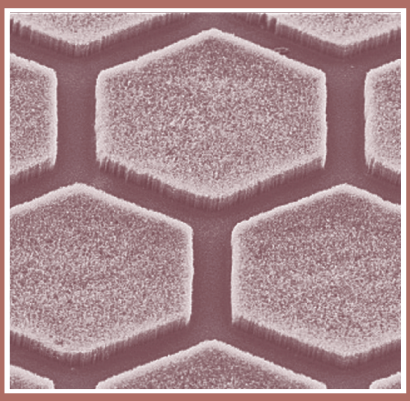
WOODHEAD PUBLISHING IN MATERIALS



Nanostructure control of materials

Edited by R. H. J. Hannink and A. J. Hill

CRC

WP

Nanostructure control of materials

Related titles:

Nanolithography and patterning techniques in microelectronics
(ISBN-13: 978-1-85573-931-4; ISBN-10: 1-85573-931-3)
Currently, surface patterning is achieved by means of optical lithographic techniques but with industry moving towards the fabrication of devices with size features of 100 nm less, the technological community is looking for alternative approaches to materials fabrication at the nanoscale. By using nanolithography, scientists can drive patterning currents through surfaces while building a 3D structure from a series of patterned layers. Electron induced chemical lithography can create ultra-high resolution templates for the site-selective immobilisation of molecules, to form functional, hierarchically organised structures. This new book will concentrate on state-of-the-art nanolithographic methods with a particular emphasis on polymer research.

Polymer nanocomposites
(ISBN-13: 978-1-85573-969-7; ISBN-10: 1-85573-969-0)
This new book concentrates specifically on the four main groups of polymer nanocomposites: layered silicates, nanotube, nanoparticle and block co-polymer systems. There is also a section on inorganic/organic hybrid systems. Each chapter gives comprehensive coverage of the dynamic properties of these materials and the various processing methods used in their production, design, performance and applications. The book is the first to give a comprehensive treatment of the subject.

Materials for energy conversion devices
(ISBN-13: 978-1-85573-932-1; ISBN-10: 1-85573-932-1)
The term electroceramic is used to describe ceramic materials that have been specially formulated with specific electrical, magnetic or optical properties. Electroceramics are of increasing importance in many key technologies including microelectronics, communications and energy conversion. This innovative book is the first comprehensive survey on major new developments in electroceramics for energy conversion devices. It presents current research from leading innovators in the field.

Details of these and other Woodhead Publishing materials books and journals, as well as materials books from Maney Publishing, can be obtained by:

- visiting www.woodheadpublishing.com
- contacting Customer Services (e-mail: sales@woodhead-publishing.com; fax: +44 (0) 1223 893694; tel.: +44 (0) 1223 891358 ext. 30; address: Woodhead Publishing Ltd, Abington Hall, Abington, Cambridge CB1 6AH, England)

If you would like to receive information on forthcoming titles, please send your address details to: Francis Dodds (address, tel. and fax as above; email: francisd@woodhead-publishing.com). Please confirm which subject areas you are interested in.

Maney currently publishes 16 peer-reviewed materials science and engineering journals. For further information visit www.maney.co.uk/journals.

Nanostructure control of materials

Edited by

R. H. J. Hannink and A. J. Hill

Woodhead Publishing and Maney Publishing
on behalf of
The Institute of Materials, Minerals & Mining

CRC Press
Boca Raton Boston New York Washington, DC

WOODHEAD PUBLISHING LIMITED
Cambridge England

Woodhead Publishing Limited and Maney Publishing Limited on behalf of
The Institute of Materials, Minerals & Mining

Published by Woodhead Publishing Limited, Abington Hall, Abington,
Cambridge CB1 6AH, England
www.woodheadpublishing.com

Published in North America by CRC Press LLC, 6000 Broken Sound Parkway, NW,
Suite 300, Boca Raton, FL 33487, USA

First published 2006, Woodhead Publishing Limited and CRC Press LLC

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library.

Library of Congress Cataloging in Publication Data
A catalog record for this book is available from the Library of Congress.

Woodhead Publishing Limited ISBN-13: 978-1-85573-933-8 (book)
Woodhead Publishing Limited ISBN-10: 1-85573-933-X (book)
Woodhead Publishing Limited ISBN-13: 978-1-84569-118-9 (e-book)
Woodhead Publishing Limited ISBN-10: 1-84569-118-0 (e-book)
CRC Press ISBN-10: 0-8493-3449-7
CRC Press order number: WP3449

The publishers' policy is to use permanent paper from mills that operate a sustainable forestry policy, and which has been manufactured from pulp which is processed using acid-free and elementary chlorine-free practices. Furthermore, the publishers ensure that the text paper and cover board used have met acceptable environmental accreditation standards.

Project managed by Macfarlane Production Services, Dunstable, Bedfordshire
(macfarl@aol.com)
Typeset by Replika Press Limited, India
Printed by T J International Limited, Padstow, Cornwall, England

Contents

Contributor contact details

(* = main contact)

Introduction

Dr A. J. Hill and Dr Richard Hannink*
Commonwealth Scientific and Industrial Research Organisation (CSIRO)
Manufacturing and Infrastructure Technology (CMIT)
Private Bag 33
Clayton South
Victoria 3169
Australia

Tel: +613 9545 2665
Fax: +613 9544 1128
E-mail: Anita.Hill@csiro.au

Tel: +613 9545664
Fax: +613 9544 1128
E-mail:. Richard.Hannink@csiro.au

Chapter 1

Philip Casey
CSIRO (CMIT)
Private Bag 33
Clayton South
Victoria 3169
Australia

Tel: 61 03 9545 2777
Fax: 61 03 9544 1128
E-mail: phil.casey@csiro.au

Chapter 2

Tomaso Aste* and Tiziana Di Matteo
Department of Applied Mathematics
RSPhysSE
The Australian National University
Canberra ACT 0200
Australia

Tel: +61 (0)2 61252846
Fax: +61 (0)2 61250732
E-mail: tomaso.aste@anu.edu.au; tiziana.dimatteo@anu.edu.au

Chapter 3

James Riches* and John Drennan
Centre for Microscopy and Microanalysis
University of Queensland
St Lucia
Queensland 4072
Australia

Tel: +61 7 3346 2935
Fax: + 61 7 3346 2101
Email: j.riches@uq.edu.au

Chapter 4

Tim Merkel* and Ingo Pinnau
Membrane Technology and Research Inc.
1360 Willow Road
Suite 103
Menlo Park
CA 94025
USA

E-mail: tcmerkel@mtrinc.com

Chapter 5

Maria Forsyth*, Josefina Adebahr, Nolene Byrne and Doug Macfarlane
Monash University
Wellington Road
Clayton
Victoria, 3800
Australia

E-mail:
maria.forsyth@spme.monash.edu.au

Chapter 6

Huajian Gao* and Haimin Yao
Max Planck Institute for Metals Research
Heisenbergstrasse 3
D-70569
Stuttgart
Germany

E-mail: hjgao@mf.mpg.de

Baohua Ji
Department of Engineering Mechanics
Tsinghua University
Beijing 100084
China

E-mail: bhji@mail.tsinghua.edu.cn

Chapter 7

Amit Misra
Materials Science and Technology Division
MS G755
Los Alamos National Laboratory
Los Alamos
NM 87545
USA

E-mail: amisra@lanl.gov

Chapter 8

Dr Georg Rixecker, Dr Zaklina Burghard and Dr Fritz Aldinger*
Max Planck Institute for Metals Research Stuttgart
Heisenbergstrasse 3
70569 Stuttgart
Germany

Tel: +49-711/689-3202
E-mail: aldinger@mf.mpg.de

Lian Gao
Shanghai Institute of Ceramics
Chinese Academy of Sciences
52 Sanlihe Road
Beijing
China 100864

Chapter 9

Roger Lumley* and Allan Morton
CSIRO Manufacturing and
Infrastructure Technology
Private Bag 33
Clayton South
Victoria 3169
Australia

E-mail: roger.lumley@csiro.au;
allan.morton@csiro.au

Ian Polmear
Department of Materials
Engineering
Monash University
Melbourne
Victoria 3168
Australia

E-mail:
Ian.Polmear@eng.monash.edu.au

Chapter 10

Kate Nairn* and Tim Bastow
CSIRO Manufacturing and
Infrastructure Technology
Private Bag 33
Clayton South
Victoria 3169
Australia

E-mail: kate.nairn@csiro.au;
tim.bastow@csiro.au

Chapter 11

Thomas Klassen
Department Powder and
NanoTechnology
Institute for Materials Research
GKSS Research Centre Geesthacht
GmbH
Max-Planck-Str. 1
D-21502 Geesthacht
Germany

Tel: +49-(0)-4152-87-2525
Fax: +49-(0)-4152-87-2636
E-mail: thomas.klassen@gkss.de

Chapter 12

Erol Harvey
MiniFAB (AUST) Pty Ltd
Caribbean Park
1 Dalmore Drive
Scoresby
Victoria 3179
Australia

Tel: +61 3 9764 2241
Fax: +61 3 9763 9867
E-mail: erolharv@tpg.com.au

Foreword

The World is confronted with a plethora of potentially disruptive technologies – some, such as nanotechnology, will inevitably challenge our cherished social, economic and industrial stability over the next two decades. Rational design and manufacture of materials properties through nanostructure control will profoundly affect the food we eat, the sources of water and power we use, our health services, our national security, housing, transportation, etc. Precisely what will be the ultimate effect of these disruptions is still unclear. However, this book is particularly timely as many new industries are already emerging from nanostructure manipulation. Smaller, First World countries, such as Australia, can play a significant role – their smaller size allows for agile adoption of radical new technologies and modification of social infrastructure, something not possible in large economies.

The idea for this book orginates from an international workshop on Nanostructure Control of Materials, held at Noosa, Queensland in 2002. The workshop was sponsored by the Australian Government research agencies and academic institutions (DEST, CSIRO and ATSE) with one of the objectives being to facilitate Australia's early entry into the marketplace with nanotechnology-enabled products. Topics of the workshop included:

- nanostructure characterisation tools
- nanoparticle fabrication and production methods
- nanocomposite production and properties
- nano self-assembly
- nanoporous materials
- nanostructured membranes
- nanodimensional sensors.

This conference showed how broad the topic of nanostructed materials was becoming at the beginning of the twenty-first century. For a country the size of Australia, which produces approximately 2% of the world's research and development, much discussion was held on exactly where to focus limited resources in this burgeoning field. As reflected by the Australian contributors to this book, Australia has focused its R&D efforts on nanoparticle production,

nanostructured materials, nanostructure characterisation techniques, and nanofabrication. All of these topics are presented at the current state of the art in this book. Significant progress has also been made by overseas colleagues on many new areas of nanostructure control of materials, and the editors have included chapters of topical interest such as hydrogen storage, membranes, laminates, and biomimetic design rules.

This book conveys the 'materials and practical manufacturing' side of the current nanoscience and nanotechnology thrust of research and development. The book is suited to those who seek an understanding and appreciation of the direction the nano-revolution is taking in materials science and technology.

T.W. Turney
Director of the CSIRO Nanotechnology Institute
Australia

Acknowledgements

First and foremost we thank DEST, ATSE, and CSIRO for funding the original workshop initiative on 'Nanostructure Control of Materials' that led to this edited book. We also thank CSIRO for providing the environment in which the pursuit of reserch in this rapidly evolving area is enabled. We thank the contributing authors for their prescient chapters on their cutting-edge research and their preparedness to share it with a wide audience. The production staff at Woodhead Publishing have been most patient and diligent in their preparation of the book. We thank the publisher, Martin Woodhead, for funding and for his belief in the vision of a book derived from the workshop. We also acknowledge colleagues from TNO-TPD and TU Delft (Hartmut Fischer and Simon de Leeuw) for hosting and funding the second workshop in this series, on 'Processing and Manufacture of Nanostructured Materials'. We thank the attendees of both of these workshops, whose contributions have allowed the editors to gauge the breadth and direction of this developing field.

Introduction: special properties resulting from nanodimensionality

ANITA HILL and RICHARD HANNINK

Man's approach to building structures has historically been a 'top-down' approach. The ability to assemble only the atoms and molecules needed to build the desired structure has been the domain of Nature and is referred to as the 'bottom-up' approach. Until we were able to 'see' the smaller detail, we did not know how Nature went about building structures. It has been the development of advanced 'measuring' instrumentation such as transmission electron- and scanning probe-microscopes that has allowed scientists and engineers to 'see' and 'manipulate' structure on the nanometer (nm) length scale (1 nm = 10^{-9} m). By controlling material structure at the nanoscale, properties can often be significantly enhanced. For example, we can exploit material properties that are much more surface-related rather than bulk controlled, including optical properties of metal oxides,[1] gas transport properties of membranes,[2] catalytic properties of nanoparticles;[3] and the utilisation of these benefits have been some of the drivers for fabrication of structure on the nanoscale.

The ability to measure and manipulate matter on the nanometer level (nanoscience) has led to the discovery of surprising material properties achievable with this ability, and to the discovery and patenting of new products and processes based on nanostructure control of materials (nanotechnology). As materials scientists and engineers, we are curious as to how the properties of matter change as its structure and composition are manipulated on the nanometre scale. We must be able to manipulate matter at this length scale. And we must be able to measure what we have done so that we may decide how we can further improve materials and products and benefit from these modifications. This book contains twelve chapters illustrating examples of nanostructure control, fabrication, modelling and utilisation. While it cannot be an all-embracing book, it sets the scene for the readers' further enquiry into this exciting field. The chapter topics have been selected to give insight into several aspects of nanostructure control of materials.

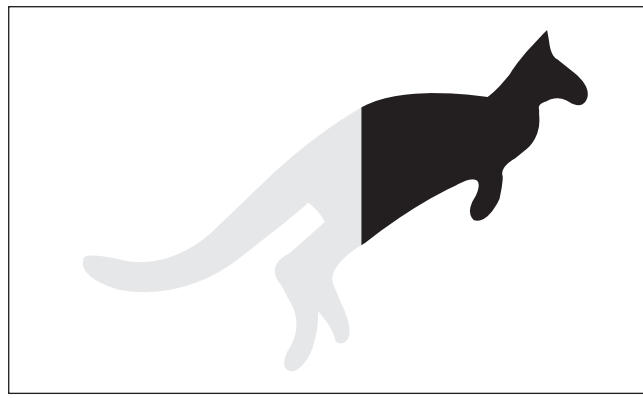
Chapter 1 deals with production technologies and applications of nano-particles, highlighting the breadth of manufacturing methods encompassing

top-down and bottom-up processes, and reviews the diversity of existing and potential applications. Currently the main applications, as described in Chapter 1, are nanoparticles used in composites, where the particles are encapsulated in a matrix. A very specific example of this technology is the 'light-strike' protective coating for transparent containers of high value-added light-sensitive products, for example pharmaceuticals, cosmetics, foods, and beverages.[4] Clear containers are transparent to damaging wavelengths of ultraviolet and short-wavelength visible violet, leaving their contents vulnerable to light-induced damage (light-strike). Light-strike is often evident by odour, taste, or change in colour of the contents.

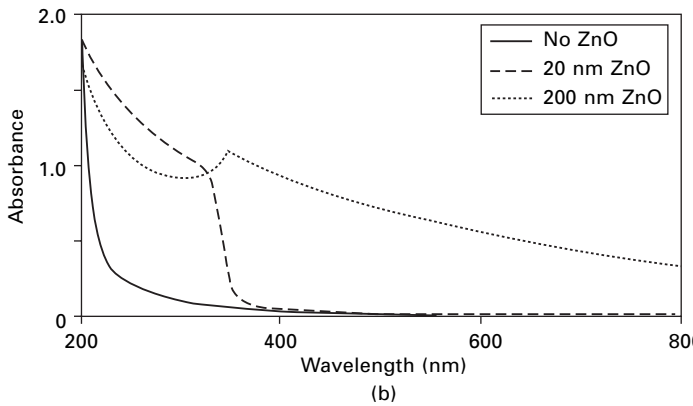
A transparent nanoparticle-containing coating can absorb wavelengths shorter than 450 nm thereby protecting the contents from damaging UV and short wavelength-visible light. The transparency allows consumers to see the contents while protecting those contents from light-strike. For glass packaging, an added advantage is that the coating provides a layer on the glass that acts as an anti-scuff scratch-resistant coating during the mass processing of the bottles. Nanoparticles can be made by top-down or bottom-up technologies, as described in Chapter 1, and the average particle diameter can be tuned to the appropriate size. The visible spectrum of light spans wavelengths between 400 and 700 nm. Bottled beverages and foods are damaged by wavelengths between 200 nm (near-ultraviolet) and 450 nm (blue).[4] An example of how visible light can be affected by particle size is shown in Fig. I.1(a). This figure shows how 30 nm diameter particles of a metal oxide encapsulated in a polymer permit the kangaroo image to be viewed clearly while 500 nm particles cause scattering and absorbance and hence a slightly less intense and obscured image. The chemical composition and weight per cent of the two components, the polymer and the metal oxide, are the same for both sides of the composite in Fig. I.1(a), only the size of the metal oxide particles is different. The uv-visible absorption spectra in Fig. I.1(b) for the polymer matrix with and without metal oxide particles illustrates the ability of the nanoparticles to absorb wavelengths up to 450 nm but, if they are small enough, still allow transparency (absorption $<<1$) in the higher visible wavelengths.[5]

Production of nanoparticles or high surface area phases is generally achieved through comminution (top-down), or self-assembly (bottom-up), approaches. In Chapter 2 Aste and Matteo use modelling to examine the preferences of atomic cluster arrangements. These arrangements can vary from nanostructural disorder to geometrical organisation. These researchers have shown that nanocluster-disorder can at times achieve a lower energy state than order.

Rigid nanoparticles are often contained within a flexible matrix as illustrated by the example of the 'light-strike' scuff resistant coating, described in Chapter 1. In this example, the composite hard/soft nanostructure controls the optical and to some degree the mechanical properties. Similarly, the transport



(a)



(b)

I.1 (a) This kangaroo image shows that 30 nm diameter ZnO particles are transparent while 500 nm particles of the same oxide in the same matrix at the same loading (60 w/o) scatter light causing the image to be diffuse and obscured (after (5)) – see text for further details. (b) Absorption spectra, ranging from ultraviolet to visible wave length light, are modified by the inclusion of ZnO particles of two different sizes in a transparent polymer matrix (after (5)) – see text for further details.

properties, such as permeability, diffusivity, and conductivity, of organic and organic-ionic materials can also be substantially changed by the addition of metal oxide nanoparticles. Chapters 4 and 5 by Merkel and Pinnau, and Forsyth *et al.*, respectively, illustrate how rigid inorganic nanoparticles, through their surface characteristics or size, can be used to manipulate the molecular and ionic transport properties of organic matrices. Application of this nanostructure approach is found in gas separation membranes (Chapter 4) as well as in ionic electrolytes for lightweight batteries and energy storage devices (Chapter 5).

The fundamental characteristic of these organic materials, that makes them suitable for applications as electrolytes or membranes, is their empty space or 'nanospace' with dimensions on the atomic and molecular scale. Nanospace controls the transport of molecules or ions through these materials as well as the interactions of ions or molecules with each other and the materials themselves during transport. This nanospace has dimensions of the order of 0.1 nanometre to a few nanometres. This unoccupied volume has static and dynamic components, both of which are critical to the transport of gas molecules and ions through the material. As shown in Chapter 4, an increase of the nanospace by less than 0.1 nm can cause dramatic increase in flux (the speed at which selected gas molecules can move through the polymer membrane).

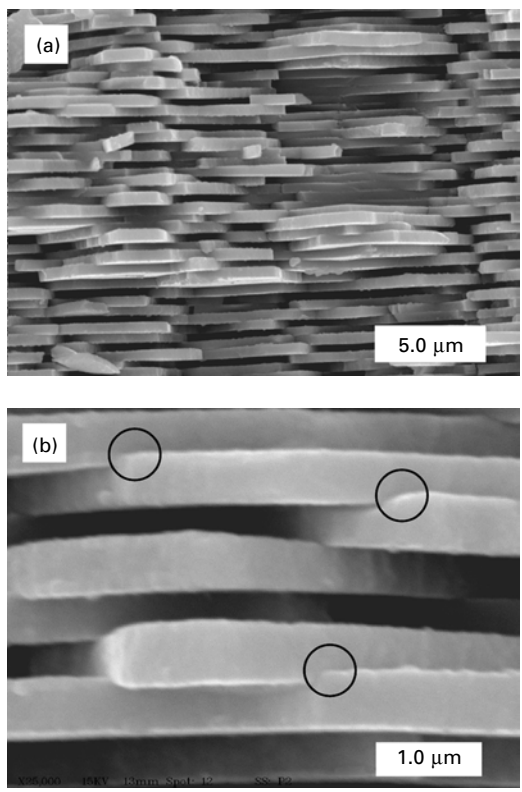
Researchers have explored hybrid or nanocomposite materials made of inorganic domains within an organic matrix to produce barrier polymers (low molecular flux) and membrane polymers (high molecular flux). The solid state ionics community has also used incorporation of nanometre-sized domains of an inorganic phase within an organic matrix to improve ionic conductivity of electrolytes.[7,8] The inability to model and predict these outcomes using current rules of mixing is indicative of the new phenomena brought about by nanodimensionality. The use of transmission electron microscopy (Chapter 3) to image nanocomposites allows knowledge of the distribution of nanoparticles within a matrix and phase content that are critical to properties. Advanced characterisation tools, which can be applied to the science of materials at small size scales encompassing a wide range of time scales, are necessary to advance from an empirical understanding of a phenomenon to a fundamental understanding. As scientists and engineers, we rely on advanced characterisation on the atomic, molecular and phase level to transition from laboratory innovations into commercial enterprises.

Nature has also used the combination of rigid and flexible materials at the nanoscale to evolve composite structures that optimise mechanical properties (fracture toughness, defect tolerance, strength) for lightweight defensive and structural components of vertebrates and invertebrates, e.g. bones, shells, etc. In Chapter 6, Gao, Ji, and Yao describe the engineering principles behind the superior properties that arise from combinations of stiff/hard and flexible/soft matter on the nanometre scale; material structures that nature has been evolving for an aeon. Scientists and engineers have long been enthralled by nature's works on the nanoscale, entailing both biological and mineralmaterials often combining the two (collagen and hydroxyapatite), for example nacre, abalone, mother of pearl, opal, eggshell, tooth enamel, etc. The desire to be able to build (self-assembled) structures at this level comes from an appreciation of their practical beauty, both aesthetic and functional, and their material utilisation efficiency by using the bottom-up approach.

Nacre is the lining of sea shells and is also known as mother of pearl.

Nacre has outstanding mechanical properties such as strength, toughness, and wear resistance. These superior properties are attributed to the arrangement of layers of stiff inorganic aragonite (a form of $CaCO_3$) and flexible organic proteins, where the *interfaces* are critical. The recent achievement of synthetic analogues of nacre and bones represents a fundamental breakthrough in nanostructured materials design.[9] This synthetic structure also leads to a high tolerance for defects generally resulting from impact and abrasion.

As discussed by Gao *et al*. in Chapter 6, modelling of the effect of nanodimensions on the mechanical properties of laminate structures such as nacre and bone is well advanced. A good example of this is the prediction of the extraordinary fracture toughness of nacre, even when the constituent components of stiff calcium carbonate and soft protein possess mechanical properties similar to their bulk phase. The mechanism for deformation is one that absorbs a lot of energy, giving the nacre its toughness. It was recently shown[10] that platelets of calcium carbonate interlock on the nanometre-



I.2 (a) Image of nacre platelets. (b) Image of the mechanical interlocking of nacre platelets of calcium carbonate (Courtesy Katti *et al.* (10)) – see text for further details.

micron scale to make the failure mode one of progressive failure of the individual interlocks (a tough, energy absorbing mode of failure) with an accompanying large deformation of the organic protein interlayer.

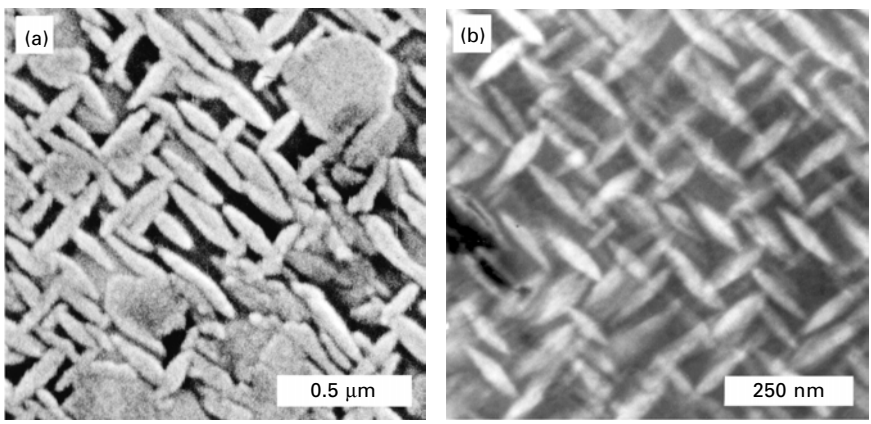
Misra in Chapter 7 has shown how metallic laminates also benefit from alternating layers of soft and hard metals. These systems can possess ultra-high strengths approaching the theoretical strength of metals when the layers are nanometre in scale with the interfaces carefully controlled.

In a similar manner to that employed by Nature, where a normally brittle matrix of calcium carbonate can be acceptably toughened through the manipulation of the structure and phase content, synthetically fabricated materials can also be optimised using this approach. In Chapter 8, Rixecker *et al*. have shown how monolithic nanocrystalline ceramics can be formed through a variety of fabrication processes to achieve a range of property improvements. The 'holy-grail' for engineering ceramics has been the improvement of both strength and toughness. As pointed out in Chapter 8, strength is generally described by the Griffith theory where the strength of a brittle material is inversely related to the square root of the dominant flaw size. In well processed ceramics, a very good correlation exists between the flaw size and the grain size, so to increase strength the incentive to pursue nanograin sized ceramics can be rewarding. This approach has been explored with limited success as explained in Chapter 8.

The development of 'fracture-toughness' for brittle ceramics has become the classic example of nanostructure control of ceramics. This approach was first highlighted by the development of a toughened ceramic known as partially stabilised zirconia (PSZ). PSZ was thought to be of little value as an engineering material due to its destructive phase transformation and its brittle behaviour under stress. Control of phase form, size and composition at the nanometre level changed this brittle ceramic to a tough ceramic engineering material. The demonstration of PSZ microstructural control was first published as a paper in 1975 under the title 'Ceramic Steel?'.[11] The approach described in this paper changed the whole concept for the future development of engineering ceramics through composition (alloying) and thermal treatment control. Hannink, Kelly and Muddle[12] recently reviewed the extent and impact of this development. The essence of the development was the subtle manipulation of the precipitation processes from a solid solution. This process resulted in the formation of five different matrix and precipitate phases whose size and proportions lead to both enhanced strength and fracture toughness. The mechanisms that have been identified to increase the functionality of such ceramics include transformation-, microcrack-, crack deflection-, and crack bridging-toughening. By far the most effective mechanism, at room temperature, is transformation toughening.

One example of the transformation toughening mechanism is found in the magnesia doped zirconia ceramic, known as a magnesia-partially stabilised

zirconia (Mg-PSZ). Mg-PSZ is a zirconia based alloy containing ~9 mol % magnesia, sintered and solution treated, with special control over the cooling and ageing cycle. When optimally aged, Mg-PSZ contains polymorphic phases of zirconia, which may include cubic, tetragonal, monoclinic and orthorhombic. The transformation toughening mechanism is invoked when tetragonal particles transform to the room temperature stable monoclinic form during the application of a stress or around a propagating crack. The phase that assists the nucleation of this transformation is called the delta phase. To achieve the optimal tetragonal precipitate size, as well as the optimal density and distribution of the delta-phase, the ceramics are 'massively-undercooled'. The undercooling process leads to the nucleation of a large number of nanosized delta-phase particles – the optimum nanosized particles leading to enhanced mechanical properties. The material is then taken back up to a suitable sub-eutectoid ageing temperature to grow these delta-phase nuclei that assist in 'destabilising' the tetragonal phase. The development of these sub-eutectoid heat treatment processes illustrate how lateral-thinking or a counterintuitive approach may lead to startling progress in the development of material properties or processes. The images in Fig. I.3 illustrate the microstructure of an optimally sub-eutectoid-aged Mg-PSZ ceramic. The scanning electron microscope (SEM) image of Fig. I.3(a) shows the polished and etched surface of Mg-PSZ where the three variants of the tetragonal precipitates occur as discs (two variants on edge and one in plan view). These precipitates are contained in a cubic stabilised zirconia matrix (black region). Figure I.3(b) is a bright field



I.3 (a) Scanning electron microscope images showing the polished and etched surface of a magnesia-partially stabilised zirconia ceramic where the three variants of the tetragonal precipitates contained in a cubic-stabilised zirconia matrix are clearly visible (two on edge, one in plan view). (b) Bright field-TEM image showing the size and density of two tetragonal precipitate variants – see text for further details.

transmission electron microscope (TEM) image of predominantly tetragonal zirconia precipitates in a cubic stabilised zirconia matrix. These elegant nanostructures epitomize the subtlety of materials control for ceramics.

In Chapter 9 the work of Lumley, Morton and Polmear, which focuses on the nanostructure control of metals, has shown that similar counterintuitive sub-eutectoid heat treatments can further enhance the properties of aluminium alloys. In particular, Chapter 9 describes how the super-saturation and precipitation of a high density of nanoparticulate second phases can significantly improve the mechanical properties of aluminium-based alloys. Lumley *et al.* also describe other alloy systems where the employment of forces under our control, e.g. temperature, time, and composition, can be used to tailor the nanostructure of metallic alloys, producing improvements in mechanical properties not previously thought achievable. The development of metallic and ceramic materials with superior properties achieved through nanostructure control has been translated to wide ranging industrial applications thereby extending the use of these materials as their areas of application become more demanding.

Understanding how these combinations of time–temperature–composition produce improvements in bulk properties arising from the nanoscale is possible using advanced characterisation tools. In Chapter 10, Bastow and Nairn show how the atoms in an alloy, with specific emphasis on light metals, e.g. aluminium and magnesium, can be tagged and followed using solid state nuclear magnetic resonance (NMR). This capability to quantitatively follow the atoms as they form nanometre-sized clusters has enabled the elucidation of the important pre-precipitate formation in these lightweight engineering alloys. Being able to see how many solute atoms are still in solid solution and available to undergo dynamic precipitation underpins the ability to design optimum nanostructures. This powerful technique can be used in conjunction with small-volume sampling techniques such as high resolution electron microscopy or atom probe field ion microscopy to obtain information on precipitate morphology and microstructures. As NMR is a rapid, large-volume sampling technique, the work of Bastow and Nairn has opened a new area of high throughput atomic and nanoscale metallurgical investigation of alloys.

Klassen (Chapter 11) has used nanostructure control of metal hydrides to facilitate the transport and storage of large volumes of hydrogen gas. One of the major challenges of the forthcoming hydrogen economy is the safe storage of large volumes of hydrogen gas for use in automotive and general transportation systems. A number of approaches look promising, which include both chemical and physical storage systems in addition to the more conventional liquid and high pressure approaches. To date, nanostructured metals and metal alloy systems, nanotubes and large surface area carbon nanoparticles are being investigated as safe hydrogen storage systems. Currently the nanocrystalline metal hydride and complex hydride route has been preferred

as the most promising immediate solution.[13] The work of Klassen and colleagues, described in Chapter 11, examines the case of nanocrystalline magnesium hydride with various catalysts to facilitate the adsorption and desorption of hydrogen in these nanostructured hydrides.

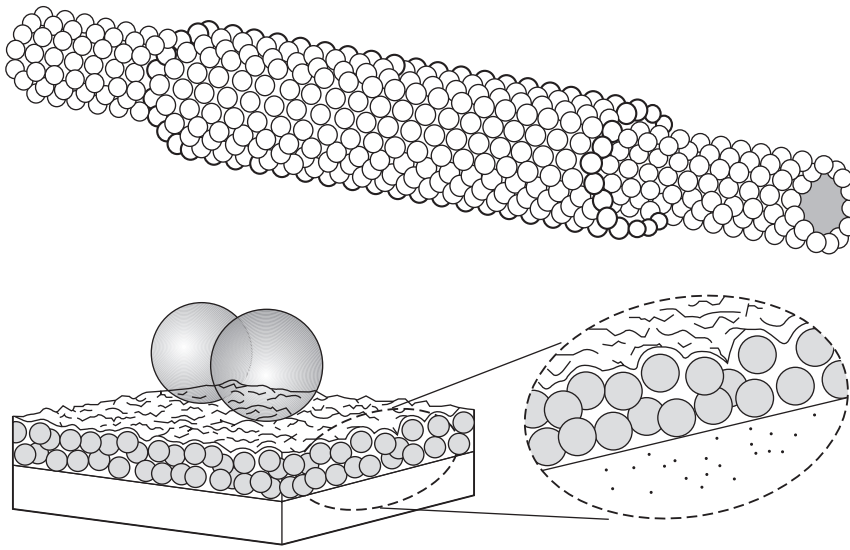
In Chapter 12, Harvey and Ghantasala bring together nanostructure control and nano- and micro-manufacturing techniques to enable functional components and systems to be made for commercial applications. The range of methods for creating individual components at the nanoscale may include molecular chemistry, biological growth, various lithographic techniques, deposition and ablation techniques, and templating. The advantage to be gained by creation and manipulation of structures at this level is that it almost invariably utilises the bottom-up approach, which translates to zero or minimal waste manufacturing.

An elegant example of nanofabrication of an engineering system is that provided by the work of Zettl and co-workers[14,15] on low friction bearings and springs fabricated from multi-wall carbon nanotubes. This example is schematically shown in Fig. I.4(a) and illustrates the potential of exploiting nanoscale surface features for the manufacture of low-friction devices. The production of inner- and outer-roller bearings and laterally sliding cages or wires are all potential configurations using this approach. A second example of the exploitation of nanoscale surface features is in the fabrication of super-hydrophobic[16,17] or -hydrophilic surfaces. These surfaces can be exploited for self-cleaning fabrics, paints, windows, and in microfluidics. The schematic illustration of Fig. I.4(b) shows how molecular clusters assembled on a rigid or flexible surface, and covered by a molecular net, can produce surfaces of an extreme hydrophobic nature. In this example the droplets with wetting angles near zero illustrate super-hydrophobicity.

These final two examples of nanofabricated structures give some indication of the future direction that manufacturing may take and the benefit that may arise from the nanostructure control of materials. The adoption and use of these technologies will help society achieve and maintain a standard of living not possible with our current manufacturing methods and present rate of natural resource consumption. Nature uses low energy manufacturing routes to produce efficient and practical nanostructured materials. This bottom-up approach to manufacturing is one in which everything can be recycled and nothing is wasted: sustainable manufacturing.

Our own vision of the future includes:

- noise- and pollution-less production systems
- transportation that has no environmental impact and is self guiding
- water in cities that is cleaned at drain and re-enters the distribution system immediately and
- sustainable energy sources that are non-polluting and highly distributed.



I.4 (a) Shows schematically how multi-wall nanotubes can be used to create a rotating shaft and bearing system for the potential manufacture of low friction devices via exploitation of nanoscale surface features (after (14,15)). (b) Illustrates how molecular nanoscale clusters, deposited on a rigid solid or flexible surface, might appear and (may be fabricated to) produce surfaces of an extreme hydrophobic nature (after (16,17)).

We feel that a better understanding and use of *Nanostructure control of materials* will help society achieve these goals.

References

1. G. Schmid (ed.), *Nanoparticles: From Theory to Application*, Wiley-VCH (2004).
2. A. Mayes, 'Nanocomposites: Softer at the boundary', *Nature Materials*, **4**, 651–652 (2005).
3. R. Glazier, R. Venkatakrishnan, F. Gheorghiu, L. Walata, R. Nash and W.X. Zhang, 'Nanotechnology takes root,' *Civil Engineering* **73** (5) 64–69 (2003).
4. Australian or International Pat Nos: WO200285992-A; WO200285992-A1; NO200304646-A, 'Coating composition for containers for storing light sensitive products, comprises carrier dispersed with pigment having ultraviolet light absorber nanoparticles, or ultraviolet and visible light absorber nanoparticles' D.R. Harris, P. Meakin, T.W. Turney and I. Lely.
5. A. Ammala, A.J. Hill and T.W. Turney, 'Invisible eternal additives', *Polym. Mater. Sci. and Engr.* **89**, 128–129 (2003).
6. T.C. Merkel, B.D. Freeman, R.J. Spontak, Z. He, I. Pinnau, P. Meakin and A.J. Hill, 'Ultrapermeable, reverse-selective nanocomposite membranes', *Science* **296** (5567) 519–522 Apr 19 (2002).
7. F. Croce, G.B. Appetecchi, L. Persi and B. Scrosati, 'Nanocomposite polymer electrolytes for lithium batteries', *Nature* **394** (6692) 456–458 Jul 30 (1998).

8. P. Wang, Q. Dai, S.M. Zakeeruddin, M. Forsyth, D.R. MacFarlane and M. Gratzel, 'Ambient temperature plastic crystal electrolyte for efficient, all-solid-state dye-sensitized solar Cell', *J. Am. Chem. Soc.* **126** (42) 13590–13591 Oct 27 (2004).
9. Z. Tang, N. Kotov, S. Magonov and B. Ozturk, 'Nanostructured artificial nacre,' *Nature Materials*, **2**, 413–418, June (2003).
10. K.S. Katti, D.S. Katti, S.M. Pradhan and A. Bhosle, 'Platelet interlocks are the key to toughness and strength in nacre', *J. Mater. Res.* **20** (5) 1097–1100 (2005).
11. R.C. Garvie, R.H. Hannink and R.T. Pascoe, 'Ceramic steel?', *Nature* **258** 703–705 (1975).
12. R.H.J. Hannink, P.M. Kelly and B.C. Muddle, 'Transformation toughening of zirconia containing ceramics', *J. Am. Ceram. Soc.* **83** (3) 461–87 (2000).
13. A.M. Seayad and D.M. Antonelli, 'Recent advances in hydrogen storage in metal-containing inorganic nanostructures and related materials', *Adv. Mater.* **16** (9–10) 765–777 (2004).
14. B.C. Regan, A. Aloni, R.O. Ritchie, U. Dahmen and A. Zettl, 'Carbon nanotubes as nanoscale mass conveyors', *Nature* **428** (6986) 924–927 (2004).
15. A.M. Fennimore, T.D. Yuzvinsky, W.Q. Han, M.S. Fuhrer, J. Cumings and A. Zettl, 'Rotational actuators based on carbon nanotubes', *Nature* **424** (6947) 408–410 (2003).
16. Australian or International Pat Nos: WO200114497-A; WO200114497-A1; AU200065497-A; 2001-290250, 'A method for forming a modified gel capable of being applied to a surface, forming a hydrophobic surface on which water has a contract angle at least 150 degrees which can be used as an anti-icing, anti-fouling and anti-corrosion coating', A.W. Jones, R.N. Lamb, H. Zhang, *et al*.
17. Australian or International Pat Nos: WO9842452-A; EP969934-A; WO9842452-A1; 1998-531767, 'Application of durable, hydrophobic surface film – comprising coating article with particles with functional groups and cross linking to bind particles to one another and article', R.N. Lamb, H. Zhang and C.L. Raston.

1

Nanoparticle technologies and applications

P C A S E Y, CSIRO, Australia

1.1 Introduction

In recent years, the nanotechnology concept has stimulated the collective imagination of the scientific, engineering and business communities. If its perceived potential can be realised then it can arguably be regarded as a 'disruptive' technology. Whether or not this occurs remains to be seen. In the meantime, the lure of what may be possible, by exploring and developing science and engineering at the nano-level, has reinvigorated the efforts of these practitioners across almost the entire spectrum of disciplines. Applications that may fundamentally impact in many aspects of our lives appear to be as broad as the imagination. While this new outlook is refreshing and exciting, what has become clear is that the breadth and depth of information surrounding it is fast becoming virtually unmanageable.

This chapter represents a general attempt to broadly scope technologies involved in nanoparticle production so that the reader has an appreciation of the range of methods used. The reader is refferred to references [1–11] as a useful starting point for further reading. Later chapters will cover specific areas in much greater detail than is presented here. As such it is a broad survey, which limits itself to the production of nanoparticles (<100 nm) and does not encompass the design, characterisation or application of what is generally termed structures, devices or systems. It focuses on 2 and 3D nano-particulate materials rather than one-dimensional materials such as films, engineered surfaces or positionally assembled nanoparticles. Table 1.1 presents examples of these.

In general, there are two approaches to nanoparticle production that are commonly referred to as 'top-down' and 'bottom-up'. 'Top-down' nanoparticles are generated from the size reduction of bulk materials. They generally rely on physical, the combination of physical and chemical, electrical or thermal processes for their production. Such methods include high-energy milling, mechano-chemical processing, electro-explosion, laser ablation, sputtering and vapour condensation. 'Bottom-up' approaches generate nanoparticles

Table 1.1 Examples of two- and three-dimensional nanoparticles

Two dimensions (few nm in diameter, length up to several cm)	Three dimensions (<100 nm in diameter)
carbon nanotubes (single and multi wall)	nanoparticles (particles
inorganic nanotubes	< 100 nm in diameter)
carbon and inorganic nanorods	fullerenes (C_{60})
nanoplatelets	dendrimers
nanofibrils	quantum dots
nanowires	
biopolymers	

from the atomic or molecular level and thus are predominantly chemical processes. Commonly used techniques are crystallisation/precipitation, sol-gel methods, chemical vapour deposition and self-assembly routes. Some processes may use a combination of both.

Either approach may be performed in all three states of matter, i.e., vapour, solid or liquid (or combination of these) and the limits to the physical size of nanoparticles produced by either approach is converging and may overlap. Consequently, the choice of particle size, from a product design perspective, is directly influenced by process economics, capability to supply and the adequacy and type of performance required in the target application. To meet performance criteria, not only does the nanoparticle material have to have (or be able to impart) desired functionality but also that functionality must be predictable and reliable. The latter is often determined by extrinsic factors such as the degree of dispersion, the level of contamination and the working environment. Apart from characteristic size range significant commonality is not apparent between the technologies used to produce nanoparticles although some techniques appear to be more flexible than others. Table 1.2 summarises the range of production processes by type and will be used as the basis for the following sections. Nanoparticles have a size dimension ranging from the atomic and molecular (~0.1 to 3 nm) up to 100 nm and thus represent a 'bridge' between the quantum and 'real' world (micro and macro).

Table 1.2 Summary of production processes by type

Process routes	
Top-down	Bottom-up
High-energy milling	Crystallisation
Chemical mechanical milling	Sol-gel
Vapour phase condensation	Chemical vapour deposition
Electro-explosion	Self-assembly
Laser ablation	
Sputtering	

Nanomaterial properties differ from their 'real' world (bulk) counterparts primarily due to changes in physical attributes (size, shape and specific surface area) and quantum effects such as quantum confinement (the organisation of energy levels into which electrons can climb or fall being squeezed into a dimension that approaches a critical quantum measurement). As a result of these nanoscale particle dimensions the properties that can change from those observed in the bulk include mechanical, physical, chemical, electronic, optical and magnetic. A change in more than one property may occur in the same particle (e.g. optical and electrical) so that particles with a range of functionalities can be engineered for target applications. Moreover, these properties may be used intentionally to directly affect bulk properties of any material system into which they are incorporated.

Some examples of property changes resulting from nanoparticle dimensionality are presented below:

Chemical: when the particle size of Au (usually regarded as inert in the bulk) is reduced to 2–10 nm the catalytic activity of supported gold catalysts can be significantly enhanced [12]. Gadolinium acetylacetonate (GdAcAc), a potential anticancer agent for neutron capture therapy (NCT), has been entrapped in stable nanoparticles. The apparent water solubility of GdAcAc increased more than 2000-fold by such entrapment [13].

Mechanical: the hardness of silicon nanospheres (20–50 nm) has been measured to be four times greater than the expected value for bulk silicon (50 Gpa *cf*. 12Gpa) [14]. Quantum effects begin to dominate at the nanoscale and are the key determinant influencing optical, electrical and magnetic properties. The exploration of how these effects influence nano-, macro- and fabricated-bulk properties is the subject of a great deal of the continuing research.

Optical: nanoparticles, with dimensions less than the wavelength of light (400–700 nm), appear transparent when well dispersed. Moreover, colour and luminescent properties can be significantly altered. For example, the optical properties of gold nanoparticle films have been fabricated with tailored optical responses, both in transmitted and reflected light, and to generate a series of materials with hybrid optical properties – i.e., possessing tunable opacity and reflectivity [15].

Electronic: the electronic structure of a nanocrystal critically depends on its very size. For small particles, the electronic energy levels are not continuous as in bulk materials, but discrete, due to the confinement of the electron wavefunction as a result of the physical dimensions of the individual nanoparticles. This confinement exhibits quantum size effects influencing properties such as electrical conductivity and magnetic susceptibility. Perhaps the most impressive example of the relationship between size and electronic

effects is that of size induced transitions from metal to non-metal in nanocrystals such as Hg, Au, Ag, Pd, Ni and Cu [16].

Magnetic: magnetic properties of nanoparticles of transition metals such as Co, Ni show marked variations with size. In the nanometric domain, the coercivity (a measure of the strength of a magnetic field) of the particles tends to zero and nanocrystals behave as superparamagnets with no associated coercivity or retentivity. This is somewhat analogous to the metal-insulator transition observed in nanocrystalline metals. The blocking temperature, which marks the onset of this superparamagnetism, also increases with the nanocrystal size. Further, the magnetic moment per atom is seen to increase as the size of a particle decreases [17].

As a particle becomes smaller it theoretically can physically fit into, be assembled or positioned onto/into or pass through significantly smaller spaces. The implications of this feature are enormous in terms of the potential applications. Perhaps the greatest realised impact of this has been in the functional quality and performance properties of bulk material systems such as polymer nanocomposites and coatings (lighter, stronger, functionally smarter); ceramic composite materials (improved fracture toughness, wear resistance, greater biocompatibility, etc.); data storage and processing systems (greater capacity to store and process at reduced size); dispersions (for chemical mechanical polishing, personal care); sensors and instrumentation, drug delivery systems and enhanced activity in catalytic processes.

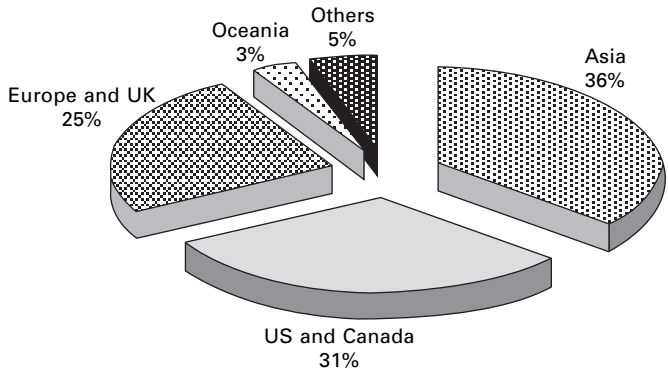
1.2 Commercial production of nanoparticles

Commensurate with the rise in scientific interest in nanotechnology there has been an increasing number of companies (now in excess of 500) claiming prime business in this area. If one allows for formalised activities such as consultancies, specialised centres, institutes, etc., then the number is even greater. A significant proportion of these are in the development phase and their ability to bring a realisable technology to the marketplace is undetermined. What is clear is that the number that produce and sell nanoparticulate products is relatively small and represents <10% of the total number of companies. A brief and by no means comprehensive survey is presented in Table 1.3 and Figs 1.1 and 1.2 where global company distribution and type of product are summarised [18].

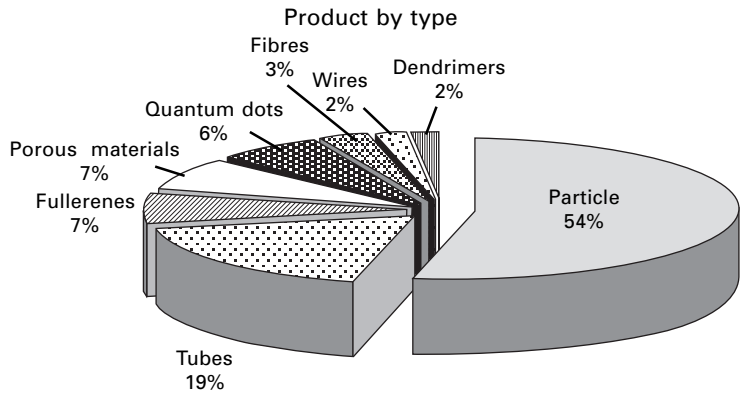
Most companies are generally small to medium enterprises (SME) and produce products for niche market areas. Their production capacity appears to be limited by market demand and ranges from grams to several kilograms to hundreds of kilograms depending on the type of product. Capacity to scale up to substantial commercial quantities (tons) is claimed in only a few instances. In recent years there are more occurrences of the larger production houses aspiring to use (or acquire) production resource to enter into nano markets.

Table 1.3 Summary of nanoparticle-producing companies together with product type

Region	Country	Number of listed nanotech concerns	Number dealing with nanomaterials	Number producing nanoparticles	Listed materials on product list
North America	US and Canada	260	98	19	oxides (Ti, Zn, Ce, Al, Zr, Y, Cu, Mg, Si), luminescent quantum dots, nanofibres, fullerenes, nanotubes, MoS_2, WS_2, clays, carbides, nitrides, metals (Al, Ni, Cu), titanates (Ba, Sr)
Asia	China	25	20	10	oxides (Zn, Ti, Ce), nanotubes, clays, nitrides, carbides, diamond
	Japan	42	14	3	oxides (Zn, Zr), nanotubes, fullerenes, silver
	Korea	18	9	7	oxides ITO, ATO, Ce, Ti, metals Ag, Pd alloys WS_2, WC/Co, carbon nanofibre
	Taiwan	10	3	2	metals Ag, Au, oxides Ti, Zn
Europe and UK	Germany	79	31	5	metals Ag, oxides, ceramics, silica, zeolites, nanotubes, titania,
	France	17	4	2	nanotubes, alumina
	UK	63	14	3	nanotubes, metals (Al, Cu, Co, Ag, W, Mo) oxides (ZnO, CuO, ZrO_2, Y_2O_3), ceramics (SiN, SiC, AlN) alloys
	Sweden	7	2	1	colloidal silica
	Slovakia	1	1	1	graphite materials, nanotubes
	Belgium	2	1	1	nanotubes
Oceania	Australia	18	9	2	oxides (Al, Zn, Ce), talc
Others	Israel	22	9	3	metal and metal oxide powders, alloys Ag Ag/Pd Ag/Pt Ag Au Ni Cu Fe Co Ba
	Cyprus	1	1	1	nanotubes



1.1 Global distribution of nanoparticle producing companies.



1.2 Product by type.

Major product groupings are primarily metal and metal oxide powders and carbon nanotubes with mixed metal oxides, non-oxides and silicates secondary. The reasons for this are largely historical as most of these materials are not 'new' materials and find themselves in existing applications (carbon nanotubes aside). As such their refined nano counterparts have followed similar paths to market. Carbon nanotubes (and fullerenes) in this context represent a 'new' material. Their discovery has occurred only in recent decades. Organic particles such as dendrimers do not as yet occupy a significant portion of the nanoparticulate production pie although the activity in this area (represented by patent activity) is substantial. In more recent years China has been the most significant mover in bringing to market production capacity especially in the area of carbon nanotube production.

1.3 Synthesis and production processes

1.3.1 Top-down processes

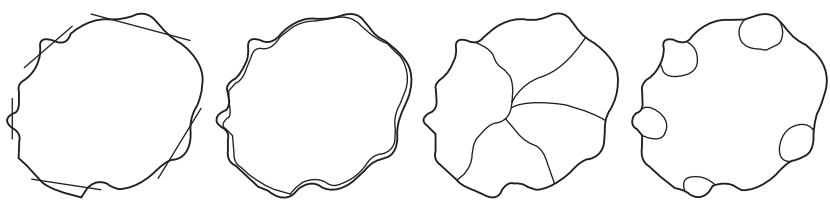
There is a range of top-down processes that can be used to produce nanoparticles. The most significant of these are physical methods (high-energy milling), the combination of physical and chemical methods (chemical-mechanical milling) and vapour phase condensation (using laser ablation, electro-explosion, sputtering and vapour condensation using thermal methods).

High-energy milling

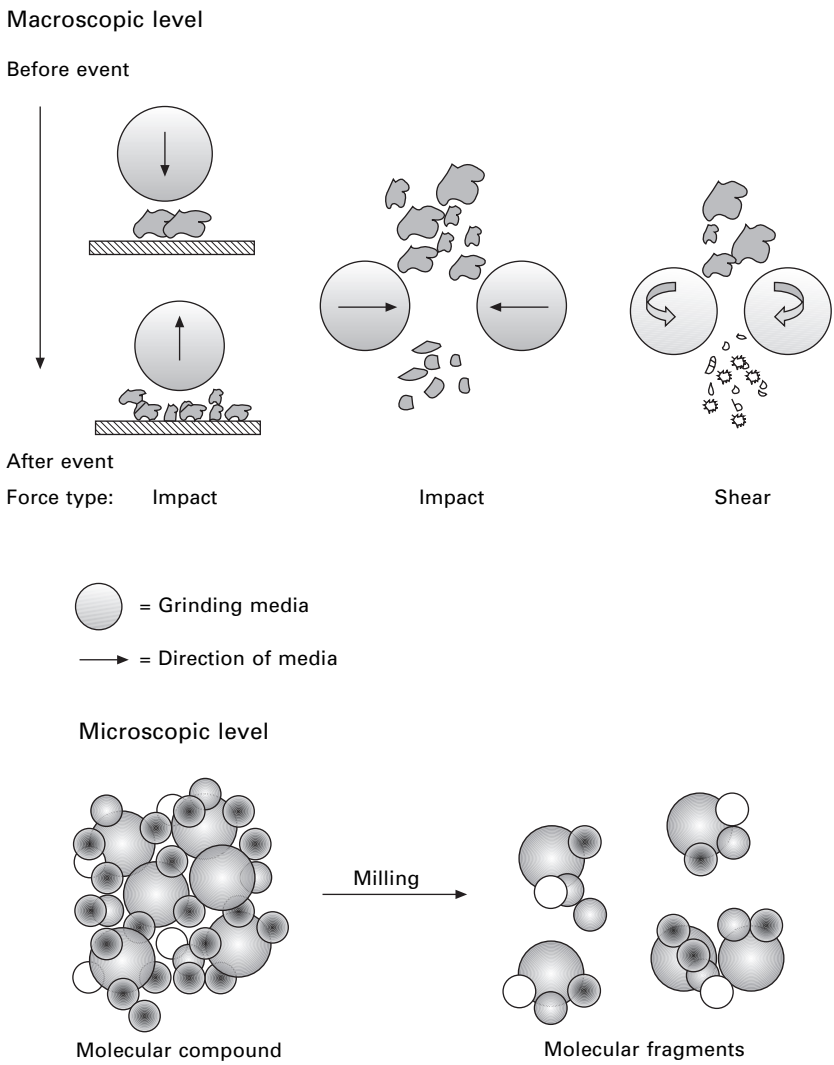
There are in general two methods by which nanoparticles can be produced using high-energy milling: (i) milling alone and (ii) combining chemistry and milling (referred to as chemical-mechanical milling or mechanochemical processing). It is suggested that these methods offer the advantage of being easily scaled. References [19 and 20] are good starting points for further reading.

High-energy milling processes involve the comminution of bulk materials. The principle of comminution is centred on applying physical forces to bulk material so as to effect breakage into smaller sizes. The forces required to effect breakage are usually a combination of either impact or shear. Material is introduced into a milling chamber in which grinding (milling) media are contained. Milling occurs when the media is made to move either by stirring (using a rotor) or by shaking/vibrating the chamber and contacts the bulk material thus imparting, depending on the milling parameters, either impact or shear forces on it. Breakage can occur through a variety of mechanisms and are generally described as attrition, abrasion, fragmentation or chipping and occur both at the macro and microscopic level [21]. This is illustrated in Figs 1.3 and 1.4

The rate at which comminution occurs is dependent on the size and frequency with which forces are applied. Breakage is influenced by both extrinsic and intrinsic factors. Intrinsic factors include such things as material properties (hardness, density, size) whilst extrinsic factors are determined by the amount of energy put into the system and the efficiency with which that energy is

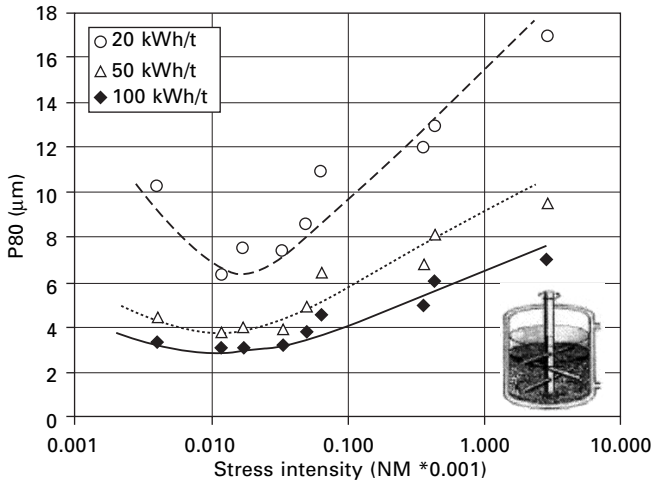


1.3 Breakage mechanisms; (a) attrition, (b) abrasion, (c) fragmentation, (d) chipping.



1.4 Breakage at the micro and macro level.

translated to the milling process. The latter is determined by variables such as vibrating frequency (in a rotorless mill), rotor speed (in a stirred mill), mill design, media size and loading, solids loading and whether the milling is performed dry or wet. These variables dictate which force regime predominates (i.e. shear or impact) which in turn dictate milling rate and efficiency. In high-speed stirred mills the effect of mill tip speed, media size and density can be evaluated simultaneously using the grinding media 'stress intensity' approach and an illustration of this is summarised in Fig. 1.5 which shows a plot of product particles size (starting size 45 μm, product size ~2–6 μm) versus stress intensity for a pin mill using a zinc concentrate.



1.5 Stress intensity plot for milling in a pin mill [22].

A variety of mills are commercially available and range from tumbling, shaker, vibratory, planetary and stirred ball mills. Production of nanoparticles using this technique is sometimes limited by the need for extended milling times, material properties and contamination issues. Attrition methods allow the production of alloys and composites that cannot be synthesised by conventional casting methods. They have also gained attention as non-equilibrium processes resulting in solid state alloying beyond the equilibrium solubility limit and the formation.

The types of nanoparticles produced by the attrition milling technique are generally alloys or single-phase powders. When a single-phase elemental powder (or intermetallic) is milled, grain size asymptotically reduces to a range of 3–30 nm [19]. For alloys produced by this method, unstable intermediate substances are formed, from mixing and diffusion as a consequence of repeated deformation and folding of the different metals. These intermediates allow the chemical reactions necessary for alloy formation to occur [11].

For non-metallic compounds (carbides, oxides, etc.) the reduction in grain size is consequent on fracturing and cold welding and the limit to minimum grain size is determined by the minimum size that does not support nucleation and propagation of cracks within the grain. For metallics, on the other hand, it is thought that the reduction in grain size is a process where localised plastic deformation is induced, subgrains are formed (by eradication of dislocations) which combine (through intimate mechanical contact) to form discrete grains. The latter process is analogous to recrystallisation observed during hot forming of metals and alloys but in these circumstances at low temperatures. In intermetallics, the process is thought to be different again in that grain formation is due to nucleation (on a nanoscale) followed by a

limited growth of the generated phase [20, 23]. There are numerous examples in the literature of alloy and mixed metal oxide production using this technique [20, 28]. Few examples can be found where single-phase powders or particles are produced at the nanoscale level [24, 25].

Chemical–mechanical milling

Mechanochemical routes to the production of nanoparticles utilise the (usually simultaneous) combination of chemical and physical forces. Variants include (i) milling under controlled atmosphere and (ii) milling precursor materials that react to form a product phase during milling. The latter has gained more attention in the literature.

Milling under controlled atmosphere

There are a few examples in the literature where milling has been performed using a reactive atmosphere in an attempt to produce nanophase alloys, referred to as CMA (Chemical Mechanical Alloying) or CRMA (Controlled Reactive Mechanical Alloying). The work is developmental and has focused primarily on the synthesis of nitrides (Ti, Al, Si, Cr, Zr, etc.) and hydrides (Mg-Fe) [26, 27]. In these processes a metal is milled under ammonia or nitrogen (for nitride formation) or hydrogen (for hydride formation). The metal itself may have been pre-milled under an inert atmosphere such as argon or helium before the reactive alloying step. Experimental studies have produced evidence that nanophase product was formed, however, the mechanisms by which this occurred are complex and not well understood. Consequently, the production of nanoparticles by this method, whilst interesting, is restricted.

Milling reactive precursors

The production of nanoparticles by milling a combination of compounds to form a new product by a solid-state displacement reaction has received increasing attention over recent years. The process is referred to as Mechano-chemical Processing (MCP^{TM}). In this process two or more materials are simultaneously milled to produce, through an exchange reaction, a nanoscale composite that can be further processed into dispersed nanopowders by removing the matrix phase. For example, ZnO powder has been produced by simultaneously milling powders of $ZnCl_2$ and Na_2CO_3 to form $ZnCO_3$ and NaCl by the following reaction:

$$ZnCl_2 + Na_2CO_3 \rightarrow ZnCO_3 + 2NaCl \quad (\Delta G = -80 \text{ kJ})$$

The nanostructured product mix is then heat treated (170–380 °C) to thermally

decompose $ZnCO_3$ to ZnO, washed (to separate the sodium chloride from ZnO) and dried. Particles produced by this method had an average particle size of 27 nm. Amounts of NaCl (excess to stoichiometry) were added to act as a diluent and assisted both in particle separation and size control [29]. A range of materials has been produced by this method commercially and in the laboratory. These include oxides, sulphides, carbonates doped metal oxides and metals. Table 1.4 presents some examples of these.

Precursors can be selected from oxides, carbonates, sulphates, chlorides, fluorides, hydroxides and reported products are not limited to those cited above but vary across a range of metals including, for example, Cu, Ni, Al, Cd, Pb and Se. Process control is dependent on a number of variables, which include milling time, the level of diluent and the choice of starting material and thermal treatment parameters [30–37].

Vapour phase processes

In many ways the vapour phase process, and in particular condensation from the vapour phase of nanoparticulate materials, is currently the most versatile synthesis process. It has the ability to generate non-agglomerated nanopowders, which are sinterable at relatively low temperatures. In this approach, a vaporisable material source is used to generate the powder particles, which are convectively transported to and collected on a cold substrate. Particles are formed in a thermal zone above the material source, due to interactions

Table 1.4 Examples of materials produced by milling reactive precursors [30–37]

Compound	Reaction	Particle size (nm)
Oxides		
ZnO	$ZnCl_2 + Na_2CO_3 \rightarrow ZnCO_3 + 2NaCl$	26
SnO	$SnCl_2 + Na_2CO_3 \rightarrow ZnCO_3 + 2NaCl$	28
Co_3O_4	$Co(NO_3)_2.6H_2O + 5NH_4HCO_3 \rightarrow Co_2(OH)_2CO_3 + 4NH_4NO_3 + NH_3{+}4CO_2 + 14H_2O$	13
ZrO_2	$ZrCl_4 + 4LiOH \rightarrow Zr(OH)_4 + 4LiCl$	5–10
In_2O_3	$InCl_3 + 3NaOH \rightarrow 3NaCl + In(OH)_3$	26.7
Sulphides		
ZnS	$ZnCl_2 + CaS \rightarrow ZnS + CaCl_2$	8
ZnS	$(CH_3COO)_2Zn + Na_2S \rightarrow ZnS + 2CH_3COONa$	5–18
Carbonates		
$CaCO_3$	$CaCl_2 + Na_2CO_3 \rightarrow CaCO_3 + 2NaCl$	80–140
Doped metal oxide		
Sm-CeO_2	$0.2SmCl_3 + 0.8Ce(OH)_4 + 0.66NaOH + 1.54NaOH \rightarrow Sm_{0.2}Ce_{0.8}O1.0\text{-}x + 7.26NaCl + 0.06NaOH$	10–13
Metals		
Fe	$FeCl_3 + 3Na \rightarrow Fe + NaCl$	10

between the hot vapour species and the much colder inert gas atoms. Ceramic powders are usually produced by a two-stage process: evaporation of a metal source, or preferably a metal suboxide of high vapor pressure, followed by slow oxidation to develop the desired n-ceramic powder particles. What varies in these approaches is usually the way in which energy source for vaporising the bulk material and generally encompasses laser light and electrical techniques.

Laser ablation

When an adequately powerful laser beam is focused on a material surface nanoparticles can be generated by evaporation of the targeted material into a surrounding gas where the vaporised material condenses to form nanoparticles. Here incident photon energy is high enough to break the chemical bonds of the target material directly, the material is dissociated into its chemical components and in some cases no liquid phase transition occurs. Nanoparticles of various compositions can be produced by this technique simply by changing the composition of the target material. Particle purity depends on the purity of the target material and the purity of the environment into which the material vaporises. Normally, a strategically located substrate would be used to collect these particles and it is therefore regarded as a potentially useful tool for manufacturing and depositing nanoparticles on surfaces. Typical particle sizes are in the 5–30 nm range and include a variety of metals (Ag, Au, Ni, Fe, Al, Ti, Si) alloys, oxides and nitrides [38, 39, 45]. More recent studies have examined target materials (usually a metal) immersed in liquids (water, alcohols, salt solutions). A variant on this technique is surfactant controlled laser ablation in solution (SC-LAS) [40–43]. The effect a different target material geometry such as microspheres (referred to as laser ablation of microspheres (LAM)) has also been examined [44].

As with milling, a distinction should be made between reactive and unreactive laser ablation. Reactive ablation occurs when the surrounding atmosphere or environment is reactive to the vaporised material. For example, metal oxide particles made by this method are a consequence of metal particles vaporising from a metal foil and subsequently reacting with an imposed oxygen environment. Laser ablation is currently not energy efficient as the energy absorbed for evaporation is typically about 0.1–5% of the laser energy.

Electro-explosion

In this process nanoparticles are prepared by the electro-explosion of metal wires (referred to as Electro-Explosion of Wires (EEW)), usually in non-oxidising environments. When a short (microsecond) high-current, high-voltage electric power impulse (1.5–2.0 times higher than the sublimation

energy) is applied to a metal wire, it causes it to explode. Extremely high temperatures (>20,000 °C) and cooling rates in the supersonic shock wave (10^6–10^8 °C/s) can be achieved. Nanoparticles formed through this method have unusual intrinsic thermal properties that give rise to 'self sintering'. This enhanced reactivity allows subsequent alloying to occur at substantially reduced temperatures. Initial metal clusters are ~8–10 nm with subsequent coalescence into crystallites of ~100 nm. The process can be applied to any ductile metal in wire form and products include transition and precious metals. The process can be modified to produce metal oxide and nitride powders [46–48].

Sputtering

Sputtering is a form of physical vapour deposition (PVD) often used for production of metal films, sputtering involves knocking metal atoms off a disc of pure metal with charged, energetic, chemically inactive atoms called ions produced in plasma. The metal ions will re-deposit onto a substrate the wafer to build up the desired metal film. Particle sizes ranged from 3–12 nm [49].

Spark discharge generation

In this process a metal is vaporised by charging electrodes until breakdown voltage is reached. The arc (spark) formed across the electrode vaporises a small amount of metal. Condensation and capture of generated particles is similar to that previously stated.

1.3.2 Bottom-up processes

Bottom-up processes produce nanoparticles by combination to generate material from the atomic or the molecular level. Consequently, synthesis using this approach relies heavily on chemical processes that are performed in the liquid and/or gas phase. The most common are chemical vapour deposition (CVD), sol-gel, and atomic or molecular condensation. The following section reviews some of these techniques.

Liquid phase

Precipitation/crystallisation

Wet chemical methods used to produce particles and powders have traditionally been precipitation and crystallisation. For example, large-scale production of alumina and titania, utilises these processes, as does the production of

catalytic materials such as zeolites. The majority of products in the chemical industry undergo crystallisation/precipitation at some stage of their manufacture and information related to these processes is largely proprietary. Crystallisation occurs when a solution is cooled to such an extent that it becomes supersaturated leading to nucleation and crystal formation. Precipitation, on the other hand, occurs when an insoluble product is formed as a consequence of a chemical reaction. In both cases the product usually requires filtration, washing and drying and sometimes calcination steps.

The most effort in using these two techniques for nanoparticle production has come from the pharmaceutical and catalyst production sectors by reason that increasing surface area will enhance chemical reactivity and dissolution; production infrastructure and know-how already exist for the particle formation route. Specialised products for application in both areas usually demand high purity with well-defined specifications and to meet these requirements well-controlled processes have been developed. The greatest challenge in using these nanoparticle production approaches for pharmaceutical and catalytic application is to control the size, shape and stability of the product. Consequently, other competing methods and technologies are being explored [50]. The use of micro-reactor technology for its potential to control product specifications such as particle size is of increasing interest [51].

Sol-gel

In simple terms sol-gel processing involves the transition from a sol (very finely divided solid particles dispersed in a liquid which will not settle out) to a gel (a dispersion of a liquid throughout a solid matrix). It is a long-established process with the key controlling processing steps being sol formation and gelling, drying, curing and sintering. Starting materials used in the preparation of the sol are usually inorganic metal salts or metal organic compounds such as metal alkoxides. Typically, the precursor is subjected to a series of hydrolysis and polymerisation reactions to form a colloidal suspension. This suspension can be further processed to make materials in different forms from thin films and aerogels to nanocomposites and nanoparticles. The advantages of the process are perceived to be (i) low processing temperatures (ii) high homogeneity and purity of products and (iii) versatility with respect to the materials that can be formed. Typical sizes of nanoparticles produced by this method are 5–30 nm [52, 53]. In particular sol gel processing offers unique opportunities to generate gels of organic-inorganic materials [54].

The majority of sol-gel focus has been on producing glass ceramic powders, coatings and thin films. However, the recognition that particles of high purity, controlled particle size and new composition may be produced, and these benefits combined with the ease by which they may be processed, has increased

interest in this technique for nanoparticle production. Particles of interest range from alumina, zirconia, titania and silica to titanates, phosphors, pigments and encapsulated pigments, optical materials, ferroelectrics and catalyst supports.

The new generation of sol-gel produced hybrid materials was pioneered by the preparation of non-crystalline organic–inorganic materials that were named Ormocers. The physical and mechanical properties of these materials can be tailored by manipulating the ratio of inorganics–organics which when applied on a substrate in gel form and dried give rise to different surface morphologies. Resultant 'super-repellent' surfaces are based on a physical structure rather than chemical character [55].

Another new group of materials that have been spawned by this preparation method are the biological composites. The demonstration that an organic dye could be entrapped in a porous silica cage led to the greater aspiration of incorporating biologically active materials in inorganic matrices to control reactions arising from the transport of materials in or out of the matrix [56].

In forming sols it is desirable to achieve a homogeneous colloid matrix. Material properties (including particle size) may be controlled by pre-reacting the slower reactive components or chelating the faster components to achieve homogeneity. Mechanical treatments and the use of ultrasound to generate a more homogeneous distribution of the nucleation sites for nanoparticles have been used. The drying stage is perhaps the most critical in influencing the properties of the final product (especially in catalyst supports, films and coatings). Past difficulties with shrinkage and cracking of coatings have to a significant degree been overcome with the use of drying control chemical additives (such as formamide) [52].

Self-assembly

Self-assembly of functional nanomaterials has received a substantial amount of attention over the last decade. It is a predominantly chemical process rather than a specific technique. A wide range of materials have been processed using self-assembly techniques. Production methods are analogous to nanoparticle production in that they are typified by two approaches (i) top-down and (ii) bottom-up. In the self-assembly context, however, the process is usually regarded as one of fabrication, on the nanometric scale, of architectured material systems (usually in two and three dimensions). The top-down approach is one where components are assembled by direct delivery to specific locations whereas the bottom-up approach directs the assembly from the bottom up usually by using a molecular template. The latter has received the most attention as it represents an opportunity to fabricate structures with precision and versatile functionality with the promise of cheaper production costs than those conventionally associated with top-down approaches.

The inspiration for self-assembly comes from nature itself and it is no coincidence that a major focus of nanoparticle self-assembly (as distinct from nanostructures or nanostructured materials) is on particles associated with biological systems including lipids, organic molecules, peptides, proteins and DNA. However, peptides and proteins are used as the main building blocks because their folding behaviour and stability is relatively well understood together with the influences on protein–protein interactions [57–61].

A simple yet elegant example of where the top-down and bottom-up approaches converge where nanometric layers of gold were successively built up on a glass substrate to a film thickness up to 2 μm. By functionalising the glass and the gold surfaces respectively with thiol and a mix of alkyl thiols and carboxylate thiol groups, the gold clusters were cross-linked by Cu^{2+} into a network and simultaneously attached to the glass substrate. By repeating these steps a multilayer film was fabricated [62].

Vapour phase

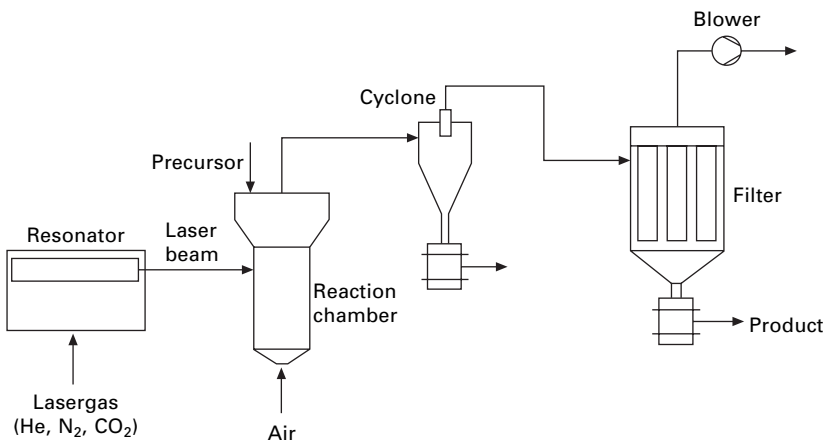
For general nanoparticles, bottom-up vapour phase processes (as distinct from top-down vapour phase processes as previously described) harness fluid mechanics and particle dynamics within a short time period (milliseconds) during which key formation mechanisms occur. In general the controlling formation mechanisms are (i) chemical reaction (ii) nucleation and (iii) growth. In these processes a solution (usually saturated or supersatured) containing specific but reactive precursors is heated so that chemical reaction (either between species or within a species, e.g., decomposition) occurs to produce a non-volatile phase. Subsequent rapid cooling of the vapour is imposed so as to control particle size, agglomeration and in some cases sintering. In controlling these reaction process parameters, a substantial degree of flexibility on controlling particle properties can be achieved and particle sizes ranging from a few nanometres to several tens of nanometres are synthesised. Control of the factors that affect the particle properties are reactor temperature profile, residence time and reactant concentration as well as mixing of reactants with carrier gases. The size of nanoparticles synthesised from the gas phase is the size that particles grow to before agglomeration occurs. Both temperature and concentration will promote agglomeration. Increasing the partial pressure reactant will necessitate a shorter residence time to produce a particular particle size and will increase the production rate.

In engineering this process for larger-scale production, various reactor configurations for controlled heating and cooling can be utilised depending on the type and form of targeted product. These range from furnaces (for producing saturated vapours for substances having a high vapour pressure, temperatures up to ~1700 °C), lasers (that selectively heat source molecules by absorption but not the surrounding (gas) phase), flame reactors (usually

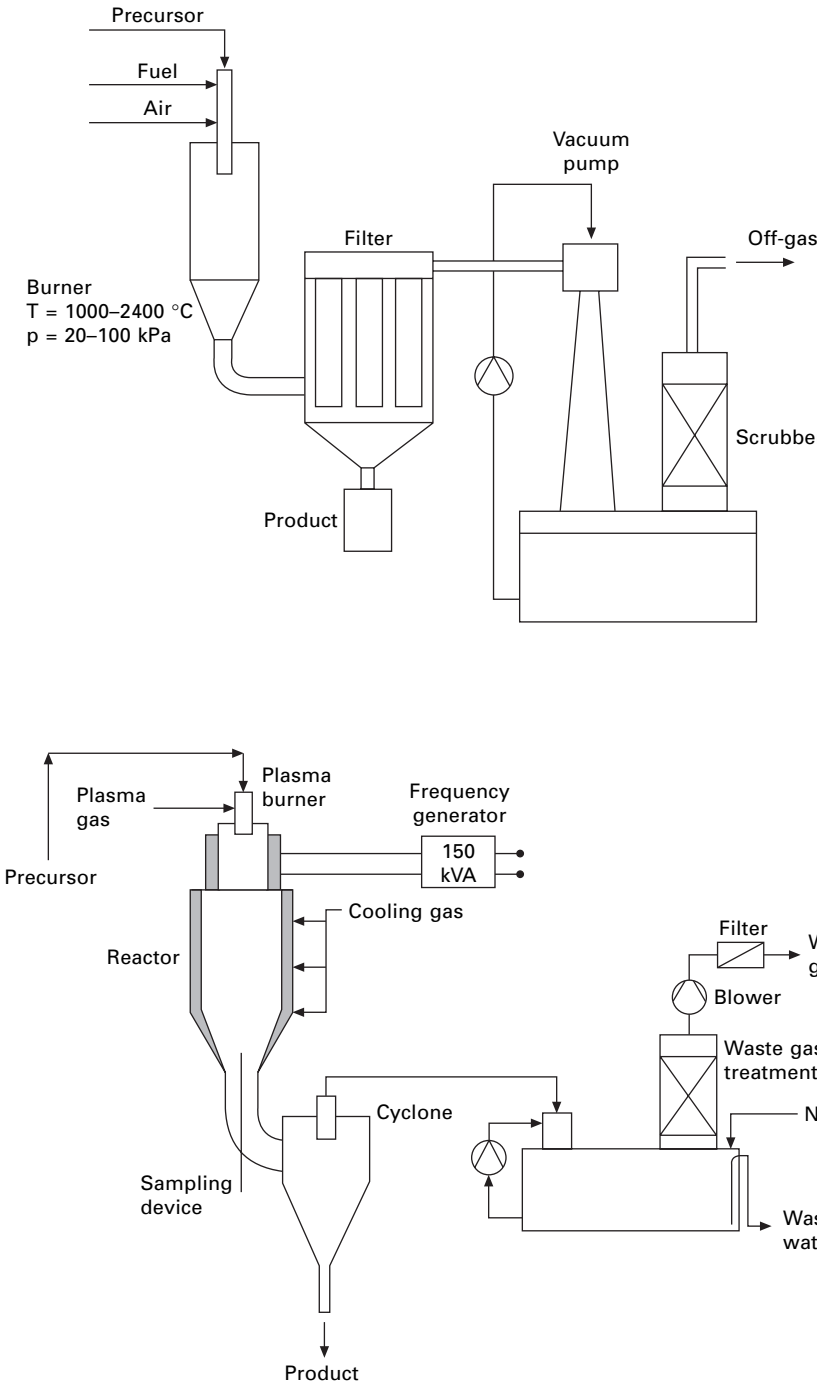
low pressure and temperatures up to ~ 2000 °C), plasma reactors (thermal or microwave, temperatures ~ 10,000 °C or 300–900 °C respectively), electrospray (where a fine aerosol is produced by electrostatic charging) and spray pyrolysis systems (usually occurring under pressure and at temperatures above 450 °C). Cooling may occur naturally (convective flow), by dilution (with an inert cooling medium) or expansion [63, 64]. Some of these are illustrated in Fig. 1.6.

The real advantage of vapour phase techniques is that they can produce nanoparticles of uniform size, shape and purity by a continuous process that is scalable. By supplying additional precursors (either simultaneously or sequentially) more complex nanoparticles may be produced (e.g., doped or coated nanoparticles). A wide range of materials have been synthesised using this approach, e.g., oxides from Al, Zr, Ce, Ti, Zn, Si, etc., metals (e.g., Fe, Co, Mo, Ni, Cu), nitrides (e.g., AlN), borides, carbides, complex oxides (e.g., $BaTiO_3$), doped oxides (e.g., Mn doped ZnO, Al doped TiO_2) and coated particles (e.g., ZrO_2 coated with Al_2O_3). Appropriately chosen precursors are used which are usually compounds such as metal chorides and organo-metallics [65–67].

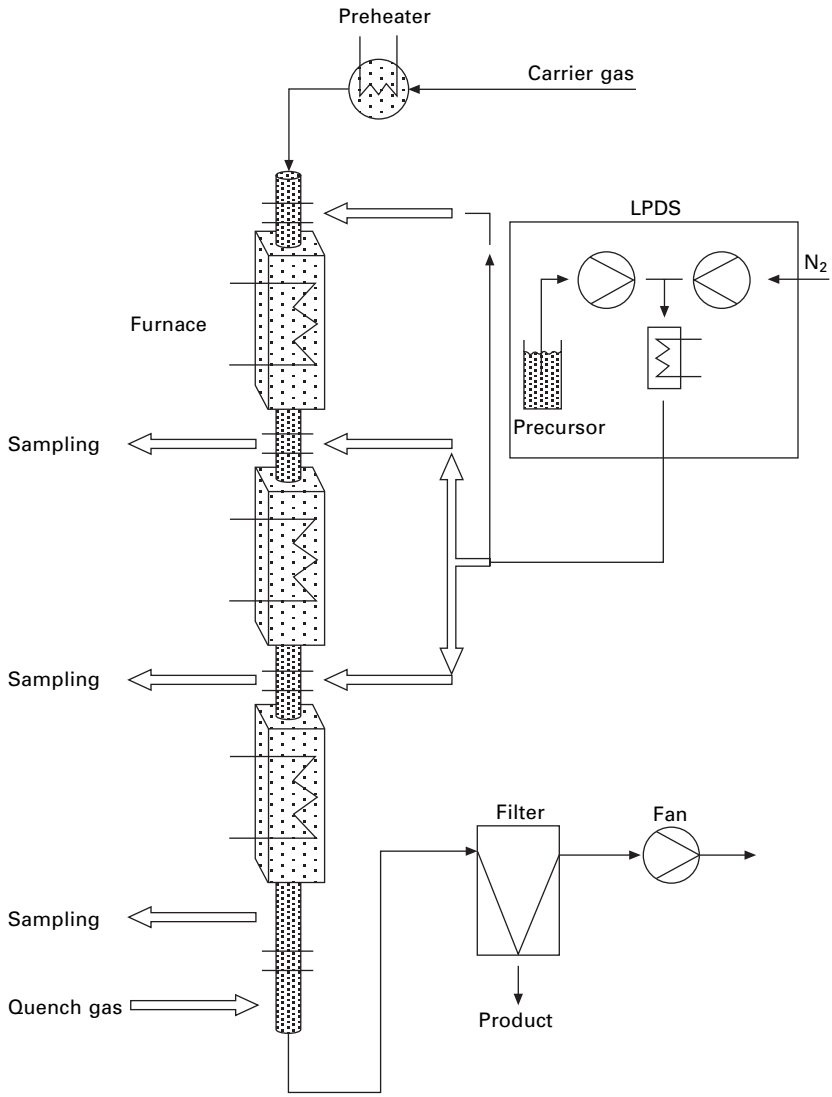
Perhaps the most significant advantage that vapour phase chemical processing has over other techniques is that it can be used (in combination with solids) to synthesise particles of complex structures. The most significant examples of this are carbon nanotubes (NT). Carbon nanotube synthesis in many ways is analogous to heterogeneous catalysis in that a reactant gas such as methane is decomposed in the presence of a catalyst (generally transition metals such as Fe, Co, Ni and their alloys) to carbon that grows in the form of nanotubes [68].



1.6 Four illustrations of reactor configurations (vapour phase nanoparticle production) [65].



1.6 (Continued)

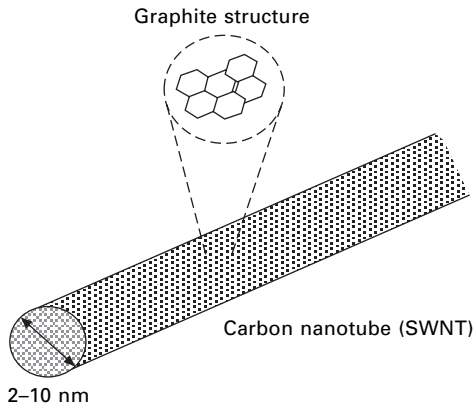


1.6 (Continued)

1.4 Examples of commercial production

1.4.1 Carbon nanotube production

Carbon nanotubes (sometimes called buckytubes) are molecules made from carbon atoms equivalent to a flat graphite sheet formed into a tube. Carbon nanotubes need a special mention because they are relatively new materials with extraordinary mechanical and electrical properties (Fig. 1.7). They



1.7 Illustration of a single-walled carbon nanotube (SWNT).

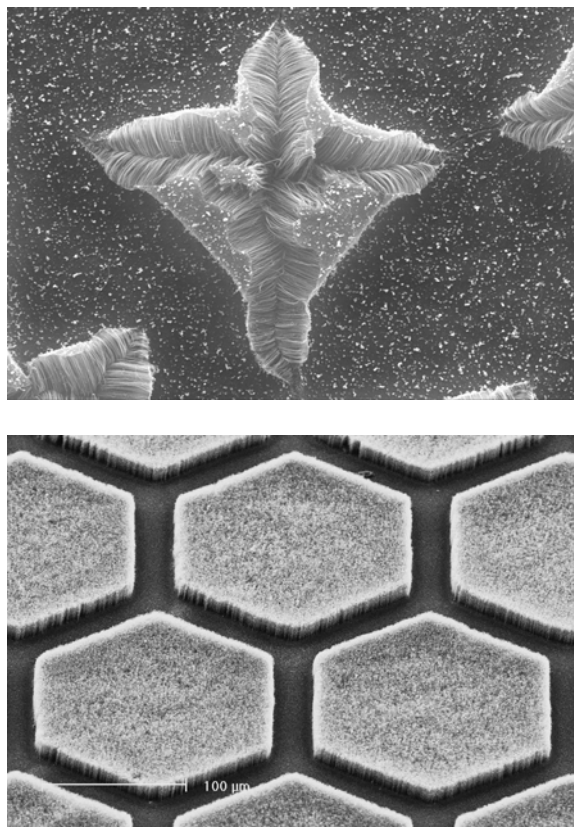
consequently have received increasing attention both commercially and developmentally. They are not a single material but vary enormously in structure, properties and ease of production. They can be long or short, open or closed, single- (SWNT), double- (DWNT) or multi-walled (MWNT). Related, but different, materials are nanofibres, nanohorns and nanorods. Depending on the production process employed, nanotube synthesis may be regarded either as a top-down or bottom-up, self-assembly process.

Commercially, carbon nanotubes (NT) are generally supplied as single (SWNT) or multiwalled (MWNT) nanotubes. There are several commercial producers throughout the world and their capacity ranges from grams to kilograms. Production routes can generally be summarised as follows (i) arc discharge and (ii) chemical vapour deposition including catalytic CVD. The type of tube produced (SW or MWNT) strongly depends on whether a catalyst is present. Where a catalyst is present, SWNT is the dominant product [69–71, 73].

Commercial processes use all of the above production routes in one form or another. SWNT are generally made either by arc discharge or catalytic-chemical vapour deposition whilst MWNT are predominantly produced by vapour deposition techniques. There are approximately 40–50 producers globally who together have an annual production capacity for SWNT of ~10–15 tons and 100–120 tons for MWNT. The US is the main producer of both nanotube forms with China and Japan being the other significant producers. Purity of product is generally reported to be >95% for MWNT but ranges from ~ 40–90% for SWNT. A qualification on that purity is often reported on a carbon and not a structural basis. The main non-carbon impurity is the catalyst itself. Dimensions typically range from a few nanometres in nanotube diameter to microns and centimetres in length. More recently claims have been made for fabricated tube fibre lengths up to one hundred metres. Prices

range from the order of tens US$/gram for MWNT for several hundreds US$/gram for SWNT [18, 74, 75].

Catalytic decomposition of a carbon containing reaction gas (acetylene, methane, xylene, etc.) to produce nanotubes represents the most significant commercial process primarily because tubes can be grown using this method in larger quantities and at low temperatures. This, however, results in lower structural perfection with poorer graphitisation of the tube walls than in other methods. An additional advantage (and perhaps a disadvantage) is that the nanotubes may be selectively grown in a predetermined manner, for example, the catalyst (e.g., Fe, Co or Ni) may be located on a substrate prior to growth enabling the growth of structured 'nanobrushes' consisting of vertically aligned nanotubes at high density (see Fig. 1.8 for some examples of how nanotubes can be grown in ordered patterns using pre-designed catalyst templates). Some recent focus has been on trying to perform the process



1.8 Examples of carbon nanotubes catalytically grown in ordered predetermined patterns (photo courtesy of Commonwealth Scientific and Industrial Research Organisation CSIRO, Australia).

continuously as, for example, in a fluidised bed [76]. At low temperatures (700 °C), the distribution of nanotube diameter appears to correlate with original particle size distribution but at higher temperatures (800 °C) diameter seems to be independent of particle size [72].

Catalytic disproportion of carbon monoxide is commercially practised to produce SWNT with high selectivity at temperatures between 700–950 °C, at elevated pressures (1–10 atm) using a Co and Mo catalyst. Catalysts such as Pd, Pt and Cr can reduced required temperatures to ~ 550 °C [77].

Non-catalytic, processes employ thermally driven decompositions of hydrocarbon liquids such as solvents like benzene, toluene, alkanes, alcohols and/or mixtures thereof. This has been done by self-regulated low-voltage contact arc discharge to produce MWNTs and is a controlled liquid pyrolysis [78].

1.4.2 Metal oxide production (ZnO)

The commercial production of nano ZnO is currently achieved by two main methods that represent 'top-down' and 'bottom-up' approaches. These are physical vapour synthesis (PVS) and mechanochemical processing (MCP^{TM}).

Mechanochemical processing (MCP^{TM})

MCP^{TM} technology, the basis of which was previously described on page 10, is a solid-state process that can be used for the manufacture of a range of nanopowders including ZnO as is practised by Advanced Powder Technologies (APT). Dry milling is used to induce chemical-exchange reactions through ball-powder collisions that result in nanoparticles formed within a salt matrix. For ZnO production the ball mill reaction is

$$ZnCl_2 + Na_2CO_3 \rightarrow ZnCO_3 + 2NaCl$$

followed by a dissolution step to remove the NaCl, drying and calcination of the $ZnCO_3$ to generate the oxide.

Particle size is defined by the chemistry of the reactant mix, milling and heat treatment conditions. Particle agglomeration during milling is minimised by the salt matrix, which is then removed, prior to calcination, by a simple washing procedure. The ball mill acts as a low temperature chemical reactor where the reaction process results from local heat and pressure at contact surfaces and occurs at the nanoscale. Reported product sizes are 25 to 30 nm [79].

Physical vapour synthesis (PVS)

In the PVS process, arc energy is first applied to a solid precursor (typically

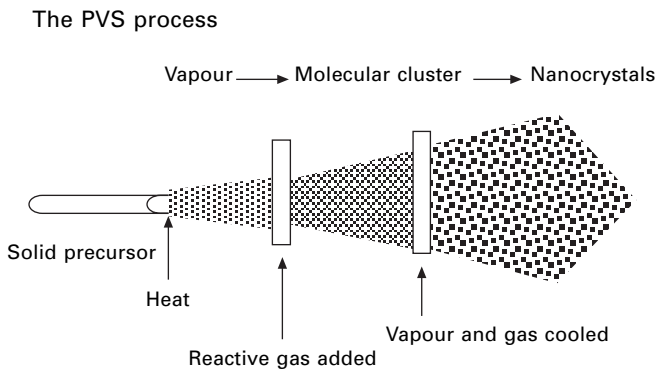
metal) in order to generate a vapour at high temperature (Fig. 1.9). A reactant gas is then added to the vapour, which is then cooled at a controlled rate and condenses to form nanoparticles. The nanomaterials produced by the PVS process consist of discrete, fully dense particles of defined crystallinity. This method typically produces particles with average sizes ranging from 8–75 nm. Nanophase Technologies (USA) uses the PVS process in the commercial scale production of NanoGard® zinc oxide and NanoTek® aluminum oxide. In addition, this process has been used to generate other materials and compounds such as a variety of doped zinc oxides, selected rare earth and transition metal oxides, and transparent conductive oxides such as antimony-tin oxide and indium-tin oxide [80].

Chemical routes

There are other ZnO production processes that use chemical routes [81]. An example is where ZnO is dissolved in a solution of ammonia/ammonium bicarbonate. The solution is then filtered and the ammonia and carbon dioxide evaporated off and recycled. A crystalline zinc subcarbonate is formed which is washed and filtered before being dried and calcined to produce nanophase ZnO.

Pyrogenic processing

Pyrogenic processing is the formation of particles by means of flame oxidation of metals, metalloids or their derivates in the gas phase. The flame with temperatures from 1000 °C up to 2400 °C provides the energy to evaporate the precursors and to drive the chemical reactions. With extremely short residence times of 10 to 100 ms in the highest temperature region, this zone is crucial for the formation of the primary particles. Various pyrogenic



1.9 Illustration of physical vapour synthesis (PVS).

processing techniques to synthesise a variety of nano-scaled particles have been employed to produce ZnO, TiO_2 and ZrO_2 [82].

1.5 Applications

Having made nanoparticles the question arises where are they used? The range of applications is broad and growing with the current main uses as functional additives or precursors for emulsions, composites and coatings. The worldwide market for polymer-nanocomposites in 2003, for example, was 90m US$ with expected growth to 211m $ by 2008 [83]. The polymer properties that show substantial performance improvements include: mechanical properties (e.g., strength, modulus and dimensional stability), decreased permeability (to gases, water and hydrocarbons), thermal and UV stability and heat distortion temperature, flame retardancy and reduced smoke emissions, chemical resistance, surface appearance, electrical conductivity and optical clarity and increased resistance to solar degradation in comparison to conventionally filled polymers.

Main application areas of nanoparticles are as additives to polymers used in the transport (automotive and aerospace) sector (vehicle parts for lighter weight and higher performance), packaging (including food and biomedical) to protect and preserve the integrity of the product by controlling the barrier, mechanical, optical and respiration properties, textiles (increased strength, water resistance, self cleaning, fade resistance) and personal care products (UV protection, deep penetration skin cream emulsions). Many of these functionalities can be interchanged from one application to another. For example, similar technology used for transparent UV protective coatings such as sunscreens in personal care products can be used for UV protection in food packaging, paints, textiles, plastics used in outdoor use and protection of wood without altering the optical properties.

In the communications and information technology sectors, nanoparticles are used for increasing the efficiency of electronics by increasing the information storage capacity whilst reducing the size and weight of devices and components. Additionally, dispersions of nanoparticles in different matrices are used for chemical–mechanical planarisation (CMP) of hard drives and high surface area carbons are used in energy storage devices such as supercapacitors [84]. Other important areas include inks and ink printable electronic circuitry. The paper industry has employed nanoparticle technology for improving fillers and coatings [85].

A current status overview of a selection of present and future applications using nanoparticles is presented in Table 1.5 [86]. As can be seen from this overview, nanoparticles are being designed and delivered into a broad and ever-increasing range of applications.

Table 1.5 A selection of current and future applications of nanoparticles [86]

Area	Under development	On the market	Well established
Power/energy	Ni and metal hydrides for batteries Dye sensitised solar cells using TiO_2 Hydrogen storage using metal hydrides Improved anode and cathode materials for solid oxide fuel cells Thermal control fluid using Cu	Environmental catalysts, ceria in diesel	Automotive catalysts
Healthcare/ medical	Nanocrystalline drugs for easier adsorption Inhalable insulin Nanospheres for inhaling drugs currently injected using biocompatible Si Bone growth promoters Virus detection using quantum dots Anti-cancer treatments Coatings for implants such as hydroxyapatite	Sunscreens using ZnO and TiO_2 Molecular tagging: quantum dots, CdSe Carriers for drugs with low water solubility	Ag antibacterial wound dressings ZnO fungicide Au for biolabelling and detection MRI contrast agents using superparamagnetic iron oxide
Engineering	Coating tool bits: WC, TaC, TiC, Co Spark plugs using nanoscale metal and ceramic powders Nanoporous silica-based on aerogels for high efficiency insulators Controlled delivery of herbicides and pesticides Chemical sensors Molecular sieves	Abrasion-resistant coatings using alumina, Y-Zr_2O_3 Nanoclay-reinforced polymer composites Lubricant/hydraulic additives: $CuMoS_2$ Pigments Improved scratch-resistant coatings Self-cleaning glass using TiO_2 Propellants using Al	Structural enhancement of polymers and composites Thermal spray coatings based on TiO_2, TiC-Co, etc. Inks: conducting, magnetic, etc. using metal powders

Table 1.5 Continued

Area	Under development	On the market	Well established
Consumer goods		Anti-counterfeit devices	Packaging using silicates Ski wax White goods Glass coatings for anti-glare, anti-misting mirrors using TiO_2 Sports goods: tennis balls, rackets using nanoclays Water anti-stain repellent textiles
Environmental		Alumina fibres for water treatment Self-cleaning glass using TiO_2-based nanostructured coatings Photo-catalyst water treatments using TiO_2 Anti-reflection coatings	Tiles coated using alumina and others Sanitary ware Soil remediation using Fe
Electronics	Nanoscale magnetic particles for high-density data storage EMI shielding using conducting and magnetic materials Electronic circuits using Cu, Al Display technologies including field-emission devices using conducting oxides	Ferro-fluids using magnetic materials Optoelectronic devices such as switches using rare-earth-doped ceramics Conductive coatings and fabrics using rare-earth-doped ceramics	CMP alumina, ceria Coatings and joining materials for optical fibres based in Si

1.6 Future challenges of nanoparticle production

The ultimate challenge for all processes involving nanoparticle production is one of realising precision particle and process engineering at an affordable price. The associated technical challenges are not trivial. As the complexity of particle composition and the demand for new or enhanced functionality together with the precise control of the particle properties increases so will the need for engineering systems to accommodate this. Control of the physics and chemistry at the nanolevel on a large scale may never be possible without the help of the materials themselves. Nanoparticle size, shape and morphology can vary within the same product. For example, ZnO can be synthesised under controlled conditions by thermal evaporation of solid powders to form a myriad of shapes and morphologies from spheres to rods to rings, springs, spirals and coils, propellers and trees. Whilst this is impressive, the corollary from a precision engineering perspective is that without total system control, there may well be numerous ways of making an undesired product using the same generic process.

The interfacial phenomena that are key controllers in determining process efficiency are influenced by both extrinsic and intrinsic factors. It is therefore not unreasonable to suspect that the development of large-scale processes will require precision from both these factors. Whilst engineering is becoming more precise in its control at the macro level the challenge for the future will be to adequately and consistently control particle properties at the nano level.

1.7 Useful websites

nanotechwire.com
www.ingenta.com
www.gii.co.jp
www.buscom.com
/www.the-infoshop.com
www.nanotech-now.com
nanovip.com
www.cientifica.com
www.solgel.com
www.forbes.com
www.bccresearch.com
www.degussa.com
www.atomate.com
www.nanophase.com
www.azom.com

1.8 References

1. *Nanoscience and Nanotechnologies: Opportunities and Uncertainties*. The Royal Society and The Royal Academy of Engineering, July 2004.
2. *New Dimensions for Manufacturing: A UK Strategy for Nanotechnology*. Taylor, J. M. Department for Trade and Industry (dti), June 2002.
3. Links provided by the Nanotechnology Business Directory (Nanovip.com).
4. Nanoparticles as building blocks? Pitkethly, M. J. *Nanotoday* December 2003, pp 36–42.
5. Technological Analysis. Industrial Application of Nanomaterials – Chances and risks. Luther, W. *Future Technologies* No. 54 Dusseldorf, August 2004. ISSN 1436-5928.
6. Nanoparticles in the Making, Yadav, T. *Chemical Engineering Progress*, AIChE, pp 40S-41S, November (2003).
7. *Bottom up Production Technologies*. Holister, P., Vas, C. R. and Harper. T., Cientifica. Technology White Paper No. 15 October 2003.
8. *Nanoparticles*. Holister, P., Vas, C. R. and Harper, T. Technology Cientifica. Technology White Paper No. 3 October 2003.
9. Nanoparticles. Building Blocks for Nanotechnology. Edited by Ritello, V. *Nanostructure Science and Technology*, Kluwer Academic/Plenum 2004.
10. *Nanoparticles. From Theory to Application*. Edited by Gunter Schmid. Wiley-VCH. 2004-10-09.
11. Morphology Control of materials and Nanoparticles Advanced Materials Processing and Characterisation edited by Waseda, Y. and Muramatsu, A. Springer. *Materials Science*, 2003.
12. Size- and support-dependency in the catalysis of gold. Haruta, M. *Catal. Today*, 1997, 36, pp 153–166.
13. Gadolinium-loaded nanoparticles engineered from microemulsion templates. Oyewumi, M. O. and Mumper, R. J. *Drug Dev Ind Pharm*. 2002 Mar 28(3): 317–28.
14. Superhard silicon nanospheres. Gerbericha, W. *et al*. *Journal of the Mechanics and Physics of Solids* 51 (2003) 979–992.
15. *Gold nanoparticle thin films Colloids and Surfaces A: Physicochemical and Engineering Aspects* 202 (2002) 119–126. Ung, T., Liz-Marza, L. and Mulvaney, P.
16. Aiyer, H. N., Vijayakrishnan, V., Subanna, G. N. and Rao, C. N. R. *Surf. Sci*. 1994, 313, 392.
17. Size-Dependent Chemistry: Properties of Nanocrystals. Rao, C. N. R., Kulkarni, G. U., Thomas, P. J. and Edwards P. P. *Chem. Eur. J*. 2002, 8, No. 1 pp 28–35.
18. *The International Nanotechnology Business Directory* (Nanovip.com).
19. Processing of advanced materials using high-energy mechanical milling. Zhang. D. L. *Progress in Materials Science* 49 (2004) pp 537–560.
20. Synthesis of Nanostructured Materials by Mechanical Milling: Problems and Opportunities. Koch, C. C. *Nanostructured Materials*, 1997, Vol. 9, pp 13–22.
21. Determining granule strength as a function of moisture content. Verkoeijen, D., Meesters, G., Vercoulen, P. and Scarlett, B. *Powder Technology*. Vol. 124. Issue 3, April 2002, pp 195–200.
22. Variables affecting the fine grinding of minerals using stirred mills. Jankovic, A. *Minerals Engineering* 16 (2003) 337–345.
23. Structural and thermodynamic properties of heavily mechanically deformed Ru and

AlRu. Hellstern, E., Fecht, H. J., Fu, Z. and Johnson, W. L. *J. Appl. Physics*. Vol. 65, Issue 1, Jan, 1989, pp 305–310.

24. Effects of high energy ball milling on ceramic oxides. Gaffet, E., Michel, D., Mazerolles, L. and Berthet, P. *Materials Science Forum*. Vols 235–238 (1997) pp 103–108.
25. Ball milled ZnO for varistor applications. Boily, S., *et al*. *Materials Science Forum*. Vols 235–238 (1997) pp 993–998.
26. The synthesis of TiN by ball milling – a neutron diffraction study. Campbell, S., Hofmann, M. and Calka, A. *Physica B* 276–278 (2000) pp 899–900.
27. Formation and environmental stability of nanocrystalline and amorphous hydrides in the 2Mg-Fe mixture processed by controlled reactive mechanical alloying (CRMA). Varin, R., Li, S., Calka, A. and Wexler, D. *Journal of Alloys and Compounds* 373 (2004) pp 270–286.
28. Formation of Ni Fe_2O_4 nanoparticles by mechanochemical reaction. Huaming Yang *et al. Materials Research Bulletin* 39 (2004), pp 833–837.
29. ZnO nanoparticles synthesised by mechanochemical processing. Tsuzuki, T. and McCormick, P. G. *Scripta Mater*. 44 (2001) pp 1731–1734.
30. Synthesis of tin oxide nanoparticles by mechanochemical reaction. Huaming Yang *et al. Journal of Alloys and Compounds* 363 (2004) pp 271–274.
31. Mechanochemical synthesis of cobalt oxide nanoparticles. Huaming *et al*. *Materials Letters* 58 (2004) pp 387–389.
32. Synthesis of nanocrystalline ZrO_2 powders by mechanochemical reaction of $ZrCl_4$ with LiOH. Dodd, A. and McCormick, P. G. *Journal of European Ceramic Society* 22 (2002) pp 1823–1829.
33. In_2O_3 nanoparticles synthesised by mechanochemical processing. Huaming *et al*. *Scripta Materialia*, 50 (2004) pp 413–415.
34. Mechanochemical Synthesis of metal sulphide nanoparticles. Tsuzuki, T. and McCormick, P. G. *Nanostrucured Materials* (1999) Vol. 12, pp 75–78.
35. Mechanochemical Route for sulphide nanoparticles preparation. Balaz, P., Boldizarova, E., Godocikova, E. and Braincin, J. *Materials Letters* 57 (2003) pp 1585–1589.
36. Mechanochemical synthesis and characterisation of nanoparticluate samarium-doped cerium oxide. Hos, J. and McCormick, P. G. *Scripta Materialia* 48 (2003) pp 85–90.
37. Mechanochemical Synthesis of ultrafine Fe Powder. Ding, J., Miao, W. F., McCormick, P. G. and Street, R. *Applied Physics Letters*, Vol. 67, Issue 25, (1995) pp 3804–3806.
38. Design of metal nanoparticle synthesis by vapour flow condensation. Wegner, K. Walker, B., Tsantilis, S. and Prastinins, S. *Chemical Engineering Science* 57 (2002) pp 1753–1762.
39. Nanoparticle formation by laser ablation. Ulmann, M., Friedlander, S. and Schmidt-Ott A. *Journal of Nanoparticle Research* 4 (2002) pp 499–509.
40. Nanoparticles produced by laser ablation of solids in a liquid environment. Simakin, A. V., Voronov, V. V., Kirichenko, N. A. and Shafee, V. G. A. *Appl. Phys. A* 79 (2004) pp 1127–1132.
41. Nanoparticles produced by laser ablation of solids in liquid environment. Dolgaev, *et al. Applied Surface Science* 186 (2002) 546–551.
42. Formation of silver nanoparticles by laser ablation of a silver target in NaCL solution. Seung Min Park *et. al. Applied Surface Science* 197–198 (2002) pp 628–634.
43. Metal nanoparticles generated by laser ablation. Becker, M. F., Glitcksmann, H. D., *et al. Nanostructured Materials* Vol. 10. No. 5 (1998) pp 853–863.

44. Generation of Metal Nanoparticles by Laser Ablation of Microspheres. Cai, H., Chaudhary, N., Lee, J., Becker, M. F., Brock, J. R. and Keto, J. W. *J Aerosol Sci*. Vol. 29. No. 5/6 (1997) pp 627–636.
45. Formation of small gold clusters in solution by laser excitation of interband transition. Fumitaka Mafune, Tamotsu Kondow. *Chemical Physics Letters* 372 (2003) pp 199–204.
46. Preparation of Cu, Ag, Fe, and Al nanoparticles by the exploding wire technique. Sen, P. *et al. Proc. Indian Acad. Sci*. (Chem Sci), Vol. 115, Nos 5&6 (2003) pp 499–508.
47. Nanosized Alumina Fibers. Tepper, F., Lerner, M. and Ginley, D. *American Ceramic Society Bulletin*, Vol. 80, No. 6 pp 57–60 (2001).
48. Active Metal powders that burn efficiently and quickly-obtd. By electro-explosion of wires composed of e.g. aluminium in non oxidising atmos. E.g. argon to produce micron sized powder. Ivanov, G .V. Patent EP718061-A.
49. Fabrication of Au nanoparticles by radiofrequency magnetron sputtering. Terauchi, S., Koshizaki, N. and Umehara, H. *Nanostructured Materials*. Vol. 5. No. 1. (1995) pp 71–78.
50. Nanoparticles in Catalysis, in *Nanoparticles: Building Blocks for Nanotechnology* Zhong, C. J., Maye, M. M., Luo, J., Han, L. and Kariuki, N. N., ed. by Rotello, V. M. Kluwer Academic Publishers. Chapter 5, pp 113–144, 2004.
51. 'Numbering Up' Small Reactors. Microreactor technology offers many benefits for process development and production. Michael Freemantle, *C&EN News*. June 16, 2003. Volume 81, Number 24, pp 36–37.
52. Nanomaterials. Top-to-bottom functional design. Walt, D. *Nature Materials*. Vol. 1, Sept. 2002 pp 17–18.
53. Past and Present of Sol-Gel Science and Technology. Zarzycki, J. *Journal of Sol-Gel Science and Technology*, 8, (1997) pp 17–22.
54. Sol-Gel technology as Reflected in the Journal of Sol-Gel Science and Technology. Sakka, S. *Journal of Sol-Gel Science and Technology*, 26, (2003) pp 29–33.
55. The Sol-gel process as a basic technology for nanoparticle-dispersed Inorganic-organic composites. Scmidt, H., Jonschker, S., Goedicke, S. and Mennig, M. *Journal of Sol-Gel Science and Technology*, 19, (2000) pp 39–51.
56. Sol-Gel Science and Technology: Current State and Future Prospects. Uhlmann, D. R. and Teowee, G. *Journal of Sol-Gel Science and Technology*, 13, (1998) pp 153–162.
57. Sol-Gel research – Achievements since 1981 and Prospects for the Future. Mackenzie, J. D. *Journal of Sol-Gel Science and Technology*, 26, (2003) pp 23–27.
58. Inorganic Self-Assembly. Depero, L. and Curri, M. L. *Current Opinion in Solid State and Materials Science*. 8 (2004) pp 103–109.
59. Physical Properties of self-assembled nanosized cobalt particles. Petit, C., Pileni, M. P. *Applied Surface Science* 162–163 (2000) pp 519–528.
60. The Fabrication and organization of self-assembled metallic nanoparticles formed in reverse micelles. Seip, C. and O'Connor, C. J. *Nanostructured Materials* Vol. 12 (1999) pp 183–186.
61. The role of soft colloidal templates in controlling the size and shape of inorganic crystals. Pileni, M. P. *Nature Materials*. Vol. 2, (2003) pp 145–150.
62. Self-assembly as a route to fibrous materials: concepts, opportunities and challenges. Viney, C. *Current Opinion in Solid State and Materials Science*. Vol. 8. (2004) pp 95–101

63. Gas Phase Production of Nanoparticles. Gutsch, A., Kramer, M., Michael, G., Muhlenweg, H., Pridohl, M. and Zimmerman, G. *Kona*. No. 20 (2002) pp 24–37.
64. Vapor Phase synthesis of Nanoparticles. Swihart, M. *Current Opinion in Colloid and Interface Science* Vol. 8 (2003) pp 127–133.
65. Synthesis of nanoparticles in the gas phase for electronic, optical and magnetic applications – A review. Kruis *et al. J. aerosol Sci*. Vol. 29 No. 5/6 (1998) pp 511–535.
66. Nanoparticle Synthesis at high production rates by flame pyrolysis. Mueller, R., Madler, L. and Prastinis, S. *Chemical Engineering Science*. 58 (2003) pp 1969–1976.
67. Gas Phase synthesis of nanocrystalline materials. Hahn, H. *Nanostructured Materials*, Vol. 9. (1997) pp 3–12.
68. Catalytic synthesized carbon nanostrucures from methane using nanocrystalline Ni. Zongquan, Li and Jiuling Chen *et al. Carbon* 40 (2002) pp 409–415.
69. Carbon nanotubes: dynamics and synthesis processes. Laplaze, D., Alvarez, L., Guillard, T., Badie, J. M. and Flamant, G. *Carbon* Vol. 40 Issue 10 (2002) pp 1621–1634.
70. Nanotubes, Berholc, J., Roland, C. and Yakobson, B. *Current Opinion in Solid State and Materials Science*. Vol. 2 (1997) pp 706–715.
71. New direction in nanotube science. Terrones, M. *et al. Materials Today*. Vol. 7 Issue 9, (2004) pp 30–45.
72. Correlation between metal catalyst particle size and carbon nanotube growth. Kukovitsky, E. F. *et al. Chemical Physics Letters* 355 (2002) pp 497–503.
73. Carbon Nanotubes: A small-scale wonder. Shelley, S. *Chemical Engineering* (Feb, 2003).
74. Nanotubes, Holister, P., Harper, T. and Roman, C. *Cientifica*. March 2004.
75. Record nanofibre teased from smokey suspension. Hogan, J. *New Scientist*. March 2004 p. 25.
76. Growth of carbon nanotubes by fluidized bed catalytic chemical vapour deposition. Venegoni, D., Serp, P., Feurer, R., Kihn, Y., Vahlas, C. and Kalck, P. J. *Phys. IV France* 12 (2002) pp 4–93.
77. SouthWest Technologies Website. www.swnano.com/technology
78. Rosseter Holdings Website www.e-nanoscience.com/process
79. Advanced Powder Technology website www.apt-powders.com
80. Nanophase Technologies website www.nanophase.com
81. Patent CN1210813A Shanxi Four High Nanotechnology Co. (www.fhnm.com).
82. Information from Degussa website www.advanced-nano.com
83. www.plasticsnews.com. Source: Business communication Co. Inc. Norwalk. Conn. March, 2004.
84. Cap-XX website www.cap-xx.com
85. Nanovation Invasion. How real is it? Martin Koepenick. *Pulp and Paper Canada*. 105 (1) 2004 pp 19–21.
86. Nanotoday, Nanoparticles as building blocks?, Pitkethly, M. J. *Materials Today* Volume 6, Issue 12, Supplement. December 2003 pp 36–42.
87. Nanostructures of zinc oxide. Zhong, L. W. *Materials Today* (2004) 7 (6), 26.

2

Nanometric architectures: emergence of efficient non-crystalline atomic organization in nanostructures

T ASTE and T DI MATTEO,
The Australian National University, Australia

2.1 Introduction

A common paradigm in science tends to assume that things are naturally inclined to pursue order and perfection. When the atomic structure of materials is concerned, such a paradigmatic order is achieved by the endless periodic repetition of local motifs in perfect crystals. Such an assumption has a quite solid background. Indeed, there are several good reasons to choose order as a constitutive paradigm. Order is *convenient*, the complete information about a crystalline structure can be encoded in a very compact way by retaining only the information about a single fundamental cell plus some translational symmetry operations. Order is *efficient*, often the most efficient structures have high degrees of symmetry and result in crystalline arrangements. But, on the contrary, disorder is *everywhere*; in nature order is more the exception than the rule. This is in particular the case when complex structures are concerned as, for instance, in several nanomaterials or in other systems such as polymers, biological systems and glasses. The actual driving mechanisms that lead to such disordered structures are very diverse and range from fast out-of-equilibrium cooling (metal glasses) to evolution (biological systems). In the context of the present book, one should remark that in small systems such as nano-assemblies, the competition between the contribution from the surface with respect to the contribution from the bulk becomes extremely important and there exists a critical size below which an *organized disorder* is more effective than order, making in this way non-crystalline structures favorite.

It must be first clarified that disorder is not necessarily randomness; we often encounter systems that are highly organized both at the local and global scale but nevertheless do not reveal any periodic (or quasi-periodic) pattern. The study of such organized but disordered structures is very challenging and the available investigation tools appear to be inadequate to capture the essential features of these structures. Indeed, a complete description of the structure of a disordered system requires information about the exact

position, orientation, shape and connectivity of each element. However, part of such information is redundant or irrelevant and several degenerate states with different microscopic realizations can share the same macroscopic properties. To determine which are the accessible configurations at the local level, and to understand which are the possible combinations which generate the global packing is of fundamental importance. In particular, finding measures for local, and global, organization is the essential starting point for understanding the basic mechanisms which form these structures. It is also an essential step in the development of technologies which enable us to control and tune the structure of amorphous materials.

In this chapter we focus on disordered sphere packings as an archetypal example. We review and discuss the most suitable paradigms and tools to investigate and classify the structure of such systems. We examine the 'typical' local configurations, their relative occurrences, their correlations, their organization and the resulting overall hierarchical structure. We show that these systems are disordered but nonetheless they are highly organized. Through several examples we show that such an organization can yield to amorphous packings which are *more* efficient than crystalline aggregates. We show that in small-scale systems, where small numbers of atoms in nano-sized clusters are involved, such disordered configurations can have relatively large packing fractions, large particle interconnectivity and small surface-to-volume ratio.

Our study is based on the analysis of the largest empirical dataset on disordered structures presently available in the literature. Such a dataset records the positions of more than 385,000 sphere centers in six different packings of equally sized spheres. It should be noted that, although these studies are specific to the case of equal spheres, the methodology that we use is based on the analysis of the properties and constraints resulting from the inescapable rules dictated by geometry and packings. Therefore the results are rather universal.

Each section in this chapter is organized to be as self-contained as possible with cross-references to other sections when necessary. Section 2.2 describes the empirical dataset and the sample preparation. In section 2.3 we discuss the structural organization of equal spheres packings by analyzing the radial distribution function. In section 2.4 we measure the local orientational symmetry and the occurrence of known crystalline configurations showing that no crystalline or poly-crystalline order is present – even at atomic level. The study of non-crystalline structures is pursued in further details in section 2.5 where several techniques to analyze local geometrical correlations between more than two particles are discussed and applied. We find that these systems can be conveniently described in terms of packings of (irregular) tetrahedra. The volume distribution of these tetrahedra, studied in section 2.6, indicates that these packings are formed by maximizing the configurational entropy under the geometrical constraints dictated by the non-overlapping condition.

In section 2.7 we discuss the local topological properties by measuring the number of bounds between neighboring particles. We point out that there is a bijective relation between the packing density and the system's topological organization. The hierarchical organization beyond first neighbors is analyzed in section 2.8 where the interconnectivity of the contact network is investigated. In section 2.9 we show that disorder can be efficient and these systems can pack more tightly and effectively than several ordered crystalline systems. Conclusions, in section 2.10, summarize the main results and perspectives.

2.2 Equal spheres packings

A static stack of particles in a container can minimize its potential energy (under gravity) by maximizing the packing fraction. Such a pursuit of maximum compaction is common to several other physical systems. Notably, at atomic level this is a feature associated with metallic bounding. From a purely geometrical perspective, it is known that the largest attainable packing fraction in a system of equal spheres is $\rho = \pi/\sqrt{18} \sim 0.74$ (Conway and Sloane 1997, Aste and Weaire 2000) which corresponds to a disposition of the spheres in an ordered stack of parallel hexagonal layers (forming the so-called Barlow packings). On the other hand, when equal spheres are poured in a container they spontaneously arrange in a disorderly fashion occupying a fraction of the total volume (called packing fraction or packing density ρ) between 0.55 and 0.64. As mentioned above, here we are discussing a paradigmatic example of a system where the most efficient global structure is ordered but, on the other hand, it spontaneously forms disordered arrangements. It is very interesting to note that such a disordered state is rather robust and resilient; operations such as gentle shaking or compression might change the packing fraction (within a range 0.61–0.64) but no spontaneous generation of configurations with packing fraction above ~0.64 has ever been reported. Such a packing fraction appears therefore to be an (empirical) upper limit and it is often referred to in the literature as the 'random close packing' limit. Larger densities can be obtained only by externally inducing (partial) crystallization.

Until now the empirical investigation of the geometrical structure of these systems has been limited by the very sparse availability of accurate experimental or numerical data. Indeed, after the seminal works of Bernal, Mason and Scott (Bernal and Mason 1960, Scott 1962, Mason 1968), only very recently the use of tomography has allowed us to 'see' three-dimensional structures and explore their geometry from the sphere level up to the whole packing.

2.2.1 Six samples

All the structures discussed in this chapter refer to a database constructed

from the empirical analysis, by means of X-ray computed tomography, of several large samples of disorderly packed monosized spheres. The experimental technique and some of the empirical results were presented in a preliminary form in Aste *et al*. (2004) and in a more complete form in Aste *et al*. (2005b). These studies are the largest and the most accurate empirical analysis of disordered packings ever performed. The entire database collects the coordinates (with precision better than 0.1% of their diameters) of more than 385,000 spheres from six samples of monosized acrylic spheres prepared in a cylindrical container.[1] In this chapter we refer to these samples as samples A, B, C, D, E and F. The geometrical investigation of the packing structure was performed over a central region ($\boldsymbol{G}$) at 4 sphere-diameters away from the sample boundaries. (Note that spheres outside $\boldsymbol{G}$ are considered when computing the neighboring environment of spheres in $\boldsymbol{G}$.) The total number of spheres in each sample, the respective numbers in $\boldsymbol{G}$ and the sample densities are reported in Table 2.1.

Table 2.1 Sample density and the interval of variation (±) within each sample; number of spheres in the sample (N); number of spheres in the central region (N_G); estimated average number of neighbors in contact (n_c), average number of neighbors at given radial distance ($n_t(r)$ with r = 1, 1.02, 1.05, 1.1 diameters)

	Density (ρ)	N	N_G	n_c	$n_t(1)$	$n_t(1.02)$	$n_t(1.05)$	$n_t(1.1)$
A	0.586 ± 0.005	102 897	54 719	**5.81** ± 0.3	3.0	5.5	6.7	7.5
B	0.596 ± 0.006	34 016	15 013	**5.91** ± 0.2	2.9	5.9	6.8	7.7
C	0.619 ± 0.005	142 919	91 984	**6.77** ± 0.2	3.5	6.4	7.5	8.4
D	0.626 ± 0.008	35 511	15 725	**6.78** ± 0.2	3.3	6.0	7.5	8.4
E	0.630 ± 0.01	35 881	15 852	**6.95** ± 0.2	3.4	6.3	7.6	8.6
F	0.640 ± 0.005	36 461	16 247	**6.97** ± 0.2	3.3	6.9	7.9	8.9

2.3 Searching for structure: the radial distribution function

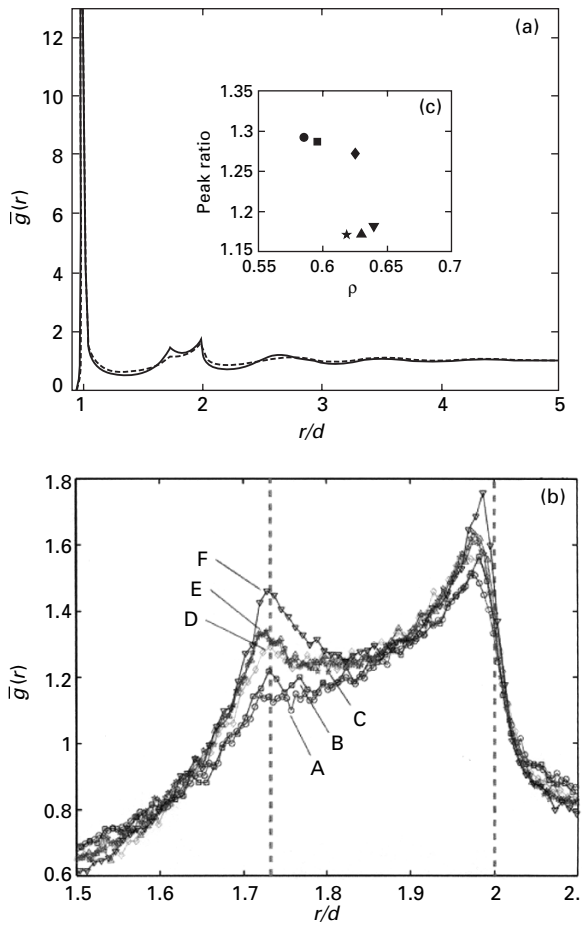
The *radial distribution function* is the probability distribution to find the *center* of a particle in a given position at a radial distance r from the center of a reference sphere. This measurement is widely used to characterize packing structures and contains information about long-range interparticle correlations and their organization (Scott 1962, Mason 1968, Bideau and Hansen 1993).

Figure 2.1 shows the normalized radial distribution functions $\tilde{g}(r)$ for the six samples A–F (Aste *et al*. 2004, Aste *et al.* 2005b). In this figure a large peak is clearly visible at r/d = 1 which is due to the contribution from the spheres in contact. The presence of other peaks after this first peak at r/d =

1. Additional material is available at: http: //wwwrsphysse.anu.edu.au/granularmatter/

1 is a clear indication that the system is organized. A structure with distinct local patterns and characteristic lengths is present. For instance, it is easy to verify that a radial distance $r/d \sim \sqrt{3}$ is consistent with configurations made by placing the centers of four touching spheres on the vertices of two in-plane equilateral triangles which share an edge. Whereas, the peak at r/d~2 is due to three or more spheres which are lying along a (rather) straight line.

The detail of the secondary peaks (Fig. 2.1b) shows that both the peaks at $r/d = \sqrt{3}$ and $r/d \approx 2$ increase in height with the packing density (although with different rates (see Fig. 2.1(c)). The presence of secondary peaks is a clear indication of structural organization and the fact that these peaks sharpen



2.1 (a) Normalized radial distribution function. (b) The detail of the two peaks respectively at $r/d = \sqrt{2}$ and $r/d = 2$ (vertical lines). (c) Ratio between the value of the peak at $r/d = 2$ and the peak at $r/d = \sqrt{3}$ vs. packing density.

with the packing fraction certainly indicates an increasing organization. On the other hand, no signs of crystallization are detected. This is discussed in the next section.

2.4 Local motifs and orientation symmetry

We have seen in the previous section that the radial distribution function clearly reveals the existence of a structural organization; some typical lengths (associated with special local configurations) appear repeatedly in the packing. However, the fact that there are typical local configurations which often repeat in space does not necessarily imply that there is any long-range order or (poly)crystalline regions. In this section we discuss a method to classify and quantify the type of local orientational order. Revealing and quantifying orientational order is a key issue in establishing the type of internal organization and in particular in determining whether there exists a 'typical' disordered state and identifying possible tendencies towards hidden symmetries.

2.4.1 Two-dimensional packings and geometrical frustration

In two dimensions, when equal disks are poured in a container, they tend to organize themselves locally into close packing configurations but, at first, such a local organization is limited to very short distances yielding to an overall disordered assembly. Some shaking can increase the density leading gradually to an overall ordered configuration. It has been argued that a similar mechanism could take place in three dimensions as well. It is, however, clear from several empirical observations that the tendency towards a crystalline order (even if short-ranged) is much less spontaneous in three dimensions than in the two-dimensional case. Such a resistance to become ordered has often been associated with 'geometrical frustration', which is present in three dimensions but not in two dimensions. Indeed, in two dimensions six equal disks can be placed in touch with a given central disk forming the densest possible local configuration. Such a configuration is compatible with translational symmetry and it can be repeated periodically in space leading to a crystalline packing (the triangular lattice). However, in three dimensions the closest possible local configuration is an arrangement of twelve equal spheres in touch with a central sphere and with centers on the vertices of a regular icosahedron. Such a packing has local density $\rho = 0.754...$ (which is more than 1% denser than the configurations in Barlow packings at $\rho = 0.740...$) (Aste and Weaire 2000). But, differently from the two-dimensional case, such a compact local icosahedral configuration is not compatible with translational symmetry and therefore it cannot be repeated in space without leaving gaps; it is geometrically frustrated.

It has often been argued that competition between the tendency to form locally compact configurations and geometrical frustration, could be the key to understanding the mechanism of formation of disordered packings and glassy structures (Aste and Coniglio 2004). If this is the case, one would expect to see, at local level, configurations with rotational symmetries characteristic of icosahedral and other closely packed structures. The study of the local packing arrangements and in particular the rotational symmetry can therefore give important insights into the mechanism of formation of these structures.

2.4.2 Local rotational symmetry

A useful method to quantify and classify rotational symmetry was introduced by Steinhardt *et al*. (1983). This method consists of the projection, on a spherical harmonic base, of the sets of versors pointing between a sphere and its neighbors. Such an overlap between the rotational symmetry of a spherical harmonic and the local symmetry is measured by the quantity:

$$Q_l^{(i)} = \sqrt{\frac{4\pi}{2l+1} \sum_{m=-l}^{l} |\langle Y_{l,m}(\vartheta(\vec{r}_{i,j}), \varphi(\vec{r}_{i,j}))\rangle_j|^2} \qquad 2.1$$

with $Y_{l,m}(\vartheta(\vec{r}_{i,j}), \varphi(\vec{r}_{i,j}))$ the spherical harmonics. Here $\vec{r}_{i,j}$ is the vector between the center of sphere 'i' and the center of a neighboring sphere 'j' and the two quantities $\vartheta(\vec{r}_{i,j})$ and $\varphi(\vec{r}_{i,j})$ are respectively the polar and azimuthal angles. In eqn 2.1 the average $\langle(\ldots)\rangle_j$ is restricted only over the local bonds between a sphere and its neighbors 'j'. In this way we can associate to each sphere in the system a value of $Q_l^{(i)}$ and local order can be singled out by counting the number of configurations with Q_l corresponding to special symmetries. In particular the two values of Q_l for $l = 4$ and six have special significance. For instance, the simple cubic lattice has $(Q_4, Q_6)^{sc} = (0.764, 0.354)$, the body centered cubic lattice has $(Q_4, Q_6)^{bcc} = (0.036, 0.511)$, the *fcc* has $(Q_4, Q_6)^{fcc} = (0.191, 0.574)$, the *hcp* has $(Q_4, Q_6)^{hcp} = (0.097, 0.485)$ and the icosahedral rotational symmetry gives $(Q_4, Q_6)^{ico} = (0, 0.663)$. Since the lowest non-zero Q_l common to the icosahedral, hexagonal and the cubic symmetries is for $l = 6$, it has been argued by several authors that the value of Q_6 is a good indicator of the degree of order in the system and it might be used as an 'order parameter' (Richard *et al*. 1999a, b, Truskett *et al*. 2000, Torquato *et al*. 2000, Kansal *et al*. 2002). Indeed, Q_6 is very sensitive to any kind of crystallization and it increases significantly when order appears (Richard *et al*. 1999a).

On the other hand, in samples A–F, we observe (Aste *et al*. 2004, 2005b) that a very large fraction of local configurations (between 23 and 47%) have local symmetry characterized by (Q_4, Q_6) with values which are rather far

from any special symmetry. This analysis reveals that there are no significant fractions of local configurations with symmetry compatible with icosahedral, simple cubic, *bcc, fcc or hcp.* These findings exclude the existence of any crystalline order (even at local level) and also discard the ideas that a tendency towards local *frustrated* icosahedral order can be responsible for the resistance to crystallize for such packings. However, these findings, although intriguing, are not conclusive because the origin and nature of the most abundant configurations with (Q_4, Q_6) ~ (0.25, 0.45) are still elusive. In the next section we clarify this point by using different techniques to allow viewing local packing organization from a different perspective.

2.5 Understanding the structure: local geometrical organization

The radial distribution function (section 2.3) clearly shows that despite the fact that these systems are amorphous and disordered, they have a characteristic structural organization which reveals itself regularly at local level. However, the study of the local rotational symmetry, discussed in the previous section, does not disclose any known special order. In this section we focus on techniques to identify local correlations which involve more than two neighboring spheres.

2.5.1 Voronoï shapes

Let us first consider the analysis of the shapes of Voronoï cells constructed around the center of each sphere. The Voronoï cell is a polyhedron the interior of which consists of all points closer to the center of the given sphere than to any other sphere center in the sample. It coincides and generalizes the Wigner–Seitz cell construction for crystalline systems. This is a powerful tool to investigate the local arrangement of packed objects in space and it can be used to classify their local organization. The advantage of such a method is that there is no arbitrary choice for the neighboring criteria. On the other hand, this is also its weakness because neighbors which are rather far from the central sphere (typically, up to distances r/d~1.4) contribute to the shape of the cell.

In (Aste *et al*. 2004, 2005b) we observe an average number of Voronoï faces which decreases with the density (see Table 2.2) from $\langle f \rangle = 14.6$ at $\rho = 0.586$ (sample A) to $\langle f \rangle$ 14.3 at $\rho = 0.64$ (sample F). Whereas the fractions of faces with 4, 5 and 6 edges (respectively f_4, f_5 and f_6) follow mixed trends (see Table 2.2). These data give a clear indication that the internal structural organization of these systems is changing with the overall packing density. However, these measures are still not sensitive enough to define clearly a precise link between the local properties of the packing structure and the overall packing density.

Table 2.2 Average number of Voronoï faces $\langle f \rangle$. Fractions of Voronoï faces with 4, 5 and 6 edges (f_4, f_5 and f_6). Fraction of configurations with four and five common neighbors (p_4 and p_5). Fraction of configurations with five bounded common neighbors disposed in a pentagonal ring (555). This configuration can be viewed as a 'piece' of icosahedron. (The threshold distance is 1.1d)

	$\langle f \rangle$	f_4 (%)	f_5 (%)	f_6 (%)	p_4 (%)	p_5 (%)	555 (%)
A	14.6	20.3	33.8	26.8	17.5	2.59	0.01
B	14.6	20.0	34.0	26.8	18.6	2.66	0.01
C	14.4	19.4	37.3	27.8	26.4	6.00	0.06
D	14.4	19.5	36.4	28.1	24.4	5.48	0.06
E	14.4	19.4	37.1	28.2	26.7	6.11	0.06
F	14.3	18.8	39.6	28.6	30.6	8.29	0.09

2.5.2 Common neighbors

An alternative interesting approach called 'common-neighbor analysis of structure' was introduced by Clarke and Jónsson (1993). This analysis is based on the study of local 'bounds' between particles. In particular, two spheres are 'bounded' if they stay within a given threshold radial distance. The study consists in analyzing the structure around couples of bounded particles, identifying common neighbors which are bounded to both the particles. In general, two touching equal spheres can have a number between 0 and 5 of common neighbors which are also in touch with both spheres. This number is rather sensitive to the local correlations in the packing. Indeed, in (Aste 2005) it is reported that the numbers of configurations with four or five common neighbors increase sensibly during compaction (see Table 2.2). In particular, when the threshold distance is fixed at 1.1d, the fraction of edges with four common neighbors (p_4) varies between 17 and 31%, whereas the configurations with five common neighbors (p_5) grow from less than 3% to above 8% (A to F). This is in agreement with observations in computer simulations and it suggests that these quantities could be profitably used as 'order' parameters to characterize the packing structure. Indeed, they are much more sensitive than the equivalent quantities obtained from the study of the shape of the Voronoï cell or the Q_4 and Q_6 measurements.

2.5.3 Searching for close local packings

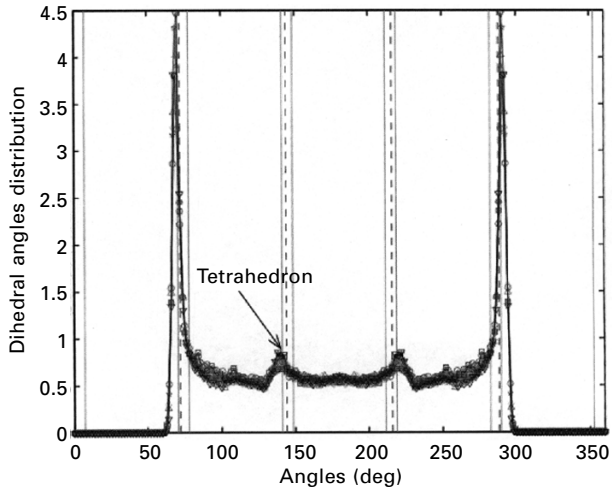
We have previously discussed the facts that the closest local arrangement of equal spheres in three dimensions is achieved by 12 spheres in touch with a central sphere and placed on the vertices of a regular icosahedron. We have also shown in section 2.4 that in the samples A–F there are no traces of icosahedral rotational symmetry. However, this does not exclude the possibility that close packed configurations made by 'pieces' of icosahedron might be present in these structures. Signatures of icosahedral order can be singled out

by analyzing in which way common neighbors arrange themselves around the two touching spheres. In particular a configuration of five bounded spheres arranged in a pentagonal ring around a couple of bounded spheres is a local signature of a (partial) icosahedral order. Such a configuration is indeed a 'piece' of icosahedron and therefore its occurrence within the samples is a good measure to understand whether icosahedral configurations might have played any significant role in the dynamic formation and in the static organization of such structures.

We observe (Aste, *et al*. 2005b) that such configurations are present only in very small fractions (from 0.01 to 0.1%, see Table 2.2), indicating therefore that this kind of close-packed local arrangement has very little statistical significance and no relevant physical importance. This observation is in agreement with the results from the (Q_4, Q_6) analysis in section 2.4. It appears therefore that, differently from the two-dimensional case (section 2.4), in three dimensions closed packing configurations are not present – even at the local level. A further demonstration of this comes from the study of the local volume fractions. Over a total number of more than 209,000 local Voronoï configurations, only 14 have volume fractions above 0.74 (the *fcc* has 0.7405... and the icosahedral arrangement has 0.75...). Similar evidence can be obtained by looking at the local density associated with the Delaunay tetrahedral decomposition. Such decomposition is constructed by choosing the vertices of each tetrahedron on the centers of four neighboring spheres under the condition that no other spheres in the pack have centers within the circumsphere of the Delaunay simplex. We find that across the samples A–F only 2% of Delaunay configurations are closed tetrahedra. This is further evidence that close-packed configurations play no relevant role in the structure of these systems. This therefore also implies that geometrical frustration, resulting from local dense packing incompatible with global translational symmetry, cannot have any relevant role in the mechanism of formation of these disordered systems.

2.5.4 Dihedral angles distribution

A further step toward the understanding of which kind of local configurations are present in these systems can be undertaken by analyzing the distribution of dihedral angles between common neighbors around a given couple of bounded spheres. Previously we have calculated the number of common neighbors and shown that the most frequent configurations are the arrangements with four and five common neighbors. Let us here study how these spheres are distributed inside this common ring. To this end, we first construct the triangle between two bounded spheres and one common neighbor and then we measure the dihedral angles between such a triangle and all the other triangles formed with the other common neighbors. The resulting distribution of angles is shown in Fig. 2.2.



2.2 Dihedral angle distribution (x-axis: angular degrees; y-axis: renormalized frequencies). The vertical lines indicate the angles $\theta = n$ arccos(1/3) (and 360-θ) with n = 1, 2, 3, 4, 5 (tetrahedral packings). The dashed lines are at the angles $\theta = n\frac{360}{5}$ (n = 1, 2, 3, 4), which will correspond to icosahedral configurations.

This figure shows a distribution which is symmetric in θ and 360-θ. In such a distribution, the largest peak is at θ = arccos(1/3) = 70.5... (and 360-arccos(1/3) = 289.4...) which is the dihedral angle of a regular tetrahedraon. Another (smaller) peak is visible at θ = 2arccos(1/3) = 141.0... (and 218.9…) which corresponds to a configuration with two touching tetrahedra. This clearly indicates that the common neighbors tend to gather together in the circular ring forming tetrahedral packings with angles θ = n arccos(1/3). Indeed, the case with a uniform angular repartition of common neighbors will lead to a rather different structure of the angular distribution with peaks at $n\frac{\pi}{3}$, $n\frac{\pi}{4}$ and $n\frac{\pi}{5}$, with n = 1, 2, 3, 4… These peaks are less visible than the 'tetrahedral' ones in the empirical distributions in Fig. 2.2. It is interesting to note that, similarly to the case for the radial distribution function, the relative heights of the peaks are sensitive to the sample density. But in this case the variations are less evident.

2.6 A packing of tetrahedra

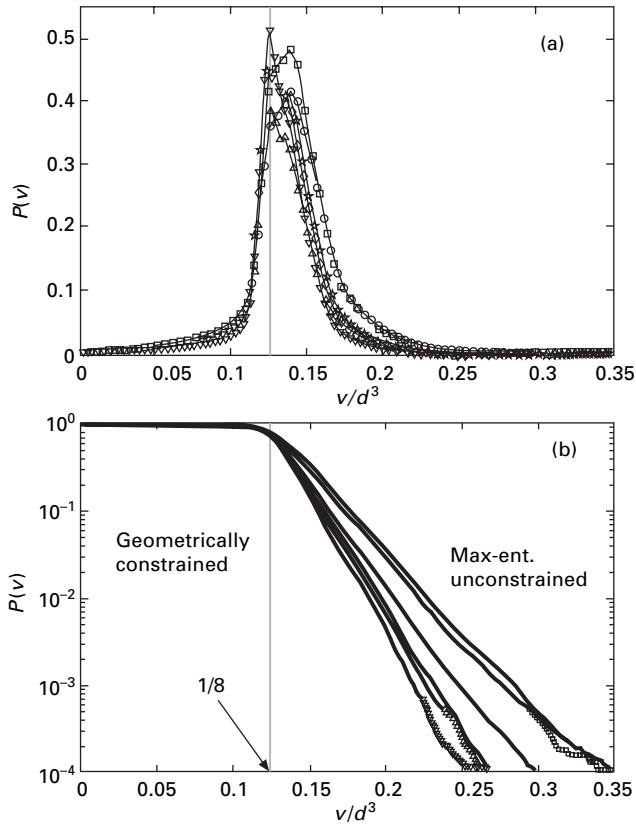
In the previous section, we discussed that these amorphous structures can be conveniently viewed as the result of a packing of tetrahedra. A natural way to study such a system of tetrahedra is the Delaunay decomposition where space is divided into minimal tetrahedra (with vertices on the centers of

neighboring spheres and chosen in such a way that no other spheres in the pack have centers within the circumsphere of each Delaunay tetrahedron). In this section we discuss how these Delaunay tetrahedra partition space and in particular we focus on the distribution of the Delaunay volumes. Indeed, in a given packing some local configurations are closer and others are looser and the whole packing is made by gluing together these local configurations in a disordered way which is compatible with mechanical stability, geometrical constraints and global density. The study of how space is shared among the packed spheres is essential for understanding how efficiently the spheres are arranged locally.

It has been observed (Aste 2005) that the volumes of the Delaunay tetrahedra follow distributions which, at large volumes, are well described by the exponential behavior $P(v) \propto \exp(-\beta v/d^3)$, where the coefficients β grow with the density ranging between $\beta \sim 44$ at $\rho = 0.586$ (sample A) to $\beta \sim 73$ at $\rho = 0.64$ (sample F) (see Fig. 2.3).

In order to understand the origin of such exponential behavior, let us consider the whole system as an ensemble of local independent cells that freely exchange volume among each other under the constraint that the sum over all the volumes must be a constant which characterizes the state of the system. If we assume that: (i) cells are indistinguishable (the shapes are ignored); (ii) there are no correlations among cells; (iii) the geometrical constraints on the packing (stability, kissing number, etc.) have no relevance, then the probability of finding a cell with a volume v must be given by: $P(v) \propto \exp(-\beta v/d^3)$, with $\beta = d^3/\langle v \rangle$ (Aste 2005). However, this simple argument predicts coefficients which are not satisfactory; we obtain $\beta \approx 36\rho/\pi$ which gives values for the coefficients which are one order of magnitude lower than the empirical coefficients. On the other hand, it is also evident from Fig. 2.3 that these distributions are not pure exponentials. Indeed, they have an exponential *tail* at large volumes but the exponential behavior ceases at smaller volumes.

Looking in detail at Fig. 2.3 we observe that the exponential trend starts at values of v/d^3 around $0.125 = 1/8$. Such a critical volume is very meaningful; it is the maximum volume for a tetrahedron made from the centers of four spheres with two spheres in touch and the other two common touching neighbors. A configuration which corresponds to a tetrahedron with five edges with length equal to d and one edge with length $\sqrt{3/2}d$. The fact that the exponential ceases around this volume indicates that the Delaunay tetrahedra can arrange freely in space only up to the limit when spheres get in contact and geometrical constraints begin to be relevant. We can therefore view these systems as comprising two phases: (i) a 'solid' phase made by compact tetrahedra ($v/d^3 \leq 1/8$) which are geometrically constrained and are responsible for the mechanical stability; (ii) a 'fluid' phase made by loose tetrahedra ($v/d > 1/8$) which are geometrically unconstrained and take volumes accordingly



2.3 (a) Distribution of Delaunay volumes in ***G***. (b) The inverse normalized cumulants show tails that decrease linearly in semi-logarithmic scale: $p(v\sim)\exp(-\beta v/d^3)$ (best-fits: β = 43.9; 45.4; 55.2; 64.6; 66.8; 72.9).

with the distribution which maximizes entropy. We verified that, by assuming that the exponential behavior is followed only above $v/d^3 = 1/8$, one can calculate analytically the coefficients at the exponent $\beta^{-1} = \frac{\langle v \rangle}{d^3} - \frac{1}{8} \approx \frac{\pi}{36\rho} - \frac{1}{8}$ which returns values rather close to the ones measured empirically.

2.7 Structural organization and packing fraction

In the previous sections we have observed that the sample density affects several local structural properties. Understanding if there is a uniform, bijective and reproducible relation between the packing properties and the sample density is of great importance because it will open up the possibility of

classifying complex structures by using only the information concerning their density, which is a simple and scalar parameter.

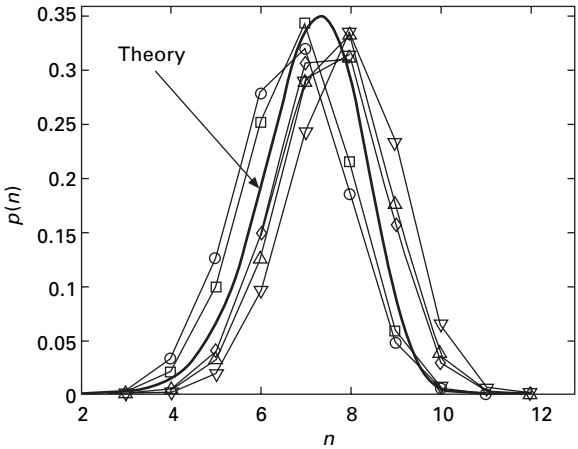
2.7.1 Neighbors in contact

In this section we address this question by looking at the average number of spheres in contact with a given sphere. This is probably the most investigated structural parameter in the literature of packings (Smith *et al*. 1929, Bernal and Mason 1960, Scott 1962, Mason and Clark 1966, Mason 1968, Steinhardt, *et al*. 1983, Seidler, *et al*. 2000, Aste and Weaire 2000, Sederman *et al*. 2001, Torquato and Stillinger 2001, Silbert *et al*. 2002, Kohonen *et al*. 2004). Indeed, this is a very simple topological quantity which gives important information about the local configurations and the packing stability. Moreover, historically this is the first topological quantity investigated in these systems (Smith *et al*. 1929, Bernal and Mason 1960, Scott 1962). Unfortunately, although simple in its definition, such a number is unavoidably an ill-defined quantity. Indeed, from a geometrical perspective, the information about the positions, and eventually the sizes of all spheres, are not sufficient to determine such a number; two spheres can be arbitrarily close but not in touch. In the literature several physical methods have been used (Smith *et al*. 1929, Bernal and Mason 1960, Scott 1962, Kohonen *et al*. 2004), but they encounter problems all essentially associated with the uncertainty on the correct threshold distance which must be used to distinguish between 'touching' and 'non-touching' spheres.

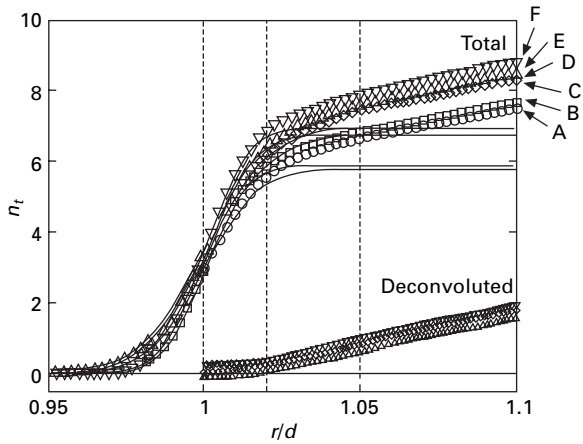
Here we discuss the results reported in Aste *et al*. (2004, 2005b) where the number of spheres in contact (n_c) and their distribution were calculated with great accuracy on the six samples A–F. Table 2.1 reports the values of the average number of neighbors (n_t) computed in $\boldsymbol{G}$ at the four different radial distances: $r = d$, $1.02d$, $1.05d$ and $1.1d$. As one can see, this number ranges between 3 and 8.9 increasing with the sample density and the threshold distance. The probability distribution for the number of spheres within a radial distance $1.05d$ is plotted in Fig. 2.4. An estimate for the actual number of spheres in contact can be calculated by deconvoluting the behavior of $n_t(r)$ (shown in Fig. 2.5) with a complementary error function normalized to n_c. Such a deconvolution extracts the information on the number of neighbors in contact from the spread in $n_t(r)$ due to the polydispersity in the sphere diameters (Aste *et al*. 2005b). The resulting values of n_c are also reported in Table 2.1.

2.7.2 Connectivity and density: a rigidity percolation?

The relation between the average number of neighbors in contact and the packing density is shown in Fig. 2.6. As one can see there is a clear and

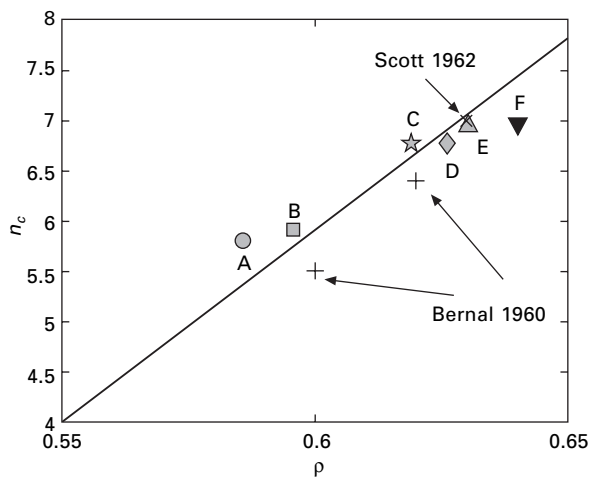


2.4 Distribution of the number of near neighbors at radial distances within 1.05 diameters. The tick line is the theoretical behavior predicted by a free-volume-like local theory: $p(n) \propto (4\pi - 4\pi\ n/n^*)^n$ with $n^* = 12.99$ (see eqn 4 in Aste 2005).



2.5 Average number of sphere centers within a radial distance r (symbols). Complementary error function, normalized to n_c by best-fitting the agreement with the data in the region $r/d < 1$ (lines). The averages ($r = d$) and the standard deviations are calculated from the probability distribution for radial distances smaller than d between pairs of centers. The re-normalized complementary error function fits well the data for $r/d < 1$. After this value near-neighbors not in contact start to contribute significantly to $n_t(r)$ and the two behaviors split. The 'deconvoluted' plots show the differences between $n_t(r)$ and the fits with the complementary error functions normalized to n_c. Data from Aste *et al*. 2005b).

consistent increasing behavior of n_c with the density. A similar increasing trend was also found in simulated packings (Clarke and Jónsson 1993, Yang *et al*. 2000). Interestingly, from Fig. 2.6 one can see that a linear extrapolation is consistent with a value $n_c = 4$ at the packing density $\rho = 0.555$. These two values have a special significance: the packing density $\rho = 0.555$ is the lowest empirically achieved and it coincides with the dilatancy onset (Onoda and Liniger 1990). On the other hand, $n_c = 4$ is the average number of neighbors required to exactly match the number of degrees of freedom with the number of constraints in these systems (Maxwell counting) (Lagrange 1788, Maxwell 1864, Aste 2005). When freedoms and constraints are exactly matched the system is isostatic (Moukarzel 1998, Edwards and Grinev 1999); the intergranular forces are uniquely determined by the balance of force and torque alone. On the contrary, an over-constrained structure can generate self-stress and the deformation of individual grains becomes relevant. In such an isostatic unstressed state, the system has zero elastic modulus (in the thermodynamic limit) (Thorpe and Chubynsky 2001, Chubynsky and Thorpe 2001), it is therefore marginally rigid and it can be seen as an intermediate state between fluid and solid (Ball and Blumenfeld 2002). The observation that such an isostatic state can be realized at densities $\rho \approx 0.55$ suggests that at such a density a *rigidity percolation* (Thorpe and Chubynsky 2001, Chubynsky and Thorpe 2001) takes place. Below $\rho \approx 0.55$ the system would be still a fluid, unable to sustain its own weight and it cannot be realized under gravity, above this density the system becomes (marginally) rigid and realizations start to be empirically observable.



2.6 Number of neighbors in contact vs sample density. The filled symbols correspond to the samples investigated in the present work. The two symbols '+' are the values from Bernal and Mason (1960) whereas the 'X' is from Scott (1962) and Mason (1968).

2.8 Packing hierarchy: the topological structure beyond first neighbors

Any mechanical stress, any perturbation or any infinitesimal local displacement must propagate from particle to particle through the network of bounded neighbors (Aste *et al*. 2002). The understanding of the hierarchical organization of such a contact network beyond first neighbors is therefore crucial. However, such analysis is not trivial. Indeed, one must consider that in a disordered packing the local environment surrounding each particle is unique. The packing itself can be considered as a gigantic three-dimensional jigsaw puzzle where each local configuration interlocks precisely with the surrounding ones in a hierarchical organization which goes beyond first neighbors.

2.8.1 Shell analysis

In this section we apply to disordered matter an approach which was originally developed for the study of crystalline systems (Brunner and Laves 1971, O'Keeffe 1991b, Conway and Sloane 1997) and disordered foams (Aste *et al*. 1996a, Aste 1997). The topological structure of crystalline frameworks has been intensely studied in terms of the number of atoms that are j bonds away from a given atom. If we start from a given 'central' atom, the first 'shell' (distance $j = 1$) is made by all the atoms in contact with the central one. The second shell (distance $j = 2$) consists of all atoms which are neighbors to the atoms in the first shell, excluding the central one. Moving outward, the atoms at shell $j + 1$ are all the ones which are bonded to atoms in shell j and which have not been counted previously.

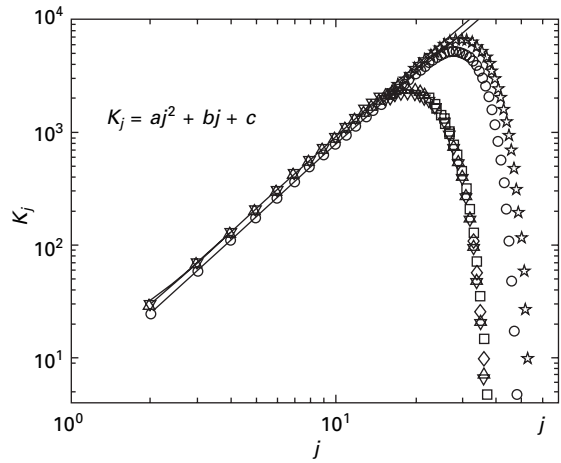
In the ideal case of infinite, periodic, crystalline structures with no boundaries, the number of atoms per shell should increase with the topological distance and it has been shown that, in several three-dimensional crystalline structures, the law of growth for the number of atoms (K_j) at shell j can be described with: $K_j = a_j j^2 + b_j j + c_j$ (where a_j, b_j and c_j coefficients might vary with j but only within a bounded finite interval) (Brunner and Laves 1971, O'Keeffe 1991b, Conway and Sloane 1997, Grosse-Kunstleve *et al*. 1996). Following the definition of O'Keeffe (1991a), the asymptotic behavior of K_j can be characterized in terms of an 'exact topological density'; $TD = \langle a_j \rangle/3$ (Grosse-Kunstleve *et al*. 1996). It has been noted that such a topological density is interestingly related with the geometrical density of the corresponding crystalline structure and it is a powerful instrument to characterize such systems. For instance, it is easy to compute K_j for the cubic lattice: $K_j = 4j^2 + 2$. Whereas spheres packed in a *bcc* (body centered cubic) crystalline arrangement have $K_j = 6j^2 + 2$ ($j > 0$). On the other hand, it has been shown (Conway and Sloane 1997) that for Barlow packings of spheres, K_j are always in a narrow range with a_j within 10 and 11.5. More generally, it has

been observed by O'Keeffe and Hyde (1996) that for *lattice sphere packings* with coordination number n, the general rule holds, $K_j = (n - 2)j^2 + 2$, implying therefore $a = n - 2$.

Beyond perfect crystalline order very few results are known either from theoretical, empirical or numerical points of view. One can argue that K_j must grow with a law comparable with the law for a spherical shell, $K_j \sim aj^2 \sim 4\pi j^2$. However, it is also clear that the shape of the growing shell and its roughness can drastically influence the value of the coefficient a (as observed in two-dimensional cases) (Aste *et al*. 1996b). Moreover, it can be shown (Aste *et al*. 2005a) that in some topological networks the law of growth can follow an intrinsic dimension which is different from the dimension of the embedding space (three in our case). This mechanism can produce a power law growth with exponents different from 2, or different behaviors such as exponential – or even faster – laws of growth. By applying such analysis to the six empirical samples (A–F) we observe that the number of spheres at a given topological distance j from a central one follows the quadratic law

$$K_j = a\,j^2 + b\,j + c \qquad 2.2$$

up to a critical distance $\hat{j}$, above which the shells hit the sample boundaries and K_j begin to decrease (see Fig. 2.7).



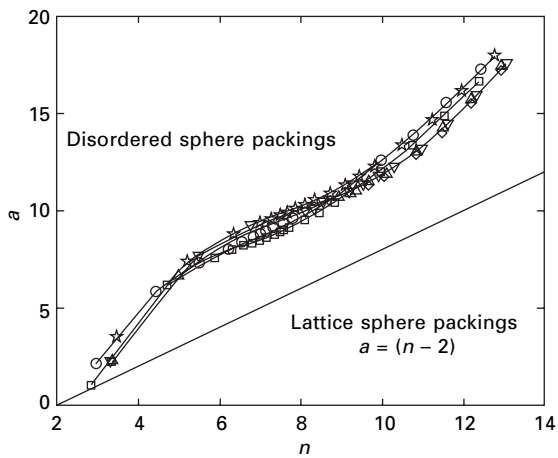
2.7 Shell occupation numbers vs topological distance. The symbols indicate the different samples (as in Fig. 2.5) and the lines are the best-fits using the polynomial form: $K_j = aj^2 + c_1 j + c_0$. The fits are between $j = 2$ and $j = \hat{j} = 10$ (for samples B, D, E, F) and $\hat{j} = 15$ (for samples A, C). The data refer to threshold distance $1.05d$.

2.8.2 Network interconnection

We have observed that the coefficient '*a*' in the polynomial growth described by eqn 2.2 depends on the criterion of the threshold distance which defines bounded neighbors. Indeed, changes in the threshold distances are unavoidably associated with changes in the contact network. Such a threshold affects also the average coordination number *n* in the contact network (Fig. 2.5). In Fig. 2.8 we show that these two quantities are positively correlated; the coefficient *a* increases monotonically with *n*. Interestingly, these disordered systems have values of the coefficient *a* which are consistently larger than $a = n - 2$ observed in lattice sphere packings. Such larger values indicate that the contact network in disordered packing is more densely inter-connected than in the analogous lattices. A topologically more compact network must correspond to structures which are more stable and robust. It appears therefore – for the first time – that disorder can, in some cases, be more efficient than order. This will be explored in further detail in the next section.

2.9 Can disorder be more efficient?

In the previous section we have observed that the contact network of a disordered sphere packing is consistently more interconnected than the equivalent network in lattice sphere packings, suggesting that disordered structures might be more stable or robust in comparison with the equivalent crystalline structures. In this section we show that disordered packings might indeed be more efficient than ordered, crystalline structures. In particular,

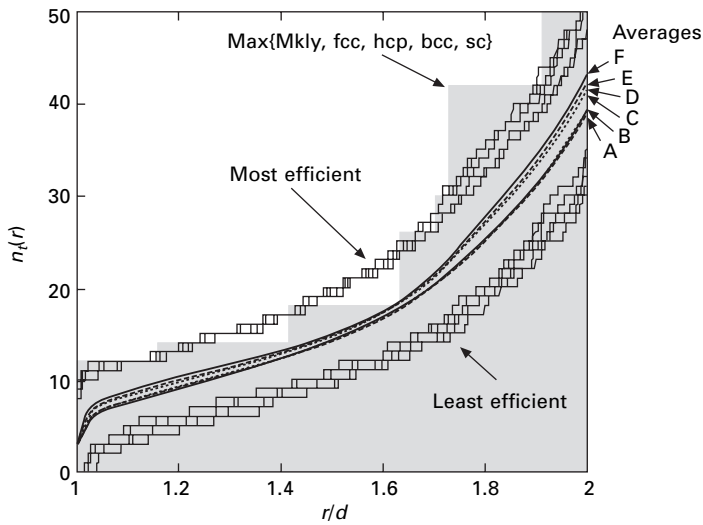


2.8 The coefficient *a,* plotted as a function of the average coordination number of the contact network (*n*), shows that disordered packings have larger topological densities in comparison with *lattice* sphere packings.

we investigate the geometrical organization in a spherical shell within a given radial distance r from the center of a particle and we compute the number of particles, the volume and the surface of the convex hull made from the particle-centers.

2.9.1 Packing efficiency

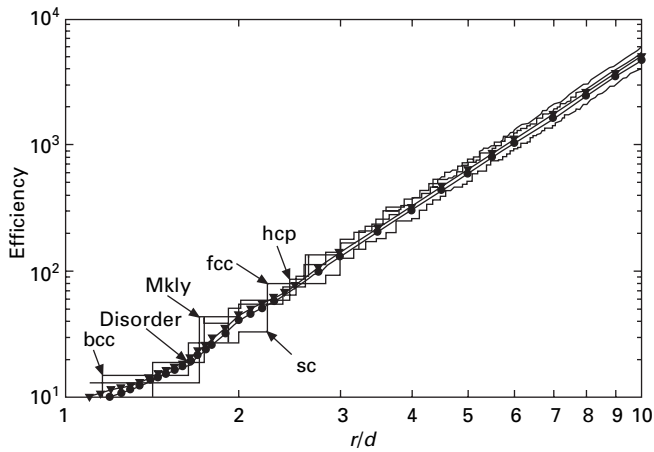
The number of sphere centers within a radial distance r from the center of a given sphere ($n_t(r)$) can be viewed as a measure of how efficiently locally dense agglomerates of spheres are formed. We call this quantity the *packing efficiency* (Aste *et al*. 2004). It is well known that no more than 12 spheres can be found in contact with one sphere (the 'Kissing number') (Aste and Weaire 2000), but the upper limit for the number of spheres within a given radial distance is, in general, unknown. Figure 2.9 shows the average, the maximum and the minimum numbers of neighbors within a given radial distance from any sphere in $\boldsymbol{G}$. Clearly, there are no neighbors up to distances close to $r \sim d$, when suddenly the number of neighbors increases very steeply and then, after this jump, it increases with the distance following a less steep trend with very comparable behaviors between all six samples A–F. It is of



2.9 Total number of sphere centers in a cluster made by all the spheres with centers within a given radial distance r from a given sphere. The thick lines are the average cluster sizes in (G) for samples A–F. The thin lines are the clusters with maximum/minimum number of spheres at a given distance. The filled area is bounded at the top by the most efficient (largest numbers of spheres) packings among the Mackay icosahedron (*Mkly*) and the lattices *fcc*, *hcp*, *bcc*, *sc*.

some interest to compare the empirical values of $n_t(r)$ obtained for these disordered samples with the known values in crystalline and other ordered structures.

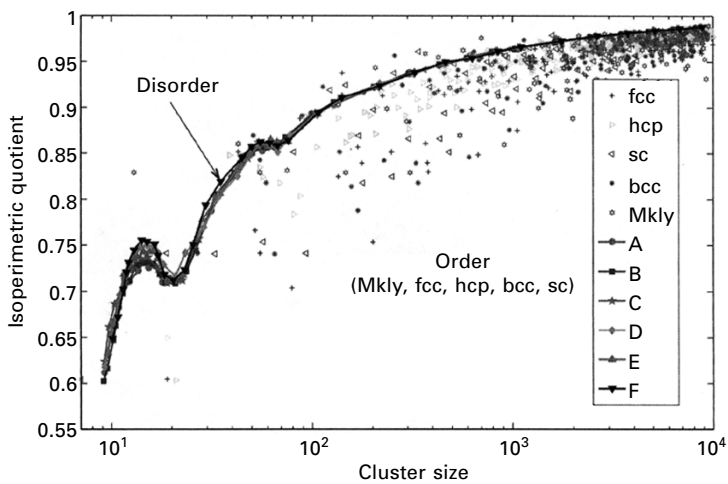
We observe that in a large range of radial distances between 1*d* and 2*d* there are some local configurations with packing efficiencies which are above the ordered ones. Moreover, in the region around $r \sim 1.6d$, disordered packings show better *average* packing efficiencies than all ordered ones (Mackay icosahedron (Mkly), face centered cubic (fcc), hexagonal closed packed (hcp), body centered cubic (bcc) and simple cubic (sc)). This is rather surprising if we consider that the *fcc* and *hcp* packings are more than 15% denser than the disordered packings. Such a high efficiency is counter-intuitive; disordered systems have a lower global density therefore, inside a given volume, there must be on average a lower number of sphere-centers than in the crystalline aggregate. On the other hand, we find that in disordered packings there are spherical regions which contain on average a larger number of spheres than the equivalent ordered packings. This, although being counter-intuitive, is not a contradiction; indeed one can observe from Figs 2.9 and 2.10 that at larger distances the crystalline packings are more efficient (*sc* excepted) than the disordered ones. A careful analysis of Fig. 2.10 shows that disordered packings have values of the efficiency which are comparable with the crystalline close packings up to radial distances of about three sphere diameters corresponding to clusters made of about 100 spheres. In other words, small aggregates of the order of 100 spheres can spontaneously pack into disordered configurations, achieving densities comparable with the Barlow packings.



2.10 Log–log plot of the average number of sphere centers within a radial distance *r* from a given sphere for the six samples (symbols). The full lines are the number of sphere centers within a radial distance *r* from a given sphere for the Mackay icosahedron (Mkly) and the lattices *fcc, hcp, bcc, sc*.

2.9.2 Isoperimetric quotient

When assemblies of relatively small numbers of spheres are concerned, the surface starts playing a very relevant role in determining the efficiency of the packing. Indeed, many physical systems such as atoms in metals can reduce their configurational energy by both increasing the packing fraction and decreasing the surface area. It is therefore interesting to calculate the surface area associated with the clusters of spheres within a radial distance r from a central sphere. A good measure of the relative weight between surface area and volume in a cluster is the 'isoperimetric quotient' which measures the ratio: $36\pi\ (volume)^2/(surface\ area)^3$, a quantity which is always smaller or equal to 1 and reaches this maximum value for a spherical shape. Figure 2.11 shows the isoperimetric quotients for the clusters resulting from all the spheres within a given radial distance from a central sphere in the samples A–F. Such a quantity is obtained by generating the convex hull from the set of sphere centers and then calculating the *volume* and *surface area* of such convex domain. From Fig. 2.11 one observes that the *isoperimetric quotient* of disordered clusters can be larger than the values for crystalline clusters with the same sizes, showing therefore that, also in this respect, disorder can be more efficient than order.



2.11 Average isoperimetric quotients $(36\pi(volume)^2/(surface\ area)^3)$ vs number of particles in the clusters formed by all the spheres within a radial distance from a given central sphere. The empty square symbols correspond to ordered and crystalline packings (*Mkly, fcc, bcc, hcp, bcc, sc*).

2.10 Conclusions

In this chapter the structural properties of disordered packings of equally sized spheres have been investigated. Our efforts have been devoted to discuss and uncover several specific features that are relevant for the understanding of the structure of amorphous systems. The most important concepts and the principal findings can be schematically summarized in the following points: (i) *disordered is not randomness*, characteristic lengths and similar local motifs are repeated (non-periodically) all over the systems; (ii) *amorphous systems are organized* and such organization can be detected and quantified with several techniques; (iii) such a structural organization *extends hierarchically* beyond first neighbors; (iv) there are *no signatures of (poly)-crystalline arrangements* – even at local level; (v) the contact network reveals a *higher degree of connectivity* than the equivalent crystalline structures, showing that disorder can be efficient and robust; (vi) the *local density* in small amorphous clusters, made by all the particles within a radial distance from a central particle, can be *higher* than in crystals and other ordered structures; (vii) such clusters can have *smaller surface extensions* (larger isoperimetric quotient) than the equivalent ordered structures.

The analyses and discussions presented in this chapter concern packings of monosized spheres. Such packings have been considered as the archetypal form for amorphous arrangements of particles in space. On the other hand, different physical systems have distinct properties that are consequences of different interactions between particles and/or are the outcomes of different dynamics. However, in this chapter we focused on some structural properties which are the results of geometrical and topological constraints only. Therefore, the laws and properties that we have uncovered are mostly determined by the rules governing the packing of objects in space. In the quest for universality, such laws and rules should be valid and applicable to a very broad range of systems.

2.11 Acknowledgments

Many thanks to T.J. Senden, M. Saadatfar, A. Sakellariou for several discussions and the tomographic data. We also thank A. Limaye for the packing images. This work was partially supported by the ARC discovery project DP0450292 and Australian Partnership for Advanced Computing National Facilities (APAC).

2.12 References

Aste, T. 1997. *The Shell Map, in Foams and Emulsions*, eds J. F. Sadoc and N. Rivier (NATO-SERIES, Kluwer Academic Publisher, Dordecht Netherlands 1997), ch. 29, 497–510.

Aste, T. 2005. Variations around disordered close packing, *J. Phys.: Condens. Matter* 17: S2361–S2390.

Aste, T., Coniglio, A. 2004. Cell theory for liquid solids and glasses: From local packing configurations to global complex behaviors, *Europhys. Lett.* **67**: 165–171.

Aste, T., Weaire, D. 2000. *The Pursuit of Perfect Packing*, Institute of Physics, Bristol.

Aste, T., Boosé, D., Rivier, N. 1996a. From one cell to the whole froth: A dynamical map, *Phys. Rev. E* **53**: 6181–6191.

Aste, T., Szeto, K. Y., Tam, W. Y. 1996b. Statistical properties and shell analysis in random cellular structures, *Phys. Rev*. E **54**: 5482.

Aste, T., Di Matteo, T., d'Agliano, E. G. 2002. Stress transmission in granular matter, *J. Phys.: Condens. Matter* **14**: 2391–2402.

Aste, T., Saadatfar, M., Sakellariou, A., Senden, T. 2004. Investigating the geometrical structure of disordered sphere packings, *Physica A* 339: 16–23.

Aste, T., Di Matteo, T., Hyde, S. T. 2005a. Complex networks on hyperbolic surfaces, *Physica A* **346**: 20–26.

Aste, T., Saadatfar, M., Senden, T. 2005b. The geometrical structure of disordered sphere packings, *Phys. Rev*. E. **71**: 061302.

Ball, R. C., Blumenfeld, R. 2002. Stress field in granular systems: Loop forces and potential formulation, *Phys. Rev. Lett.* **88**: 115505.

Bernal, J. D., Mason, J. 1960. Co-ordination of randomly packed spheres, *Nature* **188**: 910–911.

Bideau, D., Hansen, A. (eds) 1993. *Disordered and Granular Media*, North-Holland, Amsterdam.

Brunner, G. O., Laves, F. 1971. Zum problem der koordinationszahl, *Wiss. Z. Techn. Univ. Dresden* **20**: 387–390.

Chubynsky, M. V., Thorpe, M. F. 2001. Self-organization and rigidity in network glasses, *Current Opinion In Solid State & Materials Science* **5**: 525.

Clarke, A. S., Jónsson, H. 1993. Structural changes accompanying densification of random hard-sphere packings, *Phys. Rev. E* **47**: 3975–3984.

Conway, J., Sloane, N. 1997. Low dimensional lattices vii: Coordination sequences, *Proc. Royal Soc. London A* **453**: 2369.

Edwards, S. F., Grinev, D. V. 1999. Statistical mechanics of stress transmission in disordered granular arrays, *Phys. Rev. Lett.* **82**: 5397.

Grosse-Kunstleve, R. W., Brunner, G. O., Sloane, N. J. A. 1996. Algebraic description of coordination sequences and exact topological densities for zeolites, *Acta Cryst. A* **52**: 879–889.

Kansal, A. R., Torquato, S., Stillinger, F. H. 2002. Diversity of order and densities in jammed hard-particle packings, *Phys. Rev. E* **66**: 041109.

Kohonen, M. M., Geromichalos, D., Scheel, M., Schier, C., Herminghaus, S. 2004. On capillary bridges in wet granular materials, *Physica A* **339**: 7–15.

Lagrange, J. 1788. *Mécanique Analytique*, (Chez la Veuve Desaint, Paris).

Mason, G. 1968. Radial distribution function from small packings of spheres, *Nature* **217**: 733.

Mason, G., Clark, W. 1966. Fine structure in the radial distribution function from a random packing of spheres, *Nature* **211**: 957.

Maxwell, J. C. 1864. On the Calculation of the Equilibrium and Stiffness of Frames, *Philos. Mag.* **27**: 294.

Moukarzel, C. F. 1998. Isostatic phase transition and instability in stiff granular materials, *Phys. Rev. Lett.* **81**: 1634.

O'Keeffe, M. 1991a. Dense and rare four-connected nets, *Z. Kristallogr.* **196**: 21–37.
O'Keeffe, M. 1991b. N-dimensional diamond, sodalite and rare sphere packings, *Acta Cryst. A* **47**: 748–753.
O'Keeffe, M., Hyde, B. G. 1996. *Crystal Structures, Patterns and Symmetry*, Vol. 1, Mineralogical Society of America, Washington D.C.
Onoda, G. Y., Liniger, E. G. 1990. Random loose packings of uniform spheres and the dilatancy onset, *Phys. Rev. Lett.* **64**: 2727–2730.
Richard, P., Oger, L., Troadec, J.-P., Gervois, A. 1999a. Geometrical characterization of hard-sphere systems, *Phys. Rev. E* **60**: 4551.
Richard, P., Oger, L., Troadec, J.-P., Gervois, A. 1999b. Order and disorder in hard-sphere packings, *Europhys. Lett.* **48**: 415–420.
Scott, G. D. 1962. Radial distribution of random close packing of equal spheres, *Nature* **194**: 956–957.
Sederman, A., Alexander, P., Gladden, L. 2001. Structure of packed beds probed by magnetic resonance imaging, *Powder Technology* **117**: 255–269.
Seidler, G. T., Martinez, G., Seeley, L. H., Kim, K. H., Behne, E. A., Zaranek, S., Chapman, B. D., Heald, S. M., Brewe, D. L. 2000. Granule-by-granule reconstruction of a sandpile from x-ray microtomography data, *Phys. Rev. E* **62**: 8175.
Silbert, L. E., Ertas, D., Grest, G. S., Halsey, T. C., Levine, D. 2002. Geometry of frictionless and frictional sphere packings, *Phys. Rev. E* **65**: 031304.
Smith, W. O., Foote, P. D., Busang, P. F. 1929. Packing of homogeneous spheres, *Phys. Rev.* **34**: 1271–1274.
Steinhardt, P. J., Nelson, D. R., Ronchetti, M. 1983. Bond-orientational order in liquids and glasses, *Phys. Rev. B* **28**: 784.
Thorpe, M. F., Chubynsky, M. V. 2001. *Rigidity And Self-Organization Of Network Glasses And The Intermediate Phase*, Dordrecht: Kluwer Academic Press.
Torquato, S., Truskett, T. M., Debenedetti, P. G. 2000. Is random close packing of spheres well defined?, *Phys. Rev. Lett.* **84**: 2064–2067.
Torquato, S., Stillinger, F. 2001. Multiplicity of generation, selection, and classification procedures for jammed hard-particle packings, *J. Phys. Chem. B* **29**: 11849–11853.
Truskett, T. M., Torquato, S., Debenedetti, P. G. 2000. Towards a quantification of disorder in materials distinguishing equilibrium and glassy sphere packings, *Phys. Rev. E* **62**: 993.
Yang, R. Y., Zou, R. P., Yu, A. B. 2000. Computer simulation of the packing of fine particles, *Phys. Rev. E* **62**: 3900.

3

Nanostructure characterisation using electron-beam techniques

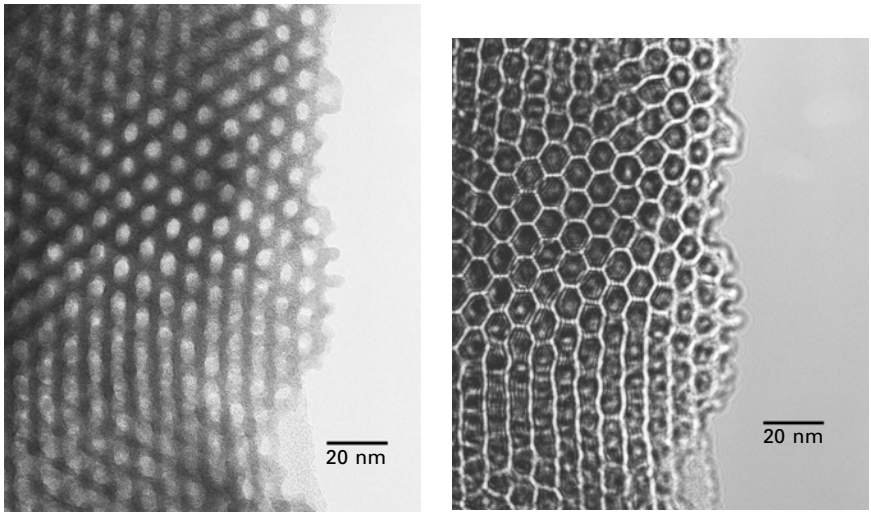
J RICHES and J DRENNAN,
University of Queensland, Australia

3.1 Introduction

The majority of papers that include the word 'nanotechnology' in the title contain somewhere in the text the mandatory micrograph, taken using an electron microscope. If the true definition of nanotechnology is adhered too, more often than not the instrument of choice is a transmission electron microscope (TEM) which is the only instrument capable of obtaining an image of the actual structure of the material. The necessity to actually 'see' the nanostructured material or nano-effect, ensures that microscopists continue to play a key role in the developing technology of nanostructural science.

However, with this demand comes a responsibility and each microscopist knows that the 'representative' picture may tell only part of the story. In addition there is a worrying trend away from any rigour in presenting electron micrographs. It is increasingly difficult to lay a ruler on a printed micrograph with any confidence – diffraction effects are not properly described and often are shown without any scale markings and the conditions under which they may be taken are lost in the prescriptive experimental conditions appearing at the beginning of the paper. In addition, phase contrast imaging is often misused in the examination of nano-scaled material. This technique can be prone to misinterpretation by inexperienced operators, as the image features can be strongly influenced by factors such as defocus and thickness. These effects can be quite complex and a full explanation is beyond the scope of this chapter, but can be found elsewhere.[1,2] A simple example of these effects is shown in Fig. 3.1, where an MCM-41 type mesoporous material has been imaged at two different focus conditions. The very different images produced by this single change clearly illustrates the challenge facing microscopists.

More importantly, as magnifications become higher and less material is examined – selectivity becomes more subjective. It is vital that researchers examine a sample thoroughly, ideally at different length scales to properly understand the presence or absence of inhomogeneities and resist the temptation to seek what they expect to find and ignore all else.

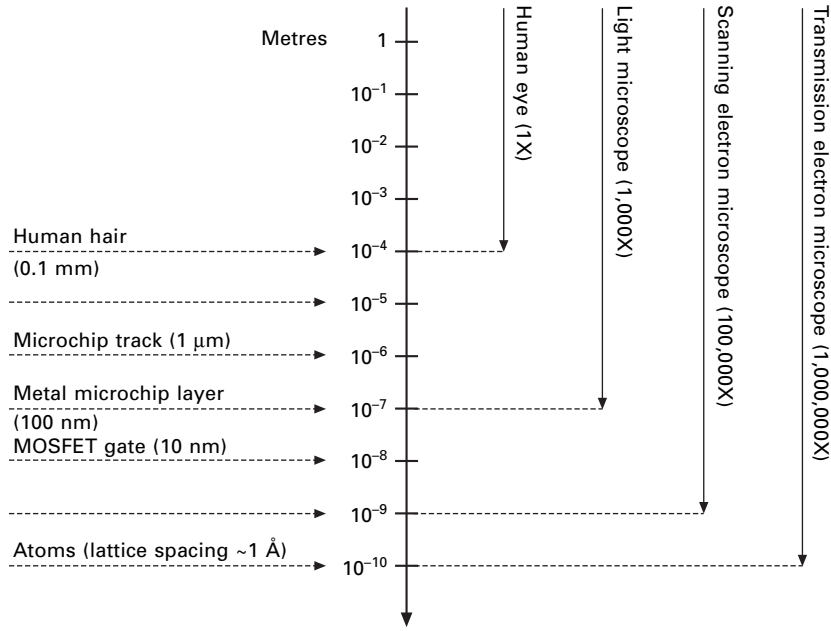


(a) (b)

3.1 The effect of varying defocus on the image of a TiO_2 MCM-41 type mesoporous material is shown here. In (a) the sample has been imaged close to focus and the circular nature of the pores is evident. In (b) the image has been taken at a relatively high degree of overfocus and the pores now appear to be hexagonal in shape.

In most cases, the microstructural image is supported by evidence obtained from other characterisation techniques such as surface area measurements or X-ray diffraction. It is the micrograph that can effectively confirm these observations by being able to provide a direct image of the structure of the material, when taken under the appropriate conditions. In this chapter, we examine some of the latest developments in examining nano-structured materials. We outline future trends and highlight some of the pitfalls of relying on the microscope to provide all the answers. Transmission electron microscopy can provide a wide range of image and analysis data and, correctly used with the images interpreted with some rigour, will continue to be the main tool of choice in examining materials at the nanoscale.

Figure 3.2 simply sets the scene for the following discussion and reminds the reader of the scales at which we are working in this new class of materials. The present resolution of TEMs is less than 2 Å and new developments in corrected systems, described in section 3.4.2, sees the capabilities being extended. Thus it is possible to directly image the lattice of a crystalline material in a suitable orientation and observe columns of atoms in the material. Interpretation of the data is important, and this chapter outlines new developments that can aid the nanotechnologists in extending the characterisation of the material in question and bring some awareness to the challenges associated with doing microscopy of this exciting new class of compounds.



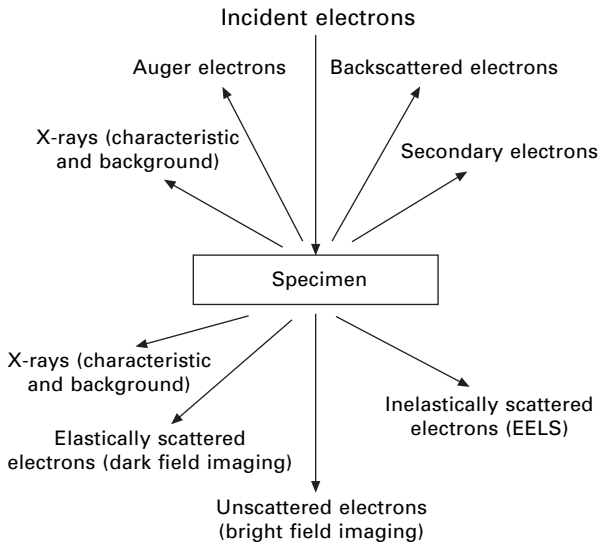
3.2 This shows the size scale corresponding to certain objects of scientific interest and the microscopes that are needed to image them. The relevance of electron microscopy to the examination of nanostructured materials is clear.

Before outlining the new developments it is worth reminding the reader of the flexibility of the modern electron microscope. Not simply a tool for observing features at high magnification, the myriad of effects that can be used when electrons interact with matter are extensive and can be used to good effect on specific challenges. Figure 3.3 graphically outlines the wide range of useful signals generated in the electron-specimen interaction. Although we will concentrate on transmission effects in this chapter, each emitted signal can be detected within the modern microscope and the ability to do this simply depends on the options that are fitted.

In summary, transmission electron microscopy can provide a powerful tool to supplement the characterisation of nano-sized materials. In some cases it is the only way the information can be obtained. However, a cautionary and sceptical eye must be kept to ensure that the interpretation of data is rigorous and representative.

3.2 Sample preparation

TEM specimens in materials science have traditionally been prepared using two techniques. The first technique, which is suitable for powders and small



3.3 This figure shows the electron specimen interactions that take place in an electron microscope, and some of the imaging modes corresponding to the transmitted electrons. While most interest focuses on the transmitted electrons and their role in providing high resolution structural information, the other signals provide equally valuable structural and compositional information.

grained materials, simply involves crushing the sample in a solvent and dispersing this on a holey carbon film. Regions that are thin enough to be electron transparent are then sought out and examined in the TEM. The second technique, which is used for bulk samples such as metals and ceramics, consists of mechanically thinning a specimen and then thinning it to electron transparency in an ion mill. In order to examine some nanostructured materials, it has been necessary to adapt these techniques, or to adopt new techniques from other fields. The two main examples that will be discussed here are the use of focused ion beams for precision sample preparation and the application of ultramicrotomy to nanostructured materials.

3.2.1 Focused ion beam system

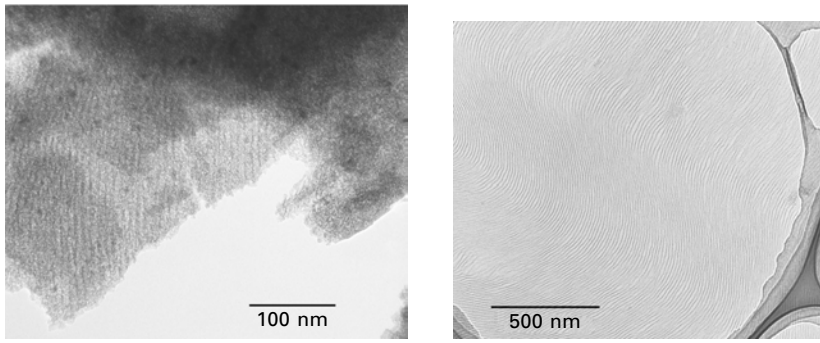
Traditional ion beam thinning, as a TEM sample preparation technique, suffers from some significant disadvantages. Firstly, it is difficult to accurately locate a region of interest on a sample and position this region to ensure that the optimal thinning occurs here. In addition, in multi-component systems preferential thinning often occurs, hindering the preparation of layered structures, interfaces and the like. These difficulties can be overcome by the use of a dual-beam focused ion beam (FIB) instrument. Originally used in

the semiconductor industry for the examination of subsurface features, they consist of an electron source, an ion source and a movable stage. The electron and ion sources operate on separate columns, which coincide at the position of the sample. The electron beam, operating as a scanning electron microscope (SEM), is used to locate the region of interest, and the ion beam is used to thin the material. The advantage here is that, due to the two independent columns, the progress of the sample preparation can be monitored in real time during the process and hence time-consuming mistakes avoided. Differences in milling rates of various components can be corrected by increasing or decreasing the time that the ion beam dwells on certain areas.

This technique has been applied to the examination of SiC and SiCN nanoparticles in nanostructured films and coatings.[3] By using the FIB it was possible to prepare a TEM specimen of these materials and retain their complex structural and chemical information. Dual beam FIBs have also been used to prepare cross-sectional TEM specimens of a specific carbon nanotube/metal junction.[4] The advantages of the FIB have been exploited in the development of a novel technique known as FIB nanotomography. This technique allows a three-dimensional image of the sample to be produced by sequential thinning and subsequent imaging of the surface of the material. This technique has been applied to porous $BaTiO_3$ ceramics[5] and was able to obtain a resolution of 7 nm in the X and Y directions and 17 nm in the Z direction. These capabilities could be effectively used in the characterisation of a wide range of macroporous and some mesoporous materials.

3.2.2 Ultramicrotomy

Ultramicrotomy has long been a standard TEM specimen preparation technique in the biological sciences and has recently been increasingly adopted by the materials science community. A good review of the process of ultramicrotomy and its application to materials science specimens is given elsewhere[6] and only a brief description of the technique will be given here. Essentially, the process consists of embedding material in resin and then cutting very thin (<70 nm) sections from this resin. These sections are generally thin enough to be electron transparent and the resin can often act to support the structure of the sample. There are a number of nanostructured materials that would benefit from this technique, including mesoporous particles and films.[7,8] Often these materials are several microns thick and attempts to produce specimens by grinding and dispersing on grids result in a very small visible area in the electron microscope. This is shown in Fig. 3.4, where the same material has been prepared in two different methods. The microtomed sample gives a very large thin area in comparison to the ground sample that has been dispersed on a carbon film. Ultramicrotomes have also been used to prepare



(a) (b)

3.4 A TEM specimen of a mesoporous TiO_2 film has been prepared in two different ways here. In (a) the sample has simply been ground in hexane and dispersed on a holey carbon grid. The variable thickness of the specimen results in limited areas around the edges being visible, with the interior being difficult to image. In (b) the film has been embedded in resin and microtomed and it is clear that large areas of the porous structure can be easily observed.

TEM specimens of Si nanodots and nanorods and this has allowed the crystal structure of a single nanorod to be accurately characterised.[9]

3.3 Nanoanalysis

While structural imaging at the nanoscale is one of the areas where electron microscopy can provide high-quality information, it is important not to overlook the compositional and crystallographic information that can be obtained from the electron specimen interactions. The three main forms of nanoanalysis used in the TEM are energy dispersive X-ray spectrometry (EDS), electron energy loss spectroscopy (EELS) and nanodiffraction. These will be briefly described, as will the application of energy filtering.

3.3.1 Energy dispersive X-ray spectrometry

The interaction between an incident electron and an atom in the specimen can result in the ionisation of the atom and the ejection of an inner shell electron. An X-ray can be produced as a secondary signal upon the transition of an outer shell electron to the now vacant inner shell. EDS works by detecting this X-ray with a cooled silicon single crystal detector and measuring its energy. The energies of these X-rays are characteristic of particular transitions within elements and hence are extremely useful for qualitatively determining the elements present within a sample. Quantitative information regarding elemental composition can also be obtained, if the appropriate standards are

used and procedures followed.[10] The technique of EDS is used in both SEM and TEM but as there are some differences between them they will be described separately.

EDS in an SEM

When performing EDS on thick samples in the SEM, the spatial resolution is considerably greater than the size of the electron beam. This is due to the scattering of the electron beam in the specimen. A single incident electron may be scattered several times, with large changes in its trajectory occurring. Therefore the characteristic X-rays emitted from later scattering events may take place some distance from the site of the electron beam. The volume from which these electron-specimen interactions take place, and from which various signals can escape, is known as the interaction volume, and its size depends upon the atomic number of the material and the energy of the incident electrons. For nanostructured materials this degradation of the spatial resolution is a serious problem and the only solution has been to reduce the accelerating voltage of the microscope and hence the energy of the incident electrons and the size of the interaction volume. This has the drawback, however, of decreasing the number of peaks in the energy spectrum that can be excited by the incident beam. There are numerous closely spaced peaks at these low energies and on conventional EDS detectors they overlap, preventing accurate identification and quantification. At higher accelerating voltages this is not a significant problem, as the position and intensity of the higher energy peaks can allow the accurate identification of the elements present. In the past this situation has forced researchers to compromise and choose either to acquire structural information at high resolution, or quality compositional information.

A new technology addressing some of these issues has recently been developed and the first commercial instrument has recently been installed.[11] Based on detecting very small changes in adsorbed energy as focused X-rays interact with a semiconductor, the X-ray microcalirometer promises both an order of magnitude increase in resolution over conventional detectors but also a lower limit for X-ray detection. This will become particularly important in analysing nanosized particles with small beam sizes of low accelerating voltage. This new detector has demonstrated an energy resolution for the Al-Kα transition at 1.5 keV of 7 eV (FWHM), which compares very favourably with the energy resolution of a conventional Si(Li) EDS detector in this range (~90 eV). The ability to obtain accurate compositional information from high-resolution SEM investigations of materials will be particularly valuable to the semiconductor industry.

EDS in a TEM

In contrast to the situation with the SEM, the spatial resolution of EDS in the TEM is essentially the same as the electron probe size itself. This is because the samples are very thin and the incident electrons have a much higher energy. Therefore there is less opportunity for multiple scattering to take place in the specimen that produces a large interaction volume. The widespread application of field emission guns (FEGs) has allowed much smaller electron probe sizes to be produced, as low as 1 Å in special cases and routinely less than 1 nm, and this has naturally improved the spatial resolution of normal EDS analyses. An interesting example of the spatial resolution and also the detectability limits obtainable with EDS in the TEM is shown by the detection of 2 Mn atoms in a Cu matrix, performed on a very thin specimen in a specially modified TEM.[12]

A further extension of the EDS method, used in TEMs, is being increasingly applied to analysing nanoscaled features. This method, developed almost a decade ago, combines crystallography with knowledge of X-ray detection to precisely locate atoms on the material under study. Atom location by channelling-enhanced microanalysis (ALCHEMI) uses the fact that the electron beam effectively sees different atomic arrangements when focused down well defined crystallographic directions.[13] In a recent example of this technique applied to a nanomaterials problem, Casey *et al.*[14,15] determined that a percentage of Co atoms were incorporated into nanosized ZnO oxide particles as opposed to being a surface coating. Subtle changes in the X-ray intensities recorded from precisely orientated ZnO oxide crystals confirmed the percentage of Co incorporated in the lattice. This is important information in that the photoactivity of the nanosized crystal particles is very dependent on size and dopant concentration.

3.3.2 Electron energy loss spectroscopy

EELS differs from EDS in that the signal that is detected is the forward scattered electron. These electrons have been inelastically scattered by the specimen and can have any energy up to the energy of the incident electrons. Scattering events that can be detected include the ionisation of atoms but also the production of plasmon or phonon oscillations within the sample. Examining the energies corresponding to the ionisation of various elements gives information on the elemental composition and this can provide similar information to EDS investigations, although the two techniques tend to be complementary in this sense, as EELS is superior to EDS at detecting light elements. In addition to obtaining compositional information it is also possible to obtain chemical bonding information from the near edge structure of these ionisation peaks. Finally the low-loss region of the energy loss spectrum can

be used to gain information about electronic states near the Fermi level in materials.

Similar to the case with EDS, the introduction of FEG electron sources and the much smaller electron probe sizes that they can generate has allowed researchers to investigate the composition and chemical structure of a variety of nanoscaled materials. This can be seen in the determination of structural[16] and electronic[17] properties of carbon nanotubes by examination of surface plasmon modes with EELS. Other examples are given in a useful overview of this nanoanalytical technique in the TEM.[18]

3.3.3 Nanodiffraction

One of the traditional strengths of TEM is its ability to investigate the crystallographic nature of materials, through the use of electron diffraction. A detailed description of the subject of electron diffraction is beyond the scope of this chapter, but good sources exist elsewhere[19–21] for those seeking a comprehensive treatment. For the purposes of this chapter it will suffice to say that the electrons are diffracted by crystalline material which satisfies certain orientational requirements and that it is possible to image this diffraction pattern in a TEM and use the pattern to obtain information about the crystal structure, lattice spacings and other useful information about the material.

The reduced electron probe sizes that the introduction of FEGs have enabled have also been beneficial for studies of the crystal structure of materials by electron diffraction. Originally, diffraction studies of materials were mainly performed using selected area electron diffraction, which used an aperture to limit the diffracting area and could only examine areas greater than 50–100 nm in diameter. Subsequent research developed the technique of convergent beam electron diffraction (CBED) which reduced the size of areas that could be examined to several nanometres. Variants of CBED, such as large angle CBED (LACBED), have also been shown to be applicable to the investigation of nanoparticles.[22] The reduced electron beam sizes that have been achieved through the use of FEGs and aberration corrections now mean that the technique of nanodiffraction can be used to extract crystallographic information from areas of less than 1 nm. A good overview of this field has been produced recently[23] but it is worth highlighting some of the specific examples.

Electron nanodiffraction is particularly suited to the examination of nanotubes, as it allows researchers to obtain diffraction patterns, and hence information about the crystal structure, from the different areas of the nanotubes (i.e. walls, interiors and also faults within the tubes).[24] This crystal structure determines the electronic properties of the nanotubes and is therefore of great interest to researchers. Also, nanotube samples are not all of the single walled variety, where the crystal structure is the same throughout the material,

but rather are often multiwalled or contain defects that affect the local crystal structure of the tube. In other cases it has been shown that the crystal structure of materials is distinctly different when the size drops below a certain threshold value than it is in the bulk material.[25] The change in properties of materials as the size is reduced is of course one of the defining characteristics of research into nanomaterials, and it is important to be able to accurately determine the size at which this effect takes place.

3.3.4 Energy filtering

The distribution of elements in nanoscaled materials often has a large effect on their properties, as can be seen in fields ranging from semiconductor or superconductor structures through to metal alloys. In the past it has been possible to produce elemental maps of materials in the TEM by focusing the beam to a small spot and acquiring an EDS spectrum from each point. Analysis of these spectra could show where particular elements were concentrated, and this information was quite valuable. However, the serial nature of the collection meant that it was necessarily a time consuming process, and issues such as sample drift made it difficult to obtain unambiguous information. In addition, focusing the beam to a spot produced high localised current densities and often resulted in problems with contamination and beam damage of the specimen.

An alternative approach to this problem that has become more prominent in recent years is the use of energy filtering. This technique takes advantage of the fact that electrons interact with the specimen and lose different amounts of energy, depending upon the nature of the interaction that they have undergone. This information can be used spectroscopically with the EELS technique described in section 3.3.2. Alternatively, an energy selecting slit can be used to produce an image using only electrons that have lost a certain amount of energy. If the energy loss is selected to correspond to an ionisation edge for a given element, one can effectively produce a map of the concentration of that element in the specimen. This is a parallel collection technique, known as energy filtered transmission electron microscopy (EFTEM), and is capable of producing a map of an area in less than a minute, as compared with the hours that are generally needed to produce high quality EDS maps. This reduced collection time minimises the effect of sample drift and provides more certainty over the location of specific elements. The technique is described in detail elsewhere[26] and has been used to locate individual Gd atoms in C60 buckminsterfullerenes, that themselves lie within a carbon nanotube.[27] The elemental distribution of oxygen monolayers in a $Al_{11}O_3N_9$ sample at 5 Å resolution has also been obtained, clearly demonstrating the high spatial resolution that is possible with this technique.[28] Another example of the power of the EFTEM technique is shown by the investigation of an amorphous

carbon layer on a diamond film. An EDS analysis of this material would have failed to distinguish between the two layers, as the elemental composition of the two are identical. However, the different bonding environments in the two produce differences in the near edge structure that can be used to show that the thickness of the amorphous layer was between one and two nanometres.[29]

3.4 Imaging

The most important application of electron microscopy to the characterisation of nanostructured materials is the direct observation of their structure. No other technique can characterise the structure of the materials with the same degree of confidence and TEM is often used to confirm the details that are inferred from other characterisation techniques. There have been two important advances in recent years to the imaging capabilities of the transmission electron microscope and these are the adoption of tomographic imaging techniques and the development of spherical aberration corrected TEMs.

3.4.1 Tomography

Control of the structure in three dimensions is very important for many nanoscaled materials. For example, mesoporous materials are used as catalyst supports as their surface area is much greater than a solid catalyst support of a similar size. However, to take advantage of this superior surface area it is imperative that metallic catalyst particles infiltrate the pore network and do not simply attach to the external surfaces. Conventional TEM techniques give a two-dimensional projection of the material, and the information in the z-direction is lost in this process. In the example mentioned above, this makes it impossible to tell from one projection whether the catalysts reside within the pores, or on the external surfaces.

One solution to this problem, practised in biological electron microscopy for many years,[30–32] is electron tomography. This technique involves taking two-dimensional projections of the material at a range of tilt values and then reconstructing the initial sample using a back-projection algorithm. The details of this technique are discussed elsewhere in more detail,[33] but briefly it involves projecting the density of the image back in the direction of the illumination. This is done for all tilt values and where several of these back-projections intersect there will be more mass.

Recently electron tomography has been applied to a range of materials science problems, including the structure of pores in zeolites[34,35] and mesoporous materials.[36,37] The three-dimensional morphology of polymer blends[38,39] has also been studied. Finally microelectronic components[40] have been investigated with electron tomography and it is expected that these

materials will be increasingly studied in this fashion, as feature sizes continue to shrink and quantum effects become more important.

While it is useful to produce three-dimensional reconstructions of various nanoscaled materials, often the composition of the materials of interest is inhomogeneous and it is important to determine the distribution of different elements throughout the sample. To achieve this feat, EFTEM is used, as described in section 3.3.4, to produce the images comprising the tilt series. The reconstruction of these images is performed in exactly the same way as for conventional electron tomography. This technique is still in early stages of development and to date it has been used to analyse the composition of grain boundaries in stainless steel[41] and also the location of nanoparticles in intermetallic alloys.[42] One drawback of this technique is the lower signal-to-noise ratio obtained with EFTEM images. To compensate for this reduced signal it is necessary to increase the electron dose that the sample is subjected to and this can cause unacceptable levels of damage to beam-sensitive materials.

Another limitation of electron tomography as described above is that the reconstruction algorithm depends on the contrast in the sample arising solely from differences in mass or thickness. Contrast arising from the diffraction of crystalline materials can be quite strong at specific tilt angles and weak at others, and this causes the reconstruction algorithm to give spurious results. This obviously limits the application of this technique to samples that do not display significant diffraction contrast, i.e., amorphous materials. While there are many useful amorphous nanoscaled materials, it would be advantageous to be able to examine crystalline materials as well. In order to achieve this, a refinement of the standard electron tomography technique has been developed, known as high-angle annular dark field scanning transmission electron microscopy (HAADF-STEM) tomography.[41] This technique is performed in STEM mode and takes advantage of the fact that the scattering of electrons to high angles is primarily dependent on the atomic number of the material and is not strongly influenced by the crystal orientation of the material. Using this technique it has been possible to image crystalline materials, such as bimetallic nanocatalysts[43] and magnetotactic bacteria,[41] that were unsuitable for imaging using conventional electron tomography.

3.4.2 C_s corrected instrumentation

The resolution obtainable in TEMs has traditionally been limited by lens aberrations such as spherical aberration and chromatic aberration. These aberrations cause points in the specimen to form a disc in the image plane and hence degrade the image that we can produce. With spherical aberration this comes about due to the inhomogeneous strength of the objective lens, with off-axis rays being deflected more than those travelling near the optic axis. With chromatic aberration the same effect occurs due to the objective

lens strength varying depending on the energy of the electrons, with less energetic electrons being deflected more. A more complete explanation of the effect of aberrations on the resolution of the instrument can be found elsewhere.[1]

In order to improve the resolution of the TEM, several groups[44–46] have developed additional elements to compensate for the spherical aberration present in the objective lens of the TEM. The chromatic aberration has also been reduced through the introduction of monochromators immediately after the electron gun.[47] Using these compensating elements it is possible to improve the resolution of the microscope to better than 1 Å. This improvement in resolution has allowed researchers to directly image light elements, including oxygen[48–50] and lithium atoms.[51] These elements, which were not previously able to be imaged due to their low atomic weight and consequent low electron scattering ability, are very important for novel energy storage materials being developed.

The correction of lens aberrations in the microscope, coupled with the use of FEGs, has also allowed researchers to produce electron beams with a diameter of less than 1 Å. This has important implications for the use of TEMs as an analytical tool, as is described in section 3.3, but it also allows the improvement of the obtainable resolution when the instruments are used in STEM mode. In this mode a very fine electron beam is produced and rastered over the sample, with the signal produced from each point on the sample that the electron beam dwells on corresponding to a pixel on the display. When operated in this mode, the resolution of the microscope is limited by the size of the electron beam that can be formed. An aberration corrected TEM/STEM has achieved a 0.8 Å sized probe and been used to reconstruct a CdSe semiconductor nanocrystal.[52] The improved resolution is extremely important in semiconductor nanocrystals, or quantum dots as they are commonly known, as the electronic and optical properties depend strongly on the size and surface structure of the materials.[53] High resolution structural information of this nature, combined with spectroscopic studies of single quantum dots,[54] should help advance fundamental understanding of the science behind these materials and also enhance their prospects for applications.

3.5 Future trends

The development of techniques and instrumentation in electron microscopy is continuing at a rapid pace, driven by the requirements of users. Three of the main areas where advances are currently being made are discussed below and these are; the improvements in resolution through aberration corrections, the automation of many of the tasks that are undertaken on an electron microscope and finally the analysis of the large quantities of data that are

produced by the microscope and the identification of valuable information within this data.

3.5.1 Improved resolution

As described in section 3.4.2, there is a great deal of development currently under way in correcting the spherical and chromatic aberration that limits the resolution of transmission electron microscopes. Sub-Angstrom sized beams are currently being achieved in development projects with further improvements being predicted[55,56] and it is expected that this development will lead to sub-Angstrom sized beams being regularly obtained in commercially available instruments. The improved resolution that this will bring about will allow technologically important materials such as lithium and oxygen to be routinely imaged and many other materials to be examined in a wider range of orientations, instead of a limited number of low angle zone axes that is currently possible. In addition, the small probe sizes that will be able to be obtained in STEM imaging will provide unparalleled spatial resolution for analytical techniques, such as EDX or EELS. Improved monochromated beams in the microscopes will improve the energy resolution of EELS and also enhance EDS and EFTEM studies of nanometre sized materials by improving the beam current of a probe of a given size.

3.5.2 Automated microscopy

One of the most notable trends in modern science is the push towards high throughput examination of samples as a means of discovery. This has become one of the primary forms of research in biological fields and it could also prove very powerful in the field of materials science. However, current methods of structural characterisation such as electron microscopy constitute serious bottlenecks in a high throughput pipeline. Typically, sample preparation can take days, the microscope requires highly skilled operators, and depending upon the sample and the information that is required, only a few samples can be examined in a day.

To overcome this stumbling block there has been a general push to automate as many aspects of TEM as possible. Software is currently available to assist with the alignment of the TEM, speeding up this process and allowing more time for the analysis of samples, and simultaneously reducing the skill level needed for operators. Other software has been written to automate the image collection process,[57] and while this is aimed at biological studies, it is sufficiently flexible to be easily adapted to certain materials problems. In addition, an ambitious project is under way to automate the sample loading of a TEM. Initial results have shown that it is possible to automatically load and examine 96 grids in just over 24 hours, producing low magnification

images useful for screening samples.[58] The introduction of this system into mainstream TEMs would greatly improve the productivity of these instruments, allowing them to be in operation essentially 24 hours a day.

Another area where automation is playing an important role is that of crystallography in the TEM. The automatic determination of the orientation of polycrystalline materials is commonplace in SEMs, using electron backscattered diffraction, but the resolution of this technique is limited to around 100 nm. With nanostructured materials, it is often the case that grain sizes of interest will be of the order of several nanometres, or less, and this technique is clearly insufficient to address these problems. To solve this, another automated crystallography method has been developed for the TEM,[59] where the spatial resolution is much better, and initial results have shown this system to be effective in studying the grain boundary segregation in Cu-Sb alloys.[60] It is expected that the use of this system will become more widespread in the future.

3.5.3 Knowledge mining

Advanced information technologies are playing a significant role in nanostructural characterisation by improving the capture, analysis and correlation of the massive digital data and image sets being generated. The application of sophisticated statistical analysis tools (such as R[61]) to existing data sets are driving experimental design, reducing duplication and ensuring only optimum data is being captured. Dynamic, flexible workflow management systems (based on Web Services and BPEL4WS[62]) are streamlining the effort in capturing and analysing experimental data and metadata, choreographing the sub-activities, notifying experimenters of uncompleted tasks, precisely recording provenance data (processing and instrument settings, time, place, etc.) and validating scientific data before saving it to a database and invoking the next task. The Internet and data interchange standards such as XML[63] are facilitating the establishment of secure online distributed databases. Metadata standards such as OME (open microscopy environment),[64] MPEG-7[65] and OWL (ontology web language)[66] are enabling the indexing, retrieval, comparison and integration of heterogeneous microscopic data sets. Large-scale, high-speed processing and interpretation of 2D and 3D images is possible through online services such as Multi-MATLAB[67] and Rules-By-Example.[68] Semantically annotated content and Semantic Web technologies[69] (e.g., RDF, OWL, RuleML) enable on-the-fly integration, reasoning and more advanced search and querying. Finally, new sophisticated, intelligent data exploration and visualisation systems such as FUSION[70] are capable of assimilating heterogeneous, multi-dimensional information and data into coherent, interactive and animated presentations – enabling scientists more easily to determine relationships between processing parameters and

nanostructure and to formulate more accurate predictive models of nanostructural behaviour.

3.6 Acknowledgements

We wish to thank Dr Jane Hunter for her valuable contribution to the manuscript, and also the staff of the Centre for Microscopy and Microanalysis, for helpful discussions.

3.7 References

1. Spence, J. C. H. *Experimental high-resolution microscopy* (Oxford University Press, New York, 1988).
2. Buseck, P. R., Cowley, J. M. C. and Eyring, L. (ed.) *High-resolution transmission electron microscopy and associated techniques* (Oxford University Press, New York, 1988).
3. Perrey, C. R. *et al*. Using the FIB to characterize nanoparticle materials. *Journal of Microscopy-Oxford* **214**, 222–236 (2004).
4. Wei, B. Q. *et al*. Selective specimen preparation for TEM observation of the cross-section of individual carbon nanotube/metal junctions. *Ultramicroscopy* **85**, 93–98 (2000).
5. Holzer, L., Indutnyi, F., Gasser, P. H., Munch, B. and Wegmann, M. Three-dimensional analysis of porous BaTiO3 ceramics using FIB nanotomography. *Journal of Microscopy-Oxford* **216**, 84–95 (2004).
6. Glanvill, S. R. Ultramicrotomy of Semiconductors and Related Materials. *Microscopy Research and Technique* **31**, 275–284 (1995).
7. Grosso, D. *et al*. Two-dimensional hexagonal mesoporous silica thin films prepared from black copolymers: Detailed characterization amd formation mechanism. *Chemistry of Materials* **13**, 1848–1856 (2001).
8. Yu, C. Z., Fan, J., Tian, B. Z. and Zhao, D. Y. Morphology development of mesoporous materials: a colloidal phase separation mechanism. *Chemistry of Materials* **16**, 889–898 (2004).
9. Teo, B. K. *et al*. Precision-cut crystalline silicon nanodots and nanorods from nanowires and direct visualization of cross sections and growth orientations of silicon nanowires. *Nano Letters* **3**, 1735–1737 (2003).
10. Williams, D. B., Goldstein, J. I. and Newbury, D. E. *X-Ray Spectrometry in Electron Beam Instruments* (Plenum Press, New York, 1995).
11. Hollerith, C., Wernicke, D. and Buhler, M. Energy dispersive X-ray spectroscopy with microcalorimeters. *Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment* **520**, 606–609 (2004).
12. Watanabe, M. and Williams, D. B. Atomic-level detection by X-ray microanalysis in the analytical electron microscope. *Ultramicroscopy* **78**, 89–101 (1999).
13. Jones, I. P. in *Advances in Imaging and Electron Physics,* Vol. 125 63–117 (2002).
14. Casey, P. S., Rossouw, C. J., Boskovic, S., Lawrence, K. and Turney, T. W. in *International Congress for Particle Technology* (Nuremberg, 2004).
15. Rossouw, C. J., Casey, P. S. and Turney, T. W. in *18th Australian Conference on Microscopy and Microanalysis* (Geelong, 2004).

16. Reed, B. W. and Sarikaya, M. Electronic properties of carbon nanotubes by transmission electron energy-loss spectroscopy. *Physical Review B* **64**, art. no.-195404 (2001).
17. Kociak, M., Henrard, L., Stephan, O., Suenaga, K. and Colliex, C. Plasmons in layered nanospheres and nanotubes investigated by spatially resolved electron energy-loss spectroscopy. *Physical Review B* **61**, 13936–13944 (2000).
18. Thomas, J. M. and Midgley, P. A. High-resolution transmission electron microscopy: the ultimate nanoanalytical technique. *Chemical Communications*, 1253–1267 (2004).
19. Cowley, J. M. *Electron diffraction techniques* (Oxford University Press, Chester, 1992–1993).
20. Edington, J. W. *Electron diffraction in the electron microscope* (Macmillan, London, 1975).
21. Williams, D. B. and Carter, C. B. *Transmission Electron Microscopy: A Textbook for Materials Science* (Plenum Press, New York, 1996).
22. Yao, Y. M. and Tholen, A. R. Characterization of contacting boundaries between nanoparticles with LACBED. *Microscopy and Microanalysis* **9**, 237–244 (2003).
23. Cowley, J. M. Applications of electron nanodiffraction. *Micron* **35**, 345–360 (2004).
24. Cowley, J. M. in *Electron Microscopy of Nanotubes* (eds Wang, Z. L. and Hui, C.) 73–93 (Kluwer Academic Publishers, Boston, 2003).
25. Mani, R. C. *et al*. Synthesis and electrochemical characteristics of a nanocomposite diamond electrode. *Electrochemical and Solid State Letters* **5**, E32-E35 (2002).
26. Reimer, L. (ed.) *Energy Filtering Transmission Electron Microscopy* (Springer Verlag, Berlin, 1995).
27. Suenaga, K. *et al*. Element-selective single atom imaging. *Science* **290**, 2280–2282 (2000).
28. Mitome, M. *et al*. Nanoanalysis by a high-resolution energy filtering transmission electron microscope. *Microscopy Research and Technique* **63**, 140–148 (2004).
29. Mayer, J. and Plitzko, J. M. Mapping of ELNES on a nanometre scale by electron spectroscopic imaging. *Journal of Microscopy-Oxford* **183**, 2–8 (1996).
30. Hart, R. G. Electron microscopy of unstained biological material: the polytropic montage. *Science* **159**, 1464–1467 (1968).
31. Baumeister, W., Grimm, R. and Walz, J. Electron tomography of molecules and cells. *Trends in Cell Biology* **9**, 81–85 (1999).
32. Frank, J. Approaches to Large-Scale Structures. *Current Opinion in Structural Biology* **5**, 194–201 (1995).
33. Frank, J. *Electron tomography: three-dimensional imaging with the transmission electron microscope* (Plenum Press, New York, 1992).
34. Janssen, A. H., Koster, A. J. and de Jong, K. P. Three-dimensional transmission electron microscopic observations of mesopores in dealuminated zeolite Y. *Angewandte Chemie-International Edition* **40**, 1102–1104 (2001).
35. Koster, A. J., Ziese, U., Verkleij, A. J., Janssen, A. H. and de Jong, K. P. Three-dimensional transmission electron microscopy: A novel imaging and characterization technique with nanometer scale resolution for materials science. *Journal of Physical Chemistry B* **104**, 9368–9370 (2000).
36. de Jong, K. P. and Koster, A. J. Three-dimensional electron microscopy of mesoporous materials – Recent strides towards spatial imaging at the nanometer scale. *Chemphyschem* **3**, 776–780 (2002).
37. Janssen, A. H., Van Der Voort, P., Koster, A. J. and de Jong, K. P. A 3D-TEM study of the shape of mesopores in SBA-15 and modified SBA-15 materials. *Chemical Communications*, 1632–1633 (2002).

38. Spontak, R. J. *et al*. Phase behavior of ordered diblock copolymer blends: Effect of compositional heterogeneity. *Macromolecules* **29**, 4494–4507 (1996).
39. Yamauchi, T., Matsuba, Y., Ohyama, Y., Tabuchi, M. and Nakamura, A. Quantum size effects of InAs- and InGaAs-quantum dots studied by scanning tunneling microscopy/spectroscopy. *Japanese Journal of Applied Physics Part 1-Regular Papers Short Notes & Review Papers* **40**, 2069–2072 (2001).
40. Stegmann, H., Engelmann, H. H. and Zschech, E. Characterization of barrier/seed layer stacks of Cu interconnects by electron tomographic three-dimensional object reconstruction. *Microelectronic Engineering* **65**, 171–183 (2003).
41. Midgley, P. A. and Weyland, M. 3D electron microscopy in the physical sciences: the development of Z-contrast and EFTEM tomography. *Ultramicroscopy* **96**, 413–431 (2003).
42. Mobus, G. and Inkson, B. J. in *Electron Microscopy and Analysis 2001* 267–270 (2001).
43. Midgley, P. A. *et al*. High-resolution scanning transmission electron tomography and elemental analysis of zeptogram quantities of heterogeneous catalyst. *Journal of Physical Chemistry B* **108**, 4590–4592 (2004).
44. Batson, P. E., Dellby, N. and Krivanek, O. L. Sub-angstrom resolution using aberration corrected electron optics. *Nature* **418**, 617–620 (2002).
45. O'Keefe, M. A. *et al*. Sub-Angstrom high-resolution transmission electron microscopy at 300 keV. *Ultramicroscopy* **89**, 215–241 (2001).
46. Pennycook, S. J. *et al*. Aberration-corrected scanning transmission electron microscopy: the potential for nano- and interface science. *Zeitschrift für Metallkunde* **94**, 350–357 (2003).
47. Mook, H. W. and Kruit, P. Construction and characterization of the fringe field monochromator for a field emission gun. *Ultramicroscopy* **81**, 129–139 (2000).
48. Jia, C. L., Lentzen, M. and Urban, K. Atomic-resolution imaging of oxygen in perovskite ceramics. *Science* **299**, 870–873 (2003).
49. Jia, C. L. and Urban, K. Atomic-resolution measurement of oxygen concentration in oxide materials. *Science* **303**, 2001–2004 (2004).
50. Kisielowski, C. *et al*. Imaging columns of the light elements carbon, nitrogen and oxygen with sub Angstrom resolution. *Ultramicroscopy* **89**, 243–263 (2001).
51. Shao-Horn, Y., Croguennec, L., Delmas, C., Nelson, E. C. and O'Keefe, M. A. Atomic resolution of lithium ions in LiCoO2. *Nature Materials* **2**, 464–467 (2003).
52. McBride, J. R., Kippeny, T. C., Pennycook, S. J. and Rosenthal, S. J. Aberration-corrected Z-contrast scanning transmission electron microscopy of CdSe nanocrystals. *Nano Letters* **4**, 1279–1283 (2004).
53. Alivisatos, A. P. Semiconductor clusters, nanocrystals, and quantum dots. *Science* **271**, 933–937 (1996).
54. Gammon, D. High-resolution spectroscopy of individual quantum dots in wells. *Mrs Bulletin* **23**, 44–48 (1998).
55. Cowley, J. M., Spence, J. C. H. and Smirnov, V. V. The enhancement of electron microscope resolution by use of atomic focusers. *Ultramicroscopy* **68**, 135–148 (1997).
56. Krivanek, O. L., Nellist, P. D., Dellby, N., Murfitt, M. F. and Szilagyi, Z. Towards sub-0.5 angstrom electron beams. *Ultramicroscopy* **96**, 229–237 (2003).
57. Carragher, B. *et al*. Leginon: An automated system for acquisition of images from vitreous ice specimens. *Journal of Structural Biology* **132**, 33–45 (2000).

58. Potter, C. S., Pulokas, J., Smith, P., Suloway, C. and Carragher, B. Robotic grid loading system for a transmission electron microscope. *Journal of Structural Biology* **146**, 431–440 (2004).
59. Wright, S. I. and Dingley, D. J. in *Texture and Anisotropy of Polycrystals* 209–214 (1998).
60. Li, C. F. and Williams, D. B. Application of automated crystallography for transmission electron microscopy in the study of grain-boundary segregation. *Micron* **34**, 199–209 (2003).
61. The R Project for Statistical Computing, http://www.r-project.org/
62. Business Process Execution Language for Web Services Version 1.1, May 2003, http://www-128.ibm.com/developerworks/library/ws-bpel/
63. World Wide Web Consortium (W3C), Extensible Markup Language (XML), August 2004, http://www.w3.org/XML/
64. Open Microscopy Environment (OME), http://www.openmicroscopy.org/
65. ISO/IEC 15938-5 FDIS Information Technology, Multimedia Content Description Interface – Part 5: Multimedia Description Schemes: MPEG7, Sydney. July 2001.
66. McGuiness, D. and van Harmelen, F., OWL Web Ontology Language Overview, W3C Recommendation, February 2004. http://www.w3.org/TR/owl-features/
67. Trefethen, A., Menon, V., Chang, C., Czajkowski, G., Myers, C. and Trefethen, L. N., MultiMATLAB: MATLAB on Multiple Processors, Cornell University, Ithaca, NY, 1996.
68. Little, S. and Hunter, J., Rules-By-Example – a Novel Approach to Semantic Indexing and Querying of Images, 3rd International Semantic Web Conference (ISWC2004). Hiroshima, Japan, November 2004. pp. 534–548.
69. World Wide Web Consortium (W3C), Semantic Web, July 2004. http://www.w3.org/2001/sw/
70. Hunter, J., Falkovych, K. and Little, S., Next Generation Search Interfaces – Interactive Data Exploration and Hypothesis Testing, 8th European Conference on Digital Libraries (ECDL2004). Bath, UK, September 2004. pp. 86–98.

4

Organic–inorganic nanocomposite membranes for molecular separation processes

T MERKEL and I PINNAU,
Membrane Technology and Research, Inc., USA

4.1 Introduction

The discovery of novel materials with unique properties can profoundly impact existing technologies, as well as enable development of entirely new products and processes. Such breakthrough potential has led to considerable research into organic–inorganic composites. Through hybridization one hopes to capture the desirable attributes of the organic and inorganic components in a single material, and possibly to discover new synergistic properties. In most cases, the best property enhancements are achieved when the inorganic and/or organic components are of sufficiently small dimensions to blend at a molecular level. Such nanocomposites frequently take the form of nanoscale inorganic particles dispersed in an organic matrix. For example, addition of nanoparticles to polymers has been found to yield nanocomposites with enhanced conductivity,[1,2] mechanical toughness,[3] and catalytic activity.[4,5]

Nanocomposites may also prove to be useful for molecular separations, a diverse field affecting processes such as biomolecule purification, environmental remediation, seawater desalination, and petroleum chemicals and fuel production.[6] Separations in these applications are often accomplished using energy-intensive, established technologies such as distillation, absorption, and adsorption. Membrane separation can be an attractive alternative because membranes are a low-cost, energy-efficient, green technology. Their widespread use in separation applications has been limited, however, by the difficulty of preparing membranes with the desirable combination of high selectivity, which yields high product purity and low operating costs, and high permeability, which reduces membrane area and capital cost. Unfortunately, as selectivity of conventional polymer membrane materials increases, permeability invariably decreases and vice versa.[7,8] Recent efforts to circumvent this fundamental limitation have explored the addition of small inorganic particles to organic polymers. While the polymer phase offers mechanical elasticity and processability, the inorganic particles can be used either to sieve molecules, for example, utilizing the spatially well-defined pores of a zeolite,[9–11] or to control packing

of polymer chains and thereby transport characteristics.[12,13] In this chapter, we review the history and current state-of-the-art for molecular separations using organic/inorganic nanocomposite membranes.

4.2 Transport in dense membranes

Gas and vapor permeation through non-porous polymer or nanocomposite membranes is usually rationalized by the solution-diffusion model.[14] This model assumes that the gas phases on either side of a membrane are in thermodynamic equilibrium with their respective interfaces, and that the interfacial sorption process is rapid compared with the rate of diffusion through the membrane. Thus, the rate-limiting step for permeation is diffusion through the membrane, which is governed by Fick's law of diffusion. In simple cases, Fick's law leads to the equation

$$P = \frac{Nl}{\Delta p} = S \cdot D \tag{4.1}$$

where P is the permeability coefficient, a pressure- and thickness-normalized transmembrane flux [cm^3(STP) · cm/cm^2 · s · cmHg], N is the volumetric flux per unit membrane area [cm^3(STP)/cm^2 · s], l is the membrane thickness [cm], and Δp is the pressure difference across the membrane [cmHg]. Permeability is frequently expressed in Barrers, where 1 Barrer = $10^{-10} \frac{cm^3(STP) \cdot cm}{cm^2 \cdot s \cdot cm\ Hg}$. As indicated in eqn 4.1, the permeability of a membrane depends on the solubility coefficient, S, of the gas or vapor in the membrane material and the permeant diffusion coefficient, D, within the membrane.

For a given membrane material, the ideal selectivity, $\alpha_{A/B}$, for gas A over gas B is defined as the ratio of gas permeability coefficients:

$$\alpha_{A/B} = \frac{P_A}{P_B} = \left(\frac{S_A}{S_B}\right) \cdot \left(\frac{D_A}{D_B}\right) \tag{4.2}$$

where P_A and P_B are the permeability coefficients of gases A and B, respectively, determined from isotropic films of known thickness. As shown in eqn 4.2, the ideal selectivity may be expressed as the product of a solubility selectivity and a diffusion selectivity. The diffusion selectivity always favors smaller molecules, whereas the solubility selectivity often favors larger, more condensable species. The overall selectivity depends on the relative contributions from these two factors.

4.3 Transport in composite materials

The effect of small inorganic nanoparticles on molecular transport through

polymer membranes depends first on whether the particles are porous or non-porous. A porous particle, such as a zeolite, may allow molecules to travel through the inside of the particle, and in fact may do so selectively, thereby sieving molecules. As will be shown later, this is one of the primary reasons for incorporating inorganic particles into a polymer membrane. In contrast, a non-porous particle is an impermeable object around which molecules must traverse.

Numerous theoretical expressions have been developed to explain transport behavior in composite polymer systems.[15–18] A simple and frequently used model, first proposed for analyzing steady-state dielectric properties of a dilute suspension of hard spheres, is that of Maxwell.[19] Applied to permeation in a binary composite system where both phases are permeable, Maxwell's model yields the following expression:

$$P_{eff} = P_C\left(\frac{P_D + 2P_C - \phi_D(P_C - P_D)}{P_D + 2P_C + \phi_D(P_C - P_D)}\right) \tag{4.3}$$

where P_{eff} is the permeability of the composite, P_C is the permeability of the continuous (polymer) phase, P_D is the permeability of the dispersed (porous particle) phase, and ϕ_D is the volume fraction of the dispersed particles. Several recent studies comparing theoretical models have concluded that Maxwell's expression is most appropriate for predicting transport properties in composite membrane materials.[18,20,21]

If the particles dispersed in a continuous polymer matrix are non-porous, and, therefore impermeable, Maxwell's model can be simplified to

$$P_{eff} = P_C\left(\frac{1 - \phi_D}{1 + \phi_D/2}\right) \tag{4.4}$$

According to eqn 4.4, the permeability of the filled polymer is always less than that of the pure polymer and decreases with increasing filler particle concentration. This reduction in permeability is related to an increase in the tortuosity of the diffusion path, as well as a decrease in penetrant solubility caused by the replacement of polymer, which can sorb penetrant, with non-sorbing filler particles. Equation 4.4 has been used successfully to describe the effect of non-porous filler content on permeability in a variety of filled polymers.[15,22]

4.4 Nanocomposite membrane research

The tremendous current interest in nanotechnology sometimes masks the fact that nanoscale materials have been studied and used for quite some time. This is certainly true for polymer/inorganic filler nanocomposites. Since nearly the beginnings of polymer science, and long before polymers were

utilized commercially for membrane applications, nanoscale filler particles have been incorporated into polymers to modify mechanical and flow properties, thermal and electrical conductivity, physical appearance, and for cost reduction.[23] Some of the more common fillers that have been used in polymers include nanoscale carbon blacks, metal oxides, and silicas, as well as micron-sized mica and aluminum powders, calcium carbonate and later, molecular sieves and zeolites.[23] These materials exhibit a wide range of sizes, shapes, surface area, and polarity, and can interact with polymers and penetrants to affect small molecule transport properties in a variety of ways. Table 4.1 lists some of these particles and their relevant properties. In this chapter, we have included zeolites as nanoscale materials even though the particle size of commercial zeolites is in the micrometer range. This classification was made because the internal pore dimensions of zeolites, which are utilized for molecular separations, are well within the nanoscale domain, typically ranging between 3–15 Å.

4.4.1 Early studies

By the 1950s, several pioneers of gas transport in polymers began systematic studies of the influence of small particle fillers on transport properties. These studies necessarily focused on natural and synthetic elastomers that were available at the time. Of particular note is the important work of van Amerongen[24,25] with natural rubber and Barrer and coworkers with silicone rubber[26] and natural rubber.[22] van Amerongen reported permeability coefficients for H_2, O_2, N_2, and CO_2 in natural rubber containing 19 different types of fillers, including most of the varieties mentioned above. In each case, compared to natural rubber without filler, permeability coefficients were reduced after

Table 4.1 Filler types and properties

Filler type	Examples	Surface chemistry	Primary particle size, nm	Common uses
Mineral oxides	Al_2O_3, TiO_2, ZnO	Hydrophilic	1–100	Catalysis, tailoring mechanical properties
Carbon blacks	Spheron 9 Spheron 4	Hydrophilic or hydrophobic (treated)	10–100	Modifying mechanical properties, toughness, coloring
Silicas	Hi-Sil Durosil Aerosil	Hydrophilic or hydrophobic (treated)	7–50	Rheology control

incorporation of filler into the polymer matrix (see Table 4.2). Furthermore, as filler content increased, gas permeability decreased monotonically. van Amerongen attributed this behavior to impermeable filler particles functioning as geometric obstructions to gas transport.[24,25] As filler content increases, the path a diffusing gas molecule follows to traverse a polymer film becomes more tortuous. Consistent with this viewpoint, in every system examined, van Amerongen measured diffusion coefficients in filled polymers to be lower than those in the unfilled polymer.[25] van Amerongen also observed that at a particular filler content, large lamellar fillers with high aspect ratios, such as mica, were more effective at blocking penetrant transport (i.e., reducing permeability and diffusion coefficients) than isolated spherical filler particles.[25]

The filler effects reported by van Amerongen were typical of the behavior observed in rubbery polymers by researchers in this time period.[22,26–28] A notable exception was the work of Melikhova *et al.*,[29] who reported that at filler loadings greater than 50% by volume, permeability in butadiene-based rubbers increased sharply. This result was attributed to discontinuities or small microscopic gaps in the polymer matrix created by the close proximity of filler particles that were not completely wetted by the polymer.[29] These

Table 4.2 The effect of different filler types on gas permeability in natural rubber[25]

	Filler	Volume fraction	Permeability ratio – filled/unfilled rubber[a]		Selectivity O_2/N_2[b]
			O_2	N_2	
Mineral fillers	Whiting	0.18	0.76	0.75	2.8
	Iron oxide	0.18	0.73	0.73	2.9
	Aluminum oxide	0.17	0.71	0.80	2.8
	Barium sulfate	0.18	0.77	0.77	2.5
Silicas	Hi-Sil	0.22	0.62	0.59	2.8
	Cal-sil	0.17	0.78	0.8	3.0
	Durosil	0.18	0.65	0.69	2.7
	Aerosil	0.18	0.65	0.66	2.6
	Ultrasilteg	0.18	0.61	0.56	2.8
Lamellar fillers	Aluminum	0.18	0.33	0.31	3.1
	Mica	0.18	0.20	0.20	3.0
	Graphite	0.19	0.48	0.48	2.8
	Bentonite	0.18	0.29	0.30	2.8
Carbon blacks	Thermax	0.22	0.74	0.72	2.7
	P 33	0.22	0.72	0.72	2.9
	Statex	0.22	0.67	0.66	2.8
	Vulcan 3	0.22	0.62	0.63	2.9
	Spheron 9	0.22	0.67	0.66	2.8
	Spheron 4	0.22	0.68	0.64	2.9

[a] Maxwell's model predicts the permeability ratio for a filler volume fraction of 0.17–0.22 to range from 0.76 to 0.70 independent of gas type (i.e., no change in selectivity with filler addition).
[b] O_2/N_2 selectivity of pure natural rubber = 2.8.

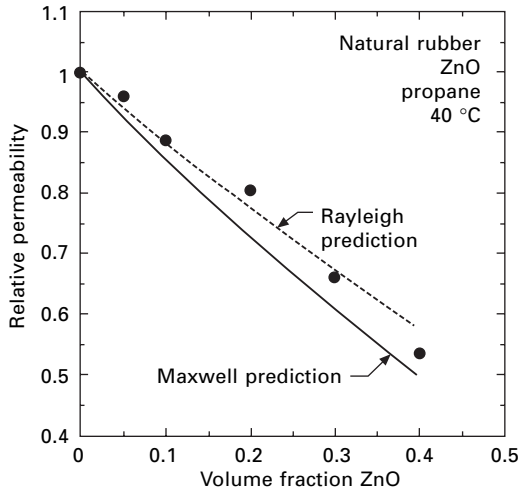
gaps increase the free volume of the polymer/filler system and present low resistance avenues for gas transport.

At nearly the same time, Barrer and coworkers[15,22,30] recognized the contribution molecular-scale gaps or voids introduced by fillers have on transport properties and incorporated this into a mathematical treatment of filled polymer systems. These researchers measured sorption isotherms for propane and benzene in natural rubber containing varying amounts of ZnO filler, as well as in the pure polymer and filler.[22] These data were analyzed using three different models corresponding to Henry's law sorption in one, two, or three distinct phases, respectively. Barrer *et al*. found that treating rubber and filler as separate phases that do not interact, where the filler has either no sorption capacity or its isolated particle capacity was insufficient to describe the experimental variation of solubility with filler content. Introduction of a third phase to include small gaps in the system model improved the agreement with experimental data, but was still recognized as an oversimplification of a complex heterogeneous system.[22]

Barrer and coworkers used these sorption models and tortuosity factors (the ratio of diffusion path length in the filled polymer to that in the unfilled material) to modify Fick's diffusion equation to apply to nanocomposite rubbery polymer/filler systems.[15,22] They found that Rayleigh's expression for the tortuosity factor, describing a cubic lattice of impermeable cylinders normal to the direction of flow, provided the best fit to their transport data in ZnO filled natural rubber. This result appears reasonable because ZnO particles are known to have a rod-like shape. The permeability data of Barrer *et al.* is reproduced in Fig. 4.1 along with the predicted permeability given by Maxwell's and Rayleigh's models.

Another important filler effect recognized by early researchers was penetrant immobilization and its influence on the diffusion time lag. Higuchi *et al*.[17] reported a dramatic increase in the diffusion time lag when a silica gel was incorporated into a white petroleum cream. Once steady-state was reached, flux through the filled cream was only slightly reduced compared to that through an unfilled cream. Most[31] observed similar behavior when he measured the time lag and permeability of ethyl *p*-aminobenzoate in silicone rubber membranes containing fumed silica fillers. For example, in a membrane with 25% silica by weight, apparent transient diffusion coefficients were reduced 15 times, whereas permeability decreased only 21% as compared to unfilled silicone rubber.[31] It was recognized that traditional time lag theory, as formulated by Daynes and Barrer,[32,33] could not be applied to systems with adsorptive fillers to evaluate diffusion coefficients.

Around this same time, Paul[34] modified time lag theory to account for immobilizing adsorption of a portion of the sorbed gas molecules in a polymer. This work was initiated primarily to explain results in polymers that existed in the glassy state at ambient conditions. Ever since the pioneering work of



4.1 Ratio of propane permeability in ZnO filled natural rubber to that in unfilled natural rubber (relative permeability or P_{eff}/P_c) as a function of filler content. The solid line represents the prediction of Maxwell's equation for a random dispersion of spherical particles and the dashed line represents Rayleigh's equation for a cubic array of cylindrical particles modified by Barrer for random dispersion.[22]

Meares[35] on polyvinyl acetate in the 1950s, a great deal of research had begun to focus on solution and diffusion in glassy polymers. In contrast to rubbers, these materials exhibit nonlinear sorption isotherms that are typically concave to the pressure axis.[36,37] These results led a number of researchers[36,38] to propose that sorption was occurring into two distinct environments in glassy polymers. It was suggested that molecular sorption in a densified glassy phase was equivalent to dissolution into a rubber or liquid, while another phase, consisting of non-equilibrium microvoids, followed Langmuir type adsorption.

Initially it was believed that molecules sorbed into Langmuir sites were immobilized and did not participate further in penetrant transport.[39] Consequently, a glassy polymer could be treated in much the same way as a rubber containing an adsorptive filler. Paul and Kemp,[40] recognizing this fact, measured time lags in a silicone rubber membrane containing adsorptive zeolite fillers and compared the results to their theoretical predictions.[34] The agreement between experiment and theory was excellent and suggested their modified time lag theory could be applied to glassy polymers if penetrant molecules were immobilized in Langmuir sites. Similar to the earlier work of Higuchi *et al*. and Most, Paul and Kemp observed large increases in time lag and minor decreases in steady-state permeation after the incorporation of filler into silicone rubber membranes. They explained the increase in time lag accompanying immobilizing adsorption as reflecting the time required for

excess penetrant to accumulate in the polymer system before equilibrium is reached.[40]

In summary, these early studies of gas transport in filled polymers necessarily dealt primarily with rubbers containing inert, nonporous particles. These fillers were added during the manufacturing process to modify properties that were important to the application for which the material was to be used. As membrane science was in its infancy during this period, filler effects on gas transport and separation properties were typically not the principal reason for adding small inorganic particles to polymers. Nevertheless, some of the work of this period, such as that of Barrer and van Amerongen and later Paul, helped elucidate the fundamental effects fillers have on transient and steady-state gas transport in polymers. This understanding was utilized, for example, in situations where it was desirable to improve the barrier properties of polymeric films. Applications such as food and beverage packaging, industrial hoses and o-rings, and protective paint coatings require polymers that have very low transmission rates for gases such as oxygen, carbon dioxide, and water vapor.[41] As the early work on filled polymers demonstrated, the presence of impermeable filler particles tends to increase the time to steady-state and decrease the steady-state flux, both desirable effects from a barrier standpoint. Later work[41–43] would expand upon van Amerongen's early findings and demonstrate that oriented, high-aspect-ratio particles yield composites that can be orders of magnitude less permeable than the base polymer.[41]

4.4.2 Modern nanocomposite membranes

After the first commercial demonstration of gas separation membranes in the 1970s,[44,45] polymer membrane research focused primarily on exploring different classes of glassy polymers in an attempt to identify materials with favorable permeability/selectivity characteristics. Initially the development of designer polymer materials significantly improved membrane performance. However, by the late 1980s enhancements in polymer permeability/selectivity properties slowed to a crawl. At this time, attention returned to filled polymer systems as a possible means of improving membrane performance. The first studies of modern nanocomposite membranes focused on incorporating highly selective zeolites into polymers in the hope that the excellent separation properties of the zeolites would be integrated into a mechanically robust polymer membrane. The promise of these so-called 'mixed-matrix membranes' stems from the fact that some zeolites have very high selectivities for industrially important gas pairs. For example, zeolite 4A has an O_2/N_2 selectivity of 37 as compared to 7–9 for the best polymeric membranes.[11]

Mixed-matrix approach

One of the first tests of mixed-matrix membranes for gas separations was conducted by Jia *et al.*[10] who incorporated silicalite, a hydrophobic crystalline molecular sieve, into silicone rubber. It was found that increasing filler loading augmented the permeability of small gas molecules (H_2, He, CO_2, O_2), and hindered the transport of larger molecules (N_2, CH_4, *n*-butane). As a result, selectivity for the gas pair O_2/N_2 increased modestly with the addition of silicalite from 2.1 in the pure polymer to 2.9 in a composite containing 70 wt% silicalite. Similar small improvements in gas separation properties were reported around this time by several researchers working mostly with rubbery polymers.[9,46] While these studies demonstrated that the approach of adding nano-porous inorganic particles to polymer films could improve selectivities, the composites formed were not competitive with commercial gas separation polymer membranes. This was because the rubbery polymers to which the zeolites were added have low intrinsic selectivities for industrially important gas pairs such as O_2/N_2 or CO_2/CH_4. For this reason, subsequent studies focused on adding zeolites to commercial gas separation polymers, which are almost exclusively rigid, glassy materials.

As a number of researchers soon reported,[11,47–50] the addition of zeolites to glassy polymers to improve separation performance proved to be quite difficult. Suer *et al.*,[48] for example, characterized gas permeation in polyethersulfone containing varying amounts of zeolites 13X and 4A. They found that for each gas considered, as filler content increased, permeability first decreased and than increased at high zeolite loadings. However, this high permeability was accompanied by low selectivities. SEM microscopy suggested the zeolites were poorly wetted by the polymer matrix, and that at high loadings, void spaces may form networks that increase gas flux and reduce selectivity. Similarly, Mahajan *et al.*[11] examined the O_2/N_2 separation properties of polyimide membranes containing zeolite 4A, carbon molecular sieves and impermeable carbon blacks. These systems failed to produce the separation performance predicted by Maxwell's equation using the transport properties of the pure polymer and filler. Consistent with the results of Suer *et al.*, it was concluded that non-selective voids formed by dewetting of polymer chains from the molecular sieve surface compromised membrane performance.

Recently, efforts have been made by several groups to address this polymer/inorganic particle wetting problem. Duval *et al.*,[50] for instance, employed silane coupling agents to tether the zeolite particles to the polymer phase. While better polymer-zeolite contact was observed in SEM images, the selectivities obtained with the silylated-zeolite/polymer composites were still poor. Duval and coworkers also tried membrane fabrication above the polymer glass transition temperature and annealing techniques with little success. Other researchers,[51–55] using similar methods to enhance the compatibility

between the polymer phase and zeolite particles, have reported improvements in mixed-matrix performance. For instance, Mahajan and Koros demonstrated an O_2/N_2 selectivity of 12.9 for a polyetherimide (Ultem®, General Electric) containing 35 vol% of zeolite 4A, as compared to 7.8 for the pure polymer.[54] Tables 4.3 and 4.4 summarize recent permeability and selectivity results reported by various researchers for mixed-matrix membranes. Through proper choice of polymer, zeolite, coupling agent and preparation/post-treatment conditions, some of these composites are starting to approach useful gas separation properties at lab scale.

Sol-gel approach

A slightly different approach to improving polymer gas separation properties by incorporation of nanosized inorganic components into the polymer matrix has been pursued by several researchers in the past ten years. This work utilizes the sol-gel technique[56] to produce hybrid inorganic–organic materials where the inorganic moiety can either be inserted into the polymer backbone to yield an entirely new network material or can exist as discrete nanodomains that modify chain packing and segmental mobility of the polymer. The latter approach was employed by Hu *et al.*[57] who reported permeability data in poly(amide-imide)-TiO_2 composite membranes. In this system, TiO_2 nanodomains interact with amide linkages in the polymer backbone to produce a stiffer material as compared to the unfilled polymer. At a TiO_2 content of 7.2% by weight, permeability coefficients for all gases studied decreased whereas size-selectivities increased. This result is consistent with the typical permeability/selectivity tradeoff relationship exhibited by size-selective materials.[7]

Smaihi *et al.*[58] examined the gas separation properties of a hybrid network material formed from tetramethoxysilane and an organoalkoxide. This membrane exhibited pore-flow-like, Knudsen selectivities at low organoalkoxide concentrations and higher selectivities as the molar ratio of organoalkoxide was increased above 0.5. Joly *et al.*[59] reported gas permeabilities in a polyimide-silica sol-gel composite membrane. These authors observed a small increase in gas permeability for a film containing 32 wt% silica as compared to a pure polyimide film. This result was attributed primarily to increased gas solubility in the composite film resulting from a lower imidization ratio and possible morphological changes induced by the silica. Kusakabe *et al.*[60] reported gas transport properties for a similar polyimide-silica sol-gel membrane. These authors observed that both CO_2 solubility and permeability increased as silica content was increased from 0 to 100 wt%. This result was attributed to the high micropore volume of the silica phase. Marand and coworkers[61,62] have also investigated polyimide/silica hybrid membranes formed via the sol-gel method. They have found that membrane performance

Table 4.3 Mixed-matrix membrane permeability summary

Polymer	Filler	Loading	Permeability (10^{-10} cm^3(STP) · cm/cm^2 · s · cmHg)					Reference
			He	H_2	CO_2	O_2	N_2	
Matrimid 5218	–	–	21	–	8.3	1.5	0.22	Yong[69]
Matrimid 5218	Zeolite 4A	1:0.43	20	–	9.4	1.9	0.46	
Matrimid 5218	Zeolite 13X	1:0.43	54	–	33.4	6.6	1.35	
PVAC	–	–	15	–	3.1	0.5	0.09	Hasse[70]
PVAC	Zeolite 4A	15%	11		2.4	0.6	0.08	
PVAC	Zeolite H-ZK-5	15%	24.6	–	4.9	0.8	0.12	
PVAC	Zeolite SAPO-34	15%	17.8	–	4.4	0.7	0.10	
PVAC	Zeolite SAPO-44	15%	–	–	4.9	0.8	0.10	
PVAC	KFI	20%	–	–	4.9	0.6	0.09	
PVAC	–	–	–	–	–	0.5	0.09	Mahajan[53]
PVAC	Zeolite 4A	15%	–	–	–	0.45	0.06	
PVAC	Zeolite 4A	25%	–	–	–	0.4	0.05	
PVAC	Zeolite 4A	40%	–	–	–	0.3	0.03	
Matrimid 5218	–	–	–	–	10.0	2.1	0.32	Vu[71]
Matrimid 5218	CMS 800-2	17 vol%	–	–	10.3	2.1	0.29	
Matrimid 5218	CMS 800-2	19 vol%	–	–	10.6	2.4	0.35	
Matrimid 5218	CMS 800-2	33 vol%	–	–	11.5	2.7	0.38	
Matrimid 5218	CMS 800-2	36 vol%	–	–	12.6?	3.0	0.38	
Ultem 1000	–	–	–	–	1.45	0.4	0.05	Vu[71]
Ultem 1000	CMS 800-2	16 vol%	–	–	2.5	0.6	0.07	
Ultem 1000	CMS 800-2	20 vol%	–	–	2.9	0.7	0.09	
Ultem 1000	CMS 800-2	35 vol%	–	–	4.5	1.1	0.14	
PMDA-PPD	–	–	–	17.7	5.2	–	0.32	Zhong[72]
PMDA-PPD	TEOS	51 wt%	–	37	12.9	–	0.56	
PMDA-PPD	SiO_2	51 wt%	–	32	10.6	–	0.53	
Polyamide-imide 6FPA/I	–	–	–	67	52.7	11.2	2.30	Hu[57]
Polyamide-imide 6FPA/I	TiO_2	7.3	–	66	44.7	9.2	1.80	

Table 4.4 Mixed-matrix membrane selectivity summary

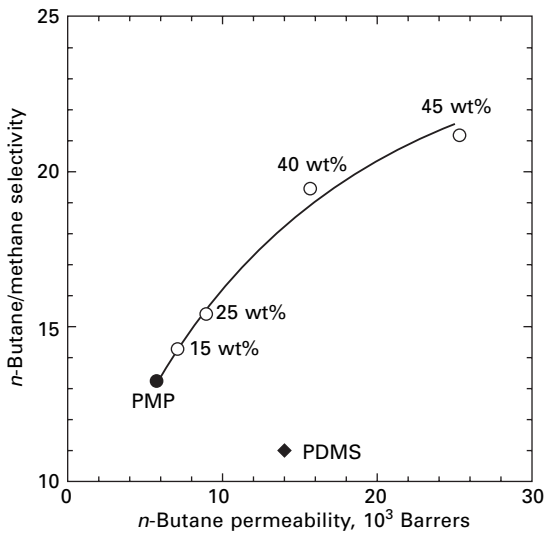
Polymer	Filler	Loading	Selectivity (gas/N_2)				Reference
			He	H_2	CO_2	O_2	
Matrimid 5218	–	–	95	–	38	6.8	Yong[69]
Matrimid 5218	Zeolite 4A	1:0.43	43	–	20	4.1	
Matrimid 5218	Zeolite 13X	1:0.43	40	–	25	4.9	
PVAC	–	–	167	–	34	5.6	Hasse[73]
PVAC	Zeolite 4A	15%	138	–	30	7.5	
PVAC	Zeolite H-ZK-5	15%	205	–	41	6.7	
PVAC	Zeolite SAPO-34	15%	178	–	44	7.0	
PVAC	Zeolite SAPO-44	15%	–	–	49	8.0	
PVAC	KFI	20%	–	–	54	6.7	
PVAC	–	–	–	–	–	5.6	Mahajan[53]
PVAC	Zeolite 4A	15%	–	–	–	7.5	
PVAC	Zeolite 4A	25%	–	–	–	8.0	
PVAC	Zeolite 4A	40%	–	–	–	10	
Matrimid 5218	–	–	–	–	31	6.6	Vu[71]
Matrimid 5218	CMS 800-2	17 vol%	–	–	36	7.2	
Matrimid 5218	CMS 800-2	19 vol%	–	–	30	7.0	
Matrimid 5218	CMS 800-2	33 vol%	–	–	30	7.1	
Matrimid 5218	CMS 800-2	36 vol%	–	–	33	7.9	
Ultem 1000	–	–	–	–	29	7.3	Vu[71]
Ultem 1000	CMS 800-2	16 vol%	–	–	32	7.9	
Ultem 1000	CMS 800-2	20 vol%	–	–	32	7.9	
Ultem 1000	CMS 800-2	35 vol%	–	–	32	8.0	
PMDA-PPD	–	–	–	55	16	–	Zhong[72]
PMDA-PPD	TEOS	51 wt%	–	66	23	–	
PMDA-PPD	SiO_2	51 wt%	–	60	20	–	
Polyamide-imide 6FPA/I	–	–	–	29	23	4.9	Hu[57]
Polyamide-imide 6FPA/I	TiO_2	7.3	–	37	34	5.1	

is strongly dependent on the polymer/silica interface and its influence on segmental mobility and free volume.

These sol-gel nanocomposite membranes exhibit tremendous promise for permitting tailored modification of polymer matrix properties or the creation of entirely new network materials with undiscovered properties. However, much like the zeolite mixed-matrix membranes, considerable work is still required before a viable commercial membrane material is produced.

Polymer/nanoparticle blend approach

Recently, an alternative means of using inorganic nanoparticles to modify transport and separation properties of polymer membranes has been identified.[12,13,63–65] In this approach, non-porous nanoparticles are physically blended into certain rigid, glassy polymers to increase matrix free volume by essentially functioning as 'nanospacers' to disrupt efficient polymer chain packing. In this way, both membrane permeability and vapor/permanent-gas selectivity can be increased. This result is completely contrary to previous studies of polymers containing non-porous fillers, which typically exhibit reduced permeability as predicted by the Maxwell model. Figure 4.2 illustrates the effect of adding 13-nm diameter fumed silica particles on the permeability and selectivity of an amorphous polyacetylene, poly(4-methyl-2-pentyne) [PMP].[12,13] The selectivity of the PMP membrane for *n*-butane over methane nearly doubles as fumed silica content is increased to 45 wt%. At the same

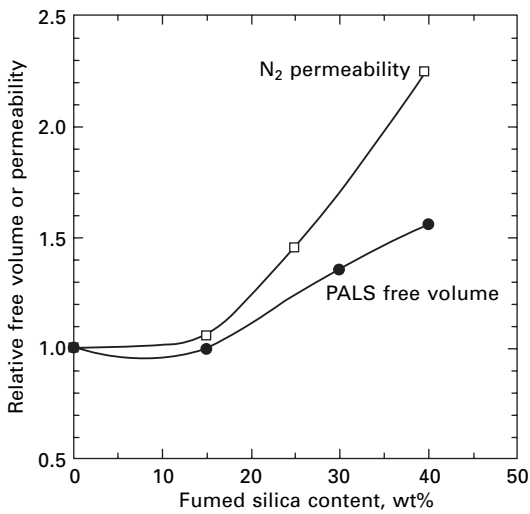


4.2 Mixed-gas *n*-butane/methane selectivity as a function of *n*-butane permeability for PMP containing various amounts of fumed silica. Data for polydimethylsiloxane [PDMS] is included for comparison.[13]

time, *n*-butane permeability increases by a factor of four as the filler content increases from 0 to 45 wt% fumed silica. As demonstrated in Fig. 4.2, the permeability/selectivity combinations achieved in the PMP/silica nanocomposites far exceed that of the commercial vapor-separation membrane material, polydimethylsiloxane [PDMS].

Because fumed silica is non-porous, it does not permeate gas molecules as a zeolite could. Rather, it appears that for stiff-chain glassy polymers such as PMP, fumed silica particles disrupt polymer chain packing, leading to an increase in system free volume. This enhanced free volume in the filled polymers has been confirmed by vapor sorption,[63] positron annihilation lifetime spectroscopy [PALS][13] and ^{128}Xe NMR.[66] By increasing free volume within the polymer membrane, fumed silica addition reduces transport resistance for permeating species. As a result, permeant diffusion coefficients, and hence, permeability coefficients increase with increasing fumed silica content (silica addition has little effect on solubility coefficients). Figure 4.3 demonstrates a rather strong qualitative correlation between free volume, as discerned by PALS measurements, and permeability in PMP containing fumed silica. As silica content in PMP increases, membrane free volume and permeability increase in a nearly identical manner.

In addition to increasing permeability, incorporation of fumed silica into PMP and other high-free-volume glasses also increases vapor selectivity (i.e., the selectivity of the membranes for large, condensable molecules, such as *n*-butane, over permanent gases, such as methane). This simultaneous increase in permeability and vapor selectivity is consistent with a reduction



4.3 Relative nitrogen permeability or PALS free volume as a function of fumed silica content in PMP.[63]

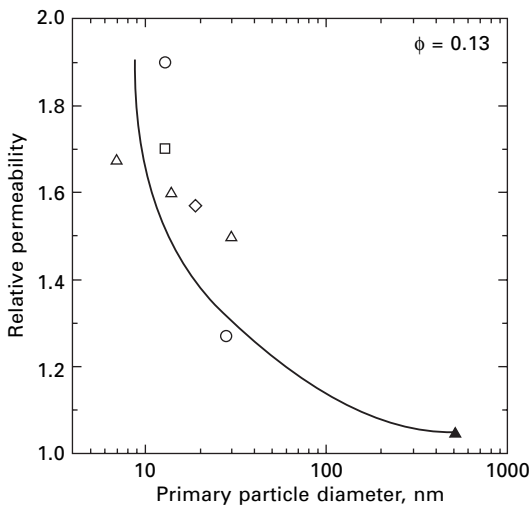
in the size-sieving ability of the polymer; that is, the difference in diffusion coefficients between large and small molecules decreases as fumed silica content increases. In general, size selectivity of polymers decreases as permeability increases, in accord with the well-known permeability/selectivity tradeoff relationship.[7] This behavior may be rationalized from free volume considerations as follows: as the free volume in a polymer matrix available for penetrant transport increases, sorbed molecules have a greater likelihood of being in the vicinity of an accessible free volume element and making a diffusive jump. Consequently, penetrant diffusion coefficients increase with increasing polymer free volume. Larger penetrants are more sensitive to free volume increases than smaller penetrants and, as a result, experience a greater percentage increase in diffusion coefficients.[67] Thus, as the free volume of a polymer is increased, it loses its ability to discriminate between molecules based on size differences. Therefore, the ratio of diffusion coefficients in eqn 4.2 approaches one, which is a desirable result for enhanced vapor/gas selectivity. The permeation data in Fig. 4.2 are consistent with the PALS findings that incorporation of fumed silica into PMP increases system free volume, thereby increasing diffusion coefficients and reducing diffusivity selectivity, which contributes to higher permeability and vapor selectivity, respectively.

These free volume considerations nicely explain the observed impact of fumed silica addition on PMP permeability and selectivity. This rationale is quite different from that used to describe transport in conventional filled polymers, where particles act as obstructions to molecular transport and reduce permeability as predicted by eqn 4.4. In these conventional systems, relatively flexible polymer chains wet or coat filler particles well enough to prevent a significant impact on free volume, and consequently geometric obstruction is the dominant effect on transport. In contrast, PMP and related polymers possess double bonds in the main chain and bulky substituents that render them extremely rigid and likely prevent them from wrapping around and efficiently coating silica particles. As a result, additional free volume is created and this effect overwhelms any obstruction to transport caused by the presence of impermeable filler particles.

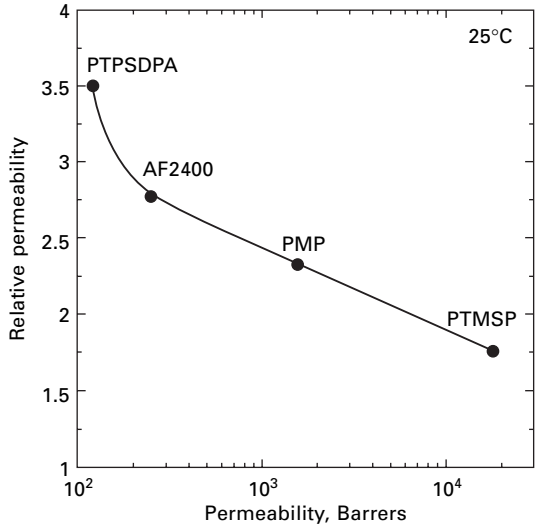
Previously, there have been a few studies where zeolite particles were added to glassy polymers and permeabilities increased. As mentioned, in these cases the unexpectedly high permeation rates were attributed to the creation of micron-scale defects at the polymer/filler interface that allow fast, but relatively non-selective, Knudsen diffusion.[11] Defects of this type are not evident in the PMP/fumed silica system. The relatively large pathways associated with Knudsen diffusion favor methane transport and, if present in PMP/silica nanocomposites, would reduce *n*-butane/methane selectivity.[11] In fact, the opposite is observed, indicating that nanoscale fumed silica particles subtly increase free volume in PMP without introducing defects large enough to permit free-phase gas flow.

One of the most intriguing aspects of molecular transport in PMP/filler nanocomposites is the effect of filler particle size, illustrated in Fig. 4.4. At a fixed volume fraction of filler, as particle size decreases, relative permeability, or the ratio of filled polymer to pure polymer permeability, increases. This behavior is unexpected based on traditional theories of transport in filled polymers that predict no effect of particle size on permeability for spherical particles. Microscopic examination of the nanocomposites reveals that the individual fumed silica particles aggregate into clusters averaging roughly 100 nm in diameter.[63] This clustering initially makes it difficult to understand the relationship between primary particle diameter and permeability enhancement. However, regardless of the extent of filler clustering, the number of filler particles inserted into a polymer matrix at a fixed volume fraction increases as particle size decreases. For example, as primary particle size decreases from 50 to 10 nm, the number of filler particles per cubic centimeter of composite increases by approximately two orders of magnitude. Correspondingly, the particle surface area also increases substantially, so that there is much more potential polymer-filler interfacial area in PMP containing 10-nm particles than in PMP containing the same volume fraction of 50-nm particles. Apparently, the more interfacial area (or volume) present, the higher is the permeability. This suggests that a relatively low-density interfacial region is the locus of rapid mass transport in these nanocomposite membranes.

In addition to the effect of filler particle size, the choice of polymer matrix also impacts achievable permeability enhancement. Figure 4.5 presents relative permeability in four different high-free-volume polymers containing



4.4 The effect of primary particle size on the relative permeability of PMP containing 13 vol% filler.[13]



4.5 Ratio of permeability in the filled polymer to that in the unfilled polymer (i.e., relative permeability) as a function of the unfilled polymer permeability. These data are for methane permeation at 25 °C and a Δp of 3.4 atm.[64]

the same volume fraction of 13-nm silica particles as a function of the unfilled polymer permeability. The data in this figure indicate that the increase in permeability obtained by adding a given amount of fumed silica decreases as the permeability of the polymer being filled increases. This behavior is probably related to the relative free volume of the polymers being filled and their ability to accommodate fumed silica particles within the polymer matrix. For example, based on transport parameters and other estimates of free volume such as PALS, poly(1-trimethylsilyl-propyne) [PTMSP] possesses the largest free volume elements among these polymers.[67,68] Consequently, PTMSP accommodates fumed silica particles with the least disruption of polymer chain packing and experiences therefore the smallest permeability enhancement. Conversely, the polyacetylene, poly(1-phenyl-2-(*p*-(triisopropylsilyl)phenyl)acetylene) [PTPSDPA] has the lowest permeability and free volume among these polymers. It experiences the greatest chain packing disruption at a given volume fraction of fumed silica, and it shows the largest permeability enhancement.

4.5 Conclusions and future trends

Nanoscale inorganic particles have been incorporated into polymers to modify properties for more than 50 years. Early studies of gas transport in these composite systems observed and described in simple mathematical terms

traditional filler effects, namely, the enhancement of polymer barrier properties through increased tortuosity and amplified steady-state time lag.

In the last 15 years, a conscious effort has been made to modify polymer separation properties through the addition of various nano-sized fillers. Conceptually this work can be divided into three separate avenues of exploration. Firstly, nano-porous molecular sieves such as zeolites have been incorporated into polymer matrices in the hope of imparting the very high selectivity of a sieve particle to the resulting mixed-matrix membrane. Secondly, hybrid organic-inorganic membranes have been created by the sol-gel process with the goal of either permitting tailored modification of polymer matrix properties or creating entirely new network materials with undiscovered properties. Both approaches, although showing glimpses of promise, have enjoyed little success to date in producing commercially viable membrane materials. Thirdly, nanoscale, non-porous inorganic fillers, such as fumed silica and carbon black, have been added to rigid, high-free-volume glassy polymers to disrupt chain packing and increase free volume by serving as nanospacers. These nanocomposites have been shown to exhibit significantly higher permeability and vapor/gas selectivity than the unfilled polymer. For practical commercial applications, however, this approach also has limitations because current filled polymer materials that exhibit the most promising transport properties show poor chemical stability.

Future work in nanocomposite membranes will no doubt continue in each of the areas outlined above with the distinctions between them becoming less obvious. Mixed-matrix membranes will require smaller sieving particles to improve mechanical properties and allow the formation of high-performance, ultra-thin-film membranes (1 μm or less). Successful mixed-matrix materials may also use sol-gel chemistry to improve organic/inorganic compatibility. In addition to zeolites, other structured nanoscale materials, such as carbon nanotubes, could find uses in polymer membranes as highly-selective molecular gates, although the difficulties of reproducibly fabricating industrial membranes with such materials appear daunting.

With the extension of industrial membrane use to more high-temperature and harsh chemical environments, sol-gel type organic–inorganic hybrids offer promise, not so much for their improved separation performance, but for their potential stability under these challenging conditions. Ideally, they can bridge the gap between inherently stable ceramics and cheap, processable polymers. However, work must be done to improve the simplicity and reliability of fabricating these materials on a large scale.

Vapor-selective membranes are seeing increasing use in petrochemical refinery applications. High-free-volume polymer/nanoparticle blends such as PMP/fumed silica offer a vapor-selective material with attractive, tunable permeability-selectivity characteristics. However, the best performing polymers tested to date are not stable enough to be used commercially. Ongoing work

continues to examine using nanoparticles to control polymer chain packing and thereby permeability and selectivity with the goal of identifying stable nanocomposite membranes with useful transport properties.

4.6 References

1. Coronado, E., Galan-Mascaros, J. R., Gomez-Garcia, C. J., Laukhin, V. *Nature* **2000**, *408*, 447–449.
2. Croce, F., Appetecchi, G. B., Persi, L., Scrosati, B. *Nature* **1998**, *394*, 456–458.
3. Pinnavaia, T. J. *Science* **1983**, *220*, 365.
4. Schmid, G., ed. *Clusters and Colloids*; John Wiley & Sons: Weinheim, Germany, 1994.
5. Sidorov, S. N., Volkov, I. V., Davankov, V. A., Tsyurupa, M. P., Valetsky, P. M., Bronstein, L. M., Karlinsey, R., Zwanziger, J. W., Matveeva, V. G., Sulman, E. M., Lakina, N. V., Wilder, E. A., Spontak, R. J. *Journal of the American Chemical Society* **2001**, *123*, 10502–10510.
6. Shiflett, M. B., Foley, H. C. *Science* **1999**, *285*, 1902–1905.
7. Robeson, L. M. *Journal of Membrane Science* **1991**, *62*, 165–185.
8. Freeman, B. D. *Macromolecules* **1999**, *32*, 375–380.
9. Kulprathipanja, S., Neuzil, R. W., Li, N., 'Separation of fluids by means of mixed matrix membranes', US Patent No. 4,740,219, 1988.
10. Jia, M., Peinemann, K. V., Behling, R. D. *Journal of Membrane Science* **1991**, *57*, 289–296.
11. Mahajan, R., Zimmerman, C. M., Koros, W. J. In *Polymer Membranes for Gas and Vapor Separation: Chemistry and Materials Science*; Pinnau, I., ed., American Chemical Society: Washington, D.C., 1999; pp 277–286.
12. Pinnau, I., He, Z. 'Filled Superglassy Membrane', U.S. Patent No. 6,316,684, 2001.
13. Merkel, T. C., Freeman, B. D., Spontak, R. J., He, Z., Pinnau, I., Meakin, P., Hill, A. J. *Science* **2002**, *296*, 519–522.
14. Graham, T. *Philos. Mag.* **1866**, *32*, 401.
15. Barrer, R. M. In *Diffusion in Polymers*; Park, G. S., ed., Academic Press: London, 1968; pp 165–217.
16. Bruggeman, D. A. G. *Ann. Phys. (Leipzig)* **1935**, *24 (Ser. 5)*, 636–664.
17. Higuchi, W. I., Higuchi, T. *Journal of the American Pharmaceutical Association: Sci. Ed.* **1960**, *49*, 598–606.
18. Petropoulos, J. H. *Journal of Polymer Science: Polymer Physics Edition* **1985**, *23*, 1309–1324.
19. Maxwell, J. C. *A Treatise on Electricity and Magnetism*; Oxford University Press: London, 1873; Vol. 1.
20. Zimmerman, C. M., Singh, A., Koros, W. J. *Journal of Membrane Science* **1997**, *137*, 145–154.
21. Mahajan, R., Koros, W. J. *Polymer Engineering and Science* **2002**, *42*, 1420–1431.
22. Barrer, R. M., Barrie, J. A., Rogers, M. G. *Journal of Polymer Science: Part A* **1963**, *1*, 2565–2586.
23. Rothon, R. N. *Advances in Polymer Science* **1999**, *139*, 67–104.
24. van Amerongen, G. J. *Rubber Chem. Tech.* **1955**, *28*, 821.
25. van Amerongen, G. J. *Rubber Chemistry and Technology* **1964**, *37*, 1065–1152.
26. Barrer, R. M., Barrie, J. A., Raman, N. K. *Polymer* **1962**, *3*, 605–614.

27. Carpenter, A. S., Twiss, D. F. *Rubber Chem. Tech.* **1940**, *13*, 326.
28. Smith, C. W. *India Rubber World* **1953**, *129*, 55.
29. Melikhova, N. A., Reitlinger, S. A., Kuzina, E. N. *Soviet Rubber Tech.* **1959**, *18*, 34.
30. Barrer, R. M., Chio, H. T. *Journal of Polymer Science: Part C* **1965**, *10*, 111.
31. Most, C. F. *Journal of Applied Polymer Science* **1970**, *14*, 1019–1024.
32. Daynes, H. A. *Proc. R. Soc., London, Ser. A* **1920**, *97*, 286.
33. Barrer, R. M. *Trans. Faraday Soc.* **1939**, *35*, 628.
34. Paul, D. R. *Journal of Polymer Science: Part A-2* **1969**, *7*, 1811–1818.
35. Meares, P. *Journal of the American Chemical Society* **1954**, *76*, 3415.
36. Michaels, A. S., Vieth, W. R., Barrie, J. A. *Journal of Applied Physics* **1963**, *34*, 1–12.
37. Vieth, W. R., Tam, P. M., Michaels, A. S. *Journal of Colloid and Interface Science* **1966**, *22*, 360.
38. Vieth, W. R., Alcalay, H. H., Frabetti, A. J. *Journal of Applied Polymer Science* **1964**, *8*, 2125.
39. Vieth, W. R., Sladek, K. J. *Journal of Colloid Science* **1965**, *20*, 1014.
40. Paul, D. R., Kemp, D. R. *Journal of Polymer Science* **1973**, *Symposium No. 41*, 79–93.
41. Cussler, E. L., Hughes, S. E., Ward, W. J., Aris, R. *Journal of Membrane Science* **1988**, *38*, 161–174.
42. Okuda, S. In *Organic Coatings: Science and Technology*; Patsis, A. V., ed., Academic Press: New York, 1982; Vol. 7, pp 285–298.
43. Murthy, N. S., Kotliar, A. M., Sibilia, J. P., Sacks, W. *Journal of Applied Polymer Science* **1986**, *31*, 2569–2582.
44. Henis, J. M. S., Tripodi, M. K. *Science* **1983**, *220*, 11–17.
45. Parkinson, G., Ushio, S., Lewald, R. *Chemical Engineering* **1984**, *April 16*, 14–19.
46. Duval, J. M., Folkers, B., Mulder, M. H. V., Desgrandchamps, G., Smolders, C. A. *Journal of Membrane Science* **1993**, *80*, 189–198.
47. Gur, T. M. *Journal of Membrane Science* **1994**, *93*, 283–289.
48. Suer, M. G., Bac, N., Yilmaz, L. *Journal of Membrane Science* **1994**, *91*, 77–86.
49. Vankelecom, I. F. J., Merckx, E., Luts, M., Uytterhoeven, J. B. *Journal of Physical Chemistry* **1995**, *99*, 13187–13192.
50. Duval, J. M., Kemperman, A. J. B., Folkers, B., Mulder, M. H. V., Desgrandchamps, G., Smolders, C. A. *Journal of Applied Polymer Science* **1994**, *54*, 409–418.
51. Vankelecom, I. F. J., Broeck, S. V. d., Merckx, E., Geerts, H., Grobet, P., Uytterhoeven, J. B. *Journal of Physical Chemistry* **1996**, *100*, 3753–3758.
52. Guiver, M. D., Thi, H. N. L., Robertson, G. P. 'Composite Gas Separation Membranes', U.S. Patent No. 6,605,140, 2003.
53. Mahajan, R., Koros, W. J. *Industrial & Engineering Chemistry Research* **2000**, *39*, 2692–2696.
54. Mahajan, R., Koros, W. J. *Polymer Engineering and Science* **2002**, *42*, 1432–1441.
55. Mahajan, R., Burns, R., Schaeffer, M., Koros, W. J. *Journal of Applied Polymer Science* **2002**, *86*, 881–890.
56. Brinker, C. J., Sherrer, G. W. *Sol-gel Science, the Physics and Chemistry of Sol-gel Processing*; Academic Press: San Diego, 1989.
57. Hu, Q., Marand, E., Dhingra, S., Fritsch, D., Wen, J., Wilkes, G. *Journal of Membrane Science* **1997**, *135*, 65–79.
58. Smaihi, M., Jermoumi, T., Marignan, J., Noble, R. D. *Journal of Membrane Science* **1996**, *116*, 211–220.

59. Joly, C., Goizet, S., Schrotter, J. C., Sanchez, J., Escoubes, M. *Journal of Membrane Science* **1997**, *130*, 63–74.
60. Kusakabe, K., Ichiki, K., Hayashi, J., Maeda, H., Morooka, S. *Journal of Membrane Science* **1996**, *115*, 65–75.
61. Cornelius, C. J., Marand, E. *Journal of Membrane Science* **2002**, *202*, 97–118.
62. Hibshman, C., Cornelius, C. J., Marand, E. *Journal of Membrane Science* **2003**, *211*, 25–40.
63. Merkel, T. C., Freeman, B. D., Spontak, R. J., He, Z., Pinnau, I., Meakin, P., Hill, A. J. *Chemistry of Materials* **2003**, *15*, 109–123.
64. Merkel, T. C., He, Z., Pinnau, I., Freeman, B. D., Meakin, P., Hill, A. J. *Macromolecules* **2003**, *36*, 6844–6855.
65. Merkel, T. C., He, Z., Pinnau, I., Freeman, B. D., Meakin, P., Hill, A. J. *Macromolecules* **2003**, *36*, 8406–8414.
66. Merkel, T. C., Toy, L. G., Andrady, A. L., Gracz, H., Stejskal, E. O. *Macromolecules* **2003**, *36*, 353–358.
67. Freeman, B. D., Hill, A. J. In *Structure and Properties of Glassy Polymers*; Hill, A. J., ed., American Chemical Society: Washington, D.C., 1998; pp 306–325.
68. Morisato, A., He, Z., Pinnau, I. In *Polymer Membranes for Gas and Vapor Separation: Chemistry and Materials Science*; Pinnau, I., ed., American Chemical Society: Washington, D.C., 1999; pp 56–67.
69. Yong, H. H., Park, H. C., Kang, Y. S., Won, J., Kim, W. N. *Journal of Membrane Science* **2001**, *188*, 151–163.
70. Hasse, D. J., Kulkarni, S. S., Sudhir, S., Corbin, D. R., Patel, A. N., 'Mixed matrix membranes incorporating chabazite type molecular sieves', US Patent No. 6,626,980, 2003.
71. Vu, D. Q., Koros, W. J., Miller, S. J. *Journal of Membrane Science* **2003**, *211*, 311–334.
72. Zhong, S. H., Li, C. F., Xiao, X. F. *Journal of Membrane Science* **2002**, *199*, 53–58.
73. Hasse, D. J., Kulkarni, S. S., Sudhir, S., Corbin, D. R., Patel, A. N. 'Mixed matrix membranes incorporating chabazite type molecular sieves', US Patent No. 6,626,980, 2001.

5

Developing fast ion conductors from nanostructured polymers

M FORSYTH, J ADEBAHR, N BYRNE
and D MACFARLANE, Monash University, Australia

5.1 Introduction

The transport of ions in solid materials is of significant technological interest as it is a key requirement for the development of solid state devices such as high energy density lithium batteries, fuel cells, electromechanical actuators, photoelectrochemical solar cells, electrochromic displays and sensors. Some of the earliest solid state conductors were based on crystalline materials whose structure consisted of channels or layers which facilitated rapid ion transport between sites of similar energy [1]; AgI, b-alumina, nasicon and lisicon are all examples of such systems. Fast ion conduction in these systems is a reflection of their structure at the nano or even sub-nano level. Fast ion conductors based on polymeric materials [2–4] have also been actively pursued due to the mechanical and processing benefits which are associated with these compared to ceramic electrolytes. In these materials the ion transport appears to be dependent on the motion of the polymer itself, although this chapter will discuss how the degree of coupling of ionic motion to the host material can be significantly influenced by the control of the morphology and structure at the nanoscopic level. This control can come about via chemical, thermal or mechanical changes, or via the incorporation of nanosized particulates which appear to change the ionic structure and/or transport mechanism, as will be discussed later. Finally, fast ion conduction in organic crystalline materials and doped organic crystals has recently been discussed and it is suggested that the high conductivities are associated with nano-sized defects such as vacancies, dislocations or grain boundaries [5–22]. This chapter will focus on the ion transport behaviour in polymer-based electrolytes, and in particular how this can be enhanced by control of the structure and morphology at the nanometer level.

5.2 Polymer electrolytes

Ion transport in polymer electrolytes has been generally accepted to occur

via the amorphous regions of the polymeric material and as such is correlated with the local polymer dynamics [23, 24]. Recently, however, it has been shown that ionic conductivity can be significant through ordered regions where ions may hop, for example through a helical structure or through lamellar layers obtained via block copolymerisation. Bruce *et al.* [25, 26] have successfully determined crystal structures of polymer electrolytes based on PEO complexed with lithium salts such as $LiPF_6$, $LiAsF_6$ and $LiSbF_6$ from powders of these materials. The stoichiometry of these complexes has all been 6:1 (i.e., six ether oxygens to each lithium ion). In such a structure the PEO adopts a helical structure forming tunnels within which the Li^+ ions reside and are coordinated to the ether oxygens. The anions generally line up between these tunnels. The tunnels can either be formed by a single PEO helix or by two separate chains folding to form half cylinders and pairs of interlocking chains coordinating the lithium ions.

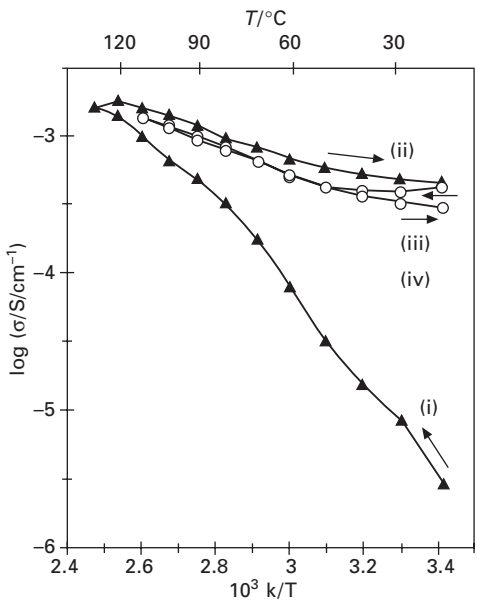
Conductivity was determined from a.c. impedance measurements in all three systems and was shown to be substantial even well below room temperature; importantly, conductivity was higher in the crystalline state than in the corresponding amorphous state, and the temperature behaviour was Arrhenius as opposed to VTF-like. Of course, the ionic transport would be highly anisotropic at a local level in these systems with the lithium ion diffusion pathway being through the tunnel. Bruce *et al*. have postulated that lithium ions move from one five-coordinate site to another via vacancy defects that are present in all crystalline materials above 0 K. The size of the crystallites (between 170 nm and 250 nm) appeared to have a great effect on the level of bulk conductivity, decreasing with decreasing crystallite size, most likely due to an increased level of disordered domains and misalignment between tunnels, which act as impediments to ion transport. This work opens up the possibility of enhancing ion transport still further by controlling the nanostructure in 2D or even 3D and possibly increasing the vacancy concentration by stoichiometry control, thereby overcoming some of the present limitations of crystalline polymer electrolytes.

Golodnitsky *et al*. have also confirmed the motion of lithium ions along the PEO helices found in LiI-PEO polymer electrolytes [27, 28]. In their work they showed that stretching these materials resulted in more ordering in the semicrystalline materials and this led to increased ionic conductivity along the stretching direction; this increase appeared to be decoupled from the polymer motions that usually govern ion transport in amorphous polymer electrolytes. SEM micrographs of the materials stretched both at room temperature and at 60 °C (near the melting point of PEO) showed a fibrous structure consisting of aligned oriented fibres with crystallite dimensions estimated between 70 and 80 nm.

A different approach to developing ordered polymer electrolyte materials is to prepare phase separated polymer blends or block copolymers whereby

the microstructure consists of non-conducting blocks separating sheets of ionically conducting domains which may contain polyether-salt type complexes [29, 30]. These materials consist of an amphiphilic helical polymer such as poly(2, 5, 8, 11, 14-pentaoxapentadecamethylene(5-hexadecyloxy-1,3-phenylene) and alkyl substituted polyether such as poly(tetramethylene oxide)-co-dodecamethylene as discussed by Wright in Chia *et al*. [29]. When these systems are fully blended and prior to phase separation, the ionic conductivity is relatively low (see Fig. 5.1) at ambient temperature, however, following the heating cycle de-blending occurs leading to conducting inter-lamella polyether/salt based regions. The oxygen rich helices in these regions are unable to reorganise as they are strongly constrained by the alkyl superstructure and thus the lithium ions are able to move through these oxygen rich spaces relatively unhindered. This leads to high lithium ion conductivity with little temperature dependence, i.e., low activation energy, as would be found in many inorganic crystalline fast ion conductors [1]. Similar approaches to decoupling lithium ion conductivity from polymer dynamics has been adopted by Ingram *et al*. [31] and Mayes *et al*. [32–34]. In the former case, side group liquid crystalline polyethers were used to constrain the ether oxygen rich regions in a bi-layer type structure.

Mayes and coworkers have developed block copolymer electrolytes based on poly (ethylene glycol)methyl ether methacrylate copolymerised with either



5.1 Conductivity plot for the polymer blend. The triangles are the first heating cycle, the circles the second. Reproduced with permission from Wright *et al.*, *Electrochimica Acta*, 48, 14–16 (2003) 1939.

lauryl methacrylate, butyl methacrylate or methyl methacrylate, whereby microphase separation within these leads to solid-like mechanical properties and high ionic conductivity (10^{-5}–10^{-6} S/cm at room temperature) confined to the ether rich phase. NMR measurements have suggested that ion transport in these systems is, however, still likely to be coupled to the chain dynamics and thus will be limited by this [32].

The idea of using phase separation at the molecular or nano-dimension to control ion transport also arises in the case of polyelectrolytes and polyelectrolyte gels. Nafion is the ultimate example of this type of material where the hydrophobic fluorinated backbone provides high mechanical stability in a hydrated membrane with interconnected nano domains rich in water and H^+ ions which lead to high proton conductivities [35]. In lithium ion conducting systems similar approaches can be taken where aprotic, device compatible solvents such as ethylene carbonate, dimethylacetamide and propylene carbonate are used to solubilise the polyelectrolyte by dissociating some fraction of the lithium ions from the backbone [36, 37]. These ions are then free to transport charge via the solvent rich domains. The material itself appears spatially homogeneous at the microscopic level, however, NMR has shown clear evidence of dynamic heterogeneity with a micelle-like phase separation suggested [37]. In the quest to further enhance ion transport in these types of materials, the use of nanosized particles as additives has been investigated. The following section discusses in detail the transport and structural implications of such composite systems.

5.3 Composite electrolytes

Ceramic particles were first introduced into polymer electrolytes to improve the mechanical stability [38]. Recently, composite electrolytes have attracted great interest due to reports of increased ionic conductivities, and a number of different filler particles, insulating (TiO_2, Al_2O_3) as well as electron (carbon) and cation (Li_2O-$3B_2O_3$) conducting, have been incorporated into mainly PEO based electrolytes [39–41]. The mechanism behind the increase in conductivity is still under debate. Various effects have been observed, the most obvious being a decreased crystallinity due to the plasticising effect of the filler particles. Only in part does this explain the increase in conductivity, as the conductivity is still an order of magnitude higher for the filled electrolytes than for electrolytes without fillers, even above the melting point of the PEO/salt complexes [39].

Due to the rather complex situation in the PEO-based electrolytes, with different crystalline and amorphous phases, various model systems have been used, such as low molecular weight poly(ethylene glycol) [42] and amorphous copolymers of PEO/PPO [43–45]. A homogenisation of the electrolytes with the addition of filler to PEO based electrolytes has been

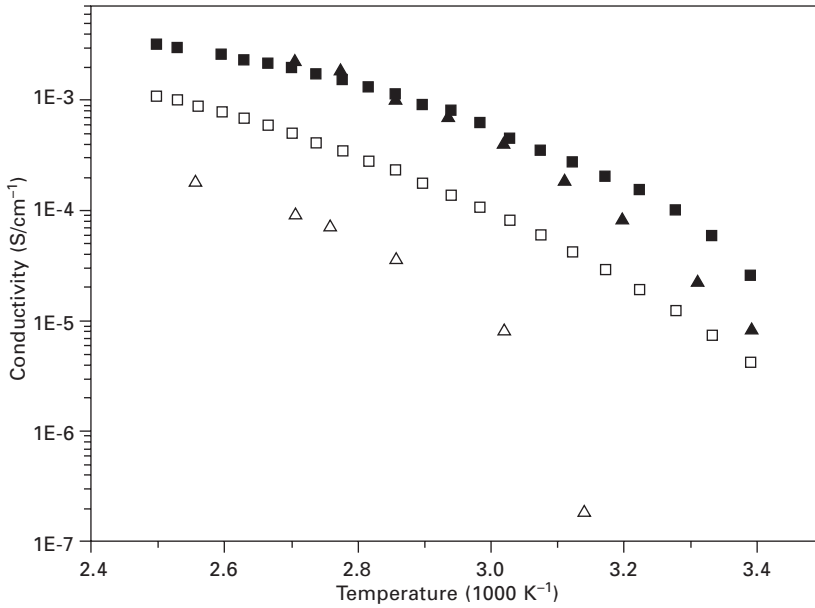
reported, [46] while the opposite has been observed for already amorphous electrolytes [44]. However, small increases in conductivity (a factor of 2–5) with the addition of nanosized filler particles to amorphous electrolytes [43] as well as to gel electrolytes [47] have been reported, pointing towards a different conduction mechanism, where the ions are less coupled to the segmental mobility of the polymer. This is further supported by an increased cationic transport number, observed for these types of systems [48]. The size, porosity, surface properties and dielectric constant of the particles also seem to be critical factors determining the conductivity increase [44, 49, 50]. These factors are discussed in more detail in the following sections.

5.3.1 Nanocomposite solid polymer electrolytes

In 1998 Croce *et al*. reported a more than one order of magnitude increase in conductivity for a PEO/$LiClO_4$ based electrolyte doped with 10 wt% nanosized TiO_2 [39]. Since then, a number of different non-conducting fillers have been reported to increase the ion conductivity in PEO based systems, such as TiO_2, SiO_2, Al_2O_3, MgO, $LiAlO_2$, $BaTiO_3$ and Li_2O-$3B_2O_3$ [39, 47, 48, 51–53]. Part of the increase in conductivity has been attributed to a decrease in crystallinity with the addition of ceramic particles, although this does not explain the enhancement in the molten state, amorphous polymer electrolytes [43, 54], or polymer gel electrolytes [55, 56].

The effect of nanosized ceramic particles on the conductivity of solid polymer electrolytes is shown in Fig. 5.2. In this case an amorphous polymer, 3PEG (a 3:1 co-polymer of ethylene glycol/propylene glycol) is compared to PEO. It can be seen that the addition of 10 wt% TiO_2 to the samples show half an order of magnitude increase in conductivity for the amorphous electrolyte, and over one order of magnitude for the PEO based electrolyte. The size of the filler is thought to play an important role. It has been shown that nanosized filler particles have a much larger effect on the conductivity than sub-micron sized, and furthermore the maximum conductivity enhancement has been found at lower loadings for smaller particles. [57–59] Other filler properties that influence the conductivity are the acid/base nature of the surface of the fillers, where neutral and acidic Al_2O_3 particles have been shown to give better electrochemical stability than basic, possibly being due to hydrogen bonds between the OH groups on the filler surface and the anion of the salt [42, 51, 60].

The TiO_2 filler has been shown to have a larger effect on the conductivity than Al_2O_3, both in the amorphous as well as in the semicrystalline polymer-based electrolytes [39, 44]. The dielectric constant of the fillers has been suggested to play an important role [44]. The interaction energy between the point charge and the filler is stronger in TiO_2 ($\varepsilon_r = 435$) than in Al_2O_3 ($\varepsilon_r = 20$, for particles with no surface treatment) due to the relative strength of



5.2 Arrhenius plot for electrolytes based on 3PEG + 1.5 mol/kg $LiClO_4$, unfilled (□) doped with 10 wt% TiO_2 (■) compared to an electrolyte based on the semicrystalline PEO_8LiClO_4, unfilled (Δ) and doped with 10 wt% TiO_2 (▲) [39, 44].

their dielectric constants [53, 61]. The TiO_2 filler will have a strong attractive effect, much greater than that of the Al_2O_3 on the other components in the electrolytes. Furthermore, the binding energy of the Li ion at the surface of a high dielectric constant filler is of the same order of magnitude as the binding energies of the Li cation with 1,2 dimethoxyethane (an analogue of PEO) [62]. The lithium ion should therefore be able to reside either on the TiO_2 or the polymer chain, providing more sites for conduction. Up to a distance of 3 Å, the lithium ion may see an even potential between the TiO_2 filler and the ether oxygen sites on the polymer, which may be of importance for improved conductivity in polymer electrolytes [44].

Raman spectroscopy

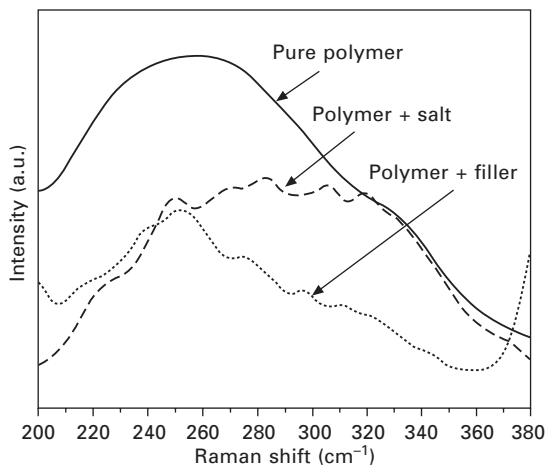
The disordered longitudinal acoustic modes (D-LAM) are found in amorphous polymers at low frequencies, below 300 cm^{-1}, and comprise a number of different backbone vibrations that in, e.g., PEO and PPO systems are a combination of O-C-C bending, C-O-C bending and O-C-C-O / C-C-O-C torsional modes [63]. The peak frequency of this broad mode is proportional to $1/n^2$ where n is the number of skeletal atoms in the chain. The D-LAM can also be used to probe conformational changes in polymers [64]. A decrease

in the flexibility of the chains has been shown to shift the peak to higher frequencies, which, for example, are the case when a salt is added to the polymer [65], see Fig. 5.3.

The band-width of the D-LAM mode has been related to the disorder of the polymer and determines the dispersion around the average conformation [66]. As shown in Fig. 5.3, the addition of nanoparticles to the polymer decreases the width of the D-LAM mode, suggesting an ordering of the polymer chains [44]. This is in agreement with recent MD simulations of PEO/$LiBF_4$/Al_2O_3, which showed that a fraction of the polymer, close to the nano-particle is immobile. According to the simulations, the immobilised chains bind a fraction of the anions to the surface of the filler, thereby giving an excess of Li^+ ions in the regions away from the surface [67]. Raman spectroscopy has also been widely used to probe the formation of free ions vs. ion pairs in electrolytes [66, 68]. Reports on the effect of filler particles on the dissociation of the salts suggest that there is no detectable change in the ion pair formation on the addition of filler particles, which has been shown in both solid polymer electrolytes as well as polymer gel electrolytes [44, 69].

Free volume – PALS characterisation

It is well established in the traditional polymer electrolyte field that ion transport is linked to polymer dynamics. These local polymer segmental motions (which also govern the glass transition temperature) are known to depend on the free volume available and hence the temperature dependent

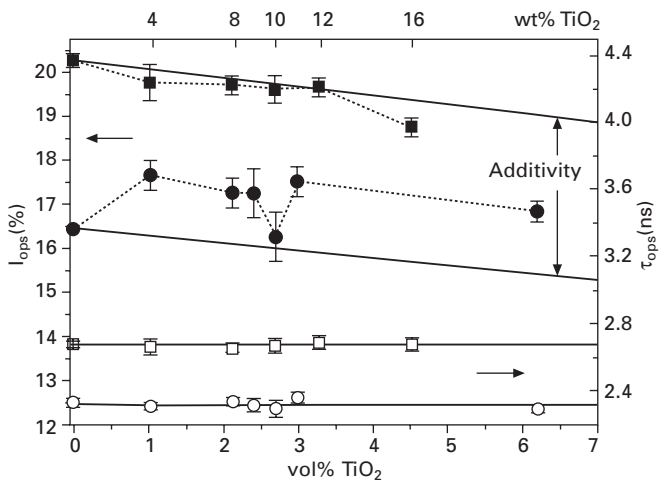


5.3 The D-LAM mode for three systems based on the amorphous 3PEG (PEO-co-PPO); pure polymer, polymer/salt and polymer/filler particles.

transport properties are usually well fitted by the Vogel–Tamman–Fulcher (VTF) equation [2]. Therefore an ability to measure free volume, to determine changes in this with composition and to correlate these with conductivity in electrolyte systems is of some significance.

Positron annihilation lifetime spectroscopy (PALS) is a technique that is capable of probing free volume and hence packing in materials such as these. Figure 5.4 shows the ortho-positronium lifetime, τ_{ops}, and intensity, I_{ops}, as a function of the amount of filler in these systems, both for the salt free and the salt-containing series. The results show that the τ_{ops} does not change with the addition of filler, indicating that the average size of the free volume holes is constant, and this seems to be independent of the presence of salt in the polymer. Addition of filler to the salt-free system decreases I_{ops} (related to the concentration of free volume) linearly and according to volume additivity, as seen in Fig. 5.4, as expected if the TiO_2 particles are acting as an inaccessible space filler. This behaviour has been observed in other polymer/filler systems, e.g., LDPE/ultrasil [70].

Addition of filler to the salt containing electrolytes shows a positive deviation in free volume concentration, I_{ops}, indicating a greater than additive number of free volume sites when TiO_2 is added to the polymer electrolyte. This result suggests that there may be competition between the polymer-Li^+ interaction and the polymer-TiO_2 interaction, and that by introducing the TiO_2 particles into the polymer electrolyte there is a change in the structure of the polymer.

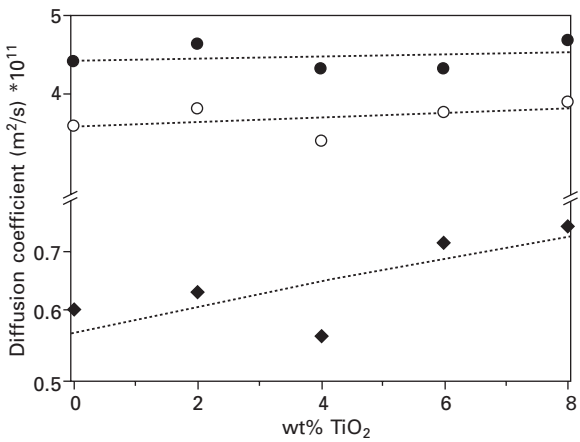


5.4 Mean orthopositronium lifetime, τ_{ops}, (□, ○) and mean orthopositronium intensity, I_{ops}, (■, ●) for pure 3PEG (squares) and 3PEG/$LiClO_4$ (circles) as a function of TiO_2 content.

5.3.2 Nanocomposite polymer gel electrolytes

Gel electrolytes are formed by immobilising a liquid electrolyte, typically a mixture of ethylene carbonate and propylene carbonate, in a relatively non-interactive polymer matrix such as PMMA, PVdF, PAN, etc. [71–76]. It has been shown, however, that although gel electrolytes appear solid-like, they exhibit flow properties [77, 78]. One way of increasing the mechanical stability of the gels is to add filler particles.The addition of insulating filler particles to polymer gel electrolytes, based on polymer matrixes such as PVC, PAN and PMMA, has also been shown to effectively control the growth of the passivation layer on the Li electrode, as well as increase its stability during long-term storage. However, the ionic conductivity has been shown to be unchanged or, for some compositions, slightly increased, but the effect is much less than in the case of solid polymer electrolytes [55, 56, 79–82].

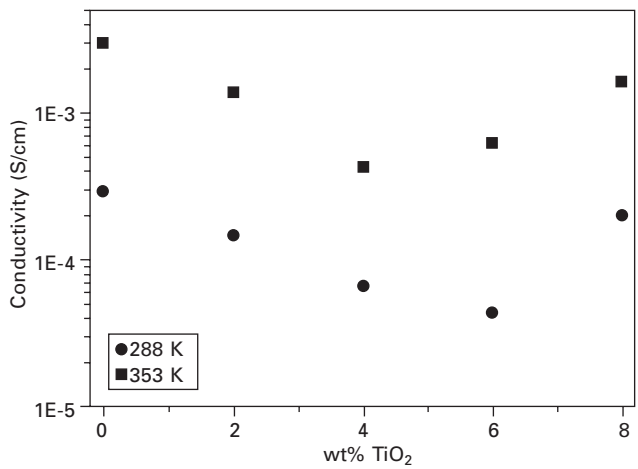
An increased diffusion coefficient for the cation has been reported for PMMA-based gels with the addition of nanosized TiO_2 (Fig. 5.5) [69]. The diffusion coefficient of the cation is about eight times lower than that of the solvent ($D_{solv}/D_{Li} \approx 8$) for the unfilled sample; this ratio decreased to less than six for the sample with the highest TiO_2 content. Other work has shown that this factor is ~25 for cations in EO_9-grafted PMMA polymers used as matrices in the gels [83]. Less interactive matrixes than PMMA, such as PVdF, have been shown to have a D_{solv}/D_{Li} ratio of around two [76], which is consistent with values for pure liquid electrolytes [83–86]. These results suggest that the addition of filler to the PMMA based gels either decreases the polymer/cation interaction (less coordinating polymer matrix), or introduces



5.5 Cation and solvent self-diffusion coefficients for PMMA based gel electrolytes, (Li:♦, EC:●, gBL:○), as a function of filler content measured at room temperature (22 °C), where the average molar ratio of EC:PC:$LiClO_4$:PMMA is 46.5:19:4.5:30.

a new pathway or transport mechanism for the lithium ions. It is also evident that the diffusion coefficients of solvent molecules remain unchanged with the addition of filler to the gel electrolyte. The increased mobility of the Li ions with increasing amount of filler is even more pronounced at elevated temperatures.

For this particular system (PMMA-EC-PC-$LiClO_4$) the conductivity has been shown to go through a minimum as a function of addition of filler particles, see Fig. 5.6. To some extent, the decrease is expected given that the filler particles must have a certain blocking effect on long-range ion transport, both filling volume and also making remaining conduction pathways more tortuous. Such an effect would not have a major impact on the NMR measured diffusivity. One hypothesis is that the filler particles interact with the anions, thereby slowing them down. For the sample with 4 wt% filler, this means that the decrease in anion mobility is larger than any increase in cation mobility. On the basis that the cation transport number is probably around 0.3 to 0.5 in the unfilled gel, this anion slow-down might explain as much as two-thirds of the discrepancy; the remainder can then be explained in terms of the volume filling and tortuosity effects referred to above. Interestingly, such an anion slow-down perforce produces an increase in the cation transport number, as has been observed in some systems. The high conductivity of the 8 wt% sample is therefore intriguing, since the decrease in mobility of the anion must be compensated by the increase in cation mobility. In fact, a percolation phenomenon has been suggested whereby new pathways for lithium ions open up on the surface of the filler particles. Although the volume content of TiO_2 is too low for a three-dimensional static percolation



5.6 The conductivity as a function of filler content measured at 25 and 80 °C for the PMMA-based gel electrolytes, where the average molar ratio of EC:PC:$LiClO_4$:PMMA is 46.5:19:4.5:30.

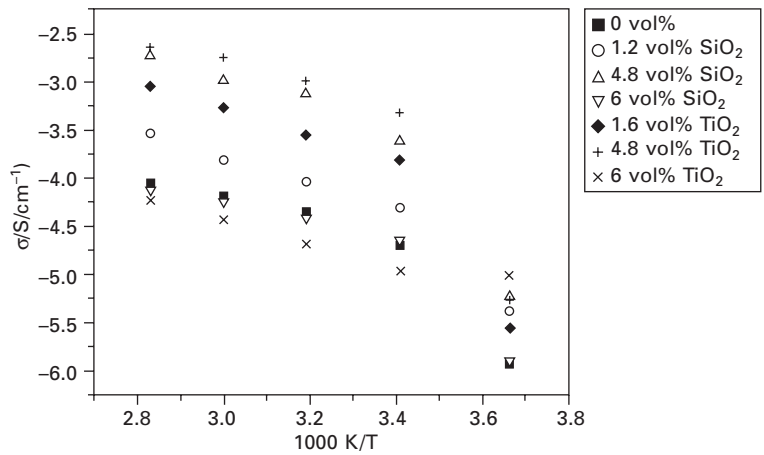
scenario, dynamic percolation can be expected to occur at very low concentration fractions [87, 88]. Experimentally, a similar scenario has also been observed in a micro-phase-separated polymer electrolyte at very low concentrations [89].

5.3.3 Nanocomposite polyelectrolyte gels

Polyelectrolyte gels have a polymeric backbone with covalently bonded ionising groups attached to them. The polymer in these gels therefore actively participates in the conductivity process in contrast to the systems described above. The overall functionality of the system depends on the ionisable groups attached to the polymer as well as the solvent component which can contribute dissociating power as well as high mobility. The main advantage of polyelectrolyte gels over the polymer gel is the absence of concentration gradients since, in principle, only a single ionic species will be responsible for the ionic conductivity. The lithium ions also have long-range order, which is absent in polymer gels. These advantages come about due to the anion being covalently bound to the polymer.

In this section the effect of adding nanosized TiO_2 and SiO_2 on the transport properties of polyelectrolytes is described. The system which has been investigated extensively is that based on the polyelectrolyte poly (lithium 2-acrylimido-2-methyl-1-propanesulfonate) [Li poly(AMPS)] and its copolymers with N,N-dimethylacryl amide (DMAA).

The effect of filler type and concentration on the ionic conductivity is shown as a function of temperature in Fig. 5.7. At all temperatures the ionic conductivity is enhanced by filler addition. This behaviour is more pronounced



5.7 Conductivity as a function for temperature for the poly(Li-AMPS)-based system with addition of TiO_2 and SiO_2 filler.

at elevated temperatures predominately due to viscosity changes in the gel, the conductivity is enhanced by almost a factor of ten at 80 °C. The increase in ionic conductivity has been postulated to be the result of the filler aiding in dissociating the lithium ions from the polymer backbone perhaps by interacting preferentially with the polymer bound sulfonate anion. This dissociation leads to a higher number of charge carriers.

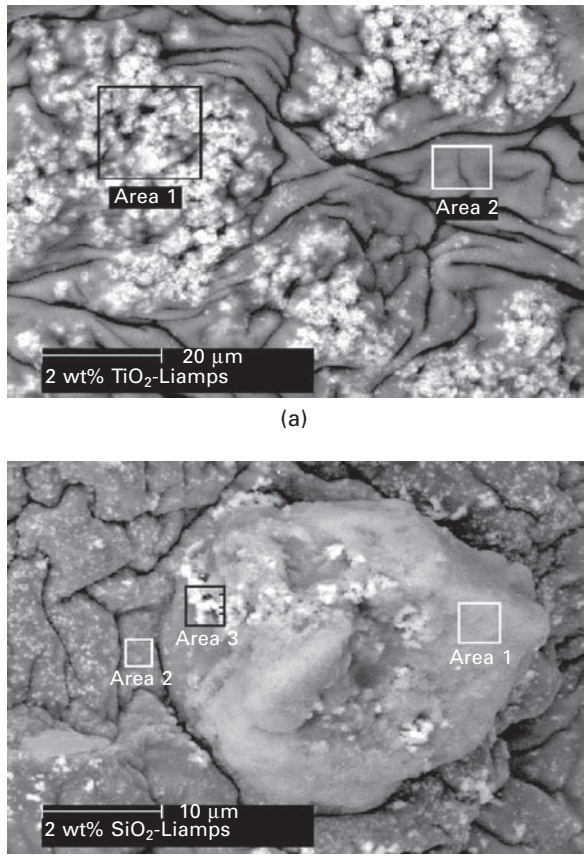
The conductivity of a material can be described by the Nernst–Einstein equation

$$\sigma_{\text{calc}} = \frac{Nq^2}{kT}[D_{\text{cation}} + D_{\text{anion}}] \qquad 5.1$$

such that higher conductivities can be achieved by either increasing the number of charge carriers and/or their diffusivity. An increase in the average diffusivity of the lithium ions was indeed found upon the addition of the nanosized inorganic fillers. The enhanced diffusivity (which would lead to a higher μ_i) is not likely to be due to an increase in the mobility of the overall system; such an enhancement would be manifested in high solvent diffusion coefficients (the measured average solvent diffusivity remained constant with filler addition) and likely lowering of Tg. In fact, neither of these appeared to change in the composite systems compared to the unfilled polyelectrolyte gel. The question then arises as to why the lithium diffusion coefficient increases. One can consider that, in the polyelectrolyte gel there are at least two possible sites for Li^+-associated to the sulfonated anion on the polymer backbone, or solvated by the solvent. The likelihood is that both sites are occupied to some extent and that rapid exchange occurs between these sites as was discussed by Every *et al*. [37]. In the absence of the filler, the associated species dominates and hence the average diffusion coefficient (as measured by pulsed field gradient NMR) is low which correlates with the lower ionic conductivities.

Electron microscopy EDX measurements (see Fig. 5.8) shows direct evidence for SO_3^--filler interactions with the simultaneous detection of Ti and S atoms throughout the sample. This could then explain the enhanced Li^+ ion transport as it would allow a greater fraction of dissociated lithium ions leading to a higher average lithium diffusion coefficient. Furthermore, this model was further strengthened by the fact that SiO_2, according to the SEM images, did not appear to interact as strongly or as readily with the sulphonate group within the polyelectrolyte which led to a lessened effect of filler on lithium ion diffusivity as well as a smaller effect on the overall conductivity in this system compared to TiO_2-based composites.

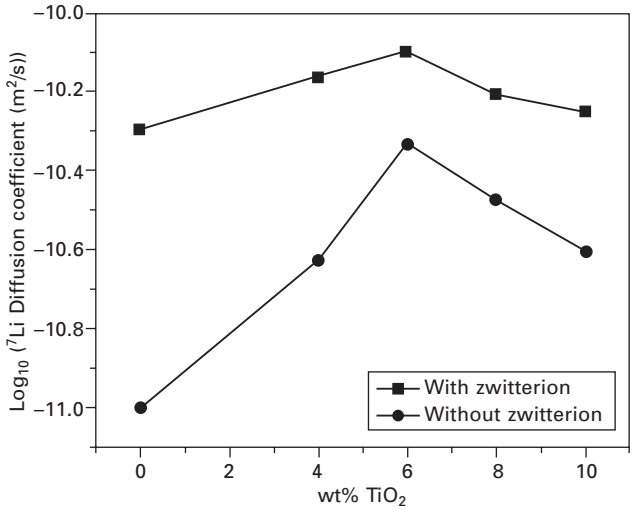
It would appear therefore, that a larger surface area of filler should maximise polymer filler interactions and lead to higher dissociation and ionic conductivity. However, it was shown that as more filler was incorporated the conductivity sharply decreased. This was suggested to be a result of agglomeration of the



5.8 Back scattered image of poly(Li-AMPS)-based system with the addition 0.8 vol% TiO_2. Area 1: TiO_2, sulphur and potassium. Area 2: carbon and oxygen; this area most likely represents the solvent rich region. (b) Back scattered image of poly(Li-AMPS)-based system with the addition of 1.2 vol% SiO_2. Area 1: SiO_2 Area 2: SiO_2, sulphur and potassium. Area 3: carbon and oxygen.

nanosized particles which led a decrease in the effective surface area. Given the high surface energy of nanosized particles it is not surprising that such agglomeration occurs.

Another recent exciting development in these composite polyelectrolyte systems is the potential of replacing some of the volatile solvents with materials of lower volatility and higher dissociating ability. Certainly ionic liquids have been suggested as possible replacements [90, 91], however, these solvents themselves contribute to the ionic conductivity since they consist of charged ionic species. On the other hand zwitterion type materials have recently found use in electrolytes as effective ion dissociators [92]. These can be best



5.9 ^{7}Li diffusion coefficient at T = 70 °C of the [Li poly(AMPS)] based gels with and without the addition of zwitterion, as a function of filler content.

described as an ionic liquid whereby the cationic unit is covalently tethered with an anionic unit [93]. The advantage of zwitterion over traditional ionic liquids is that, unlike these systems, the component ions cannot migrate with the potential gradient making them an excellent solvent for electrochemical applications [94].

MacFarlane *et al*. [92] have shown that these zwitterions are themselves excellent additives which encourage the dissociated state of the lithium from the polyelectrolyte backbone thereby influencing the local gel structure. In addition, when a composite polyelectrolyte gel is formed with the incorporation of TiO_2 nanosized particles, the enhancement of ionic conductivity and lithium ion diffusivity is even greater. A morphological or structural understanding of the additive behaviour reported here is still lacking, however, the potential for the development of these materials in devices presents exciting opportunities.

5.4 Conclusions

We have seen how modifying the local polymer morphology by controlling the phase behaviour, for example, crystalline versus amorphous, phase separated polymer blends, chemical control of copolymer composition and the use of nanosized inorganic additives has led to improved ionic transport of target ions in polymer electrolytes. The ability to control the structure of solid electrolyte materials at the nanodomain will be a powerful tool in the design of new materials with desired ion transport properties which will facilitate new electrochemical device technologies.

5.5 References

1. P.G. Bruce, ed. (1995), *Solid state electrochemistry*, Cambridge University Press: Cambridge; New York, NY, USA.
2. F.M. Gray. *Polymer Electrolytes*, ed. J.A. Connor. (1997), Canterbury: The Royal Society of Chemistry.
3. J.R. MacCallum and C.A. Vincent, eds (1987), *Polymer electrolytes rivers 1 and 2,* Elsevier Applied Science Publishers Ltd.
4. P.V. Wright, *MRS Bulletin*, **27**, 8 (2002) 597.
5. D.R. MacFarlane, P. Meakin, J. Sun, N. Amini, and M. Forsyth, *Journal of Physical Chemistry B*, **103**, 20 (1999) 4164.
6. D.R. MacFarlane, J.H. Huang, and M. Forsyth, *Nature*, **402**, 6763 (1999) 792.
7. J.H. Huang, M. Forsyth, and D.R. MacFarlane, *Solid State Ionics*, **136**, (2000) 447.
8. M. Forsyth, J. Huang, and D.R. MacFarlane, *Journal of Materials Chemistry*, **10**, 10 (2000) 2259.
9. D.R. MacFarlane and M. Forsyth, *Advanced Materials*, **13**, 12–13 (2001) 957.
10. D.R. MacFarlane, P. Meakin, N. Amini, and M. Forsyth, *Journal of Physics – Condensed Matter*, **13**, 36 (2001) 8257.
11. J. Golding, N. Hamid, D.R. MacFarlane, M. Forsyth, C. Forsyth, C. Collins, and J. Huang, *Chemistry of Materials*, **13**, 2 (2001) 558.
12. S. Forsyth, J. Golding, D.R. MacFarlane, and M. Forsyth, *Electrochimica Acta*, **46**, 10–11 (2001) 1753.
13. A.J. Hill, J. Huang, J. Efthimiadis, P. Meakin, M. Forsyth, and D.R. MacFarlane, *Solid State Ionics*, **154**, (2002) 119.
14. J. Efthimiadis, S.J. Pas, M. Forsyth, and D.R. MacFarlane, *Solid State Ionics*, **154**, (2002) 279.
15. J. Golding, S. Forsyth, D.R. MacFarlane, M. Forsyth, and G.B. Deacon, *Green Chemistry*, **4**, 3 (2002) 223.
16. J. Sun, D.R. MacFarlane, and M. Forsyth, *Solid State Ionics*, **148**, 1–2 (2002) 145.
17. M. Forsyth, J. Pringle, D.R. MacFarlane, and W. Ogihara, ion conduction in organic plastic crystals in *Ionic Liquid: The Front and Future of Material Development.* H. Ohno, editor. (2003), CMC Publishing Co., Ltd: Tokyo. p. 144.
18. A. Seeber, M. Forsyth, C. Forsyth, S. Forsyth, G. Annat, and D.R. MacFarlane, *Physical Chemistry Chemical Physics*, **5**, (2003) 2692.
19. J. Efthimiadis, G.J. Annat, M. Forsyth, and D.R. MacFarlane, *Physical Chemistry Chemical Physics*, **5**, 24 (2003) 5558.
20. J. Efthimiadis, M. Forsyth, and D.R. MacFarlane, *Journal of Materials Science*, **38**, 15 (2003) 3293.
21. J. Adebahr, P. Johansson, P. Jacobsson, D.R. MacFarlane, and M. Forsyth, *Electrochimica Acta*, **48**, 14–16 (2003) 2283.
22. S. Long, D.R. MacFarlane, and M. Forsyth, *Solid State Ionics*, **161**, 1–2 (2003) 105.
23. C. Berthier, W. Gorecki, M. Minier, M.B. Armand, J.M. Chabagno, and P. Rigaud. *Solid state ionics*, **11**, (1983) 91.
24. M.A. Ratner, Aspects of the Theoretical Treatment of Polymer Solid Electrolytes: Transport Theory and Models, in *Polymer electrolyte reviews 1*, J.R. MacCallum and C.A. Vincent, ed (1987), Elsevier Applied Science Publishers Ltd. p. 173.
25. Z. Gadjourova, Y.G. Andreev, D.P. Tunstall, and P.G. Bruce, *Nature*, **412**, 6846 (2001) 520.

26. Z. Stoeva, I. Martin-Litas, E. Staunton, Y.G. Andreev, and P.G. Bruce, *J. Am. Chem. Soc*., **125**, (2003) 4619.
27. D. Golodnitsky, E. Livshits, Y. Rosenberg, I. Lapides, and E. Peled, *Solid State Ionics*, **147**, (2002) 265.
28. D. Golodnitsky, E. Livshits, and E. Peled, *Macromolecular Symposia*, **203**, (2003) 27.
29. F. Chia, Y. Zheng, J. Liu, N. Reeves, G. Ungar, and P.V. Wright, *Electrochimica Acta*, **48**, 14–16 (2003) 1939.
30. Y.G. Zheng, F.S. Chia, G. Ungar, T.H. Richardson, and P.V. Wright, *Electrochimica Acta*, **46**, 10–11 (2001) 1397.
31. C.T. Imrie, and M.D. Ingram, *Electrochimica Acta*, **46**, 10–11 (2001) 1413.
32. D.J. Harris, T.J. Bonagamba, K. Schmidt-Rohr, P.P. Soo, D.R. Sadoway, and A.M. Mayes, *Macromolecules*, **35**, 9 (2002) 3772.
33. P.P. Soo, B.Y. Huang, Y.I. Jang, Y.M. Chiang, D.R. Sadoway, and A.M. Mayes, *Journal of the Electrochemical Society*, **146**, 1 (1999) 32.
34. A.V.G. Ruzette, P.P. Soo, D.R. Sadoway, and A.M. Mayes, *Journal of the Electrochemical Society*, **148**, 6 (2001) A537.
35. K.D. Kreuer, *Chem. Mater*., **8**, 3 (1996) 610.
36. J. Travas-Sejdic, R. Steiner, J. Desilvestro, and P. Pickering, *Electrochimica Acta*, **46**, 10–11 (2001) 1461.
37. M. Forsyth, H.A. Every, D.R. MacFarlane, J. Travas-Sejdic, and R. Steiner, *Macromolecular Materials and Engineering*, **287**, 8 (2002) 523.
38. J.E. Weston and B.C.H. Steele, *Solid State Ionics*, **7**, 1 (1982) 75.
39. F. Croce, G.B. Appetecchi, L. Persi, and B. Scrosati, *Nature*, **394**, (1998) 456.
40. G.B. Appetecchi and S. Passerini, *Electrochimica Acta*, **45**, 13 (2000) 2139.
41. P. Mustarelli, E. Quartarone, C. Tomasi, and A. Magistris, *Solid State Ionics*, **86–8**, (1996) 347.
42. M. Marcinek, A. Bac, P. Lipka, A. Zalewska, G. Zukowska, R. Borkowska, and W. Wieczorek, *Journal of Physical Chemistry B*, **104**, 47 (2000) 11088.
43. A.S. Best, A. Ferry, D.R. MacFarlane, and M. Forsyth, *Solid State Ionics*, **126**, (1999) 269.
44. A.S. Best, J. Adebahr, P. Jacobsson, D.R. MacFarlane, and M. Forsyth, *Macromolecules*, **34**, 13 (2001) 4549.
45. J. Adebahr, A.S. Best, N. Byrne, P. Jacobsson, D.R. MacFarlane, and M. Forsyth, *Physical Chemistry Chemical Physics*, **5**, 4 (2003) 720.
46. Y. Dai, S. Greenbaum, D. Golodnitsky, G. Ardel, E. Strauss, E. Peled, and Y. Rosenberg, *Solid State Ionics*, **106**, (1998) 25.
47. M. Morita, T. Fujisaki, N. Yoshimoto, and M. Ishikawa, *Electrochimica Acta*, **46**, 10–11 (2001) 1565.
48. C. Capiglia, P. Mustarelli, E. Quartarone, C. Tomasi, and A. Magistris, *Solid State Ionics, Diffusion & Reactions*, **118**, 1–2 (1999) 73.
49. E. Quartarone, P. Mustarelli, and A. Magistris, *Solid State Ionics*, **110**, (1998) 1.
50. M. Forsyth, D.R. MacFarlane, A.S. Best, J. Adebahr, P. Jacobsson, and H. A.J., *Solid State Ionics*, **147**, (2002) 203.
51. W. Wieczorek, J.R. Stevens, and Z. Florjanczyk, *Solid State Ionics*, **85**, 1–4 (1996) 67.
52. B. Kumar, L. Scanlon, R. Marsh, R. Mason, R. Higgins, and R. Baldwin, *Electrochimica Acta*, **46**, 10–11 (2001) 1515.

53. H.Y. Sun, H.J. Sohn, O. Yamamoto, Y. Takeda, and N. Imanishi, *Journal of the Electrochemical Society*, **146**, 5 (1999) 1672.
54. W. Wieczorek, P. Lipka, G. Zukowska, and H. Wycislik, *Journal of Physical Chemistry B*, **102**, 36 (1998) 6968.
55. S. Rajendran and T. Uma, *Materials Letters*, **44**, 3–4 (2000) 208.
56. K.H. Lee, Y.G. Lee, J.K. Park, and D.Y. Seung, *Solid State Ionics*, **133**, (2000) 257.
57. Z.Y. Wen, T. Itoh, T. Uno, M. Kubo, and O. Yamamoto, *Solid State Ionics*, **160**, 1–2 (2003) 141.
58. M. Dissanayake, P. Jayathilaka, R.S.P. Bokalawala, I. Albinsson, and B.E. Mellander, *Journal of Power Sources*, **119**, (2003) 409.
59. C. Capiglia, J. Yang, N. Imanishi, A. Hirano, Y. Takeda, and O. Yamamoto, *Solid State Ionics*, **154**, (2002) 7.
60. C.H. Park, D.W. Kim, J. Prakash, and Y.K. Sun, *Solid State Ionics*, **159**, 1–2 (2003) 111.
61. J. Israelachvili, 2nd edn ed. *Intermolecular & Surface Forces*. (1991), London: Academic Press. p. 191.
62. A. Sutjianto and L.A. Curtiss, *J Phys Chem*, **102**, 6 (1998) 968.
63. R.G. Snyder, N.E. Schlotter, R. Alamo, and L. Mandelkern, *Macromolecules*, **19**, (1986) 621.
64. H. Nishide, M. Ohyanagi, O. Okada, and E. Tsuchida, *Macromolecules*, **19**, (1986) 496.
65. S. Shashkov, S. Wartewig, B. Sandner, and J. Tubke, *Solid State Ionics*, **90**, 1–4 (1996) 261.
66. L.M. Torell and S. Schantz, in *Polymer electrolyte reviews 2*, J.R. MacCallum and C.A. Vincent, eds (1989), Elsevier Science Publishers Ltd.
67. H. Kasemagi, M. Klintenberg, A. Aabloo, and J.O. Thomas, *Solid State Ionics*, **147**, 3–4 (2002) 367.
68. S. Schantz, L.M. Torell, and J.R. Stevens, *J Appl Phys*, **64**, 4 (1988) 2038.
69. J. Adebahr, N. Byrne, M. Forsyth, D.R. MacFarlane, and P. Jacobsson, *Electrochimica Acta*, **48**, 14–16 (2003) 2099.
70. J. Tino, J. Kristiak, Z. Hlouskova, and O. Sausa, *Eur. Polym. J.*, **29**, 1 (1993) 95.
71. E. Cazzanelli, G. Mariotto, G.B. Appetecchi, F. Croce, and B. Scrosati, *Electrochimica Acta*, **40**, 13 (1995) 2379.
72. D. Ostrovskii, A. Brodin, L.M. Torell, G.B. Appetecchi, and B. Scrosati, *Journal of Chemical Physics*, **109**, 17 (1998) 7618.
73. D. Ostrovskii, L.M. Torell, G.B. Appetecchi, and B. Scrosati, *Solid State Ionics, Diffusion & Reactions*, **106**, 1–2 (1998) 19.
74. J. Adebahr, P. Gavelin, D. Ostrovskii, L.M. Torell, and B. Wesslen, *Journal of Molecular Structure*, **482**, 483 (1999) 487.
75. G.B. Appetecchi, F. Croce, G. Dautzenberg, F. Gerace, S. Panero, E. Spila, and B. Scrosati, *Gazzetta Chimica Italiana*, **126**, (1996) 405.
76. I.M. Ward, M.J. Williamson, H.V.S.A. Hubbard, J.P. Southall, and G.R. Davies, *J Power Sources*, **81–82**, (1999) 700.
77. Y.K. Yarovoy, H.P. Wang, and S.L. Wunder, *Solid State Ionics*, **118**, (1999) 301.
78. O. Bohnke, G. Frand, M. Rezrazi, C. Rousselot, and C. Truche, *Solid State Ionics*, **66**, 1–2 (1993) 97.
79. S. Slane and M. Salomon, *Journal of Power Sources*, **55**, 1 (1995) 7.
80. G.B. Appetecchi, P. Romagnoli, and B. Scrosati, *Electrochemistry communications*, **3**, (2001) 281.

81. E. Morales and J.L. Acosta, *Solid State Ionics*, **111**, 1–2 (1998) 109.
82. K.M. Kim, N.G. Park, K.S. Ryu, and S.H. Chang, *Polymer*, **43**, 14 (2002) 3951.
83. J. Adebahr, M. Forsyth, P. Gavelin, P. Jacobsson, and G. Orädd, *J Phys Chem*, **106**, 47 (2002) 12119.
84. M. Clericuzio, W.O. Parker, M. Soprani, and M. Andrei, *Solid State Ionics*, **82**, 3–4 (1995) 179.
85. G. Fleischer, H. Scheller, J. Kärger, A. Reiche, and B. Sandner, *Journal of Non-Crystalline Solids*, **235–237**, (1998) 742.
86. K. Hayamizu, Y. Aihara, S. Arai, and W.S. Price, *Electrochimica Acta*, **45**, 8–9 (2000) 1313.
87. R. Granek and A. Nitzan, *J Chem Phys*, **90**, (1989) 3784.
88. A.L.R. Bug and Y. Gefen, *Phys Rev A*, **35**, (1987) 1301.
89. A. Ferry, P. Jacobsson, and L.M. Torell, *Electrochimica Acta*, **40**, (1995) 2369.
90. M. Forsyth, J. Sun, F. Zhou, and D.R. MacFarlane, *Electrochimica Acta*, **48**, 14–16 (2003) 2129.
91. J. Sun, D.R. MacFarlane, and M. Forsyth, *Solid State Ionics*, **147**, 3–4 (2002) 333.
92. C. Tiyapiboonchaiya, J.M. Pringle, J.Z. Sun, N. Byrne, P.C. Howlett, D.R. Macfarlane, and M. Forsyth, *Nature Materials*, **3**, 1 (2004) 29.
93. H. Ohno, M. Yoshizawa, and W. Ogihara, *Electrochimica Acta*, **48**, 14–16 (2003) 2079.
94. M. Yoshizawa, M. Hirao, K. Ito-Akita, and H. Ohno, *Journal of Materials Chemistry*, **11**, (2001) 1057.

6

Nanostructures in biological materials

H GAO and H YAO,
Max Planck Institute for Metals Research, Germany and
B JI
Tsinghua University, China

6.1 Introduction

Through millions of years of natural evolution, living organisms have developed elaborate bulk and surface nanostructures to improve and/or optimize various functions of life such as mechanical support, locomotion and prey capturing. Bone, tooth, and shell are hierarchical materials with nanostructures made of mineral crystals and protein. Wood and spider silk have nanostructures of crystalline and amorphous bio-polymers. While different biological materials differ widely in their hierarchical structures and can in many cases dynamically adapt to changing environments, at the nanostructure level they exhibit convergent evolution in the form of a network of interspersed hard and soft phases; the hard phase provides a basic structural scaffold for mechanical stability and stiffness, while the soft phase absorbs mechanical energy and provides toughness, support, and a damage buffer for the composite. Nature has also evolved various forms of surface nanostructures, such as the adhesive hairy structure on the foot of gecko, the anti-adhesive structures of plant leaves and insect bodies, and the hydrophobic hairy structure which allows insects like the water strider to walk on water via surface tension. The biological nanostructures can be assumed to have gradually developed and improved over the long course of natural evolution to maximize the chance of survival of various living creatures. In this evolution, mechanical forces have no doubt played critical roles in shaping biological systems into what they have become today.

This chapter discusses how nanostructures of biological systems are related to their mechanical properties. For the convenience of discussion, we will refer to materials with a nanostructure similar to that of bone as bone-like materials and a 'hairy' biological surface nanostructure as a gecko-like biological system. We begin by reviewing some typical nanostructures of biological systems, focusing on their geometries and characteristic length scales. We then discuss the mechanical properties of biological nanostructures, especially those related to their stiffness, toughness and adhesion. We show

that the nanometer scale plays a key role in allowing bone- and gecko-like biological systems to achieve their superior properties. The results suggest that the principle of flaw tolerance may have had an overarching influence on the evolution of the bulk nanostructure of bone-like materials and the surface nanostructure of gecko-like animal species. The nanoscale sizes allow the mineral nanoparticles in bone to achieve optimum fracture strength and the spatula nanoprotrusions in geckos to achieve optimum adhesion strength.

It is emphasized that, in both systems, strength optimization is achieved by restricting the characteristic dimension of the basic structure components to nanometer scale so that crack-like flaws do not propagate to break the desired structural link. The mechanical properties of biological nanostructures are strongly anisotropic due to the unidirectional alignment and large aspect ratios of their geometry. The large aspect ratio and optimized fracture strength of mineral crystals are found to be critical in allowing the soft protein in bone to support large mechanical loads with relatively small stress and to effectively dissipate large amounts of fracture energy via domain unfolding and slipping along the protein–mineral interface. Similarly, the large aspect ratio of nanoprotrusions allows the hairy nanostructure of geckos to dissipate adhesion energy via elastic instabilities upon detachment from a substrate.

Nanotechnology promises to enable mankind to eventually design materials using a bottom-up approach, i.e., to construct multi-functional and hierarchical material systems by tailor-designing structures from atomic scale and up. Currently, there is no theoretical basis on how to design a hierarchical material system to achieve a particular set of functions. The studies outlined here are part of a broader effort aimed to extract from convergent evolutions some basic principles of multiscale and multifunctional materials design. It is hoped that this chapter could stimulate some further interest in this emerging research field.

6.2 Nanostrucures of biological materials

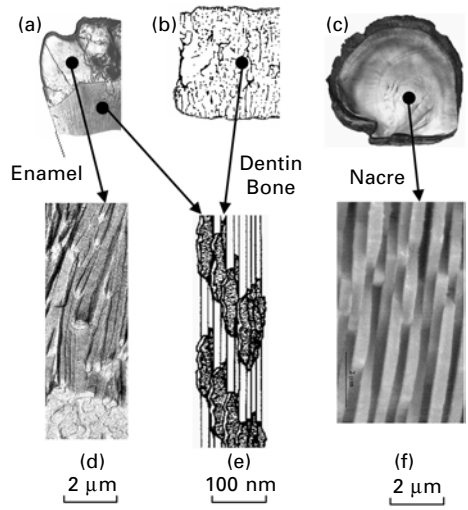
One of the important objectives of materials science and engineering has been to understand the relationships between the microstructure of materials and their macroscopic mechanical properties. From this point of view, knowledge about the structures and mechanical properties of biological systems are significant to materials research. In principle, living organisms should strive for efficient use of materials through optimized structure design; they should not waste valuable energy resources on developing/maintaining structures or materials which either function inefficiently or have no function at all. Structures that function abnormally or inefficiently are gradually eliminated or replaced by more efficient ones during the course of evolution. Those that are found to work are continuously optimized, allowing the

organisms to adapt better to the environment. Convergent evolution occurs either when genetically unrelated organisms are found to have adopted the same strategies to achieve the same objectives/properties/functions, or when genes associated with a given set of properties or functions of the same organism are found to be strongly conserved over the course of evolution. Convergent evolution is generally regarded as a strong hint that some properties important for the survival of an organism have been optimized.

6.2.1 Convergent bulk nanostructures of bone-like materials

Bone-like materials, including shell (Currey, 1977; Menig *et al*., 2000, 2001; Kamat *et al*., 2000; Jackson *et al*., 1988; Taylor, 1973; Currey and Taylor, 1974), tooth (Tesch *et al*., 2001; Weiner *et al*., 1999) and bone (Rho *et al*., 1998; Weiner and Wagner, 1998; Landis, 1995), are representatives of hard and tough materials in nature. These materials exhibit many levels of hierarchical structures from microscopic to macroscopic length scales. It has been argued (Gao *et al*., 2003, 2004) that the nanometer length scale plays a critically important role in such materials. The enamel of tooth (Figs 6.1(a) and 6.1(d) is made of long, more or less needle-like crystals about 15–20 nm thick embedded in a soft matrix (Tesch *et al*., 2001; Warshawsky, 1989). The nanostructure of bone (Figs 6.1(b) and 6.1(e) consists of mineral crystal platelets with thickness around a few nanometers embedded in a collagen matrix (Landis, 1995; Landis and Hodgens, 1996; Roschger *et al*., 2001). Nacre has the 'brick and mortar' structure (Figs 6.1(c) and 6.1(f) in which aragonite bricks with thickness around a few hundred nanometers are staggered and glued together by protein (Currey, 1977; Menig *et al*., 2000; Jackson *et al*., 1988). The mineral crystals in bone, tooth and nacre all have slender geometry with large aspect ratios. While the composite structures of bone-like materials have attracted broad interest, a fundamental question is why the nanometer scale is so important for these materials.

As nanocomposites of protein and mineral, bone-like materials have a similar stiffness to mineral but their fracture energy can be several orders of magnitude higher than that of mineral. For example, nacre, as a composite of $CaCO_3$ and protein, has a fracture energy about 3000 times higher than that of the monolithic $CaCO_3$ (Jackson *et al*., 1988). Generally, bone-like biocomposites exhibit a well-defined organization of components in the form of interlaced bricks separated by soft layers of protein. Mineral platelets are separated and glued together by protein. The high toughness of protein-mineral biocomposites have been attributed to various mechanisms (Ji and Gao, 2004a) including their hierarchical structures (Menig *et al*., 2000, 2001; Kamat *et al*., 2000), the effects of mechanical properties of protein on dissipating fracture energy (Smith *et al*., 1999), protein–mineral interface roughness



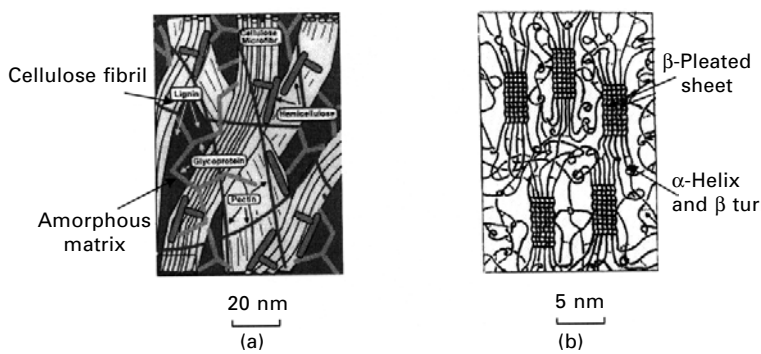
6.1 Bulk nanostructures of bone-like materials. Many hard biological tissues, such as tooth (a), vertebral bone (b) and shells (c) are made of nanocomposites with hard mineral platelets in a soft (protein) matrix. Enamel (d) is made of long, more or less needle-like crystals about 15–20 nm thick and 1000 nm long, with a relatively small volume fraction of a soft protein matrix. Dentin and bone (e) are made of plate-like crystals (2–4 nm thick, up to 100 nm long) embedded in a (collagen-rich) protein matrix. The volume ratio of mineral to matrix is of the order of 1:2. Nacre (f) is made of plate-like crystals (200–500 nm thick and a few micrometers long) with a small amount of soft matrix in between. All the composites share the structural feature of hard platelets with a very large aspect ratio, arranged parallel in a brick-and-mortar-like fashion. Figure adapted from Gao *et al.* (2003).

(Wang *et al.*, 2001) and reduction of stress concentration at a crack (Okumura and de Gennes, 2001).

The 'hard-nanoparticles-interspersed-in-soft-matrix' structure is a popular design in nature. Wood and spider dragline silk also exhibit such structures. In contrast to bone and nacre which are made of mineral and protein, the chemical constituents of wood are mainly cellulose, hemicellulose, and lignin. Wood has a cellular structure designed to provide mechanical support using minimum materials. The wood cell wall is a fiber composite consisting of partly crystalline cellulose fibrils in an amorphous hemicellulose-lignin matrix. The fibers are wound around the tube-shaped wood cells in several concentric layers with different helical orientations and different helical pitches (Brett and Waldron, 1981; Fengel and Wegener, 1984; Greil, 2001). The structure of the multiphase cell wall can be compared to the 'brick and mortar' structure in bone and shells where cellulose serves as the hard inclusions, while hemicelluloses, lignin and pectin serve as the soft matrix (see Fig. 6.2(a). A

key component in the structure of the cell wall is cellulose, a long-chain linear condensation polymer of β-D-glucose with three free hydroxyl groups on each monomeric unit, resulting in strong inter- and intramolecular hydrogen bonds. Because of the hydrogen bond network and restricted rotation around the polymeric 1,4-β-linkage, cellulose is stiff enough to provide mechanical stiffness and stability to the cell wall. The intermolecular hydrogen bonds in cellulose coupled with the inherent stiffness of the polymeric chains result in high levels of crystallinity. The cellulose molecules aggregate into microfibrils with sizes ranging around several tens of nanometers.

Some web-spinning spiders produce a variety of high-performance structural fibers. X-ray diffraction studies have shown that virtually all silk contains protein crystals, mostly in the form of β-pleated sheet crystals from tandemly repeated amino acid sequences rich in small amino acid residues (Gosline *et al*., 1999). The β-pleated sheet crystals are relatively hard and stiff and can occupy 40–50% of the total volume of the silk fiber (Termonia, 1994). The remainder consists of much less ordered, amorphous protein chains, α helices and β turns. This kind of arrangement again exhibits the hardness and toughness combination: hard β-sheet crystals among soft α helices and β turns form a crystal-cross-linked and crystal-reinforced polymer network (see Fig. 6.2(b)). X-ray studies indicate that the characteristic dimensions of crystals are approximately several nanometers (Gosline *et al*., 1999). The amorphous chains, which interconnect the crystals, are estimated to be 16–20 amino



6.2 The nanostructures of wood and spider silk. (a) In the nanostructure of wood cell wall, cellulose serves as the hard and stiff inclusion while lignin, heicellulose and pectin form the soft matrix. The diameter of the cellulose fibril is 20–30 nm. (Figure reproduced with permission from Greil, *J. Euro. Ceram Soc.* 21, 105 (2001).) (b) Composition of the nanostructure of spider dragline silk in which β-pleated sheet serves as hard and stiff crystals while α-helix and β turn serve as the soft, amorphous matrix. The characteristic dimension of the β-pleated sheet crystals is several nanometers. (Figure reproduced with permission from Termonia, *Macromolecules* 27 7378 (1994).)

acid residues long. The crystals are believed to be made of hydrophobic polyalanine sequences arranged into hydrogen-bonded β-pleated sheets which run parallel to the fiber axis (Fraser and MacRae, 1973; Hayashi *et al*., 1999; Heslot, 1998; Gosline *et al*., 1999). The amorphous chains, on the other hand, consist of kinetically free oligopeptide chains rich in glycine, such as pentapeptide GPGGX (G denotes glycine; P, proline; X can be alanine, serine, valine or tyrosine) and glycine-rich motif GGX (Termonia, 1994).

6.2.2 Convergent surface nanostructures in nature

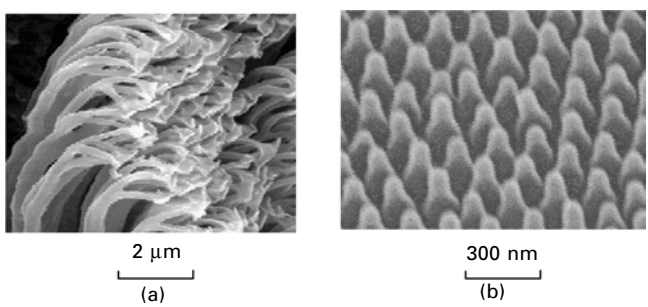
In addition to the bulk nanostructures of bone-like materials which are designed for mechanical support of living organisms, many other nanostructures have been evolved to achieve different mechanical functions. Geckos and many insects have evolved surface nanostructures in their foot hair for adhesion to walls and ceilings (Arzt *et al*., 2002, 2003; Scherge and Gorb, 2001). The surfaces of leaves and of insect bodies exhibit anti-adhesive surface microstructures with trichomes, cuticular folds, wax crystals, etc., presumably for self-cleaning, e.g., Lotus effect (Barthlott, 1990; Barthlott and Neinhuis, 1997; Wagner *et al*., 1996). Water striders and some insects have hydrophobic surface nanostructures on their feet which allow them to walk on water via the surface tension of water (Suter *et al*., 1997; Hu *et al*., 2003).

Attachment structures have independently developed in animal evolution (Breidbach, 1980; Schliemann, 1983). Setose or hairy systems of various animal groups, such as insects, spiders, and geckos, contain surfaces covered by fine patterns of protuberances of different origin. These animals possess the ability to adhere to vertical surfaces and ceilings. Geckos have hundreds of thousands of keratinous hairs or setae on their feet each seta is 30~130 μm long and contains hundreds of protruding nanoscale structures called spatula (Fig. 6.3(a). Possible mechanisms of biological attachment could include mechanical surface interlocking, fluid secretion (capillarity and viscosity) and molecular adhesion (van der Waals interaction). It is only until recently that the developments of MEMS techniques have allowed the adhesive force of geckos to be measured at the level of a single seta (Autumn *et al*., 2000), with evidence that the dominant adhesion mechanism is the van der Waals interaction (Autumn *et al*., 2002). Apparently, the surface nanostructures play a key role in allowing the gecko-like animals to achieve strong adhesion via weak van der Waals forces.

On the other hand, insects and leaves have developed an unwettable surface nanostructure (Wagner *et al*., 1996). In comparison with smooth surfaces, those with a rough microsculpture are found to have a larger contact angle with water droplets (Holloway, 1969a, b; 1994; Wagner *et al*., 1996). Unwettable leaf surfaces have a ‘self-cleaning effect’, due to the fact that water drops running off wash away dirt particles. It was supposed that such

'Lotus effect' is present in all terrestrial organisms exposed to sun, dust and rain (Barthlott, 1990; Barthlott and Neinhuis, 1997; Wagner *et al.*, 1996). Their surfaces usually show very fine microsculptures in the nanometer to submicrometer range (see Fig. 6.3(b)). The influence of the microsculptures on the wettability and contamination of leaf and insect surfaces has been investigated with various methods (Holloway, 1969a, b; 1994; Wagner *et al.*, 1996). The hairy structures on the legs of water striders and fisher spiders can also make the surface highly hydrophobic so that surface tension can supply these insects with sufficient vertical and lateral resistance to facilitate their walking on the surface of water (Denny, 1993; Vogel, 1994; Suter *et al.*, 1997; Hu *et al.*, 2003).

Apparently, biological systems have developed convergent designs of bulk and surface nanostructures with characteristic sizes and geometries through natural evolution. For example, the convergent nanostructure of bone-like materials exhibits a 'soft-hard' combination, with nanoscale hard crystals aligned in a soft matrix. The convergent surface nanostructure of gecko-like animals consists of slender keratinous hairs with diameters in the submicron range. A basic question is why the nanoscale is so important to biological systems. What are the basic mechanisms and principles behind the convergent nanostructure evolution of biological systems? The purpose of this chapter is to examine some of the mechanical bases for the convergent designs of biological nanostructures. We will investigate issues such as flaw tolerance, robustness, strength, stiffness, fracture energy and adhesion energy. We hope that the present study can stimulate broader studies from the scientific community on the mechanics of biological nanostructures.



6.3 The surface nanostructures of geckos and leaves. (a) Geckos have evolved surface nanostructures for adhesion. The toe of a gecko contains a terminal nanostructure called a spatula which is about 200–500 nm in diameter. (Courtesy of Dr S. Gorb of MPI-MF in Germany.) (b) *Tettigia orni* has evolved a surface nanostructure for self-cleaning (Lotus-effect). There are extremely fine regular surface nanostructures (about 100 nm) on the forewing of *Tettigia orni*. (Figure adapted from Wagner *et al., Acta Zoologica* 77, 213 (1996).)

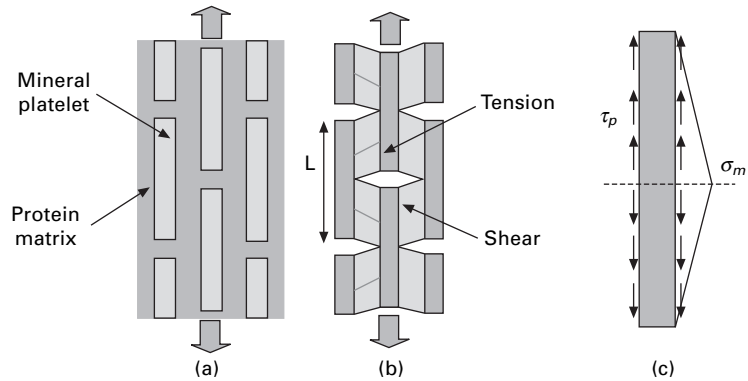
6.3 Mechanics of bulk nanostructures of bone-like materials

6.3.1 Stiffness of the bone nanostructure and a tension–shear chain (TSC) model

Bone-like biological materials are made of mineral and protein, an interesting combination of a hard phase interspersed in a very soft phase. The modulus of protein is usually three orders of magnitude lower than that of mineral. It might seem at first sight that the lack of stiffness of protein would present a severe challenge to bone-like materials which must be hard enough to provide mechanical support for the animals. Nature's solution to this problem can be of great interest to materials scientists. Experimental observations (Kamat *et al*., 2000; Wang *et al*., 2001; Fratzl *et al*., 1997; Tesch *et al*., 2001; Landis, 1995; Landis and Hodgens, 1996) have shown that bone-like materials exhibit a convergent nanostructure consisting of staggered mineral platelets embedded in a soft matrix, as shown in Fig. 6.4(a). Based on the prior work of Jaeger and Fratzl (2000), Gao *et al*. (2003) presented a simple tension–shear chain (TSC) model (Fig. 6.4) to estimate the stiffness of bone. Under an applied tensile stress, the path of load transfer in the mineral–protein composite nanostructure follows a tension–shear chain (Fig. 6.4) where the mineral platelets carry tensile load and the protein transfers load between mineral crystals via shear. In this model, the mineral–protein composite structure is simplified to a network of one-dimensional chains consisting of tensile springs (mineral) interlinked by shear springs (protein). This simple model captures the essential features of load transfer in the protein–mineral biocomposite and provides a useful first-order description of biocomposites. Ji and Gao (2004b) have used the TSC model of Fig. 6.4(b) to estimate the composite stiffness, with results in reasonable agreement with the full finite element analyses of biocomposites. Gao *et al*., (2004) and Ji and Gao (2004b) have further generalized the TSC model to estimate the fracture energy of biocomposites by accounting for nonlinear behaviors of protein deformation.

Kotha *et al*. (2000) previously employed a more complicated shear lag model together with FEM calculations to study stress distribution in mineral and protein matrix. Their results showed that the shear stress distribution along the surface of mineral platelets is approximately constant and the tensile stress in the mineral platelet varies somewhat linearly with the distance from the center of the platelet. The TSC model captures these essential features and simplifies the shear lag model of Kotha *et al*. (2000). In the free body diagram of the TSC model (Fig. 6.4(c)), the tensile stress in a mineral platelet varies perfectly linearly with the distance from the center of the platelet,

$$\tilde{\sigma}_m(x) = \sigma_m \frac{L - 2|x|}{L} \qquad 6.1$$



6.4 A simple tension–shear chain model of biocomposites. (a) A schematic diagram of staggered mineral crystals embedded in a soft (protein) matrix. (b) The path of load transfer in the mineral–protein composites. The mineral platelets carry tensile load and the protein transfers loads between the platelets via shear. (c) The free body diagram of a mineral platelet showing the shear stress applied on the surface of the platelet by protein. Figure adapted from Gao *et al.* (2003).

The maximum stress occurs at the center of the platelet and is related to the shear stress of protein as,

$$\sigma_m = \rho\tau_p, \tag{6.2}$$

where $\rho = L/h$ is the aspect ratio of the platelet. The average stress in the composite is

$$\sigma = \frac{1}{2}\Phi\sigma_m = \frac{1}{2}\rho\Phi\tau_p \tag{6.3}$$

where Φ is the volume concentration of mineral. The average strain in the composite can be related to the deformation of protein and mineral as

$$\varepsilon = \frac{\Delta_m + 2\varepsilon_p h(1-\Phi)/\Phi}{L} \tag{6.4}$$

where Δ_m is the elongation of the platelet, ε_p is the shear strain of protein and L is the length of mineral shown in Fig. 6.4(b). Equations (6.1–6.4) remain valid even to the nonlinear regimes of deformation as long as the shear strain in protein can be regarded as approximately uniform.

The composite stiffness (Young's modulus) E along the longitudinal direction of mineral platelets can be obtained by considering infinitesimal deformations in the composite, in which case

$$\Delta_m = \frac{\sigma_m L}{2E_m}, \quad \varepsilon_p = \frac{\tau_p}{G_p} \tag{6.5}$$

where E_m is the Young's modulus of mineral and G_p is the shear modulus of protein. Inserting eqn 6.5 into eqn 6.4 while making use of eqn 6.3 yields an estimate for the composite stiffness (Gao *et al.*, 2003),

$$\frac{1}{E} = \frac{4(1-\Phi)}{G_p \Phi^2 \rho^2} + \frac{1}{\Phi E_m} \qquad 6.6$$

Ji and Gao (2004b) have verified that eqn 6.6 agrees well with finite element calculations. Generally, G_p can be much smaller (up to three orders of magnitude) than E_m. For small aspect ratios, the second term can be neglected so that the composite stiffness behaves as $E \propto G_p \rho^2$, indicating that the composite stiffness can be rapidly amplified with respect to that of protein via mineral aspect ratio. The simple expression in eqn 6.6 indicates that the large aspect ratio of mineral particles can effectively compensate for the low modulus of protein since it is the combination $\rho^2 G_p$ which contributes to the composite stiffness. In other words, biocomposites are stiffened relative to protein by the square of the aspect ratio of mineral crystals. An aspect ratio of 30:40 would provide a magnification three orders of magnitude over the stiffness of protein and bring the composite stiffness close to that of mineral.

As the mineral aspect ratio increases, the composite stiffness estimated in eqn 6.6 approaches the upper bound defined by the Voigt limit of protein–mineral biocomposites (Ji and Gao, 2004b). At the same time, the Young's modulus in the transverse direction of mineral platelets approaches the lower bound of composite stiffness defined by the Reuss limit. The protein–mineral nanostructure is thus highly anisotropic. Higher-order hierarchical structures must be developed to allow biocomposites to carry complex three-dimensional loads. In the past, studies on bone remodeling and growth have been mostly based on top-down approaches, with macroscopic phenomenological descriptions based on Wolfe's law (Wolff, 1892). It will be interesting in the future to study the growth and remodeling of bone-like materials from a bottom-up approach, with microscopic evolution laws based on the nanoscale mechanical properties and growth mechanisms of bone.

6.3.2 Strength of mineral platelets

In the nanostructure of bone-like materials, the mineral platelets take most of the load acting on the composite structure and the protein transfers load between adjacent mineral platelets via shear deformation, as illustrated by the TSC model in Fig. 6.4. The integrity of the composite chain structure shown in Fig. 6.4(b) is hinged upon the strength of mineral platelets since breaking of the platelets would destroy the critical structural links in the composite, leading to disintegration of the protein–mineral network. It will be discussed later (section 6.3.3) that the strength of mineral platelets plays

a crucial role in the fracture energy of the composite. In order to achieve high fracture energy, the mineral platelets must be able to sustain large tensile stress without fracture, while the protein and the protein–mineral interface must be able to absorb and dissipate a large amount of energy by undergoing large deformation under significant stress (Gao *et al*., 2004).

The mineral platelets in bone must be very strong in order to maintain the structural integrity of protein–mineral networks while allowing the biocomposite to absorb large amount of fracture energy through protein deformation. Gao *et al*. (2003) pointed out that the nanometer scale is the key to optimizing the strength of mineral platelets. A perfect, defect-free mineral crystal should be able to sustain mechanical stress near the theoretical strength σ_{th} of the material. However, the strength of brittle solids like bone mineral is usually determined by pre-existing flaws. It is practically impossible to produce defect-free mineral crystals. Protein molecules trapped within the mineral platelets during the biomineralization process are mechanically equivalent to embedded microcracks and may precipitate brittle fracture. If we consider a thumbnail crack in the mineral particle as shown in Fig. 6.5(a), the strength of this 'cracked' mineral platelet can be calculated from the Griffith criterion of fracture as

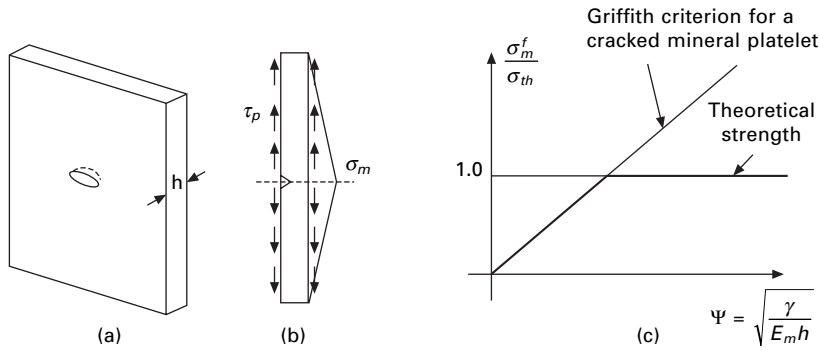
$$\sigma_m^f = \alpha E_m \Psi, \quad \Psi = \sqrt{\frac{\gamma}{E_m h}} \tag{6.7}$$

where γ is the surface energy and h is the thickness of the platelet. The parameter α depends on the crack geometry and is approximately equal to $\sqrt{\pi}$ when the maximum crack depth equals one half of the platelet thickness. Comparing the strength of the cracked crystal with that of a defect-free crystal, Fig. 6.5(c) shows that there exists a transition in failure mode between crack propagation governed by the Griffith criterion and uniform failure near theoretical strength at a critical length scale (Gao *et al*., 2003)

$$h_{cr} \approx \alpha^2 \frac{\gamma E_m}{\sigma_{\text{th}}^2} \tag{6.8}$$

At this length scale, the fracture mechanism changes from energy governed flaw propagation to strength governed homogeneous structure failure.

Below the critical length of eqn 6.8, the fracture strength of a cracked crystal approaches that of a perfect crystal because the Griffith criterion for crack propagation cannot be satisfied until the applied stress reaches the theoretical strength of material. Taking a rough estimate $\gamma = 1$ J/m^2, $E_m = 100$ GPa, and $\sigma_{\text{th}} = E_m/30$, h_{cr} is found to be around 30 nm for a half-cracked platelet. The nanometer scale allows the strength of mineral particles to be optimized near the theoretical strength and also renders these particles insensitive to crack-like defects (Gao *et al*., 2003). For structures larger than



6.5 A length scale for optimized strength in mineral platelet. (a) A schematic diagram of mineral platelet with a surface crack. (b) The load distribution on the cracked mineral crystal. (c) Comparison of the strength of a cracked mineral platelet calculated from the Griffith criterion with the strength of a perfect, defect-free crystal. Figure adapted from Gao *et al.* (2003).

the critical length, conventional engineering concepts apply; the strength is determined by pre-existing flaws and the failure occurs by propagation of flaws under stress concentration. As the structure size drops below the critical length, the strength of a perfect mineral platelet is maintained despite the defects; the failure of structure is governed by the theoretical strength rather than by the Griffith criterion, and the material becomes insensitive to flaws. Based on this argument and further analysis, Gao *et al.* (2003) advanced the hypothesis that the nanometer size of mineral crystals in biocomposites is selected to ensure optimum fracture strength and flaw tolerance. In other words, the nanometer sized mineral crystals are assumed to result from a convergent evolution for mechanical property optimization. So far, both continuum (Gao *et al.*, 2003; Gao and Ji, 2003) and atomistic (Gao *et al.*, 2004) simulations have been performed to investigate the concept of nanoscale flaw tolerance in nominally brittle solids like biomineral particles. These simulations have provided further support for the conjecture of Gao *et al.* (2003) that materials become insensitive to flaws below a critical nanometer length scale.

The bio-inspired length scale for optimum mineral strength

$$\gamma E_m / \sigma_{\text{th}}^2 \tag{6.9}$$

is an intrinsic material parameter which measures the size of fracture process zone in a brittle material. Depending on the material, this can range from a few nanometers to a few tens of nanometers. As the structural size reaches this critical length, materials become insensitive to flaws in the sense that pre-existing cracks no longer propagate in the structure. Instead, the materials fail by homogeneous disintegration near the theoretical strength.

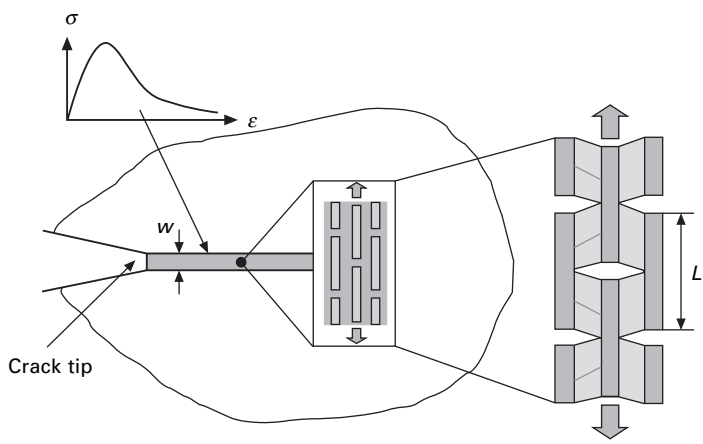
6.3.3 Fracture energy of biocomposites

We have shown that the low toughness problem of mineral crystals can be alleviated by restricting the crystal size to nanoscale. How does the strength of mineral platelets influence the fracture toughness of the protein–mineral composite? Following Gao *et al*. (2004) and Ji and Gao (2004b), we consider a crack propagating in the composite as shown in Fig. 6.6. For simplicity, we assume that dissipation of fracture energy is concentrated within a strip of localized deformation along the prospective crack path (Fig. 6.6), similar to the classical Dugdale model (Dugdale, 1960) of plastic yielding near a crack in a ductile sheet. In this cohesive strip model, the fracture energy can be calculated from the integral (Klein and Gao, 1998; Gao and Ji, 2003)

$$J_c = w \int \sigma(\varepsilon) d\varepsilon \qquad 6.10$$

where w is the width of the localization strip and $\sigma = \sigma(\varepsilon)$ is the continuum cohesive law of the biocomposite. To evaluate the integral in eqn 6.10, we generalize the TSC model of Fig. 6.4(b) to accounting for large deformation of protein within the fracture process zone.

Assuming mineral particles are strong enough to retain their integrity during composite fracture, the width of the fracture localization zone w should be larger than, and proportional to, the length L of mineral particles,



6.6 A Dugdale-type estimate for the fracture energy of biocomposites. The fracture energy dissipation is assumed to concentrate within a strip of localized deformation with width *w*. The tension–shear chain model is used to estimate the fracture energy in the localization strip. The stress and strain relation of material within the localization strip is assumed to obey a cohesive law. Figure adapted from Gao *et al*. (2004).

$$w = \xi L,\ \xi \geq 1 \tag{6.11}$$

Inserting eqns 6.3 and 6.4 into the integral of eqn 6.10 leads to

$$J_c = \xi L \int \sigma d\varepsilon = \frac{1}{2}\Phi\xi \int \sigma_m d\Delta_m + \xi(1-\Phi)L \int \tau_p d\varepsilon_p \tag{6.12}$$

where the first term corresponds to the contribution of mineral and the second term that of protein to the overall fracture energy of the biocomposite. The contribution from the deformation of brittle mineral particles to the composite fracture energy is expected to be small compared to that due to protein deformation. We neglect the first term in eqn 6.12 and write

$$J_c = \xi(1-\Phi)L \int \tau_p d\varepsilon_p \xi(1-\Phi)L\tau_p^* \varepsilon_p^f \tag{6.13}$$

Here, we introduce ε_p^f as the total effective shear strain of protein during fracture and τ_p^* is the associated effective shear stress. The effective strain ε_p^f should include the deformation of protein molecules as their domains unfold as well as slipping along the protein–mineral interface. On the other hand, the effective shear stress in protein can be expressed in the form

$$\tau_p^* = \min\left(\tau_p^f, \tau_{\text{int}}^f, \sigma_m^f/\rho\right) \tag{6.14}$$

where τ_p^f denotes the stress needed to operate domain unfolding, τ_{int}^f is the strength of interface and σ_m^f is the tensile strength of mineral.

Therefore, a simple estimate of the fracture energy of biocomposites is (Gao *et al*., 2004; Ji and Gao, 2004b)

$$J_c = \xi(1-\Phi)L\varepsilon_p^f \min\left(\tau_p^f, \tau_{\text{int}}^f, \sigma_m^f/\rho\right) \tag{6.15}$$

This suggests that the fracture toughness of bio-composites increases with the volume fraction of protein $(1-\Phi)$, the length of the mineral particle L, the effective strain of protein ε_p^f during the fracture process, and the effective stress of protein which is bounded by the stress τ_p^f associated with protein domain unfolding, the protein–mineral interface strength τ_{int}^f, and the mineral strength σ_m^f. Among these parameters, the effect of volume fraction of protein is clear; the more protein, the more volume of materials for energy dissipation. Assuming mineral platelets are strong enough not to break during protein deformation, the platelet length then sets an intrinsic length scale for strain localization near the crack tip. The longer the mineral, the more delocalized the crack-tip deformation and the larger the fracture energy. The total shear strain of protein during the fracture process is a key parameter for fracture energy. The large deformation associated with protein domain unfolding (Smith *et al*., 1999) allows these molecules to endure a very long range of deformation during the fracture process. Slipping along the protein–mineral interface further increases the effective range of strain. The effective shear

stress of protein is determined by the lower bound of protein strength, interfacial strength and mineral strength. In other words, the protein stress is limited by the weakest of three different deformation mechanisms, protein unfolding, interface slipping, and mineral fracture. A key to increase the first two can be linked to Ca^{++} induced sacrificial bonds in bone (Thompson *et al.*, 2001). In an ideal situation, the strengths governing each of the three mechanisms would be identical to each other, i.e.,

$$\tau_p^f = \tau_{\text{int}}^f = \sigma_m^f / \rho \tag{6.16}$$

Assuming the mineral strength obeys the Griffith criterion $\sigma_m^f = \sqrt{\alpha^2 \gamma E_m / h}$, an optimum aspect ratio is implied by eqn 6.16 as (Gao *et al.*, 2003)

$$\rho^* = \frac{\sigma_m^f}{\tau_p^f} = \frac{1}{\tau_p^f}\sqrt{\frac{\alpha^2 \gamma E_m}{h}} \tag{6.17}$$

This optimum aspect ratio of platelets is inversely proportional to the square root of the mineral thickness; the smaller the platelets, the larger the optimal aspect ratio. The mineral crystals in bone have a thickness of the order of a few nanometers and an aspect ratio of 30:40 and those in nacre have thickness of the order of a few hundred nanometers and an aspect ratio around ten, which roughly corresponds to the scaling law predicted by eqn 6.17.

Equations 6.15 and 6.16 also show why it is necessary to have very strong mineral crystals to achieve high fracture energy. Since large aspect ratio $\rho = 30 \sim 40$ is needed to compensate for the low modulus of protein, in order to avoid breakage of platelets during fracture of the composite, the mineral strength must exceed

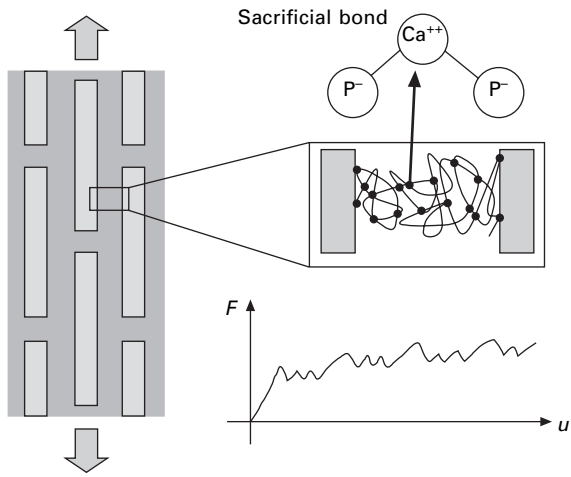
$$\sigma_m^f \geq (30 \sim 40)\,\tau_{\text{int}}^f \tag{6.18}$$

The organic molecules and the inorganic mineral crystals are locally polarized and the interface strength is dominated by electrostatic interactions. Taking τ_{int}^f to be around (20–50) MPa, we can immediately estimate from eqn 6.18 that the mineral strength σ_m^f needs to be of the order of a few GPa, which is near the theoretical strength of the mineral crystal. This explained, from a different perspective, why it is important to have nanometer-sized mineral platelets in biological materials. High strength, flaw tolerant mineral crystals are crucial to operate protein deformation mechanisms and slipping along the protein–mineral interface, which, as shown by eqn 6.15, are the key factors to achieving high fracture energy in the biocomposite.

The molecular design of proteins is ideally suited for absorbing fracture energy. Proteins deform by gradual unfolding of their domain structures. It can take a large amount of deformation before the primary structure of

protein, the peptide backbone, is stretched. In some cases, large deformation alone is not sufficient as it is the area under the stress-strain curve which defines the fracture energy. According to eqn 6.15, a large effective stress in protein is also needed for the high toughness of biocomposites. The usual entropic elasticity of chain polymers involves relatively small effective stress. Cross-linking mechanisms can significantly increase the protein stress and enhance the energy dissipation capability of proteins. For example, an important cross-linking mechanism has been identified in bone, in which Ca^{++} ions cross-link peptides with negative electric charges, forming relatively strong sacrificial bonds with strength up to 30% of the covalent bonds of the peptide backbone (Thompson *et al.*, 2001). The sacrificial bonds bind functional groups along different segments of protein and along the protein–mineral interface, increasing the effective stress in protein and leading to high fracture energy.

The sacrificial bonds not only lead to a large effective stress in protein but also allow protein deformation and interface slipping to occur simultaneously under similar stress levels, resulting in maximum possible shear deformation. Figure 6.7 schematically shows the mechanism of sacrificial bonds in bone.



6.7 A schematic illustration of protein modules deforming between mineral platelets in the biological nanostructure and the inherent force-extension relation of protein with cross-linking mechanism of Ca^{++} formed sacrificial bonds. The sacrificial bonds are formed by Ca^{++} ions linking negatively charged functional groups along the peptide chain. The sacrificial bonds effectively convert the usual entropic elasticity behaviors of biopolymers to one that resembles metal plasticity. The long flat tail and the sawtooth undulation of protein deformation (Smith *et al.*, 1999; Thompson *et al.*, 2001) are due to breaking of sacrificial bonds and protein unfolding in the modules instead of the molecular backbone. Figure adapted from Gao *et al.* (2004).

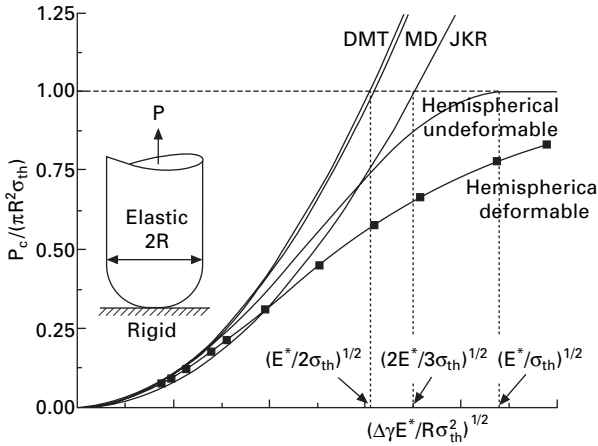
The protein molecules between mineral platelets deform as sacrificial bonds formed by Ca^{++} ions breaking sequentially, resulting in a very long flat tail of the force-extension (F-u) curve with saw-tooth undulation (Thompson *et al.*, 2001). Cross-link mechanisms such as Ca^{++} induced sacrificial bonds, and hydrogen bonds in general, effectively convert entropic elasticity behavior of polymers to one that resembles metal plasticity, enhancing the capability of polymers to absorb fracture energy. The combination of a hard phase providing structural stability and a strongly dissipative soft phase absorbing fracture energy can be used as a guiding principle in designing organic–inorganic composite materials.

6.4 Mechanics of surface nanostructure of gecko-like materials

6.4.1 Johnson–Kendall–Robert (JKR) type contact mechanics models

The density of surface hairs on the attachment pads of biological systems increases with the body weight of the animal, with gecko having the highest hair density among all animal species that have been studied (Scherge and Gorb, 2001). Such hairy or fibrillar attachment systems are being investigated by contact mechanics models (Persson, 2003; Hui *et al.*, 2002). At the single asperity level, dry adhesion between solid objects has been an active topic of contact mechanics since the Johnson–Kendall–Roberts (JKR) model (Johnson *et al.*, 1971). The JKR model has been recently used to show that splitting of a single contact into multiple smaller contacts results in enhanced adhesion strength (Arzt *et al.*, 2002, 2003; Autumn *et al.*, 2002), thus providing a theoretical basis for understanding the hairy attachment system. A paradox of the JKR type models (Johnson *et al.*, 1971; Derjaguin *et al.*, 1975; Maugis, 1992) is their prediction that the spatula structure of gecko can be split *ad infinitum* to support arbitrarily large body weight, which is clearly impossible as the adhesion strength cannot exceed the theoretical strength of van der Waals interaction (Gao *et al.*, 2004, 2005).

The apparent paradox of the JKR type models can be related to linear elastic fracture mechanics which predicts unlimited increase in strength as a structure is reduced in size (Fig. 6.8). To understand this paradox, consider a cylindrical hair in contact with a smooth surface with adhesive interaction forces. The hair is assumed to have a hemispherical tip with diameter $2R$, as shown in the inset of Fig. 6.8. The profile of the hemispherical tip can be described by the function $z = R - \sqrt{R^2 - r^2}$ where z measures the height and r the planar radius of a point on the tip. For simplicity, let us first consider a DMT (Derjaguin *et al.*, 1975) type model in which the elastic deformation of the contacting objects is neglected. Assuming a simple Dugdale



6.8 Prediction of JKR type contact mechanics models for the adhesion strength of an elastic cylinder with a hemispherical tip on a rigid substrate. The classical model of contact mechanics (JKR, MD, DMT) make incorrect predictions at very small sizes. The cohesive solutions for a hemispherical tip are labeled as 'hemispherical deformable' if elastic deformation is taken into account. Unlike the JKR type models, the cohesive solutions show saturation at the theoretical strength of van der Waals interaction at very small sizes. Figure adapted from Gao *et al.* (2005).

type interaction law where the force is equal to σ_{th} within a critical interacting distance $\Delta\gamma/\sigma_{th}$ and zero beyond this distance, $\Delta\gamma$ being the van der Waals interaction energy, the pull-off force P_c could be calculated following Bradley (1932) as

$$P_c = \int_0^R 2\pi r\sigma_0 dr \tag{6.19}$$

The relation between the surface separation z and radius r is given by

$$z = R - \sqrt{R^2 - r^2} \tag{6.20}$$

Therefore

$$(R - z)\, dz = rdz \qquad \left(0 \le z \le \frac{\Delta\gamma}{\sigma_{th}}\right) \tag{6.21}$$

Substituting eqn 6.21 into eqn 6.19 and using the Dugdale interaction law, we immediately obtain

$$\frac{P_c}{\pi R^2\sigma_{th}} = \begin{cases} 2\eta - \eta^2 & \eta < 1 \\ 1 & \eta \ge 1 \end{cases} \tag{6.22}$$

where $\eta = \Delta\gamma/(R\sigma_{th})$. This result is plotted in Fig. 6.8 as the curve labeled 'hemispherical undeformable'. There is clearly a saturation of adhesion strength below a critical size. When the radius of spatula is smaller than $\Delta\gamma/\sigma_{th}$, the contact achieves its maximum strength, which is equal to the theoretical strength of the van der Waals interaction σ_{th} (see Fig. 6.8). The critical size for strength saturation $\Delta\gamma/\sigma_{th}$ is the effective range of van der Waals interaction, which is typically of the order of a few atomic spacings. If elastic deformation is considered in the analysis, an explicit expression for the pull-off force cannot be obtained. We have developed a numerical method to treat a deformable hemispherical tip similar to the previous studies (Greenwood, 1997; Johnson and Greenwood, 1997), but with an important difference in that the hemispherical geometry is now exactly taken into account (Gao *et al*., 2005). Assuming a Lennard–Jones type interaction law (Greenwood, 1997), the pull-off force for a deformable hemisphere-tipped cylinder is calculated and plotted in Fig. 6.8 as the curve labeled 'hemispherical deformable'. The pull-off force also asymptotically approaches the theoretical strength of the van der Waals interaction as the size decreases. The calculations confirmed that the maximum adhesive strength is the theoretical strength of van der Waals interaction. It is interesting that the classical models of JKR (Johnson *et al*., 1971), DMT (Derjaguin *et al*., 1975) and MD (Maugis, 1992) in contact mechanics are unable to capture the process of strength saturation. This has been attributed in part to the approximation of the tip geometry as a parabola adopted in these models (Gao *et al*., 2005).

For the hemispherical tip, the adhesive strength does not reach the maximum strength until the tip size falls below a few atomic spacings. This suggests that a hemispherical tip is a relatively poor design for adhesion. Indeed, broad studies of biological attachment devices (Scherge and Gorb, 2001) revealed that the spherical shape of contact element is a very rare case in nature and usually consists of an extremely compliant material (Spolenak *et al*., 2004).

For a given contact area A, the theoretical pull-off force $\sigma_{th}A$ is the upper limit of adhesion strength. Since the hemispherical shape turns out to be a poor design for adhesion, a question naturally arises: is there an 'optimal' shape that would give rise to the theoretical pull-off force $\sigma_{th}A$. Gao and Yao (2004) addressed this question and showed that it is always possible to design an optimal shape of the tip of a fiber to achieve the theoretical pull-off force. However, it was found that such design tends to be unreliable at the macroscopic scale because the pull-off force is sensitive to small variations in the tip shape of the fiber. A robust design of optimal adhesion becomes possible when the diameter of the fiber is reduced to length scales of the order of 100 nanometers. At this critical size, the pull-off force becomes insensitive to small variations in tip shape. In nature or in engineering, optimal adhesion could be achieved by a combination of size reduction and shape

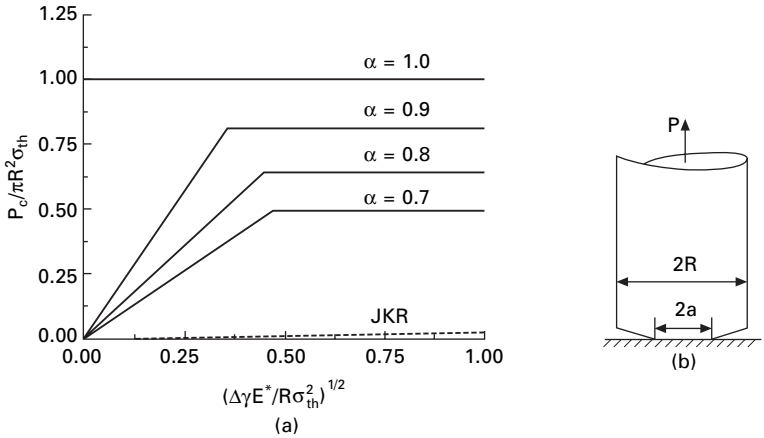
optimization (Gao and Yao, 2004). The smaller the shape, the less important the shape optimization. At the macroscopic scale, it is theoretically possible to achieve theoretical pull-off force but the shape has to be manufactured with sufficiently high precision. The optimal adhesion at nanoscale provides a plausible explanation for the convergent design of surface nanostructures in biology.

Here we consider the idealized case of a cylindrical hair in frictionless contact with a rigid substrate, resembling a soft biological tissue in contact with a hard material. In this case, the optimal shape is simply a flat punch; if there are no defects over the contacting region, the adhesion strength would be identical to the theoretical strength of the van der Waals interaction. In other words, a flat-ended cylindrical hair adhering to a flat rigid substrate would have the theoretical adhesion strength irrespective of the size of contact. The problem with such a design is that the strength would be highly sensitive to small variations in tip geometry (Gao and Yao, 2004) and/or adhesion defects in the contact region (Gao *et al.*, 2004, 2005). For example, surface roughness could result in poor adhesion and crack-like defects within the contact region, and the defects tend to reduce the actual contact area, inducing stress concentration along the edge of the contact region and causing adhesion failure as the crack-like flaws spread. We show in the following that the concept of nanoscale flaw tolerance is also applicable in the design of surface structure of biological systems.

6.4.2 Saturation of adhesion strength of spatula at nanoscale

Why is spatula, the most terminal (smallest) structure of a gecko's attachment mechanism (Fig. 6.3(a)), designed in the nanometer (200–500 nm in diameter) range? This could be understood using the similar idea of nanoscale flaw tolerance as in the bone-like materials. We have modeled the spatula as an elastic flat-ended cylinder with radius R (Gao *et al.*, 2004, 2005). Consider the adhesive contact of such a hair with a rigid substrate. To test the ability of the flat-ended cylindrical hair to adhere in the presence of adhesive flaws, imperfect contact between the spatula and substrate is assumed such that the radius of the actual contact area is $a = \alpha R$, $0 < \alpha < 1$, as shown in Fig. 6.9(b); the outer rim $\alpha R < r < R$ represents flaws or regions of poor adhesion. The strength of such an adhesive joint can be calculated by treating the contact problem as a circumferentially cracked cylinder, in which case the stress field near the edge of the contact area has a square-root singularity with stress intensity factor (Tada *et al.*, 2000)

$$K_I = \frac{P}{\pi a^2}\sqrt{\pi a}\, F_1(\alpha) \qquad 6.23$$



6.9 Adhesion of a flat-ended cylinder to a rigid substrate. The actual contact area is assumed to be smaller than the total area of the punch due to imperfections along the outer rim of the punch. This contact model is elastically equivalent to a cracked cylinder. (a) Variations of the apparent adhesion strength for different actual contact areas according to the Griffith criterion and theoretical strength. The JKR prediction of a hemispherical tip is plotted as a dashed line for comparison. The plot shows that a flat punch induces much larger adhesive forces in comparison with a hemispherical tip, and that the adhesion strength reaches the theoretical strength of van der Waals interaction at a critical contact size. (b) The geometry of a flat punch partially adhering to a substrate. Figure adapted from Gao *et al.* (2004).

where $F_1(\alpha)$ varies in a narrow range between 0.4 and 0.5 for $0 \le \alpha \le 0.8$ ($\alpha = 1$ corresponds to perfect, defect-free contact). Substitute eqn 6.23 into the Griffith condition

$$\frac{K_I^2}{2E^*} = \Delta\gamma \tag{6.24}$$

where $E^* = E/(1 - \nu^2)$ and the factor of 2 is due to the rigid substrate. The apparent adhesion strength normalized by the theoretical strength for adhesion, $\hat{\sigma}_c = P_c/(\sigma_{\text{th}}\pi R^2)$, is found to be

$$\hat{\sigma}_c = \beta\alpha^2\psi \tag{6.25}$$

where

$$\psi = \sqrt{\frac{\Delta\gamma E^*}{R\sigma_{\text{th}}^2}} \tag{6.26}$$

$$\beta = \sqrt{2/(\pi\alpha F_1^2(\alpha))} \tag{6.27}$$

The adhesive strength is a linear function of the dimensionless variable ψ with slope $\beta\alpha^2$. The maximum adhesion strength is achieved when the pull-off force reaches $P_c = \sigma_{th}\pi a^2$, or $\hat{\sigma}_c = \alpha^2$, in which case the traction within the contact area uniformly reaches the theoretical strength σ_{th}. This saturation in strength occurs at a critical size of the contact area

$$R_{cr} = \beta^2 \frac{\Delta\gamma E^*}{\sigma_{th}^2} \tag{6.28}$$

Figure 6.9(a) plots the apparent adhesive strength for $\alpha = 0.7$, 0.8 and 0.9, together with the case of flawless contact ($\alpha = 1$). The corresponding result of a hemispherical tip based on the JKR model is plotted as a dashed line for comparison. (In plotting the JKR curve, we have taken E^*/σ_{th} to be 75). The flat-ended spatula achieves the maximum adhesion strength much more quickly than the hemispherical configuration.

The critical contact size for saturation of adhesion strength can be estimated as follows. Assume the actual contact area is about 50% of the total area available for contact, corresponding to $\alpha \cong 0.7$ Take the parameters for the van der Waals interaction and the Young's modulus of spatula (keratin) as follows:

$$\sigma_{th} = 20 \text{ MPa}, \Delta\gamma = 0.01 \text{ J/m}^2, \frac{\Delta\gamma}{\sigma_{th}} \cong 0.5 \text{ nm}, E^* = 2 \text{ GPa} \tag{6.29}$$

This gives the critical size for adhesive strength saturation as

$$R_{cr} \cong 225 \text{ nm} \tag{6.30}$$

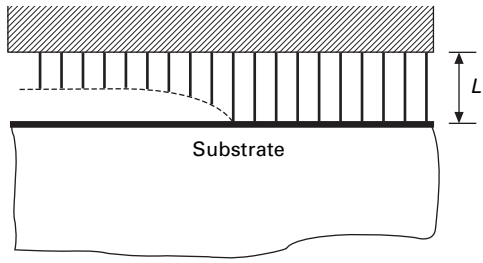
Interestingly, the radius of a gecko's spatula is typically around 100–250 nm. This analysis suggests that the nanometer size of the spatula structure of geckos may have been evolved to achieve optimization of adhesive strength in tolerance of contact flaws.

6.4.3 Adhesion energy of spatula arrays

To relate the adhesion strength of a single spatula to the adhesion energy of an array of spatulas, consider a fibrillar structure in adhesive contact with a substrate, as shown in Fig. 6.10. The Young's modulus and adhesion strength of a single spatula being E and σ_c, the adhesion energy of the spatula array analogous to eqn 6.15 is

$$J_c = \varphi\Delta\gamma + L\int \sigma\varphi d\varepsilon = \varphi\Delta\gamma + \frac{L\sigma_c^2\varphi}{2E} \tag{6.31}$$

where the first term $\Delta\gamma$is the energy from interface separation and the second term represents the elastic energy stored in the fiber array which is lost



6.10 Detachment of an array of adhesive fibers from a substrate. The adhesion energy of the array can be estimated by calculating energy absorbed in the layer as individual hairs detach from the substrate. Figure adapted from Gao *et al.* (2004).

through irreversible heat dissipation after the detachment; L is the length of fiber and φ is the areal density of contact regions. Thus the adhesion energy of spatula structure is proportional to the square of the adhesion strength of a single spatula. It is thus strongly desirable to reduce the size of spatula near the critical size so as to maximize the adhesion strength and to increase the adhesion energy of an array. In the case of a Dugdale-like interaction law, the energy from interface separation is only $\varphi\Delta\gamma = \sigma_{th}\varphi\delta$ where δ is the range of van der Waal interaction typically of the order of 1 nm. Assuming that the adhesion strength of individual fibers is optimized by restricting the fiber size to nanoscale, $\sigma_c \approx \sigma_{th}$, the ratio between the adhesion energy of the fiber array and work of separation is

$$\frac{J_c}{\varphi\Delta\gamma} = 1 + \frac{\sigma_{th}}{2E}\frac{L}{\delta} \qquad 6.32$$

We see that the adhesion energy between the fiber array and the substrate can be magnified by the ratio L/δ. If L is selected to be much larger than $(2E/\sigma_{th})\delta$ (estimated to be around 50–100 nm), the elastic energy stored in the fibers can contribute significantly to the total adhesion energy of the array. It is therefore possible to enhance adhesion by increasing the aspect ratio of the fiber. On the other hand, the slender geometry of the fiber may induce self-bunching among neighbouring fibers (Hui *et al.*, 2002; Persson, 2003; Gao *et al.*, 2004, 2005). A solution to the problem of self-bunching without diminishing adhesion energy is to adopt a tree-like design for the seta (where the spatulas are located), with fine, branched tips to achieve optimal adhesion strength and thick base trunks to counter self-bunching. Such a design, as observed in the case for gecko, allows the fibers to maintain a large aspect ratio in order to store large amounts of elastic energy to enhance adhesion.

Note that the mechanism of fracture energy dissipation in the spatula array is very different from that in the protein–mineral biocomposite. In the latter case, fracture energy is predominantly dissipated through shear

deformation in protein and slipping along the protein–mineral interface. Optimization of mineral platelets near the theoretical strength plays a crucial role in allowing a significant stress to be built up in protein via cross-link mechanisms such as Ca^{++} induced sacrificial bonds in bone. In the case of the spatula structure, the adhesion strength of individual spatulas directly affects the adhesion energy of the array. The energy dissipation during the detachment of spatulas consists of not only the van der Waals work of adhesion but also the elastic energy stored in the nanofibers upon their detachment from the substrate. In this process, most of the elastic energy stored in the fiber cannot be restored and is instead dissipated as heat. It is interesting to observe that, in both systems, strength optimization is achieved by restricting the characteristic dimension of the basic structure components to nanometer scale so that crack-like flaws do not propagate to break the desired structural link; at the same time, the size in other dimensions is enlarged to achieve a large aspect ratio to enhance the fracture energy in the case of protein-mineral composite or adhesion energy in the case of spatula array.

6.4.4 Anti-bunching condition of the spatula structure

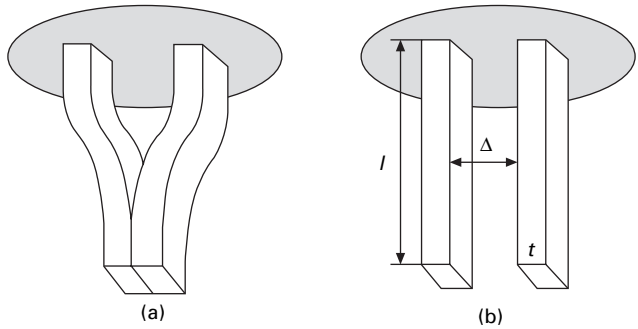
Lateral van der Waals interaction may cause structural instability, and results in clustering or bundling of adjacent fibers due to their relatively slender geometry with large aspect ratios. Lateral stability of fibers against bundling is a necessary condition for their viability as an adhesion structure. Recent study (Geim *et al*., 2003) has demonstrated that bunching leads to a reduction of adhesive strength in the microfabricated artificial gecko structure made of polyimide micro-hairs. To gain some insight into this issue, we have modeled the fiber as a cantilever beam with square cross-section (Gao *et al*., 2005). The configuration of a pair of adjacent fibers in the free-standing state is shown in Fig. 6.11(b), and that in the state of clustering is shown in Fig. 6.11(a). Stability against bunching is interpreted as such that the bundled state of Fig. 6.11(a) should be unstable and will spontaneously detach if formed. This problem can be treated as a crack problem with detachment condition (Tada *et al*., 1985; Hui *et al*., 2002; Gao *et al*., 2005)

$$E^* \geq \frac{8\gamma l^4}{3\Delta^2 t^3} \tag{6.32}$$

where γ is the surface energy of spatula; the geometrical parameters l, Δ, t are shown in Fig. 6.11(b).

A different condition of bunching instability has been derived by Persson (2003) who compared the energy of fibers bent over a curvature to that of straight fibers. As a rough estimate we may assume

$$\Delta = 0.6\ \mu m,\ t = 0.2\ \mu m,\ l = 2\ \mu m,\ \gamma = 0.01 J/m^2 \tag{6.33}$$



6.11 Lateral bundling of two adjacent elastic protrusions. (a) The free standing and (b) bundled configurations of two adjacent protrusions. The stability of an array of these protrusions is defined such that the bundled configuration would spontaneously separate into the free standing configuration. Figure adapted from Gao *et al.* (2005).

The stability condition, eqn 6.32, suggests a minimum Young's modulus of 0.15 GPa. Biological materials such as keratin have Young's modulus of the order of a few GPa and thus meet the stability condition. For a hemispherical tip, the critical size to achieve theoretical strength is of the order of $t = 1$ nm. At this length scale, the stability condition requires a minimum Young's modulus of the order of 10^5 GPa, which is impossible to meet. The advantage of the optimal shape design of fiber over the hemispherical design is thus twofold. The optimal shape (a flat-ended punch in the case of an elastic cylinder in frictionless contact with a rigid substrate) not only allows the system to maximize adhesion strength and fracture energy more efficiently, but it is also required from the point of view of structural stability.

6.5 Conclusions

In this chapter, we have reviewed some typical bulk and surface nanostructures of biological systems which have been developed over the course of natural evolution for the purpose of providing living organisms with various mechanical functions/properties that are important for their survival. We have developed the concept of flaw tolerance of bone-like bulk nanostructures and gecko-like surface nanostructures. The chapter is intended to provide some unified viewpoints on how mechanical forces may have played an overarching role in the evolution of various convergent designs of biological structures. It is noteworthy that biological systems exhibit extremely simple and elegant nanostructures, although on larger scales they are really very complex hierarchical systems involving significant differences in material constituents, geometries and functions.

A tension–shear chain (TSC) model was introduced to capture the essential features of load transfer and deformation in the protein–mineral nanostructure

of the bone-like materials. This simple model has been particularly helpful in identifying important parameters responsible for the superior mechanical properties of bone. The TSC model has been used to estimate the stiffness of the protein–mineral nanostructure where the large aspect ratio of mineral platelets is shown to compensate for the low modulus of protein, allowing the biological materials to achieve a stiffness close to that of mineral. In the surface nanostructure of gecko, the large aspect ratio of hairy protrusions seems to be critical in allowing them to achieve conformal contact with a rough surface and to achieve high adhesion energy via dissipation of irreversible elastic energy stored in the spatulas upon detachment from a substrate. On the other hand, the slender geometry of hairs could render them vulnerable to self-bunching among the neighboring hairs under van der Waals interactions. A compromise is to adopt a tree-like structure for the attachment terminals with thick trunk and branched tips to achieve enhanced adhesion while avoiding self-bunching.

The selection of nanometer scale is shown to be the key for mineral platelets in bone-like materials (Gao *et al*., 2003) to achieve optimum strength in the presence of pre-existing cracks, and for the spatula structure of gecko to achieve optimum adhesion strength in the presence of adhesion flaws (Gao *et al*., 2004, 2005). At a nanometer critical length determined by fracture energy, Young's modulus and theoretical strength, the mineral platelets become insensitive to pre-existing crack-like flaws and their strength is maintained near the theoretical strength of the material despite defects. Following the same principle, the nanometer size of spatulas, the most terminal structure of gecko, allows them to achieve maximum adhesion strength and become insensitive to variations in the tip shape (Gao and Yao, 2004) and crack-like adhesion flaws in the contact regions (Gao *et al*., 2004, 2005).

The fracture energy of bone-like materials is shown to be a combined property of ultra strong slender mineral platelets interspersed among tough protein molecules which can undergo extraordinarily large deformation via domain unfolding and interface slipping under a large effective shear stress. Cross-link mechanisms such as Ca^{++} induced sacrificial bonds in bone increase the shear stress in protein and along the protein–mineral interface, significantly enhancing the energy absorbing capability of the biocomposites. Mechanisms like this not only build up a large effective stress in protein but also allow protein deformation and interface slipping to occur simultaneously under similar stress levels, making it possible to engineer a long range of deformation to maximize energy absorption. Optimization of mineral platelets near theoretical strength allows a large aspect ratio and a large effective stress to be built up in protein via cross-link mechanisms. In this way, the mineral platelets provide the required stiffness for mechanical support of biological systems and the protein matrix is put in charge of absorbing and dissipating fracture energy. In the case of detachment of hairy structures of gecko from

substrates, the adhesion energy can be significantly enhanced by elastic instabilities during the dynamic process of detachment in which the elastic energy stored in the hairs is largely dissipated as heat to the environment. Significant energy dissipation can be achieved by enlarging the length (i.e. aspect ratio) of the hairy structure. In both bone and gecko systems, the strength optimization is achieved by restricting the characteristic dimension of the basic structure components to nanometer scale so that crack-like flaws do not propagate to break the desired structural link, and large aspect ratios have been adopted to enhance fracture or adhesion energy, as shown by eqns 6.15 and 6.31. The anisotropic materials designs seem to endow the biological systems with major advantages in achieving optimized stiffness, strength and toughness.

The protein–mineral structure of biocomposites is consistent with the ancient Chinese philosophy that a combination of 'yin' and 'yang', things of complementary nature or properties, results in perfection and harmony in nature. In biological materials, one may identify many complementary properties such as: stiff versus soft, hard versus gentle, brittle versus ductile, strong versus weak, long versus short, tension versus shear, dissipative versus non-dissipative, etc. In biological materials, the mineral platelets are the'yang' phase which provides the stability and stiffness of the structure with the 'yang' characters like stiff, hard, brittle, non-dissipative and non-yielding. The protein acts as the 'yin' phase with characteristics like soft, gentle, ductile, dissipative, flexible, and plays the main role of dissipating fracture energy.

6.6 Future trends

The engineering principles of biological systems can play an important role in helping address some of the major challenges in materials science and engineering in the twenty-first century. The development of nanotechnology will eventually open up enormous possibilities in developing multi-functional and hierarchical materials systems through nanostructure control. What we have discussed in this chapter regarding the bulk and surface nanostructures of biological systems is only a small step in this grand endeavor. Convergent designs in nature provide fertile opportunities for scientific research in the coming decades. In some cases, convergent evolutions occur when genetically unrelated biological systems have adopted the same strategies to achieve the same objectives or properties or functions. In other cases, genes associated with a given set of properties or functions of the same biological system may be found to be strongly conserved over the history of evolution. Whenever convergent evolutions occur, we may look for hints on property optimization. We have discussed stiffness, strength, toughness and adhesion in connection with the principle of nanoscale flaw tolerance. Similar investigations may be

conducted with respect to other properties including friction, fatigue, corrosion and hydrophobicity. There is still a long way to go before the complexities and basic principles of hierarchical structures of biological systems and their associated functions are fully understood.

Understanding the optimization in biological systems requires systematic studies on how biological structures adapt to internal and external forces. A strong interdisciplinary approach is needed to extract the basic concepts and principles from the apparent complexity of nature. The investigations in this chapter underline the importance of a coherent study of mechanics and chemistry of biological systems. Nature does not distinguish between material and structure. Through the examples of biological nanostructures, we wish to emphasize that mechanics (structure) and chemistry (material) can be used with equal importance to achieve optimized materials properties in nanoscale materials engineering.

6.7 Acknowledgement

We are grateful to Dr Stanislav Gorb of the Max Planck Institute for Metals Research for helpful comments during the preparation of this manuscript. Support for this work has been provided by the Max Planck Society, the National Science Foundation of China and the Chang Jiang Scholar program through Tsinghua University.

6.8 References

Arzt E, Enders S and Gorb S (2002), Towards a micromechanical understanding of biological surface devices, *Z. Metallk.*, 93, 345–351.

Arzt E, Gorb S and Spolenak R (2003), From micro to nano contacts in biological attachment devices, *Proc. Natl. Acad. Sci*. USA, 100, 10603–10606.

Autumn K, Liang Y A, Hsieh S T, Zesch W, Chan W T Kenny T W, Fearing R and Full R J (2000), Adhesive force of a single gecko foot-hair, *Nature*, 405, 681–685.

Autumn K, Sitti M, Liang Y A, Peattie A M, Hansen W R, Sponberg S, Kenny T W, Fearing R, Israelachvili J N and Full R J (2002), Evidence for van der Walls adhesion in gecko setae, *Proc. Natl. Acad. Sci. USA*, 99, 12252–12256.

Barthlott W (1990), Scanning electron microscopy of the epidermal surface in plants. In: Claugher D (ed.) *Scanning electron microscopy in taxonomy and functional morphology*. Clarendon Press, Oxford, pp. 69–94.

Barthlott W and Neinhuis C (1997), Purity of the sacred lotus, or escape from contamination in biological surfaces. *Planta* 202, 1–8.

Bradley R S (1932), The cohesive force between solid surfaces and the surface energy of solids. *Phil. Mag*. 13, 853–862.

Breidbach O (1980), Die Tarsen von Insekten – ein schönes Beispiel von Konvergenz, *Mikrokosmos*, 69, 200–201.

Brett D and Waldron K (1981), *Physiology and biochemistry of plant cell walls*, Chapman & Hall, London.

Currey J D (1977), Mechanical properties of mother of pearl in tension, *Proc. R. Soc. London B*, 196, 443–463.

Currey J D and Taylor J D (1974), The mechanical behavior of some molluscan hard tissues, *J. Zool., Lond*., 173, 395–406.

Denny M W (1993), *Air and water: The biological and physics of life's media*, Princeton: Princeton University Press.

Derjaguin B V, Muller V M and Toporov Y P (1975), Effect of contact deformations on the adhesion of particles, *J. Colloid Interface Sci*, 53, 314–326.

Dugdale D S (1960), Yielding of steel sheets containing slits, *J. Mech. Phys. Solids*, 8, 100–104.

Fengel, D and Wegener G (1984), *Wood chemistry, ultrastructure, reaction*, Walter de Gruter, Berlin.

Fraser R D B and MacRae T P (1973), *Conformation in fibrous proteins*, Academic press, New York.

Fratzl P, Jakob H F, Rinnerthaler S, Roschger P and Klaushofer K (1997), Position resolved small-angle X-ray scattering of complex biological materials, *J. appl. Cryst*., 30, 765–769.

Gao H and Ji B (2003), Modeling fracture in nano-materials via a virtual internal bond method, *Engng. Frac. Mech*., 70, 1777–1791.

Gao H and Yao H (2004), Shape insensitive optimal adhesion of nanoscale fibrillar structures, *Proc. Natl. Acad. Sci. USA*, 101, 7851–7856.

Gao H, Ji B, Jaeger I L, Arzt E and Fratzl P (2003), Materials become insensitive to flaws at nanoscale: lessons from nature, *Proc. Natl. Acad. Sci. USA*, 100, 5597–5600.

Gao H, Ji B, Buehler M J and Yao H (2004), *Flaw tolerant bulk and surface nanostructures of biological systems, Mechanics and Chemistry of Biosystems*, 1, 37–52.

Gao H, Wang X, Yao H, Gorb S and Arzt E (2005), Mechanics of hierarchical adhesion structure of gecko, *Mech. Mater*., 37, 275–285.

Geim A K, Dubonos S V, Grigorieva I V, Novoselov K S and Zhukov A A (2003), Microfabricated adhesive mimicking gecko foot-hair, *Nature Mat*., 2, 461–463.

Gosline J M, Guerette P A, Ortlepp C S and Savage K N (1999), The mechanical design of spider silks: from fibroin sequence to mechanical function, *J. Exp. Biol*., 202, 3295–3303.

Greenwood J A (1997), Adhesion of elastic spheres. *Proc. R. Soc. Lond. A* 453, 1277–1297.

Greil P (2001), Biomorphous ceramics from lignocellulosics, *J. Euro. Ceram. Soc*. 21, 105–118.

Griffith A A (1921), The phenomena of rupture and flow in solids, *Phil. Trans. R. Soc. London A*, 221, 163–198.

Hayashi C Y, Shipley N H and Lewis R V (1999), Hypotheses that correlate the sequence, structure, and mechanical properties of spider silk proteins, *Int. J. Biol. Macrom*., 24, 271–275.

Heslot H (1998), Artificial fibrous proteins: A review, *Biochimie*, 80, 19–31.

Holloway P J (1969a), Effects of superficial wax on leaf wettability, *Ann. Appl. Biol*. 63, 145–153.

Holloway P J (1969b), Chemistry of leaf waxes in relation to wetting, *J. Sci. Food Agr*. 20, 124–128.

Hu D L, Chan B and Bush J W M (2003), Hydrodynamics of water strider locomotion, *Nature*, 424, 663–666.

Hui C Y, Jagota A, Lin Y Y and Kramer E J (2002), Constraints on Microcontact printing imposed by stamp deformation, *Langmuir*, 18, 1394–1407.

Jackson A P, Vincent J F V and Turner R M (1988), The mechanical design of nacre, *Proc. R. Soc. London B*, 234, 415–440.

Jaeger I and Fratzl P (2000), Mineralized collagen fibrils: A mechanical model with a staggered arrangement of mineral particles, *Biophys. J.*, 79, 1737–1746.

Ji B and Gao H (2004a), A study of fracture mechanisms in biological nano-composites via the virtual internal bond model, *Mater. Sci. Engng. A*, 366, 96–103.

Ji B and Gao H (2004b), Mechanical properties of a generic nanostructure of biomaterials, *J. Mech. Phys. Solids*, 52, 1963–1990.

Johnson K L and Greenwood J A (1997), An adhesion map for the contact of elastic spheres. *J. Colloid Interface Sci.* 192, 326–333.

Johnson K L, Kendall K and Roberts A D (1971), Surface energy and the contact of elastic solids, *Proc. R. Soc. London A*, 324, 301–313.

Kamat S, Su X, Ballarini R and Heuer A H (2000), Structural basis for the fracture toughness of the shell of the conch strombus gigas, *Nature*, 405, 1036–1040.

Klein P and Gao H (1998), Crack nucleation and growth as strain localization in a virtual-bond continuum, *Eng. Fract. Mech.*, 61, 21–48.

Kotha S P, Kotha S and Guzelso N (2000), A shear-lag model to account for interaction effects between inclusions in composites reinforced with rectangular platelets, *Com Sci. Tech*, 60, 2147–2158.

Landis W J (1995), The strength of a calcified tissue depends in part on the molecular structure and organization of its constituent mineral crystals in their organic matrix, *Bone*, 16, 533–544.

Landis W J and Hodgens K J (1996), Mineralization of collagen may occur on fibril surfaces: evidence from conventional and light-voltage electron microscopy and three-dimensional imaging, *J. Struct. Biol.*, 117, 24–35.

Maugis D (1992), Adhesion of spheres: the JKR-DMT transition using a Dugdale model. *J. Colloid Interface Sci*, 150, 243–269.

Menig R, Meyers M H, Meyers M A and Vecchio K S (2000), Quasi-static and dynamic mechanical response of haliotis rufescens (abalone) shells, *Acta Mater.*, 48, 2383–2398.

Menig R, Meyers M H, Meyers M A and Vecchio K S (2001), Quasi-static and dynamic mechanical response of strombus gigas (conch) shells, *Mater. Sci. Engng. A*, 297, 203–211.

Okumura K and de Gennes P -G (2001), Why is nacre strong? Elastic theory and fracture mechanics for biocomposites with stratified structures, *Eur. Phys. J., E* 4, 121–127.

Persson B N J (2003), On the mechanism of adhesion in biological systems, *J. Chem. Phys.*, 118, 7614–7621.

Rho J Y, Kuhn-Spearing L and Zioupos P (1998), Mechanical properties and the hierarchical structure of bone, *Med. Engng. Phys.*, 20, 92–102.

Roschger P, Grabner B M, Rinnerthaler S, Tesch W, Kneissel M, Berzlanovich A, Klaushofer K and Fratzl P (2001), Structural development of the mineralized tissue in the human L4 vertebral body, *J. Struct. Biol.*, 136, 126–136.

Scherge M and Gorb S N (2001), *Biological Micro and Nano-Tribology*, Springer-Verlag, New York.

Schliemann H (1983), Adhesive organs – Frequently occurring convergences, *Funkt. Biol. M*ed., 2, 169–177.

Smith B L, Schaeffer T E, Viani M, Thompson J B, Frederick N A, Kindt J, Belcher A, Stucky G D, Morse D E and Hansma P K (1999), Molecular mechanistic origin of the toughness of natural adhesive, fibres and composites, *Nature*, 399, 761–763.

Spolenak R, Gorb S, Gao H and Arzt E (2004), Effects for contact shape on biological attachments, *Proceedings of the Royal Society*, in press.

Suter R B, Rosenberg O, Loeb S, Wildman H and Long J H (1997), Locomotion on the water surface: propulsive mechanism of the fisher spider, *Dolomedes Triton, J. Exp. Biol.*, 200, 2523–2538.

Tada H, Paris P C and Irwin G R (2000), *The stress analysis of cracks handbook*. ASME Press, New York.

Taylor J D (1973), The structural evolution of the bivalve shell, *Palaeontology*, 16, 519–534.

Temonia Y (1994), Molecular modelling of spider silk elasticity, *Macromolecules*, 27, 7378–7381.

Tesch W, Eidelman N, Roschger P, Goldenberg F, Klaushofer K and Fratzl P (2001), Graded microstructure and mechanical properties of human crown dentin, *Calcif. Tissue Int.*, 69, 147–157.

Thompson J B, Kindt J H, Drake B, Hansma H G, Morse D E and Hansma P K (2001), Bone indentation recovery time correlates with bond reforming time, *Nature*, 414, 773–776.

Wagner T, Neinhuis C and Barthlott W (1996), Wettability and contaminability of insect wings as a function of their surface sculptures, *Acta Zoologica* (Stockholm), 77, 213–225.

Wang R Z, Suo Z, Evans A G, Yao N and Aksay I A (2001), Deformation mechanisms in nacre, *J. Mater. Res.*, 16, 2485–2493.

Warshawsky H (1989), Organization of Crystals in Enamel. *Anat. Rec*. 224, 242–262.

Weiner S and Wagner H D (1998), The material bone: Structure-mechanical function relations, *Ann. Rev. Mater. Sci.*, 28, 271–298.

Weiner S, Veis A, Beniash E, Arad T, Dillon J W, Sabsay B and Siddiqui F (1999), Peritubular dentin formation: crystal organization and the macromolecular constituents in human teeth, *J. Struct. Biol.*, 126, 27–41.

Wolff J (1892), *Das Gesetz der Transformation der nochen*. Berlin: A Hirchwild, *The law of bone remodeling*. Translated by Maquel P and Furlong R, Berlin: Springer, 1986.

Vogel S (1994), *Life in moving fluids*, Princeton: Princeton University Press.

7

Mechanical behavior of metallic nanolaminates

A MISRA, Los Alamos National Laboratory, USA

7.1 Introduction

Nanolaminate materials, also referred to as 'superlattices' or 'multilayers', represent a class of composite materials that are made up of alternating nanometer-scale layers of two different materials, where the individual layer thickness may vary from a few atomic layers to a few tens of nanometers. Nanolaminates have been the subject of significant recent research worldwide due to the novel mechanical and physical properties that emerge as the individual layer thickness is reduced to nanometer-scale [1–4]. For example, a nanolaminate comprising two soft metals may possess several GPa level strengths at bilayer periods of a few nanometers, with the peak strengths within a factor of two to three of a lower bound estimate of the theoretical strength ($\approx E/25$ where E is the Young's modulus) [5]. In contrast, the constituent pure metals in the bulk form may have yield strengths in the range of a few tens of MPa. In addition to mechanical properties, novel electronic, magnetic and optical behaviors also result from nano-layering, making such materials attractive for a number of potential applications such as hard, wear resistance coatings, X-ray optics, thin film magnetic recording media, microelectromechanical systems (MEMS) and self-supported high-strength foils for a variety of structural components. While nanolaminates may be metal–metal, metal–intermetallic, metal–ceramic, ceramic–ceramic or crystalline amorphous, the focus of this chapter will be on mechanical behavior of metallic nanolaminates. Besides the technological applications, metal–metal systems are studied extensively as model systems for fundamental research on the effects of nanometer length scales and the interface structures on the mechanical properties of nanolaminates.

The chapter starts with a brief description of the commonly used methods of synthesis. A brief overview of the strengthening mechanisms relevant to nanolaminates is presented next. Experimental data on the strength as a function of bilayer period of the nanolaminates and the interpretation of these data in terms of the dislocation pile-up based Hall–Petch model is

presented next. Single dislocation-based deformation models are developed to account for the increasing strength with decreasing layer thickness at length scales where the continuum Hall–Petch model is not applicable. Atomistic modeling that predicts the limiting strength value for these nanolaminates is described next. The deformation behavior of nanolaminates subjected to large plastic strains is discussed next to highlight the dislocation storage, work hardening and texture evolution in these materials. The chapter concludes with a summary and comments on future trends in this topic.

7.2 Methods of synthesizing metallic nanolaminates

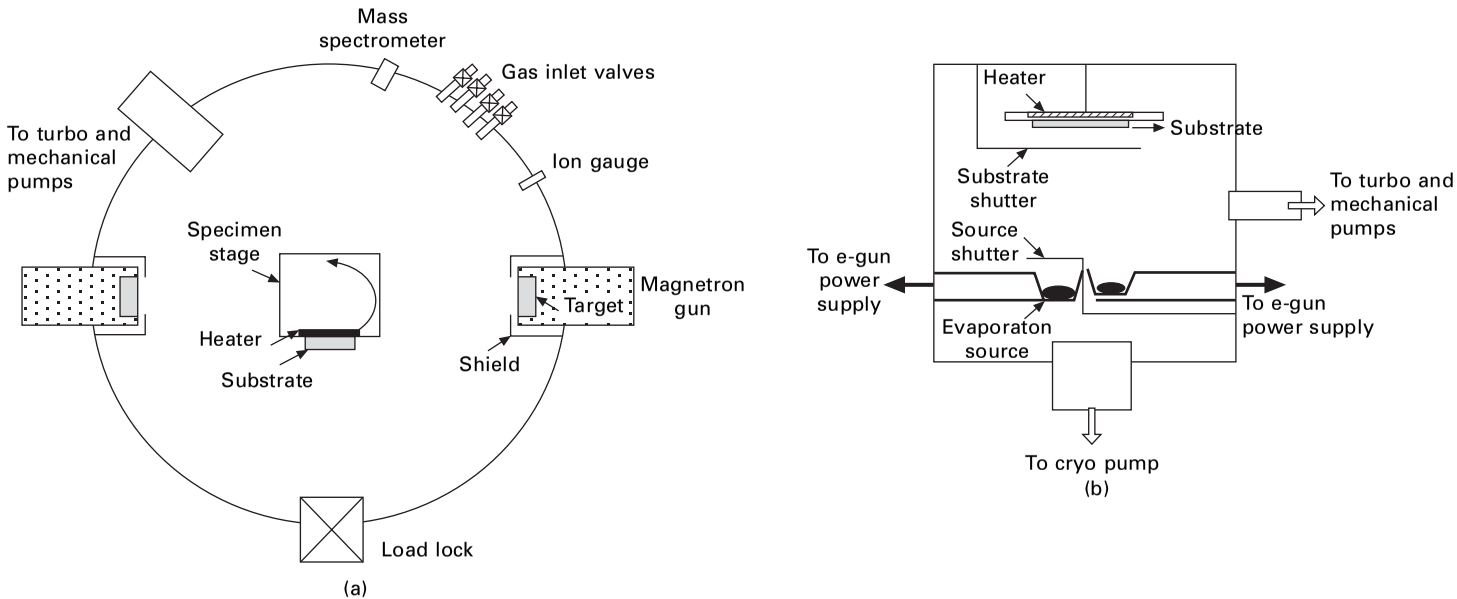
There are two commonly used synthesis methods that are discussed here. First, is a 'top-down' type approach involving severe plastic deformation (SPD) of bulk sheets [6–10]. Conceptually the approach is similar to the SPD techniques such as equal-channel angular pressing or high-pressure torsion employed to synthesize ultra-fine grained single-phase metals or repeated wire drawing to make composite wires. Shingu and co-workers [6, 7] used repeated hot pressing–cold rolling–annealing cycles to produce bulk nanolaminates starting with tens of micrometers thick individual metal sheets. A typical processing cycle employed for Fe/Ag multilayers is described below. A hundred sheets (13 mm × 18 mm) each of 50 μm thick Fe and 30 μm thick Ag were stacked together and pressed in a controlled temperature and atmosphere to bond the metal sheets. The bonded stack was then rolled in ambient atmosphere to reduce the thickness of the layers. The rolled sheets were then annealed at 873 K to reduce the hardness, while maintaining the layered microstructure, before the next pressing–rolling cycle. After the second pressing–rolling step, the average Ag layer thickness was 20 nm. Since annealing after cold rolling may have lead to some coarsening, eliminating the annealing step resulted in an average Ag layer thickness of 4 nm. Similar results were shown for Fe-Cu multilayers as well. This press-rolling synthesis approach could effectively reduce the metal layer thickness from tens of micrometers to a few nanometers (a factor of 10^4), although the resulting nanolaminates did not have very uniform inter-lamellar spacing.

The second synthesis approach that is commonly used is 'bottom up', i.e., via deposition. This may involve electro-deposition or physical vapor deposition (PVD) techniques such as magnetron sputtering or electron beam evaporation. Electrodeposition synthesis of nanolaminates is accomplished by using an electrolyte from which two different metals can be deposited by varying the applied voltage [11–13]. For example, Cu-Ni nanolaminates with a total foil thickness of 50 μm were electrodeposited from a nickel sulfamate bath containing Cu. Layered deposits were formed by periodically switching between two sets of deposition conditions under computer control: positive cathode

potentials and high rotation rates produce pure Cu layers while negative cathode potentials and lower rotation rates produced almost pure Ni layers [12]. The constraint of finding an electrolyte from which two different metals can be electro-deposited limits the number of nanolaminate systems that can be synthesized by electrodeposition. Magnetron sputtering is widely used to synthesize nanolaminates [14–16]. Different designs are possible, the one used in the author's lab is shown schematically in Fig. 7.1(a) (for simplicity, only the key elements are shown). The base pressure in the chamber is typically $\sim 1 \times 10^{-8}$ torr. The specimens are loaded through a load lock onto a specimen stage that contains a substrate heater. A few millitorr pressure of argon is introduced into the chamber and d.c. power of a few hundred volts is used to ignite the plasma at either the left or the right gun. A stepper motor is used to rotate the sample stage from facing the load lock (for loading/unloading the samples) to facing either the left or the right gun for deposition.

A typical deposition sequence would be as follows: turn on one gun, rotate sample stage towards the gun for the desired time, turn gun off, power on the other gun, rotate sample again to face the gun that is on, turn gun off. The entire computer-controlled cycle is repeated to produce the desired number of bilayers. The typical deposition rates vary from a few to a few tens of Å/s. Thousands of bilayers can be easily deposited to produce tens of micrometer thick samples in a few hours. To make nitride layers, reactive sputtering is used with a Ar/N_2 gas mixture while sputtering from the metallic target. For insulating materials, r.f. power instead of d.c. is applied to the target. Although the substrate is usually kept at ground potential, a negative d.c. bias can be applied to the substrate to induce Ar^+ bombardment of the growing film. A limitation of the design shown in Fig. 7.1(a) is that co-sputtering cannot be done, i.e., materials from both left and right guns cannot be simultaneously deposited on the substrate.

A different design is shown in Fig. 7.1(b). In the author's lab, this is the design used for electron beam evaporators [17] although it can be used in magnetron sputtering chambers as well. Here the two guns sit side-by-side while the sample faces down on the guns. To synthesize a multilayer film, both electron guns are kept on simultaneously, at the respective deposition rates, and through a computer-controlled stepper motor, the shutters are alternated to sequentially deposit the two layers. The substrate can be kept at ambient temperature or heated. Unlike magnetron sputtering that involves a certain level of energetic particle bombardment either via negative substrate bias or reflected Ar neutrals or the energy of the depositing atoms (especially at low Ar partial pressures), the depositing flux in evaporation is essentially thermalized. For substrates at ambient temperature, non-energetic e-beam evaporation may lead to films that are less than fully dense, especially for high melting materials. A separate ion-gun added in the evaporator (not



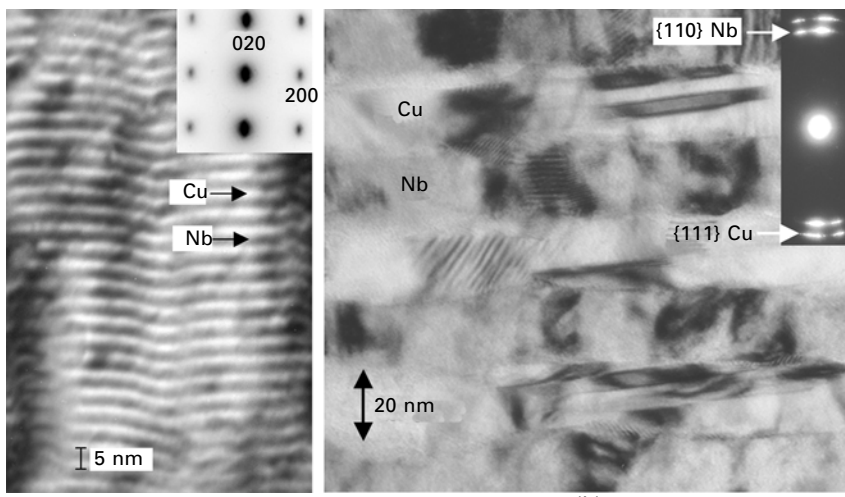
7.1 Schematic diagrams (not to scale) showing the key elements of (a) magnetron sputtering system (top view), and (b) electron beam evaporator (side view) used for nanolaminate synthesis.

shown in Fig. 7.1(b) can be used to bombard the growing film (ion-beam assisted deposition). Both sputtering [18] and evaporation [19, 20] techniques have been used to prepare nanolaminates with total thicknesses in the several hundred micrometer range.

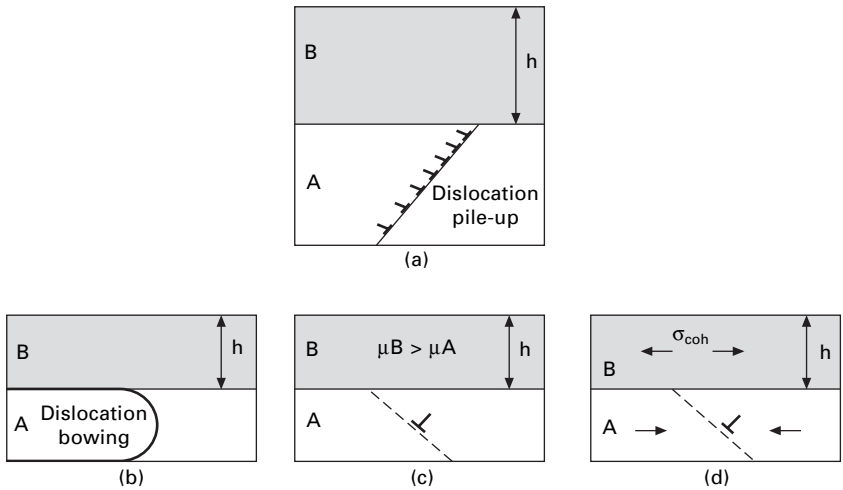
Examples of nanolaminates synthesized by PVD techniques are shown in Fig. 7.2 as cross-section transmission electron microscope (TEM) micrographs. The microstructures produced could be single crystalline or textured polycrystalline. A single crystal cube-on-cube oriented Cu-Ni nanolaminate is shown in Fig. 7.2(a). This nanolaminate has a bilayer period (λ) of 5 nm [17]. The thickness (h) of the individual Cu and Ni layers is 2.5 nm each. Throughout this article, the nanolaminates will be identified by the individual layer thickness (h) when the constituent layers have equal thicknesses. In Fig. 7.2(b), textured polycrystalline structure of Cu-Nb nanolaminate with h = 20 nm is shown. The interface plane is {110}bcc-Nb // {111}fcc-Cu. The in-plane grain size (d) increases with increasing h.

7.3 Overview of strengthening mechanisms

The dislocation mechanisms that are relevant to the deformation of nanolayered materials are briefly outlined here, and shown schematically in Fig. 7.3.



7.2 Cross-section TEM micrographs showing microstructures of (a) single crystalline Cu-Ni (h = 2.5 nm) and (b) textured polycrystalline Cu-Nb (h = 20 nm) nanolamintes. The respective diffraction patterns in the inset indicate a cube-on-cube orientation relation for Cu-Ni and {110} Nb//{111} Cu texture.



7.3 Schematic of the key unit processes considered in interpreting the deformation behavior of nanolaminates. (a) Dislocation pile-up-based Hall–Petch model at sub-micron to micron length scales; (b)–(d) single dislocation behavior that could be confined layer Orowan bowing (b) or interface crossing of single dislocations overcoming barriers such as image force on the glide dislocation from a higher modulus layer (c) or alternating tension-compression coherency stresses (d).

7.3.1 Hall–Petch model [21]

This model was originally applied, empirically, to strengthening from grain boundaries in single-phase metals. Later, it was shown that the model applies to laminate metallic composites as well [22]. The physical mechanism underlying this scaling law is that dislocations pile up at grain or interphase boundaries till the applied stress plus the stress concentration due to the pile-up can exceed the barrier strength and transmit slip across the boundary (Fig. 7.3(a)). At smaller grain sizes, the number of dislocations in the pile-up is smaller and hence, the resulting stress concentration is small and a larger applied stress is needed for slip transmission across boundaries. Quantitatively, as derived below, this results in the following scaling: $\sigma_{ys} \propto h^{-1/2}$, ($\sigma_{ys}$ is the yield strength and h is the layer thickness or grain size).

The number of dislocations (N) in a pile-up scales with the distance between source and obstacle (L) as [23]:

$$N = \frac{\pi L(1-\nu)(\tau - \tau_o)}{\mu b} \qquad 7.1$$

where ν is the Poisson ratio, μ is the shear modulus, b is the Burgers vector, τ is the applied shear stress and τ_o is the lattice friction stress for the motion of dislocations in the crystal. The stress concentration produced by the pile-

up is given by the product of N and the effective shear stress in the glide plane pushing the pile-up against the grain boundary. At the point when the leading dislocation in the pile-up is just able to cross the interface:

$$N(\tau - \tau_o) = \tau^* \tag{7.2}$$

where τ^* is the barrier strength of the boundary for slip transmission.

Eliminating N from the above equations, we get:

$$\tau = \tau_o + \left(\frac{\tau^* \mu b}{\pi(1-\nu)L}\right)^{1/2} \tag{7.3}$$

Eqn 7.3 can be re-written as:

$$\tau = \tau_o + kL^{-0.5} \tag{7.4}$$

where $k = \left(\frac{\tau^* \mu b}{\pi(1-\nu)}\right)^{0.5}$ is known as the Hall–Petch slope. Equation 7.4 is commonly referred to as the Hall–Petch model. Typically, the pile-up length is on the order of the grain size (or, the layer thickness for nanolaminates), so the appropriate microstructural dimension is used in eqn 7.4 instead of L. Simply, strength scales with the inverse square root of the grain size. The Hall–Petch slope determines the rate at which strength increases with grain size refinement [24–28].

7.3.2 Single dislocation behavior

At nanometer length scales, dislocation pile-ups may not form and the deformation behavior is controlled by the nucleation and motion of single dislocations. Two kinds of single dislocation behavior are considered here. The first kind is referred to as 'confined layer slip', and involves propagation of single dislocation loops parallel to the interfaces via an Orowan bowing type mechanism [24, 29–34]. To have slip propagate through the nanolaminate structure, both metallic layers will deform by confined layer slip of Orowan loops (Fig. 7.3(b)). For this mechanism, strength will still increase with decreasing layer thickness: $\sigma_{ys} \propto h^{-1}\ln(h)$.

The second kind of single dislocation behavior relevant to nanolaminates involves 'interface crossing'. Here, the critical event that determines the yield strength is the stress needed to push a glide dislocation from one side of the interface to the other [35, 36]. Quantitatively, this is of the order of τ^* that appears in eqn 7.3 above. Several factors contribute to the interface resistance to single dislocation transmission:

- **Koehler stress** [37]: for multilayers with large shear modulus mismatch between layers, a dislocation in the lower modulus (hence, lower line energy) phase may need to overcome a significant repulsive image stress from the higher modulus phase before slip can be transmitted across

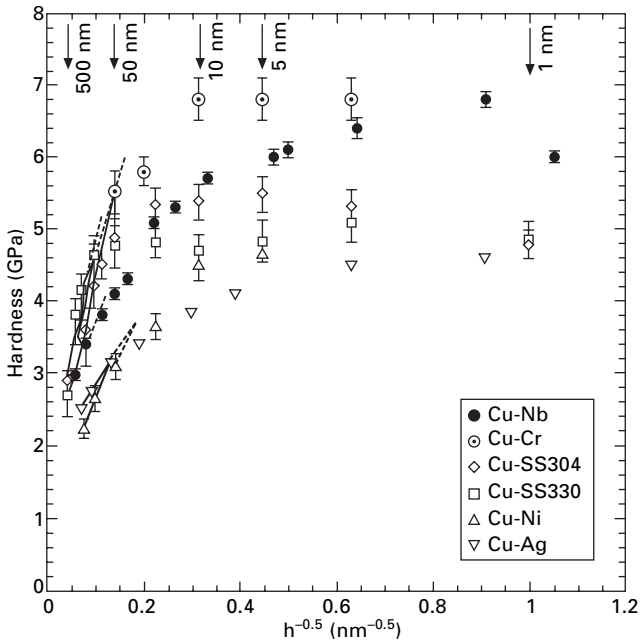
layers (Fig. 7.3(c)). This image force will primarily depend on the shear modulus mismatch and to a first approximation, will be independent of h.

- **Coherency stress**: very thin lattice matched multilayers have large in-plane coherency stresses that alternate from tensile to compressive between layers, giving rise to a periodic resistance to dislocation motion across layers (Fig. 7.3(d)) [35]. Above the critical thickness for coherency loss, interfaces are semi-coherent and have dislocation arrays to accommodate the misfit between layers. The stress field of the misfit dislocation array then poses a barrier to dislocations crossing the interface.
- **Other factors**: crystallographic mismatch in the slip systems across the interface, and stacking fault energy differences between the layers can also contribute to the interface barrier strength [36]. The role of these mechanisms in interpreting the measured mechanical properties of multilayers is discussed later in this article.

7.4 Dependence of nanolaminate strength on layer thickness

Figure 7.4 shows the nanoindentation measured hardness (H) of a variety of Cu-based multilayers as a function of $1/\sqrt{h}$ where h is the individual layer thickness, equal for both layers [4, 17, 33, 38–41]. For all systems, H scales linearly with $1/\sqrt{h}$ for $h > \sim 50$ nm, consistent with dislocation pile-up based Hall–Petch (H–P) model. This indicates that the H–P model for increase in strength with microstructural refinement as observed in micron-scale materials is also valid in these vapor deposited nanolaminates but only at the sub-micron layer thicknesses. In the nanometer-regime, (typically, $h < \sim 50$ nm), this scaling law is not applicable since continuum-scale dislocation pile-ups cannot form at such fine length scales and the behavior must now be interpreted in terms of single dislocations as discussed in section 7.3.2 above. As shown in Fig. 7.4, when the H–P model is not applicable, three kinds of functional dependence of H on h are observed: (i) at a few tens to a few nanometers range, H still increases with decreasing h but not by the $1/\sqrt{h}$ scaling, (ii) a peak in H is reached when h is of the order of a couple nanometers, (iii) at $h <$ a couple nanometers, some systems exhibit a small drop in hardness below the peak value. The maximum hardnesses achieved in all these systems are significantly higher than the rule-of-mixtures estimates.

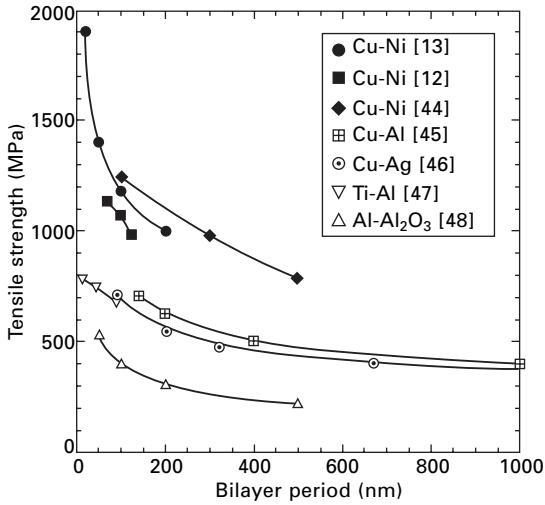
Limited studies have been performed to evaluate the tensile properties of nanolaminates. In many of these investigations tensile strengths were not reported as a function of layer thickness, rather the strength values for only a couple of different thicknesses were measured. For example, Mara *et al.* [42] measured a tensile strength of ~1.1 GPa in sputter-deposited Cu-Nb with h = 75 nm. The same sample (Fig. 7.4) had hardness of 3.6 ± 0.2 GPa which is in reasonable agreement with flow stress in uniaxial loading being



7.4 Hall–Petch plot of the nanoindentation measured hardness of a number of Cu-based nanolaminates (h = layer thickness or one-half of the bilayer period). For each set of data, a linear fit (solid line) is consistent with Hall–Petch behavior of hardness scaling with $1/\sqrt{h}$. For all systems, in the few to a few tens of nanometers range the Hall–Petch scaling law is not applicable, although hardness still increases with decreasing h.

a factor of 3 lower than hardness. Nieh *et al*. [43] reported tensile strength of 1120 MPa for a sputter-deposited Cu-amorphous Cu_4Zr_3 multilayer with bilayer period of 40 nm. Similarly, GPa level tensile strengths were reported for nanolaminates prepared by repeated rolling and bonding process. For Fe-Cu with bilayer period of 50 nm, tensile strength of 1570 MPa was reported [6], and for Ag-Ni with Ag layer thickness of 2.6 nm, tensile strength was ~1150 MPa [9]. In some other investigations, tensile properties were measured as a function of bilayer period and the data are compiled in Fig. 7.5. Since some of the data in Fig. 7.5 are on systems where the individual layer thicknesses were not equal, the bilayer periods (sum of the individual layer thickness of the two layers) are plotted instead of h.

In some of the data shown in Fig. 7.5 (e.g., electrodeposited Cu-Ni), a drop in strength was noted below bilayer period of 20 nm. This has been excluded from the data shown in Fig. 7.5 since as shown in later sections, a drop in strength of Cu-Ni is not expected until the layer thickness approaches the dislocation core dimensions of about a nanometer. Also, recent experimental



7.5 Tensile strengths of several metallic nanolaminates as a function of the bilayer repeat length.

work has shown that in the formation of nodules or discontinuous Cu layers during electrodeposition of Cu-Ni may be a factor in a drop of strength at layer thicknesses of a few tens of nanometers [49]. Clearly, more work is needed in evaluating tensile properties of nanolaminates. In some investigations, fracture prior to significant plasticity in tension, particularly at length scales of a few nanometers made it difficult to ascertain the yield behavior in tension [46]. In spite of the limited plasticity in tension, these materials exhibited ductile (shear) type fracture indicating that at very high strengths, the work hardening rate is lower than the flow stress resulting in a geometrical softening due to localized necking.

The data shown in Fig. 7.4 are summarized in Table 7.1 along with information on other metallic multilayers [50–53]. The data compiled in Table 7.1 show that the peak strength for most systems is typically of the order of a few GPa, and expressed as a fraction of the composite Young's modulus (E_c) is typically within a factor of 2–4 of a lower-bound theoretical strength limit (~E/25). For systems where data are available at larger bilayer periods where H–P behavior is observed, the H–P slope is calculated (divided by 3 for hardness data) and also shown in Table 7.1. This is because for single-phase metals, H–P slope is shown to scale with the modulus and the grain boundary resistance to slip transmission. For a given type of material systems, e.g., polycrystalline fcc-fcc such as Ag-Cu, Cu-Ni and Cu-SS330, note that the H–P slope increases as E_c. A similar trend is observed in polycrystalline fcc–bcc systems (Cu-Nb, Cu-Cr and Pt-Fe). Higher modulus and higher interface resistance to slip transmission produce higher H–P slopes and thus, higher strength at a given length scale in the H–P regime.

Table 7.1 Some mechanical properties of metallic nanolaminates

Multilayer system	σ_{max} (GPa)	h (nm) at σ_{max}	σ_{max} as a fraction of E_c	H–P slope (MPa$\sqrt{m}$)	Reference
fcc–fcc					
Cu-Ni	1.6	5	E_c /100	0.14	10
				0.10 *	12
Ag-Cu	1.5	1.2	E_c /70	0.11	
				0.11*	46
Au-Ni	2.1	1–2	E_c /70	nr	14
Ag-Ni	2.08	5	E_c /70	nr	15 and 52
Cu-SS330	1.7	5	E_c /100	0.204	
Al-Al_3Sc†	1	1.5 (Al_3Sc); 50 (Al)	E_c /75	nr	
fcc–bcc					
Ag-Cr	1.63	2.5	E_c /100	nr	1
Cu-Nb	2.2	2.5	E_c /55	0.18	10
Cu-Cr	2.3	2.5–10	E_c /80	0.30	10
Pt-Fe	3.2	2	E_c /65	0.23	1
Cu-304SS	1.8	5	E_c /95	0.30	
bcc–bcc					
Fe-Cr (sc)	2.8	2	E_c /80	nr	1
Mo-W	6.5	5	E_c /55	nr	

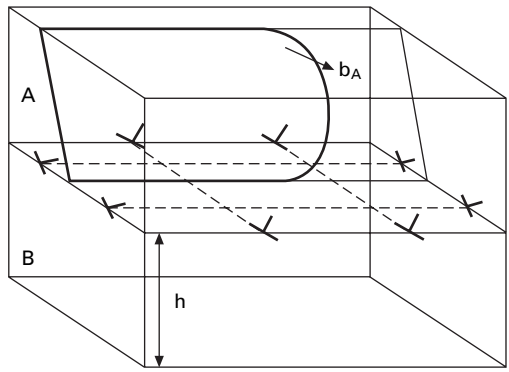
σ_{max} is the maximum strength for that system estimated as nanoindentation measured hardness divided by a factor of 3.
E_c is the rule-of-mixtures Young's modulus for the multilayer.
H–P slope refers to the Hall–Petch slope obtained from data at larger h which hardness scales with $h^{-0.5}$.
* slope from tensile data.
nr (not reported).
†Al_3Sc is fcc-based ordered $L1_2$ structure.

7.5 Modeling of single dislocation behavior

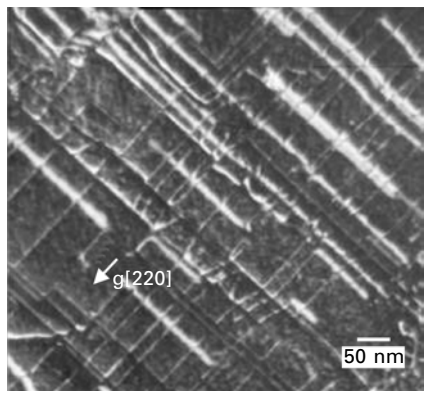
7.5.1 Confined layer slip (CLS)

As mentioned in section 7.3, CLS involves the propagation of single dislocation loops parallel to the interfaces via an Orowan bowing type mechanism and results in an increase in yield strength with decreasing layer thickness. Experimentally, CLS has been observed in *in situ* straining TEM experiments [24, 54, 55, 56]. CLS mechanism has also been inferred from post-mortem TEM studies. Consider the case of a thin Ni film deposited on a much thicker {001} Cu substrate. Initially, Ni is coherent with a tensile strain of ~2.6%. However, as the Ni thickness exceeds a critical value (~ 2.5 nm in this case [57]), the grown-in threading dislocations are observed to glide in the Ni overlayer, depositing misfit dislocations at the Cu-Ni interface. This CLS in Ni is shown schematically in Fig. 7.6(a) and the corresponding plan view TEM micrograph is shown in Fig. 7.6(b).

The Orowan stress (τ_{Orowan}) to propagate the loop shown in Fig. 7.6(a) is given as:



(a)



(b)

7.6 (a) Schematic illustration of confined layer slip in layer A in an A/B nanolaminate. The propagation of single dislocation loops confined within a layer leads to the deposition of misfit type dislocations at the interface. An example of such a slip mode is observed during coherency loss of a thin film on a thick substrate, as shown in (b) for {001} Ni-Cu interface. This plan view TEM image for a 5 nm thick Ni on 100 nm thick Cu shows dislocations at the interface with Burgers vectors in the {111} glide planes of Ni.

$$\tau_{\text{Orowan}} = \frac{2\,W_D}{bh} \qquad 7.5$$

where b is the Burgers vector, h is the layer thickness and W_D is the self-energy per unit length of each interface dislocation segment deposited. To calculate W_D of the glide dislocation which is 60° from screw, we resolve its Burgers vector into three components along the orthogonal x, y and z directions as listed in Table 7.2 [33].

The self-energies of the resolved components along x, y and z directions respectively are [33]:

Table 7.2 Line directions, Burgers vectors and character of glide dislocation resolved along three orthogonal directions for calculation of self-energy of a glide dislocation loop

	Glide dislocation (**b**)	Resolved **b** along x-direction	Resolved **b** along y-direction	Resolved **b** along z-direction
Line direction	[101]	[101]	[101]	[101]
b	1/2[110]	1/4[10$\bar{1}$]	1/2[010]	1/4[101]
Character	60°	edge	edge	screw

(note: 1/2[110] = 1/4[10 $\bar{1}$] + 1/2[010] + 1/4[101]; also, $b_x = b_z = b/2$ and $b_y = b/\sqrt{2}$)

$$\left(\frac{W}{L}\right)_x = \frac{\mu b^2}{8\pi(1-\nu)} \ln\frac{h_p}{b} \tag{7.6a}$$

$$\left(\frac{W}{L}\right)_y = \frac{\mu b^2}{4\pi(1-\nu)} \ln\frac{h_p}{b} \tag{7.6b}$$

$$\left(\frac{W}{L}\right)_z = \frac{\mu b^2}{8\pi} \ln\frac{h_p}{b} \tag{7.6c}$$

where h_p is the layer thickness measured parallel to the glide plane. For the case shown in Fig. 7.6, the angle between the {111} glide plane and {001} interface plane is 54.7°. Thus, $h_p = h/\sin 54.7°$.

Summing up the three equations above, we get:

$$W_D = \frac{\mu b^2(4-\nu)}{8\pi(1-\nu)} \ln\frac{h_p}{b} \tag{7.7}$$

Note that the denominator b in the ln term implies that the lower cut-off in the self-energy calculation is equal to the magnitude of b. In fact, the core width may be three to four times the Burgers vector in metals such as Cu, Ni, etc. [35]. Thus '3b' is used in the denominator of the ln term to calculate W_D for the case of Cu-Ni.

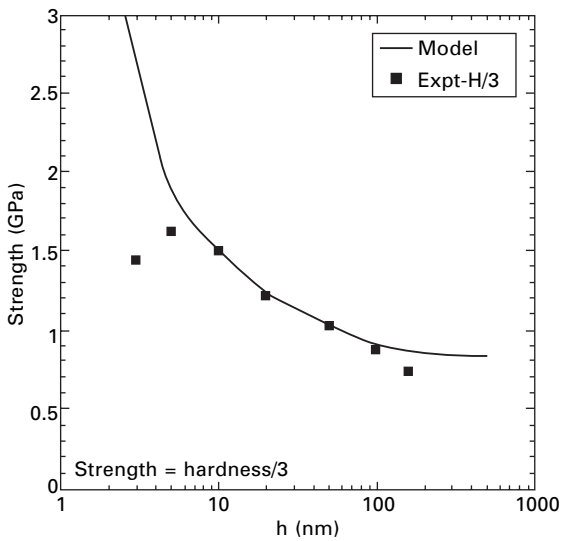
Substituting eqn 7.7 into 7.5 gives the Orowan stress for confined layer slip. We note that the presence of residual stresses in the layers (e.g., alternating tensile-compressive coherency stress between the layers) could affect the Orowan stress for glide. For example, under applied tension, the layer with residual tensile stress may yield first at an applied stress given as τ_{Orowan} from eqn 7.5 minus the in-plane residual stress ($\tau_{residual}$). Finally, the array of misfit dislocations at the interface has a stress field associated with it that is an additional barrier to the confined layer slip. From this stress tensor, a shear stress component resolved onto the glide system can be computed (τ_{array}) at the interface. The total stress for yield via CLS (τ_{cls}) is, therefore [33]:

$$\tau_{cls} = \tau_{Orowan} - \tau_{residual} + \tau_{array} \quad 7.8$$

The calculated τ^* is compared with experimental data on Cu-Ni system in Fig. 7.7. Note that model prediction compares favorably with the measured values of strength at nanometer length scales where the Hall–Petch extrapolation is not valid. However, at a layer thickness of a couple of nanometers, the strength reaches a peak whereas the CLS model predicts a continuous increase in strength with decreasing layer thickness. The cross over from the CLS to the interface crossing mechanism presumably defines the peak in strength. In other words, at h of a couple nanometers, the interface barrier to slip transmission (τ^*) is lower than τ_{cls}, and therefore, the composite strength is determined by the stress needed to transmit a glide dislocation across the interface as opposed to confined layer slip. The calculation of τ^* is described next.

7.5.2 Interface crossing

In our discussion of the stress needed for single dislocation transmission across interfaces, we consider two kinds of interfaces. The first kind is where the slip planes and slip directions are continuous across the interfaces and



7.7 Confined layer slip model prediction for Cu-Ni. Experimental data points are nanoindentation hardness divided by a factor of 3. The model prediction (solid line) fits the data well in the layer thickness range of a few nm to a few tens of nm. Note that the behavior at h ≈ couple nanometers, i.e., a peak in strength followed by a drop in strength as h approaches the dislocation core dimensions, is not predicted by the model.

both layers have the same crystal structure (iso-structural case), e.g., fcc-fcc Cu-Ni, Cu-Ag systems. The second kind is where the slip planes and slip directions are discontinuous across the interfaces and the two layers have different crystal structures (non-iso-structural case), e.g., fcc-bcc Cu-Nb systems. Atomistic calculations using the molecular dynamics (MD) approach for these two categories of the interfaces are described next.

For the coherent Cu-Ni case, in spite of the continuity of slip, dislocations are not able to easily move from one layer to the next because of several factors, the principal one being the alternating compression-to-tension in-plane coherency stresses [35, 36, 58]. These coherency strains enable matching of the lattice parameters in the two adjacent lattices. Coherency strains vary periodically with the multilayer wavelength, are elastic, and, can be quite large (e.g., 1–2%) when compared with the elastic strain at yield of conventional bulk forms of the constituents. Other factors that influence forces on and mobility of glide dislocations are image stresses due to modulus mismatch (Koehler force), step creation on crossing interface, and cutting of misfit dislocations that are on the interface. All of these factors have been the topics of study by means of atomistic simulations of fcc–fcc systems, namely Cu/Ni and Cu/Ag [35, 36, 58]. The results suggest that, for iso-structural, coherent systems, the most important effect on strength derives from the coherency strains. Hoagland *et al*. [58] found that the magnitude of the in-plane applied normal stress needed to cause the lead partial to cross a coherent interface in an infinite bilayer Cu/Ni model was about 2.6 GPa. Comparing this stress to the biaxial in-plane coherency stress in the Cu and Ni that are –2.4 GPa and +2.4 GPa, respectively, it appears that the critical applied stress is about equal to the coherency stress. The reason is that the magnitude of the applied in-plane stress needs to be this large in order to zero the coherency stress in one of the layers. Thus an applied stress of +2.6 GPa would be required for a dislocation to cross from the Ni to the Cu and –2.6 GPa for an opposite sign dislocation to cross from Cu to Ni. The small difference between the coherency stress and the critical stress arises from contributions from nonlinear elastic effects, Koehler forces, and step creation.

Given the coherency stress dominated behavior for isostructural systems revealed in atomistic modeling, Hoagland *et al*. [58] derived an expression for the peak strength (σ_f) attainable assuming it to be equal to coherency stress:

$$\sigma_f = \frac{C^A C^B}{C^A + C^B} \varepsilon_m \qquad 7.9$$

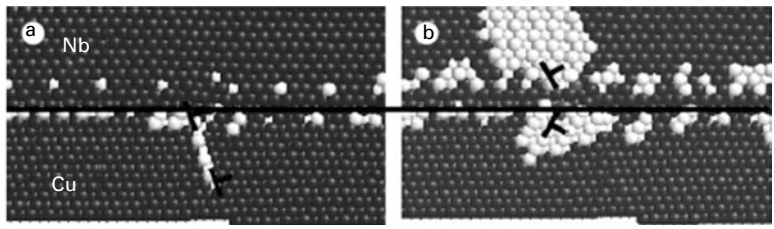
where C^A and C^B are the effective biaxial elastic constants for layers A and B respectively.

The right side of eqn 7.9 is the maximum yield strength of the composite if it is fully coherent and therefore, the individual layers are quite thin. For

the Cu/Ni embedded atom method (EAM) potentials, Ccu = 133 GPa, CNi = 212 GPa, εm = 0.0266, a strength of 2.2 GPa is calculated in reasonable agreement with the MD simulation results. However, this estimate is somewhat larger than the peak strength values of about 1.7 GPa measured experimentally for Cu/Ni in Fig. 7.7 at h ~ 5 nm. At h = 5 nm, partial loss of coherency at the Ni-Cu interface may reduce the level of coherency stress and hence σ_f. Therefore, eqn 7.9 should be regarded as a guide since it is based on an assumption of perfectly coherent interfaces. Accordingly, it shows that higher strengths should be achieved in higher modulus materials, combinations of materials with higher misfit strains, and with moduli that maximize the coefficient in eqn 7.9.

For Cu/Ag the misfit strain is 0.12, significantly larger than for Cu/Ni. However, this misfit is too large to achieve coherency in this system because the misfit dislocations are so close together that the cores of the adjacent dislocations at the interface overlap [58]. Consequently, the maximum strength in Cu/Ag is less than that of Cu/Ni. In Cu/Ni, the experimental results show that the peak strength develops at about 5 nm layer thickness and then decreases for thinner layers. MD simulation results obtained from models with very thin layers also show a corresponding decrease in the stress needed to transfer slip across the interface as the layer thickness is reduced. There are two explanations for this behavior: (i) the image forces on glide dislocations due to modulus mismatch decrease with layer thickness, and (ii) the dislocation core size becomes a significant fraction of the layer thickness.

For the non-iso-structural case, the transfer of slip across an interface between two different phases must, in general, involve dislocation multiplication because of discontinuity of slip plane and slip direction. Furthermore, the interface may have limited shear strength, a property that leads to attraction of glide dislocations to the interface and core spreading for dislocations in the interface. Such interfaces are likely to be incoherent, and, although periodic structures might occur, they will not sustain the large stresses that can develop in the transparent interface systems. The factors involved in slip transfer across non-iso-structural interfaces are more complex than those discussed in the previous section. Hoagland [58] employed a Cu-Nb EAM potential to examine the behavior of slip through Cu/Nb interfaces. Figure 7.8(a) shows results from an unstressed Cu/Nb model in which a dissociated glide dislocation was placed in the Cu layer near the interface. In this particular case, the model is oriented such that one set of slip planes in the Cu and another set in the Nb are misoriented by a small tilt and their traces on the interface are normal to the page. The lead partial of a dissociated glide dislocation in the Cu was attracted to the interface while the trailing partial remained some distance away. At a relatively small applied stress of about 0.2 GPa, the trailing dislocation also entered the interface. Disregistry analysis shows that the in-plane components of these dislocations spread significantly in the



7.8 A molecular dynamics simulation of slip transmission through a Cu/Nb interface. Energetic atoms associated with dislocations and interfaces appear light. In (a) the model is unstressed and one of the Shockley partials has entered the interface. In (b) an in-plane strain of 3.5% has been applied sufficient to cause the emergence of a glide dislocation in the Nb layer. Another Shockley partial can be seen emerging in the Cu layer [58].

interface. Very large strains are required for the slip to transfer into the Nb. Figure 7.8(b) shows a glide dislocation emerging in the Nb on a {110} slip plane at an in-plane strain of 3.5% (equal in both layers), producing an applied tensile stress of about 7 GPa in the Cu and 4.5 GPa in the Nb. There is also evidence of a dislocation emerging on another {111} plane in the Cu layer at this huge stress. The results shown here derive from calculations near 0 K and it is very likely that thermal activation would assist dislocation nucleation into the Nb layer at higher temperatures and at lower stresses.

In a different simulation, the shear strength of the interface in the Cu/Nb system was measured by applying an increasing homogeneous shear strain accompanied by energy minimization. The resulting shear stress-strain curve for the interface indicates that the interface begins to undergo stick-slip sliding at a shear stress of about 0.55 GPa. Relative to conventional strength levels, this interfacial shear strength is large but it is very much less than the strength one obtains for homogeneous shear parallel to the closest packed planes, the theoretical shear strength, which is about 5 GPa for Cu and somewhat larger for Nb. A limited interfacial strength has important implications for the interactions of dislocations with the interface and the behavior of dislocations that enter the interface.

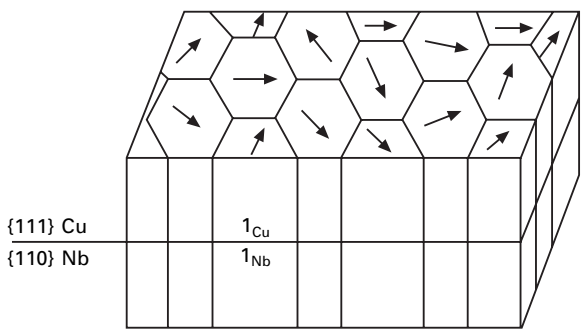
7.6 Plastic stability of nanolaminates

While the previous section discussed the length-scale dependence of yield strength of nanolaminates, this section describes the post-yield behavior. As the nanolaminates are deformed to large plastic strains, several new questions can be asked about the mechanical behavior. For example, what are the dislocation storage and recovery mechanisms? How do the work hardening rates in nanolaminates compare to microlaminates? What are the fracture mechanisms and strains to failure? For polycrystalline nanolaminates, do

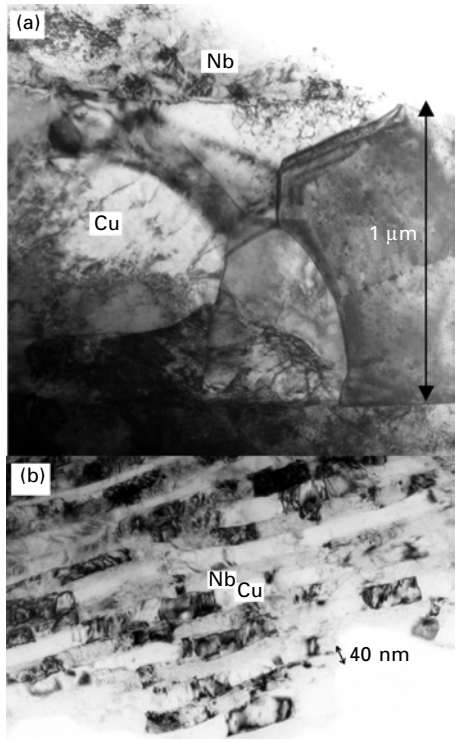
any preferred grain orientations (textures) develop after plastic straining and how are such deformation textures different from bulk metals?

Due to the limited deformability in tension, and the difficulty of performing uniaxial compression tests on thin foils, rolling (plane strain compression) of nanolaminates is a way of subjecting these materials to large plastic strains. Such experiments have been attempted on sputter deposited Cu-Nb nanolaminates (h = 75 μm, total foil thickness = 15 μm) as well as sputter deposited microlaminates (h = 4 μm or 2 μm, total foil thickness = 16 μm) for comparison [59, 60]. In both cases, the initial microstructure was as shown in Fig. 7.2(b), polycrystalline with columnar grains with Kurdjumov–Sachs (K–S) orientation relationship: {110}bcc-Nb // {111}fcc-Cu, and <111>Nb // <110>Cu. The K–S relationship holds within a Cu/Nb bicrystal. However, the microstructure is polycrystalline, so we need to understand the texture in the sample composed of a lot of Cu/Nb bicrystals each having K–S relationship. The texture deduced from pole figure measurements is shown schematically in Fig. 7.9. In the sputter deposited Cu-Nb nano- or micro-laminate foils, the Cu/Nb bicrystals are strongly fiber-textured such that {110}bcc-Nb // {111}fcc-Cu // interface plane. However, from one bicrystal to the adjacent, the in-plane K–S directions (<111>Nb // <110>Cu) are randomly oriented [60].

The evolution of the dislocation substructures in cold rolled Cu-Nb laminates is shown in Fig. 7.10 [59, 60]. For the case of microlaminates, Fig. 7.10(a), (h = 2 μm reduced by rolling to h = 1 μm), the layer thickness reduction was non-uniform. A very high density of dislocation tangles/networks was noted in the Nb layer, whereas in the Cu layer the dislocation structure was recovered to form a sub-micrometer scale cell structure comprising dislocation boundaries with a significantly lower dislocation density within the cells. The behavior



7.9 Schematic illustration of the texture inferred from X-ray pole-figure measurements of sputter deposited Cu-Nb nanolaminates. The films exhibit a fiber texture such that the interface plane is {111}Cu//{110}Nb for the majority of the Cu/Nb column grain pairs. The arrows denote the in-plane Kurdjumov–Sachs directions, <111>Nb//<110>Cu, for each of the columnar Cu/Nb grain pairs. Thus, from one Cu/Nb bicrystal to another there is no preferred alignment of <111>Nb//<110>Cu in the interface plane.



7.10 TEM images showing microstructures after rolling in Cu-Nb multilayers; (a) 2 μm initial layer thickness reduced to 1 μm, and (b) 75 nm initial layer thickness reduced to 40 nm. Note cell structure in (a) and absence of dislocation tangles or cell structures in (b).

shown in Fig. 7.10(a) is typical of that of bulk metals. In other words, pure Cu or Nb rolled to 50% reduction would develop a cell structure quite similar to the meso-scale Cu and Nb layers in the multilayer. However, the nanolaminate with initial h = 75 nm did not exhibit any cell sub-structure or tangles/networks of dislocations within the layers, even after ~50% reduction in rolling, although the Cu and Nb layers were reduced in thickness uniformly (Fig. 7.10(b)). While a few dislocations could be discerned threading the layer thickness, such threading dislocation segments were observed in the as-deposited samples as well. Statistically, no significant change in the threading dislocation density was observed even after 50% reduction. Hence, despite the very different flow strengths and hardening rates of Cu and Nb single-phase materials, the alternating layers of Cu and Nb in nanoscale multilayers respond to room temperature rolling with a uniform reduction in layer thickness. This is an indication that the length scale is small enough so that the plastic properties of the constituent phases in the bulk form do not play a role in determining the deformation behavior.

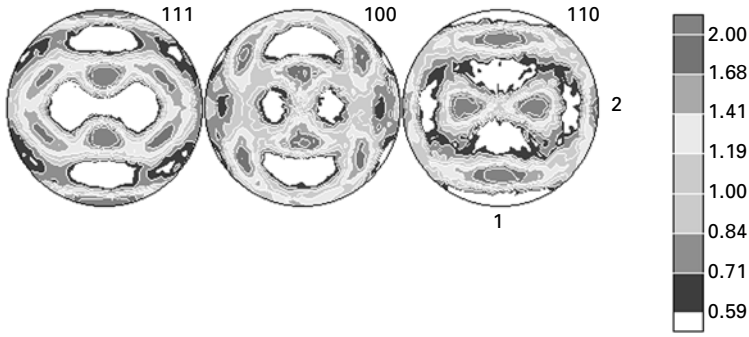
In addition to the significant difference in the dislocation storage in micro vs. nanolaminates, the deformation textures were also completely different [61, 62]. Figure 7.11(a), (b) show that typical fcc and bcc bulk rolling textures develop for samples with a 4 μm initial layer thickness, so that the Kurdjumov-Sachs orientation relations are lost during deformation. In contrast, Fig. 7.11(c), (d) show that the local Kurdjumov–Sachs relation along the interface normal is maintained in nanolaminates. The pole figure measurements also showed that for a majority of the grains there is a co-rotation of the in-plane <111> Nb // <110> Cu directions to about 55° from the rolling direction. From the KS orientation relationship, this implies that the rolling direction is <001> Nb. This in-plane texturing that develops during rolling is illustrated schematically in Fig. 7.11(e). Note that the layers do not become mono-crystalline after rolling, there is simply a tendency of the grains to rotate in-plane to the stable end orientation of <111> Nb // <110> Cu // 55° from RD.

The observations of dislocation substructures and deformation textures (Figs 7.10, 7.11) indicate that for the same material systems synthesized and tested the same way, reducing the length scale from micro- to nanometer scale introduces qualitatively new behavior. This length scale induced new behavior must be consistent with the change in strengthening mechanism discussed earlier (i.e., from Hall–Petch to confined layer slip). Using the CLS dislocation mechanism, the following specific observations need to be explained: (i) lack of dislocation cell structure formation, (ii) plastic stability in rolling (i.e. uniform reduction to large plastic strains), (iii) no out-of-plane rotations (i.e. preservation of the (110) Nb // (111) Cu interface plane), and (iv) in-plane co-rotation of the grains such that the rolling direction is [001] Nb.

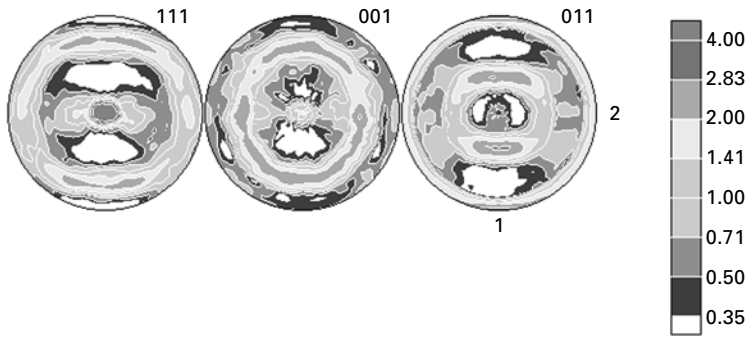
In bulk materials, the formation of dislocation cell structures involves interaction between pile-ups of dislocations on parallel (closely spaced) or intersecting slip planes. At length scales where CLS is observed, dislocation cell structures are unlikely to form since the layer thickness is less than the typical cell size in bulk materials. In CLS, glide dislocations propagate as single loops within each layer, depositing misfit type dislocations at the interface. Thus, interfaces act as sources and sinks of dislocations, and hence all the dislocation storage is at the interfaces, as opposed to bulk metals where the storage is at cell walls created during deformation.

The CLS mechanism depicted schematically in Fig. 7.12 illustrates the uniform reduction of both layers with no out-of-interface plane lattice rotation [60]. Figure 7.12 (a)–(c) schematically illustrates the evolution of the dislocation segments deposited at the interface with increasing strain (for simplicity, only one layer is shown). Figure 7.12 also shows the side view of the process illustrated in Fig. 7.6(a). Figure 7.12(a) shows that glide on a plane labeled xx′ has deposited dislocation segments at the interface while the bowed out gliding segment has moved parallel to the interface (and normal to the plane

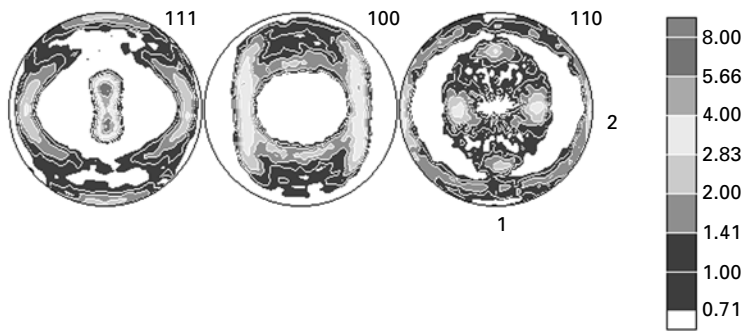
of view in Fig. 7.12) till it became absorbed in a columnar grain boundary or exited at the side surface of the sample. The dislocation segments (which have the Burgers vector of a perfect glide dislocation) at the interface will repel another nearby glide dislocation on the same xx′ plane, with the repulsion increasing the closer the spacing. The implication is that the next plasticity events will occur (stochastically) at other locations along the length of the interface, as shown in Fig. 7.12(b). The regions along the interface where the



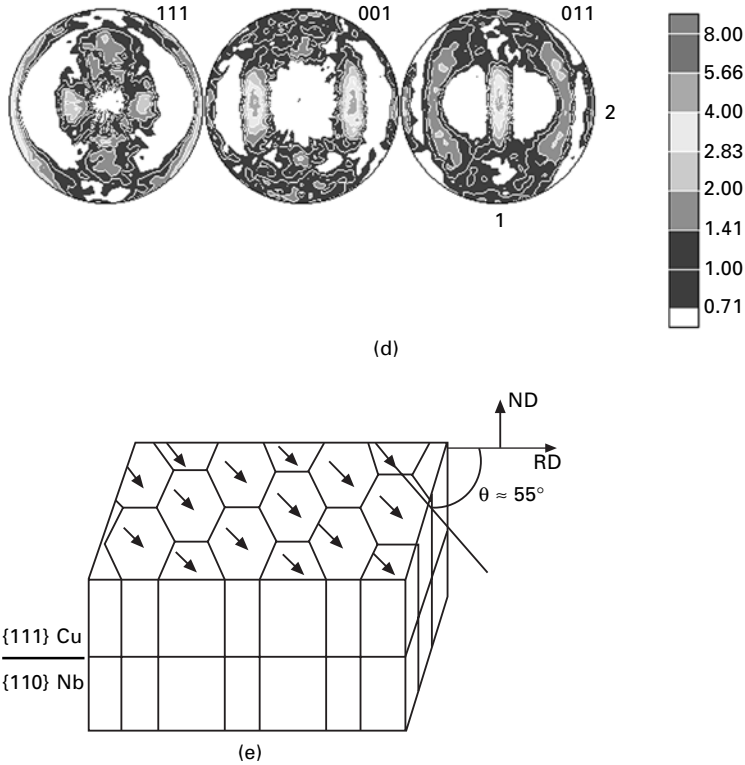
(a)



(b)

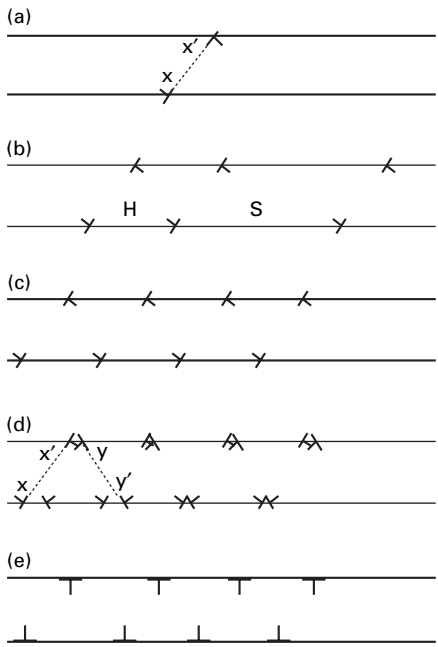


(c)



7.11 X-ray pole figures of Cu-Nb micro and nanolaminates rolled to 50% reduction in layer thickness. (a) <111>, <001> and <110> pole figures for Cu from a microlaminate with h = 4 μm rolled to h = 2 μm; (b) <111>, <001> and <110> pole figures for Nb from a microlaminate with h = 4 μm rolled to h = 2 μm; (c) <111>, <001> and <110> pole figures for Cu from a nanolaminate with h = 75 nm rolled to h = 40 nm; (d) <111>, <001> and <110> pole figures for Nb from a nanolaminate with h = 75 nm rolled to h = 40 nm. Direction 1 (vertical) is the rolling direction and direction 2 (horizontal) is the transverse direction. Comparing the <111> Cu poles in (a) and (c) shows that in microlaminates there is significant rotation of the <111> poles away from the interface normal while this rotation is largely suppressed in nanolaminates. Comparing the <110> Nb poles in (b) and (d) shows that in microlaminates there is significant rotation of the <110> poles away from the interface plane normal while this rotation is also largely suppressed in nanolaminates. In fact, the interface plane normal for Nb layers switches from <110> to <111> after rolling in microlaminates whereas it remains unchanged in nanolaminates. Thus, nanolaminates deform while preserving the initial Kurdjumov–Sachs orientation relationship whereas the initial orientation relationship is lost in microlaminates during deformation. For nanolaminates there is also an in-plane rotation such that most grains become aligned with <111>Nb//<110>Cu directions at about 55° to the rolling direction which becomes <001> Nb, shown schematically in (e). Compare (e) with Fig. 7.9 [61].

spacing between the segments deposited at the interface is smaller (such as the region label H) will be relatively harder as compared to regions with larger spacing (label S) between the dislocation segments deposited at the interface. In other words, a gliding hairpin dislocation loop also sees repulsion from nearby interface dislocations. The pair-wise repulsive force scales inversely with the spacing between interface dislocations. The result is that new plasticity events will preferentially occur at regions such as S and the deformation will tend to be homogeneously spread along the interface. Thus, the Orowan glide mechanism leads to a uniform reduction in layer thickness and plastic stability, as opposed to preferential thinning of the layer at a few



7.12 Schematic illustration of uniform straining in the layers, symmetric slip, and recovery of interface dislocations during single dislocation glide. For simplicity only one layer is shown in (a)–(e). (a)–(b) the repulsion associated with the dislocations deposited at the interface following glide on plane xx′ leads to spread of plasticity events along the interface. (c) regions such as H are locally harder for a new glide loop to propagate than those such as S that have a larger spacing of interface dislocations, leading to uniform distribution of slip events at the interface. (d) symmetric slip: lattice rotation about the interface plane by xx′ slip may be canceled by yy′ slip, preserving the interface plane orientation during plastic deformation. (e) nearby dislocation segments at the interface may react to form dislocations with Burgers vectors in the interface plane.

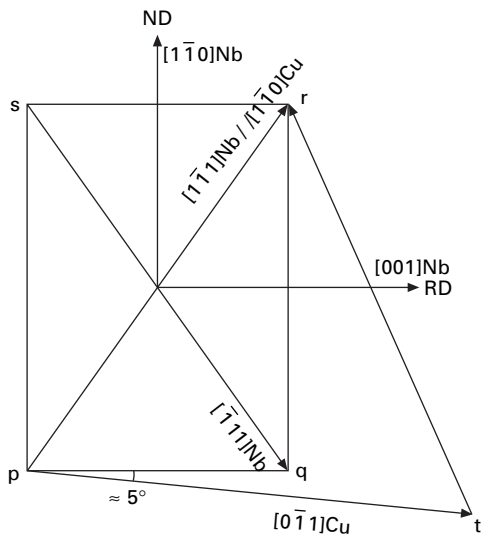
locations where the strain is concentrated. This local hardening at the site of a given depositing dislocation provides a resistance to shear instability and allows the observed extended plastic deformation with little macroscopic hardening. Note also that the CLS stress scales inversely with the layer thickness. Thus, for every plastic strain increment, as the layer thickness reduces, the applied stress to continue CLS increases. This is another source for work hardening in addition to the repulsion from the interface dislocation segments deposited.

The next observation to explain is the no out-of-plane rotations (i.e. preservation of the {110}Nb//{111}Cu interface plane) after large plastic strains in nanolaminates. The concept for this is illustrated in Fig. 7.12(d) where duplex slip is considered by activating the plane yy′ in addition to xx′. The grouping of interface segments of a specific $\mathbf{b_A}$ that is locally hardening for other dislocations with the same $\mathbf{b_A}$, as just discussed, is softening for the symmetric Burgers vector $\mathbf{b_B}$. The softening occurs because the two dislocations attract and, indeed if close enough, would react according to $\mathbf{b_A} + \mathbf{b_B} = \mathbf{b_C}$. This provides an energetic driving force that gives rise to a tendency for equal glide on the symmetric systems such as those in Fig. 7.12(d). Thus, the lattice rotation about the interface plane due to glide on xx′ is effectively canceled by an equal and opposite lattice rotation due to glide on yy′ and within each layer the single dislocation slip operates stably on symmetric glide systems. With continued straining, we consider the rearrangement of dislocations deposited at the interface. A possible scenario is shown in Fig. 7.12 (d) and (e). Here two neighboring interface dislocations have reacted to form another with Burgers vector in the (110) Nb // (111) Cu interface plane [63]. A similar process on the other side of the interface would result in opposite sign dislocations at the interface (not shown). The specific dislocations at the interface would annihilate resulting in the creation of a new interfacial area. Even if the reactant dislocations do not exactly coincide as in Fig. 7.12(d) and (e), short range attraction would tend to cause the reaction to occur by local climb in the interface, which would be spontaneous if the dislocation spacing were small enough. Thus, part of the plastic work done is stored in the material as the energy of the newly created interfacial area. A 50% rolling reduction results in a factor of two increase in interface area. A detail regarding the annihilation mechanism is that the Burgers vectors of the dislocations in the two layers will typically not be equal. Hence, the annihilation will tend to restore coherency and the accompanying local coherency strain and strain energy in the vicinity of the annihilation event.

Finally, we consider the in-plane lattice rotations such that the rolling direction (RD) ends up being <001> Nb, and correspondingly, ~5° from <011>Cu to maintain K–S orientation relation. Detailed crystal plasticity analyses have shown that the interface crystallography of Fig. 7.13 represents a stable end orientation that requires minimum work for further deformation.

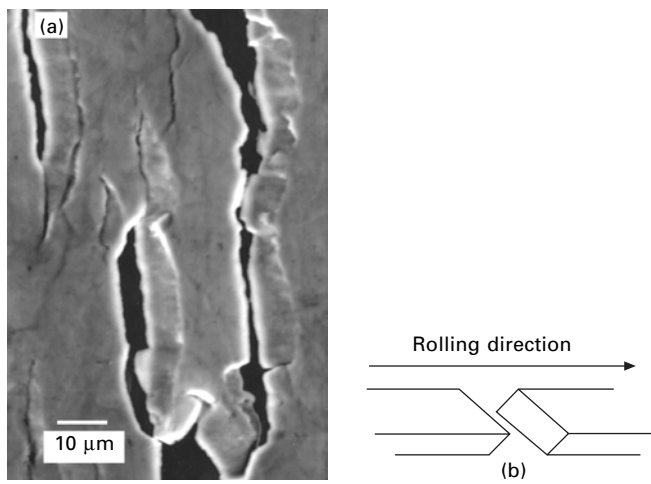
This is the view looking down on the (110) Nb//(111) Cu interface with the rectangle 'pqrs' representing (110) Nb and triangle 'prt' representing (111) Cu. In the interface plane, [1 $\bar{1}$ 1] Nb//[1 $\bar{1}$ 0] Cu represents the K–S direction. If the Nb grain rotates such that [001] ends up being parallel to RD and the K–S direction is maintained, then the following are obvious from Fig. 7.13: (i) the in-plane K–S direction is at an angle of 54.7° from the RD (i.e., the angle between <001> and <111>), as also shown in Fig. 7.11(e), and (ii) the RD in Cu is about 5° off from [0 $\bar{1}$ 1] in Cu. Crystal plasticity analysis also showed that the applied stress for plane strain compression of the Nb layer was a strong function of the angle between the in-plane K–S direction and the RD, and a clear minimum was observed when the RD was [001] in Nb. However, for the Cu layer, the applied stress was independent of the angle between the in-plane K–S direction and the RD. Thus, in-plane rotation will be dominated by Nb and Cu will follow along to maintain the K–S interface. In a macroscopic sense, the concept of plastic constraint is considered as one where a rigid phase exerts tractions that alter the plasticity of the neighboring phase. But in the nanolaminates described here, there is an interfacial constraint in which the pattern of plastic flow and the choice of slip systems are determined by the need to preserve a specific low energy interface crystallography.

Finally, fracture of the nanolaminate foils during rolling is described. Fig.



7.13 Schematic illustration of the crystallography at the interface after rolling looking down on the parallel (110) Nb (shown as rectangle pqrs) and (111) Cu (shown as triangle prt) planes. For <001> Nb//RD, the in-plane K–S direction [1 $\bar{1}$ 1] Nb // [1 $\bar{1}$ 0]Cu is ~ 55° from RD.

7.14 is a plan view SEM micrograph showing through thickness cracks in a rolled foil [59]. This Cu-Nb nanolaminate had an initial layer thickness of 15 nm and was rolled to 50% reduction. For the case shown in Fig. 7.14, the cracks were spaced about 50–100 μm (the total initial sample thickness was 7.5 mm) and were about 200 μm long, on average. With such a distribution of micro-cracks, the multilayered foil was perforated but the sample was still in one piece. Perhaps the most significant observation in Fig. 7.14 is that the fracture appears to be by shear cracks. The crack plane is not normal to the rolling direction, rather inclined and roughly along the planes of maximum shear stress. For initial layer thickness of 4 nm, shear cracks were observed at rolling reductions of 3–5%. However, samples with initial layer thickness of 75 nm could be rolled to reductions > 60% without any cracking. This behavior could be interpreted as follows. For initial layer thickness of a couple to a few nanometers, hardness reaches a peak with little or no dependence on layer thickness. This is consistent with a transition from CLS to interface crossing deformation mechanism as discussed in the previous section. For single dislocation transmission across the interface, as opposed to CLS, the work hardening rate would primarily result from the residual Burgers vector left behind at the interface. Thus, a second dislocation transmission event on the same glide plane would be harder due to repulsion experienced by the glide dislocation from the residual Burgers vector at the interface. Since the magnitude of the residual dislocation ($\mathbf{b}_{Nb}$-$\mathbf{b}_{Cu}$) is only a small fraction of the perfect glide dislocation deposited at the interface during CLS (Fig. 7.12(a)), this repulsion force would be relatively weaker. Nevertheless, this local

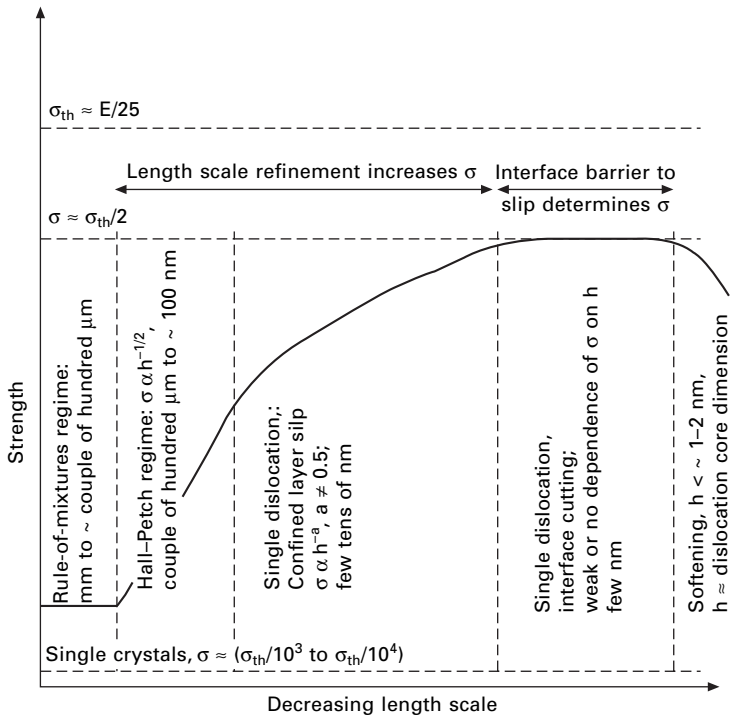


7.14 (a) SEM image of the surface of a Cu/Nb multilayer with an initial layer thickness of 15 nm rolled to 50% reduction; (b) schematic illustration of the shear cracks.

work hardening would initially tend to spread the slip along the interface, analogous to the picture shown in Fig. 7.12 for CLS except that dislocations deposited at the interface are now ($\mathbf{b}_{Nb}$-$\mathbf{b}_{Cu}$) for each slip transfer event, and not $\mathbf{b}_{Nb}$ or $\mathbf{b}_{Cu}$ as in CLS. However, as the repulsion from the residual dislocations is overcome, slip tends to be localized with continued interface crossing at the same location. Each slip transfer event creates a step at the interface with magnitude of the order of the edge component of the residual dislocation. As the magnitude of the step approaches the layer dimension (which is only a couple of nanometers), the layer is effectively fractured by shear.

7.7 Conclusions

Many of the concepts presented in this chapter are shown in a schematic diagram of the strength of layered composites as a function of layer thickness in Fig. 7.15. First, we note that the theoretical strength limit of most materials is of the order of $\mu/10$ where μ is the shear modulus. Also, since $\mu \approx 0.4$ E for most metals where E is the Young's modulus. Thus, theoretical strength limit (σ_{th}) in uniaxial loading is of the order of E/25, shown as a dotted line in Fig. 7.15. Single crystalline metals have yield strengths that are a factor of 1,000 to 10,000 lower than σ_{th}. Composites of two metals with layer thickness of millimeters to hundreds of micrometers would have strengths given the rule-of-mixtures, i.e., length scale independent behavior, as shown in Fig. 7.15. The design philosophy here is that increasing the volume fraction of the hard phase would increase the composite strength. As the layer thickness is decreased below a couple of hundred micrometers, strength becomes dependent on the length scale of the material via the Hall–Petch model, $\sigma \propto h^{-1/2}$. Dislocations pile up at the interfaces that are barriers to slip transmission and the stress concentration from the pile-up eventually overcomes the interface barrier to propagate slip through the composite. With reducing layer thickness, the pile-up length gets shorter and so does the stress concentration at the tip of the pile-up, thereby resulting in a higher applied stress for slip transmission. This continuum model is applicable to sub-micrometer length scales. As the layer thickness in a metallic laminate is decreased below a few tens of nanometers, the Hall–Petch scaling law is no longer applicable implying that continuum-scale dislocation pile-ups cannot form in nanoscale multilayers, and the deformation behavior must be interpreted in terms of the nucleation and motion of single dislocations. For such nanolaminates, strength increase with decreasing layer thickness in the few to a few tens of nanometers range is interpreted as confined layer slip of single Orowan loops (Fig. 7.15). As the layer thickness is decreased to a couple of nanometers, the strength reaches a peak value independent of layer thickness. This peak strength is interpreted as the stress needed to transmit



7.15 Schematic illustration of the evolution of strength of metallic laminate composites with the layer thickness. The mechanisms that operate at different length scales are shown.

a single glide dislocation across the interface, and can be as high as $\sigma_{th}/2$. The peak strength can be tailored by interface structure, e.g., degree of coherency, interface shear strength, etc. Finally, as the layer thickness decreases below the typical dislocation core dimensions of ~1 nm, the interface barrier stress to single dislocation transmission decreases.

The operation of single dislocation based deformation mechanisms, confined layer slip and interface crossing in nanolaminates also has implications on the post-yield behavior. Nanolaminates deformed to large plastic strains did not show any dislocation of cell structures, layered morphology was maintained (plastic stability) with uniform reduction in both layers, and even the interface crystallography (no significant lattice rotations about the interface plane normal indicating symmetric slip) was maintained. These effects have been discussed in terms of dislocation storage and recovery mechanisms at the interface.

7.8 Acknowledgments

This research was funded by the US Department of Energy, Office of Science, Office of Basic Energy Sciences. The author acknowledges discussions with H. Kung, R.G. Hoagland, J.P. Hirth, J.D. Embury, T.E. Mitchell, F. Spaepen, X. Zhang, D. Mitlin and M. Nastasi.

7.9 References

1. B.M. Clemens, H. Kung and S.A. Barnett, *MRS Bulletin*, **24**, 20, February (1999).
2. G.S. Was and T. Foecke, *Thin Solid Films*, **286**, 1 (1996).
3. S.A. Barnett and M. Shinn, *Ann. Rev. Mater. Sci*., **24**, 481 (1994).
4. *Scripta Materialia* Viewpoint Set No. 34 titled *Deformation and Stability of Nanoscale Metallic Multilayers*, edited by A. Misra, J.D. Embury and H. Kung, vol. 50, No. 6, (2004).
5. A. Misra and H. Kung, *Advanced Engineering Materials*, **3**, 217 (2001).
6. B. Huang, K.N. Ishihara and P.H. Shingu, *J. Materials Science Letters*, **20**, 1669 (2001).
7. K. Yasuna, M. Terauchi, A. Otsuki, K.N. Ishihara and P.H. Shingu, *Mat. Sci. Engg. A*, **285**, 412 (2000).
8. T. Yoshioka, M. Yasuda, H. Miyamura, S. Kikuchi and K. Tokumitsu, *Materials Science Forum*, **386–388**, 503 (2002).
9. S. Kikuchi, H. Kuwahara, N. Mazaki, S. Urai and H. Miyamura, *Mat. Sci. Engg. A*, **234–236**, 1114 (1997).
10. F. Bordeaux and R. Yavari, Z. *Metallkde*., **81**, 130 (1990).
11. T. Foecke and D.S. Lashmore, *Scripta Metallurgica et Materialia*, **27**, 651 (1992).
12. D.M. Tench and J.T. White, *J. Electrochem. Soc*., **138**, 3757 (1991).
13. S. Menezes and D.P. Anderson, *J. Electrochem. Soc*., **137**, 440 (1990).
14. W.D. Sproul, *Science*, **273**, 889 (1996).
15. S.A. Barnett, in *Physics of Thin Films: Mechanic and Dielectric Properties*, **17**, edited by M.H. Francombe and J.L. Vossen, 1 (1993).
16. B.J. Daniels, W.D. Nix and B.M. Clemens, *Thin Solid Films*, **253**, 218 (1994).
17. M. Verdier, M. Niewczas, J.D. Embury, M. Nastasi and H. Kung, *Mat. Res. Soc. Sym. Proc*., **522**, 77 (1998).
18. S.D. Dahlgren, *J. Vac. Sci. Technol*., **11**, 832 (1974).
19. R.F. Bunshah, R. Nimmagadda, H.J. Doerr, B.A. Movchan, N.I. Grechanuk and E.V. Dabizha, *Thin Solid Films*, **72**, 261 (1980).
20. R.L. Bickerdike, D. Clark, J.N. Easterbrook, G. Hughes, W.N. Mair, P.G. Partridge and H.C. Ranson, *Int. J. Rapid Solidification*, **1**, 305 (1984–85).
21. N.J. Petch, *J. Iron Steel Inst. London*, **173**, 25 (1953).
22. J.D. Embury and R.M. Fisher, *Acta Met*., **14**, 147 (1966).
23. J.P. Hirth and J. Lothe, *Theory of Dislocations*. Malabar, FL: Krieger (1992).
24. P.M. Anderson, T. Foecke and P.M. Hazzledine, *MRS Bulletin*, **24**, 27, February (1999).
25. P.M. Anderson and C. Li, *NanoStructured Materials*, **5**, p. 349 (1995).
26. L.H. Friedman and D.C. Chrzan, *Physical Review Letters*, **81**, 2715 (1998).
27. R.W. Armstrong, Y.T. Chou, R.M. Fisher and N. Louat, *Phil. Mag*., **14**, 943 (1966).

28. C.S. Pande, R.A. Masumura and R.W. Armstrong, *NanoStructured Mat*., **2**, 323 (1993).
29. J.D. Embury and J.P. Hirth, *Acta Met*., **42**, 2051 (1994).
30. W.D. Nix, *Scripta Mat*., **39**, 545 (1998).
31. W.D. Nix, *Mat. Sci. Eng*., **A234-236**, p. 37 (1997).
32. L.B. Freund, *Journal of the Mechanics and Physics of Solids*, **38**, 657 (1990).
33. A. Misra, J.P. Hirth and H. Kung, *Philos. Mag. A*, **82**, 2935 (2002).
34. M.A. Phillips, B.M. Clemens, W.D. Nix, *Acta Materialia*, **51**, 3157 (2003).
35. R.G. Hoagland, T.E. Mitchell, J.P. Hirth and H. Kung *Philos. Mag A*, **82**, 643, (2002).
36. S.I. Rao and P.M. Hazzledine, *Philos. Mag. A*, **80**, 2011 (2000).
37. J.S. Koehler, *Phys. Rev. B*, **2**, p. 547 (1978).
38. A. Misra, M. Verdier, Y.C. Lu, H. Kung, T.E. Mitchell, M. Nastasi and J.D. Embury, *Scripta Mat*., **39**, 555 (1998).
39. J. McKeown, A. Misra, H. Kung, R.G. Hoagland and M. Nastasi, *Scripta Materialia*, **46**, 593 (2002).
40. X. Zhang, A. Misra, H. Wang, T.D. Shen, J.G. Swadener, J.D. Embury, H. Kung, R.G. Hoagland and M. Nastasi, *Journal of Materials Research*, **18**, 1600 (2003).
41. X. Zhang, A. Misra, H. Wang, T.D. Shen, M. Nastasi, T.E. Mitchell, J.P. Hirth, R.G. Hoagland and J.D. Embury, *Acta Materialia*, **52**, 995 (2004).
42. N. Mara, A. Sergueeva, A. Misra and A.K. Mukherjee, *Scripta Materialia*, **50**, 803 (2004).
43. T.G. Nieh and J. Wadsworth, *Scripta Materialia*, **44**, 1825 (2001).
44. D.M. Tench and J.T. White, *Metall. Trans. A*, **15**, 2039 (1984).
45. S.L. Lehoczky, *J. Appl. Phys*., **49**, 5479 (1978).
46. H. Huang and F. Spaepen, *Acta Materialia*, **48**, 3261 (2000).
47. D. Josell, D. vanHeerden, D. Read, J. Bonevich and D. Shechtman, *Journal of Materials Research,* **13**, 2902 (1998).
48. A.T. Alpas, J.D. Embury, D.A. Hardwick and R.W. Springer, *Journal of Materials Science*; **25**, 1603 (1990).
49. Ebrahimi, F. and Kong, D. *Scripta Materialia*, **40**, 609 (1999).
50. S.P. Baker and W.D. Nix, *J. Mater. Res*., **9**, 3131 (1994).
51. H. Geisler, K.O. Schweitz, J. Chevallier, J. Bottiger and K. Samwer, *Phil. Mag. A*, **79**, 485 (1999).
52. K.O. Schweitz, J. Chevallier, J. Bottiger, W. Matz and N. Schell, *Phil. Mag. A*, **81**, 2021 (2001).
53. L. Geyang, X. Junhua, Z. Liuqiang, W. Liang and G. Mingyuan, *J. Vac. Sci. Technol. B*, **19(1)**, 94 (2001).
54. D.E. Kramer and T. Foecke, *Phil. Mag. A*, **82**, 3375 (2002).
55. G. Dehm and E. Arzt, *Applied Physics Letters*, **77**, 1126 (2000).
56. M. Legros, G. Dehm, R.M. Keller-Flaig, E. Arzt, K.J. Hemker and S. Suresh, *Materials Science and Engineering A*; **309**, 463 (2001).
57. D. Mitlin, A. Misra, V. Radmilovic, M. Nastasi, R.G. Hoagland, J.D. Embury, J.P. Hirth and T.E. Mitchell, *Philosophical Magazine*, **84**, 719 (2004).
58. R.G. Hoagland, R.J. Kurtz and C.H. Henager, Jr., Scripta Materialia, vol. 50/6 (2004) pp. 775–779.
59. A. Misra, H. Kung, D. Hammon, R.G. Hoagland and M. Nastasi, International Journal of Damage Mechanics, v. 12 (2003), pp. 365–376.

60. A. Misra, J.P. Hirth, R.G. Hoagland, J.D. Embury and H. Kung, *Acta Materialia* (2004), in press.
61. P.M. Anderson, J.F. Bingert, A. Misra and J.P. Hirth, *Acta Materialia*, Vol. 51/20 pp. 6059–6075 (2003).
62. K. Al-Fadhalah, C.N. Tomé, A.J. Beaudoin, I.M. Robertson, J.P. Hirth and A. Misra, *Philosophical Magazine*, under review.
63. C.H. Henager, Jr and R.G. Hoagland, *Scripta Materialia*, **50**, 701 (2004).

8

Preparation of monolithic nanocrystalline ceramics

G RIXECKER, Z BURGHARD and F ALDINGER,
Max Planck Institute for Metals Research, Germany and
L GAO, Chinese Academy of Sciences, China

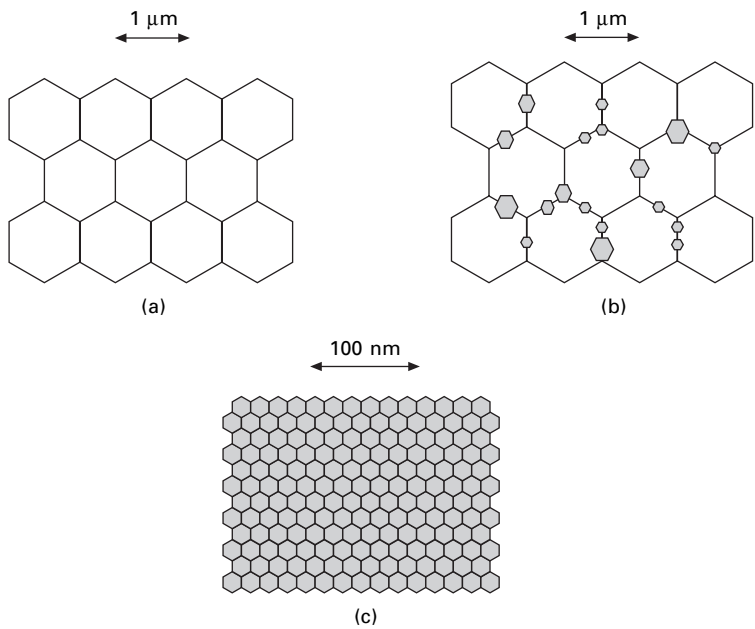
8.1 Introduction

As pointed out by Gleiter [1], the volume fraction of atoms residing in the grain boundaries of a polycrystalline solid, which is normally negligible, enters the percent range if the grain size is reduced to nanometer dimensions. Owing to the increase in average bond length and the reduction of the co-ordination number of atoms that are located in boundary planes, in general, the misorientation of the adjacent crystallites will lead to an excess free volume at the grain boundaries. Both modifications will also lead to a weakening of interatomic bonding forces, and it may be expected that numerous macroscopic properties of solids are affected [2]. This is the main reason for the widely accepted notion, that the term 'nanocrystalline material' should be used for solids with a mean grain size less than 0.1 μm. It is noteworthy, however, that 'microcystalline' or even single-crystalline solids may also contain high volume densities of nanometer-sized features. An example are two-dimensional lattice defects such as stacking faults and antiphase domains [3, 4], which may or may not cause macroscopic property variations.

In the case of ceramic materials, another important category of nanosized features are the amorphous grain boundary films which readily form in liquid-phase sintered materials (Fig. 8.1(a)). In liquid-phase sintering (LPS), a melt phase supplies pathways for matter transport. Upon solidification, the melt is converted to a wetting intergranular phase, with thin films one to several nanometers wide remaining at the grain boundaries. Such interfaces represent a volume fraction of 5–15% of the solid and have a strong impact on a variety of important properties of these materials. For example, the fracture behavior (transgranular or intergranular crack propagation) as well as the high temperature plasticity of liquid-phase sintered, silicon-based ceramics such as Si_3N_4 and SiC are controlled to a large extent by the presence or absence of vitreous grain boundary films [5, 6]. The same is true for the electrical properties of ferroelectric dielectrics and varistor materials based on $BaTiO_3$ and ZnO, respectively, and even typical technical qualities

of the most frequently used engineering ceramics, Al_2O_3, tend to contain such vitreous films owing to the presence of impurity or dopant elements.

Beside these prominent examples, a whole range of nanostructured features is present in the microstructures of many modern engineering materials. In micro/nano composites, which are microcrystalline materials containing a certain volume fraction (generally 5–30%) of nanometer-sized precipitates located either at grain boundaries or in the interior of the matrix grains (Fig. 8.1(b)), the precipitates have been demonstrated to enhance fracture properties by influencing internal stress states and crack propagation. Moreover, they may improve the creep resistance drastically by blocking grain boundary sliding [7]. Important composite systems, where these effects are exploited, are Al_2O_3/ SiC [8–10] and Si_3N_4/ SiC [11–14]. Beside these particle–particle systems, long-fiber reinforced ceramic matrix composites (CMC) form another important class of composite materials. It is quite clear, for both types of composites, that nanocrystalline matrices offer great advantages in terms of processing (lower temperatures, improved microstructural homogeneity) and potentially also fracture mechanical properties [15]. In the case of CMC materials reinforced by third-generation silicon carbide fibers, which exhibit



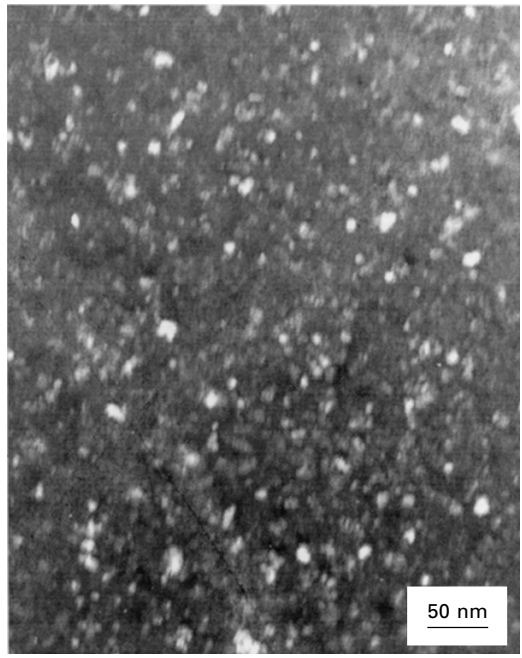
8.1 Types of superfine ceramic microstructures; (a) microcrystalline with nanometer-thick grain boundary phase; (b) micro/nano composite with intergranular precipitates; (c) nanocrystalline material.

high-temperature stability up to ~1800 °C, a highly dense, nanocrystalline matrix can be formed by combined pyrolysis of a polymer precursor and liquid-phase sintering of a nanocrystalline mixture of silicon carbide and oxidic additive powders [16].

Granular thin films and coatings may be synthesized by a variety of methods such as vapor phase deposition [17, 18], liquid phase deposition [19, 20], sol-gel methods [21, 22], pyrolysis of polymer precursors [23], and by sintering nanocrystalline powders [24]. They possess an inherently nanocrystalline grain structure which can have beneficial influences on film adhesion and wear properties of protective coatings and give rise to special electrical and magnetic effects as, for example, the weak localization of conduction electrons in half-metallic (La,Ca)MnO_3 films [25] and superparamagnetism in spinel-ferrite (MFe_2O_4) thin films [26, 27].

The remainder of this chapter, however, will be concerned mainly with monolithic materials that consist of crystallites in the size range around 100 nm and below (Fig. 8.1(c)). Traditionally, monolithic ceramics have been fabricated by powder technological methods, and the powder route will also be in the center of interest of the present discussion. It should be pointed out, however, that there are also versatile powder-free techniques for the fabrication of bulk nanoceramics. One of them is based on pre-ceramic polymers that are converted to inorganic networks upon thermolysis at intermediate temperatures (900 to 1400 °C). For instance, if polycarbosilanes, carbon-containing polysilazanes or polyorganoborosilazanes are utilized as molecular precursors, the thermolysis products attain compositions of Si-C, Si-C-N, or Si-B-C-N [28]. Further heat treatment of these amorphous materials, which tend to have extraordinarily high thermal stability, leads to microstructures consisting of crystallites of the equilibrium phases (e.g., SiC and Si_3N_4) which are typically below 50 nm in size [29] and may be embedded in amorphous or turbostratic matrices (Fig. 8.2). Owing to the largely covalent bonding in the oxygen-free systems, these materials are expected to have very high creep resistance in spite of their ultrafine grain size [30].

Nanocrystalline glass ceramics with oxidic composition represent a direct microstructural counterpart of the precursor-derived materials. In this case, the amorphous state is obtained by quenching from the melt, and controlled crystallization subsequently leads to nanocrystals embedded in a residual amorphous phase. Technologically important glass ceramics pertain to various material systems and are in use or being considered for a multitude of functional applications [31–33]. For example, the microstructural homogeneity and isotropy on length scales comparable to the wavelength of light leads to good optical transparency; the extremely low value of thermal expansion in some lithium aluminium silicate (LAS) compositions has prompted their use in applications ranging from boilerplates and cookware (owing to excellent thermal shock resistance) to large telescope mirrors; and glass ceramic surfaces



8.2 Si-C-N ceramic synthesized by polymer thermolysis and subsequent crystallization for 240 h at 1400 °C in a compression creep experiment, consisting of Si_3N_4 and SiC crystallites and turbostratic C in a residual amorphous matrix. Dark field TEM image courtesy of A. Bauer.

can be polished to extreme smoothness for applications as high-damage-threshold laser mirrors. Glass ceramics doped with transition metal ions lend themselves for non-linear [34] and laser optical applications, and ferromagnetic crystallites embedded in a glass matrix show ultrafine-particle effects such as superparamagnetism [35].

For categorizing ceramics in terms of fineness of microstructure, an argument related to their synthesis by powder technology gains additional weight. In a typical manufacturing process a powder compact, called greenbody, is densified by sintering, that is by thermal treatment at a temperature where solid state diffusion is activated. Alternatively, a few volume per cent of melt phase are present in liquid-phase sintering. Since grain growth is equally caused by diffusional processes, sintering will generally be accompanied by coarsening of the microstructure. In practice, this leads to a lower limit in excess of 100 nm for the grain size that can be achieved in dense bodies, irrespective of the fineness of the starting powders, if standard processing methods are used. However, by optimized processing techniques, dense sintered microstructures with grain sizes below 50 nm have recently been obtained [37–39].

If ultrafine powders with grain sizes below 100 nm are to be used for sintering, avoiding grain growth becomes particularly challenging because its driving force is proportional to the curvature of the particle surfaces, thus inversely proportional to the grain size (cf. section 8.4.1). Fortunately, there are several strategies available which permit, at least to some extent, decoupling of the kinetics of sintering and grain growth. Among the sintering methods that have gained a lot of attention recently are pressure-assisted techniques such as uniaxial hot pressing with sample heating by pulsed electric fields (pulsed electric current sintering) [40, 41] and hot pressing without a lateral constraint on sample deformation (sinter forging) [42]. On the other hand, by using careful pre-processing of the nanopowders and special time-temperature protocols, pressureless sintering has also been demonstrated to allow retention of a nanocrystalline structure in fully densified compacts both in solid-state sintering [43] and in liquid-phase sintering [44, 45].

8.2 Synthesis of nonmetallic-inorganic nanoparticles

Although it was only several decades ago that nanoparticles started to attract the attention of researchers, they have existed in the world since long ago. In fact, many natural things contain nanometer-scale features, for example human teeth, bones and sea algae. Micro-organisms, and also the natural weathering of iron oxides and silicate minerals, can produce colloids. Meteorites may even contain nano-diamonds [46, 47].

The history of nanoparticles synthesis and utilization by man can be traced back to over one thousand years ago. Ancient Chinese produced ink based on ultrafine powders by collecting smoke resulting from candle burning. The anti-corrosive layer on ancient copper mirrors is, in fact, a thin nano-SnO_2 film. However, systematic studies on nanoscale materials started only some decades ago. The first artificial nanopowder syntheses were reported in Japan in the 1960s. Nanostructured metal particles were obtained by using a gas-condensation method, i.e., an evaporation and condensation process in an inert gas atmosphere. From then on, the preparation of nanopowders received more and more attention, motivated by their potentially useful properties in many technological respects.

8.2.1 Methods capable of producing larger quantities of powders

Most of the raw materials used today in the ceramic industry are obtained by conventional ceramic methods, which involves the three consecutive steps of mixing, solid state reaction, and milling. For example, Ni-Zn ferrites are produced by mixing stoichiometric amounts of zinc, nickel and iron oxides

or carbonates, and then firing the mixture at 1000 °C to finish the solid-state reaction. Finally, the resultant ferrite solid is milled to reduce the particle size [48]. During the milling process, grinding aids are often added to modify the surface of the powders and prevent them from re-agglomeration [49]. In spite of the simplicity of this method, the produced powders are on the order of micrometers. They are agglomerated, which necessitates high sintering temperatures and significantly affects the properties of the final sintered bodies.

In the past several decades, the outstanding properties of ultrafine particles were a stimulus for finding new effective methods of nanopowder preparation. The primary goal of particle synthesis is to develop methods to control size, assure clean surfaces and obtain non- or weakly agglomerated powders in large quantities. Up to now, there are basically two classes of synthetic techniques for nanocrystalline powders, namely, physical and chemical methods. Physical methods currently in use mainly include physical vapor deposition (PVD) and high-energy ball milling.

High energy ball milling is able to prepare nanocrystalline particles through collisions and shearing interactions between the hard balls and the raw materials, which are brought about by rotational or vibrating movement of the milling vial. Starting from the late 1980s, the method has been used successfully to synthesize nanopowders of metals, alloys, ceramics, ceramic-metal composites, etc. [50–52]. The main disadvantage of high energy ball milling is the introduction of impurities into the product resulting from contamination during the milling process and the inhomogeneous size distribution of the obtained powders. But the high yield and the simple procedure of the process make it a widely employed method to prepare nanopowders. In addition, this method can yield high melting point metals and alloys which cannot be produced by other traditional methods.

Chemical methods have played a major role in developing new materials with novel and technologically important properties; their main advantages are the versatility of chemical syntheses in designing new materials and the fact that chemical processes offer good homogeneity due to mixing at the molecular level [53]. The chemical synthesis methods mainly involve chemical vapor deposition, pyrolysis, precipitation, sol-gel, and hydrothermal processes.

In chemical vapor deposition (CVD), a precursor is converted to the gas phase and it then undergoes decomposition at either low or atmospheric pressure to generate the nanostructured particles. These particles are then transported out of the reaction zone by a carrier gas and collected on a cold substrate. Depending on the material and reaction conditions, amorphous, polycrystalline or single-crystalline powders can be prepared. CVD processes can be used to produce a variety of ceramic oxides and non-oxides. For instance, silicon nitride powders are prepared by CVD using silane (SiH_4) and ammonia (NH_3) mixtures at a deposition temperature of 1000 °C [54].

The preparation of a promising new material which has great technological potential – carbon nanotubes – by the CVD method is currently receiving a great deal of attention [55]. The powders prepared by vapor methods often have a small size, narrow size distribution, high purity and good dispersibility.

Pyrolysis may well be the most frequently used method for the preparation of nanopowders. During pyrolysis, chemical precursors decompose into one solid compound under proper thermal conditions, while unwanted waste products evaporate. Frequently used precursors are metal carbonates, oxalates, carbonyls, nitrates, citrates, and alkoxides. To get uniform nanometer-sized powders, various modifications have been made to the pyrolysis process, for example, using decomposable additives to disperse and protect precursors and as-prepared nanoparticles, atomizing the precursor solutions, etc. In recent years, advanced heat sources such as lasers, ultrasound, electric discharges, microwaves, and plasmas, were employed to heat up the precursors or increase the reaction temperature locally. Nanosized Al_2O_3, TiO_2, and ZrO_2 powders have been successfully prepared using this method. An important non-oxide ceramic compound, SiC, has been synthesized by pyrolyzing poly(methylsilasesquicarbodiimide) to nanoscale silicon carbonitride at 700 °C, followed by a crystallization step at >1400 °C [56]. Synthesis by pyrolysis is a powerful method in many specific systems. It has the advantages that the reactions are easy to control and that the products are of high purity. The drawback is that the high temperature may lead to wide size distributions and particle agglomeration [57].

Chemical precipitation is widely used in industry and research to synthesize nanometer-sized oxide powders. The process starts from dissolving salts, which contain metal ions, in a liquid medium. The solution is then mixed with a solution of dissolved precipitating agent, such as oxalic acid or ammonium hydroxide, in order to precipitate the metal oxalates or hydroxides. The final crystalline oxide is then obtained by firing the precipitates at a higher temperature. By the *in-situ* generation of precipitating agents throughout the liquid phase, for example, by thermal decomposition of urea, the precipitation method can be extended to homogeneous precipitation, which is useful to prepare monodispersed nanoparticles. Co-precipitation may be brought about by using salt solutions containing mixed metal ions, a method widely used to synthesize complex oxides such as $BaTiO_3$ and $Y_3Al_5O_{12}$ (yttrium aluminum garnet, YAG), and others [58].

A significant problem which has to be to overcome in chemical precipitation is the agglomeration of the particles in the solution. Thus, dispersants are often added in the reaction process to provide repulsive (electrostatic or steric) interactions between the particles to prevent them from adhering to each other. To eliminate subsequent neck formation and aggregation in the calcination and firing steps, organic solvents such as ethanol may be added in the final washing step to replace water adsorbed on the powder surface.

The advantages of the precipitation method include technical simplicity, low manufacturing cost, high reproducibility and fine particle size. Disadvantages are the difficulty to control the final chemical composition of the products and upscaling issues. Also, the repeated washing and separation steps make precipitation a time-consuming method.

Sol-gel techniques have been extensively used for the production of nanophase powders during the past three decades. In a sol-gel process, the inorganic or organic precursors are subjected to hydrolysis and polycondensation in organic solvents and then, by gelling, form an interconnected 3-D network. The gel is dehydrated and calcined at elevated temperatures to form nanopowders. The method is not only applicable for materials with oxide compositions, but also permits the production of new hybrid organic-inorganic materials which do not exist naturally. The sol-gel process is very promising in preparing multi-component powders since the chemical homogeneity can be controlled at an atomic scale. The greatest limitation of the sol-gel processing is the high cost of the precursors, and especially that of alkoxides [59].

Hydrothermal synthesis is another promising chemical method for the preparation of nanosized powders. In this case, a suspension or sometimes a homogeneous solution is subjected to elevated temperatures (<1000 °C) and pressures (<100 MPa). The combined effect of pressure and temperature can reduce the free energy of phases that are not stable under atmospheric conditions. Hydrothermal synthesis is always interpreted as a dissolution-reprecipitation process. The suspended particles are dissolved in the solution. Eventually, the solution is supersaturated, which causes nucleation and growth into crystalline particles with a specific size. Solid powders and freshly prepared gels are often used as the precursors for hydrothermal synthesis. Hydrothermal processes can yield highly crystalline ceramic powders directly and do not require high temperature calcination steps. Therefore, hydrothermal synthesis is an energy-saving method. In addition, powders prepared by hydrothermal processes have a small size and good sinterability. Size, shape and crystalline phases can be easily controlled by adjusting the experimental parameters [60]. However, hydrothermal synthesis demands high temperature and high pressure equipment [61].

8.2.2 Powder characteristics that are important for bulk properties

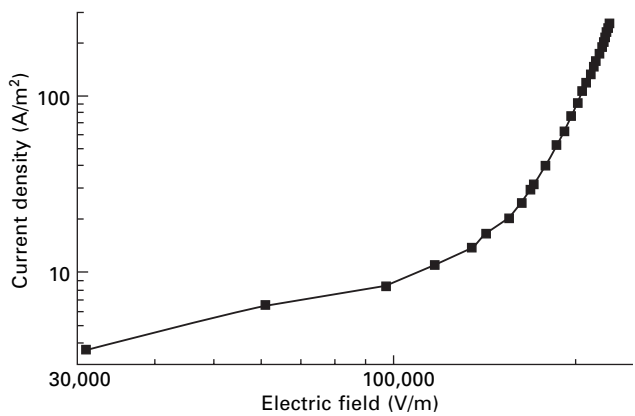
The bulk properties of ceramics are closely related to the characteristics of the initial powder used. Among them, particle size, particle size distribution, chemical purity, and degree of agglomeration are important factors [61, 62]. Fine particles have a high surface area and can be densified at lower sintering temperatures than large particles. The sintered ceramics often have a small

grain size and good mechanical and functional properties. For example, ZnO ceramics with a grain size of ~100 nm are obtained using nanopowder as initial material. In contrast to conventional pure ZnO ceramics sintered from sub-micrometer powders, the as-prepared nanoceramics exhibit nonlinearity of the electrical conduction [63] (Fig. 8.3).

The shape of the particle size distribution also plays an important role for the properties of the final product. Higher green densities can be achieved with powders having a narrow size distribution, although in certain cases the highest packing density is achieved when two types of particles are combined to give a binary size distribution. Such green bodies with a high density are easy to be sintered at lower temperatures, which significantly inhibits abnormal grain growth. The strength of advanced ceramics is sensitive to flaws in the microstructure, which serve as fracture origins. Common defects are agglomerates in the powder. Therefore, milling to break up the agglomerates is a most useful processing technique.

Chemical purity also affects the bulk properties of ceramics. During the sintering process, the impurities may diffuse into the grains, forming separate phases or segregating to the grain boundaries. Impurities that diffuse into the grains are not particularly harmful to mechanical properties but may affect optical or electronic properties. Separate impurity phases are often facture origins that lower the strength of ceramics, and the glassy phase at grain boundaries may lower the creep resistance. In some cases, however, impurities can act as sintering aids and reduce the sintering temperature. For example, alumina-yttria-silica glasses are used for the sintering of Si_3N_4 ceramics.

In one word, ideal physicochemical characteristics of a ceramic powder include ultrafine particle size, narrow size distribution, minimum degree of agglomeration, and a high degree of purity. Ceramics sintered from such powders often exhibit good mechanical and functional properties.



8.3 Voltage and current characteristics of pure nanocrystalline ZnO ceramics densified by pulsed electric current sintering at 500 °C [63].

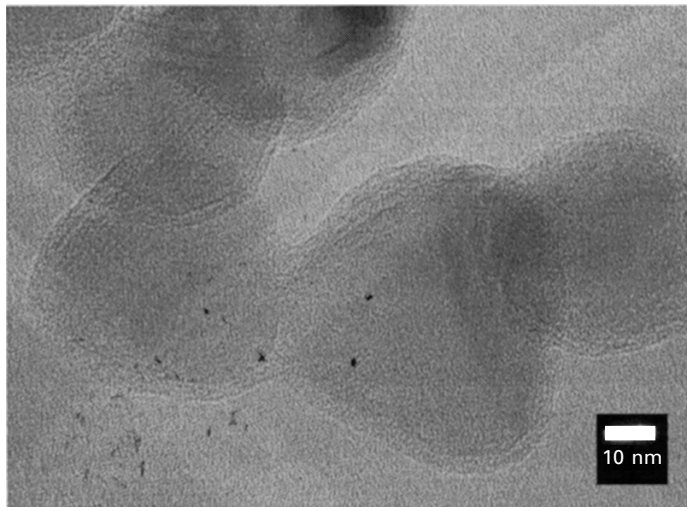
8.3 Green shaping of nanoceramic bodies

The processing of non-metallic ultrafine powders to form ceramic compacts has mostly followed the routes previously used for micrometer-sized powders, including the application of standard green forming methods such as uniaxial pressing [64–74], cold isostatic pressing (CIP) [64, 69, 75, 76], slip casting [77, 78], pressure filtration [79–82] and electrophoretic deposition (EPD) [83, 84]. However, there are several items which require special consideration in the case of nanocrystalline powders.

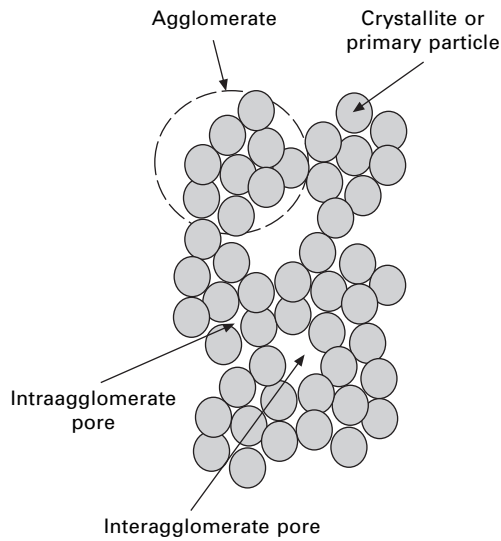
8.3.1 Specific problems in nanopowder processing

In the first instance, a strong tendency for agglomeration exists in nanopowders, which is induced by the Van-der-Waals forces acting between the individual particles. The morphology of such agglomerates may vary from chain-like (1-D) to heavily aggregated (3-D). This alone would not constitute a major problem since soft agglomerates held together by Van-der-Waals forces only are easily broken down in processing steps such as dry pressing, dispersion in a liquid medium or ultrasonication. However, additional chemical interactions between the contacting particles oftentimes lead to hard agglomerates which resist standard processing routines. Rigid particle–particle contacts may be formed due to partial hydrolysis or passive oxidation of the powder surface; in the worst case, neck formation by diffusion may have occured at elevated temperatures during powder synthesis already. The term ‘aggregate’ has been introduced especially to describe hard agglomerates in which the particles are joined by sintering necks (Fig. 8.4). The internal friction caused by these hard, fractal-like agglomerates may be exceedingly high and prevent an efficient packing of the nanoparticles during compaction, resulting in green densities that are considerably lower than those obtained for micrometer-sized particles and in inhomogeneities that typically occur on the length scale of agglomerates.

The uniformity of particle packing has a pronounced influence on the densification during sintering. It has been argued that the final grain size that may be obtained by free sintering is limited by the starting agglomerate size rather than the primary particle size, due to the much faster sintering kinetics inside the agglomerates (Fig. 8.5). The following discussion draws heavily on the excellent review article by Mayo [85]. Her conclusion, that fully dense ceramics with grain diameters below 100 nm can only be obtained by pressureless sintering if non- or weakly agglomerated powders are used, is based on the experimental evidence up to 1996 [70, 72, 73, 86, 87]. It is worthwhile mentioning that it remains fully valid if more recent literature is taken into account [39]. Although high-quality compacts derived from non-agglomerated nanopowders still have lower densities than conventional



8.4 TEM image showing vapor phase synthesized SiC nanopowder with a SiO_2 layer up to 5 nm thick. The partly hydrolyzed surface oxide forms rigid necks between the particles.



8.5 Schematic sketch of an agglomerated powder [85]. Elimination of the larger interagglomerate pores requires longer sintering times and/or higher temperatures than elimination of the small interagglomerate pores.

greenbodies – for reasons to be discussed below – such green compacts may show the remarkable property of being transparent if all inhomogeneities are smaller than about 100 nm. This has been demonstrated for different materials

that are transparent as single crystals: CeO_2 [39], ZrO_2 [87], Y_2O_3 [88], and Si_3N_4 [89].

A second specific problem associated with nanocrystalline starting powders is the high level of hydroxide and, in the case of non-oxides, oxygen impurities resulting from the large specific surface area. Since the chemical potential inceases as the radius of curvature becomes small, the thickness of such surface layers is oftentimes larger than in coarse powders of the same chemical composition [8].

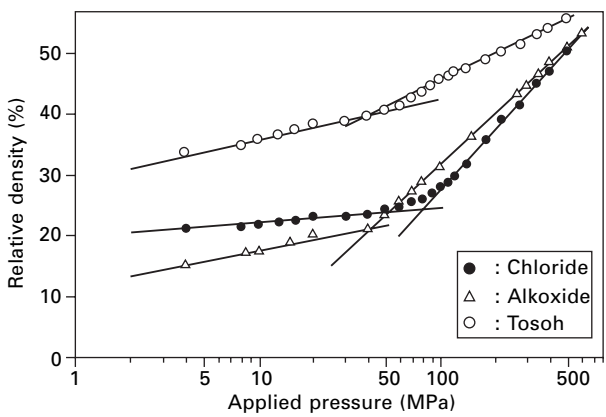
There are numerous examples of extra processing steps that were introduced in view of these specific difficulties of nanopowder processing. For example, either a relatively high oxygen content or free carbon is usually present in SiC powders synthesized by chemical vapor condensation, depending on the reaction conditions. The former situation was successfully dealt with by washing oxygen-rich SiC powder with 50 nm average particle size with concentrated hydrofluoric acid. Thereby, the oxygen content could be reduced from about 10 wt% to less than 1 wt%, and a substantial reduction in the degree of agglomeration was also obtained [90]. On the other hand, the high amount of free carbon in a vapor phase synthesized powder of 90 nm particle size could be reduced from 4.1 to 1.9 wt% by centrifugal sedimentation, which also had the effect of removing large agglomerates [91]. Large shearing forces, as they are required for the cleavage of sintering necks between particles, can be introduced by high-energy milling [65]. Accordingly, significant increases in green density were achieved by a short (10 min) planetary ball milling treatment, e.g., from 58% to 65% for Al_2O_3 powder with a median particle diameter of 30 nm, which was uniaxially compacted at 440 MPa. While centrifugation, ultrasonication and wet milling of nanopowders were proposed as viable methods to separate or break up agglomerates [7], the drying method of aqueous slurries was identified as a critical step to avoid their formation, implying that hydroxide bridges between grains are an important reason for their occurrence. Freeze drying [7] and, in particular, replacing the water by an organic solvent such as ethanol prior to drying [73, 87] were identified to be crucial processing steps.

8.3.2 Shape forming by dry pressing methods

There are further complications in the shaping of nanoparticles, which are not related to either agglomeration or impurity content. For example, in the case of uniaxial pressing [66, 71, 85], a large number of particle–particle contacts exists per unit volume, each of which is a source of frictional forces. Consequently, there is less sliding among the particles for a given applied stress and the resistance to compaction is much larger than for micrometer-sized powders. With unagglomerated 3Y-ZrO_2 (ZrO_2 doped with 3 mol% Y_2O_3 for stabilization of the tetragonal polymorph) nanopowders having a

narrow size distribution, uniaxial compaction at 483 MPa led to a green density of 45 ± 2% [86]. Very large applied stresses in the GPa range can be utilized to overcome this [37, 67, 69, 73, 74, 89, 92]; at a pressure level of ~1 GPa, green densities of 51% for 6–13 nm ZrO_2 particles [67] and 62% for 14.8 nm TiO_2 particles [74] were achieved. A TiO_2 powder, with 20 nm crystallite size and 50–80 nm agglomerate size, was uniaxially pressed at up to 1.14 GPa [73]. In addition to a shift of 400 K between the sintering curves of this nanocrystalline powder and microcrystalline titania, a pronounced correlation was found between the green density and the final grain size at a density of ~98%: for green densities of 47%, 50% and 58%, the final grain size was restricted to 165 nm, 115 nm and 80 nm, respectively. However, these high uniaxial pressures induce large residual stresses which may lead to cracking during unloading [64] or sintering. Lubrication is possible to some extent, e.g., by liquefying an adsorbed nitrogen layer at cryogenic temperature [89]. Lubrication by process additives has also been proposed [87] but is not likely to be very practical because the large specific surface makes it difficult to wet a sufficiently large fraction of the surface with an acceptable amount of organic additives. Therefore, cold uniaxial pressing turns out to be one of the least satisfactory compaction methods for ceramic nanopowders.

On the other hand, cold isostatic pressing (CIP) has been shown to lead to more homogeneous particle packing than uniaxial pressing at the same level of applied stress. Although it is generally limited to about 600 MPa by commercially available instrumentation, CIP has been used frequently and successfully in the manufacture of nanocrystalline ceramic bodies (Fig. 8.6)



8.6 Cold isostatic compaction behavior of different nanocrystalline, weakly agglomerated ZrO_2 powders [76]. The soft agglomerates are broken down in a pressure range between 50 and 100 MPa. Note that the x-axis is scaled logarithmically, i.e., the compaction efficiency decreases with increasing pressure.

[64, 69, 73, 75, 76]. In an attempt to quantify the advantages of CIP over uniaxial pressing, X-ray tomography was applied to reveal the spatial distribution of green density inside compacts of 6Y-ZrO_2 with 50 to 100 nm particle size [65]. The width of the density distribution of uniaxially pressed green bodies increased as the density increased, and high uniaxial pressure led to the formation of cracks upon unloading. The density distribution after CIP was much narrower, but CIP, too, did not fully transmit the pressure into the core of the sample. The mean density of crack-free compacts could not be increased above 47% by either uniaxial pressing or CIP, and combining uniaxial and isostatic pressure components by applying the two techniques sequentially proved not to be advantageous. The result, that a good sintering behavior may be obtained by CIP although the green densities are rather moderate, has been corroborated for SiC particles containing 1 wt% of carbon and boron sintering additives [75]. In this case, a pronounced dependence on the size range of the particles was also reported, with particle sizes of 15–20 nm and 25–75 nm leading to green densities of 38–43% and 56–57%, respectively. In the case of a 3Y-ZrO_2 powder with 15 nm particle size and only soft agglomerates [73], CIP led to a monomodal pore size distribution centered about 4 nm. Under these favorable conditions, a sintered density of 99.9% and a grain diameter of 85 nm were reached.

8.3.3 Wet processing and wet shaping methods

For conventional ceramic powders, it was established that wet shaping techniques such as tape casting, slip casting, pressure filtration and centrifugation typically produce a more homogeneous particle packing as compared to dry pressing methods, leading to green bodies with better sinterability. This is despite a slightly lower green density, and despite the fact that the effective compaction pressures are lower by about two orders of magnitude [85]. On the other hand, wet shaping requires that the particles be dispersed in a slurry that has to be stabilized against sedimentation utilizing either electrostatic repulsion between the particles or steric hindrance due to adsorbed macromolecules. Combining both principles leads to so-called electrosteric stabilization, which is efficiently accomplished with polyelectrolyte dispersants such as polyacrylic acid (PAA), polymethacrylic acid (PMAA), ammonium polymethacrylate (PMA-NH_4^+) and polyethylenimine (PEI). This concept was proven for a variety of sub-micrometer particles in aqueous suspensions, like ZrO_2, 3Y-ZrO_2, Al_2O_3, SiO_2, and SiC [93–97]. High solid contents (e.g., 60 vol% for an α-Al_2O_3 slurry [94]) and good long-term stability against sedimentation (2–3 weeks for a ZrO_2 slurry containing 40 vol% particles [93]) can be reached in this way. The maximum solid content is determined by the increasing viscosity of a highly loaded suspension. For most casting operations, the viscosity cannot be higher than 1 Pa·s.

More recently, the principle of electrosteric stabilization has been successfully transferred to nanocrystalline powder suspensions with sub-100 nm particle sizes [77, 82, 98–101]. However, there is a limiting factor for the stabilisation of colloidal suspensions. Since the interaction distance of the attached macromolecules remains the same as the particle size is reduced [99], the maximum solid loading that can be achieved is substantially lower than for sub-micrometer particles. This has been verified experimentally. For example, ZrO_2, $BaTiO_3$ and TiN suspensions in the particle size range from 20 to 60 nm already showed shear thinning and viscosities in the vicinity of 1 Pa·s at solid contents above 10 vol% [77, 82, 100, 101]. Shear thinning, i.e., the deviation from Newtonian flow behavior, which is characterized by a high viscosity of the stagnant suspension and decreasing viscosity at increasing shear rate, typically occurs close to the maximum solid loading of powder suspensions. There are few examples where the solid loading could be increased to higher values: a suspension with 23.8 vol% solid loading, containing 30 nm SiC particles and the dispersant PEI, still exhibited Newtonian flow behavior if the initial degree of ionization of the polyelectrolyte was chosen appropriately [98]. For a 20 nm SiO_2 suspension dispersed with PEI, Newtonian behavior was retained up to 25 vol% [99].

While chemisorbed macromolecules will also represent a large volume fraction of the compacted green body due to the large specific surface, electrostatic repulsion between charged particle surfaces – most commonly by pH-dependent protonation or de-protonation of OH groups – also leads to colloidal stabilization and has the advantage of introducing a smaller amount organic additives. It facilitates the sliding of nanoparticles past each other, enabling them to pack more densely during consolidation. While electrostatic stabilization is quite straightforward in aqueous suspensions [39], there are also concepts for non-aqueous suspensions which were designed especially to circumvent the problem of hydrolysis of the particle surfaces. For example, electrostatic dispersion of Si_3N_4 particles with a size on the order of 20 nm was achieved in an aprotic solvent, N,N-dimethyformamide (DMF), at high Lewis basicity brought about by adding 1 wt% of the organic base triethanolamine (TEA) [102]. Extended pre-soaking times in the DMF/TEA solution of up to one month markedly improved the stability of the suspensions against sedimentation.

In wet shaping, compaction may be brought about by removing the dispersion medium from the slurry, such as in tape casting or the slurry casting process described in ref. 103. By this somewhat simplistic approach, which consists in casting a ceramic suspension into a rubber mold and vacuum drying, a relative density of 97.4% was reached upon microwave sintering, starting from an aqueous suspension of 3Y-ZrO_2 powder with 40 nm particle size. For comparison, only 95.2% of the theoretical density could be reached upon uniaxial pressing up to 500 MPa [103]. As an alternative to drying,

several other forming methods such as gelcasting, freeze casting, direct coagulation casting (DCC), hydrolysis-assisted solidification (HAS), and temperature-induced forming (TIF), are based on modifying the physical or chemical characteristics of highly loaded suspensions in such a way that solidification takes place even in the presence of the dispersion medium [104]. However, owing to the limited solid content of nanoparticle slurries, all of these techniques have the drawback that substantial shrinkage occurs during drying, which causes high internal stresses and usually leads to the destruction of the green bodies. A third class of forming methods, which includes slip casting, pressure filtration, electrophoretic deposition, centrifugal consolidation and osmotic consolidation [39, 72, 73, 77–79, 81–86, 105, 107–109], are based on the deposition of green sheets or filter cakes under the action of gravitational, capillary, electrostatic, centrifugal, or osmotic forces. All of these techniques, except osmotic consolidation, have served as forming methods in non-nanocrystalline ceramic processing; their capability of dealing with intermediate to low particle concentrations constitutes a clear advantage with regard to nanoparticles.

Shaping of nanopowders by slip casting in macroporous plaster of Paris molds was demonstrated for polyelectrolyte-stabilized aqueous slurries [77, 78]. Particle packing seems to be better for higher solid loadings since the green density was seen to be roughly proportional to the powder content of the slurry; at a maximum solid loading of 40 vol%, a green density of 52% was obtained. In the study [79], pressure filtration at 5 MPa piston pressure was applied to a dilute (5 vol%) slurry of ultrafine particles (8 nm) of 3Y-ZrO_2. Very uniform packing and a green density of 40% was reported, comparable to samples dry pressed under 150 MPa. After sintering, a relative density of 99% and particle sizes ≤ 150 nm were obtained, whereas bimodal particle size distributions occurred in sintered bodies based on dry pressed compacts. Green compacts with *circa* 50% relative density were produced by filter pressing of 30–40 nm TiN particles [82]. Sintering of the green compacts to 96% of the theoretical density was possible at 1200 °C with negligible grain growth, but further densification was accompanied by exaggerated grain growth (to 400 nm at 1400 °C).

Superior pressureless sintering behavior of 3Y-ZrO_2 upon centrifugal consolidation was demonstrated, with a sintered density of 99.9% and a grain size of 80 nm, which was comparable to the results of 99.9%, 85 nm obtained by uniaxial pressing at 0.5 to 1 GPa [72, 73, 85]. The work was performed with the aim to corroborate and extend earlier investigations using ZrO_2 [86] and TiO_2 [105] powders with especially narrow size distributions, which were equally obtained by centrifugal sedimentation. In another recent example, the sinter activity of nanocrystalline CeO_2 with 8.4 nm crystallite size according to X-ray diffraction (XRD) and 40 nm agglomerate size according to dynamic light scattering was studied [39]. Compaction by

centrifugation led to optically transparent green bodies with a truncated pore size distribution, i.e., there was no appreciable fraction of pores larger than a few nanometers present in the samples. Compared to CIP at 200 MPa, colloidal processing led to the same density values at about 100 K lower free sintering temperatures. By comparison with literature values, the sintering temperature of the centrifuged nanopowder was 750 K lower than that of uniaxially pressed microcrystalline ceria powders. For CeO_2 doped with 7 (10) mol% of Y_2O_3, a sintering temperature of 900 °C was required to obtain relative densities of 99.1 (98.5)% at 35 (25) nm XRD grain size while retaining the optical transparency. From the dependency on the Y_2O_3 content, the unusually small grain sizes of the fully dense ceramics were concluded to be stabilized by solute drag. The conclusion, that grain growth in the last sintering stage is controlled by impurity drag, was also reached in other studies [70].

Electrophoretic deposition of ceramic particles on a substrate under the action of a static electric field is a versatile forming method for both aqueous and non-aqueous powder suspensions [106]. In the latter case, polar solvents have to be utilized in order to generate adequate electrostatic forces. On the other hand, special precautions must be taken to obtain bubble-free deposits from aqueous suspensions if deposition voltages above the electrolysis voltage of water (~1.2 V) are used. For example, the use of hydrogen-absorbing palladium cathodes has been demonstrated to yield high-quality green films from suspensions of 40 nm ZnO particles stabilized with PEI [83]. Another promising approach is the deposition of green bodies on an ion-permeable membrane (dialysis membrane), which is spatially separated from the electrode [84]. In this way, nanoscale SiO_2 (fumed silica) and ZrO_2 components were made. From suspensions with 30 wt% solids content, homogeneous green bodies with densities of 39.4% in the case of SiO_2 and 32% in the case of ZrO_2 were obtained. By sintering the fumed silica green bodies, fully dense and fully transparent silica tubes were obtained. In the case of ZrO_2, the relatively low green density led to cracking problems during sintering owing to the very large shrinkage [84]. Electrophoretic deposition was also successfully applied for the coating of carbon and stainless steel fibres with nanoscale Al_2O_3 or TiO_2 powders [107], for the particle-infiltration of ceramic matrix composites [108, 84], and it was used in combination with metal electroplating to co-deposit reinforcing ceramic particles such as SiC [109].

Osmotic consolidation has not been used or investigated extensively up to now, although it was successfully applied for the compaction of ZrO_2 particles in a size range from 8 to 140 nm and 1.3 μm Al_2O_3 particles [81]. 8 nm ZrO_2 particles packed to 47.6% green density, 139 nm particles to over 55%, and coarse Al_2O_3 powders packed to over 70%, which is an outstanding result. A dialysis membrane of 6000–8000 Dalton (~2 nm pore size) was used together with a highly concentrated polymer solution of PEO in water, which displays a very high osmotic potential. The equilibration time was 12 h, and the

polymer solution was exchanged one or two times towards the end of the equilibration time in order to have a defined osmotic potential. The effective pressure was calculated to be 12 MPa, which is comparable to pressure filtration and about 70 times higher than in slip casting. However, the cohesion of the green bodies from osmotic consolidation was better than in the case of pressure filtration.

8.4 Densification

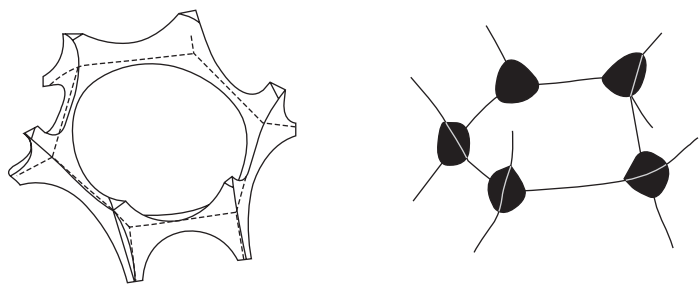
8.4.1 Sintering strategies for limited grain growth

Since matter transport by diffusion is responsible for sintering as well as grain growth, the driving forces for both phenomena are the same to a first approximation. In the case of nanocrystalline materials, in particular, the driving force for grain growth is very high since it depends directly on the convex curvature of the grains. As a consequence, the rate of grain growth may be assumed to be inversely proportional to the grain size:

$$dr/dt \propto 1/r^n,\ n \approx 1 \qquad 8.1$$

The question arises whether, in practice, grain growth can be kinetically decoupled from densification. In solid state sintering, open porosity exists in the early and intermediate stages of sintering. Grain growth is impeded in the beginning of densification because the particles are only connected by isolated grain boundary areas ('sintering necks'). In the intermediate stage of sintering, the porosity still forms a network of interconnected channels which represent an obstacle to triple point movement (Fig. 8.7(a)). The transition from open (cylindrical, one-dimensional) to closed (spherical, zero-dimensional) porosity (Fig. 8.7(b)) takes place at a density of about 92%. Thereafter, rapid grain growth starts due to the loss of grain boundary pinning at the pores.

Obvious measures against accelerated grain growth would be to restrict the time spent in the final sintering stage to a minimum, or to lower the



8.7 Pore geometries (a) in the intermediate and (b) in the final stage of solid state sintering.

temperature during the final stage in order to decrease the driving force for grain growth. Most of the densification has already happened at this time, and there may be sufficient grain boundary diffusion to eliminate the residual pores while grain growth is not activated. Application of lower temperatures during the final stage leads to two-step sintering schemes which have been probed recently as a means of preventing grain growth in the sintering of nanocrystalline systems [43]. The validity of this concept has also been demonstrated for the case of liquid-phase sintering [44, 45]. Rapid grain growth in the final sintering stage may be also impeded by impurity segregation or solute drag on grain boundaries [39, 70] or by second phase particles in the so-called nano/nano composites [15]. On the other hand, precipitates also obstruct densification and are generally not effective in achieving a homogeneous nanocrystalline structure. All these techniques have the advantage of allowing pressureless densification, and we shall resume the discussion of these methods below.

It is well known that the kinetics of densification can also be accelerated by applying external stresses, since this may increase the driving force for pore elimination. In both uniaxial hot pressing (HP) and hot isostatic pressing (HIP) processes, the sintering temperature can usually be lowered which leads to reduced grain growth. However, any pressure assisted sintering technique can only be efficient if the additional stress surpasses the internal, microscopic driving force for pore annihilation – the sintering stress – or at least reaches the same magnitude. The sintering stress is defined by

$$\sigma_S \approx 2\gamma / r_P \qquad 8.2$$

where γ is the surface tension associated with the pore/solid interface and r_P is the pore radius. According to Fick's first law, the diffusional flow of vacancies away from a pore scales with the concentration gradient, and so does the rate with which the pore is eliminated. The Kelvin–Thompson equation yields the inverse dependence of the concentration of vacancies emitted from a pore surface on the radius of curvature of the pore. Assuming $\gamma = 1$ J/m^2 and $r_P = 5$ nm, the sintering stress σ_S will be about 400 MPa [42]).

According to reference 110, the densification rate $d\varepsilon/dt$ observes the proportionality

$$d\varepsilon/dt \propto (\sigma/3 + \sigma_S)^q, \qquad 8.3$$

where σ is the applied stress, σ_S is the internal sintering stress, and q is the stress exponent of densification. Consequently, an externally applied pressure must be on the order of 10^2 MPa to have any effect in the densification of nanoceramics.

8.4.2 Pressureless sintering

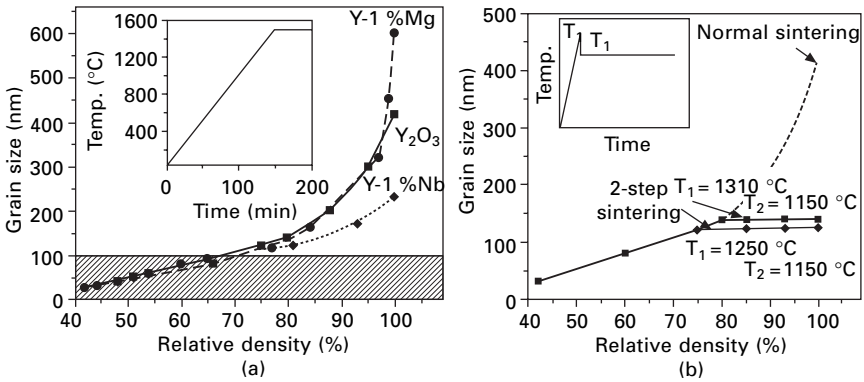
In a few cases, it has been possible by pressureless sintering to retain grain

sizes below 100 nm at a relative density ≥98% (see Table 8.1). This requires particularly rapid diffusion kinetics, and it has already been pointed out that non- or weakly agglomerated starting powders are also a precondition for preventing differential sintering. In view of the wide range of nanocrystalline ceramics that have been studied to date, the list of materials for which this has been achieved is rather short, and only pure and differently doped ZrO_2 and CeO_2 materials were studied in greater detail (Table 8.1). Especially for nanoscale Al_2O_3, obtaining full density has met great difficulties, mostly due to interference by the phase transformation from γ- and θ-Al_2O_3, which are the polymorphs present in nanopowders, to stable α-Al_2O_3. Seeding with fine α-Al_2O_3 particles has at least allowed the reaching of high sintered densities at temperatures as low as 1400–1450 °C; however, the final grain size obtained in this way has been restricted to about 1 μm [111–113].

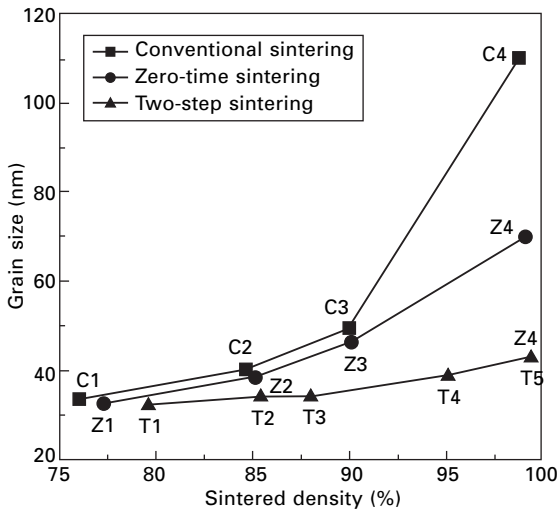
Advanced pressureless sintering methods, which make use of special temperature profiles, have been proposed for kinetic separation of the final stage densification from grain growth. For the example of Y_2O_3, Chen and Wang [43] demonstrated that sintering with restricted grain growth in the final stage is possible by a two-step sintering schedule, i.e., by reducing the temperature by about 10% immediately after reaching the peak sintering temperature (Fig. 8.8). They attributed this result to the existence of a kinetic window; according to their tentative explanation, the kinetics of grain–boundary diffusion and grain–boundary migration are different due to the slow mobility of grain junctions at lower temperatures. Departing from this view, Cahn [117] invoked segregation (i.e., solute drag) as the possible grain-growth limiting mechanism. More recently, Lee and co-workers [44, 45] were able to extend this two-step sintering strategy to a liquid-phase sintering system, namely, SiC with Al_2O_3-Y_2O_3-CaO sintering additives (Fig. 8.9).

Table 8.1 Nanocrystalline ceramics by pressureless sintering

Starting material	Dopants	Sintering temp. (°C)	Final grain size (nm)	Final density (%)	Ref.
ZrO_2	–	950	60	100	114
ZrO_2	2.6 mol% Y_2O_3	1125	60	100	67
ZrO_2	3 mol% Y_2O_3	1050	85	99.9	73
ZrO_2	3 mol% Y_2O_3	1070	72	99.9	115
ZrO_2	3–5 mol% Al_2O_3	1000	40–45	100	38
CeO_2	–	900	60	>98	39
CeO_2	7–10 at% Y	900	35–25	99.1–98.5	39
CeO_2	6 at% Ca	1350	30	full	111
TiO_2	–	600	<60	99	67
ZnO	–	700	66 (sub-grains)	98	116



8.8 (a) Density–grain size relationship for conventional sintering with dopants and (b) two-step sintering of nanocrystalline Y_2O_3. [43]



8.9 Two-step sintering applied to the liquid-phase forming system SiC/Al_2O_3-Y_2O_3-CaO additive [44]. The holding times necessary for near-theoretical density were: C4, 1750 °C/0.5 h; Z4, 1850 °C/0 h; T5: 1750 °C/0 h + 1550 °C/8 h.

8.4.3 Hot isostatic pressing

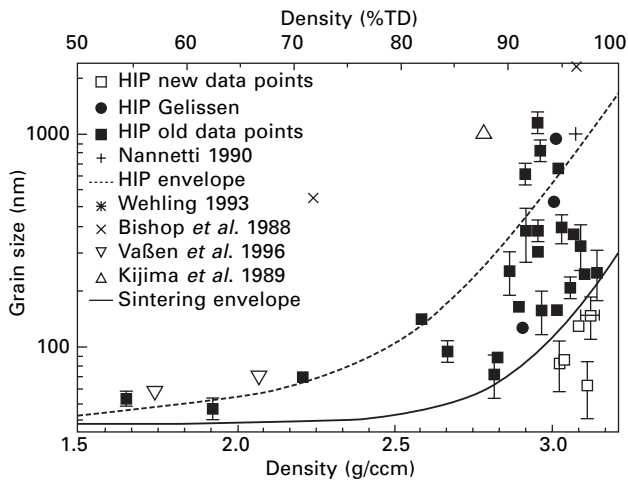
In hot isostatic pressing (HIP), inert gas pressures of up to 200 or even 400 MPa may be achieved at temperatures up to 2000 °C, depending on the equipment. In order to prevent the pressurized gas from penetrating the pore network of the sample, a hermetic, plastically deformable container is used to transmit the gas pressure to the outer surface of a sample. Depending on temperature range and sample material the container may be fabricated from glass, stainless steel, or a refractory metal such as tantalum. For oxide ceramics,

grain sizes significantly below 100 nm were obtained at near-theoretical density. In the case of TiO_2 [118], HIP at 354 MPa at 700–800 °C in vacuum yielded grain sizes of 55–67 nm at 99.5–100% density. During the process, the initial anatase modification was converted to rutile.

For the non-oxide material silicon carbide, which has a high degree of covalency, hot isostatic pressing with sintering additives (boron and carbon) at 1600–1650 °C and 350 MPa resulted in minimum average grain sizes below 150 nm at relative densities ≥ 96% (Fig. 8.10) [36]. Application of even higher pressure was also attempted as a means to shift the densification to lower homologous temperatures. HIPing SiC at 1600 °C and 980 MPa, with only some oxygen and carbon impurities originating from the starting powder, was claimed to lead to an average grain size of 30 nm [119]. However, this number is likely to represent the diffracting domain size in XRD rather than the actual grain size. Judging from a scanning electron microscopic image provided in that paper, the fineness of the microstructure seems to be rather comparable to the study mentioned above [36]).

8.4.4 Uniaxial hot pressing

In contrast to HIP, uniaxial hot pressing (HP) is usually limited to a pressure range of up to 50–100 MPa, which is significantly less than the internal sintering stress in nanocrystalline materials. The reason is that, in the temperature range needed to densify ceramic materials, there are not many alternatives to using plungers and dies made of graphite or fiber-reinforced



8.10 Grain growth as a function of density for sintered and HIPed SiC. At relative densities ≥ 96%, minimum average grain sizes ≤ 150 nm are obtained. Details of the references cited in the plot are produced in [36].

carbon materials. However, there are a few examples of uniaxial hot pressing at higher pressures. In a study [120], a relative density of 98% and a final grain size of 30 nm were reached by subjecting an ultrafine-grained (12 nm), Y_2O_3-doped TiO_2 powder to a uniaxial pressure of 1 GPa at temperatures between 450 and 550 °C. Similar results were obtained using pure TiO_2, Y_2O_3 and ZrO_2 nanopowders. For comparison, pressureless sintering of pure TiO_2 powder to a relative density of 98% led to a final grain size of 1 μm [120]. By studying the effects of high-pressure induced polymorphic phase transformations on densification at low temperatures (400–800 °C) and very high pressures (1–8 GPa), Kear and co-workers were able to reach grain sizes below 100 nm at almost theoretical density in various oxide ceramics, including <20 nm at 97% T.D. in TiO_2 [121] and 49 nm at 98.2% T.D. in hard-to-densify γ-Al_2O_3 [122, 123].

8.4.5 Hot pressing in the presence of a liquid phase

The pressure range up to 100 MPa, which can be attained with graphite tooling, may suffice to significantly aid in the densification of nanoceramics by hot pressing with a liquid phase. In this case, the pressure affects the densification mainly by facilitating grain rearrangement and pore closure by liquid intrusion [124]. Both processes are comparatively fast and happen in the initial stage of sintering. The intermediate stage of liquid phase sintering is kinetically controlled by dissolution in the melt and re-precipitation on particle surfaces, provided that some solubility of the particles in the melt exists. According to Kingery *et al.* [125], there is an external pressure contribution to this process as well, which is additive to the capillary force:

$$(\Delta L/L_0)^3 \propto (2\gamma/r + \sigma), \tag{8.4}$$

where $\Delta L/L_0$ means the linear shrinkage of the sintering body, γ is the liquid-vapor surface energy, r is the pore radius, and σ is the applied stress. Finally, for diffusion-controlled densification in the late stage of sintering, Wang and Raj [126] suggest a creep-related shrinkage mechanism. In this case, a direct applied-stress dependence of the linear shrinkage rate is observed according to

$$d(\Delta L/L_0)/dt \propto \sigma^n, \text{ with } n \approx 1. \tag{8.5}$$

For silicon carbide with liquid-phase-sintering additives (Al_2O_3–Y_2O_3–CaO), uniaxial hot pressing at a temperature of 1750 °C and a moderate applied pressure of 20 MPa [91, 127] resulted in a minimum average grain size of about 100 nm at a relative density ≥ 96%. In a more recent example [128], HP at 30 MPa and 1860 °C allowed fully densifying SiC with 10 wt% of Al_2O_3–Y_2O_3 additives while retaining an average grain size of 78 nm.

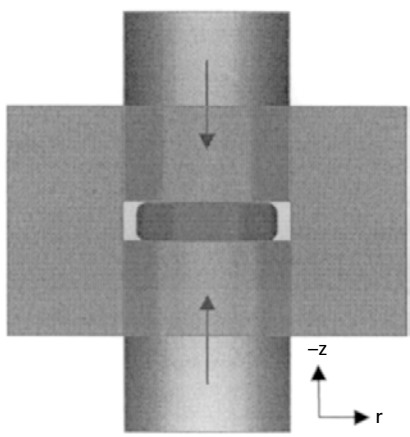
8.4.6 Sinter forging

A special variant of uniaxial hot pressing, where pressures up to 100 MPa may suffice to densify nanoceramics without extensive grain growth even in the solid state, is termed sinter forging (SF). In this process, densification is brought about by uniaxial pressing without lateral confinement, i.e., the sample (a green body or a compact which has been pre-sintered to 70–75% of the theoretical density in order to improve its mechanical strength) is allowed to flow in a radial direction without constraints (Fig. 8.11). Thereby, simultaneous densification and (creep) deformation of the sample occurs, which introduces large shear forces. The total shrinkage during SF is the result of two mechanisms: pore filling by stress-assisted diffusion (or, in LPS systems, by viscous flow of the matrix due to the hydrostatic component of the applied stress) and strain-controlled pore elimination (flattening). Sinter forging was first used by Rahaman and co-workers and by Venkatachari and Raj [129–132]. The mechanisms that may contribute during sinter forging operations have all been modeled; for an overview, see reference 42.

The strain components of volumetric densification, ε_v, and creep deformation, ε_c, may be separated in analogy to Raj's treatment [133]:

$$\varepsilon_c = 2/3\ (\varepsilon_z - \varepsilon_r), \quad \varepsilon_v = \ln\ (V/V_0) = \varepsilon_z - 2\varepsilon_r, \qquad 8.6$$

where ε_z is the strain component in the applied stress direction and ε_r is the radial strain. The driving force for the volumetric contribution has two components, namely the externally applied stress and the internal sintering stress (cf. eqn 8.3). Densification by stress-assisted diffusion requires sintering periods on the order of 10^3–10^4 s. Therefore, if the SF-specific contribution due to creep deformation is to be effective, the sample must tolerate deformation rates of 10^{-4}–10^{-3} s^{-1} by grain boundary sliding. This means that the



8.11 Schematic drawing of the die setup in sinter forging.

densification by sinter forging is closely connected with the phenomenon of superplasticity, or in other words, ceramics with ultrafine microstructures are favored for sinter forging operations.

The pore size distribution of nanocrystalline materials is usually bimodal, consisting of intercrystalline pores on the scale of the grain size and larger interagglomerate pores. Plastic flow can only close pores larger than a characteristic length scale of the elementary deformation process which exists in a particular material. For metals, where deformation occurs by dislocation activity (i.e., on an atomic length scale), essentially all pores are accessible to elimination by flattening. For ceramics, where it proceeds by grain boundary sliding, the grain size acts as a lower limit for efficient pore filling by deformation. Large interagglomerate pores can be closed in this way, whereas intercrystalline pores rely on stress-assisted diffusion for their elimination. In pressureless sintering, it is the largest pores which are hard to eliminate by diffusion alone, whereas the closure of small pores by diffusion is an efficient process. It has been established that prolonging the sinter forging beyond the point where the small pores close will result in a penalty with respect to grain size, with additional static (time- and temperature-dependent) and dynamic (plastic-strain-dependent) grain growth occurring in the final stage [42]. This means that the deformation rate should be as high as possible in order to eliminate large pores before grain growth accelerates. However, the maximum deformation rate for a given material is limited by the critical stress concentration at microstructural defects.

For yttrium stabilized t-ZrO_2, the stress dependence of the deformation rates was determined experimentally as

$$d\varepsilon_c/dt \sim \sigma^3 \quad \text{and} \quad d\varepsilon_v/dt \sim \sigma^2 \tag{8.7}$$

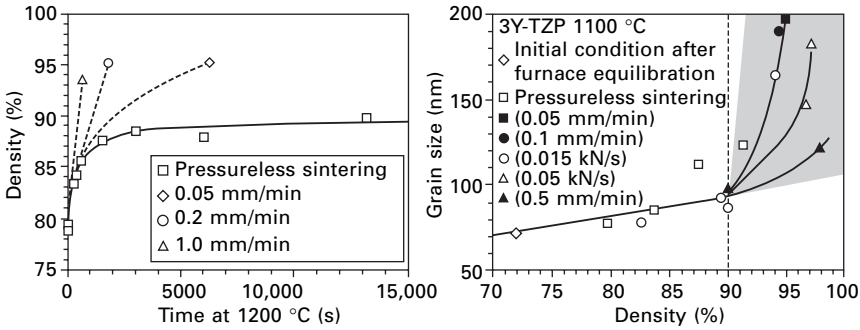
where σ is the applied stress [110]. For oxide ceramics, the specific sinter-forging effect was demonstrated to become effective at applied pressures ≥ 40 MPa.

Sinter forging experiments can be conducted using different load application modes. In the simplest case, the applied force is kept constant, which implies that with increasing radius of the sample the initial pressure decreases. This is not very efficient in terms of minimizing the sintering time in the last stage of the densification process. Two alternative load application modes were shown to be more efficient, i.e., constant loading rate and constant deformation rate. Typically, loading rates in the range of 5 to 100 N/s and constant deformation rates in the range of $1 \cdot 10^{-4}$ to $7 \cdot 10^{-4}$ s^{-1} (5 to 500 μm min^{-1}) have been described in the literature [42, 134–137].

Table 8.2 provides examples for sinter forging applied to oxide ceramics, with the case of yttrium-stabilized t-ZrO_2 [135] being further illustrated in Fig. 8.12. It can be seen that sinter forging enables densification at low homologous temperatures, in exceptional cases [134] at $T_{SF} \leq 0.5\ T_m$. Even

Table 8.2 Nanoceramics by sinter forging

Starting material	Dopants	Green/pre-sintered density (%)	Sinter forging temp. (°C)/ pressure (MPa)	T_{SF}/T_m	Final grain size (nm)	Final density (%)	Ref.
θ-Al_2O_3	1.5 wt% α- Al_2O_3	49	1150/190	0.64	190 (α-Al_2O_3)	99.5	136
Al_2O_3 – 25 wt% ZrO_2	2 mol% Y_2O_3	54.7 (pre-sintered)	1400/50	0.75	260 (Al_2O_3 grains) 150 (ZrO_2:Y grains)	>99	144
Al_2O_3 – 25 wt% ZrO_2	8 mol% CeO_2	55.6 (pre-sintered)	1400/50	0.75	260 (Al_2O_3 grains) 110 (ZrO_2:Ce grains)	>99	144
ZrO_2	–	45–50	950/100	0.43	45	100	143
t-ZrO_2	3 mol% Y_2O_3	39.5	1350/25	0.57	180	fully	110
t-ZrO_2	3 mol% Y_2O_3	38	1100/280	0.48	120	98	135
TiO_2	–	≈70 (warm-pressed)	650/93	0.40	60	97	134



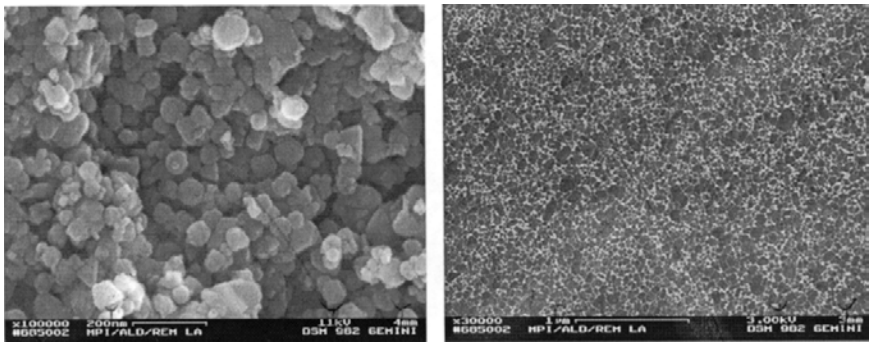
8.12 Sinter forging of nanocrystalline 3Y-TZP (ZrO_2-3 mol% Y_2O_3) [135]: (a) kinetics of densification as a function of crosshead speed at 1200 °C, (b) grain size–density relation at 1100 °C, showing that the influence of a high deformation rate becomes beneficial after pore closure.

more important is the reduction in maximum flaw size that can be achieved because large pores are preferentially eliminated by sinter forging [138]. According to a viscoelastic model [139] cited in reference 140, the necessary deformation for eliminating spherical pores is $\varepsilon_z \geq 0.64$. The validity of this model was confirmed for alumina, cubic stabilized zirconia and alumina-toughened zirconia [141, 142].

For the case when a liquid phase is present, there are indications that larger total deformations are necessary for complete removal of the pores by sinter forging. For a SiC powder with 50 nm mean particle size and Al_2O_3-Y_2O_3 sintering aids, final grain sizes between 55 and 70 nm were achieved in fully dense samples (≥ 99% T.D.) (Fig. 8.13). Full densification required a longitudinal strain of $\varepsilon_z = 0.88$, which was independent of the choice of the temperature and pressure conditions between 1660 °C/90 MPa and 1720 °C/40 MPa [90]. More specifically, the final density vs. longitudinal strain curves collapsed into one for relative densities above the presintered density of ≈75% (Fig. 8.14). The higher ε value required for complete pore elimination may be related to the internal gas pressure which builds up during pore shrinkage. One may reason that this counteracts the flattening of pores more effectively in LPS than in solid-state sintering, and that it may even lead to a re-spheroidization of the pores. In the present system (SiC/Al_2O_3-Y_2O_3), this deleterious effect is alleviated because the internal pressure may be released gradually by dissolution of the nitrogen gas in the liquid phase [145].

8.4.7 Pulsed electric current sintering

Another modified hot-pressing method effective in restricting grain growth also starts from a uniaxial pressing setup. In this case, pulsed DC currents of



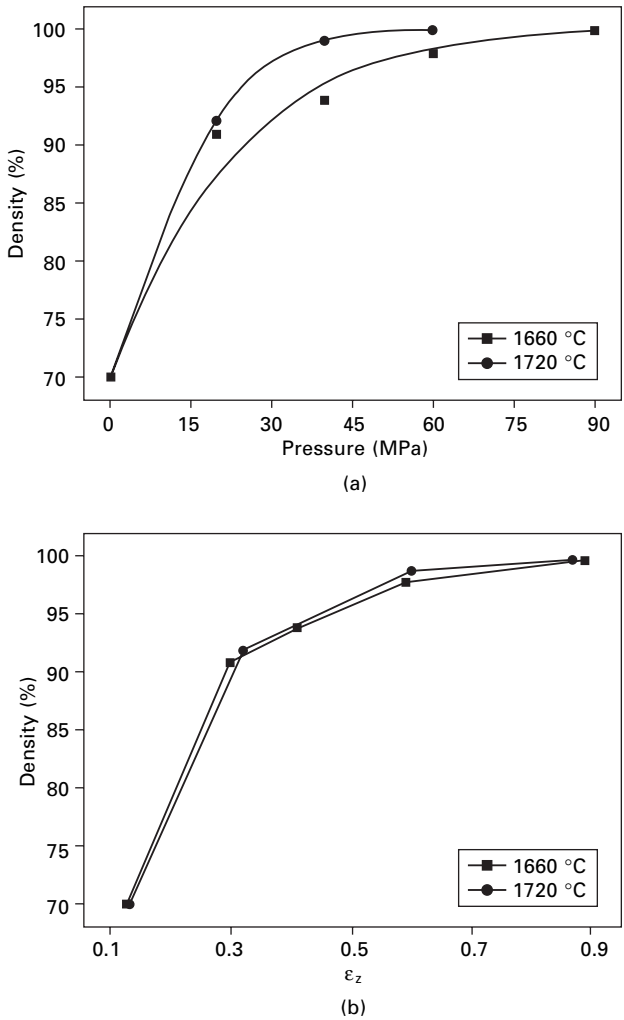
(a) (b)

8.13 (a) SEM image of a compact pre-sintered to 75% T.D. by pressureless sintering at 1600 °C. The starting powder is an attrition-milled mixture of β-SiC with 10 vol% of Al_2O_3 and Y_2O_3 with an average particle size of ≈50 nm, i.e., there is little or no grain growth during pre-sintering. (b) Near-theoretical density and d_{50} = 55 nm after sinter forging at 1660 °C/90 MPa (courtesy K. Wetzel).

some 10^3 A are applied to heat the powder via a graphite punch. Rapid heating rates of about 100 Ks^{-1} may be generated in this way, which is a unique characteristic of this so-called pulsed electric current sintering (PECS). Two more abbreviations, SPS (spark plasma sintering) and PAS (plasma activated sintering) were introduced [40, 41] for proprietary processes which differ only in the time periods during which a pulsed current is applied. Both groups of authors [40, 41] propose that spark discharges or even plasma discharges are responsible for rapid densification by promoting transfer of material across cleaned and activated interfaces. In the further discussion, we prefer to use the term pulsed electric current sintering (PECS).

By PECS, near-theoretical density has been achieved for a variety of ceramics (Table 8.3) and composites (Table 8.4) within extremely short dwell times of several minutes at the sintering temperature. However, there is still uncertainty about the important questions, whether electrical discharges generated inside the densifying ensemble of particles really contribute to the densification process and whether there is a special PECS effect on the minimum achievable grain size. For many non-conductive ceramic materials, it may be stated at least that the current flow is mainly through the graphite plunger and die and not through the powder up to reasonably high temperatures.

Mishra *et al*. [146] verified experimentally that densification is enhanced by DC pulses. While Shen *et al*. [147] state that a temperature window exists within which fully dense compacts can be obtained with very limited grain growth, Munir and Schmalzried [148] argue that the application of an electrical field enhances not only densification but also the grain growth kinetics.



8.14 Pressure–density (a) and longitudinal strain–density (b) relations for LPS-SiC sinter-forged in a temperature range of 1660–1720 °C, showing temperature-independent behavior as a function of the longitudinal deformation.

Using a widespread commercial SPS setup, Suganuma *et al.* [149] determined temperature differences larger than 100 K between graphite die and sample, which led them to the conclusion that one frequently claimed PECS effect, namely that of a significantly lower densification temperature, might actually be an artifact due to incorrect temperature measurement. Furthermore, they directly examined the impact of current passing through the sintering powder by inserting BN disks between the graphite plungers and the sample. Although

Table 8.3 Ceramics with sub-micrometer grain size by PECS

Material	Additives/ 2nd phase	Particle size (nm)	Final grain size (nm)	Final density (%)	Ref.
ZrO_2	Y_2O_3	100 nm	sub-μm	99	169
Al_2O_3	–	sub-μm	0.5/1/1.5 μm	95.6/99.7/99.9	158
Al_2O_3	–		350	99.8	165
Al_2O_3	–	400	500–600	98–100	166
Al_2O_3	0.046 mol% MgO	300–500 (agglom. size)	450	near-theoretical	172
TiO_2	–	20	170–200	97–99	163
ZnO	–	20	100	98.5	167
$BaTiO_3$	–	60	270	97	168
$BaTiO_3$	–	30–40	50 (250)	97	161
$BaTiO_3$	–	20	200–500	95.6–97.8	162
(Pb, Zr, Sn) TiO_3	La_2O_3		300	99.5	160
Si_3N_4	3 wt% Al_2O_3–5 wt% Y_2O_3	200	200	99.6	149
a-Si-C-N-(O)	Y_2O_3		38–150	2.96–3.09 g/cm^3	164
Si_3N_4-SiC (polymer-derived)	3/5 mol% Y_2O_3	amorphous	56/86	93.8/93.3	170, 171

the BN remained isolating at the sintering temperature, little or no influence on the microstructure could be detected with this modification. This was taken as a further strong hint that special effects related to electrical discharges among the powder particles were absent.

Nevertheless, during the densification by PECS, Suganuma *et al*. [149] observed virtually no grain growth if a heating rate of 100 K/min was employed and no holding time at the peak temperature of 1840 °C was allowed. In contrast, considerable grain growth was present after uniaxial HP with a heating rate of 15 K/min, a plateau temperature of 1850 °C and a holding time of 15 minutes. In both cases, theoretical density was reached. During prolonged holding times in PECS, the third power of the grain size was seen to scale linearly with time, suggesting that the grain growth was diffusion-controlled [150, 151]. Generally, the PECS results reported in the literature reveal large variations in the amount of grain growth during densification. While there are many researchers who claim that nanocrystalline microstructures are readily obtained by this method due to the very short sintering times [152–156], there are also many observations of rapid grain growth leading to a final grain size in the sub-micrometer and low micrometer range (cf. Table 8.3). In a work by Zhou *et al*. [157], nanometer-sized SiC powder with 2.5 wt% of Al_4C_3 and B_4C additives was densified at 1600 °C/47 MPa with heating rates of 100–400 K/min and holding times of 2–5

Table 8.4 Nanocomposites by PECS

Matrix	Reinforcement	Particle size matrix (nm)	Particle size reinforcement (nm)	Final grain size matrix (nm)	Final grain size reinforcemet (nm)	Final density %	Ref.
Al_2O_3	$ZrO_2(Y_2O_3)$			96	265	fully dense	174
Al_2O_3	14 vol% $Nd_2Ti_2O_7$	325 mesh	50		230	99.5	177, 178
Al_2O_3	5.7/10 vol% CNT	≤ 40		100/200	unchanged	99/100	179
MgO	20 vol% Fe	100		1300	< 50 (intragrain) 400 (intergrain)	>98	180
$Y_3Al_5O_{12}$	5 vol% SiC		70	~ 1 μm	unchanged	> 98	175
$Si_2Al_6O_{13}$	10 vol% SiC	100	70	280	unchanged	97	176
WC	6 wt% ZrO_2	200	27	300–400	100	99.9	173

minutes. The sintered specimens reached near-theoretical density but consisted of large elongated grains. Enormous grain growth took place in all cases studied.

In most reports, no stringent analysis of the kinetics of densification vs. grain growth has been provided. Shen *et al.* [158] started from a sub-micrometer Al_2O_3 powder and investigated a variety of influences on the densification and grain growth kinetics. With the parameters 1200 °C/50 MPa, a grain size of 0.5 μm could be retained at a final density of 95.6% (no holding time), 1 μm at 99.7% (3 minutes holding time), and 1.5 μm at 99.9% (10 minutes). An optimized heating rate of 120 K/min and an optimized pulse sequence were used. It was concluded that the PECS technique enhances both densification and grain growth in Al_2O_3. Retention of the starting grain size might be possible for relative densities in excess of 99%, depending on the parameters.

According to reference 159, an advantage of PECS is that compacts of complex geometry and pieces with diameters larger than 150 mm can be produced. Additionally, monophasic samples may be obtained in systems where one of the constitutents has an anomalously high vapor pressure or where there is decomposition at high temperature. Nygren and Shen [159] describe the example of hydroxyl apatite (HAP)/ZrO_2 composites, where decomposition into calcium phosphate and a solid solution of oxides starts already at 950°C.

8.5 Specific properties and applications

A great deal of interest in nanocrystalline ceramics arose because improved mechanical properties were predicted to exist in the sub-micrometer regime. This expectation was to some extent based on the Griffith concept for the strength of brittle materials. In this theory, the fracture strength of a tested sample is inversely proportional to the square root of the size of the most severe flaw in it. Since a strong correlation exists between the grain size and flaw size distributions of a ceramic microstructure, lowering the average grain size will generally lead to an increased strength. In practice, however, high strength was found difficult to realize in nanoceramics. This is because conducting each manufacturing step adequately becomes a more and more complex task as the grain size is reduced below a few hundred nanometers. On the other hand, it has been possible to achieve a significant reduction of diffuse light scattering in fine-grained ceramics, indicating that virtually all flaws can be eliminated by careful preparation and that the remaining defects are smaller than the wavelength of light. Moreover, the refractive anisotropy of non-isotropic materials such as α-Al_2O_3 vanishes for small crystallites. The in-line transmission becomes an inverse function of the grain size at zero porosity [181], which means that optical transparency, as opposed to

translucency, can be obtained at small grain sizes even in non-cubic materials. This effect is already significant at grain sizes below 1 μm, i.e., in a region which is quite accessible to the powder technological synthesis route [182–184].

Another remarkable mechanical property of nanocrystalline materials is superplasticity, which may be defined as the ability to undergo large elongations prior to failure at elevated temperatures [185]. A nanoscale grain size is required for superplastic deformation, because groups of grains must rearrange or reconstruct collectively in order to avoid grain boundary cavitation. In nanoceramics, which are intrinsically brittle materials, moderately high strain rates up to 10^{-3} s^{-1} are tolerated without generating damage to the microstructure; this puts superplastic deformation on the verge of being usable as an economically viable forming process. For this reason, the superplasticity of nanostructured ceramic materials has repeatedly been the subject of review articles [185–187].

Among the projected applications of porous ceramics, membranes for filtration gain increasing importance due to their higher mechanical, thermal and chemical stability as compared to polymer membranes [188]. Pore sizes of 20–50 nm have been realized by spin coating of suspensions of ceramic nanopowders like ZrO_2, TiO_2, γ-Al_2O_3 and TiN_x on coarser substrates, followed by partial sintering. For ultrafiltration membranes with pore sizes of 3–20 nm, supported layers of sol-gel derived nanoparticles were developed. $BaTiO_3$ nanoparticles synthesized by a wet chemical method were tested as a model perovskite material for membrane applications [189]. Alternatively, a membrane filter for airborne particles was created by transporting ultrafine particles to the interior of a heated porous substrate by means of a carrier gas. By the substrate heating, *in-situ* sintering of the ultrafine powder to the substrate surface was achieved [190]. It was also found that nanocrystalline grain structures can impart enhanced bio-compatibility to sintered materials for prostheses and implants, such as Al_2O_3, TiO_2, and hydroxyapatite, $Ca_5(PO_4)_3OH$ [191–194].

Electrical effects are very much dependent on microstructural details on the length scale of coulomb interactions, e.g., the grain boundaries. Accordingly, there has been strong scientific interest in the impact of nanocrystallinity on electrical effects. Moreover, technological impulses such as the ongoing miniaturization of multilayer capacitors and the attempt to embed capacitances in printed circuit boards have constituted a strong driving force for the reduction of the grain size of electrical ceramics to the nanometer range. In ferroelectrics, the dielectric constant decreases upon lowering the grain size to sub-micrometer dimensions and, ultimately, superparaelectric relaxation is expected to occur. However, a nearly dense (97%), translucent, 50 nm $BaTiO_3$ material densified by PECS still showed a maximum permittivity of ~1100 at 390 K. Polarization switching was found to occur locally, indicating

that the critical grain size for the disappearance of the ferroelectric behavior was somewhat less than the average grain size of 50 nm [195], i.e., in a size range which is hard to achieve in dense sintered nanoceramics. While the varistor behavior of ZnO-based ceramics was completely suppressed in pure and Bi-doped ZnO materials with grain sizes between 16 and 80 nm [196], nonlinearity of the electrical conduction in ZnO ceramics with a grain size of ~100 nm was found in another study [63] (cf. Fig. 8.2).

In semiconducting gas sensors, such as SnO_2 and TiO_2, particularly high sensitivities occur if the grain size is on the same order of magnitude as the size of the space charge zones (<50 nm). This leads to a situation where the overall resistivity of the porous sintered bodies is controlled by the neck regions between the grains which, in turn, are strongly affected by adsorbed gaseous species [197, 198]. However, a serious drawback of nanocrystalline sensors is their lack of microstructural stability at moderately high temperatures. Partially sintered, nanocrystalline TiO_2 has also been used as porous electrode material for electrochemical photovoltaic cells [199].

Diffusion and ionic conduction properties of nanocrystalline oxide ceramics were also thoroughly studied, with the thickness reduction to less than 1 μm of solid electrolyte films for sensing and energy conversion applications being a major driving force [200]. Although complete understanding of the observed phenomena has not been reached so far, an enhanced grain boundary diffusivity was generally recognized to be present [201]. For example, electrical conductivity measurements performed on nanocrystalline yttrium-stabilized ZrO_2 materials with grain sizes down to 25–50 nm yielded a specific grain boundary conductivity 1–2 orders of magnitude higher than in microcrystalline samples [202].

Last but not least, the magnetic properties of ferrite ceramics are extremely sensitive to the grain size. For example, nanoscale grain size leads to a reduction of the coercive field owing to the so-called random anisotropy effect [203]. Enhanced magnetoresistance has been found in nanocrystalline α-Fe_2O_3 [204], and ferromagnetic precipitates in a non-magnetic matrix exhibit superparamagnetic behavior at ambient temperature [35].

8.6 Conclusions and future trends

Due to their wide range of specific properties and potential applications, nanocrystalline ceramics continue to be an exciting subject for research. Within a range of synthesis methods, the processing and sintering of ultrafine powders occupies a prominent position in spite of the inherent difficulty that densification and grain growth are based on the same class of thermally activated processes. It was demonstrated by many examples that grain sizes below 100 nm can be achieved by powder technological methods if attention

is paid to the fact that non-agglomerated starting powders and very homogeneous green microstructures are required for sintering.

A great deal of motivation for devising novel sintering methods stems from the requirement of making larger parts with a larger freedom of shape and better economy. This favors free sintering techniques over pressure-assisted methods such as sinter forging and pulsed electric current sintering, and recent work suggests that there may actually be good prospect for advances in this direction. As suggested by Cahn [117], a deepened understanding of grain boundary segregation phenomena during the densification of nanocrystalline powders may be the key to categorizing the results obtained by advanced sintering techniques and, eventually, devising novel methods for decoupled densification and grain growth kinetics. This conjecture rests on several arguments. For example, an energetic consideration by Weissmüller [205] leads to the conclusion that systems with a large enthalpy of grain boundary segregation may exhibit a reduced or even vanishing excess Gibbs free energy of the grain boundaries and, therefore, a largely reduced driving force for grain growth. For metallic alloy systems, this finding was corroborated by experiments.

Beyond the addition of precipitates, which obstruct the grain boundary motion by pinning but also hinder densification, there are also novel morphological concepts for stabilizing the grain size. By analyzing the kinetics of grain boundary movement in a tri-crystal arrangement, Gottstein and Shvindlerman [206] found that triple junctions can exert a drag force on the adjoining grain boundaries. The kinetics of the system may actually be controlled by the mobility of the triple junctions, provided that this quantity is much smaller than the mobility of grain boundaries divided by a measure directly related to grain size. This finding may partly resolve the controversy concerning the stabilizing effect on the grain size of a reduced holding temperature in the last sintering stage [43, 117].

8.7 References

1. H. Gleiter, in: *Proc. 2nd Risø International Symposium on Metallurgy and Materials Science*, eds N. Hansen *et al*., Risø National Laboratory, Roskilde (1981), pp. 15–21.
2. H. Gleiter, in: *Mechanical Properties and Deformation Behaviour of Materials*, eds M. Nastasi *et al*., Kluwer Academic Publishers, Amsterdam (1993), pp. 3–35.
3. G. Rixecker, *Hyperfine Interact.* **130** (2000) 127–50.
4. K. Hilfrich, W. Kölker, W. Petry, O. Schärpf, E. Nembach, *Z. Metallkd.* **84** (1993) 255–258.
5. P.F. Becher, E.Y. Sun, C.S. Hsueh, G.S. Painter, K.L. More, *Key Engg. Mater.* **175–176** (2000) 97–106.
6. S.M. Wiederhorn, B.J. Hockey, J.D. French, *J. Eur. Ceram. Soc.* **19** (1999) 2273–2284.

7. M. Sternitzke, *J. Eur. Ceram. Soc*. **17** (1997) 1061–1082.
8. D. Sciti, J. Vicens, A. Bellosi, *J. Mater. Sci*. **37** (2002) 3747–58.
9. T. Ohji, Y.-K. Jeong, Y.-H. Choa, K. Niihara, *J. Am. Ceram. Soc*. **81** (1998) 1453–1460.
10. L.P. Ferroni, G. Pezzotti, *J. Am. Ceram. Soc*. **85** (2002) 2033–2038.
11. P. Šajgalík, M. Hnatko, F. Lofaj, P. Hvizdoš, J. Dusza, P. Warbichler, F. Hofer, R. Riedel, E. Lecomte, M.J. Hoffmann, *J. Eur. Ceram. Soc*. **20** (2000) 453–462.
12. R.F. Krause, S.M. Wiederhorn, C.-W. Li, *J. Am. Ceram. Soc*. **84** (2001) 2394–2400.
13. J.-F. Yang, T. Sekino, Y.-H. Choa, K. Niihara, T. Ohji, *J. Am. Ceram. Soc*. **84** (2001) 406–412.
14. P. Rendtel, A. Rendtel, *J. Eur. Ceram. Soc*. **22** (2002) 2061–2070.
15. J.D. Kuntz, G.-D. Zhan, A.K. Mukherjee, *MRS Bulletin* **2004/1**, 22–27.
16. S. Dong, Y. Katoh, A. Kohyama, *J. Am. Ceram. Soc*. **86** (2003) 26–32.
17. *CVD of Nonmetals*, ed. W. S. Reese, Wiley-VCH, Weinheim (1996).
18. A.C. Jones, P. O'Brien, *CVD of Compound Semiconductors*, Wiley-VCH, Weinheim (1997).
19. T.P. Niesen, M.R. de Guire, *J. Electroceram*. **6** (2001) 169–207.
20. J. Bill, R.C. Hoffmann, T.M. Fuchs, F. Aldinger, *Z. Metallkde*. **93** (2002) 478–489.
21. C.J. Brinker, G.W. Scherer, Sol-Gel Science: *The Physics and Chemistry of Sol-Gel Processing*, Academic Press, San Diego (1990).
22. A.C. Pierre, *Introduction to Sol-Gel Processing*, Kluwer Academic Publishers, Boston (1998).
23. P. Colombo, A. Martucci, O. Fogato, P. Villoresi, *J. Am. Ceram. Soc*. **84** (2001) 224–226.
24. X. Yang, M.N. Rahaman, *J. Eur. Ceram. Soc*. **17** (1997) 525–535.
25. M. García-Hernández, F. Ginea, A. de Andrés, J.L. Martínez, C. Prieto, L. Vázquez, *Phys. Rev. B* **61** (2000) 9549–9552.
26. D. Kumar, J. Narayan, T.K. Nath, A.K. Sharma, A. Kvit, C. Jin, Sol. *State Comm*. **119** (2001) 63–66.
27. S.H. Yu, M. Yoshimura, *Adv. Funct. Mater*. **12** (2002) 9–15.
28. M. Weinmann, F. Aldinger, in *Handbook of Advanced Ceramics*, Vol. 1, eds S. Somiya *et al*., Elsevier Academic Press, London (2003), pp. 265–367.
29. H. Schmidt, G. Borchardt, A. Müller, J. Bill, J. Non-Cryst. Solids **341** (2004) 133–140.
30. Ravi Kumar N.V., *The Influence of Crystallization on the High Temperature Deformation Behavior of Precursor Derived Si-B-C-N Ceramics*, Ph. D. Thesis, University of Stuttgart (2005).
31. G.H. Beall, L.R. Pinckney, *J. Am. Ceram. Soc*. **82** (1999) 5–16.
32. L. Stoch, *Optica Applicata* **33** (2003) 115–123.
33. W. Holand, V. Rheinberger, M. Schweiger, *Phil. Trans. Roy. Soc*. (London) A **361** (2003) 575–588.
34. R. Sakai, Y. Benino, T. Komatsu, *Appl. Phys. Lett*. **77** (2000) 2118–2120.
35. M. Pal, P. Brahma, D. Chakravorty, D. Bhattacharyya, H.S. Maiti, *J. Magn. Mater*. **164** (1996) 256–260.
36. R. Vassen, D. Stöver, *J. Am. Ceram. Soc*. **82** (1999) 2585–2593.
37. G. Skandan, H. Hahn, M. Roddy, W.R. Cannon, *J. Am. Ceram. Soc*. **77** (1994) 1706–1710.
38. V.V. Srdic, M. Winterer, H. Hahn, *J. Am. Ceram. Soc*. **83** (2000) 1853–1860.
39. J. Markmann, A. Tschöpe, R. Birringer, *Acta Mater*. **50** (2002) 1433–1440.

40. M. Tokita, *J. Soc. Powder Technol*. Jpn. **30** (1993) 790–804.
41. C.H. Shan, S.H. Risbud, K. Yamazaki, K. Shoda, *Mater. Sci. Eng. B* **26** (1994) 55–60.
42. D.C. Hague, M.J. Mayo, *Mater. Sci. Eng. A* **204** (1995) 83–89.
43. I.-W. Chen, X.-H. Wang, *Nature* **404** (2000) 168–171.
44. Y.-I. Lee, Y.-W. Kim, M. Mitomo, D.-Y. Kim, *J. Am. Ceram. Soc*. **86** (2003) 1803–1805.
45. Y.-I. Lee, Y.-W. Kim, M. Mitomo, *J. Mater. Sci*. **39** (2004) 3801–3803.
46. Z.R. Dai, J.P. Bradley, D.E. Brownlee, D.J. Joswiak, in *Lunar and Planetary Science XXXIV* (2003), downloaded from www.lpi.usra.edu/meetings/lpsc2003 /pdf/ 1121.pdf on 12 Dec 2003.
47. I. Amato, *Science* **286** (1999) 1059–1061.
48. E.C. Snelling, *Soft Ferrites: Properties and Applications*, Iliffe Books, London (1988), pp. 1–17.
49. G.C. Lowrison, *Crushing and Grinding*, CRC Press, Cleveland (1979), p. 60.
50. F.H. Froes, O.N. Senkov, E.G. Baburaj, *Mater. Sci. Eng. A* **301** (2001) 44–53.
51. J.S. Jiang, X.L. Yang, L. Gao, J.K. Guo, J.Z. Jiang, *Nanostruct. Mater*. **12** (1999) 143–146.
52. W. Schlump, H. Grewe, *Int. J. Mater. Product Technol*. **5** (1990) 281–292.
53. H.S. Nalwa, *Handbook of Nanostructured Materials and Nanotechnology*, Vol. 1: Synthesis and Processing, Academic Press, San Diego (2002).
54. C. Gomez-Aleixandre, J.M. Albella, J.M. Martinez-Duart, *J. Mater. Res*. **7** (1992) 2864–2868.
55. S. Iijima, *Nature* **354** (1991) 56–58.
56. R. Riedel, A.O. Gabriel, *Adv. Mater*. **11** (1999) 207–209.
57. *Nanophase and Nanostructured Materials*, eds Z.L. Wang *et al*., Kluwer Academic Press, Beijing (2002), p. 37.
58. H.Z. Wang, L. Gao, K. Niihara, *Mater. Sci. Eng. A* **288** (2000) 1–4.
59. A.C. Pierre, *Introduction to Sol-Gel Process*, Kluwer Academic Publishers, Boston (1998), p. 7.
60. Q.H. Zhang, L. Gao, J. Sun, S. Zheng, *Chem. Lett*. **31** (2002) 226–227.
61. X. Kuang, G. Carotenuto, L. Nicolais, *Adv. Perform. Mater*. **4** (1997) 257–274.
62. A.G. King, *Ceramic Technology and Processing*, Noyes Publications, New York (2002).
63. L. Gao, Q. Li, W.L. Luan, *J. Am. Ceram. Soc*. **85** (2002) 1016–1018.
64. W. Li, J.J. Lannutti, *J. Mater. Res*. **17** (2002) 2794–2801.
65. H. Ferkel, R.J. Hellmig, *Nanostruct. Mater*. **11** (1999) 617–622.
66. Y. Liang, N. Gregory, J.G.P. Binner, *Key Engg. Mater*. **264–268** (2004) 2319–2322.
67. G. Skandan, *Nanostruct. Mater*. **5** (1995) 111–126.
68. Y.C. Zhou, R.J. Phillips, J.A. Switzer, *J. Am. Ceram. Soc*. **78** (1995) 981–985.
69. R.A. Andrievski, *J. Mater. Sci*. **29** (1994) 614–631.
70. G.S.A.M. Theunissen, A.J.A. Winnubst, A.J. Burggraaf, *J. Eur. Ceram. Soc*. **11** (1993) 315–324.
71. A.J.A. Winnubst, G.S.A.M. Theunissen, M.M.R. Boutz, A.J. Burggraaf, in *Structural Ceramics: Processing, Microstructure and Properties*, eds J.J. Bentzen *et al*., Risø National Laboratory, Roskilde (1990), pp. 523–527.
72. M.J. Mayo, *Mater. & Design* **14** (1993) 323–329.
73. D.L. Chen, M.J. Mayo, *Nanostruct. Mater*. **2** (1993) 469–478.

74. W. Wagner, R.S. Averbach, H. Hahn, W. Petry, A. Wiedenmann, *J. Mater. Res*. **6** (1991) 2193–2198.
75. R. Vassen, D. Stöver, *Mater. Sci. Eng. A* **301** (2001) 59–68.
76. A.J.A. Winnubst, G.S.A.M. Theunissen, A.J. Burggraaf, in *Euro-Ceramics – Proc. 1st Eur. Ceram. Soc. Conf*. (Maastricht, 1989), Vol. 1: Processing of Ceramics, eds G. de With, R. A. Terpstra, R. Metselaar, Elsevier, London (1989), pp. 1393–1397.
77. H. Shan, Z. Zhang, *Brit. Ceram. Trans*. **95** (1996) 35–38.
78. D.M. Liu, *Ceram. Int*. **25** (1999) 107–113.
79. Ł. Zych, K. Haberko, *Key Engg. Mater*. **264–268** (2004) 2323–2326.
80. I.A. Aksay, in *Advances in Ceramics, Vol. 9: Forming of Ceramics*, ed. J. Mangels, American Ceramic Society, Columbus (1984), pp. 94–104.
81. K.T. Miller, C.F. Zukoski, *J. Am. Ceram. Soc*. **77** (1994) 2473–2478.
82. R. Nass, S. Albayrak, M. Aslan, H. Schmidt, in *Ceramic Transactions, Vol. 51: Ceramic Processing Science and Technology*, eds H. Hausner, G.L. Messing, S. Hirano, Am. Ceram. Soc., Westerville (1995), pp. 591–595.
83. F. Tang, T. Uchikoshi, Y. Sakka, *J. Am. Ceram. Soc*. **85** (2002) 2161–2165.
84. J. Tabellion, R. Clasen, *J. Mater. Sci*. **39** (2004) 803–811.
85. M. Mayo, Intern. *Mater. Rev*. **41** (1996) 85–115.
86. W.H. Rhodes, *J. Am. Ceram. Soc*. **64** (1981) 19–22.
87. M.J. Readey, R.R. Lee, J.W. Halloran, A.H. Heuer, *J. Am. Ceram. Soc*. **73** (1990) 1499–1503.
88. G. Skandan, H. Hahn, J.C. Parker, *Scr. Metall. Mater*. **25** (1991) 2389–2393.
89. A. Pechenik, G.J. Piermarini, S.C. Danforth, *J. Am. Ceram. Soc*. **75** (1992) 3283–3288.
90. K. Wetzel, G. Rixecker, G. Kaiser, F. Aldinger, *Adv. Engg. Mater*. **7** (2005) 520–524.
91. M. Mitomo, Y.W. Kim, H. Hirotsuru, *J. Mater. Res*. **11** (1996) 1601–1604.
92. L. Gao, W. Li, H.Z. Wang, J.X. Zhou, Z.J. Chao, Q.Z. Zai, *J. Eur. Ceram. Soc*. **21** (2001) 135–138.
93. D.M. Liu, *J. Mater. Sci. Lett*. **17** (1998) 207–210.
94. J. Cesarano, I.A. Aksay, *J. Am. Ceram. Soc*. **71** (1988) 1062–1067.
95. J.F. Kelso, T.A. Ferrazzoli, *J. Am. Ceram. Soc*. **72** (1989) 625–627.
96. S. Baklouti, C. Pagnoux, T. Chartier, J.F. Baumard, *J. Eur. Ceram. Soc*. **17** (1997) 1387–1292.
97. F. Shojai, A.B.A. Pettersson, T. Mäntylä, J.B. Rosenholm, *J. Eur. Ceram. Soc*. **20** (2000) 277–283.
98. X. Zhu, F. Tang, T.S. Suzuki, Y. Sakka, *J. Am. Ceram. Soc*. **86** (2003) 189–191.
99. A. Dietrich, A. Neubrand, *J. Am. Ceram. Soc*. **84** (2001) 806–812.
100. Z.G. Shen, J.F. Chen, H. Zou, J. Yun, *J. Coll. Interf. Sci*. **275** (2004) 158–165.
101. Z.G. Shen, J.F. Chen, H. Zou, J. Yun, *Colloids and Surfaces A* **244** (2004) 61–66.
102. R. Zhou, Y. Chen, Y. Liang, F. Zheng, J. Li, *Ceram. Int*. **28** (2002) 705–709.
103. B. Vaidhyanathan, J.G.P. Binner, *Key Engg. Mater*. 264–268 (2004) 239–242.
104. W.M. Sigmund, N.S. Bell, L. Bergström, *J. Am. Ceram. Soc*. **83** (2000) 1557–1574.
105. M.F. Yan, W.W. Rhodes, *Mater. Sci. Eng*. **61** (1983) 59–66.
106. P. Sarkar, P.S. Nicholson, *J. Am. Ceram. Soc*. **79** (1996) 1987–2002.
107. C. Kaya, P.A. Trusty, C.B. Ponton, *Br. Ceram. Trans*. **97** (1998) 48–54.
108. A.R. Boccaccini, U. Schindler, H.-G. Krüger, *Mater. Lett*. **51** (2001) 225–230.
109. S.C. Wang, W.C.J. Wei, *J. Am. Ceram. Soc*. **84** (2001) 1411–1414.
110. D.M. Owen, A.H. Chokshi, *Acta Mater*. **46** (1998) 719–729.

111. M.N. Rahaman, in *Sintering Technology*, eds R.M. German, G.L. Messing, R. Cornwall, Marcel Dekker, New York (1996) 93–100.
112. S.J. Wu, L.C. De Jonghe, M.N. Rahaman, *J. Am. Ceram. Soc*. **79** (1996) 2207–2211.
113. C.S. Nordahl, G.L. Messing, *J. Eur. Ceram. Soc*. **22** (2002) 415–422.
114. V.V. Srdiç, M. Winterer, H. Hahn, *J. Am. Ceram. Soc*. **83** (2000) 729–736.
115. P. Durán, M. Villegas, J.F. Fernández, F. Capel, C. Moure, *Mater. Sci. Eng. A* **232** (1997) 168–176.
116. A.P. Hynes, R.H. Doremus, R.W. Siegel, *J. Am. Ceram. Soc*. **85** (2002) 1979–1987.
117. R. Cahn, *Materials Today* **4** (2001) 13.
118. H.G. Kim, K.T. Kim, *Acta Mater*. **47** (1999) 3561–3570.
119. Y. Shinoda, T. Nagano, F. Wakai, *J. Am. Ceram. Soc*. **82** (1999) 771–773.
120. H. Hahn, *Nanostruct. Mater*. **2** (1993) 251–265.
121. S.-C. Liao, J. Colaizzi, Y. Chen, B.H. Kear, W.E. Mayo, *J. Am. Ceram. Soc*. **83** (2000) 2163–2169.
122. S.-C. Liao, Y.-J. Chen, B.H. Kear, W.E. Mayo, *Nanostruct. Mater*. **10** (1998) 1063–1079.
123. B.H. Kear and W.E. Mayo, *Mater. Sci. Forum* **437–438** (2003) 399–402.
124. R.M. German, *Liquid Phase Sintering*, Plenum Press, New York (1985), pp. 163–164.
125. W.D. Kingery, J.M. Woulbroun, F.R. Charvat, *J. Am. Ceram. Soc*. **46** (1963) 391–395.
126. J.G. Wang, R. Raj, *J. Am. Ceram. Soc*. **67** (1984) 399–409.
127. Y.W. Kim, M. Mitomo, H. Hirotsuru, *J. Am. Ceram. Soc*. **78** (1995) 3145–3148.
128. D. Sciti, J. Vicens, N. Herlin, J. Grabis, A. Bellosi, *J. Ceram. Proc. Res*. **5** (2004) 40–47.
129. M.N. Rahaman, L.C. DeJonghe, *J. Am. Ceram. Soc*. **67** (1984) C205–C207.
130. M.N. Rahaman, L.C. DeJonghe, R.J. Brook, *J. Am. Ceram. Soc*. **69** (1986) 53–58.
131. K.R. Venkatachari, R. Raj, *J. Am. Ceram. Soc*. **69** (1986) 499–506.
132. K.R. Venkatachari, R. Raj, *J. Am. Ceram. Soc*. **70** (1987) 514–520.
133. R. Raj, *J. Am. Ceram. Soc*. **65** (1982) C46.
134. M. Uchic, H.J. Höfler, W.J. Flick, R. Tao, P. Kurath, R.S. Averback, *Scr. Metall. Mater*. **26** (1992) 791–796.
135. D.C. Hague, M.J. Mayo, *J. Am. Ceram. Soc*. **80** (1997) 149–156.
136. C.S. Nordahl, G.L. Messing, *J. Am. Ceram. Soc*. **79** (1996) 3149–3154.
137. L. Kim, T. Klassen, R. Averback, C. Altstetter, *Mater. Sci. Forum* **235–238** (1997) 881–886.
138. Y.J. He, J.A. Winnubst, H. Verweij, A.J. Burggraaf, *J. Mater. Sci*. **29** (1994) 5868–5874.
139. B. Budiansky, J.W. Hutchinson, S. Slutsky, in *Mechanics of Solids,* eds H.G. Hopkins, M.J. Sewell, Pergamon Press, Oxford (1982), pp. 13–45.
140. O.H. Kwon, C.S. Nordahl, G.L. Messing, *J. Am. Ceram. Soc*. **78** (1995) 491–494.
141. B.J. Kellet, F.F. Lange, *J. Am. Ceram. Soc*. **71** (1988) 7–12.
142. B.J. Kellet, F.F. Lange, *J. Am. Ceram. Soc*. **69** (1986) C172–C173.
143. G. Skandan, H. Hahn, B.H. Kear, M. Roddy, W.R. Cannon, *Mater. Lett*. **20** (1994) 305–309.
144. J. Kishino, A. Nishiyama, T. Sakuma, *J. Mater. Sci*. **31** (1996) 4991–4995.
145. H. Ye, G. Rixecker, S. Haug, F. Aldinger, *J. Eur. Ceram. Soc*. **22** (2002) 2379–2387.
146. R.S. Mishra, S.H. Risbud, A.K. Mukherjee, *J. Mater. Res*. **13** (1998) 86–89.
147. Z. Shen, M. Johnsson, Z. Zhao, M. Nygren, *J. Am. Ceram. Soc*. **85** (2002) 1921–1927.

148. Z. A. Munir, H. Schmalzried, *J. Mater. Synth. Proc*. **1** (1993) 3–16.
149. M. Suganuma, Y. Kitagawa, S. Wada, N. Murayama, *J. Am. Ceram. Soc*. **86** (2003) 387–394.
150. C. Wagner, *Z. Elektrochem*. **65** (1961) 581–591.
151. K.R. Lai, T.Y. Tien, *J. Am. Ceram. Soc*. **76** (1993) 91–96.
152. T. Nishimura, M. Mitomo, H. Hirotsuru, M. Kawahara, *J. Mater. Sci. Lett*. **14** (1995) 1046–1047.
153. N. Murayama, *Bull. Ceram. Soc. Jpn*. **32** (1997) 445–449.
154. M. Yoshimura, T. Ohji, M. Sando, K. Niihara, *J. Mater. Sci. Lett*. **17** (1998) 1389–1391.
155. T. Takeuchi, M. Tabuchi, H. Kageyama Y. Suyama, *J. Am. Ceram. Soc*. **82** (1999) 939–943.
156. L. Gao, Z.J. Shen, H. Miyamoto, M. Nygren, *J. Am. Ceram. Soc*. **82** (1999) 1061–1063.
157. Y. Zhou, K. Hirao, M. Toriyama, *J. Am. Ceram. Soc*. **83** (2000) 654–656.
158. Z. Shen, M. Johnsson, Z. Zhao, M. Nygren, *J. Am. Ceram. Soc*. **85** (2002) 1921–1927.
159. M. Nygren, Z. Shen, *Sol. State Sci*. **5** (2003) 125–131.
160. L. Zhou, Z. Zhao, A. Zimmermann, F. Aldinger, M. Nygren, *J. Am. Ceram. Soc*. **87** (2004) 606–611.
161. M.T. Buscaglia, V. Buscaglia, M. Viviani, J. Petzelt, M. Savinov, L. Mitoseriu, A. Testino, P. Nanni, C. Harnagea, Z. Zhao, M. Nygren, *Nanotechnol*. **15** (2004) 1113–1117.
162. B. Li, X. Wang, L. Li, H. Zhou, X. Liu, X. Han, Y. Zhang, X. Qi, X. Deng, *Mater. Chem. Phys*. **83** (2004) 23–28.
163. Y.-I. Lee, J.-H. Lee, S.-H. Hong, D.-Y. Kim, *Mater. Res. Bull*. **38** (2003) 925–930.
164. J. Wan, M.J. Gasch, A.K. Mukherjee, *J. Am. Ceram. Soc*. **86** (2002) 526–528.
165. G.-D. Zhan, J. Kuntz, J. Wan, J. Garay, A.K. Mukherjee, *Scr. Mater*. **47** (2002) 737–741.
166. Z. Shen, M. Johnsson, Z. Zhao, M. Nygren, *J. Am. Ceram. Soc*. **85** (2002) 1921–1927.
167. L. Gao, Q. Li, W. Luan, H. Kawaoka, T. Sekino, K. Niihara, *J. Am. Ceram. Soc*. **85** (2002) 1016–1018.
168. T. Takeuchi, C. Capiglia, N. Balakrishnan, Y. Takeda, H. Kageyama, *J. Mater. Res*. **17** (2002) 575–581.
169. T. Takeuchi, I. Kondoh, N. Tamari, N. Balakrishnan, K. Nomura, H. Kageyama, Y. Takeda, *J. Electrochem. Soc*. **149** (2002) A455–A461.
170. M.J. Gasch, J. Wan, K.C. Liu, E. Lara-Curzio, A.K. Mukherjee, in *Creep Deformation: Fundamentals and Applications* (2002 TMS Annual Meeting), eds R.S. Mishra, J.C. Earthman, S.V. Raj, The Minerals, Metals & Materials Society, Warrendale (2002), pp. 247–56.
171. J. Wan, M.J. Gasch, A.K. Mukherjee, *J. Am. Ceram. Soc*. **85** (2002) 526–528.
172. S.H. Risbud, C.H. Shan, A.K. Mukherjee, M.J. Kim, J.S. Bow, R.A. Holl, *J. Mater. Res*. **10** (1995) 740–742.
173. B. Basu, J.-H. Lee, D.-Y. Kim, *J. Am. Ceram. Soc*. **87** (2004) 317–319.
174. G.D. Zhan, J.D. Kuntz, J. Wan, J. Garay, AK. Mukherjee, *J. Am. Ceram. Soc*. **86** (2003) 200–203.
175. L. Gao, H. Wang, H. Kawaoka, T. Sekino, K. Niihara, *J. Eur. Ceram. Soc*. **22** (2002) 785–789.

176. L. Gao, X. Jin, H. Kawaoka, T. Sekino, K. Niihara, *Mater. Sci. Engg. A* **334** (2002) 262–266.
177. G.D. Zhan, J. Kuntz, J. Wan, J. Garay, A.K. Mukherjee, in *Proc. Ultrafine Grained Materials II*. (TMS Annual Meeting), eds Y.T. Zhu, T.G. Langdon, R.S. Mishra, S.L. Semiatin, M.J. Sharan, T.C. Lowe, The Minerals Metals & Materials Society, Warrendale (2002) 219–224.
178. G.D. Zhan, J. Kuntz, J. Wan, J. Garay, A.K. Mukherjee, *Scr. Mater*. **47** (2002) 737–741.
179. G.-D. Zhan, J.D. Kuntz, J.E. Garay, A.K. Mukherjee, *Appl. Phys. Lett.* **83** (2003) 1228–1230.
180. T. Nakayama, B.-S. Kim, H. Kondo, Y.-H. Choa, T. Sekino, M. Nagashima, T. Kusunose, Y. Hayashi, K. Niihara, *J. Eur. Ceram. Soc*. **24** (2004) 259–264.
181. R. Apetz, M.P.B. van Bruggen, *J. Am. Ceram. Soc*. **86** (2003) 480–486.
182. O.-H. Kwon, C.S. Nordahl, G.L. Messing, *J. Am. Ceram. Soc*. **78** (1995) 491–495.
183. D. Godlinski, M. Kuntz, G. Grathwohl, *J. Am. Ceram. Soc*. **85** (2002) 2449–2456.
184. A. Krell, P. Blank, H. Ma, T. Hutzler, M.P.B. van Bruggen, R. Apetz, *J. Am. Ceram. Soc*. **86** (2003) 12–18.
185. M.J. Mayo, in *Mechanical Properties and Deformation Behavior of Materials Having Ultra-Fine Microstructures*, eds M. Nastasi *et al*., Kluwer Academic Publishers, Dordrecht (1993), pp. 361–380.
186. F. Wakai, N. Kondo, Y. Shinoda, *Curr. Opinion Sol. State Mater. Sci*. **4** (1999) 461–465.
187. I. Amato, *Mater. Sci. Forum* 455–456 (2004) 550–555.
188. S. Benfer, P. Árki, G. Tomandl, *Adv. Eng. Mater*. **6** (2004) 495–500.
189. G. Battilana, A. Bottino, G. Capannelli, P. Nanni, *J. Mater. Sci*. **37** (2002) 4343–4347.
190. Y. Kinemuchi, T. Suzuki, W. Jiang, K. Yatsui, *J. Am. Ceram. Soc*. **84** (2001) 2144–2146.
191. T.J. Webster, C. Ergun, R.H. Doremus, R.W. Siegel, R. Bizios, *Biomater*. **22** (2001) 1327–1333.
192. T.J. Webster, C. Ergun, R.H. Doremus, R.W. Siegel, R. Bizios, *Biomater*. **21** (2000) 1803–1810.
193. S.A. Catledge, M.D. Fries, Y.K. Vohra, *et al., J. Nanosci. Nanotechnol*. **2** (2002) 293–312.
194. T.J. Webster, *Am. Ceram. Soc. Bull*. **82** (2003) 23–28B.
195. M.T. Buscaglia, V. Buscaglia, M. Viviani, J. Petzelt, M. Savinov, L. Mitoseriu, A. Testino, P. Nanni, C. Harnagea, Z. Zhao, M. Nygren, *Nanotechnology* **15** (2004) 1113–1117.
196. C.-W. Nan, A. Tschöpe, S. Holten, H. Kliem, R. Birringer, *J. Appl. Phys*. **85** (1999) 7735–7740.
197. O.K. Varghese, C.A. Grimes, in *Encyclopedia of Nanoscience and Nanotechnology*, Vol. 5, ed. H.S. Nalwa, American Scientific Publishers, Stevenson Ranch (2004), pp. 505–521.
198. E. Traversa, O. Schäf, E. Di Bartolomeo, P. Knauth, in Nanocrystalline Oxides, eds P. Knauth, J. Schoonman, Kluwer Academic Publishers, Norwell (2002), pp. 189–207.
199. B. Levy, *J. Electroceram*. **1** (1997) 239–272.
200. D. Perednis, L.J. Gauckler, *Sol. State Ionics* **166** (2004) 229–239.
201. P. Heitjans, S. Indris, *J. Phys. – Cond. Matter* **15** (2003) R1257–R1289.

202. P. Mondal, A. Klein, W. Jaegermann, H. Hahn, *Sol. State Ionics* **118** (1999) 331–339.
203. G. Herzer, *Mater. Sci. Engg. A* **133** (1991) 1–5.
204. M. Venkatesan, S. Nawka, S.C. Pillai, J.M.D. Coey, *J. Appl. Phys*. **93** (2003) 8023–8025.
205. J. Weissmüller, in: *Proc. 22nd Risø International Symposium on Materials Science*, Risø National Laboratory, Roskilde (2001), pp. 155–176.
206. G. Gottstein, L.S. Shvindlerman, *Z. Metallkd*. **95** (2004) 219–222.

9

Nanoengineering of metallic materials

R LUMLEY and A MORTON,
CSIRO, Australia and I POLMEAR, Monash University, Australia

9.1 Introduction

Nanoengineering, by definition, is the science and technology involved with modifying or manipulating the microstructure of *bulk* materials on a nanoscale. Such materials may display enhanced mechanical, physical or chemical properties, and hence nanoengineering is of paramount importance to the evolution of new metallic materials. Rather than dealing with milligrams or grams of material, tonnes of materials are processed, at prices that are economical and sustainable for modern manufacturing.

In this chapter, the principles underlying the nanoengineering of metallic materials are first considered in terms of those criteria that have the most significant effect on the mechanical properties of metals and alloys namely, the effect of grain size and the role of obstacles that retard deformation by the process of crystallographic slip. Relationships between processing, microstructure, properties and cost are discussed, followed by an introduction to manufacturing routes used in the processing of metallic materials. Four examples are then described in which microstructures are manipulated on the nanoscale in order to improve properties. These examples are concerned with the precipitation strengthening of aluminium alloys, the development of high strength low alloy steels, mechanical alloying, and rapid solidification processing.

9.2 Nanoengineering of metallic materials

Traditional metallurgy has been involved with mineral-bearing ores being mined, refined, and processed into metals that are used in day-to-day living. The metals are melted, alloyed with other elements, cast, and often fabricated into wrought products by hot and cold working methods such as rolling, extrusion or forging. It is during these various processes that changes to the microstructure of the metals and alloys occur that are often on a nanoscale.

In metals and alloys, improved strength is commonly achieved through control of grain size and shape, and by introducing into the grains themselves obstacles that restrict the movement of defects (dislocations) that are both generated and propagated during plastic deformation. In a practical sense, it is the yielding behaviour of metals and alloys that is a major design criterion and, because of this, special attention is paid to characterising material performance using the stress–strain curve.[1]

9.2.1 The Hall–Petch effect

The Hall–Petch effect is a direct relationship between grain size and strength of a metal or alloy, and is described by:[2,3]

$$\sigma_y = \sigma_i + kd^{-1/2}$$

where σ_y is the yield strength; σ_i is the overall resistance of the lattice to dislocation movement; k is the grain boundary locking term which measures the relative hardening contribution by the grain boundaries, and d is the grain diameter.

As the grain sizes of metallic materials approach the nano-scale, the strength increases arising from the Hall–Petch relationship can become quite dramatic. This occurs because of the inherently high proportion of grain boundaries at which dislocations may become immobilised. However, once the grain size of metals or alloys passes below a critical size (~10 nm), the strength of the material may decrease because the proportion of grain boundaries now becomes so high that the grains can move readily with respect to each-other. This 'grain size softening' is known as the inverse Hall–Petch Effect.[4–6] It can, however, be utilised as the materials may display abnormal (super-plastic) deformation so that extensions exceeding 1000% are possible without failure occurring.[7,8]

9.2.2 The role and manipulation of nanoscale obstacles

The engineering of nanoscale obstacles into the microstructure of metals and alloys increases strength by impeding the motion of defects (dislocations) during deformation. Such obstacles are:

1. Other dislocations
2. Vacancies (vacant lattice sites)
3. Solute atoms
4. Precipitates formed through the decomposition of supersaturated solid solutions
5. Fine intermetallic compounds or ceramic particles.

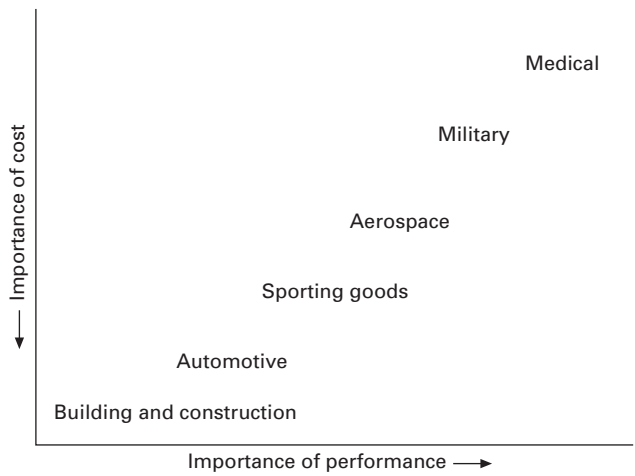
All of these obstacles (with the exception of ceramic particles) have associated strain fields, which distort the crystal lattice and further impede the motion of dislocations. Solute atoms may segregate to dislocations, which also impedes their motion, whereas precipitates or other particles must either be cut or by-passed by the dislocations before these defects can continue to pass through the material.

9.2.3 Relationships between processing/structure/properties/cost

The means by which a nanoengineered metal or alloy is produced, and the properties that it displays, are closely related to the method of processing, which also influences the final cost. The relationship between material cost and performance for different industries can be represented in the manner shown in Fig. 9.1. In relation to nanoengineered metals and alloys, the general rules are that the more difficult it is to control the structure of the metal, the more unusual is the component and the more stringent the quality criteria, the more expensive it is to produce. Frequently, the main cost factor is the number of processing steps involved in producing a component.

9.2.4 Manufacture of nanoengineered metals and alloys

Through the traditional metallurgical processing routes mentioned in the introduction, fine, nanoscale microstructures can be obtained in bulk materials, and many tonnes may be produced daily. Components can also be mass-



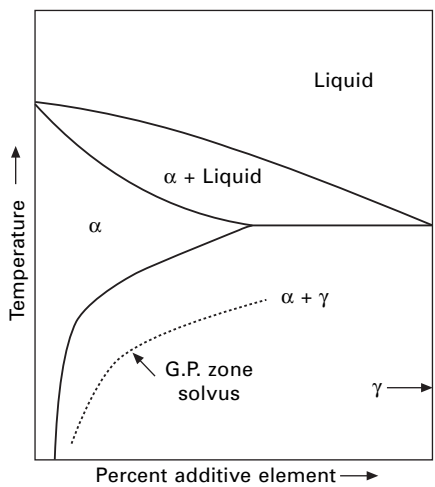
9.1 The relationship between cost and performance in manufacturing industries.

produced from metal powders where the advantage lies in being able to achieve an intricate product shape in one step. More recent techniques in which nanoscale microstructures and improved properties can be obtained are mechanical alloying and rapid solidification processing. However, these methods are costly and so far have been applied only to small-scale manufacture of products.

9.3 Age hardened alloys

Age hardening was first reported in a ternary Al-Cu-Mg alloy in 1911 by Alfred Wilm.[9] Accounts of the discovery record that, after the Al-Cu-Mg alloy had been held at high temperature and quenched to room temperature late one week, it was found to be soft. However, when examined again on the following Monday it had undergone significant hardening during the weekend. The alloy (now known as Duralumin) was quickly adopted for the structure of the early Zeppelin airships and for the first all-metal passenger aircraft, the Junkers F13, that first flew in 1919.

It was also in 1919 that it was first realised the phenomenon of age hardening was associated with the decreasing solid solubility of alloying elements with decreasing temperature, as shown in Fig. 9.2. This led to the hypothesis that quenching from a high 'solution treatment' temperature (Ta) could suppress the equilibrium separation of a second phase, as expected from the phase diagram, and result in the formation of an unstable supersaturated



9.2 Model phase diagram for an alloy capable of undergoing precipitation strengthening. Solution treatment is conducted within the single phase α region of the phase diagram, prior to quenching. Upon heat treatment, the alloy is again heated into the α+ γ phase region, where γ is typically an intermetallic phase (e.g. A_2B).

solid solution (SSSS).[10] It was proposed that 'ageing' the quenched alloy for sufficient time may then allow a submicroscopic dispersion of the second phase (γ) to form, which acted as an obstacle to crystallographic slip during deformation.[11]

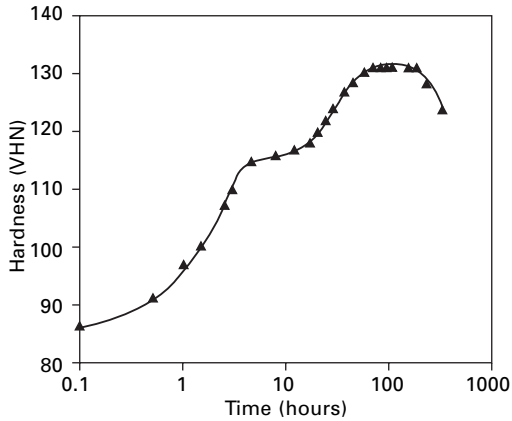
Subsequent investigations showed that the complete process of age or precipitation hardening may involve several stages commencing with the formation of fine, coherent zones that were first detected in 1938 by X-ray diffraction techniques.[12,13] However, it was not until the development of transmission electron microscopy in 1956 that much of the fine detail associated with precipitation hardening was revealed. Precipitation hardening became recognised as the most common method for strengthening by heat treatment and has also been applied to alloys based on other metals including copper, titanium, nickel, magnesium and lead. The major challenge in using age hardening for strengthening of alloys is to promote the formation of precipitates that are large enough to resist being sheared by dislocations, yet are too finely spaced to be by-passed. This critical condition is often difficult to produce in practical alloys, but a near approximation is often able to be achieved.

Historically, special attention has been paid to studying the mechanisms of precipitation hardening in the Al-Cu system in which the complete precipitation sequence has been identified as:

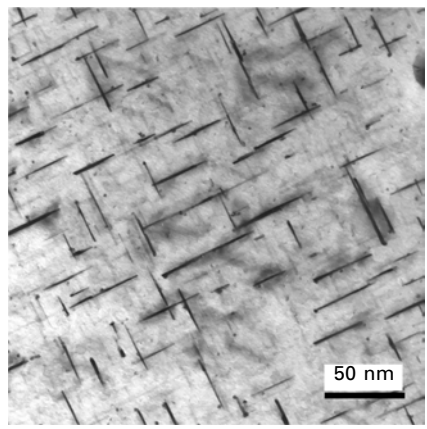
$$\text{S.S.S.S} \rightarrow \text{GP zones} \rightarrow \theta'' \rightarrow \theta' \rightarrow \theta.$$

GP (Guinier–Preston) zones are ordered, solute-rich clusters of atoms that are coherent with the parent lattice and are of only one or two atomic planes in thickness. They may be present in densities as high as $10^{18}/\text{cm}^3$. Although they are readily sheared by moving dislocations, they are still able to cause significant strengthening because of the strain they impart into the surrounding lattice. The phase θ'' is also coherent with the aluminium lattice, and can be regarded as a higher order zone. The intermediate phase θ' is semi-coherent and occurs as plates that are typically much larger than the θ''. The equilibrium precipitate θ (Al_2Cu) is incoherent with the aluminium lattice and, because it is normally coarsely dispersed, causes little strengthening. In this condition an alloy is termed overaged. The hardness–time curve for an Al-4Cu alloy aged at 150 °C is shown in Fig. 9.3(a). The microstructure of the alloy at the peak aged condition in shown in Fig. 9.3(b), and contains both the θ'' and θ' phases. Typically this alloy may display a 0.2% proof stress of ~240 MPa, a tensile strength of ~325MPa, and an elongation at failure of 5%.

Quenching from a high temperature also results in a supersaturation of vacancies that are available to facilitate solute diffusion during ageing. With Al-Cu alloys, for example, which age harden at room temperature, it has been estimated that diffusion rates are 10^7 faster than expected because of the presence of vacancies retained by quenching from high temperature.[14]



(a)



(b)

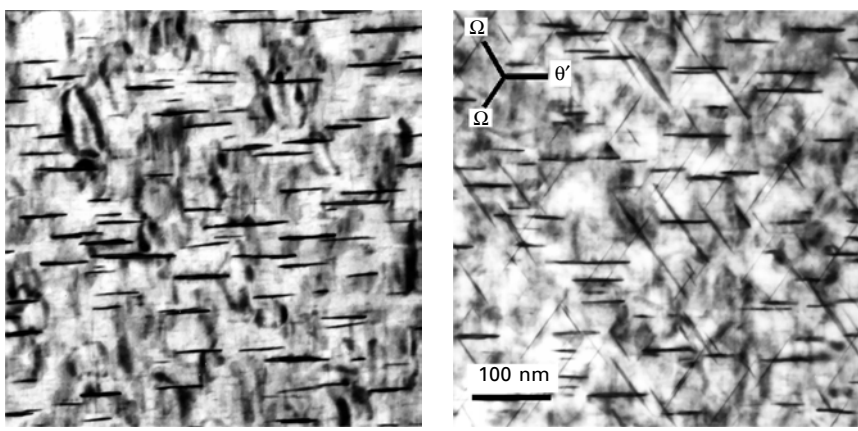
9.3 Transmission electron migrograph showing hardening diagram of Al-4Cu aged at 150 °C (a) and the corresponding precipitate structure at the peak hardened condition in an [001]α crystallographic orientation (b). Fine particles of θ″ and θ′ are shown within the material edge on, that provide strengthening to the alloy.

Vacancies also promote clustering of solute atoms and this phenomenon may contribute to precipitation processes in some alloys. Although it is not possible to observe vacancies directly in aluminium alloys using techniques such as transmission electron microscopy or atom probe field ion microscopy, positron annihilation spectroscopy (PAS) now allows quantitative measurements to be made of solute atom/vacancy interactions.[15]

9.3.1 Chemistry and alloy design in precipitation strengthened aluminium alloys

Wrought aluminium alloys that are amenable to precipitation strengthening may be classed according to the major alloying elements that are present, and are divided into three groups. These are the 2000 series alloys (Al-Cu-Mg), the 6000 series (Al-Mg-Si), and the 7000 series (Al-Zn-Mg).[16] The microstructures of wrought materials can be more readily refined than cast alloys because of the working processes and annealing treatments to which they are subjected. Both hot and cold working causes fracture and dispersal of the intermetallic particles that form during casting; grain shape is changed and dynamic recrystallisation may occur that greatly increases or refines the grain size.

For the 2000 series alloys, the ratio of Cu:Mg is key to their ageing behaviour because the nanoscale precipitates that form are mixtures of phases based on the compounds (Al_2Cu and Al_2CuMg) and the proportion of each depends on the alloy composition. Microalloying additions of Ag may promote the formation of a range of three additional fine precipitates, typically known as Ω, X' and Z phases,[17] and examples of peak aged Al-Cu-Mg and Al-Cu-Mg-Ag alloys are shown in Fig. 9.4. Because such additions can modify precipitation processes, they may exert a disproportionate effect on strengthening. Some alloys of the 2000 series also contain Li, such as the composition Al-6.3Cu-1.3Li-0.4Mg-0.4Ag-0.18Zr, that can result in yield strengths exceeding 700MPa.[18] This is impressive considering that pure, annealed aluminium has a yield strength of only ~20MPa,[19] and the theoretical maximum value is 900 MPa.



9.4 Peak aged microstructures of (a) an Al-Cu-Mg and (b) an Al-Cu-Mg-Ag alloy aged to peak properties, in [101]α orientations displaying major precipitate phases. The Ag containing alloy exhibits plates of the Ω phase, which modifies the mechanical behaviour of the alloy.

Precipitation in 6000 series alloys is particularly complex as GP zones may be preceded by the formation of disordered clusters containing less than 40 atoms.[20] Later precipitates have compositions that progressively approach the equilibrium phase Mg_2Si. Similar reactions occur in the 7000 series alloys based on the Al-Zn-Mg system in which the GP zones and an intermediate precipitate have Mg and Zn atoms in the approximate ratio of 1:1.[21] The equilibrium precipitate is $MgZn_2$. Typical precipitation sequences for 2000, 6000 and 7000 series alloys are given in Table 9.1.[22]

Cast metals and alloys generally cost less than wrought materials, but there are limited opportunities to produce highly refined microstructures. Nevertheless, some refinement is often possible with cast aluminium alloys that are amenable to precipitation hardening, and these are used for a range of automotive applications. For the casting alloys 356 and 357, for example, which are commonly used for wheels, the major addition is silicon (6.5–7.5 wt%), which produces high fluidity in the molten alloy allowing the casting of intricate components. However, the binary alloy is too brittle to be used because of the presence of an abnormal eutectic containing large plates of silicon. Some grain refinement is possible by adding small amounts of titanium (e.g. 0.05%) in the melt but the major improvement comes from the refinement (modification of the eutectic) of the silicon plates by the addition of strontium or sodium prior to casting. Further strengthening may be achieved by adding ~0.5% Mg, which induces fine, nanoscale precipitates of MgSi and Mg_2Si to form through the α aluminium grains, if the components are subsequently age hardened.

9.3.2 Heat treatment

Apart from modifying the sequences of precipitates that may occur through changes in the chemistry of an alloy, it is the ageing heat treatment that controls their size and distribution which then affects mechanical behaviour. For example, the precipitation process occurring in an alloy that is solution treated, quenched, and aged at room temperature (termed natural ageing, or a T4 temper) is less advanced than that for the same alloy is aged at 150 °C (termed a T6 temper). Many different means have been investigated by which heat treatment of aluminium alloys may modify precipitation processes in order to benefit the mechanical, thermal or chemical behaviour of alloys.[22, 23] All have the common theme of improving the range and application of alloys in engineering service through control of their nanostructures.

One recent development in the heat treatment of aluminium alloys has shown that nucleation of GP zones may be made to occur more than once by a process referred to as interrupted ageing.[24] By this method, a solution treated and quenched alloy is initially aged for a relatively short period at an elevated temperature to produce an underaged microstructure. If the alloy is

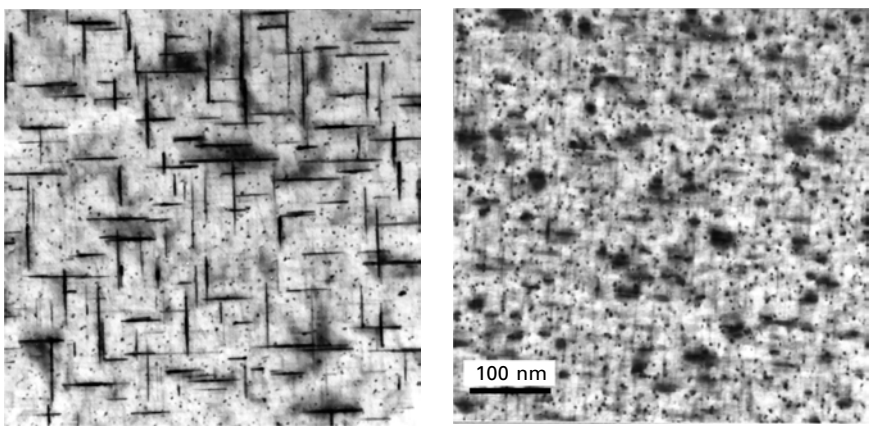
Table 9.1 Precipitation sequences of 2000, 6000 and 7000 series aluminium alloys[22]

Alloy		Precipitates	Comment
Al-Cu	Precipitation sequence ↓	Coherent GP zones as thin plates on $\{001\}_\alpha$	Cu monolayers on $\{001\}_\alpha$.
		Coherent θ″ (Al_2Cu) on $\{001\}_\alpha$	May nucleate at GP zones. Two layers of Cu separated by three layers of Al atoms.
		Semicoherent θ′ (Al_2Cu) on $\{001\}_\alpha$	Tetragonal; may nucleate from dislocations.
		Incoherent θ	May nucleate at θ′. Body centred tetragonal.
Al-Cu-Mg (2000 series)		Mg/Cu Clusters	No crystalline structure but cause rapid strengthening.
	Precipitation sequence ↓	GPB (Cu, Mg) zones as fine rods along $<100>_\alpha$	(Note precipitates from binary Al-Cu alloys can also form where Cu:Mg ratio is high.)
(If Ag is added, Ω phase may form)		Semicoherent S′ (Al_2CuMg) as laths in $\{210\}_\alpha$ along $<001>_\alpha$	Orthorhombic; may nucleate at dislocations.
		Incoherent S (Al_2CuMg)	Orthorhombic; probably transforms from S′.
Al-Mg-Si (6000 series)		Clusters of discrete Mg or Si atoms	Clustering of Si may be detrimental to strength.
	Precipitation sequence ↓	Co-clusters of Mg and Si atoms	May be GP zones resulting from ordering.
		Coherent β″ (Mg_2Si / MgSi) as needles along $<100>_\alpha$	Monoclinic.
		Semicoherent β′ (Mg_2Si) as rods along $<100>_\alpha$	Hexagonal, may form from β″.
		Semicoherent B′ (Mg_2Si) as laths along $<100>_\alpha$	Hexagonal, forms together with β′, favoured by high Si:Mg ratios.
		β (Mg_2Si), as platelets on $\{100\}_\alpha$	Face centred cubic; may transform directly from β′.
Al-Zn-Mg-(Cu) (7000 series)		Two types of GP zones as spheres; clusters may also form in Cu-containing alloys,	May have variable composition from MgZn to $MgZn_2$.
	Precipitation sequence ↓	Semicoherent η′ ($MgZn_2$) along $\{111\}_\alpha$	Hexagonal; may form from GP zones when Zn:Mg>3:1.
		Incoherent η ($MgZn_2$)	Hexagonal; 9 orientation relationships possible.
		Semicoherent T′ $Mg_{32}(Al,[Zn/Cu])_{49}$	Hexagonal; may form instead of η in alloys with high Mg:Zn ratios.
		Incoherent T $Mg_{32}(Al,[Zn/Cu])_{49}$	Cubic; may form from η if ageing temp. >190 °C, or from T′.

then quenched again, secondary precipitation may occur at a lower temperature that appears to continue indefinitely. Secondary precipitation has several effects on the microstructure of the alloy. The pre-existing precipitates that are formed during the initial ageing period may be chemically modified, and new GP zones may form which provides additional opportunities to control the microstructure. An example of this effect in an Al-Cu-Mg alloy is shown by comparing Figs 9.5(a) and 9.5(b). Precipitates in the conventionally aged alloy (Fig. 9.5(a)) are substantially coarser than those present after interrupted ageing (Fig. 9.5(b)), in which there is a 75% increase in the proportion of the phase Al_2CuMg, that is present (shown here as rods viewed end on, i.e., as 'dots'). By this means, it is possible to tailor the nature of the precipitate particles on the nanoscale such that an alloy may display improved service performance. It may also be noted that heat treatment processes incorporating secondary precipitation may produce simultaneous improvements in yield strength and fracture toughness in aluminium alloys, whereas these properties are normally inversely related. Table 9.2 provides a comparison of the tensile properties of a range of common precipitation hardened aluminium alloys subjected to either the common artificial ageing temper (T6), or to the modified, interrupted aged temper, referred to as T616 temper

9.4 High strength low alloy (HSLA) steels

For almost all of the 3400 years since the Hittites of ancient Mesopotamia



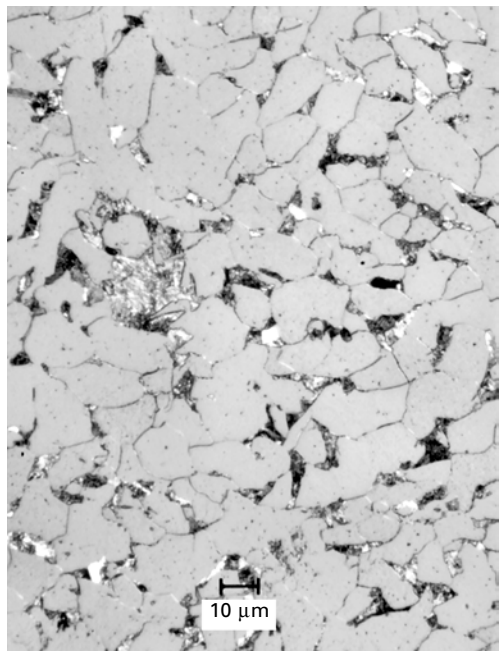
9.5 Peak aged microstructures of an Al-Cu-Mg alloy in (a) the conventionally aged T6 condition and (b) the novel interrupted aged condition in $[001]_\alpha$ orientations displaying major precipitate phases. θ′ precipitates are viewed as perpendicular plates edge on, and S phase is viewed end on as rods less than 5 nm in diameter ('dots'). Not only is the microstructure shown in (b) much finer, the proportion of precipitate phases present has been altered.

Table 9.2 Tensile properties of alloys aged by T6 and T6I6 treatments[24]

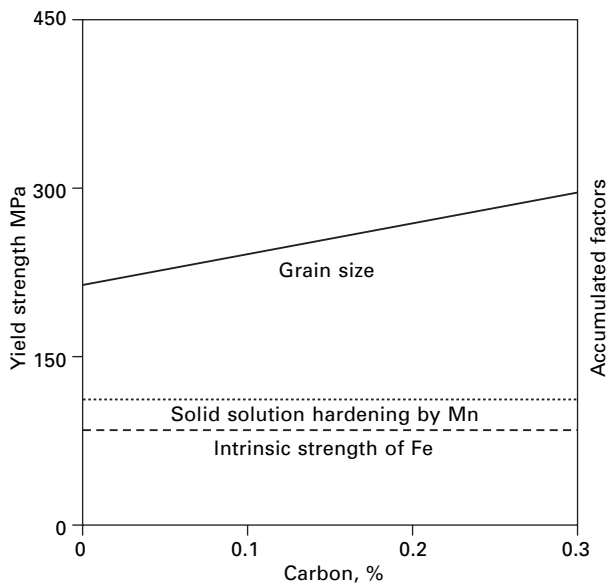
Ageing treatment ⇒	T6			T6I6		
Alloy	0.2% Proof stress (MPa)	Tensile strength (MPa)	Elongation (%)	0.2% Proof stress (MPa)	Tensile strength (MPa)	Elongation (%)
Al-4Cu	236	325	5	256	358	7
2014	414	488	10	436	526	10
Al-Cu-Mg-Ag	442	481	12	502	518	7
6061	267	318	13	299	340	13
6013	339	404	17	380	416	15
7050	546	621	14	574	639	14
7075	505	570	10	535	633	13
8090	349	449	4	391	512	5
357	287	340	7	341	375	5

became the first civilised people to smelt, heat treat and use a primitive form of steel, carbon has served as the main alloying element. Furthermore, the achievement of adequate strength and hardness have been the major objectives during much of this time. Currently, iron and steel still account for over 90% of all metals produced and more than 80% of this is the comparatively simple, plain carbon alloy known as mild steel (nominal composition 0.1%C–0.8%Mn). It is the small amount of carbon that controls the basic microstructure which develops from the high temperature phase austenite, when mild steel ingots are hot worked, e.g., by rolling, and cooled to room temperature. Some 85–90% of this microstructure consists of grains of nearly pure iron, ferrite, and the remainder comprises colonies of a lamellar ferrite/cementite (Fe_3C) eutectoid known as pearlite (Fig. 9.6). Some of the manganese present dissolves in the iron and strengthens the ferrite, although its major role is to stabilise sulphur impurities as relatively innocuous MnS particles, and to assist in de-oxidising molten steel during melting and casting into ingots. Factors that contribute to yield strength in mild steel are the intrinsic strength of iron, solid solution strengthening by Mn, and grain size.[25] Grain size is the most important factor and, since yield strength is inversely related to the square root of grain size according to the Hall–Petch relationship discussed earlier, significant increases may be achieved if grain size is reduced (Fig. 9.7). However, there is a limit to which fine grain sizes can be achieved in mild steels due to the dynamic recrystallisation and grain growth that occurs in the high temperature phase austenite during hot working. Typically, the yield strength of a hot rolled and air cooled (normalised) mild steel is relatively low and lies in the range of 225–250 MPa.

Although increasing strength is an ongoing goal in metallurgical alloy development, more attention is now focused on improving other critical



9.6 Typical microstructure of mild steel. Ferrite grains are surrounded by lamellar pearlite.



9.7 Factors contributing to the yield strength of mild steel. Adapted from Irvine and Pickering.[25]

properties such as toughness and weldability, each of which is adversely affected by the presence of carbon.[26, 27] In 1936 it was found that microalloying additions of the metal niobium could strengthen mild steels although the underlying mechanism could not be identified at that time. This discovery was patented by the Union Carbide Corporation in the USA, but further development was not pursued largely because of the high cost of niobium.[26] Subsequently, in the late 1950s, an interesting interplay of technology and economics occurred when the price of niobium fell and there was a simultaneous demand for improved strength, toughness and weldability for steels used in gas pipelines. This led to the development of the so-called high strength low alloy (HSLA) steels in which the role of microalloying additions (e.g. 0.05 wt%) of niobium and/or vanadium has played a vital role. Although these additions may have differing effects on microstructure and mechanical properties, all are associated with precipitation of nanoscale particles within the matrix of the steels. Some of these effects are discussed below.

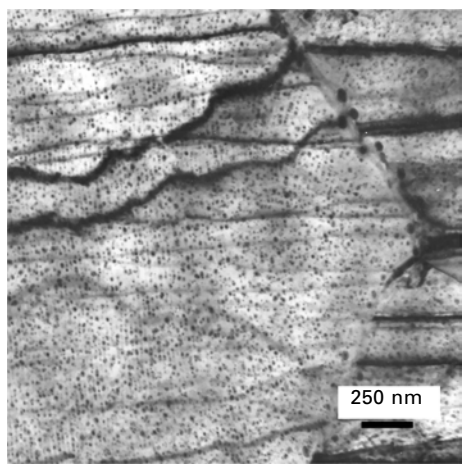
9.4.1 Grain refinement and the role of Nb and V in modifying the properties of HSLA steels

As mentioned above, refining grain size in metals and alloys increases strength as defined by the Hall–Petch relationship. Finer grain sizes also increase toughness, which is an uncommon coincidence since strength and toughness are normally inversely related as mentioned earlier. Grain refinement in HSLA steels may be accomplished by microalloying additions of niobium and vanadium each of which is a stronger carbide/nitride former than iron. In most instances these elements, together with carbon and nitrogen, are in solution in the high temperature phase austenite at the beginning of hot working operations. As the temperature falls, precipitation of nanoscale particles of Nb(C, N), V_4C_3 and/or VN may occur in the austenite grain boundaries, which act as obstacles to their migration and therefore hinder grain growth. Furthermore, during hot working at still lower temperatures, these particles may inhibit actual recrystallisation of the deformed austenite grains so that, when austenite transforms to ferrite and pearlite, the ferrite grains are much reduced in size. Of the two microalloying additions, niobium is favoured because the solubility of Nb(C, N) in austenite is less than that of V_4C_3 or VN. As a consequence, more precipitates form when niobium is used so that grain refinement can be achieved with a smaller addition of this element, thereby reducing costs.

Another benefit arising from the formation of nanoscale niobium and vanadium carbides is that less carbon is available to form eutectoid pearlite. This reduces the amount of pearlite in the microstructure of HSLA steels and each 1% reduction in volume is known to reduce the temperature of the ductile/brittle transition by some 2 °C.[25]

9.4.2 Precipitation strengthening of HSLA steels

Although grain refinement has become the preferred mode of strengthening HSLA steels because it also improves toughness, precipitation strengthening can also be very important. The particles that precipitate in austenite grain boundaries during hot working at high temperatures, which have been described as being effective in controlling grain growth, do not themselves cause strengthening because they are relatively large and widely spaced. Fine and more densely dispersed particles can be formed, however, if the cooling rate during hot working is increased so that precipitation of nanoscale particles occurs at lower temperatures. These particles then form at the ferrite–austenite interfaces during the phase transformation of austenite, by a mechanism that differs from the classical precipitation reaction described earlier in this chapter, and which has been termed interphase precipitation.[28, 29] Interphase precipitation is promoted by first hot working at higher temperatures (e.g. 1100–1200 °C), at which the niobium and vanadium carbides/nitrides are dissolved; for this purpose, vanadium is favoured because its carbides and nitrides have a greater solubility in austenite than those of niobium. If these elements are retained in solid solution until the pro-eutectoid transformation of austenite to ferrite occurs, then the carbides and nitrides will precipitate because their solubility in ferrite is much less. These carbides/nitrides form as discontinuous bands of fine, nanoscale particles as the ferrite/austenite boundaries sweep through the steel. Particle sizes and spacing vary with transformation temperature (which is normally in the range 850 °C to 650 °C), and may be as small as 10 nm in diameter with spacings of 20 nm (e.g. Fig. 9.8). A schematic model



9.8 Interphase precipitation of carbides and nitrides within an Fe-0.08C-0.38V alloy, solution treated at 1250 °C and isothermally transformed at 700 °C. The maximum dimension of the particles is ~50 Å, and the sheet spacing is 210 Å. Photo courtesy of R.W.K. Honeycombe.

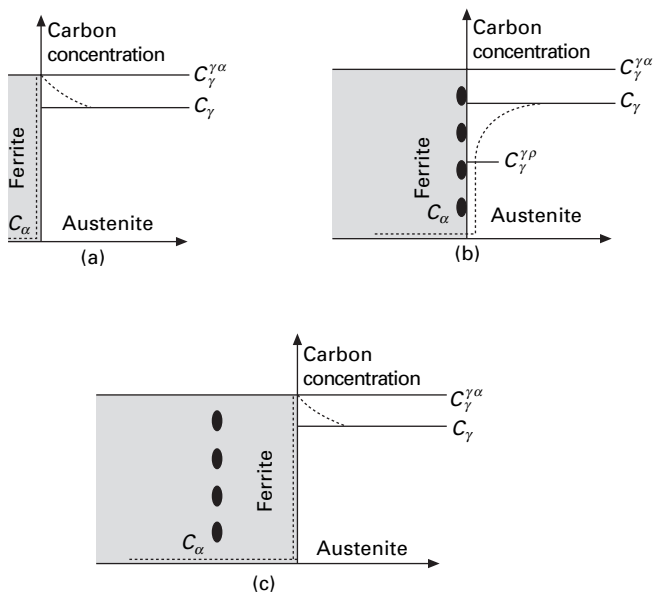
for this interphase precipitation has been proposed by Davenport *et al.* is shown in Fig. 9.9.[29]

From the practical viewpoint, Fig. 9.10 shows that the yield strengths of HSLA steels may exceed 450 MPa if interphase precipitation is promoted, despite the fact that grain size has increased because the carbides and nitrides are no longer available to pin austenite grain boundaries during hot working. Larger grain size means reduced toughness, however, which in practical terms means that each 15 MPa gain in yield strength through interphase precipitation raises the ductile/brittle transition temperature by 4 °C.[25] For this reason, invoking interphase precipitation to increase yield strength is usually confined to HSLA steels that are grain refined by other means.

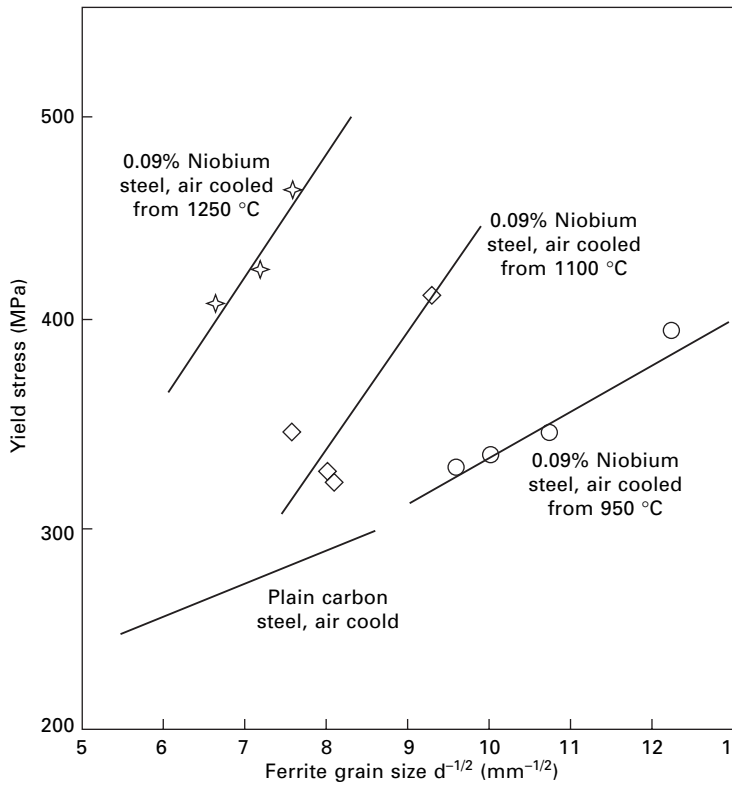
9.5 Mechanical alloying

9.5.1 Mechanical alloying

Mechanical alloying is a high-energy, powder manufacturing technique. Constituent powders and sometimes a process control agent (such as a wax-

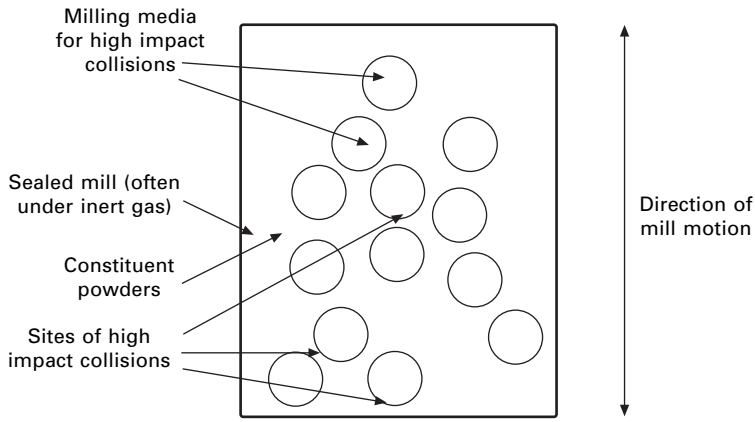


9.9 Model of interphase precipitation (adapted from Leslie).[26] (a) The carbon concentration in austenite increases at the ferrite–austenite boundary. (b) the carbon concentration promotes carbide nucleation on the ferrite side of the boundary, pinning the boundary and depleting the carbon concentration in the austenite. The depletion of carbon increases the driving force for the transformation of austenite, moving the boundary away from the precipitate particles. (c) The process may begin again.



9.10 The effect of austenitising temperature on the yield strength of a 0.1C-0.6Mn-0.09Nb (wt%) steel. Adapted from Gladman *et al.*[30]

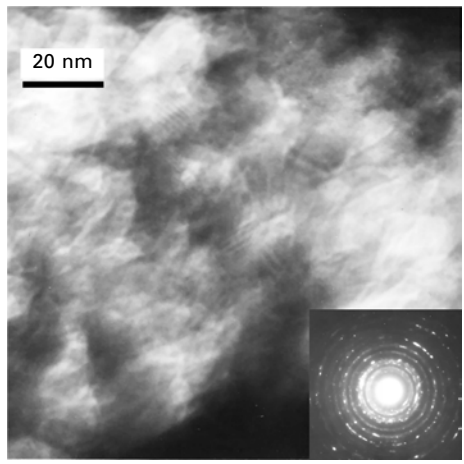
like stearic acid) are loaded into a mill under a controlled atmosphere, along with an impact imparting media, such as milling balls made from a ceramic or metallic material. The ratio of the mass of the balls to the mass of the powder is referred to as the charge ratio, and the higher this value, the greater is the collision frequency and hence speed of the milling process.[31] During milling, the balls and powders continually collide causing particles to fracture, weld, agglomerate, and then fracture again, and this production technique is shown schematically in Fig. 9.11. Several kinds of impact may occur that lead to alloys being formed. These are essentially due to impacts by compression and shear between ball–powder–ball or ball–powder–mill. Often, when one or more of the constituent powders is a relatively soft metal, agglomerates of elements form on the surfaces of the balls, and become sites of impact, or are sheared or broken off the surface. When these agglomerates fracture, they may leave the surface of the ball, and the process begins again. Several kinds of mills exist, like shaker mills or rolling mills such as those used in mineral comminution, but all are similar in their use and product.



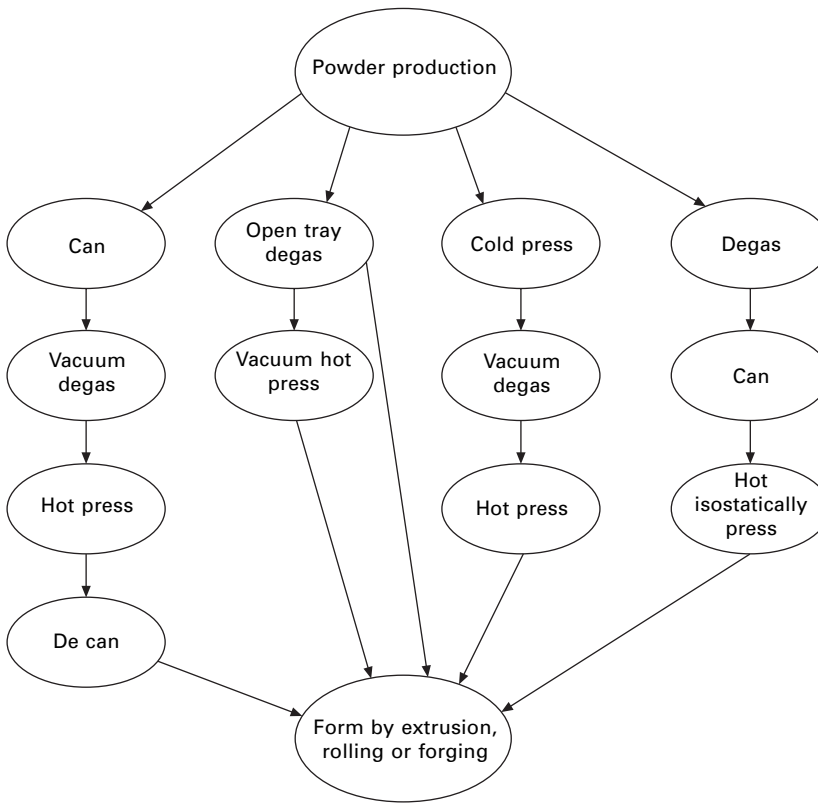
9.11 Schematic representation of mechanical alloying process. Powders are alloyed by high impact collisions between the milling media, which serve to fracture and re-weld particles, thus causing a nanoscale grain size to form or, alternatively, nanoscale particles to be incorporated into the base material.

The rate at which mechanical alloying occurs during milling is related exponentially to the rate at which the powder constituents accumulate lattice strain, and in turn, is dependent on the kinetic energy of the milling balls. The minimum requirement for mechanical alloying is for the forces of the impacts to exceed the yield stress of the powders, during processing.[31] These phenomena may produce extended solid solutions of constituent elements in the base material, cause the formation of stable and metastable intermetallic particles, introduce fine-scale oxides and carbides into the base material, or produce nanoscale grain sizes. Additionally, exothermic reactions may occur during the milling process, which may lead to the formation of intermetallic phases.

The structure of mechanically alloyed materials is, in essence, highly disordered. Grain size is reduced and the grain boundary area is greatly increased during the process, so that the disordered regions take on a greater and greater proportion of the volume fraction of the material as milling proceeds. Eventually, the material may become amorphous. Typically in the milled condition, powder agglomerates may display an internal grain size of, for example, 5–20 nm, as shown in Fig. 9.12. These powders are then prepared as final product by one or more applications of temperature and pressure (e.g. Fig. 9.13).[32] The primary goal in the consolidation and manufacture of nano-engineered, mechanically alloyed powdered materials is to retain as fine a grain size as is possible in the final material, while achieving as close as possible to 100% density. This may be achieved by hot isostatic pressing the powders into a solid bulk material. Powders are sealed inside ‘cans’ and heated at high temperature and pressure, which serves to consolidate the



9.12 Many beam TEM micrograph of an as-milled powder blend of WC-Ni-Cr displaying grain sizes of 5–20 nm. Corresponding diffraction rings are typical of nanocrystalline grain sizes in metallic materials, and are indicative of a large number of grains with differing orientations within the material.



9.13 Production routes for the preparation of solid materials from powder constituents. Adapted from *Metals Handbook*, 9th edn.[32]

material. Here it should be noted that the application of temperature causes recovery (of dislocations) and recrystallisation within the fine scale microstructure. Grain growth may be considerable, increasing grain diameter by up to 1000 times or more, such that the final grain size may or may not still be nanocrystalline.

Of the many current or potential uses for mechanically alloyed materials, the application for which the process was originally designed remains one of the most common in use today.[33] This is to provide dispersion strengthening of alloys through the introduction of nanoscale oxide and carbide particles. Several iron and nickel-base superalloys make use of these hard particles, mostly in applications for which resistance to deformation and oxidation resistance at elevated temperatures are most important. The compositions of a selection of commercially produced iron and nickel based oxide dispersion strengthened alloys, are shown in Table 9.3. Some aluminium alloys making use of hard carbide or oxide particles have also been used in high-performance forgings. Other potential applications of mechanically alloyed materials include magnetic materials for electric motors, rock drill buttons, high performance Ti-based intermetallics for gas turbines, and components for hydrogen storage devices.

9.5.2 Structure type 1: Production of nanocrystalline or amorphous powders that improve subsequent processing where the final material is not nanocrystalline

The purpose of producing material in this way is to be able to use the extraordinarily fast diffusion rates of constituent species or atoms within the material, for processes or products that may exploit this feature.[34] One manufacturing process that makes advantageous use of high diffusivity is that of sintering, where a pre-compacted powder (either ceramic or metals and alloys) is further densified by holding at elevated temperature. Sintering in the solid state can roughly be described as being a process in which the change in volume (shrinkage) is directly proportional to the diffusivity of the chemical species present, and to time. As the temperature increases, the diffusivity of species across the interfaces and through the metal lattice increases exponentially, as it is temperature activated in an Arrhenius relationship. A longer time at lower temperature will produce the same effect as a shorter time at higher temperature. As a general rule, densification via solid state sintering of metallic and metal carbide materials will only occur at temperatures higher than half the melting point of the alloy. However, mechanical alloying techniques produce powder pre-mixes that may display solid state sintering at considerably lower temperatures due, almost entirely, to the abnormally high diffusion rates involved. The final material, when

Table 9.3 Compositions (wt%) of some typical mechanically alloyed oxide dispersion strengthened nickel and iron alloys (adapted from Capdevila and Bhadeshia)[33]

Alloy designation	Ni	Fe	C	Cr	Al	Mo	Ti	N	W	Ti_2O_3	Y_2O_3
MA6000	Bal	1.5	0.06	15.0	4.5	–	2.3	0.2	3.9	–	1.1
MA758	Bal	–	0.05	30.0	0.3	–	–	–	0.5	–	0.6
MA760	Bal	1.2	0.06	19.5	6.0	–	–	0.3	3.4	–	1.0
PM1000	Bal	3.0	–	20.0	0.3	–	0.5	–	–	–	0.6
DT2203Y05	–	Bal	–	13.0	–	1.5	2.2	–	–	–	0.5
DT	–	Bal	–	13.0	–	1.5	2.9	–	–	1.8	–
DY	–	Bal	–	13.0	–	1.5	2.2	–	–	0.9	0.5
ODM061	–	Bal	–	20.0	6.0	1.5	0.6	–	–	–	0.5
ODM331	–	Bal	–	13.0	3.0	1.5	0.6	–	–	–	0.5
ODM751	–	Bal	–	16.5	4.5	1.5	0.6	–	–	–	0.5
MA956	–	Bal	0.01	20.0	4.5	–	0.5	0.045	–	–	0.5
MA957	–	Bal	0.01	14.0	–	0.3	1.0	0.012	–	–	0.27
PM2000	–	Bal	<0.04	20.0	5.5	–	0.5	–	–	–	0.5
PM2010	–	Bal	<0.04	20.0	5.5	–	0.5	–	–	–	1.0

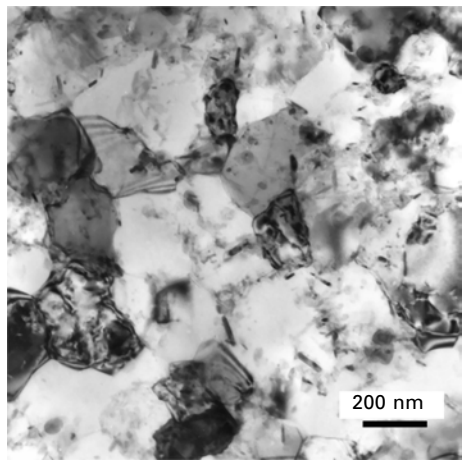
fully sintered, is not necessarily nanocrystalline, but the extra shrinkage that occurs via the solid state at abnormally lower temperatures reduces the number of steps required to produce a fully dense material,[35] thereby saving furnace time and cost.

9.5.3 Structure type 2: Production of nanocrystalline or amorphous powders where the final material is nanocrystalline and structurally homogeneous

This is a more difficult application of nanoengineering using mechanically alloyed powders. Because of the high diffusivity conditions of the powders, preventing or controlling grain growth can be an extremely important factor in retaining a nanostructure. Contamination may also be a significant factor in not achieving a homogeneous microstructure. The aim therefore when producing these kinds of materials is to retain the nanoscale grain sizes, by selective processing and structure stabilisation to impede grain growth. Frequently this is achieved by including inert particles in the milling process that act as pinning sites for grain boundaries. Some metallic materials, such as titanium aluminides, as well as some aluminium and iron alloys, have been successfully produced in this manner and are reported to display grain sizes of the order of 40–600 nm, depending on the alloy composition, final temperature and duration of the consolidation step.[36, 37] One significant advantage of production of such fine grained metallic and intermetallic materials, is that the inherent problems of poor formability and poor ductility may be largely overcome. Some mechanically alloyed materials may become superplastic at comparatively low temperatures or at much higher strain rates than those of micro-sized grain materials.

9.5.4 Structure type 3: Production of nanocrystalline or amorphous powders where the final material is structurally heterogeneous, with nanocrystalline or amorphous regions

This type of microstructure is the most commonly produced in mechanically alloyed materials, and encompasses many cases where oxides, carbides, or other hard particles are introduced into the material. These materials use these particles, for example, to inhibit dislocation motion in the same way as for precipitation hardened materials. Their major advantage, however, is that the oxides or carbides rarely coarsen, meaning that service temperatures may be significantly increased, without softening occurring. An example of an aluminium alloy produced for this purpose is shown in Fig. 9.14 and has the composition Al-4Mg-0.8O-1.1C. Apart from having a grain size of less than



9.14 Transmission electron migrograph showing grain structure of IN9052 aluminium alloy (courtesy J. Weber).

500 nm, the alloy contains a dispersion of oxide and carbide hard particles as finer nanoscale particles. It also has the advantage that magnesium causes solid solution strengthening of the aluminium matrix. This alloy typically displays a tensile strength of 550 MPa and 8% elongation at room temperature. Similar to this is the alloy IN9021, (Al-4Cu-1.5Mg-0.8C-1.1C) where, rather than solution strengthening, the matrix is strengthened by densely dispersed precipitates of Al_2CuMg that are often rods less than 40 nm in length with a diameter of <5 nm. Alloy IN9021 displays a typical tensile strength of 600 MPa, 11% elongation at room temperature, and shows high creep resistance. This alloy is perhaps one of the more highly nanoengineeered materials in use today in that it incorporates several aspects of nanoengineering to produce its property advantages.

9.6 Amorphous solids and controlled crystallisation through rapid solidification

As indicated in the earlier sections, manipulation of engineering materials at the nanometre scale is most readily achieved by starting with a material that is very far from its equilibrium state. Processing routes for the production of 'extreme non-equilibrium' materials range from the solid state quench producing, for example, the supersaturated solid solution discussed in the previous section, to mechanical milling, rapid solidification and even condensation from the vapour phase. Each technique allows the development of very extended solid solutions, whereas the latter three also may induce the formation of new crystalline phases or amorphous solids.[38] Of these processes rapid solidification processing (RSP), in particular, has provided the ability

to produce significant quantities of 'tailored nanostructure materials' for engineering applications.

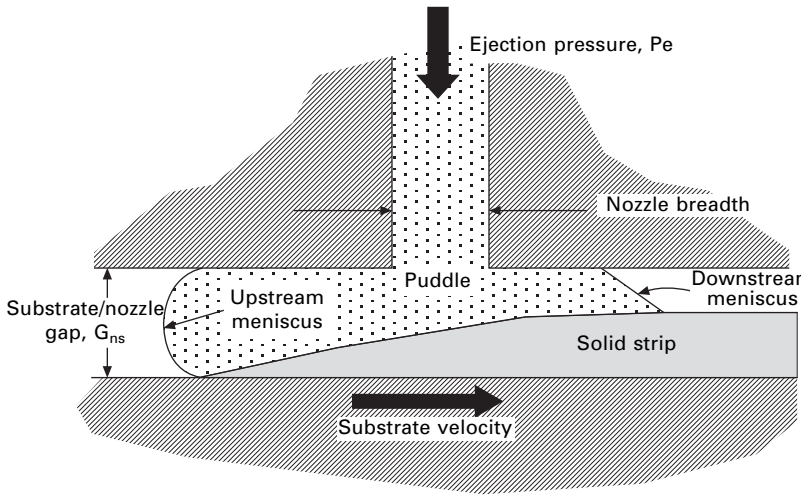
In their review of non-equilibrium processing of materials Gibson and Morton[39] discussed the various techniques for RSP of metallic materials and pointed out the developing thrust towards net shape processing capability, i.e., processes that result in semi- or even fully finished components. More recently Inoue,[40–42] Warlimont,[43] Lu[44] and Dubois [45] have provided excellent insights into the production and applications of amorphous and nanocrystalline alloys and the remarkable progress in the development of bulk amorphous alloys.

9.6.1 RSP techniques

The RSP techniques that have been used to produce engineering quantities of nanocrystalline or amorphous alloys are planar flow casting or PFC,[46] melt extraction, melt drag, melt overflow,[47] spray casting[48] and the compaction of gas atomised powders.[49] Except for gas atomisation, these processes all rely on contacting a low dimensional melt with a larger cold quench surface. In the case of PFC the molten metal is ejected through a rectangular slit nozzle onto a rapidly spinning chilled roll surface. Typically the metallic ribbon produced has a thickness of 40–100 microns and is produced at speeds up to 50 metres per second. This process can produce nanocrystalline or amorphous ribbon of very high dimensional stability and is the basis for the very significant quantities of Metglas series alloys produced originally by Allied Signal, now by Metglas Solutions, part of Hitachi Metals America Ltd.[50]

The concept of PFC is illustrated in Fig. 9.15. The figure shows the very close proximity of the slit-shaped nozzle to the chill wheel surface that is critical to the development of the stable flow and cooling conditions that are the key to the production of high-quality cast ribbon, whether crystalline or amorphous. Maintenance of stable flow and cooling conditions across a very wide ribbon requires innovative slit designs and sophisticated temperature sensing and flow control but production of 300 mm wide strip is relatively routine. The other melt chill processes are also characterised by contacting the molten metal with a rotating chill surface, a textured wheel in the case of melt extraction. However, the contact is simply achieved by dipping into the melt surface or by controlled overflow from the melt container. The melt extraction process has been utilised successfully to produce industrial quantities of stainless fibres for concrete reinforcement.[51]

The gas atomisation and thermal spray techniques also offer the possibility for the very high solidification rates necessary for far from equilibrium processing. In the latter process fine liquid (or perhaps semi-solid) droplets are ejected at high velocity onto a chilled substrate where they form a splat



*9.15 Sch*ematic representation of the planar flow casting process showing the nozzle/substrate configuration (after reference 39).

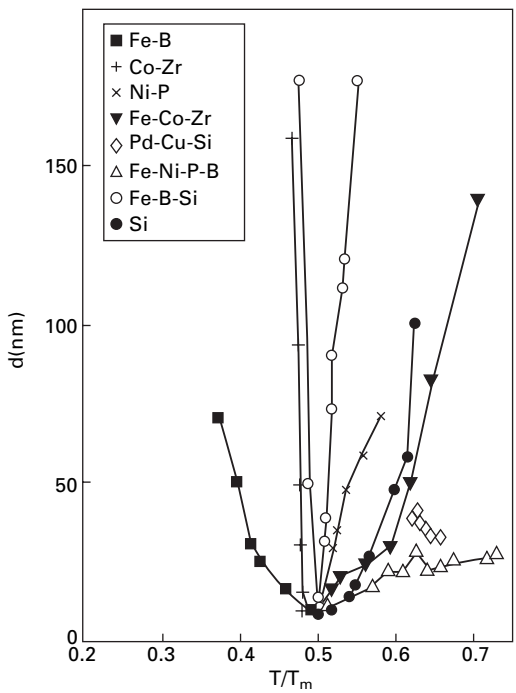
and solidify or complete their solidification very rapidly. This process is capable of producing strip and sheet but with poorer dimensional control than via PFC. However, provided that the cooling rates necessary for the achievement of nanocrystalline or amorphous structures is not extreme, the spray casting routes offered the possibility of more general net shape nanocrystalline and amorphous component production. In practice, the continuing increase in thermal barrier with component growth has prevented this process from becoming an effective net shape forming process for nanocrystalline and amorphous alloys. The recent development of bulk amorphous alloy systems[52] offers new opportunities for net shape forming processes.

Gas atomised powders will usually exhibit a range of microstructures with particle size but it is possible to produce large quantities of powder with significant amorphous phase content. Depending on the glass transition temperature, these powders may be compacted by extrusion or other means to bulk form as indicated in Fig. 9.13.[53]

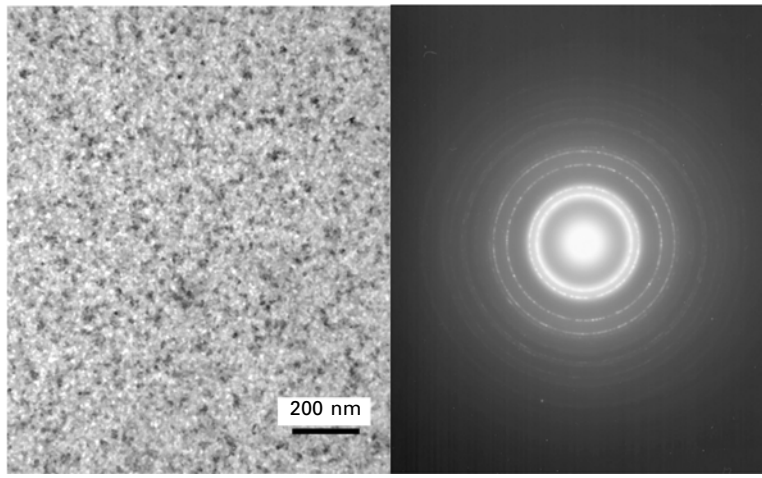
In the production of nanoscale structural materials via the rapid solidification techniques two possible process routes are available: (i) direct quenching to the desired nanocrystalline structure, and (ii) 'overquenching' to develop a fully amorphous structure and development of the final nanocrystalline structure by controlled nucleation and growth from the amorphous state. Of these two routes direct quenching has proved less flexible and more difficult to implement. Almost invariably the overquenching route with controlled crystallisation from the amorphous state is adopted to develop the desired nanocrystalline state. This has been aided in the past decade by advances in 'alloy design' allowing the preparation of the so-called bulk amorphous alloys in sheet millimetres in thickness rather than 50–100 microns.[52]

A successful instance of the direct production of a nanocrystalline material was for the alloy $Hf_{11}Ni_{89}$.[53] This alloy is one of many alloys with good glass forming ability based on a composition close to the eutectic composition for a late transition metal (Fe, Co, Ni) with an early transition metal (Zr, Hf). Nanocrystalline material produced by PFC exhibited a grain size less than 10 nm and this grain structure was very stable with the grain size after annealing at 740 K reaching only 12 nm. Studies of this material have provided useful insight into the nature of the grain boundaries in as prepared and annealed nanocrystalline alloys.[54]

The significant advantage of the overquench and crystallise route is the greater ability this route provides for control of the final microstructure at the nanoscale. For example, the proportion and grain size of the nanoconstituents can be varied over the range from a few nanometres to microns or even larger. The final grain size is dependent upon the balance between the nucleation rate and the grain growth rate. As shown in Fig. 9.16, for many highly stable glass forming alloys, the minimum grain size is produced by crystallising at or near to 0.5 Tm where Tm is the melting point of the alloy.[55] A typical nanocrystalline developed by this route is shown in Fig. 9.17 . The original amorphous AlNiY alloy was annealed at 200 °C for one hour, resulting in an



9.16 Variation of nanocrystalline grain size with crystallisation temperature from various amorphous materials (after Lu[43]).



9.17 Bright field electron micrograph and diffraction pattern showing 15 nm aluminium crystals produced after annealing an amorphous Al86Ni10Y4 alloy at 300 °C for 20 mins (M.A. Gibson – private communication).

nanocrystal grain size of 45 nm.[56] Partial crystallisation can be used to develop a duplex amorphous/nanocrystal structure and this may have significant property benefits, as is the case for some of the soft magnetic alloys. Another major benefit of the overquench route is the flexibility it provides with the possible development of new non-equilibrium phases during crystallisation and/or control over the growth habit of the nanocrystalline phases.[55]

9.6.2 Applications of amorphous alloys

This control of microstructure at the nanometre scale applies even to the amorphous alloys or metallic glasses where, for example, the lack of grain boundaries and the uniformity of composition at the nano as well as the large scale can provide excellent corrosion resistance. An application for amorphous alloys that perhaps does not rely on the control of microstructure at the nanoscale is the development of brazing foils.[57] These materials have found ready acceptance in critical applications areas such as in aero-engines where they provide fluxless brazing capability with close joint fit up allowing excellent tolerances to be maintained. However, a key property of these foils that does rely on microstructure is their ductility. Because of their ductility the foils can be cut to close tolerances and maintain their integrity during fit up and heating to the brazing temperature.[50]

As early as 1988 amorphous Al-based alloys had been produced in wire and strip form for structural applications. They exhibited high strength but

limited ductility[57] and have not reached engineering application but again, with the developments in bulk amorphous alloys, the possibility of significant structural applications for amorphous alloys cannot be ruled out. Amorphous alloys have also found application as catalysts. Many conventional catalysts are composed of nanometre scale metallic particles distributed throughout a fine porous ceramic support structure, the fine metal particles providing the active sites for the chemical activity. For amorphous catalysts the reason for their chemical activity is not completely understood but there is speculation that it is related to the presence of a uniform and very high density of low co-ordination sites at the surface of the amorphous structure.[58]

9.6.3 Applications of nanocrystalline alloys

While nanostructured materials might be relatively new they have already found engineering applications requiring production in significant tonnage quantities. More than 30,000 tonnes per year of the soft magnetic nanocrystalline alloys are produced by crystallisation of RSP amorphous ribbon.[42] Nanocrystalline WC-Co composites were also an early success, with their superior hardness, toughness and wear resistance providing high-performance machining tools.[59] Further, the very fine structure of the nanocrystalline tools offers the promise of far superior fine drilling and cutting performance in the micro-machining area. In structural applications the increasing competitiveness of carbon fibre composites has stimulated the demand for high-performance light alloys for aerospace applications. High-strength aluminium alloys based on the Al-V-Fe and Al-Fe-Mo systems have been produced from PFC ribbon by milling and compaction by extrusion. These alloys have good combinations of strength and thermal stability but to date have seen only limited application.[60]

Nanocomposite aluminium and magnesium alloys containing a high volume fraction of nanoscale precipitates also show very high strengths with moderate ductility.[61–63] However, these materials have been produced in only relatively small quantities and have not reached commercial application. Nanocrystalline materials developed from bulk amorphous alloys show high strength with moderate toughness but further development is required to achieve the ductility and toughness levels required for critical engineering applications.[64] As for the amorphous alloys, the dominant usage to date and probably for the foreseeable future for nanocrystalline alloys has been for soft magnetic applications. As stated above the very desirable combination of magnetic and electrical properties of these materials has seen them adopted widely for high performance applications as indicated in Table 9.4.

A specific example of the control of microstructure at the nanoscale is represented by the FINEMET series of alloys based around Fe74.5Nb3Si13.5B9 with minor additions of Cu replacing some of the Fe to attain the best

Table 9.4 Applications of amorphous and nanocrystalline metallic materials (after Warlimont (42))

Properties applied	Field of application	Specific application
Soft magnetic	• Higher frequency and pulse power electronics	Cores of transformers, chokes and magnetic amplifiers for SMPS up to the kilowatt range Cores of common-mode chokes in RFI suppression filters Cores in ground fault interrupters and current transformers Gate triggering transformers for power semiconductors
	• Mains frequency transformers	50/60 distribution transformers
	• Communication and sensor technology	Signal transformers and chokes in ISDN and other telecommunication networks Electronic article surveillance: security labels Pulse sensors GMR and magneto-impedance sensors Giant magnetostrictive actuators
Hard magnetic	• Magnetic shielding	Flexible shielding
	• Permanent magnets	$Fe_{14}Nd_2B$ flakes for bonded and sintered magnets
Mechanical		Metal joining alloys Alloys for cladding and surface alloying Spray deposited amorphous wear resistant coatings Flexible filler materials for cladding and surface alloying Reinforcement fibres for ceramic building materials Golf clubs
Chemical		Electrodes Electrodeposited coatings Heterogeneous catalysts

properties.[65] These alloys are prepared as amorphous ribbon by PFC and after crystallisation at temperatures between 500 and 600 °C small Fe-Si grains (10–15 nm) and Cu-rich clusters develop within the remaining amorphous matrix as indicated in Fig. 9.16. This nanostructure produces very low coercivity – 5–10 A/cm, high permeability – 100,000, essentially zero magnetostriction, and low core loss due to the high electrical resistivity; all ideal characteristics for soft magnetic cores in transformers and other electrical components. More recently a new series of alloys with higher Fe content has been developed, allowing even higher saturation but retaining the other desirable characteristics of FINEMET.[66]

Nanoscale materials with good hard magnetic properties have to date

mainly been produced by mechanical milling routes to eliminate the deleterious effects of texture on the attainable energy product. However, a recent approach has been to use the exchange coupling between hard and soft magnetic components at the nanoscale. Crystallisation of amorphous Fe-rich amorphous alloys such as $Fe_{90}Nd_7B_3$ result in a nanoscale two-phase mixture of $Fe_{14}Nd_2B$ and ∝-Fe that delivers a remanent magnetisation significantly greater than that for an isotropic solid, 0.5 Ms. This combined with the inherent cleanliness of the PFC material compared to mechanically milled material shows real promise for this route to the production of high-performance hard magnetic alloys.[67]

9.7 Future trends

The continuing thrust for the development of nanoengineered metals and alloys has come from the remarkable properties that they can attain by control of structure at the nanoscale. In the realm of nanoengineered metals and alloys, there are now a multitude of different alloys and materials with enhanced properties, and the greatest challenge is to make what already exists as laboratory scale experimental materials cost effective in mass production. Many advanced production routes for nanoengineered metals and alloys are available, but they are still considered to be too expensive for industrial applications. Nevertheless, nanoengineered alloys such as precipitation hardened aluminium alloys have already become multi-billion dollar per annum industries. The major challenges facing these established technologies is in finding new applications where the performance of these materials has evolved beyond that of those currently used, or where new emphasis is placed on one or more of the advantages of these alloys, such as improved strength:weight ratios, and where cost may be decreased. Increasing competitiveness against other metallic and non-metallic materials, such as polymers, ceramics and hybrid materials that have also been structurally engineered for service performance will also be a major focus.

9.8 References

1. R.W. Hertzberg, *Deformation and Fracture Mechanics of Engineering Materials*', 4th edn, New York, John Wiley and Sons Inc., 1996.
2. E.O. Hall, 'The deformation and ageing of mild steel: III discussion of results', *Proc. Phys. Soc. B*, 1951, **64**, 747–753.
3. N.J. Petch, 'The cleavage strength of polycrystals', *JISI*, 1953, **173**, 25–28.
4. R.W. Siegel and G.E. Fougere, 'Mechanical properties of nanophase metals', *NanoStruct. Mater*., 1995, **6**, 1–4, 205–216.
5. H. Conrad and J. Narayan, 'On the grain size softening in nanocrystalline materials', *Scripta Mat*, 2000, **42**, 11, 1025–1030.
6. H. Conrad and J. Narayan, 'Mechanism for grain size softening in nanocrystalline Zn', *Appl. Phys. Lett*., 2002, **81**, #12, 2241–2243.

7. C.E. Pearson, 'The viscous properties of extruded eutectic alloys of lead-tin and bismuth-tin', *J. Inst. Met*. 1934, **54**, #1, 111–124.
8. O.D. Sherby, 'Superplasticity', *Sci. J.*, **5**, #6, 75–80, 1969.
9. A. Wilm, 'Physical metallurgical experiments on aluminium alloys containing magnesium', *Metallurgie*, 1911, Vol. 8, 223.
10. P.D. Merica, R.G. Waltenberg and H. Scott, 'Heat treatment of Duralumin', *Bull. AIME*, 1919, 913., *Chem. & Met. Eng*., **21**, 551, *et seq*.
11. Z. Jeffries and R.S. Archer, 'The slip interference theory of the hardening of metals', *Chem. & Met. Eng*., 1921, **24**, 1057.
12. A. Guinier, 'Structure of age-hardened aluminium-copper alloys', *Nature* 1938, Vol. 142, 569–570.
13. G.D. Preston, 'The diffraction of X-rays by age hardening aluminium copper alloys', *Proc. Roy. Soc*., 1938, A167, 526.
14. D.G. Altenpohl, *Aluminium: technology, applications, and environment*, 6th edn, Washington D.C., The Aluminum Association, 150, 1998.
15. A. Dupasquier and A.P. Mills Jr, eds, *Positron Spectroscopy of Solids*, Amsterdam, IOS, 1995.
16. *International Alloy Designations and Chemical Composition Limits for Wrought Aluminium and Wrought Aluminium Alloys*, Washingtom D.C., The Aluminum Association, 2–9, 1998.
17. I.J. Polmear, 'Control of precipitation processes and properties in aged aluminium alloys by microalloying', *Mat. Forum*, 1999, **23**, 117–135.
18. J.R. Pickens, F.H. Heubaum, T.J. Langan and L.S. Kramer, 'Al-(4.5-6.3)Cu-1.3Li-0.4Ag-0.4Mg-0.14Zr Alloy Weldalite 049', *Aluminium Lithium Alloys*, Williamsburg, Materials and Component Engineering Publications 1397–1414, 1989.
19. E.A. Brandes and G.B. Brook, eds, *Smithells Metals Reference Book*, 7th edn, Oxford, Butterworth Heinmann, 22–3, 1998.
20. G.A. Edwards, K. Stiller and G. Dunlop, 'APIFM investigation of fine scale precipitation in aluminium alloy 6061', *Appl. Surf. Sci*, 1994, **76**, 1–4, 219–225.
21. S.K. Maloney, K. Hono, I.J. Polmear and S.P. Ringer, 'The chemistry of precipitates in an aged Al-2.1Zn-1.7Mg at% alloy', *Scripta Mat*., 1999, **41**, 1031–1038.
22. I.J. Polmear, *Light Alloys: From Traditional Alloys to Nanocrystals*, 4th edn, Butterworth Heinmann, 2006.
23. *Aluminum Standards and Data, 1998 Metric SI*, Washington D.C., The Aluminum Association, pp. 3–12 to 3–17, 1998.
24. R.N. Lumley, I.J. Polmear and A.J. Morton, 'Interrupted ageing and secondary precipitation in aluminium alloys', *Mat. Sci. Tech*., 2003, **19**, 11 1483–1490.
25. K.J. Irvine and F.B. Pickering, 'Low carbon steels with ferrite-pearlite structures', *J. Iron & Steel Inst*., 1963, **201**, 944–959.
26. W.C. Leslie, *The Physical Metallurgy of Steels*, Tokyo, McGraw-Hill, 1981.
27. R.W.K. Honeycombe and H.K.D.H. Bhadeshia, *Steels: Microstructure and Properties*, 2nd edn, London, Edward Arnold, 1995.
28. J.M. Gray and R.G.B. Yeo, 'Columbium carbonitride precipitation in low alloy steels with particular emphasis on precipitate row formation', *Trans Amer. Soc. Met*., 1968, **61**, pp. 255–269.
29. A.T. Davenport, F.G. Berry and R.W.K. Honeycombe, 'Interphase precipitation in iron alloys', *Met. Sci. J*., 1967, **2**, 104–106.
30. T. Gladman, I.D.M. McIvor and D. Dulieu, 'Structure–Property Relationships in Micro-Alloyed Steels', Union Carbide Corporation, Washington, *Micro-alloying*, 1975, 75.

31. G.B. Schaffer and J.S. Forrester, 'The influence of collision energy and strain accumulation on the kinetics of mechanical alloying', *J. Mat. Sci*., 1997, **32**, pp. 3157–3162.
32. *Metals Handbook*, 9th edn, Vol. 7, *Powder Metallurgy*, Ohio, ASM, 527–528, 1988.
33. C. Capdevila and H.K.D.H. Bhadeshia, 'Manufacturing and microstructural evolution of mechanically alloyed oxide dispersion strengthened superalloys', *Adv. Eng. Mat*., 2001, **3**, #9, pp. 647–656.
34. O. Dominguez and J. Bigot, 'Material transport mechanisms and activation energy in nanometric Fe powder based on sintering experiments', *Nanostr. Mater*., 1995, **6**, 877–880.
35. R.N. Lumley and G.B. Schaffer, 'The application of mechanical alloying to the production of sintered WC-Ni-Cr hardmetals', *P/M Tech. Briefs*, 2001, **3**, #2, pp. 9–12.
36. F.H. Froes, O.N. Senkov and E.G. Baburaj, 'Synthesis of nanocrystalline materials – an overview', *Mat. Sci. Eng*., 2001, **A301**, pp. 44–53.
37. F. Zhang, L. Lu, M.O. Lai and F.H. Froes, 'Grain growth and recrystallisation of nanocrystalline Al3Ti prepared by mechanically alloying', *J. Mat. Sci.*, 2003, **38**, pp. 613–619.
38. C. Suryanarayana and C.C. Koch, 'Nanocrystalline materials – Current research and future directions', *Hyperfine Interactions* 2000, **130,** 5–44.
39. M.A. Gibson and A.J. Morton, 'Non-Equilibrium Processing of Metals to near Net Shape' *Mat. Forum*, 1996, **20**, 93–120.
40. A. Inoue, 'Amorphous, Nanoquasicrystalline and Nanocrystalline Alloys in Al-based Systems', *Prog. Mater. Sci*., 1998, **43**, 365–520.
41. A. Inoue and K. Hashimoto, eds, *Amorphous and nanocrystalline materials. Preparation, properties and application*, Berlin, Springer, 2001.
42. A. Inoue, 'Bulk amorphous and nanocrystalline alloys with high functional properties', *Mat. Sci. Eng*., 2001, **A304–306**, 1–10.
43. H. Warlimont, 'Amorphous metals driving materials and process innovations', *Mat. Sci. Eng*., 2001, **A304–306**, 61–67.
44. K. Lu, 'Synthesis of Nanocrystalline Materials from Amorphous solids', *Adv. Mater*., 1999, **11**, #13, 1127–1128.
45. J-M. Dubois (2000), 'New prospects from potential applications of quasicrystalline materials', *Mat. Sci. Eng.,* 2000, **A294–296**, 4–9.
46. M.C. Narasimhan, *Continuous casting method for metallic strips* US Patent No. 4, 142, 571, 1979.
47. R.E. Maringer, 'Solidification on a substrate', *Mater. Sci. Eng*., 1988, **98**, #1–2, 13–20.
48. A.R.E. Singer, *Casting of Near Net Shape Products*, Y. Sahai *et al*. eds, Warrendale PA, The Materials Society, 245, 1988.
49. A. Inoue, K. Ohtera, K. Kita and T. Masumoto, 'Aluminum and magnesium based amorphous alloys with high strength'; in *Proc. 1st Japanese International SAPME Conference*, N. Igata *et al*, eds, 1989, Nikkan Kogyo Shimbun, Tokyo, 7–12.
50. A. Rabinkin, 'Overview: Brazing with (NiCoCr)-B-Si Amorphous Brazing filler metals; Alloys, Processing, Joint Structure, Properties and Applications', submitted to *Science and Technology of Welding and Joining*, 2003.
51. T.A. Gaspar and L.E. Hackman, *Casting of Near Net Shape Products*, Y. Sahai *et al*., eds, 1988, Warrendale, PA., The Materials Society, 613.
52. A. Inoue, 'Bulk Amorphous Alloys': in *Amorphous and Nanocrystalline Materials.*

Preparation, properties and application, A. Inoue and K. Hashimoto, eds, Berlin, Springer, pp 1–51, 2001.
53. I. Bakoni and A Cziraki, 'Nanocrystalline-Forming Ability of Alloys by Melt-Quenching', *Nanostruct. Mater*., 1999, **11**, 9–16.
54. S. Mentese, J.B. Suck and S. Janssen, 'Atomic dynamics of rapidly quenched and annealed nanocrystalline $Ni_{89}Hf_{11}$', *Physica B*, 2002, **316–317**, 438–440.
55. K. Lu, 'Nanocrystalline metals crystallized from amorphous solids: nanocrystallization, structure and properties', *Mat. Sci. Eng. R*, 1996, **16**, 161–221.
56. M.A. Gibson – private communication.
57. A. Rabinkin, 'Brazing with amorphous foil preforms', *Adv. Mat. Proc*., 2001, **159 (6)**, 65–68.
58. G.J. Shiflet, Y. He and S.J. Poon (1988), 'Mechanical properties of a new class of metallic glasses based on aluminium', *J. Appl. Phys*., 1988, **64**, (12), 6863–6865.
59. J.-F. Deng, H. Li and W. Wang, 'Progress in the design of amorphous alloy catalysts' *Catal. Today*, 1999, **51**, 113–125.
60. L.E. McCandish, V. Kevorkian, K. Jia and T.E. Fischer, 'Nanostructured WC-Co composite powders', *Advances in powder metallurgy and particulate materials – 1994*, 1994, **5**, 329–337.
61. Y. Barbaux and G. Pons, 'New rapidly solidified aluminium alloys for elevated temperature applications on aerospace structures', *J. de Phys. IV, Coll. C7* (supplement to *J. de Phys. III*), 1993, **3**, 191–196.
62. Y.-H. Kim, A. Inoue and T. Masumoto,'Increase in mechanical strength of Al-Y-Ni amorphous alloys by dispersion of nanoscale FCC aluminium particles' *Mat. Trans – JIM*, 1991, **32,** #4, 331–338.
63. A. Inoue, M. Watanabe, H.M. Kimura, F. Takahashi, A. Nagata and T. Masumoto, 'High mechanical strength of quasicrystalline phase surrounded by FCC aluminium phase in rapidly solidified Al-Mn-Ce alloys', *Mat. Trans.– JIM*, 1992, **33**,#8, 723–729.
64. S.G. Kim, A. Inoue and T. Masumoto, 'Increase of mechanical strength of a $Mg_{85}Zn_{12}Ce_3$ amorphous alloy by dispersion of ultrafine hcp magnesium particles', *Mat. Trans*.– JIM, 1991, **32**, 875–878.
65. A. Inoue and K. Hashimoto, eds, *Amorphous and nanocrystalline materials. Preparation, properties and application*, Berlin, Springer, pp. 45–48, 2001.
66. Y. Yoshizawa, S. Ogums and K. Yamauchi, 'New iron based soft magnetic alloys composed of ultrafine grain structure', *J. Appl. Phys*., 1988, **64**, #10–11, 6044–6046.
67. A, Makino, A. Inoue and T. Masumoto, 'Soft magnetic properties of nanocrystalline Fe-M-B (M=Zr, Hf, Nb) alloys with high magnetisation', *Nanostruct. Mater.*, 1995, **6**, 985–988.
68. P.A.I. Smith, J. Ding, R. Street and P.G. McCormick, 'Mechanically alloyed Sm-(Co-Fe) permanent magnets', *Scripta Mat*., 1996, **34**, #1, 61–66.

10

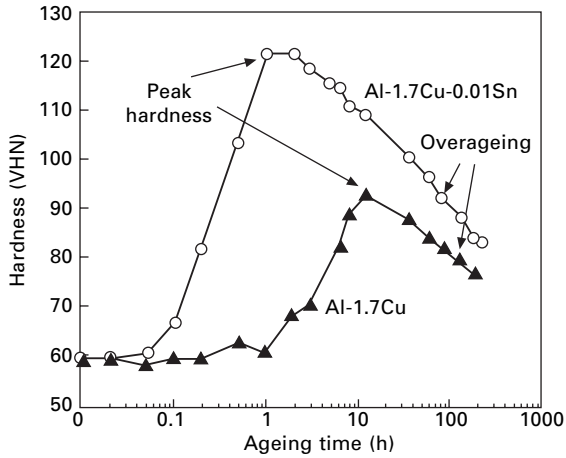
Using magnetic resonance to study nanoprecipitation in light metal alloys

K NAIRN and T BASTOW
CSIRO Manufacturing and Infrastructure Technology, Australia

10.1 Introduction to age hardening of light metal alloys

Both aluminium cooking foil and aerospace aluminium consist of more than 95% Al atoms. However, the differences in strength and durability are very significant. These differences arise as a result of precise alterations in the nanostructure of the metal, through addition of small amounts of other elements (a process known as alloying), through application of mechanical force, and through heat treatment. In Chapter 9 of this volume, Lumley and co-workers present images of various nanostructures found in alloys, showing the variety of small precipitates that are responsible for increasing the strength in the alloy. Recently, high-resolution microscopy has been used to look at precipitates on an atomic level, including imaging the interface between the matrix and the precipitate (Son, 2005; Hutchison, 2001). Three-dimensional atom probe techniques have been used to evaluate the composition of individual precipitates (Maloney, 2001).

The general process of increasing the hardness of aluminium or magnesium through alloying with small amounts of certain other elements, in combination with heat treatment is known as age hardening. Textbooks discussing the general principles of age hardening are available (Polmear, 1995; Porter, 1992). The goal of age hardening is to produce finely divided, high aspect-ratio, coherent particles (or precipitates). Coherent precipitates are those which have an atomic structure continuous with the matrix in which they are located – this enables the precipitates to block dislocation motion within the material, and hence increase strength and hardness by reducing deformation. The dislocation blocking occurs at the interface between the precipitate and the matrix – hence the desirability of small, elongated, precipitates. Fig. 10.1 shows age hardening when alloys of composition Al-1.7at.% Cu and Al-1.7at.% Cu-0.01at.% Sn are heat treated. Both peak hardness and overageing (associated with coarsening of particles and formation of incoherent precipitates)

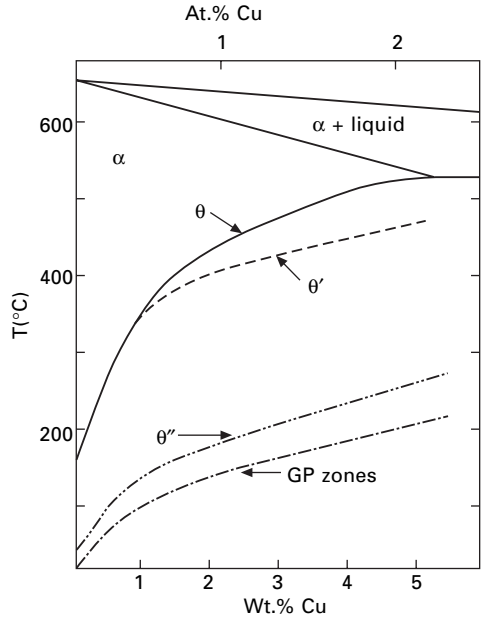


10.1 Hardness as a function of ageing time at 200 °C for Al-1.7 at.% Cu alloys. Adapted from Bastow (2005).

are indicated in the figure. Microalloying with tin accelerates the hardening process for this alloy.

Precipitates evolve through quite complicated processes involving non-equilibrium phases prior to the final equilibrium phase. Control of the non-equilibrium phases is needed to control the final microstructure. In some cases the nature of these phases is controversial (for example, the exact nature and variety of the 'GP zones' that form in the early stages of ageing in Al-Cu alloys is still a matter of debate (Son, 2005; Konno, 2001)); in other cases the phases have not actually been observed – their existence has been inferred from transitions in thermal analysis traces (Noble, 1995), anomalous hardness behaviour and positron behaviour (Dupasquier, 2004).

As a relatively simple example of this behaviour, the aluminium rich end of the equilibrium phase diagram for Al-Cu is shown in Fig. 10.2. In general the formation of a particular equilibrium phase X will take the form X″->X′ >X. However, quenching is frequently used to trap non-equilibrium phases within the structure. The initial stage of precipitate formation is believed to involve the lining up of a few copper atoms on the [100] or [200] planes of the aluminium to form a Guinier–Preston zone, or 'GP zone' (Preston, 1938a, 1938b and 1938c; Guinier, 1938). Aggregation of GP zones leads to an arrangement of planes of copper atoms separated by three planes of aluminium atoms, forming a θ″ precipitate. Upon further treatment, θ″ may transform into θ′ and finally the equilibrium phase, θ. θ precipitates are incoherent, and hence do not provide strengthening – therefore their formation is to be avoided in alloy ageing processes.



10.2 Al-rich end of the Al-Cu phase diagram, showing the non-equilibrium GP zone, θ″ and θ′ dissolution lines, along with the regions of α and θ equilibrium phases. Adapted from Porter (1992).

10.1.1 Methods for studying age-hardening mechanisms

In order to understand and better control age-hardening processes, a detailed understanding of the precipitation processes and their kinetics is needed. Historically, X-ray diffraction (XRD) has been used to determine the structures of many of the precipitate phases, particularly the final equilibrium phases (Silcock, 1953). Unfortunately, XRD is not very effective at detecting small amounts of impurity phases in a large background of matrix phase – a situation which commonly occurs during age hardening. Nor is it suited to detection of very small (nanosized) particles or poorly ordered phases.

Transmission electron microscopy (TEM) is one method suited to evaluation of the structure of small precipitate particles, their size and arrangement within an alloy (again, refer to Chapter 9 of this volume for detailed discussion of TEM in alloys). TEM can identify the various precipitates (such as θ″, θ′ and θ – see Fig. 9.3(b), Chapter 9 for example). The main weakness of TEM lies in evaluating the overall sample composition: only a small area can be analysed at a time. Evaluation of the concentration of alloying elements remaining in the matrix is also difficult. Optical microscopy provides a simpler way of evaluating overall sample composition and microstructure, with the drawback that very small precipitates (those of most importance in these particular alloys) will not be observed.

Differential scanning calorimetry (DSC) is a technique for measuring uptake and release of energy by a sample as it undergoes physical or chemical transitions. Many of the phase transitions within alloy systems are accompanied by detectable exotherms. DSC is generally used in conjunction with other techniques, since it cannot provide spatial, or specific chemical information – in other words, DSC indicates that a transition has occurred, but not the nature of the transition. An example of the use of DSC to study alloy behaviour is given in Noble (1994).

10.2 NMR – a novel method for studying precipitation in alloys

Nuclear magnetic resonance (NMR) provides detailed, quantitative, chemical information on the nature of phases within the entire sample. The main advantage of NMR is that there is no restriction on the size of precipitate that can be detected – conversely, the main disadvantage is that NMR cannot provide any spatial information. Identification of precipitate phases with NMR relies on previously determined XRD data, and on NMR measurements on model compounds.

10.2.1 What is NMR?

NMR spectroscopy is a form of Zeeman spectroscopy made possible by the existence of nuclear spin postulated by W. Pauli in 1924 to explain anomalies in optical spectra. The existence of a nuclear spin entails the existence of a nuclear magnetic moment which allows the nucleus to interact with the magnetic field.

Most of the elements in the Periodic Table have one or more stable nuclear isotopes that possess nuclear spin. For the isotopes of some important elements, such as ^{12}C, ^{16}O and ^{28}Si, the nuclear spin is zero and such isotopes are magnetically, and therefore NMR, impotent. In the simplest magnetic case, e.g., ^{1}H, ^{13}C, ^{19}F, ^{29}Si, ^{31}P, ^{89}Y, $^{107,109}Ag$, $^{119,121}Sn$, ^{195}Pt, the nuclear spin I = 1/2 and this spin can be either up or down with respect to the applied magnetic field. Or in quantum-mechanical terms, when the magnetic field vector **B** is along the z-axis, the z-component $\mathbf{I_z}$ of the nuclear angular momentum **I** can take values m = + 1/2 and m = –1/2. These two states are separated in energy by an amount ΔE proportional to the magnitude of the magnetic field and the gyromagnetic ratio γ of the nucleus:

$$\Delta E = h\nu = \hbar\omega = \hbar\gamma B; \text{ the angular frequency } \omega = 2\pi\nu. \qquad 10.1$$

Transitions between the two levels can be induced by electromagnetic radiation of frequency ν such that $h\nu = \Delta E$.

The largest number of stable magnetic isotopes have spin I > 1/2. For

example ^{7}Li, ^{23}Na, ^{63,64}Cu have I = 3/2; ^{25}Mg, ^{27}Al, ^{47}Ti, ^{67}Zn, ^{91}Zr have I = 5/2, and elements with higher spin nuclei I = 7/2 and 9/2 are relatively common. For such nuclei the magnetic quantum number M can take the half integer values –I, –(I–1),…–1/2, +1/2,…I–1, I. The energy levels corresponding to these m values, $E_m = -\gamma\hbar Bm$, are equally spaced in the absence of other interactions like chemical shift anisotropy and nuclear quadrupole interaction. Such interactions lead, in the case of a single crystal, to unequally spaced transitions with spacing dependent on the orientation of the crystal in the magnetic field and thus a powder lineshape. However, these interactions are not invoked in the following discussion, and can be pursued by the interested reader in specialised texts like Slichter (1996). With the selection rule for magnetic transitions $\Delta m = 1$ this leads to 2I + 1 possible transitions of equal frequency and thus a single resonance line. Even in the presence of the additional interactions, provided they are not too strong, the central transition (–1/2, 1/2) often remains sharp.

For an isolated nucleus the transition would have essentially no frequency width. However, in an aggregate of nuclei, such as in a molecule or solid, there are magnetic dipole interactions between a nuclear spin (i) with neighbours (j) at a distance r_{ij} each contributing an energy of order r_{ij}^{-6} and which lead to what is termed a dipolar linewidth (Slichter, 1996). For many metals (e.g. Na, Al, Cu, V) this leads to a linewidth of around 10 kHz, independent of the magnetic field and which can, if required be readily calculated if the spatial configuration of the surrounding nuclei is known (as in a molecule or crystalline solid).

For a molecule or condensed phase there is also the phenomenon of additional contributions to the main magnetic field from the fields caused by circulating (Amperian) electron currents of the valence electrons of the nucleus of the atom in question and its near neighbours. This additional magnetic field, typically for metals a factor 10^{-3} times the magnitude of the main field, leads to a shift in frequency of the resonance.

This frequency shift is known as the chemical shift (or in metals the Knight shift) and depends crucially on the spatial configuration of the surrounding atoms. It is the sensitivity to the near neighbour atomic surroundings which gives NMR spectroscopy its analytical power. These shifts are proportional to the applied magnetic field and are often given in dimensionless form by dividing the frequency shift in Hz by the resonance frequency of the nucleus in a stated reference compound, and expressed as a shift in parts per million (ppm). This means that the measured frequency shift is then independent of the magnetic field of the spectrometer and that shifts observed on different instruments may be sensibly compared.

For metals, the (dimensionless) Knight shift, which characterises different metals and metallic phases is given by the following formula (see Slichter (1996) for an extensive analysis of its formulation):

$$K = (8\pi/3)\chi_e < |\psi_k(0)|^2 >_{EF} \qquad 10.2$$

where the electronic spin susceptibility $\chi_e = \mu_B^2 N(E_F)$ and $\psi_k(0)$ is the amplitude of the wave function at the nucleus, the square modulus of which is evaluated for electrons with energy equal to the Fermi energy E_F. The quantity $N(E_F)$ is the electronic density of states at the Fermi surface and μ_B is the Bohr magneton. The Knight shift is a consequence of the Fermi contact interaction between the nucleus and the s-component of the electron wavefunction $\psi_k(r)$. This leads to line shifts of one to two orders of magnitude greater than that for the metal in an insulating (dielectric) compound.

The quantity K is unfortunately not at present calculable *a priori* even for a pure metal such as Al or Cu, and has to be determined empirically in each case. For example, relative to Cu in CuCl, the ^{63}Cu Knight shift for Cu in Cu metal is 2400 ppm while its value for Cu in the dilute alloy Al-0.2at.% Cu is 3450 ppm. The empirical procedure for allotting the ^{63}Cu Knight shift of a line in the ^{63}Cu NMR spectrum to a particular precipitate phase in Al-Cu-Mg alloys has been given (Bastow, 2005a). In simpler cases, such as dilute alloys of Sc in Al, the ^{45}Sc NMR lines may be identified by the order in which they are observed (evolve) on progressive thermal ageing of the specimen and reference to the Al-Sc phase diagram.

The NMR spectrometer consists of a magnet (usually a superconducting solenoid) and a high-power radio frequency (rf) transmitter capable of delivering pulses of power up to a kilowatt (kW) and with duration down to 1 microsecond (μs). The pulses of rf power are delivered to a small coil containing the specimen with the axis of the coil (i.e. the axis of the rf magnetic field which induces the transition between the nuclear m-states) perpendicular to the magnet field direction. The same coil is then used to detect the free induction decay or echo signal that follows the pulse or pulse sequence. This pulse or pulse sequence is then regularly repeated and the response signal is coadded until a noise averaged signal with acceptable signal-to-noise ratio is obtained. This smoothed signal is then Fourier transformed to give a spectrum with one or more lines, depending on the complexity of the alloy (material) investigated. More detailed description of the hardware and philosophy behind basic pulse sequences is available (Fukushima, 1981). While NMR of ^{1}H and ^{13}C nuclei is a routine technique for chemical analysis of molecules, many other nuclei also possess magnetic moments, and are hence able to be studied by NMR. Some NMR-accessible nuclei of interest as alloying elements in light metals are: ^{7}Li, ^{45}Sc, ^{27}Al, ^{63}Cu, ^{119}Sn, ^{25}Mg and ^{67}Zn. ^{25}Mg and ^{67}Zn are more difficult to examine than the other nuclei.

A major difference between routine NMR of molecules in solution and NMR of alloy powders is that the lines obtained in the metals may be thousands of times broader than those observed in solution. To irradiate these lines evenly requires high-powered rf amplifiers, and correspondingly robust probes and detectors. To collect the corresponding signal requires a fast digitiser

and often also the use of echo techniques to separate the residual transmitter pulse from the NMR signal. Another challenge is that some of the important nuclei for alloy structure have comparatively low resonance frequencies and also low natural abundances. This means that the signals can be both broad and weak, requiring prolonged spectral averaging to obtain a clear spectrum. Because electromagnetic radiation in the radiofrequency range is highly attenuated at the surface of metals, only the outside 20 microns or so of a metal piece is able to absorb the rf radiation and hence generate an NMR spectrum. NMR studies are therefore carried out on finely powdered or thin foil samples.

Historically, NMR has been used to study metals and alloys since the 1950s (see Cohen (1957) for a review of some of the early work). Initial studies of the alloys were performed on cubic metals, such as Al containing ~1 at.% Cu, where the breakdown of the symmetric environment around the Al atom as a result of alloying was observed as a sharp decrease, or 'wipe out' of a large fraction of the ^{27}Al signal as the concentration of Cu increased. Little further work was then reported in this field until the 1990s, when, with access to much higher fields, it became possible to use nuclei of the dilute alloying elements (e.g. Cu) as probes and to study the formation of precipitate phases cleanly without a large and obscuring background signal from the nucleus of the host lattice (e.g. Al).

This change in strategy was enabled by improvements in NMR instrumentation, in particular the availability of superconducting magnets with higher magnetic fields. High powered rf transmitters also became widely available, enabling broad, weak signals to be completely excited and accurately observed, and hence allowing the study of a wider range of nuclei, at a wider range of concentrations (Fukushima, 1981).

The NMR spectrum of the probe nucleus may be used to establish the nature of the precipitate phases, both equilibrium and non-equilibrium. The precipitate structure in general contains more than one element, in which case it may be possible to obtain more than one characteristic NMR spectrum. NMR can be used to quantify the fraction of each phase present at various stages of the treatment process and hence to give details of the kinetics of precipitation processes. It is also able to monitor the initial stages of nanoprecipitate formation.

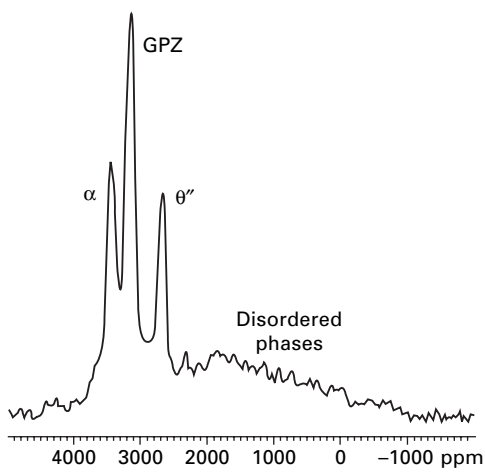
10.3 NMR spectra of alloys

The NMR technique for phase identification in alloys is based on making a one-to-one correspondence between the dominant phase known to exist in the alloy, after a certain heat treatment, from X-ray and electron diffraction studies, and the dominant line in the NMR spectrum for the alloy in that same state. For all such investigations the alloy is put into a standard initial

state by solution treatment, i.e., holding the alloy at a temperature just below the melting point and then quenching rapidly in iced water. This procedure puts all the solute atoms into random substitutional positions in the host matrix. The precipitate phase content of the alloy is then probed after specific ageing intervals at which structural studies have revealed a dominant phase presence.

For example, in the case of dilute Al-Cu alloys containing 1–2 at.% Cu, the ^{63}Cu line that appears immediately after solution treatment and quenching is identified as Cu in random substitutional sites in α–phase aluminium. Next, XRD investigations (Silcock, 1953) have established that ten days ageing of, for example, (Al, 1.7 at.%Cu) at 130 °C, 190 °C or 370 °C will produce a dominant precipitate phase of θ″, θ′ or θ respectively, and this knowledge is sufficient to identify the dominant peak appearing in the ^{63}Cu NMR spectrum for similarly treated alloys.

Figure 10.3 shows an NMR spectrum of copper in an aluminium-copper alloy (Al, 1.7 at.% Cu; heated to just below its melting point, quenched, and then aged at 130 °C for 95 hours). Each peak corresponds to a different environment of the nucleus being measured. For this particular alloy, the copper spectrum corresponds to copper atoms in three significant phases (α, which is a solid solution of copper atoms in aluminium; GP zones, which are mostly copper atoms collected together to form 40–60 Å diameter disks on (100) lattice planes; and θ″ platelets, which are larger than GP zones, and consist of stacked layers – with the disks of copper atoms separated by three planes of aluminium atoms). The area under each peak corresponds to the number of copper atoms within that particular phase. Important information



10.3 ^{63}Cu NMR spectrum of Al-1.7 at.% Cu, heated to just below its melting point, quenched, and then aged at 130 °C for 95 hours (Bastow, 2005).

about the growth and transformation of the phases within the alloy can hence be obtained from comparing the peaks. In particular, the concentration of copper remaining within the matrix is an important piece of information that is not readily obtained by other techniques. The broad underlying peak, and overlap of the other peaks lead to some uncertainty in determining the exact concentrations of phases. Details of how the peaks are assigned have been reported (Bastow, 2003).

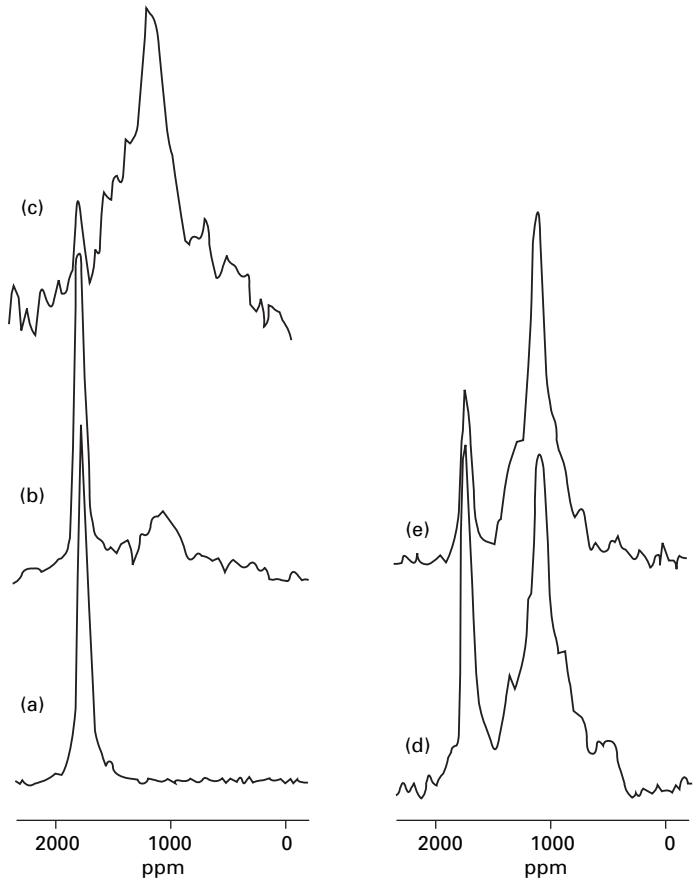
10.3.1 Effect of ageing temperature

Since ageing involves thermally activated processes, both the rate of precipitate phase development and the nature of the precipitates that form depend strongly upon the temperature regime used. An example of this is shown in Fig. 10.4, which compares the ^{45}Sc spectra of Al-0.06 at.% Sc specimens that have been solution treated, quenched and then aged at either 300 or 400 °C (Celotto, 2000). The peak at about 1800 ppm represents Sc in solid solution in the Al matrix, whilst the peak around 1000 ppm corresponds to Sc in Al_3Sc precipitates. Precipitation is much more rapid at 400 °C, and the precipitates that form also appear to be more ordered (i.e. exhibit a narrower, less noisy, NMR peak) at the higher temperature. These finely divided Al_3Sc precipitates are a key to the usefulness of this alloy (Davydov, 2000). They act as nucleation sites for Al crystallisation, and hence lead to a final Al alloy with fine grains even at the high cooling rates characteristic of welding (Hyde, 2001). Fine grains generally lead to improved mechanical properties in alloys (see Chapter 9 of this volume for further discussion of this effect).

10.3.2 Effect of microalloying

Addition of small amounts of other elements can have a dramatic effect on the kinetics of age-hardening processes. As an example of this, Fig. 10.5 compares the spectra of Al-1.7 at.% Cu with and without the addition of 0.01at.% Sn, aged at 200 °C. The tin addition initially stabilises the α phase (a GPZ peak is observed for the 'as quenched' sample without tin, but not in the spectrum for the tin-containing sample), and then speeds up phase transitions. For example, the amount of θ' phase formed at 200 °C after one hour in the alloy containing tin is greater than the amount present after three hours in the tin-free alloy. After 101 hours, both alloys have very similar phase contents, as demonstrated by comparing the two spectra of Fig. 10.5(d).

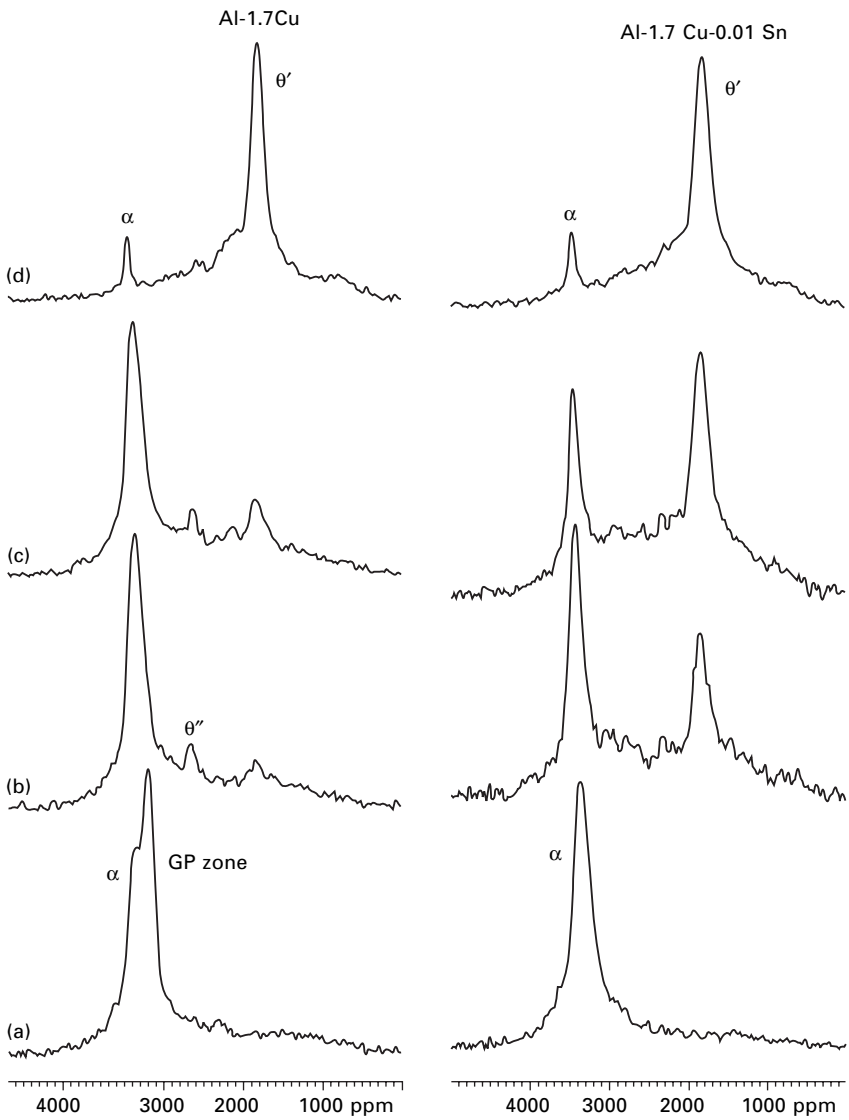
Not all copper spectra are as distinct as those observed in Al-1.7 at.% Cu. Figure 10.6 shows the ^{63}Cu spectrum of Al-1.05 at.% Cu-1.7 at.% Mg, solution treated, quenched and then aged at 150 °C for 371 hours. This alloy age hardens via a different path to Al-1.7 at.% Cu, forming Guinier Preston Bagaryatsky Zones (GPBZ) and then S phase. S phase has the composition Al_2CuMg, while the atomic structure of the GPBZ is not certain.



10.4 ^{45}Sc spectra of Al-0.06 at.% Sc: (a) is solution treated and quenched; (b) and (c) are aged at 300 °C for 2 and 24 hours respectively; (d) and (e) are aged at 400 °C for 2 and 24 hours respectively. Adapted from Celotto (2000).

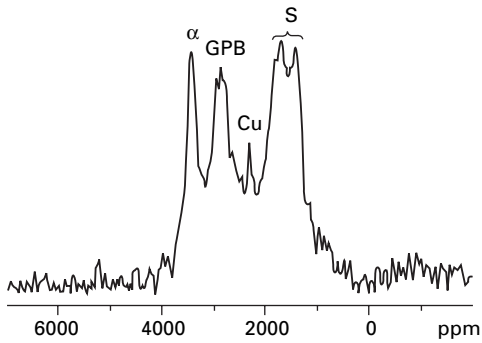
An interesting feature of the Al-Cu-Mg alloys in the $\alpha + \theta$ phase field is the effect of microalloying with silver (Polmear, 1995) which produces a different precipitate termed Ω phase. This phase has the metallurgical advantage of resistance to coarsening when operating at high temperatures. It has been proposed to be a distorted version of θ phase (Al_2Cu) oriented along (111)) planes in the host Al lattice (see, for example, Hutchison, 2001). In the ^{63}Cu NMR spectrum an additional signal to the usual θ' is observed indicative of copper in non-axially symmetric sites (Bastow, 2005a). This signal was attributed to Ω phase and is not observed in the spectrum of the same composition but without Ag.

The strength and light weight of Al-Li based alloys makes them desirable



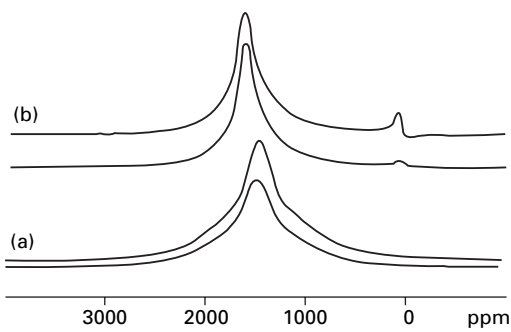
10.5 Comparison of ^{63}Cu NMR spectra from Al-1.7 at.% Cu and Al-1.7 at.% Cu-0.01 at.% Sn. Both alloys were heated to just below their melting points, quenched, and then aged at 200 °C. (a) is measured immediately after quenching; (b), (c) and (d) have been aged for 1, 3, and 101 hours respectively. Adapted from Bastow (2005).

for aerospace applications (Rioja, 1998; Williams, 2003). However, the lithium atoms may migrate to the surface of the alloy during some processing stages and rapidly oxidise, leading to poor surface properties for the alloy. NMR has been used to monitor this surface depletion (Ba, 2002). The specific



10.6 ^{63}Cu spectrum obtained during ageing of Al-1.05 at.% Cu-1.7 at.% Mg at 150 °C. Adapted from Bastow (2005a).

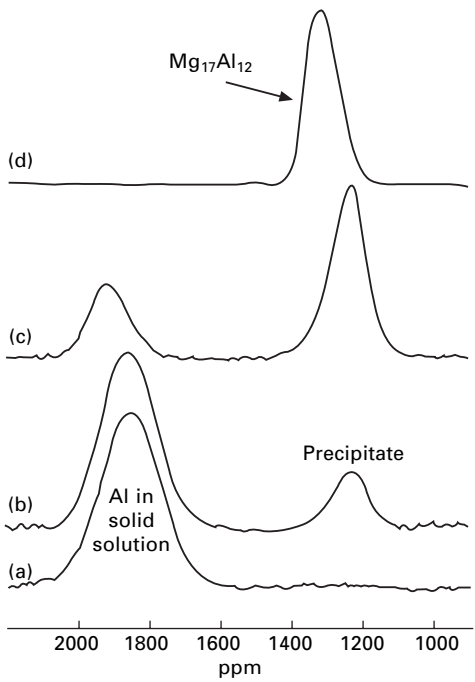
technique involved comparing the intensity of the aluminium signal when the measurement was performed in two different ways: (i) a conventional echo method and (ii) a double resonance method (SEDOR) which allows transfer of magnetisation from the aluminium to any nearby lithium nuclei. In the as-received alloy, the presence of lithium nuclei near the aluminiums resulted in transfer of magnetisation from the aluminium to the lithiums during SEDOR, and hence a substantial difference in intensity between the conventional and SEDOR spectra (pair (a) of Fig. 10.7). As the samples were heat treated, lithium was lost from the sample. This meant that there were far fewer lithium atoms near the aluminiums, and hence very little difference in signal intensity between the conventional and SEDOR spectra (pair (b) of Fig. 10.7).



10.7 ^{27}Al spectra of a sample of Al-Li 8090, a commercial alloy. Pair (a) is of as-received alloy, whilst pair (b) is of alloy that has been heated 600 °C in air for 20 minutes. The lower spectrum in each pair has been measured using the 'SEDOR' technique, which promotes transfer of magnetisation from the Al to nearby Li nuclei. The resulting decrease in signal depends upon the number of nearby lithium nuclei, and can hence be used to monitor the loss of lithium from the alloy at high temperature. Adapted from Ba (2002).

Precipitation hardening is also crucial to the performance of magnesium alloys. The usual alloying elements to increase strength in magnesium are Al, Zn and Mn. It is generally preferred to study the Al nucleus, which gives a strong signal. Examples of NMR of Al in Mg alloys include Celotto (2001), Song (2004) and Bowles (2000). A ^{25}Mg and a ^{27}Al spectrum of γ phase ($Mg_{17}Al_{12}$), which is the dominant precipitate in MgAl alloys, have been reported (Bastow, 1995). AZ91 is a Mg alloy which finds wide use, in areas as diverse as automotive components and casings for consumer electronics. Its approximate composition is 9 wt% Al, 0.5–1 wt% Zn and 0.3 wt% Mn. ^{27}Al spectra of this alloy after various thermal treatments are shown in Fig. 10.8, along with the reference spectrum of pure $Mg_{17}Al_{12}$ (Celotto, 2001).

As heat treatment proceeds, the peak corresponding to Al in solid solution decreases in intensity, and also shifts to higher ppm values. Both the area under the peak and the peak position can be used to determine the amount of Al remaining in solid solution within the alloy. The position of the NMR peak corresponding to the precipitate is not strongly dependent upon thermal treatment. The presence of Zn within the precipitates in the alloy moves the peak to lower frequencies.



10.8 ^{27}Al spectra of AZ91, a commercial Mg-Al-Zn alloy: (a) solution treated and quenched; (b) aged at 200 °C for two hours; (c) aged at 200 °C for 144 hours and (d) pure $Mg_{17}Al_{12}$. Adapted from Celotto (2001).

By combining NMR and corrosion rate measurements, Song and co-workers found that precipitate formation is a very important factor in the corrosion resistance of the industrial Mg alloy AZ91D (Song, 2004). Ageing of AZ91D (a version of AZ91 intended for die-casting applications) produced an initial increase in corrosion resistance, followed by a gradual increase in corrosion rate. This initial behaviour was attributed to the formation of a barrier layer of the precipitate at the grain boundaries, whilst the later increase in corrosion rate was attributed to the gradual decrease in aluminium content of the matrix as ageing proceeded, leading to the matrix being more reactive. NMR was used to identify both the amount of precipitate formed and the aluminium content in the matrix. An earlier study (Bowles, 2000) used NMR to evaluate precipitate formation in magnesium alloys, and correlate it with tensile properties and microstructural data.

10.4 Conclusions

Nuclear magnetic resonance is a technique which identifies different environments around atomic nuclei. Applied to light metal alloys, NMR enables the identification of precipitates and characterisation of the kinetics of their formation. The precipitates within light metal alloys control important properties such as strength, hardness and corrosion resistance. NMR is able to reveal the initial stages of precipitation, where atoms cluster together to form nanoprecipitates. In addition to this, NMR yields information on the concentration of alloying elements within the matrix solid solution. This compositional information, when combined with spatial information from electron microscopy, gives a much move detailed understanding of the alloy, and how microstructure and nanostructure may be modified to improve alloy performance.

10.5 References

Y. Ba and J.A. Ripmeester, (2002) *Magn. Reson. Chem*. 40, 81.
T.J. Bastow and M.E. Smith, (1995) *J. Phys.: Condens. Matter* 7, 4929.
T.J. Bastow and S. Celotto, (2002) *Acta mater*. 51, 4621.
T.J. Bastow, L. Bourgeois and M. Forsyth. (2005) *Phil. Mag.* 85, 2757.
T.J. Bastow, (2005a) *Phil. Mag*. 85, 1053.
A.L. Bowles, T.J. Bastow, C.J. Davidson, J.R. Griffiths and P.D.D. Rodrigo, (2000) p. 295 in *Magnesium Technology 2000* (eds H.I. Kaplan, J.H. Hryn and B.B. Clow), The Minerals, Metals and Materials Society, 2000.
S. Celotto and T.J. Bastow, (2000) *Phil Mag. A.* 80, 1111.
S. Celotto and T.J. Bastow, (2001) *Acta mater*. 49, 41.
M.H. Cohen and F. Reif, (1957) *Solid State Physics*, 5, 321.
V.G. Davydov, T.D. Rostova, V.V. Zakharov, Yu. A. Filatov and V.I. Yelagin, (2000) *Mater. Sci. Eng. A* 280, 30.

A. Dupasquier, R. Ferragut, M.M. Iglesias, C.E. Macchi, M. Massazza, P. Mengucci, G. Riontino and A. Somoza, (2004) *Mater. Sci. Forum* 445–446, 16.
E. Fukushima and S.B.W. Roeder, (1981) *Experimental Pulse NMR, a Nuts and Bolts Approach*, Addison-Wesley, Reading, Mass.
A. Guinier, (1938) *Nature* (London), 142, 569.
C.R. Hutchison, X. Fan, S.J. Pennycook and G.J. Shiflet, (2001) *Acta mater*. 49, 2827.
K.B. Hyde, A.F. Norman and P.B. Pragnell, (2001) *Acta mater*. 49, 1327.
T. Konno, M. Kawasaki and K. Hiraga, (2001) *J. Electron Microsc*. 50, 105.
S.K. Maloney, K. Hono, I.J. Polmear and S.P. Ringer, (2001) *Micron*, 32, 741.
B. Noble and A.J. Trowsdale, (1995) *Scripta Metall. et Mater*. 33, 33.
I.J. Polmear, (1995) *Light alloys: metallurgy of the light metals*, 3rd edn, J. Wiley & Sons, New York.
D.A. Porter and K.E. Easterling, (1992) *Phase transformations in metals and alloys*, 2nd edn, Chapman & Hall, London; New York.
G.D. Preston, (1938a) *Nature* (London), 142, 570.
G.D. Preston, (1938b) *Proc. Roy. Soc. A*, 167, 526.
G.D. Preston, (1938c) *Phil. Mag*. 26, 855.
R.J. Rioja, (1998) *Mater. Sci. Eng. A* 257, 100.
J.M. Silcock, T.J. Heal and H.K. Hardy, (1953) *J. Inst. Metals*, 82, 239.
C.P. Slichter, (1996) *Principles of Magnetic Resonance*, Springer Series in Solid State Sciences.
S.K. Son, M. Takeda, M. Mitome, Y. Bando and T. Endo, (2005) *Mater. Lett*., 59, 629.
G. Song, A. Bowles and D.H. St John, (2004) *Mater. Sci. Eng. A*, 366, 74.
J.C. Williams and E.A. Starke, Jr, (2003) *Acta mater*. 51, 5775.

11

Nanocrystalline light metal hydrides for hydrogen storage

T K L A S S E N, GKSS Research Centre Geesthacht GmbH, Germany

11.1 Introduction

Considering the increasing pollution and indiscriminate exploitation of fossil energy resources, new energy concepts are essential for the future of industrial society. Renewable sources have to replace fossil fuels that produce carbon dioxide upon combustion, which as a greenhouse gas is largely responsible for global warming. Hydrogen is the ideal means of energy storage for transportation and conversion of energy in a comprehensive clean-energy concept. Hydrogen can be produced from water using electricity, ideally produced from regenerative sources, such as wind, solar or water. Upon reconversion into energy, only water vapor is produced, leading to a closed energy cycle without harmful emissions. Apart from stationary applications, hydrogen energy is especially suited for mobile applications, e.g., zero-emission vehicles. However, appropriate storage facilities, both for stationary and mobile applications, are complicated, because of the very low boiling point of hydrogen (20.4 K at 1 atm) and its low density in the gaseous state (90 g/m^3). Furthermore, the storage of hydrogen in liquid or gaseous forms imposes safety problems, in particular for mobile applications, such as in the future zero-emission vehicles.

Metal hydrides are a safe alternative for hydrogen storage due to their long-term stability and low hydrogen pressures. In addition, they have a high volumetric energy density that is about 60% higher than that of liquid hydrogen [1]. Lightweight metal (Mg- and Al-) based hydrides have a high storage capacity by weight and are therefore favored for automotive applications. However, so far light metal hydrides have not been considered competitive because of their rather sluggish sorption kinetics. Filling a tank could take several hours. A breakthrough in hydrogen storage technology was achieved by preparing nanocrystalline hydrides using high-energy ball milling [2]. These new materials show very fast absorption and desorption kinetics within minutes, thus qualifying lightweight Mg- or Al-based hydrides for storage applications. For automotive application, a range of 500 km is desirable and

the lifetime of a tank system should exceed 500 cycles, corresponding to a total of 250 000 km.

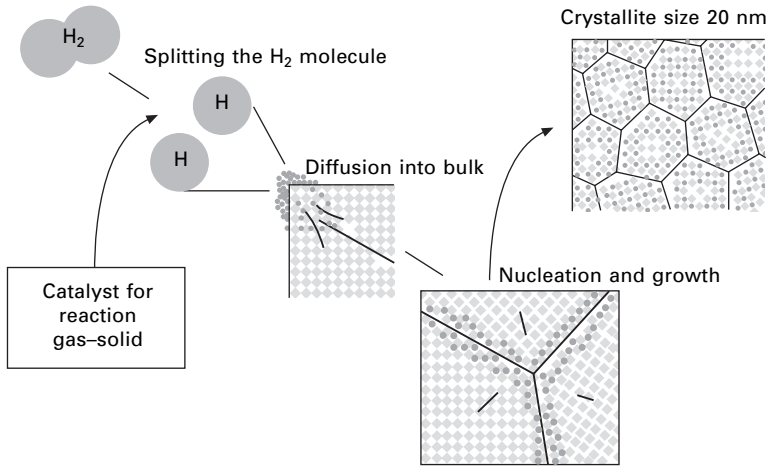
11.2 Production of nanocrystalline light metal hydrides

Different methods have been developed for the production of nanocrystalline metal hydrides, including classical ingot metallurgy [3], chemical routes [4] as well as high-energy milling [2]. For Mg-based alloys, different milling procedures have been reported. In principle, elemental powders can be blended in the desired overall composition, and milled together until chemical homogeneity and nanocrystalline microstructures are reached. The milling time may be shortened significantly by using pre-alloyed crushed ingots, thus avoiding time-consuming chemical homogenization. Some compositions, however, are difficult to process due to their high ductility, which in extreme cases may lead to complete cold-welding of the powder material to the milling tools, reducing the output. In these cases, it is favorable to hydride some or all of the components prior to milling, leading to the desired embrittlement of the powder blend for efficient milling. Accordingly, all materials presented in this study were produced using the respective available hydrides instead of pure materials for milling [5]. Alternatively, milling may be carried out under hydrogen atmosphere, forming the hydrides *in situ* [5].

11.3 Processes of hydrogen absorption and desorption

The reaction of metallic elements with hydrogen consists of distinct steps: physisorption, chemisorption (recombination), surface penetration, diffusion, and hydride formation (decomposition) by nucleation and growth (Fig. 11.1). The slowest step determines the overall kinetic rate, and the experimentally measured kinetics represents the kinetics of the slowest part of the reaction. The kinetic curve at each step has a characteristic form, which can be formulated as equations relating the transformed phase fraction to time. Thus, it is possible to deduce the rate-limiting step of the kinetics, if a good fit of experimental data with a specific kinetic equation can be obtained. The different steps can be influenced by certain measures. Regarding the reaction between the storage material and gaseous hydrogen, suitable catalysts may enhance the physical adsorption as well as the chemical absorption (desorption), i.e., the splitting (recombination) of the hydrogen molecule. Catalysts may also improve the surface penetration of hydrogen into the bulk, which is often hindered by inactive surface layers, e.g., thick oxides.

Diffusion within the bulk is usually faster along grain boundaries, where



11.1 Schematic of the different steps of hydrogen absorption and hydride formation in a solid storage material.

the packing density of atoms is less. Grain boundaries are also favorable nucleation sites for the formation and the decomposition of the hydride phase. Therefore, nanocrystalline microstructures are expected to accelerate these reaction steps, as they provide a much higher density of fast diffusion paths as well as preferred nucleation sites. In the following sections, materials with different microstructures as well as selected catalyst additions will be compared with respect to their kinetic properties.

11.4 Nanocrystalline Mg-based hydrides

11.4.1 Development of microstructure, minimum grain size

Analysis of X-ray diffraction patterns for pure MgH_2 after different milling times proves the microstructural refinement process, starting from (38 ± 10) nm initially and reaching the minimum value of (12 ± 5) nm just after 20 h of milling. Longer milling times of even up to 200 h do not lead to any further significant microstructural refinement, as documented by the grain size evaluation (Table 11.1). While the initial material consists only of tetragonal β-MgH_2, one hour of milling is sufficient to form a significant amount of the

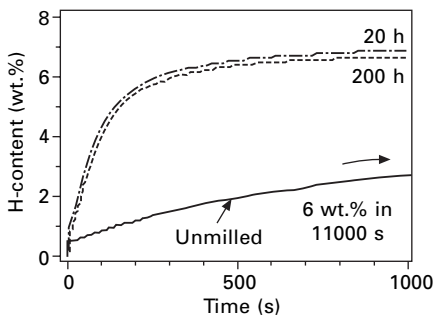
Table 11.1 Crystallite sizes of MgH_2 after different milling times, as estimated by the Scherrer method

Milling time (h)	0	1	2	5	10	20	50	100	200
Crystallite size (nm)	38	25	23	14	14	12	11	11	11

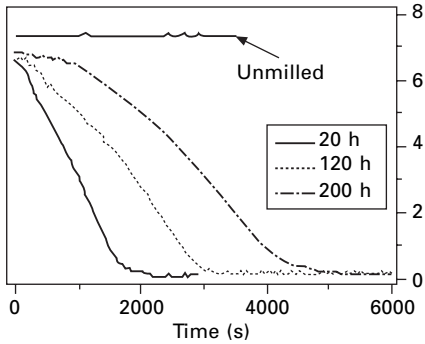
high-pressure phase γ-MgH_2, which is metastable and transforms into the equilibrium β-MgH_2–phase at 350 °C under ambient pressure. The amount of the γ-MgH_2 increases slightly with milling time and saturates as well after 20 h of milling (Fritsch® P5 planetary ball mill). For other milling machines or milling parameters, different fractions of γ-MgH_2 may be obtained. For example, the more aggressive milling conditions in a Spex® mixer mill or a Zoz Simoloyer® yields higher fractions of the γ-MgH_2 phase [6].

11.4.2 Effect of microstructural refinement on reaction kinetics and comparison to conventional coarse-grained hydrides

Figure 11.2 shows kinetic measurements at 300 °C of pure Mg milled for different times. While the reaction of the initial unmilled material is rather slow, 20 h of milling is sufficient to achieve a substantial increase in the kinetics. This is ascribed to the microstructural refinement upon milling that facilitates a higher diffusion rate due to the larger grain boundary density as well as a high density of heterogeneous nucleation sites for the hydride phase, as discussed above. Furthermore, it is known that diffusion of hydrogen through MgH_2 [7] is more than a factor of 5000 slower than diffusion through Mg [8] at 350 °C. As the hydride first forms along the grain boundaries, this leads to a rather slow filling of the interior of the crystallites, because hydrogen has to diffuse through the already formed hydride layers. In nanocrystalline materials, the volume fraction of the grain boundaries is much higher and the fraction inside the grains is much smaller, thus transformation into the hydride is significantly faster. Similarly, the desorption of hydrogen is also accelerated (Fig. 11.3). While the un-milled material does not desorb any hydrogen within one hour, desorption becomes more and more accelerated with increasing milling time. After 200 h of milling, the hydride can be fully transformed within 1800 seconds. These advantages of a nanocrystalline over coarse-



11.2 Absorption kinetics of pure Mg at 300 °C as a function of milling time.



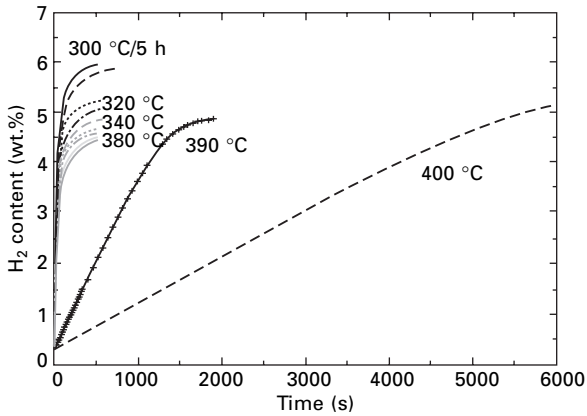
11.3 Desorption kinetics of pure Mg at 300 °C as a function of milling time.

grained microstructure have also been observed by other groups [2, 9, 10]. In addition, the particle size reduction upon milling also has a positive influence on the reaction kinetics. In general, the lower rate of desorption in comparison to absorption can be explained by the lower driving force ln p/p_{eq} [11], where p is the hydrogen pressure for absorption and desorption respectively. The equilibrium pressure p_{eq} depends on thermodynamic properties of the metal-hydrogen reaction (see reference 12) for details).

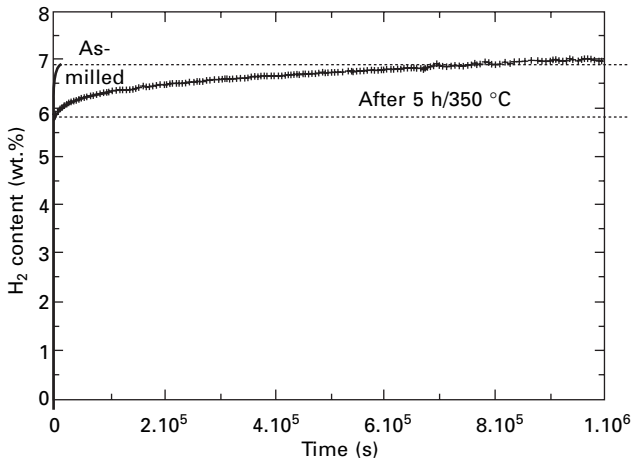
11.4.3 Thermal stability of the nanocrystalline microstructure

For technical applications of nanocrystalline metal hydrides, it is important to evaluate the thermal stability of the nanocrystalline microstructure. Due to the projected operation temperature of 300 °C to 320 °C, coarsening effects may occur and deteriorate the desired properties of the material. Therefore, the as-milled material was annealed at temperatures between 300 °C and 400 °C for 5 h, and the resulting sorption properties at 300 °C were determined and compared to the as-milled state.

Figure 11.4 shows the absorption curves of nanocrystalline MgH_2 after annealing at different temperatures. While the kinetics seem to remain as fast as for the initial material up to 380 °C, the storage capacity decreases continuously with increasing temperature. Above 380 °C, a dramatic slow-down of the reaction kinetics is also observed. However, the curves at 390 °C and 400 °C level off at higher storage capacities again than the one obtained at 380 °C. In order to clarify the issue of changing capacities, a long time measurement was performed after annealing for 5 h at 350 °C. After short measuring time, the curve levels off at 6 wt.% compared to 7 wt.% for the as-milled material (Fig. 11.5). Interestingly, absorption still continues at a much slower rate, and after long measurement time, the identical storage



11.4 Absorption kinetics of nanocrystalline MgH_2 after annealing treatments for 5 h each at the temperatures indicated.



11.5 Long-time measurement of absorption kinetics after annealing treatment for 5 h at 350 °C. The curve for the as-milled MgH_2 material is added for comparison.

capacity is obtained. This demonstrates that the actual storage capacity does not change with increasing temperature. However, the absorption process is divided into two distinct regimes. Within the first regime, the kinetics is as fast as in the initial as-milled state, while within the second regime, kinetics is as slow as in conventional, unmilled material. With increasing milling time, the fraction of material converted within the second regime continuously increases at the expense of the fraction converted within the first regime.

The results point towards the special kinetic processes in nanocrystalline hydrogen storage materials. As already mentioned above, grain boundaries are favorable nucleation sites for the phase transformation into the hydride

and vice versa. Therefore, hydrides tend to form at the surfaces and interfaces first. Afterwards, the hydride has to grow, which means that hydrogen has to diffuse through the existing hydride layers into the grains and fill the volume. However, diffusion of hydrogen within the hydride phase is substantially slowed down as compared to the unhydrided phase, e.g., by a factor of 5000 for Mg. In conventional coarse-grained materials, the hydrides formed at the interfaces thus act as diffusion barriers towards the transport of hydrogen into the interior of the grains. Therefore, overall reaction kinetics of coarse-grained material is slow. In contrast, for nanocrystalline material, the volume of each grain is much smaller and, thus, more easy to fill, because diffusion distances within the grains are short. During annealing, the grains continuously grow and diffusion distances increase accordingly. Therefore, the fraction related to the first fast kinetics regime corresponds to the hydride formation along the grain boundaries, while the fraction related to the slow kinetics regime can be attributed to the phase transformation within the grains, requiring diffusion through the existing hydride layer.

11.4.4 Additional effect of catalysts on reaction kinetics of nano-Mg

As mentioned above, the reaction of hydrogen gas molecules at the surface of the hydrogen storage material, i.e., physical van der Waals adsorption, chemical absorption reaction as well as diffusion into the bulk and vice versa, may all be accelerated using suitable catalysts, which provide favorable adsorption sites and may trigger electronic exchange reactions with the hydrogen gas molecule. This is especially important for Mg, because the probability of the adsorption of a H_2-molecule on the Mg surface is only 10^{-6} [13]. The main challenge is the identification and optimization of efficient catalytic materials. In the next sections, different materials will be discussed with respect to their catalytic activity for promoting the reaction of hydrogen gas molecules with Mg or Mg-based alloys and their hydrides as a model for sluggish light metal storage materials. Several aspects will be presented that characterize good catalysts and lead to further optimization of the kinetics of hydrides for hydrogen storage.

Comparison of different catalysts: metals, oxides, carbides, and nitrides

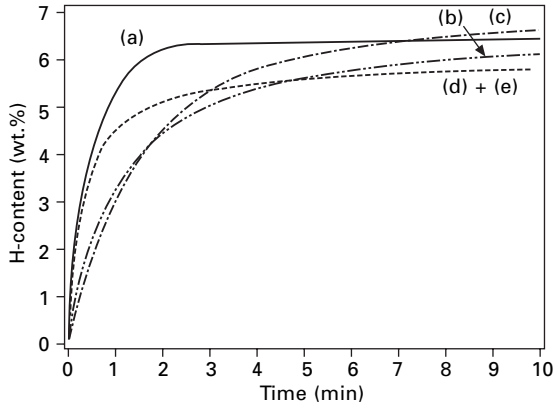
Absorption and desorption measurements demonstrate that V-containing compounds (V_2O_5, VN, VC) significantly enhance the absorption and desorption kinetics of nanocrystalline MgH_2. The absorption rates between zero and 3 wt.% capacity are about equal for all three catalysts and amount to approximately 1.6 wt.%/s (= 200 kW/kg), which is about three times as much as for pure nanocrystalline Mg. The maximum storage capacity is

typically reached after one minute. Differences with respect to the maximum hydrogen content are related to the different densities of the vanadium compounds, which were added in equal mole fractions of 5 mol%. Thus, the highest capacity is reached using VN (5.3 wt.%), followed by VC (4.9 wt.%), and V_2O_5 (4.0 wt.%). For V_2O_5, the capacity is lower than expected though possibly due to the formation of some inactive MgO. Neither vanadium oxide, nor vanadium nitride, nor vanadium carbide absorbs hydrogen under the chosen experimental conditions. The catalytic effect is even more pronounced upon desorption. Full desorption is feasible within two to 15 minutes, depending on the starting hydrogen content and the used vanadium compound. This has to be compared to a time of 30 minutes, needed for full desorption of hydrogen in pure nanocrystalline MgH_2, which, however, starts at a higher hydrogen content. To allow for a direct comparison, desorption rates between 80 and 20% of the respective maximum capacity have been calculated and are listed in Table 11.2. The values are significantly higher for the materials containing vanadium compounds. While differences upon absorption are marginal for all the catalytic additions, the most effective catalytical enhancement upon desorption is obtained using V_2O_5.

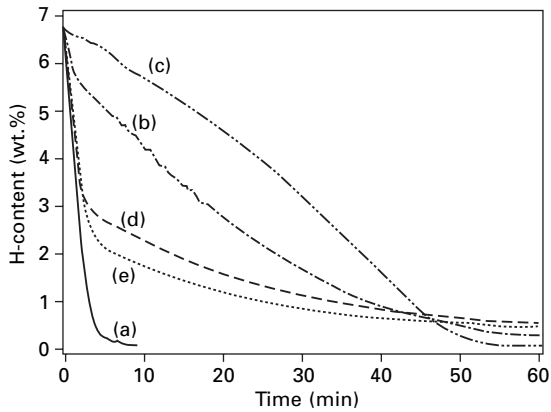
Table 11.2 Hydrogen desorption rates of nanocrystalline MgH_2 containing 5 mol% of V_2O_5, VC, and VN at 300 °C

MgH_2 + catalyst:		None	VC	VN	V_2O_5
Desorption rate	(wt.%/s)	–0.037	–0.07	–0.25	–0.31
	(kW/kg)	4.1	7.7	28	34
Density	(kg/m^3)		~4900	6095	3370

To compare the effect of an oxide catalyst with a pure metallic catalyst, two different powders were produced under identical parameters: one with 1 mol% V metal catalyst addition (oxygen < 180 ppm by weight), and the other with the addition of the related metal oxide, viz., V_2O_5. The respective absorption and desorption kinetics are compared to the nanocrystalline MgH_2 without catalyst in Figs 11.6 and 11.7. Except for the overall capacities, the absorption curves show no significant difference between $MgH_2/V_{0.01}$ and pure nanocrystalline MgH_2. The smaller capacity for the system $MgH_2/V_{0.01}$ is related to the inactive V that is not hydrided under the chosen experimental conditions. Obviously, metallic V has no accelerating effect for the hydriding of Mg. In contrast, V_2O_5 provides fast absorption of hydrogen even in this reduced proportion and full capacity is reached within two minutes. Upon desorption, the differences become even more significant: V_2O_5 facilitates complete dehydrogenation within six minutes, whereas around 50 minutes are required for $MgH_2/V_{0.01}$ and pure Mg. Again, kinetics rates were calculated



11.6 Hydrogen absorption curves of as-synthesized (a) $MgH_2/(V_2O_5)_{0.01}$, (b) $MgH_2/V_{0.01}$, (c) pure MgH_2, (d) the same material $MgH_2/V_{0.01}$ as in (b), but stored under Ar for 11 months, and (e) the same material $MgH_2/V_{0.01}$ as in (d) subsequently exposed to air for 17 h; each at 300 °C under a hydrogen pressure of 8.4 bar.



11.7 Hydrogen desorption curves of as-synthesized (a) $MgH_2/(V_2O_5)_{0.01}$, (b) $MgH_2/V_{0.01}$, (c) pure MgH_2, (d) the same material $MgH_2/V_{0.01}$ as in (b), but stored under Ar for 11 months, and (e) the same material $MgH_2/V_{0.01}$ as in (d) subsequently exposed to air for 17 h; each at 300 °C in vacuum.

Table 11.3 Hydrogen desorption rates of nanocrystalline MgH_2 containing 1 mol% of V_2O_5, 1 at.% V, and without any additions at 300 °C

MgH_2 + catalyst		None	V	V_2O_5
Desorption rate	(wt.%/s)	–0.037	–0.04	–0.31
	(kW/kg)	4.1	4.4	34

for better comparison and are listed in Table 11.3. Apparently, metallic V does not significantly influence the hydrogen sorption kinetics under these conditions.

One reason for this observation might be the different milling behavior of metals and oxides, which may lead to a finer microstructure and finer distribution in the case of oxide additions compared to pure metal additions. Metallic catalysts might be more difficult to disperse due to their ductile behavior limiting the fracture of fine particles. However, in case of the metallic catalyst Pd, Zaluski *et al*. have found by means of TEM investigations that the microstructures of the high-energy ball milled storage alloys Mg_2Ni, FeTi and $LaNi_5$ with and without catalyst do not differ from each other [14]. In particular, the nanocrystalline powder was not covered by a Pd-layer, but small Pd-clusters were cold-welded onto powder particles. Similar microstructures have also been observed by Song *et al*. for Mg/Ni-composites [15] and by Liang *et al*. for MgH_2/V-composites [16] after high energy ball milling. We therefore conclude that by optimizing the processing conditions metal as well as oxide catalysts can be homogeneously distributed in the material.

To demonstrate that microstructural reasons are not responsible for the poor catalytic effect of V in the present study, the material $MgH_2/V_{0.01}$ has been stepwise exposed to air. If a coarser distribution of the metallic V was responsible for the non-catalytic activity, kinetics should stay slow even upon oxidation/nitridation. Figs 11.6 and 11.7 also show absorption and desorption curves of the same $MgH_2/V_{0.01}$ material, stored under Ar atmosphere for eleven months (curve d) and the same material after additional exposure to air for 17 hours at room temperature (curve e). The amount of oxygen picked up in the Ar atmosphere is unknown, while the oxygen (O_2) absorbed during the additional treatment in air was estimated to be 0.65 wt.% based on the gas volume (stochiometric $MgH_2/(V_2O_5)_{0.01}$ contains about 2.9 wt.% oxygen). A strongly different behavior compared to the 'as-synthesized' $MgH_2/V_{0.01}$-powder was observed. Absorption and desorption of hydrogen starts with similar sorption rates as have been found for the $MgH_2/(V_2O_5)_{0.01}$ composite material. In case of the 'Ar-exposed' material, the desorption rate decreases abruptly at a hydrogen content of 3 wt.%, whereas the desorption rate of the 'air-treated' material drops down to a hydrogen content of 2.3 wt.% before levelling off. In both cases, the calculated desorption rate of this slower part of the desorption process is as low as –0.01 wt.%/s (= 1.1 kW/kg), which is in the range of desorption rates observed for materials with pure V.

Obviously, the V-particles are oxidized or nitrided during 11 months storage under Ar-atmosphere and during exposure to air, respectively. Oxidation leads to enhanced sorption kinetics and therefore demonstrates the catalytic behavior of metal oxides. Apparently, the point at which the desorption rate slows down is related to the amount of oxygen absorbed by the material.

After exposure to air, it can be assumed, that V becomes locally oxidized at the surface, while some metallic V remains within Mg particles which, therefore, is catalytically inactive. In conclusion, our results show that the microstructural distribution is not responsible for the observed differences in catalytic activity between V and V_2O_5/VN/VC.

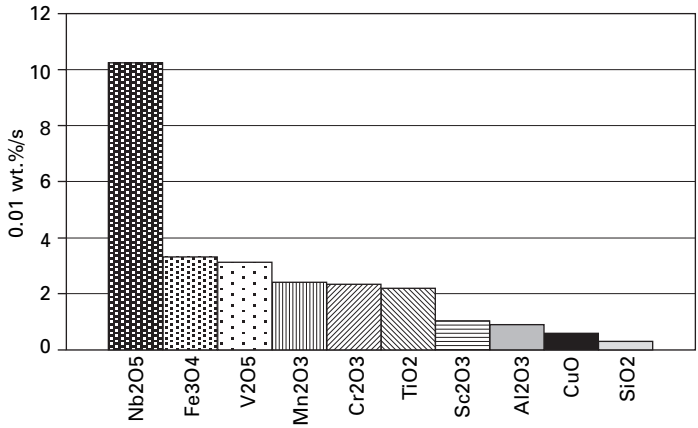
Finally, one may raise the question, whether metallic V has a catalytic effect at all. In case of V, the enhanced absorption kinetics are ascribed to a spill-over effect involving the formation and decomposition of vanadium hydride, thus acting as a hydrogen pump. However, Tanguy *et al*. [17] and partly Liang *et al*. [16, 18] have chosen temperature and pressure conditions for the absorption of hydrogen, at which the equilibrium pressure of the investigated room temperature hydrides, in particular VH_2, is much higher than the applied hydrogen pressure, so that VH_2 is not stable [16–18]. Therefore the formation of room temperature hydrides acting as 'hydrogen pumps' has to be reconsidered carefully.

In contrast to the present investigations, Tanguy *et al*. [17] have shown an improved H_2-sorption in microcrystalline Mg/$V_{0.1}$-composites. However, the alloy was prepared in air and so reaction with oxygen cannot be ruled out. Therefore, it can be assumed that vanadium oxide rather than pure V is causing the catalytic effect; the more so as already small amounts of 0.2 mol% show a significant effect [19]. More care to avoid oxidation was taken in recent investigations by Liang *et al*. [16, 18], demonstrating catalytic improvements for 5 at.% of the transition metal with a purity of 99%. Considering a typical impurity level of 1 wt.% oxygen for V as an example, 5 at.% of metal catalyst would correspond to 0.03 mol% of V_2O_5, which might cause the catalytic behavior. Respective experiments are designed to check the catalytic effect of oxides in a further reduced volume fraction and are presented on page 277.

Summarizing the results, all investigated compounds V_aX_b = V_2O_5, VN, VC act as catalysts for the magnesium–hydrogen reaction. V_2O_5 yields the fastest desorption kinetics followed by VN and VC. Ultra-pure metallic V does not show any improvement over pure Mg with respect to reaction kinetics. However, kinetics can be successively improved towards Mg + VN/V_2O_5/VC by stepwise exposure to small amounts of air. Based on these results, it is worthwhile to look at other oxides with respect to their potential catalytic properties. Comparisons may be helpful to draw conclusions about the mechanisms of catalysis and for further optimization.

Comparison of different oxide catalysts

To give a general overview of the catalytic activity of different oxides, the respective desorption rates are compared in Fig. 11.8. The fastest desorption kinetics was achieved for Nb_2O_5, yielding a desorption rate of 0.102 wt.



11.8 Comparison of the desorption rates of MgH_2 with different metal oxide catalyst additions at 300 °C into vacuum.

%/s. The calculated activation energy for hydrogen desorption of magnesium catalyzed with 0.2 mol% Nb_2O_5 was calculated to 62 kJ/mol, which is substantially lower compared to pure nanocrystalline Mg (120 kJ/mol).

Influence of the oxide catalyst content

To determine the best compromise between high efficiency of the catalyst and maximum capacity, the influence of different amounts of catalysts was investigated, using magnesium with different Nb_2O_5 contents in the range of 0.05 to 1 mol% as an example. Although differences are rather small, the absorption kinetics is slightly better using higher Nb_2O_5 content. In general, about 7 wt.% of hydrogen are absorbed within one minute at 300 °C and 8.4 bar hydrogen pressure. Differences in the desorption kinetics are more pronounced.

In general, an increase of Nb_2O_5 content again has a positive effect on the desorption kinetics. However, a saturation limit is reached for 0.5 mol% of Nb_2O_5, and kinetics does not accelerate further if more catalyst is present. With 0.5 mol% Nb_2O_5, magnesium hydride desorbs the full 7 wt.% of hydrogen within 90 s. To our knowledge, this is the fastest desorption kinetics for magnesium hydride reported so far. It is worth noting that low catalyst contents also have a significant effect on kinetics. For example, with only 0.1 mol% Nb_2O_5, full desorption takes place in 240 s. Even using only 0.05 mol% of Nb_2O_5, full desorption is possible within 500 s. Similar results are obtained at 250 °C. Absorption kinetics is almost independent of the amount of catalyst used, and even 0.05 mol% leads to fast absorption of about 6 wt.% of hydrogen within 1 min.

The desorption kinetics are again more critical and the optimum kinetics

Table 11.4 Reaction rates (wt.%/s) calculated between 20% and 80% of maximum capacity

Nb_2O_5 (mol%)	Absorption rate at 250 °C	Absorption rate at 300 °C	Desorption rate at 250 °C	Desorption rate at 300 °C
0.05	2.30	1.46	0.03	0.16
0.10	3.00	1.52	0.08	0.31
0.20	3.15	1.85	0.12	0.50
0.50	3.80	2.05	0.14	1.02

are achieved with catalyst contents of 0.5 to 1 mol%. Analogous to the results at 300 °C, the desorption curve for 0.2 mol% Nb_2O_5 is only slightly slower by about 20%. All reaction rates are listed in Table 11.4 for quantitative comparison. Compared to other catalysts, Nb_2O_5 is superior even at 250 °C for both absorption and desorption. The calculated kinetic rates in Table 11.4 show that the absorption rate at 250 °C is nearly twice as fast at 300 °C, independent of the catalyst content. This result can simply be explained by the higher thermodynamic driving force for the absorption reaction at lower temperatures at identical hydrogen pressures, according to the van't Hoff equation. This is a further proof that the kinetic barriers are efficiently reduced by Nb_2O_5 additions. Thus, the activation energy is lower and the barriers can be overcome even at lower temperatures. Consequently, the influence of the thermodynamic driving force increases and the reaction rate at 250 °C is faster than at 300 °C.

Compared to studies using pure metallic Nb [20], striking differences are observed with respect to the amount of catalyst required for equally fast kinetics. For full desorption within 250 s at 300 °C, 5 mol% of pure Nb is necessary, whereas only 0.1 mol% of Nb_2O_5 is sufficient to desorb even more hydrogen in the same time. At first glance, this proves that the catalytic effect of the metals in the pure state is much weaker than in the respective oxide form. However, as already discussed in the previous sections, the question may be posed whether pure metallic Nb is catalytically active at all. Considering the high affinity of niobium to oxygen, it may be reasonable to conclude that at least a large part – if not all – of the reported catalytic effect of pure metals is due the presence of oxides in the samples. If one assumes that the Nb used in other studies has technical purity, i.e., the material contains about 1 wt.% oxygen, an amount of 5 mol% Nb would include 0.0581 mol% of Nb_2O_5. In addition, if the catalytic effect of pure metals is studied, special care has to be taken not to expose the specimen to ambient air at any time; otherwise, formation of additional oxides cannot be excluded. In this regard, it is also worth noting that not only oxides but also nitrides show a good catalytic effect, which may add to the observed catalytic effect, if commercially pure metals are used.

In order to elucidate the mechanism of catalysis of transition metal oxides, tentative activation energies for the desorption reaction of magnesium catalyzed with different Nb_2O_5 contents are calculated using the kinetic rate constants at the two different temperatures. The calculated value of 62 kJ/mol is in good agreement with the results for pure metallic Nb [20], where an activation energy of 64 kJ/mol was determined. This may be taken as a further indication that identical phases are present and identical processes take place.

Influence of oxide catalysts on kinetics of Mg-based alloys and compounds

To investigate the catalytic effect of metal oxides on the hydrogen sorption characteristics of Mg-based alloys, the absorption and desorption behaviour of nanocrystalline $Mg_2Ni/(Mn_2O_3)_{0.01}$ was measured at 300 °C. Pure nanocrystalline Mg_2Ni powder absorbs hydrogen fast by itself and therefore additions of the metal oxides do not lead to any further significant acceleration of the hydrogen absorption. However, a reaction rate twice as fast is obtained for the release of hydrogen. Full dehydrogenation is reached within less than 60 seconds. In contrast to nanocrystalline Mg_2Ni without catalyst, ball-milled Mg_2Ni with 1 mol% catalyst releases hydrogen even at a temperature as low as 210 °C, with full desorption obtained after about 60 minutes.

11.4.5 Factors determining good catalysts for hydrogen sorption of nanocrystalline Mg

In order to clarify, what the actual mechanism of catalysis for hydrogen sorption in either case is, a detailed analysis of the kinetic data has been performed. Different aspects related to catalysts have been considered and investigated regarding their potential influence on the reaction.

Rate-limiting reaction step

As mentioned above, the reaction of hydrogen with metals and alloys consists of several distinct steps. The slowest step determines the overall kinetic rate, and the experimentally measured kinetics represents the kinetics of the slowest part of the reaction. The kinetic curve for each step has a characteristic form, which can be formulated as equations relating the transformed phase fraction to time. Thus, it is possible to deduce the rate-limiting step of the kinetics if a good fit of experimental data with a specific kinetic equation can be obtained. Using this method, the rate limiting steps of the reactions for different Nb_2O_5 contents at different temperatures are identified and the results are listed in Table 11.5. In detail, the results can be interpreted as follows:

Table 11.5 Rate-limiting step in desorption reactions with various Nb_2O_5 concentration. K is a constant, t represents time, and α represents the transformed phase fraction

Used kinetic equation	Best fit found for reaction	Rate-limiting step
$1 - (1 - \alpha)^{1/2} = kt$	at 250 °C and 0.2 mol% at 300 °C and 0.2 mol% at 250 °C and 0.5 mol% at 300 °C and 0.5 mol% at 250 °C and 1 mol% at 250 °C and 1 mol%	Contracting volume in 2-D [11]: i.e., contracting area, growth with constant interface velocity, (interface controlled)
$\alpha = kt$	at 300 °C and 0.05 mol% at 300 °C and 0.1 mol%	Surface controlled (chemisorption) [11]
$[-\ln(1 - \alpha)]^{1/3} = kt$	at 250 °C and 0.05 mol% at 250 °C and 0.1 mol%	Johnson–Mehl–Avrami (JMA) kinetics in 3-D [11,12]: growth of existing nuclei with constant velocity

1. At 250 °C and with catalyst contents lower than 0.2 mol%, the kinetic curve is best fitted by a three-dimensional Johnson–Mehl–Avrami (JMA) approach, i.e., kinetics is growth-controlled. This can be explained by the slow diffusion at this low temperature. In addition, the slow hydrogen draining due to a lack of catalyst and the related slow recombination rate leads to rather small concentration gradients, reducing the driving forces for diffusion.
2. At 250 °C and with catalyst contents above 0.2 mol%, the recombination rate is high enough and the kinetic rate-limiting step is changed to interface-controlled.
3. For a reaction temperature of 300 °C, with a catalyst content lower than 0.2 mol%, the kinetics is clearly surface controlled. Due to easier diffusion of hydrogen atoms at 300 °C, transport of hydrogen does not play an important role any more. Now, the reaction rate is determined only by the slow gas–solid reaction due to the low catalyst content.
4. At 300 °C and catalyst contents of 0.2 mol% or more, the recombination rate of hydrogen atoms increases up to a limit where it is not rate-determining any more and the reaction becomes interface controlled again.

Therefore it can be concluded from the above analysis that the oxide catalyst indeed improves mainly the chemisorption step of the overall kinetics, i.e., the surface reaction between the hydrogen gas molecule and the solid hydrogen storage material. If the catalyst content reaches 0.5 mol%, the overall reaction is not limited by the catalyst reaction any more.

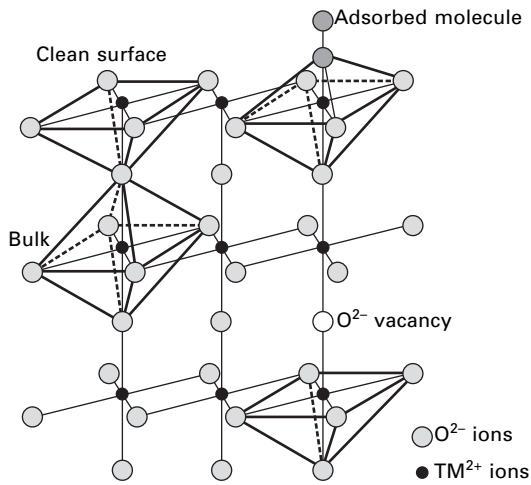
Mechanisms of reaction catalysis

Because of their distinct physical and chemical properties, transition metal (TM) oxides have been the focus of extensive research over the last sixty

years. They have been used in many technological applications for different purposes, including catalysis. These compounds have been known to be efficient catalysts in many chemical reactions involving gases and many investigations have been conducted to explain these catalytic effects. There are some interesting studies that contributed to clarification of the mechanisms.

As investigated by Fromme [21], TM ions on the surface and in the bulk of oxides may experience different crystal fields because of missing oxygen ions at the surface of these oxides. This causes a splitting of the electronic 3d state of these ions and can explain the catalytic behavior of TM oxides in adsorption of gas molecules. This phenomenon is not only applicable to the surfaces of these oxides but also to any defect in the bulk, which can lead to the same effects. As an example, Fig. 11.9 shows a generic crystal structure of transition metal oxide (like CoO). In the bulk, each TM ion is surrounded by six oxygen ions, but within the surface, one oxygen ion is missing, which causes a change in the electronic structure of the metal ion. Based on the experiments, Fromme concludes that the change in electronic structure of the TM ions is the main factor causing the high catalytic activity of oxide catalysts.

Usually in chemistry, pure metal catalysts are used in the form of thin films or clusters supported on ceramic substrates, mainly oxides, to provide a higher surface area for reaction, while keeping the amount of expensive noble metal catalyst small. Although initially the oxide supports were assumed



11.9 Illustration of the differences in symmetry of the crystal field experienced by bulk and surface transition-metal ions. Bulk ions are octahedrally surrounded by six O^{2-} ions. At the clean (100) surface (left), one O^{2-} ion of the octahedron is assumed to be missing and the surface transition-metal ions are tetragonally surrounded by five O^{2-} ions only, leading to a C4v symmetric crystal field. Bulk transition-metal ions located next to an O^{2-} vacancy are exposed to a very similar crystal field [21].

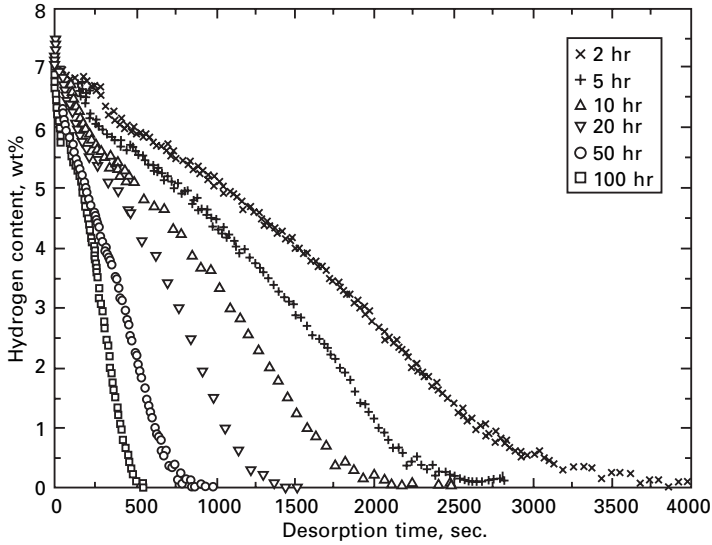
to be catalytically inert, further investigations have revealed that, depending on the support, the catalytic activity of the metal catalyst varies. The interface reaction between metal catalysts and oxide supports and the morphological and electronic structure of the support can have a remarkable influence on the catalytic activity and selectivity of the supported TM particles. This phenomenon (metal–support interaction) was first emphasized by Schwab [22], and is supported by recent experimental results [23]. Klie *et al*. analyzed the Pt/SiO_2 system at 350 °C. They observed that the electronic structure of Si is different in the interface as compared to the bulk of SiO_2, which can be attributed to the interaction of Pt atoms with the oxygen and/or Si ions of the support at the interface. In turn, the electronic structure of Pt is also modified, which improves its catalytic activity.

As another example, the effect of the support on the catalytic activity in the system MgO/Pd was investigated by Hammer [24]. The results show that there is a bond between the adsorbed molecules and the support, which is a general phenomenon for these systems. This means, that the electronic structure of the TM in the oxide may be favorably modified by interaction of the oxygen atoms with Mg, leading to excellent catalytic properties.

Factors affecting chemisorption on TM oxides

As mentioned above, the key to the catalytic effect of TM oxides is the electronic structure of the TM ions, and factors affecting it will determine the catalytic activity of the catalyst. Four potential factors are discussed below.

1. Structural defects of oxides
   As mentioned above (Fig. 11.9), structural defects in the oxides, like lack of oxygen ions on the surface, will change the electronic structure of the TM and will promote the adsorption of gas molecules. Milling is well known to introduce a high number of crystal lattice defects into the material. Thus it is interesting to study the efficiency of the catalyst with increasing milling time, i.e., with increasing defect density. In order to separate the effect of microstructural refinement of the MgH_2, the hydride was first milled separately until the final minimum grain size was reached. Afterwards, the catalyst was added and milling was resumed for the times indicated. As Fig. 11.10 demonstrates, kinetics indeed improves continuously with milling time.
2. Interaction of Mg with oxygen ions of oxides
   Another factor that can change the electronic structure of TM is the interaction of magnesium atoms with oxygen ions of the catalysts. In principle, magnesium may bond to one oxygen atom of the oxide structure due to its higher thermodynamic stability (Table 11.6) and the lowest chemical potential for oxygen. This creates a defect in the TM oxide, changing in turn the electronic structure of the TM. Obviously the stronger



11.10 Effect of milling time on the desorption kinetics of Mg-H reaction catalyzed with 0.05 mol% Nb_2O_5 at 300 °C.

Table 11.6 The thermodynamic properties of oxide catalyst and experimental desorption rates

Compound	TM hydride enthalpy of formation (kJ/mol-H)	TM compound enthalpy of formation (kJ/mol-O)	TM valence state per TM atom	Desorption rate (wt.%/s)
Nb_2O_5	–46	–379	5	1.02
V_2O_5	–36	–310	5	0.60
Ta_2O_5	–40	–409	5	0.36
NbO_2	–46	–398	4	0.35
Mn_2O_3	–11	–319	3	0.25
NbO	–46	–405	2	0.20
TiO_2	–65	–472	4	0.19
Cr_2O_3	–16	–380	3	0.19
Al_2O_3	–15	–558	3	0.07
SiO_2	+361	–455	4	0.02
MgO	–	–635	–	–

the Mg-O interaction compared to TM-O interaction, i.e., the higher the difference in the chemical potential for oxygen, the more significant the change of the electronic structure of the TM will be. In other words, the chemical potential of oxygen μ(O) in MgO is compared with the chemical potential μ(O) in the TM oxide per oxygen atom. As a matter of fact, MgO is more stable than all effective oxide catalysts and provides the lowest chemical potential for oxygen. Based on thermodynamics, a complete

reduction of the TM oxide by Mg is expected, forming TM and MgO. However, this is generally not observed, which may be due to the experimental procedure, i.e., milling MgH_2 instead of pure Mg. Then, the hydride would have to decompose, representing an additional barrier. Nevertheless, in cases where the TM oxide is too weak and the described reduction reaction indeed occurs already upon milling, any catalytic effect diminishes. Furthermore, in reference 25 the effects of V and V_2O_5 are compared and only marginal catalytic effect is found for V, which demonstrates the important contribution of oxygen.

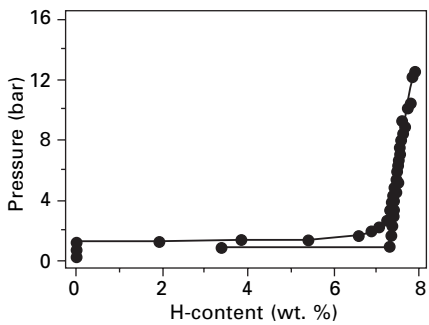
An example of a weak compound is Re_2O_7. Compared to other oxides, Re_2O_7 is rather unstable, thus it reacts immediately with magnesium and no catalytic effect can be observed. Consequently, efficiency of the catalyst increases only up to a certain extent with increasing chemical potential for oxygen, i.e., for decreasing thermodynamic stability in relation to the respective Mg compound. If the critical compound stability is exceeded, the respective Mg compound is formed and the catalytic effect vanishes. In order to test this conclusion, reactions of magnesium and hydrogen catalyzed with $TiCl_3$ and $TiCl_4$ are investigated and compared. The enthalpy of formation of $TiCl_3$ and $TiCl_4$ are –240.3 and –201.05 kJ/g-atom, respectively, i.e., $TiCl_3$ is more stable than $TiCl_4$ and already provides a lower chemical potential for oxygen. Therefore, Mg atoms can more easily change the electronic structure of Ti in $TiCl_4$ rather than $TiCl_3$. According to above conclusion, $TiCl_4$ should exhibit a better catalytic effect compared to $TiCl_3$, which is indeed in agreement with our experimental results.

3. Valence of the TM atom in the oxide
   As previously published [26], the multivalency of the TM within the compound is critical for being an efficient catalyst. To separate the different effects, the catalytic efficiency of the three different Nb oxides, NbO, NbO_2, and Nb_2O_5, is investigated. While the thermodynamic interaction of Nb with H is the same and the stability per oxygen atom is almost identical, these three compounds differ with respect to the oxidation state of Nb. As expected, our experiments indeed show an increase in the catalytic activity with the valance state of Nb. This can be explained by the ability of Nb in Nb_2O_5 to take or supply more electrons for the catalytic reaction as compared to NbO_2 or NbO.
4. The chemical interaction parameter between the TM atom and hydrogen
   The above conclusion, that hydrogen atoms are adsorbed first on the TM ions, further implies that a stronger TM-H interaction will improve the catalytic activity. To test this hypothesis, desorption kinetics of Mg with Nb_2O_5 and Ta_2O_5 are compared. As expected, the reaction rate using Nb_2O_5 is significantly higher than for Ta_2O_5. Because the thermodynamic stability of both is almost equal and the stoichiometry is identical, this difference is mainly attributed to the stronger interaction of Nb with H.

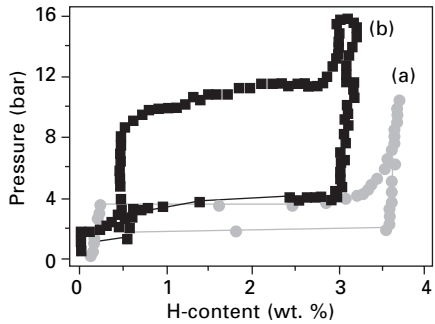
Summarizing the results, the catalytic activity of TM oxide catalysts is influenced by several factors: 1. defect density of the oxide catalyst; 2. thermodynamic stability of the TM oxide with respect to MgO; 3. the affinity of the TM atom to hydrogen, i.e., the stability of the respective hydride; 4. the valence state of the TM ions in catalyst, i.e., the number of electrons available for bonding. Factor 1. should depend on the milling conditions and can be assumed to be similar for the investigated oxides. It is evident that, if two of the remaining factors are equal, the better catalytic activity is determined by the third factor. The results facilitate further optimization of catalysts in view of a technical application.

11.4.6 Thermodynamic properties of Mg and Mg-based alloys

Figure 11.11 shows the pressure–composition–temperature (pcT) diagram for nanocrystalline Mg at 300 °C. At a pressure of 2 bars, the hydride is formed and the corresponding pressure plateau extends to 7.6 wt.%, which corresponds to full transformation of the Mg into MgH_2. After that, the hydrogen content cannot increase any further with increasing pressure, because the solubility of hydrogen in the hydride phase is low. If the pressure is decreased again, a second plateau at lower pressure, here 1 bar, is reached, at which the hydride decomposes again. The hysteresis has previously been ascribed to mechanical stress between the metal and the hydride phase [27], as the volume expansion upon hydriding amounts to about 40%. However, in this case it is most probably due to kinetic constraints during the measurement, because the driving force is low in the plateau region, where the pressures are close to the equilibrium pressure. For the nanocrystalline hydride Mg_2NiH_4 (Fig. 11.12a) the pressures are about twice as high as for Mg at this temperature. By replacing half of the Ni with Cu (Fig. 11.12b), a further increase in the plateau pressures to 11 and 4 bar, respectively, is achieved, indicating a substantial thermodynamic destabilization of the hydride phase with respect



11.11 pcT-diagram of pure nanocrystalline Mg at 300 °C.



11.12 pcT-diagrams of (a) nanocrystalline Mg_2Ni and (b) nanocrystalline $Mg_2Ni_{0.5}Cu_{0.5}$ at 300 °C.

to MgH_2 and Mg_2NiH_4. For application, a pressure of at least 1 bar is required, which is reached at 300 °C for pure Mg [28]. For Mg_2Ni, the same pressure is already available at 240 °C [29, 30], and for $Mg_2Ni_{0.5}Cu_{0.5}$ the temperature is further decreased to 230 °C [31]. The lower desorption temperature and related lower heat of formation leads to a higher efficiency of the hydride, because in turn also less heat is required for desorption. However, at the same time the capacity has decreased from 7.6 wt.% over 3.6 wt.% to about 3.1 wt.%. Similar results on Mg-based alloys have also been obtained by other groups [32–34]. More work is necessary to further reduce the temperature, while keeping a high capacity. Promising approaches will be discussed in the last section of this chapter.

11.4.7 Summary

The excellent results on Mg and Mg-based materials demonstrate that reaction kinetics of light metal hydrides can be drastically enhanced by using nanocrystalline microstructure and transition metal oxide catalysts. However, thermodynamics still requires high temperatures of operation for Mg and Mg-based hydrides. Thus, other hydrides and hydride phases have to be considered to identify materials with lower reaction heat and still sufficient storage capacity by weight.

11.5 Nanocrystalline alanates

Alanates are compounds that consist of a positively charged metallic ion, e.g., Na^+ or Mg^{2+}, and AlH_4^--groups. The alkaline alanates $LiAlH_4$ and $NaAlH_4$ are especially interesting due to their high mass-related capacities of 7.9 and 5.6 wt.%, respectively. Thermodynamically, the hydrogen bonds are rather weak, therefore hydrogen release should be possible at moderate temperatures slightly above room temperature. However, alanates decompose in two steps upon dehydriding according to:

$$3NaAlH_4 \leftrightarrow Na_3AlH_6 + 2Al + 3H_2\uparrow \leftrightarrow 3NaH + 3Al + 9/2H_2\uparrow \quad 11.1$$

This requires the diffusion of metal atoms, which leads to rather slow kinetics. Furthermore, rehydriding from the gas phase requires the recombination of the metal atoms again and could not be obtained until recently. Bogdanovic *et al.* [35] were able to demonstrate complete reversibility, if nanocrystalline alanates were used and $TiCl_3$ was added as a promoter agent. Meanwhile, a number of different promoters were tested with respect to their effects on kinetics. Sodium alanate can be readily reformed within ten minutes at around 100 °C, however, at rather high pressures of 80 bars or more [36, 37].

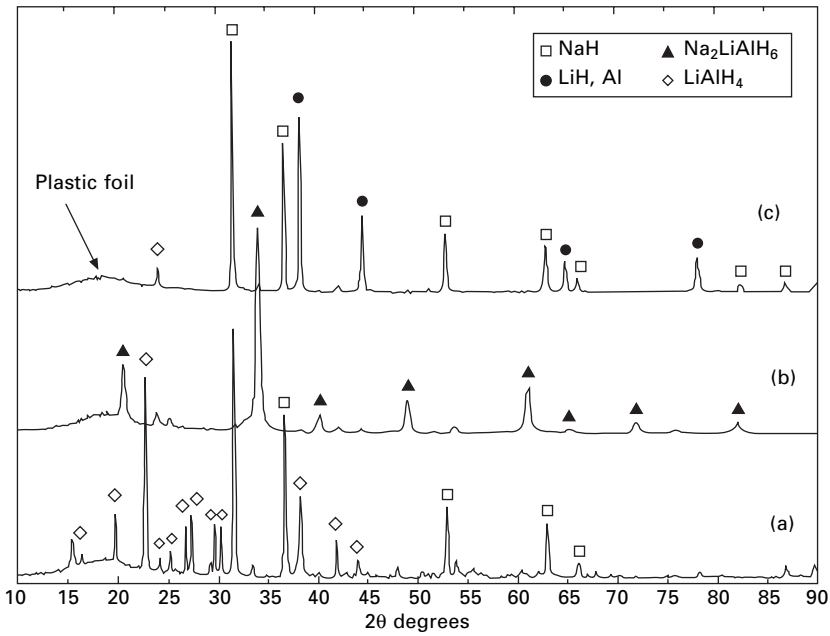
For $LiAlH_4$, a proof of reversibility is still lacking. To increase storage capacity, it is therefore worthwhile to only partially replace Na by Li, and thereby try to achieve higher capacities as well as full reversibility. As kinetics of the second decomposition step and its reversal is more sluggish than the first, this study focuses on the kinetics and thermodynamics of Na_2LiAlH_6. As a first attempt, the catalysts used to Mg are investigated regarding their effect on alanates. Here, TiO_2 is a promising material, because $TiCl_3$ has already been proven to be efficient for alkaline alanates, indicating favorable interaction of Ti with alanates.

To obtain the desired mixed alanates, powder blends consisting of the different respective hydrides have been milled together according to reference 5. Figure 11.13 confirms the formation of single-phase Na_2LiAlH_6 after milling of $2NaH + LiAlH_4$ powder blends. As proven by the XRD pattern after several cycles, the hydrogen desorption and absorption is fully reversible according to:

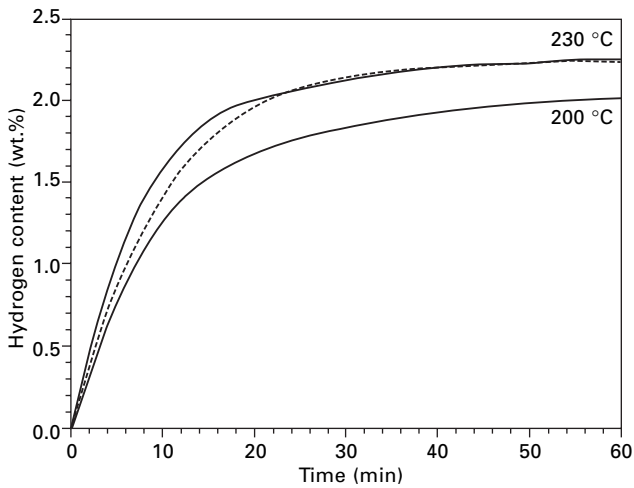
$$Na_2LiAlH_6 \leftrightarrow 2NaH + LiH + Al + 3/2H_2\uparrow \quad 11.2$$

Theoretically, the respective reversible hydrogen storage capacity amounts to 3.5 wt.%.

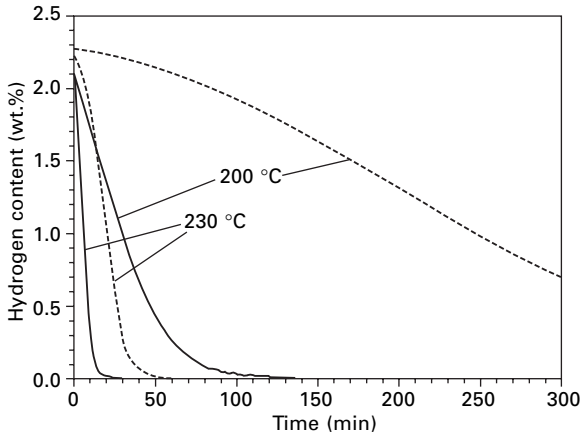
Absorption and desorption measurements for Na_2LiAlH_6 at 230 and 200 °C without and with TiO_2 as catalyst additions are shown in Figs 11.14 and 11.15. For both temperatures, the effect of the catalyst is obvious, though it is more pronounced upon desorption, analogous to the results for Mg-based hydrides. A capacity of 3 wt.% upon absorption at 230 °C, for example, is achieved in 20 minutes using the oxide catalyst, while it takes twice as long without the catalyst. Upon desorption at 230 °C, 3 wt.% H_2 is released in ten minutes versus 30 minutes without the catalyst. Similar improvements are achieved at lower temperatures. In order to determine the thermodynamic properties of Na_2LiAlH_6, pressure–composition isotherms were determined at different temperatures. Figures 11.16 and 11.17 show pcT-diagrams for Na_2LiAlH_6 with and without catalyst at 230 °C, and 200 °C for comparison. The faster kinetics achieved by catalyst additions leads to less pronounced



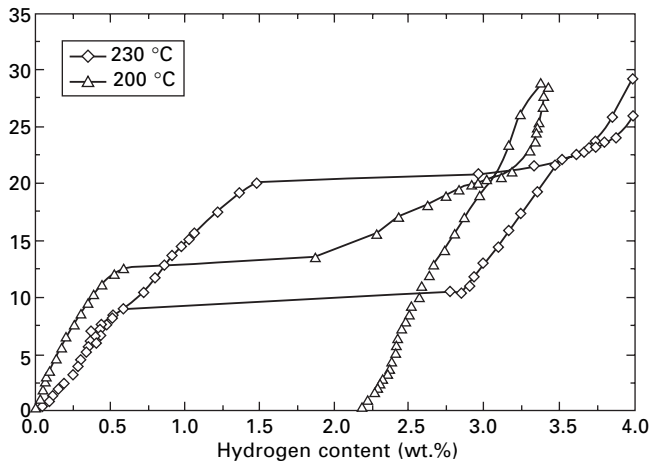
11.13 X-ray diffraction patterns of Na_2LiAlH_6. (a) Starting powder blend of 2NaH + $LiAlH_4$, (b) after 100 h of milling time, and (c) after several cycles of hydrogen sorption at elevated temperatures.



11.14 Hydrogen absorption curves of nanocrystalline Na_2LiAlH_6 with 5 mol% TiO_2 as a catalyst (solid line) and without catalyst (dashed line) at 230 °C and 200 °C.

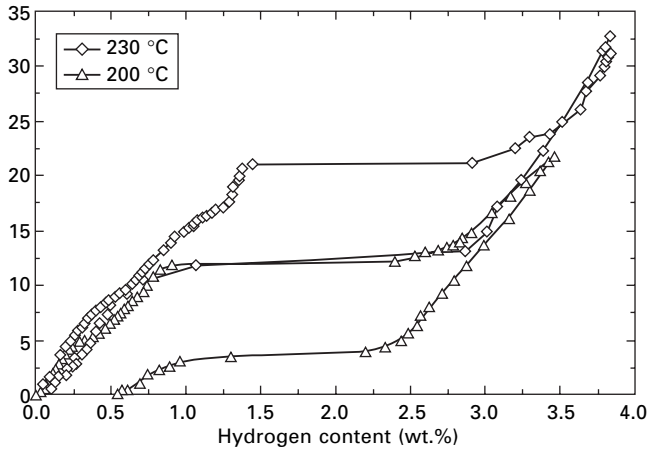


11.15 Hydrogen desorption curves of nanocrystalline Na_2LiAlH_6 with 5 mol% TiO_2 as a catalyst (solid line) and without catalyst (dashed line) at 230 °C and 200 °C.



11.16 Pressure–compositions isotherms of Na_2LiAlH_6 without catalyst at temperatures of 230 and 200 °C.

hysteresis (as already mentioned above for Mg-based alloys). This is reflected particularly by higher desorption pressure plateaus. While desorption at temperatures lower than 230 °C could not be measured for the pure alanate, reversibility is shown even for 200 °C using catalysts. At 230 °C, absorption and desorption plateaus are determined to be 21 and 10 bars, respectively. At 200 °C, a desorption pressure plateau of 3.8 bars is obtained. This indicates that there is still a potential for further reduction of temperature in view of the technical application of alanates.



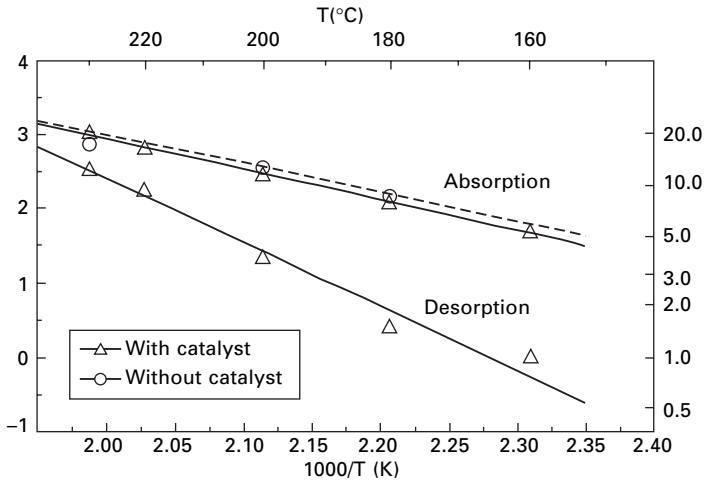
11.17 Pressure–compositions isotherms of Na_2LiAlH_6 with catalyst at temperatures of 230 and 200 °C.

Using the experimental pcT data, apparent van't Hoff plots were determined according to:

$$\ln\ P_{eq} = \frac{\Delta H}{RT} - \frac{\Delta S}{R} \qquad 11.3$$

where p_{eq} corresponds to the pressure plateaus. The calculated absorption and desorption enthalpies amount to ΔH_{abs} = 23 kJ/mol H_2 and ΔH_{des} = 48 kJ/mol H_2, respectively, for Na_2LiAlH_6 with catalyst. ΔH_{abs} = 22 kJ/mol H_2 was determined for absorption of the alanate without catalyst. Figure 11.18 shows the respective van't Hoff plots calculated from experimental results.

The plateau pressures for absorption and desorption show a large difference. This is attributed to kinetic constraints during the measurement. As can be seen in Figs 11.15 and 11.16, kinetics at reduced temperatures are rather slow and show periods of time with near zero slope, i.e., negligible hydrogen absorption and desorption, while the final state is not reached yet. During the automated pcT-measurement, the set conditions for termination of the measurement may be met in these 'near-zero slope regions', as hydrogen pick-up/release below a given limit within a certain time limit is the criterion that equilibrium has been reached and the next pressure value will be adjusted. Upon absorption, this leads to lower apparent concentrations and, consequently, higher apparent pressure plateaus. For the same reasons, the apparent pressure plateaus upon desorption are lower than the actual pressure plateaus. In the present case, the assumed deviation of the measured apparent plateau pressures from the actual pressure plateaus increases with decreasing temperature, as kinetics slow down significantly. This would lead to a reduced slope in the van't Hoff diagram for absorption and an increased slope for desorption,



11.18 van't Hoff plots calculated for Na_2LiAlH_6 without (dashed line) and with catalyst (solid line) from the experimental pcT measurements.

thus giving a large difference in apparent enthalpies. Although less pronounced, hysteresis is also observed in other systems, e.g., Pd-H, and there is still controversy regarding the origin of the hysteresis [38]. Under the assumption that the effects of kinetic constraints during absorption and desorption are comparable, the arithmetic average would be closest to the actual enthalpy of formation. In the present study, this would yield an enthalpy of about 35 kJ/mol H_2. There were no experimental enthalpies reported for Na_2LiAlH_6. For comparison, Bogdanovic *et al*. [39] have reported values of ΔH_{abs} = 45.93 kJ/mol H_2 and ΔH_{des} = 47.37 kJ/mol H_2 for Na_3AlH_6.

Summarizing the results on alanates, partial replacement of Na by Li is feasible and thus higher hydrogen storage capacities are accessible. The nanocrystalline microstructure ensures reversibility with respect to the decomposition/recombination reaction during hydrogen desorption/absorpion reaction. Metal oxides also show a significant enhancement of the reaction kinetics, indicating transferability of the good results obtained for Mg-based hydrogen storage materials. Obviously, these materials also improve the reaction of hydrogen gas molecules with alanates. However, further improvements of reaction kinetics require measures for enhancing diffusion of metal atoms, such as doping for the creation of additional vacancies in the crystal lattice.

11.6 Evaluation of technical potential of nanocrystalline hydrides

As the properties of nanocrystalline alanates are still under investigation and

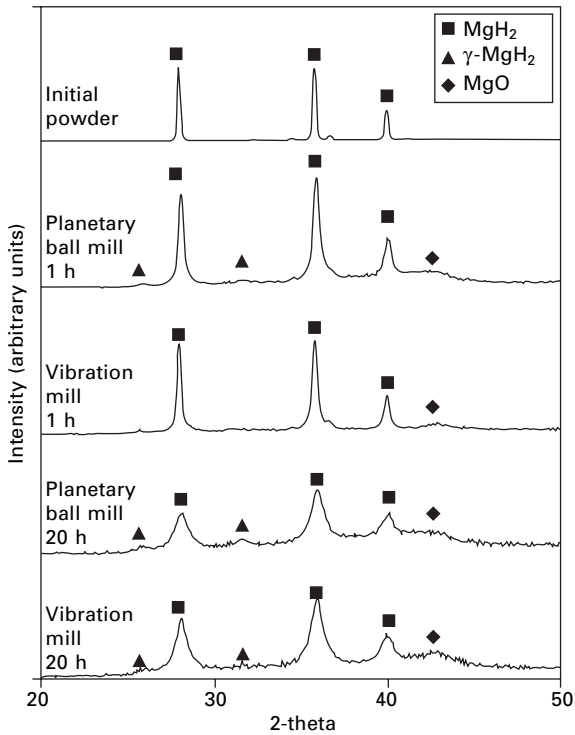
kinetics are not yet satisfactory, evaluation of the technical potential is carried out for Mg-based storage materials only. However, the evaluation regarding upscaling is generally transferable to any other hydride as well. Furthermore, some aspects for alanates are discussed in the last section of this chapter.

11.6.1 Upscaling of high-energy milling to produce larger quantities

So far the production of nanocrystalline powder has been mainly carried out in a planetary ball mill or laboratory vibration mills [40], where only amounts of less than 200 grams, mostly even less than ten grams, were processed at rather long milling times until saturation was reached. For the development of a prototype tank, however, it is necessary to provide large quantities of powder. For a future production process short milling times and low power consumption are required. Therefore, other types of mills have to be considered. A number of experiments have been carried out with horizontal attrition mills [41–44]. However, investing costs for horizontal attrition mills are relatively high and the amount of powder per batch is rather low for a technical production process.

In this section, investigations on the production of nanocrystalline powders in a common industrial tube vibration mill (Siebtechnik, type ESM 236-1bs) are presented. These mills are well established for comminution and activation of minerals [45]. Therefore, mills of this type are available up to tonnage loads and can be run in continuous mode. Although the kinematics and the kinetics of this kind of mill with respect to the milling intensity, mixing of the milled material and energy consumption have been optimized [46, 47], the influence of some critical parameters on the crystallite size reduction is not clear yet. Therefore, milling time, shape of grinding bodies and ratio of powder to grinding bodies have been systematically varied and optimized with respect to efficient crystallite refinement, in view of the desired application properties. In this regard, it is of particular interest whether the excellent powder properties achieved on the laboratory scale can be reproduced on a pilot-plant scale.

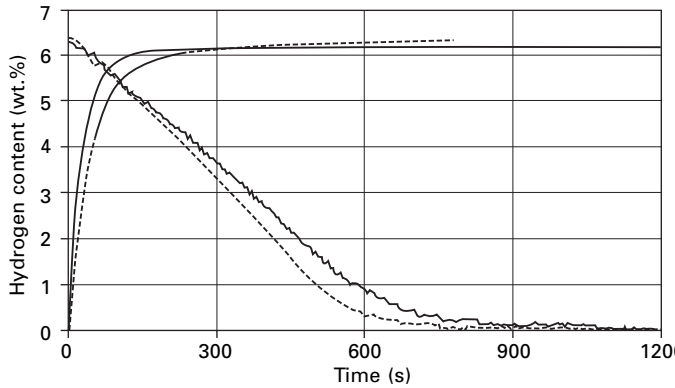
Using optimized parameters, the crystallite size refinement rate approximately corresponds to the one in the planetary ball mill [48], so that the product powders can be easily compared. The phase formation and phase composition is compared in Fig. 11.19. In both cases, the initial MgH_2 phase is the predominant phase and is refined at comparable rates with increasing milling time, documented by the broadening of peaks. The minimum saturation crystallite size is the same in both cases and is estimated to be ≈11 nm by the Scherrer method. Also, contamination by oxygen is comparable in both cases, resulting in the formation of MgO. The contamination with iron is too small in both cases to be resolved by X-ray diffraction. Slight differences between



11.19 X-ray diffraction patterns of MgH_2 as-received and milled with the planetary ball mill and the vibration mill.

the effects of the two types of mill can be observed with respect to the formation of the high temperature MgH_2 phase (γ-MgH_2). Although the overall fraction is very small, it is more significant in the material milled in the planetary ball mill, which may be due to the higher energy input. This can be compensated by the use of higher powder filling ratios in the vibration mill.

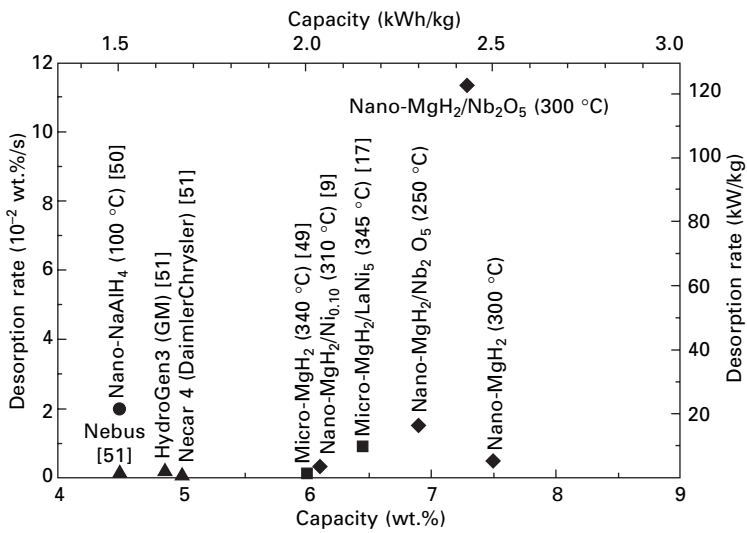
To evaluate the influence of the different amounts of the γ-phase, the adsorption and desorption curves are compared in Fig. 11.20. There is no significant difference in the total hydrogen storage capacity of the powder processed in the planetary ball mill (6.3 wt.%) and the vibration mill (6.2 wt.%) if milled with the same amount of oxide catalyst to about the same crystallite size of $\approx$ 11 nm. The slight differences may be due to differences in iron contamination, which are below the detection limit in the X-ray diffraction patterns. Kinetics are only slightly different. The absorption kinetics are a little faster for the material milled in the vibration mill, while desorption is slightly faster using the planetary ball mill, possibly because of the lack of γ-MgH_2 in the material milled in the vibration mill.



11.20 Hydrogen absorption and desorption curves of magnesium milled for 20 h by the planetary ball mill (dashed line) and the vibration mill (solid line).

11.6.2 Comparison of kinetic properties with requirements

With respect to both capacity and kinetics, the tanks with nanocrystalline MgH_2 are superior to current requirements with respect to kinetics, as indicated by the values for DaimlerChrysler Prototypes (Fig. 11.21), even at 250 °C.



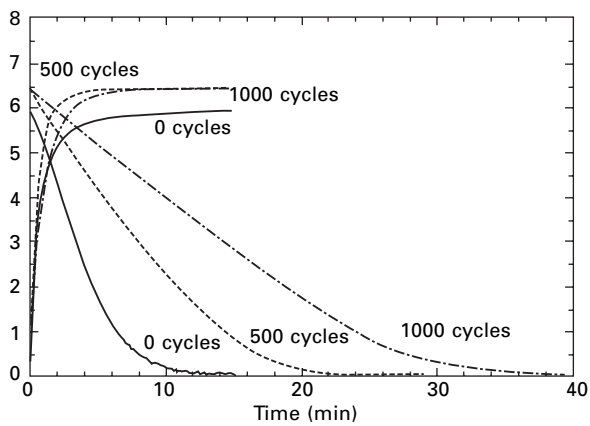
11.21 Comparison between literature data [9, 17, 49, 50], primary power data for the General Motors HydroGen 3 and the DaimlerChrysler Necar 4 (liquid hydrogen) and NEcBUS (compressed hydrogen) prototypes [51] and results of this work with respect to capacity and desorption rate.

This temperature is lower by about 100°C as compared to data for coarse-grained conventional material.

11.6.3 Long-term cycling stability

With respect to long-term cycling stability, the properties of Mg with oxide catalysts after 1000 cycles were also investigated [47]. For technical applications, a temperature of 300 °C is sufficient to get a hydrogen pressure of 1 bar. Therefore, long-term cycling experiments were conducted at 300 °C. Figure 11.22 shows corresponding dynamic hydrogen absorption and desorption curves, before and after 500 and 1000 cycles. Changes in the absorption behavior are only minor. The absorption rate after 1000 cycles amounts to 2.37 wt.% of hydrogen per minute, i.e., 47 kW/kg or 4.4 ltr-H_2/(s kg). In contrary to the absorption behavior, cycling has a stronger influence on the desorption characteristics of the composite. A systematic decrease in hydrogen desorption flow rate is observed after 500 and 1000 cycles. The desorption rate decreases from about 0.83 wt.% hydrogen per minute, i.e., 16.6 kW/kg or 1.5 ltr-H_2/(s kg), to about 0.23 wt.% hydrogen per minute, i.e., 4.5 kW/kg or 0.4 ltr-H_2/(s kg). The X-ray pattern of MgH_2-0.2 mol% Cr_2O_3 after 1000 cycles is characterized by significantly sharper peaks as compared to the as-milled material, indicating crystallite growth. The values of the crystal size estimated from the broadening of a single well-resolved peak of Mg using the Scherrer formula [53] yield values of 21 nm and 84 nm before and after cycling, respectively.

The results demonstrate that the nanostructured MgH_2-Cr_2O_3 composite has very good stability upon prolonged cycling at 300 °C. The absorption



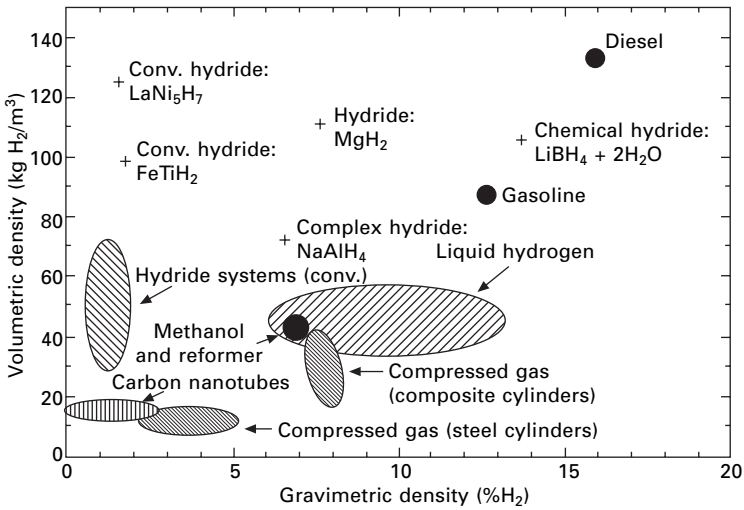
11.22 Absorption and desorption kinetics for nanocrystalline MgH_2 with 0.2 mol% Cr_2O_3 catalyst addition in the as-milled state and after 500 and 1000 cycles.

behavior after 1000 cycles remains almost as fast as in the first cycle after milling. The desorption rate decreases by a factor of 3 within 1000 cycles, however, still fulfilling the technical requirements. No capacity loss occurs. Instead, hydrogen storage capacity increases slightly by 8%, i.e., 0.5 wt.% hydrogen, to 6.45 wt.% hydrogen within the first 500 cycles. The decreased desorption rate as well as the increased capacity can both be explained by the microstructural coarsening. On the one hand, kinetics are slowing down successively with increasing crystallite size, ending up in very low sorption rates for coarse-grained materials. On the other hand, larger crystallite sizes mean that there is a higher volume of periodic atomic arrangement and less disordered crystallite boundary phase. The grain boundary regions are similar to amorphous phases, which store only small amounts of hydrogen at low pressures. On the contrary, crystalline regions reach their full storage capacity at 300 °C already at lower pressures of less than 2 bars, as observed in corresponding pcT diagrams. These results are impressive if compared to secondary batteries, in which a capacity loss of 20% is tolerated after 800 cycles. Regarding technical application though, one should keep in mind that UHP hydrogen (99.999%) was used in the present cycling tests. If lower purity hydrogen were to be used, the impurities would form MgO or $Mg(OH)_2$ with part of the material. These phases remain stable upon cycling and would thus reduce the overall storage capacity.

For technical applications, a temperature of 300 °C would be sufficient. At this temperature, the thermal stability as well as the cycling stability of the investigated material are excellent. Assuming a range of 500 km for a car with MgH_2-based storage tank, the lifetime of the storage material would exceed 500,000 km. The observed results compare well with the excellent results obtained for MgH_2–5at.% V composites published previously [54], where also enhanced storage capacity, no slow-down of the fast absorption kinetics and only a slight decrease in the desorption rate were observed.

11.6.4 Comparison with other storage alternatives

Some complex and chemical hydrides consisting of the lightest elements in the periodic table, for example lithium, boron, sodium and aluminium, form stable and ionic compounds with hydrogen contents of up to 18 wt.% for $LiBH_4$. However, in many cases for desorption very high temperatures are needed, and the reversibility of these reactions is not clear for all systems. Another possibility is the on-board generation of hydrogen by a chemical reaction. Light metal complex hydrides readily react exothermally with water to form oxides or hydroxides and hydrogen [55]. A comparison of the different storage methods concerning the volumetric and gravimetic hydrogen storage density is given in Fig. 11.23. Metal hydrides are superior regarding their volumetric energy densities, however, only light metal hydrides are competitive



11.23 Comparison of different hydrogen storage alternatives. System data are represented by the marked areas. Data for hydrogen storage capacities of pure storage materials are indicated by + signs.

regarding gravimetric storage capacities. The complete system masses are dependent on the size of the tank, i.e., the desired amount of stored hydrogen, and also on the dimensions of the heat exchanger. Efficiencies of the different storage alternatives are not taken into account. For example, the regeneration of chemical hydrides [56] requires more than 50% of the energy content, which makes this solution rather unlikely for a widespread application. Liquefaction consumes about one-third of the energy content, compression to 350 bars consumes about 18%. For reversible metal hydrides, efficiencies depend on the reaction enthalpies, i.e., on the bonding strength of hydrogen, and range from 77% for pure MgH_2 to more than 90% for the conventional room temperature hydrides $FeTiH_2$ and $LaNi_5H_7$. Here again, lower reaction enthalpies are favored.

11.6.5 Integration of hydrides for complete system solutions

Contrary to other storage alternatives, hydrogen extraction from a metal hydride is endothermic. The amount of heat supplied has to at least match the chemical enthalpy of the hydrogen bonds. Furthermore, the stronger the chemical enthalpy of the hydrogen bond the higher the temperature has to be so that the entropy can finally overcome the chemical enthalpy, supplying sufficient hydrogen pressure for the propulsion system, i.e., internal combustion engine or fuel cell and electric motor. For Mg, this leads to rather high temperatures of operation of 300 °C or more. At lower temperatures, a

compressor would be required. In principle, this is feasible. However, part of the system would be under vacuum leading to hazardous conditions in case of a leak.

Ideally, the required heat for desorption would be covered by waste heat from a combustion engine or the fuel cell system. For Mg hydride, fuel cells running at 300 °C or more would be required. Regarding fuel cells, this would necessitate solid oxide or molten carbonate fuel cells running at temperatures above 600 °C. These types of fuel cells are applied in larger stationary energy plants, however, they are currently not considered for automotive applications, because they would require a long start up time from cold, and they are prone to failure upon prolonged thermal cycling. Combustion engines would run in the right temperature range. However, modern optimized engines with high efficiencies provide less waste heat. Therefore, the available heat between the minimum temperature of 300 °C for Mg hydride and 400 °C, as obtained inside the combustion engine, may not be sufficient to maintain the decomposition reaction. Consequently, lower efficiencies would have to be tolerated for a combination with Mg hydride.

Combinations with less stable metal hydrides thus are more promising for mobile systems. For alanates, temperatures lie in the range between 100 °C and 150 °C, thus combinations with combustion engines are technically feasible. Further support for these medium-temperature hydrides may also originate from the new generation of polymer electrolyte (PEM) fuel cells, which are based on polybenzimidazol (PBI) [57]. These systems can be operated at temperatures up to 180 °C, thus providing sufficient waste heat to desorb hydrogen from hydrides at higher temperatures as compared to conventional Nafion®-based systems operating at 80 °C.

11.7 Future trends

As presented in the previous sections, nanocrystalline materials are promising for application as hydrogen storage materials. However, more work is necessary to meet the requirements as posted by the automotive manufactures. For Mg-based alloys, while kinetics are excellent, the temperatures of operation are still rather high, as already mentioned in the previous section. The variety and potential to find new Mg-based binary or ternary compounds with lower reaction enthalpies with hydrogen is rather limited, because Mg has negative enthalpy of mixing with only a few other metals. A number of respective compounds have already been investigated. While some of them show only small capacities by weight, others do not absorb any hydrogen at all. In addition, ternary phases tend to decompose into two binary phases upon cycling, which are more stable. There are basically three new potential ideas for the destabilization of Mg-H bonds in nanocrystalline Mg:

One new and promising approach is immiscible Mg-based systems, e.g., Mg-Fe [58]. Upon hydrogenation, Mg and Fe combine with hydrogen to form Mg_2FeH_6. Thus, the binding energy of the hydride is reduced by the repulsive interaction of Mg and Fe. Upon desorption, the enthalpy of decomposition of Mg and Fe is released, therefore, less heat is required for desorption, and the temperature of operation is effectively reduced. In addition, complex phase reactions may be exploited to obtain total reaction enthalpies in the range between 30 and 45 kJ/mole-H_2. The reaction of Mg_2Cu with hydrogen may be taken as an example:

$$2\ Mg_2Cu + 3\ H_2 \leftrightarrow 3\ MgH_2 + MgCu_2 \qquad 11.4$$

For absorption, the following partial reactions have to be considered:

$$2\ Mg_2Cu \rightarrow MgCu_2 + 3\ Mg \qquad 11.5$$

$$3\ Mg + 3\ H_2 \rightarrow 3MgH_2 \qquad 11.6$$

Here, reaction (11.5) is endothermic and thus counteracts the exothermic heat release for partial reaction (11.6). Therefore, the total reaction enthalpy of reaction (11.4) is effectively reduced, i.e., MgH_2 is destabilized by the phase transformation between Mg_2Cu and $MgCu_2$. Thereby, the temperature for 1 bar hydrogen pressure is also reduced. A number of other reactions between Mg-based intermetallics may be explored to resemble this reaction pattern.

Another promising magnesium-based hydride is magnesium alanate $Mg(AlH_4)_2$, which contains 7.9 wt.% hydrogen. First promising results have been presented, showing a partial hydrogen release already at rather low temperatures [59]. Similar to alkaline alanates, a decomposition reaction is observed upon dehydrogenation. Nanocrystalline microstructure and suitable promoter additions may facilitate reversibility, which has not been demonstrated yet. All of these approaches are very promising for nanocrystalline microstructures, because the diffusion of metallic atoms is required, and, therefore, diffusion distances need to be as short as possible, if reversibility is to be attained.

For alkaline alanates, the main future challenges are faster kinetics at even lower temperatures as well as reversibility of Li-rich or Li-based alanates. Thus, pressure and/or temperatures for refueling could be reduced and capacities could be increased. As pointed out above, these challenges can be met only by nanocrystalline microstructures to provide short diffusion distances for metallic atoms. In addition, favorable promoters have to be identified, which enhance bulk diffusivity of the metal atoms even at reduced temperatures. A comprehensive understanding of the currently known promoters and their function is necessary for further breakthroughs. Furthermore, there are other hydrogen containing materials and phase reactions, which may be utilized for reversible hydrogen storage. Recently, Cheng *et al.* showed that the

transformation of imides to amides and vice versa may store and release up to 2 wt.% hydrogen [60]. This approach is similar to the phase transformation reactions according to reaction (11.4), and demonstrates that unconventional materials may also be considered for reversible hydrogen storage. Recently, Orimo *et al*. [61] were able to further improve this kind of reaction by adding Mg, thus achieving full reversibility at 200 °C and 50 bars of hydrogen.

11.8 References

1. J.J. Reilly, G.D. Sandrock, *Scientific American*, 242 (1980) 5118.
2. R. Schulz, S. Boily, L. Zaluski, A. Zaluska, P. Tessier, J.O. Ström-Olsen, *Innovations in Metallic Materials* (1995) 529.
3. J.J. Reilly, R.H. Wiswall, *Inorg. Chem*. 13 (1974) 218.
4. B. Bogdanovic, M. Schwickardi, *J. Alloys Comp.* 253 (1997) 1.
5. T. Klassen, W. Oelerich, R. Bormann, V. Güther, *Deutsche Offenlegungsschrift DE* 197 58 384 A 1 (1999).
6. R. Schulz, J. Huot, G. Liang, S. Boily, G. Lalande, M.C. Denis, J.P. Dodelet, *Mat. Sci. Eng*. A267 (1999) 240.
7. J. Renner, H. Grabke, *Z. Metallk*. 69 (1978) 639.
8. J. Topler, H. Buchner, H. Saufferer, K. Knorr, W. Prandl, J. *Less-Common Metals* 88 (1982) 397.
9. M. Song, *Int. J. Hydrogen Energy* 20 (1995) 221.
10. A. Zaluska, L. Zaluski, J.O. Ström-Olsen, *J. Alloys and Compounds* 288 (1999) 217.
11. P.S. Rudman: *J. Less-Common Metals* Vol. 89 (1983) pp. 93 ff.
12. W.M. Mueller, J.P. Blackledge, G.G. Libowitz: *Metal Hydrides* (Academic Press, New York, London, 1968).
13. A.L. Reimann: *Phil. Mag*. (7) **16** (1933) 673–686.
14. L. Zaluski, A. Zaluska, P. Tessier, J.O. Strom-Olsen, R. Schulz, *J. Alloys Comp*. 217 (1995) 295.
15. M.Y. Song, *Int. J. Hydrogen Energy* 20 (1995) 221.
16. G. Liang, J. Huot, S. Boily, A. Van Neste, R. Schulz, *J. Alloy Comp*. 291 (1999) 295.
17. B. Tanguy, J.-L. Soubeyroux, M. Pezat, J. Portier, P. Hagenmuller, *Mater. Res. Bull*. 11 (1976) 1441.
18. G. Liang, J. Huot, S. Boily, A. Van Neste, R. Schulz, *J. Alloys Comp*. 292 (1999) 247.
19. G. Barkhordarian, T. Klassen, R. Bormann, *J. Alloys Comp*. 364 (2004) 242–246.
20. J. Huot, J.F. Pelletier, L.B. Lurio, M. Sutton, R. Schulz, *J. Alloys Comp*. 348 (2003) 319.
21. B. Fromme, d-d Excitations in Transition-Metal Oxides, A Spin-Polarized Electron Energy-Loss Spectroscopy (SPEELS) Study, *Springer Tracts in Modern Physics*, Volume 170, 2001.
22. G.M. Schwab, *Disc. Far. Soc*. 8 (1950) 166.
23. R.F. Klie , M.M. Disko, N.D. Browning, *Journal of Catalysis* 205 (2002) 1.
24. B. Hammer, *Phys. Rev. Lett*. 89 (2002) – 016102-1.
25. W. Oelerich, T. Klassen, R. Bormann, *J. Alloys Comp*. 322 (2001) L5.
26. W. Oelerich, T. Klassen, R. Bormann, *J. Alloys Comp*. 315 (2001), 237.
27. R. Schwarz, A. Khachaturyan, *Phys. Rev. Lett*. Vol. 74 (1995) pp. 2523 ff.

28. K. Zeng, T. Klassen, W. Oelerich, R. Bormann: *Int. J. Hydrogen Energy* Vol. 24 (1999) pp. 989 ff.
29. J.J. Reilly, R.H. Wiswall: Inorg. Chem. Vol. 7 (1968) pp. 2254 ff.
30. K. Zeng, T. Klassen, W. Oelerich, R. Bormann: *J. Alloys and Compounds* Vol. 283 (1999) pp. 213–224 ff.
31. T. Klassen, W. Oelerich, K. Zeng, R. Bormann: in *Magnesium Alloys and their Applications* ed. by B.L. Mordike and K.U. Kainer (Werkstoff-Informationsgesellschaft mbH, 1998) pp. 307 ff.
32. G. Liang, S. Boily, J. Huot, A. Van Neste, R. Schulz, *Materials Science Forum* Vol. 269–272 (1998) pp. 1049 ff.
33. P. Selvam, C.S. Viswanathan, C.S. Swamy, V. Srinivasan: *Int. J. Hydrogen Energy* Vol. 11 (1986) pp. 169 ff.
34. J.P. Darnaudary, B. Darriet, M. Pezat: *Int. J. Hydrogen Energy* Vol. 8 (1983) pp. 705 ff.
35. Borislav Bogdanovic and Manfred Schwickardi: *Journal of Alloy and Compounds* 253–254 (1997) 1–9.
36. A. Zaluska, L. Zaluski, J.O. Strom-Olsen, *J. Alloys Comp*. 298 (200) 125.
37. R.A. Zidan, S. Takara, A.G. Hee and C.M. Jensen, *J. Alloys Comp.*, 285 (1999) 119.
38. T. Flanagan, and W. Oates, Hydrogen in intermetallic compounds I (Springer 1988), *Topics in Applied Physics* 63, Chapter 3, 49–86.
39. B. Bogdanovic, R.A. Brandt, A. Marjanovic, M. Schwickardi, Joachim Tölle: *Journal of Alloys and Compounds* 302 (2000) 36–58.
40. L. Zaluski, A. Zaluska, J.O. Strom-Olsen, *J. Alloys Comp*. 217 (1995) 245.
41. R. Bohn, M. Oehring, T. Pfullmann, F. Appel, R. Bormann, in: *Properties and Processing of Nanocrystalline Materials*, eds C. Suryanarayana, J. Singh, F.H. Froes, P.A. TMS, Warrendale (1996) 355–365.
42. H. Heegn, C. Bernhardt, J. Gottschalk, K. Husemann, Aktivierungseffekte bei der mechanischen Bearbeitung von Eisenpulver in verschiedenen Labormühlen, *Chem. Techn*. 27 (1975) 12, 728–731.
43. H.-P. Heegn, *Synthese intermetallischer Phasen im System TiAl durch aktivierende Mahlung*, Abschlußbericht zum BMBF-Projekt 422-4003, Teilvorhaben 03M 3053 F, 1997.
44. W. Schlump, J. Willbrand, *Berichte zum BMFT-Verbundvorhaben: Grundlagen des mechanischen Legierens und des Intensivmahlens zur Herstellung metastabiler Legierungen*, Essen 1994.
45. G. Eberhard and K.-E. Kurrer, Eccentric vibratory Mills – Theory and practice, *Powder Technology*, Volume 105, Issues 1–3, 1 November (1999) pp. 302–310.
46. L. Rolf, E. Gock, Untersuchungen zur Optimierung der Schwingmahlung, *Chemie-Anlagen-Verfahren* (1976) pp. 27–31.
47. K.-E. Kurrer, E. Gock, Einfluß der Betriebsbedingungen auf die Bewegungs-und Stoßvorgänge in Rohrschwingmühlen, *Aufbereitungstechnik* 10 (1986) pp. 546–554.
48. T. Klassen, W. Oelerich, R. Bormann, Nanocrystalline Mg-based Hydrides: Hydrogen Storage for the Zero-Emission Vehicle, *Materials Science Forum* 360–362 (2001) 603–608, *Journal of Metastable and Nanocrystalline Mat*. 10 (2001) 603–608.
49. B. Vigeholm, J. Kjoller, B. Larsen, A.S. Pedersen, *J. Less-Common Metals* 89 (1983) 135.
50. M. Fichtner, J. Engel, O. Fuhr, O. Kirchner, O. Rubner, *Mat. Sci. Eng. B* 108 (2004) 42.
51. DaimlerChrysler, General Motors, private communication.

52. Z. Dehouche, T. Klassen, W. Oelerich, J. Goyette, T.K. Bose, R. Schulz, Cycling and Thermal Stability of Nanostructured MgH_2-Cr_2O_3 Composite for Hydrogen Storage *Journal of Alloys and Compounds* 347 (2002) 319–323.
53. C. Suryanarayana, M.G. Norton, *X-Ray Diffraction: A Practical Approach*, ISBN: 030645744X, Perseus Publishing (1998).
54. Z. Dehouche, J. Goyette, T.K. Bose, J. Huot, S. Boily, R. Schulz, *Proceedings of the International Hydrogen Energy Forum 2000*, Vol. II, 11–15 September 2000, Munich, Germany, pp. 99–107.
55. H.I. Schlesinger, H.C. Brown, A.E. Finholt, I.R. Gilbreath, H.R. Hoekstra, E.K. Hyde, *J. Amer. Chem. Soc.* 75 (1953) 215.
56. US Patent 6,534,033 B1.
57. http://www.pemeas.com/Celtec.e.htm
58. J. Huot, S. Boily, E. Akiba, R. Schulz, *J. Alloys Compounds* 280 (1998) 306.
59. M. Fichtner, O. Fuhr, O. Kircher, *J. Alloys and Compounds* 356–357 (2003) 418.
60. P. Cheng, Z. Xiong, J. Luo, J. Lin, K.L. Tan, *Nature* **420** (2002) 302.
61. Y. Nakamori, G. Kitahara, K. Miwa, S. Towata, S. Orimo, *Appl. Phys. A* (2004) DOI: 10.1007/s00339-004-3002-6.

12
Nanofabrication

E HARVEY and M GHANTASALA,
MiniFAB, Australia

12.1 Introduction

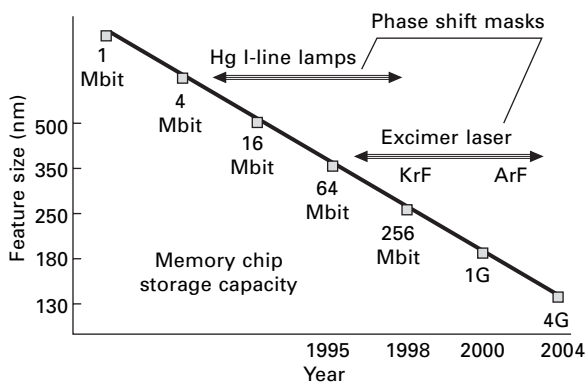
The term nanofabrication implies the creation of artefacts whose scale is in the nano domain. Fabrication is a term borrowed from macro engineering in which bridges, ships and complex structures are designed, materials of the correct properties are pre-shaped and the subcomponents assembled to create a finished piece. The microelectronics industry uses fabrication to describe the creation of complex, highly integrated circuits. Again this is a process of choosing materials of the desired properties, depositing them and patterning them in a sequence of steps designed to create an integrated circuit. Fabrication therefore describes a sequence of processes and operations that are performed in order to fashion discrete artefacts or devices from raw materials. In this chapter we use this definition in order to help us focus on engineering processes associated with nanofabrication. Molecular chemistry and biomole-cular growth may also be considered methods for creating individual components on the nano scale, however, since these are topics described by their own disciplines we choose not to cover them in depth here. A fascinating future may be one in which the engineering fabrication processes are hybridized with chemistry and biology to create new methods for fabrication, however much remains to be achieved before these dreams become reality.

In our opinion, the advanced microelectronics sub-micron technologies have sown the seeds required for the birth of nanotechnology. So process intensive is nanoscale microelectronic fabrication, that some have suggested that microtechnology should be considered a collection of processing techniques (also applicable at the nanoscale) whereas nanotechnology is principally about the creation of materials whose properties are dominated by the dimensions of the subunits. In any case, the first-generation nanofabrication technologies have their roots from within the microfabrication technology developed for further miniaturization of the then existing microelectronic circuits. Gordon Moore (Moore, 1965) predicted a logarithmic trend in the reduction of feature sizes and increase in device densities with successive

generations of integrated circuits – a prediction that has accurately described trends for the past 40 years. The graph in Fig. 12.1 clearly shows that the feature sizes have touched the 0.09 μm (90 nm) mark as predicted by the curve, and in terms of transistors per chip, current technology trends exceeded the predicted values (Bohr, 2002). The fabrication of superlattices and patterning of sub-micron interconnects and associated active and passive components with an ever increasing component/device density may be quoted as excellent examples of the nanotechnology at its infancy. The maturity in the understanding and further development of this technology in the past decade has led to the distinct classification of the processing approaches as top-down or bottom-up.

12.2 Nanomaterials – top-down and bottom-up approaches

Nanofabrication technology is often conveniently classified on the basis of two basic approaches: top-down or bottom-up. Top-down approach is exemplified by conventional microelectronic processing technologies, while bottom-up approach is uniquely associated with the rapidly growing suite of alternative nanofabrication methods. Top-down approach has its origins in the thoughts of Richard Feynmann back in the 1950s, according to whom, larger machines are used to build smaller ones and the small ones will be useful in building tiny machines and so on. Modern microlithography uses a different top-down approach based on techniques such as photolithography, and other additive and subtractive processing technologies. The top-down methods start with the use of bulk materials or a set of thick/thin layers, which are patterned to construct the required nanostructures. In this approach, we start with a larger structure (silicon wafer) and with patterning and etching or



12.1 Microelectronics: feature size and memory chip capacity, and the light sources used for photolithography.

machining small features with different nanostructures are created. A remarkable variety of nanoelectronic components and devices are fabricated using this approach. A great advantage of this approach is its use of batch processing, fabrication of multiple structures in a single step, and the ability to create ever complex patterns simply by changing the image used to photoexpose light-sensitive layers. However, the sizes of the features and structures that can be fabricated using this method are limited by the capability and resolution of the tools used.

In the bottom-up approach, the atoms and molecules are assembled one at a time in a predetermined manner. This means bottom-up manufacturing provides components made of single atoms or molecules, which are held together by covalent bonds that are far stronger than the forces experienced in macroscale components. This process takes advantage of physical or chemical interactions to assist in the ordering of the components to create complex structures, and these processes can usually also be used in batch mode.

Although these two methods are distinctly different in their approaches, often the formed structures or components are similar. Interestingly, recent trends show that a combination of these two approaches can also be used for the fabrication of molecular nanostructures. Further, these two approaches mainly use the conventional or a modified version of additive and subtractive fabrication methods. Hence, it is important to understand some of the relevant additive or subtractive fabrication technologies outlined below.

12.3 Fabrication technologies – additive/subtractive methods

In the macro world, while the painting of a wall can be taken as an example of an additive process, sculpting in wood or stone is considered as a simple subtractive process. In a microscopic context, deposition or addition of a thin layer on a selected substrate is an additive process, whereas selective etching of a layer or bulk micromachining of a particular section of the substrate are said to be good examples of subtractive processes. Traditional microlithography employs a combination of both these techniques in patterning a desired micro- or nanostructure, following the top-down approach. However, the major limitation of optical lithography is the wavelength of the source, which determines the minimum feature size achievable. The minimum feature size (MFS) obtainable using a particular exposure tool is given by

$$\text{MFS (resolution)} = (\lambda/\text{NA})\ K \tag{12.1}$$

K is a constant depending on the contrast and other properties of the photosensitive resist used. λ is the wavelength of the radiation, NA represents the numerical aperture, which is defined as the sine of the capture angle of the lens, and is a measure of the size of the lens system.

Furthermore, depending on the complexity of the pattern, the number of exposures required in a multilayered pattern can also affect the feature or edge sharpness and therefore the overall feature size. Typically, in the production of a complex modern integrated circuit, lithographic exposures are carried out at least 24 to 30 times, with each exposure requiring multiple processing steps such as spinning, prebake, exposure, postbake, development and final removal of the resist, which means approximately 150 to 200 processing steps. This results in a significantly high manufacturing cost, taking into account the additional cost of all sophisticated tools and the clean room environment required for using them. In addition, further reduction in feature sizes at the nm level necessitates the use of phase shifting masks and lasers of much lower wavelengths (e.g. Excimer laser types Krypton Fluoride (KrF) – 248 nm, Argon Fluoride (ArF) – 193 nm or Fluorine (F_2) – 157 nm) or extreme ultraviolet based advanced lithographic tools (including X-ray lithography) of the next generation.

In contrast, a pure additive process based on the bottom-up approach seems to provide a better, more viable alternative using many novel nanostructured materials. A number of additive processes such as self-assembly, soft lithography and contact printing for example, can potentially revolutionize the nanofabrication technologies. Nanofabrication technologies currently employed in the manufacturing of a variety of nanostructures may mainly be classified as lithography oriented and additive-based technologies.

12.4 Lithography-based technologies

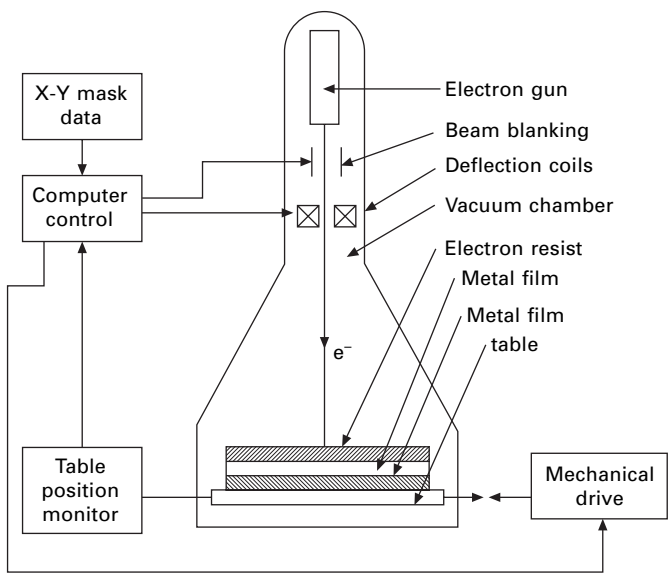
Lithography-based techniques are classified according to the tools used and the method of image transfer. These are mainly either direct write type or transfer of a pattern through a photo mask using conventional exposure and development routes.

12.4.1 Electron beam lithography

Electron beam lithography (EBL) is frequently used in the indirect exposure and pattern transfer mode as well as the direct write mode (http://dot.che.gatech.edu/henderson/introduction_to_electron_beam_lithography. htm). In this technique, a highly focused electron beam is used to write the fine line patterns on semiconductor substrates covered by a resist material. The direct write mode is often used for making photolithographic masks. As the beam energy is increased, it provides a beam at a much shorter wavelength, which helps in improving the resolution. However, the resolution or feature size that is achievable using this e-beam process is limited to around few nms, due to the resist limitations rather than the wavelength of the radiation. Though it is possible to focus the e-beam onto the sample with a resolution

of approximately 0.5 nm, the resist itself limits the final resolution achievable to approximately 5 nm. Further, an electron beam has such a short wavelength that diffraction no longer defines the lithographic resolution. In this lithographic process, the e-beam is focused onto the substrate coated with resist in a high vacuum environment (10^{-5} torr). The electron beam is moved across the surface under computer control using a pattern generator. After completely defining the pattern in the resist onto the substrate, it is removed from the system and subsequently developed following the conventional lithographic procedures. A typical e-beam system is shown in Fig. 12.2. These systems are basically scanning electron microscopes designed for the extreme precision and stability required to write large and complex patterns with nanometer linewidths.

A major commercial application of EBL systems is photomask production. These masks are made on a quartz substrate, which provides ideal transmission characteristics in the ultraviolet region, a wavelength range normally used in photolithography. Masks are made by depositing a chrome layer on the quartz glass plate. An electron beam sensitive resist layer coated on the top will subsequently be exposed and developed to generate the required pattern. The chrome layer is then selectively etched in the patterned area to provide the mask. The ability to meet stringent line width control and pattern placement makes this technique ideal for mask making. The advanced prototyping of integrated circuits and manufacture of small-volume and high-resolution products such as gallium arsenide integrated circuits and optical waveguides are some of the other important applications.



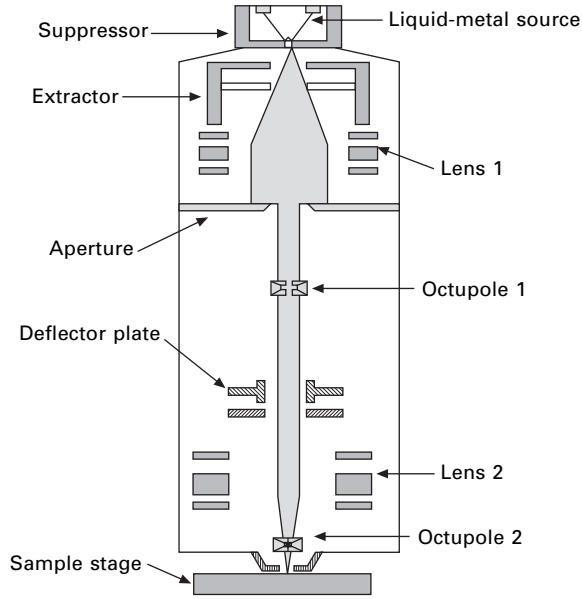
12.2 Schematic of the electron beam lithographic system.

Recently, projection and proximity type EBL systems have also been developed (Heritage, 1975; Frosien, 1979; Nakasuji *et al*., 1993). As the electrons have short penetration length, solid substrates like quartz cannot be used for the mask. This necessitates a very thin membrane mask or a stencil mask with cutouts through which the beam can pass. While these difficulties make this form of EBL unattractive, the length of time required to write the pattern in a direct write system makes it relatively expensive and commercially unviable. Another problem with this technique is the proximity effects caused due to the large degree of elastic scattering of the high-energy electrons in the resist. This tends to blur the pattern as it affects a much larger region of the resist area than the exposed feature size. Though it could be possible to correct such effects, it can significantly add to the increased costs via higher computer times, etc. Research into more advanced e-beam projection lithographic systems, which can overcome some of the difficulties mentioned above, is in progress at AT&T labs and others (Tennant, 1998; Berger *et al*., 1994).

12.4.2 Focused ion beam lithography

Ion beams have long been used for surface modification of materials and thin films by either doping or etching. Ion beams are better suited for nanofabrication though these are not considered as a fabrication tool for mass production of nano devices. One of the major advantages with focused ion beams is that the ions suffer very little back scattering compared to that of electrons, which is one of the limitations with e-beam lithography that essentially limited the minimum feature size possible.

Focused ion beam (FIB) lithography uses positively charged ions accelerated to energies in the range of a few keV to around 75 keV, before reaching the sample to etch. The most commonly used ion species is Gallium due to its long lifetime of up to 1500 hours (http://www.calpoly.edu/~sdygert/IonBeam.-html) as below. The FIB column is similar to the electron beam column used in a scanning electron microscope or electron beam lithographic system, except for the polarity of the voltages applied to the accelerating and focusing electrodes. The schematic of a focused ion beam system is shown in Fig. 12.3. Features sizes of the order nm are patterned using this technique by narrowly focusing the beam at the same time scanning it on the surface, depending on the required pattern. Beam focusing is always a challenge due to the fact that all the ions having the same charge tend to spread out. This limits the beam diameter to a few nm. However, the advantage with using ions is that they provide much more efficient energy transfer due to their heavier mass. As the ions have a tendency to completely disintegrate resists, these are not commonly used in the etching or milling of surfaces or patterns using this technique.



12.3 Schematic of a focused ion beam system.

The drive to push the feature sizes from the sub μm regime to nm regime necessitated the development of alternative means to conventional optical lithography in the fabrication of complex integrated circuits. Focused ion beam technology appears to be one of the promising candidates, as ion beams are inherently free from the limitations of optical interference. IMS nanofabrication is an example in this direction, with the development of a new technology, 'Projection Focused Ion multi-Beam tools' (PROFIB). (Fantner and Loeschner, 2003). Nano PROFIB operating in a high-resolution mode is expected to resolve less than 10 nm features within a 20 μm exposure field using stencil mask openings of 1 μm. The most exciting aspect of this project is that it facilitates the exposure of 1 million 10 nm dots in less than 0.5 sec even for a dose of 10^{16} ions/cm^2 in a single shot. The tool can be used in lithography mode as well as for direct milling or structuring of the substrate itself, which means it can be used for ion milling and ion beam assisted deposition as well as ion beam modification or implantation in one instrument.

While the application of FIB technology in nanoelectronics is still some time away, ultra-high-density magnetic recording technology seems to have benefited from using FIB in the fabrication of different components. A number of longitudinal and perpendicular nanoscale magnetic storage devices were recently fabricated using FIB technology (Khizroev *et al.*, 2002, 2004). Fabrication of prototype devices having ordered arrays of metallic nanoclusters using a set of techniques including FIB has also been reported (Weller *et al.*, 2003). One of the major advantages with FIB compared to an electron beam,

especially in patterning nanomagnetic devices, is that the electron beam tends to show physical shifting to an extent of 100 nm due to the interaction between the beam of charged electrons and the magnetic thin film on the substrate to be patterned, which is almost negligible with an ion beam. Apart from these, the heavier mass of the ions allows direct etch or deposition, thus avoiding the tedious patterning steps used in optical or electron beam lithography.

12.4.3 Laser-assisted lithography

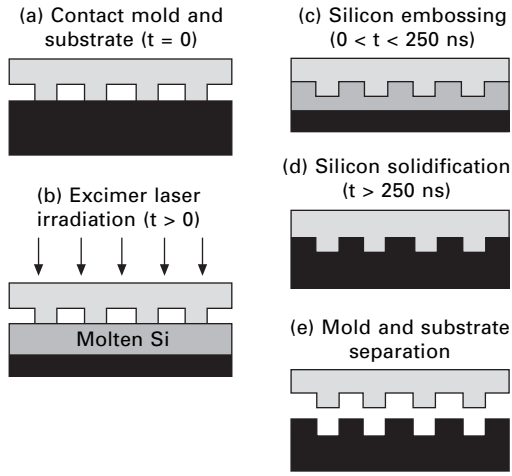
Laser-assisted lithography has been moving to the forefront of the advanced lithographic technologies, where the reduction of feature size accompanied by enhancing the resolution has become the major goal. As shown by eqn 12.1, the minimum feature size can be reduced by using radiation of lower and lower wavelengths or by increasing the numerical aperture (NA). However, higher NA optics necessitate the use of an exposure light source that produces a very narrow spectral radiation, without which excessive optical aberrations would be produced. Excimer lasers have recently become popular as lithography tools especially for producing the features in the tens of nms region. KrF-based 248 nm microlithographic scanner systems have already been extensively employed in the semiconductor industry. However, 193 nm ArF laser systems with super high numerical aperture systems are being developed (Paetzel *et al*., private communication, undated).

Future development in this area points towards the development of extreme ultraviolet lithography, through the reduction of feature sizes by another order. The wavelength of the light required for EUV lithography will be approximately 10 to 20 times shorter than that of excimer lasers, which may again depend on laser excitation or gas discharge plasma excitation.

Another interesting and a more recent technology based on lasers is the 'laser assisted direct imprint technique' (Chou *et al*., 2002). Though it is not a direct lithography technique, the capabilities of the technique in imprinting patterns with a minimum feature size of 10 nm are quite impressive. In this technique, an excimer (XeCl) laser pulse (308 nm with 20 ns pulse duration) passes through a quartz mold and melts a thin surface layer of silicon substrate within picoseconds. This molten silicon layer (typical thickness of around 300 nm) is then embossed by the quartz mold. A typical process schematic along with a 300 nm period silicon grating fabricated using this process is shown in Fig. 12.4.

12.4.4 Interferometric lithography

Deep and extreme ultraviolet radiation-based interferometric lithography techniques were developed in order to realize sub-100 nm structures for



12.4 Laser assisted direct imprint technique process schematic.

various applications. These have been extensively used in the characterization of different varieties of photoresists, for example, chemically amplified resists (Fig. 12.5). The interferometric lithography system comprises two coherent light sources. The light beams from these sources are placed at an angle of 2θ to produce a standing wave pattern on the substrate. The period of the sinusoidal wave pattern is given by

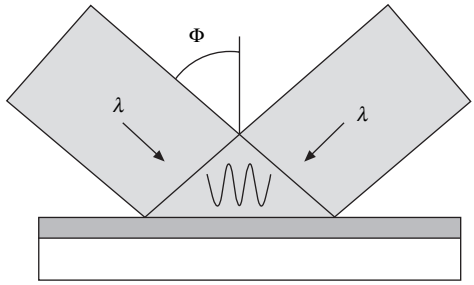
$$D = \lambda/2 \sin\theta \qquad 12.2$$

where λ is the exposing wavelength, and θ is the angle of incidence of the beams from the normal. This means the minimum period of the wave pattern that can be achieved is $\lambda/2$. This may be considered as the best resolution achievable for any lithographic system with a source that provides a light of wavelength λ. For example, using light beams with a wavelength of 248 nm, a line/space array of period of 125 nm is calculated from equation 2, taking an incident angle of 160°. This means the period of the pattern can easily be varied by changing the incident angle. A typical pattern thus obtained by Hinsberg *et al*. is shown in Fig. 12.6, with a light beam of $\lambda = 257$ nm, while exposing a Shipley UVII-HS positive tone chemically amplified photoresist.

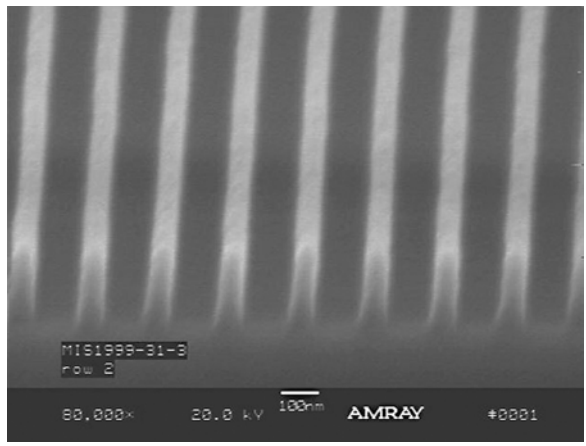
The modulation of the intensity is given by the fringe visibility

$$\mu = (I_{max} - I_{min})/(I_{max} + I_{min}) \qquad 12.3$$

where I_{max} and I_{min} are the maximum and minimum light intensities in the standing wave pattern. Fringe visibility is mainly dependent on the degree of coherence between the two beams and their relative strength (Solak, 1999). μ equals 1, when $I_{min} = 0$, which occurs only when there is completely destructive interference. This is possible only when both the beams are perfectly coherent and have equal strength.



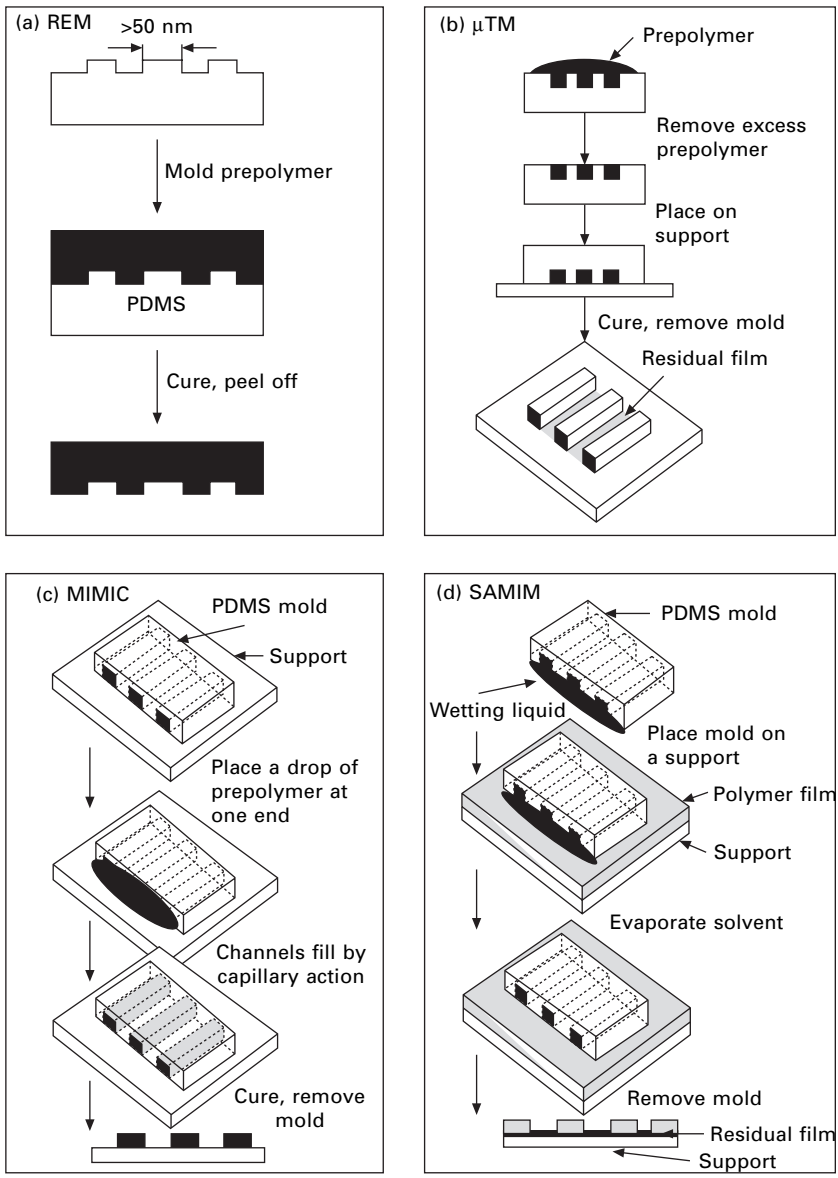
(a)



(b)

12.5 (a) Optical arrangement of IL. (b) SEM of a resist image using 257 nm beam (Hinsberg *et al.,* 1998).

The biggest advantage with this technique is that it is a mask-less process, which minimizes processing costs. Further, the depth of field is so large, that it can be considered as infinite for all practical purposes. Solak *et al*. have experimented using extreme UV light from a synchrotron radiation source and they were able to demonstrate a line resolution of 19 nm (with a pitch of 37 nm), which is considered to be one of the best obtainable values reported in the literature using optical lithography. Indeed, with a λ of 13.4 nm it is possible, in principle, to pattern features smaller than 4 nm. Fabrication of a nanofluidic chip and other prototype devices were reported recently (Brien II *et al*., 2003). A combination of traditional optical lithography along with interferometric lithography can yield a complete range of features from mm to nm on a single device. Holographic lithography based on similar principles has also been developed for achieving the sub-100 nm feature sizes (Korda *et al*., 2002). Hence, these techniques appear to be significant developments and may be expected to play an increasingly prominent role in nanofabrication technology.



12.6 Soft lithographic techniques. (a) Replication molding (b) microtransfer molding (c) micromolding in capillaries (d) solvent assisted micromolding.

12.4.5 Soft lithography

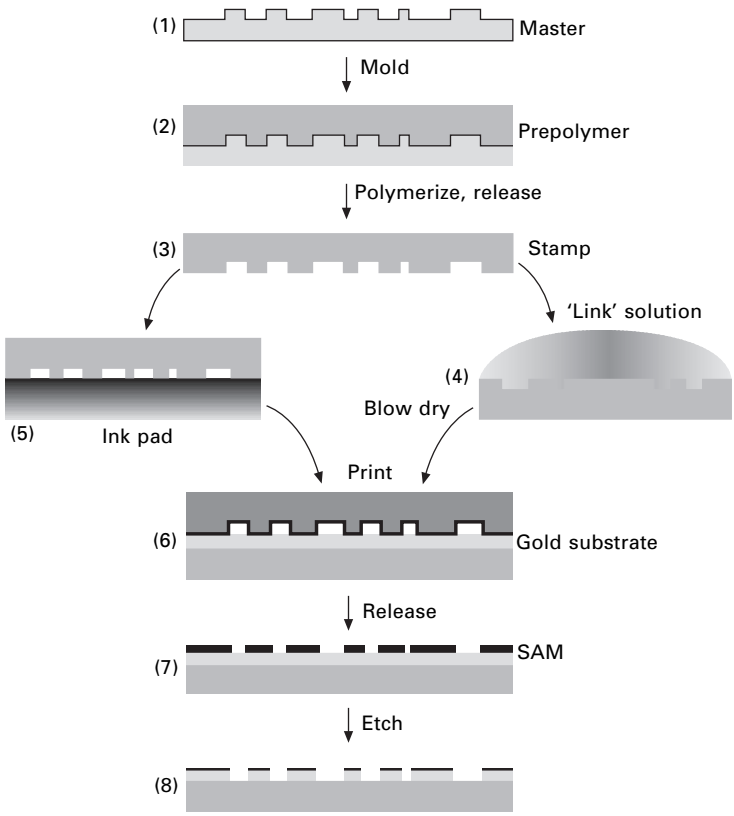
Soft lithography is a new high-resolution patterning technology, first proposed by George Whitesides at Harvard University to categorize the technique

which used a soft elastomeric stamp or mould (usually poly(dimethylsiloxane) known as PDMS) that transfers the pattern to a substrate (Kumar *et al*., 1993). Soft lithography is a broader term that covers a wide range of techniques, namely, microcontact printing, replica molding, microtransfer molding, micromolding in capillaries, solvent assisted micromolding, phaseshift photolithography, cast molding, embossing and injection molding (Xia and Whitesides 1998a, b). The major advantages of the soft lithographic techniques are the ease of fabrication and their biocompatibility. Further, the softness of the materials used in the process allows the fabrication of features two orders of magnitude smaller than those fabricated by conventional silicon micromachining and lithographic technologies. More importantly these are not limited by diffraction effects as commonly seen in conventional photolithography, which extends the attainable minimum feature sizes below 30 nm level. However, it may be noted that the master used in all the soft lithographic techniques needs to be produced using EBL or any other high-resolution photolithographic technique.

The first of the soft lithographic techniques to be discussed is microcontact printing (μCP). This printing differs from other printing technologies mainly in the use of self-assembled monolayers (SAMs) to form the required micro/nanostructures of different materials. Self-assembly is the spontaneous organization of molecules or objects into stable, well-defined structures by non-covalent forces. SAMs are good examples of non-biological self-assembling systems. Formation of double helix DNA structures, folding of proteins and tRNAs can be said to be best examples of self-assembled structures in a biological domain. Self-assembled monolayers readily form on the appropriate substrates by chemisorption and self-organization of functionalized long-chain organic molecules. These are normally prepared by immersing a substrate in the solution containing a ligand that is reactive towards the surface or by exposing the substrate to the vapor of the reactive species. A list of combinations of ligand or precursor along with the substrate, known to give SAMs, was provided by Xia and Whitesides (Xia and Whitesides, 1998a).

Microcontact printing

Microcontact printing (Fig. 12.7) relies on the replication of a patterned elastomeric stamp, usually made in PDMS or silicone rubbers, from a master to form an elastic stamp, that can be inked with a monolayer-forming ink using either wet inking or contact inking. This inked stamp is then used to transfer molecules of the ink to the surface of the selected substrate by contact printing, which protects the noble metal substrate surface during the subsequent etching process. Stamps are usually formed by pouring liquid PDMS on the master and curing it at temperatures between 20 °C and 80 °C



12.7 Microcontact printing process (michel *et al.* 2001).

for up to 48 hours. Accurate replication of features smaller than 500 nm however, necessitate the use of harder stamp materials (Schmid *et al*., 2000). A detailed study on the analysis of stability and integrity of topographic structures on the stamps was studied by various researchers and for more comprehensive information one can refer to the review by Xia and Whitesides (Xia and Whitesides, 1998a, b). Microcontact printing was first demonstrated for SAMS of alkanethiolates on gold (Kumar *et al*., 1993). The advantage with alkanethiols is that they self-assemble on noble metal surfaces such as Au, Ag, and Cu to form dense ordered monolayers. These monolayers allow control over wettability, adhesion, chemical reactivity, electrical conduction and mass transport to the underlying metal (Michel *et al*., 2001). These are later extended to many other material–substrate combinations.

Replica molding

The second method in this cluster of soft lithographic techniques is replication molding, which is a simple manufacturing concept involving the duplication

of the structures providing a faithful reproduction of complex microstructures present on the surface of a mold with nanometer resolution. The three major steps involved in this process are represented in Fig. 12.6(a). The PDMS molds are prepared by casting using a master as described in the microcontact printing process. The use of such a flexible mold, usually a thermoplastic polymer, enables the release of minute and fragile structures in a simple and yet reliable and inexpensive way. The reproduced structures in the cured polymers (for example, polyurethane (PU)) possess almost the same dimensions and topologies as the features in the PDMS mold. In a nanometer scale replication, it was shown that 13 nm thick chromium lines in the master were duplicated to the heights of 8 nm in PU.

The analysis of the quality of a typical molded structure replicated from a single master or a mold showed that these can be repeatedly used up to ten times without any visible damage to the master or any degradation in the quality of the PU replicas. This technique can also be used to produce nanostructures having feature sizes and shapes different from the mold by bending, compressing or stretching the PDMS mold during replication in a controlled manner. However, the extent of change in the replicated features is limited by the design constraints and the flexibility of the mold. It has been shown that the dimensions of the features can easily be reduced from 50 to 30 nm, while replicating the features from a PDMS mold onto a PU surface (Xia *et al*., 1997).

Microtransfer molding (μTM)

Microtransfer molding (μTM) is useful to produce patterned microstructures on a variety of polymers over relatively larger areas within a short period of time. In this process, as shown in Fig. 12.6(b), a drop of a liquid prepolymer is applied and allowed to form a thin layer on the surface of a patterned PDMS mold. The excess liquid on the top is removed by scraping with a flat PDMS block or by blowing off with a stream of nitrogen. The filled mold is then placed in contact with the required substrate and is cured by either UV illumination or heating depending on the polymer. After the completion of the curing process, the mold is peeled off carefully, leaving a patterned microstructure on the surface of the substrate. The usefulness of this process has been demonstrated in the fabrication of optical couplers, waveguides, and micromagnetic systems for various applications (Zhao *et al*., 1996; Deng *et al*., 2001; 2002).

Micromolding in capillaries (MIMIC)

Micromolding in capillaries is another variation of the microtransfer molding process mainly differing in the way the prepolymer filling of mold structures

is carried out. In this process, a PDMS mold is placed on the surface of a substrate such that it makes a conformal contact. When a low-viscosity fluid is injected into the channel through the open end of a network of channels, the liquid spontaneously fills the channels by capillary action. The PDMS mold is removed after curing the prepolymer, leaving a network of polymeric structures on the substrate surface, as shown in Fig. 12.6(c). This imposes constraints on the type of features that can be replicated – it prevents the replication of unconnected or isolated smaller features as the method relies on the flowing of liquids within the capillaries of the mold. The rate of capillary filling is determined by the surface tension (ν) and viscosity (η) of the liquid, the radius of the capillary[R] and the length of the filled section of the capillary, z, as indicated by the following equation:

$$dz/dt = R\nu_{sl} \cos\theta/4\ \eta z = R(\nu_{sv} - \nu_{sl})/4\eta z \qquad 12.4$$

This equation shows that the rate of filling of the capillaries is a function of their length as well as the width and depth of the channels. The rate of filling decreases in relatively longer (>1 cm) channels due to viscous drag of the fluid in the capillary. However, the strong dependence of the rate of filling of the capillaries on their dimensions severely limits their application with different prepolymers below 50 nm diameters. Although the MIMIC process was initially developed for use with a prepolymer having no solvents, it has been later extended to different polymers with solvents. The solvents are evaporated after the liquids have filled the channels.

Solvent-assisted micromolding

The last of the soft lithographic techniques to be discussed in this chapter is the solvent-assisted micromolding (SAMIM), technique which can be used to fabricate quasi three-dimensional microstructures on the surface of polymeric substrates. A PDMS mold is wetted with a solvent (for a polymer), and is brought into contact with the surface of the polymer as shown in Fig. 12.6(d). The solvent dissolves (or swells) a thin layer of the substrate (typically an organic polymer) and the resulting fluid or gel is molded against the structures in the mold. As the solvent dissipates and evaporates, the fluid solidifies and forms a microstructure, which is complementary to that of the mold. This shares the principles of embossing except for the fact that softening of the polymer is achieved through the use of a solvent rather than temperature. Another major difference between the two techniques is that SAMIM uses an elastomeric PDMS mold compared to the rigid (often metal) master used in embossing. In general, the solvent used in this case should have a high vapor pressure with a reasonable value of surface tension to facilitate the quicker evaporation of the excess solvent and causing minimal swelling of the PDMS mold. The key to the success of this process entirely depends on

the selection of solvent. However, this technique can replicate patterns of complex topology in a single step.

Soft lithography has become crucial in the realization of nanostructures for various applications ranging from optical components to a different kind of nanotechnology-based sensors. The progress realized in this area can be attributed mainly to these fabrication technologies, especially in the context of developing cheaper yet reliable methods. However, it has become difficult to accommodate an in-depth discussion on each of the soft lithographic techniques here. A detailed discussion on soft lithography can be seen in the review papers of this subject (Xia and Whitesides, 1998a, b).

As seen so far, lithography-based fabrication technologies have been the key for the successful production of complex nanostructures. In this context, additive deposition technologies have also played an equally important role in micro and nanofabrication. In an effort, to produce a balanced view of these technologies, some of the important deposition technologies, which contributed towards the realization of nanosensors and actuators will be discussed further.

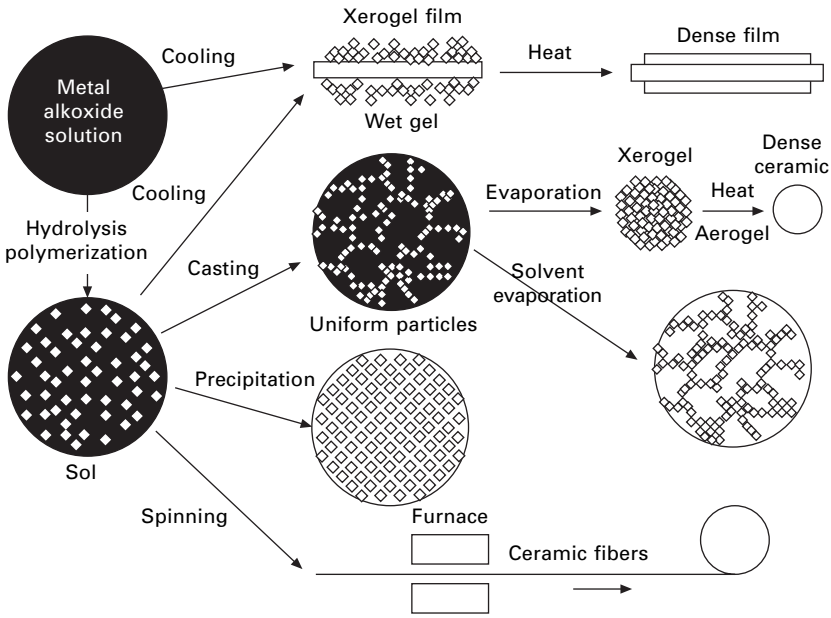
12.5 Deposition techniques

The key deposition techniques which have been frequently used in nanofabrication are the sol-gel, atomic layer deposition, electrochemical deposition and chemical vapor deposition techniques. Sol-gel and atomic layer deposition were used for the fabrication of porous nanostructures for sensing applications whereas atomic layer deposition was handy in realizing highly ordered, defect-free layered structures. A variation of conventional electroplating technologies was proven to be useful in the fabrication of nanowires and nanorods by selective electrochemical deposition. Chemical vapor deposition has extensively been employed in the fabrication of carbon nanotubes.

12.5.1 Sol-gel deposition

Sol-gel is one of the simplest of a wide range of deposition processes currently used for the fabrication of nanoceramic and glass coatings. This process can produce ultra fine or spherical-shaped powder to inorganic membranes containing micropores. The sol-gel process involves the transition of a system from a liquid ‘sol’ (mainly colloidal) into a solid ‘gel’ phase. The schematics in Fig. 12.8 show the possible routes of manufacturing using sol-gel for different applications (www.chemat.com).

In the sol-gel process, the precursors (starting materials) used in the preparation of the ‘sol’ are usually inorganic metal salts or metal organic compounds such as metal alkoxides. Metal alkoxides consist of a metal or metalloid atoms surrounded by an organic ligand. For example, titanium



12.8 Sol-gel process and possible product applications.

dioxide is prepared using the precursor titanium butoxide (Ti $(OC_4H_9)_4$). Typically, the precursor is subjected to a series of hydrolysis and polymerization reactions to form the colloidal suspension 'sol'. Further processing of the 'sol' is required to produce materials in different forms. A 'gel' is formed by either casting the 'sol' into a mold or by depositing it onto a substrate. Further drying or heat treatment is performed, as per the requirements of a given application. Highly porous and extremely low density 'aerogels' are formed by removing the liquid in a wet 'gel' under supercritical conditions (shown in Fig. 12.8).

Normally alkoxides are dissolved in alcohol and hydrolyzed by adding water. In the case of coatings, films are formed by the deposition of alkoxides followed by exposure to moisture. These reactions are well explained in many textbooks (Brinker *et al.*, 1990; Pierre, 1998). Deposition is usually carried out either using spinning or dip coating techniques. Roll/gravure (a process extensively used in the printing industry) coating is often employed, when thin coatings of low viscosity are to be deposited at very high speeds. Precursors play a very important role in the sol-gel process by directly influencing the porosity, refractive index, hardness and other mechanical and electrical properties of the films deposited. The most expensive parts of the process are the material costs and their handling procedures (as metal oxides are very moisture sensitive, these require special handling environments). However, the polymeric metal alkoxides synthesized relatively recently are

more environmentally friendly (and not moisture sensitive). A list of polymeric precursors along with their applications can be found in other references including www.ceramicindustry.com. A complete bibliography and latest developments in this technology including the Ph.D theses in this field can be seen on the website which, in our opinion, is by far the most comprehensive website on any one particular subject. This can be found at www.solgel.com/Resources/biblio.

One of the major advantages of the sol-gel process is its capability to tailor the properties of materials and films to suit a wide range of applications. Another interesting aspect of this process is its ability to facilitate doping of different elements in a host matrix film. The most important applications with respect to the nanostructured films are optical and gas sensing. Nanostructured thin films of TiO_2, WO_3, TiO_2 doped with pt and NbO_2, MoO_3-TiO_2, Ga_2O_3, and Ga_2O_3-CeO_2 and a host of many other materials have been sol-gel deposited to form thin and porous nanostructured films and have been tested for sensing gases like oxygen, carbon dioxide, oxone, nitrous oxide, alcohol and hydrogen (Atashbar, 1998; Li *et al*., 2001; Cantalini *et al*., 1999; Trinchi *et al*., 2004).

12.5.2 Chemical vapor deposition – carbon nanotubes (CNT)

Carbon nanotubes (CNTs), tiny hollow tubes with a diameter of only a few nms and lengths in the range of 100s of nms to μms, have been an integral part of nanotechnology. Hence, this section deals with one of the well-established fabrication technologies frequently used for preparing the CNTs, namely, chemical vapor deposition (CVD). After the discovery of CNTs by Sumio Iijima of NEC corporation in 1991 in the soot of an arc discharge, (Iijima, 1991), they have been considered as key elements of nanotechnology, mainly because of their interesting electrical and mechanical properties. Carbon nanotubes are commonly produced in two configurations, single wall and multiple wall. A single wall CNT (SWCNT) is a tubular shell of graphene sheet made up of benzene type hexagonal rings of carbon atoms, while a multi-walled CNT (MWCNT) is a stack of graphene sheets rolled up into concentric cylinders. As in-depth discussion on carbon nanotubes is carried out in a previous chapter, these aspects are not treated further here. Though initially arc discharge and laser ablation were used as possible routes for the production of CNTs, CVD and in particular plasma enhanced CVD (PECVD), have been found to be more favorable processes recently. The inherent difficulty in scaling up laser ablation as a manufacturing process on large substrates, has been the major constraint in its development as a potential fabrication technology, whereas the presence of amorphous carbon impurities in the arc deposited CNTs is the drawback with the arc discharge methods (Meyyappan *et al*., 2003).

CVD can be categorized into thermal and non-thermal methods. Thermal CVD use heat to activate growth and it plays a major role in the growth kinetics of the micro/nano structures on the substrate surface and this method has been successful in producing SWNTs. Thermal CVD can be divided into two sub-categories, hot wall CVD and cold wall CVD. The entire chamber is heated (using a tubular furnace) in a hot wall CVD system whereas the sample/substrate alone is heated to the required temperature (using a simple resistive heating setup in a cold wall CVD system. In both processes the chamber or substrate temperatures are normally maintained at around 900 °C during the deposition. The ramping to the required temperature in the hot wall CVD is much slower compared to cold wall CVD because of the thermal inertia of the furnace and to avoid breakage of the reactor tube (Cassel *et al*., 1999; Finnie *et al*., 2004). An interesting aspect of thermal CVD conditions used for CNT growth is that the growth temperature is maintained below the pyrolysis temperature of the precursor hydrocarbon gas to minimize gas phase disassociation. The precursor disassociation is normally carried out on the surface of the catalyst producing the carbon required for nanotube growth. However, so far, thermal CVD is proven to provide at best a marginal alignment of the grown CNTs, whereas plasma-based CVD techniques are relatively successful in producing better vertically aligned nanotubes.

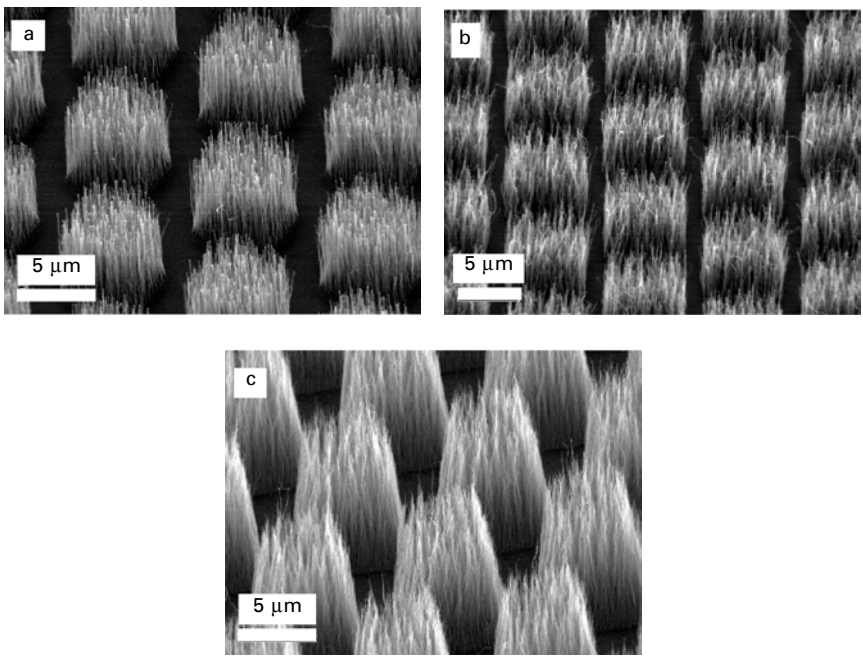
A metal catalyst usually of a transition metal layer (e.g. Ni, Co, Fe, Mo.) in the form of particles rather than a continuous film is used for seeding the CNT growth on any substrate for the fabrication of nanotubes using CVD techniques. These catalysts can be applied either using solutions containing them or can be directly deposited using physical vapor deposition techniques. Different catalysts and preparation methods were employed depending on the substrates and deposition conditions for CNTs (Meyyappan *et al*., 2003; Su *et al*., 2000; Merukulov *et al*., 2001; Delzeit *et al*., 2002).

Plasma enhanced CVD (PECVD) configuration used for the growth of the carbon nanotubes is similar to the one conventionally used in semiconductor film fabrication. However, the gas combinations and chemical gas phase reaction growth kinetics vary in different experiments depending on the deposition conditions. A number of gas combinations are used in both thermal CVD as well as PECVD techniques. Typical hydrocarbon sources used in plasma based growth of CNTs are methane, ethylene and acetylene. As a large number of gas phase reactions occur in the plasma creating many free radicals, the use of a pure hydrocarbon source alone may lead to a significant amount of amorphous carbon deposition. Hence a second gas such as argon, hydrogen or ammonia is used to dilute the primary hydrocarbon gas with the total reactor pressure being maintained in the range 1 to 20 Torr (Meyyappan *et al*., 2003). A wide number of plasma configurations DC (Han *et al*., 2000), RF (Ho *et al*., 2001) microwave or ECR (Tsai *et al*., 2002), inductively coupled plasma sources (Delzeit *et al*., 2002; Tsakadze *et al*., 2003) and RF

with magnetic enhancement (Satake *et al*., 2002) were used in the PECVD growth of CNTs. In many of these high-density plasma sources, the substrate holders are also reported to be biased by dc or rf power, which can potentially reduce the growth temperature.

The effect of the substrate bias on the growth characteristics of CNTs in a DC plasma assisted hot filament CVD system is shown in Fig. 12.9. In this technique, a dense plasma is created with the help of a filament placed between a meshed anode and the substrate. The reactive gas mixture is introduced into the chamber above the mesh anode. One of the most interesting aspects of this study is the observation of a transition of these multi-wall carbon nanofibers at a substrate bias voltage of –550 V to a well aligned exclusive growth of multi-wall carbon nanotubes. This is demonstration beyond any doubt. These are grown in an acetylene and ammonia gas mixture on patterned 20 nm thick nickel catalyst film on a layer of 100 nm thick chromium adhesive.

Since their discovery the CNTs have received considerable attention because of their interesting fundamental science as well as the potential for exciting



12.9 Effect of substrate bias on the carbon nanostructures (carbon nanofibers to carbon nanotubes in a dc plasma assisted hot filament CVD. (a) Well aligned MWCNFs at a bias voltage of –550 V (360 W, 670 mA). (b) A mixture of MWCNFs and MWCNTs grown at –575 V bias (400 W, 710 mA). (c) A well aligned exclusive growth of MWCNTs at a bias voltage of –600 V (470 W, 780 mA).

applications such as field emitters, sensors, transistors, electromechanical systems and above all as high-strength structural materials. Though chemical vapor deposition has been proven to be the best fabrication technology for growing single-wall as well as multi-wall carbon nanotubes beyond doubt, the number of configurations possible at times leads to confusion in the selection of a particular method over others. Many of the questions raised by Meyyappan and his colleagues still require intense research and further understanding in standardizing this technology for the routine fabrication of carbon nanotubes (Meyyappan *et al*., 2003).

12.6 Fabrication of nanowires – different technologies

Nanowires similar to nanotubes, have been another innovative set of components grouped within the cluster of nanotechnologies, which have tremendous potential not only for their possible applications, but also to scientific phenomena behind their conductivity and other related properties at quantum levels. Even the science behind the development of fabrication technologies has been intriguing and more understanding from current levels is required in order to improve these methods from research to industrial manufacturing. Like their nanotube counterparts, nanowires or rods are typically many nms in diameter and are 100s of nms to tens of micrometers long. These are often referred to as one-dimensional nanostructures (Xia *et al*., 2003). The major advantage with these nanowires is that they can be grown from a variety of different materials (from semiconductors to metal oxides to metals) with their chemistry being relatively easy to tailor. A novel set of applications such as 16 kilo bits of self-assembled nanocomputer memory, integrated nanoplanar electro-optical networks, nanowire-based high-definition (10^{11} pixels/sq.cm) television screens (which is an order of magnitude beyond what we have today) and novel chemical and biological sensors are forecasted (Appel, 2002) for nanowires of different materials. Solution-phase chemical or photochemical synthetic methods, template synthesis and physical vapor deposition onto carbon nanotubes or step edges have successfully been used for the fabrication of nanowires or nanorods (Jana *et al*., 2001a,b, Wang *et al*., 1999).

12.6.1 Template-based synthesis

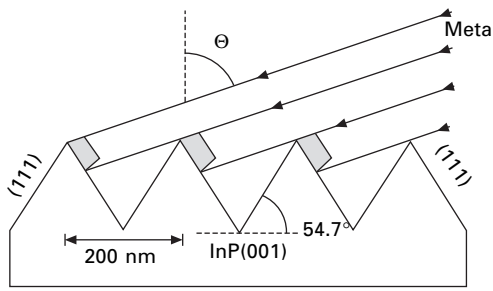
A template-based synthesis is considered as one of the most convenient methods for the fabrication of nanowires. Template-based synthesis mainly relies on different varieties of shadowing or nucleating structures on the substrate depending on the deposition technology employed. Four such examples, the shadow evaporation, deposition at the bottom of the V-grooves,

cleaved edge over growth and templating using step edges, were presented in a review recently (Xia *et al*., 2003). In all these cases, structures were fabricated by templating against features on solid substrates.

Jorritsma and co-workers prepared metallic nanowires using a shadow evaporation technique with a line width of 20–120 nms on single crystal (100) InP substrate. The V-shaped grooves were fabricated by laser interferometry and the metallic wires were deposited by placing the substrate at an oblique angle as represented in Fig. 12.10. The V-grooves have a period of approximately 200 nm, with the array of Tantalum wires prepared on a single substrate. Use of patterned V-groves in the fabrication of nanowires was reported by a number of research groups. Deposition of quantum wires by low-pressure organometallic chemical vapor deposition (LP-OMCVD) at the bottom of a V-grooved GaAs substrate (with a pitch of 3.5 or 0.5 μm) and results of characterization were reported by another group (Gustafsson *et al*., 1995). Sputtering and MBE techniques have also been used for the fabrication of nanowires using similar methods. Another method, which needs mention among this class of techniques, is membrane-based synthesis (Martin, 1994, 1995, 1996). This technique utilizes nanoporous membranes, quite often made in polycarbonate or anodic aluminum oxide, with a pore diameter even up to 10 nm having density in the range 10^9 to 10^{11} pores/cm^2 as templates. Electroless plating or the electroplating process chemically reduces cations in the pores of the template to form the nanowires.

12.6.2 Electrochemical step edge decoration

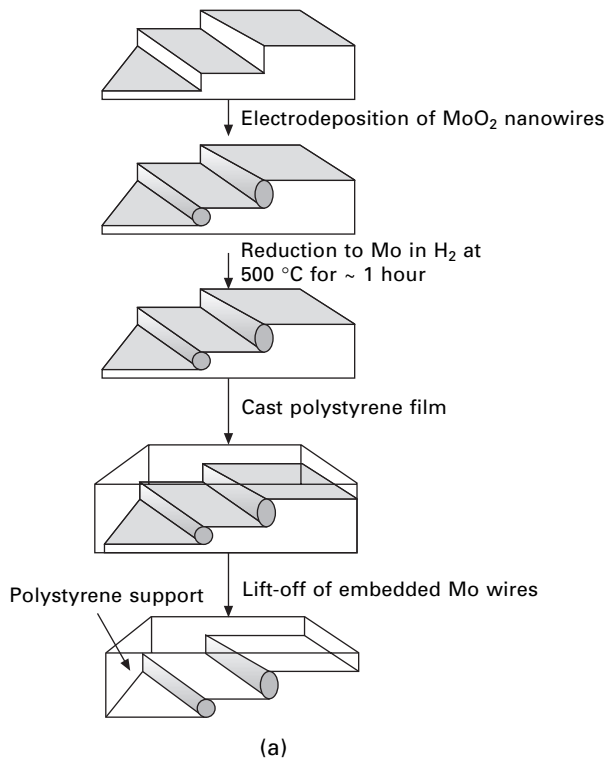
Another interesting technology that has recently evolved in this category is the electrochemical step edge decoration method developed by Penner's group at University of California, Irvine (Zach *et al*., 2000; Penner, 2002; Walter *et al*., 2002). In this process, the steps present in highly oriented pyrolitic graphite are used as the templates for the selective deposition of metal or metaloxide nanowires. Metallic nanowires can be fabricated in two different routes as shown in Fig. 12.11(a) and 12.11(b). The method involves



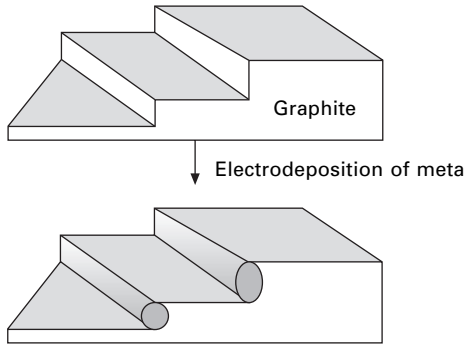
12.10 Schematic of a V-groove array on the substrate.

the electrodeposition of MoO_2 selectively at step edges on a graphite surface, and the subsequent reduction of these MoO_2 nanowires to molybdenum metal nanowires at 500 to 600 °C in hydrogen. However, this method cannot be used to prepare nanowires of metals that do not have a stable and electronically conductive oxide including noble metals. Figure 12.11(a) also depicts a method to release the nanowires from the conducting graphite surface to an insulating polystyrene surface using a simple lift-off procedure, which facilitates testing of these wires for conductivity and other properties. Figure 12.11(c) is the typical example of fabricated MoO_2 nanowires after plating for 16 seconds. Similarly, fabrication of palladium and silver wires has been reported recently using the direct process depicted in Fig. 12.11(b), though the latter experiment utilized aluminum oxide templates for its initial growth (Atashbar *et al.*, 2004; Cheng *et al.*, 2004).

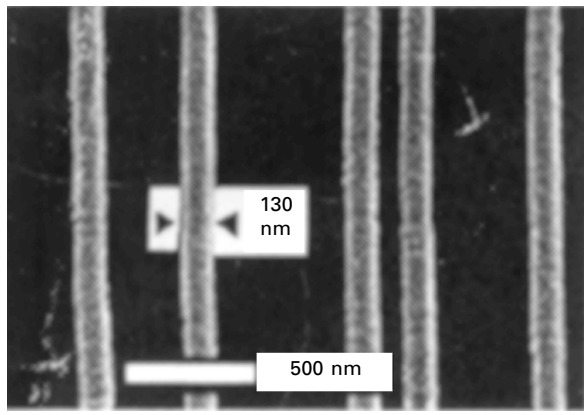
The initial nucleation to the latter plating processes is still being researched.



12.11 Electrochemical step edge decoration process. (a) Electrodeposition of metal oxide nanowires followed by reduction. (b) Direct electrodeposition of noble and coinage metal nanowires. (c) MoO_x nanowires after deposition for 16 s (Zach *et al.*, 2000).



(b)



(c)

12.11 contd

The plating reaction for the deposition of MoO_2 following the first method is

$$MoO_4^{2-} + 2H_2O + 2e^- \Leftrightarrow MoO_2 + 4OH^-$$

The defect sensitive electrodeposition of MoO_2 nanowires occurs at low current densities of 5 to 40 μA cm^{-2}, and growth times range from a few seconds for 10 to 20 nm diameter wires to hundreds of seconds for 0.5 μm diameter nanowires. The capability of this technique is demonstrated in fabricating the high aspect ratio (length to diameter ratio) nanowires with diameters as small as 13 nm and up to a length of 1 μm and more. Nanowire electrodeposition is proposed to occur in three stages. Formation of metal nuclei along the step edge in the first stage is followed by coalescence with the nearest neighbors to form 'beaded' nanowires. The final stage sees the smoothing of the continuous nanowire surface. While the transition from stage 2 to stage 3 seems to occur at around 10 to 15 nm, the smoothing process continues until around 60 nm. Some evidence for this kind of hypothesis

is shown by Penner's group (Zach *et al*., 2000) although these boundaries are still fuzzy and only a range of values may be taken for these transitions. The process underlying preferential nucleation and further plating at the step edge is also not very clear. However, so far to date, nanowires of noble metals (e.g. Pd, Cu, Au and Ag) and those of electronically conductive metal oxides (e.g. MoO_2, Cu_2O and Fe_2O_3) have been deposited using this method, while research on the fabrication of nanowires of other metals is still in progress.

Electrochemical deposition may also potentially be used in a localized manner using different electrode arrangements in order to produce micro/nanowires (Said, 2003). Fabrication of semiconductor (e.g. silicon) nanowires has been accomplished using different techniques (Milanovic, 2002; Hoffman *et al*., 2003), as silicon will have an advantage of being classified as a direct band gap semiconductor at the nanometer size due to quantum confinement. For this reason, silicon nanowires have applications both in electronics as well as optics.

12.7 Future trends

Recently, the rapid progress achieved in the development of micro- and nanotechnologies for various applications has led to the evolution of new fabrication technologies. These techniques will find increasing application in value-added products, provided the volume of product required can be synthesized at reasonable cost. The scaling up of the processes will be a challenge facing any group seeking to commercialize these technologies. Since the properties of the nanostructures are so dependent upon their size, any volume manufacturing process must produce large numbers of components within tightly controlled parameters without the need for slow measurement and inspection for quality control purposes.

An additional challenge relates to the precision with which the nanotools can produce the structures. Generally the tolerances for a manufactured part must be at least an order of magnitude smaller than the dimensions of the part itself. In many macro-world examples this ratio between part size and tolerance is usually around 1:1000, and in the case of optical devices this often exceeds 1:1,000,000. Clearly such tolerances on 100 nm structures begin to exceed the size of individual atoms, which is unrealistic, however, increasing control of tolerances and designs of systems that are able to cope with relatively poor dimensional precision will be a major challenge for any nanosystem designer.

A great prospect for new manufacturing strategies lies in the combination of biologically inspired methods with nano- and microstructured devices. Biological systems have evolved design strategies that are tolerant to the variation in size and quality of the structures, yet still achieve the desired function when assembled into larger units. For example, no two nanohairs in

a gecko's feet are identical, however, in bulk they still enable the reptile to climb up walls as smooth as glass.

Finally the integration of bio-micro-nanosystems will enable a whole new class of devices and systems to be designed and manufactured in a way that has never been possible before. A truly remarkable aspect of the emerging nanotechnology is its ability to allow engineers to design devices and solve problems using tools and processes that work at the same scale as many biological systems. Its impact on our lives, both in a positive and a negative sense, has enormous potential.

12.8 References

A comprehensive Sol-Gel Gateway web site: http://www.solgel.com/Resources/biblio

Appel D, *Nature*, 2002 419 553.

Atashbar M Z, 1998, Ph.D. thesis, RMIT University, Australia.

Atashbar M Z, Banerji D, Singamaneni S, Bliznycuk V, 'Deposition of parallel arrays of Pd nanowires and electrical characterization using microelectrode contacts', *Nanotechnology*, 2004 15 374–378.

Berger S *et al*., 1994, *SPIE Proc*. 2322 434.

Bohr M T, 'Nanotechnology Goals and Challenges for electronic applications', *IEEE Transactions on Nanotechnology*, 2002 1(1) 56.

Brien II M J O, Bisong P, Ista L K, Rabinovich E M, Garcia A L, Sibbett S S, Lopez G, Pand P, Brueck S R J, 'Fabrication of an integrated nanofluidic chip using interferometric lithography', *J. vac. Sci. Technol*. 2003 B21(6) 2941–2945.

Brinker C J, Scherer G W, 1990, *Sol-gel science: the physics and chemistry of sol-gel Processing*, Academic press, USA.

Cantalini C, Atashbar M Z, Li Y, Ghantasala M K, Santucci S, Wlodarski W, Passacantando M J, *Vac. Sci. Technol.*, 1999, A17(4) 1873–1879.

Cassel A M, Raymakers J A, Kong J, Dai H, *J. Phys. Chem.*, 1999 B–103, 6484.

Cheng Y H, Cheng S Y, 'Nanostructures formed by Ag nanowires', *Nanotechnology*, 2004 15 171–175.

Chou S Y, Kelmet C, Gu J, 'Ultrafast and direct imprint of nanostructures in Silicon', *Nature*, 2002 417 835–837.

Delzeit L, Nguyen C V, Stevens R M, Han J, Meyyappan M, *Nanotechnology*, 2002, 13, 280.

Deng T, Pretiss M, Whitesides G M, 'Fabrication of microfiltration systems using soft lithography', *App. Phys. Lett*., 2002 80(3) 461–3.

Deng T, Whitesides G M, Radhakrishnan M, Zabow G, Prentiss M, 'Manipulation of magnetic microbeads in suspension using micromagnetic systems fabricated with soft lithography', *App. Phys. Letters*, 2001 78(12) 1775–77.

Fantner E J, Loeschner H, *New tools for Nanotechnology*, Paper presented at ME-2003 Informationstagung Mikroelektronik, Oct. 1–2, 2003, Vienna, Austria (web site address: www.me-tagung.at).

Finnie P, Bardwell J, Tandev I, Tomlinson M, Beaulieu M, Fraser J, Lefebvre J, *J. Vac. Sci. Technol*. 2004 A22(3) 747–751.

From the web site: http://www.calpoly.edu/~sdygert/IonBeam.html.

Frosien J, Lischke B, Anger K, *J. Vac. Sci. Technol*., 1979 16 1827.

Gustafsson A, Reinhardt F, Biasiol G, Kapon E, 'Low pressure organometallicchemical vapor deposition of quantum wires on V-grooved substrates', *App. Phys. Lett*., 1995 67(25) 3673.

Han J, Yang W S, Yoo J B, Park C Y, *J. App. Phys*., 2000 88 7363.

Heritage M B *J. Vac. Sci. Technol.*, 1975 12 1135.

Hinsberg W, Houle F A, Hofnagle J, Sanchez M, Wallraff G, Morrison M, Frank S, 'Deep ultraviolet Interferometric lithography as a tool for assessment of chemically amplified photoresist performance', *J. Vac. Sci. Technol*. 1998 B16(6)

Ho G W, Wee A T S, Lin J, Tjiu W C, *Thin solid films*, 2001 388 73.

Hoffman S, Ducati C, Neill R J, Piscanec S, Ferrari A C, Geng J, Dunin-Borkoswski R E, Robertson J, *Journal of Applied Physics*, 2003 94(9) 6005–6012.

Iijima S, *Nature*, 1991 354, 56.

Introduction to Electron Beam lithography on the web: http://dot.che.gatech.edu/henderson/introduction_to_electron_beam_lithography.htm.

Jana N R, Gearheart L, Murphy C J, *J. Chem. Comm*., 2001b 617.

Jana N R, Gearheart L, Murphy C J, *J. Phys. Chemistry-B*, 2001a 105 4065.

Khizroev S, Bain J A, Litvinov D, 'Fabrication of nanomagnetic probes via FIB etching and deposition', *Nanotechnology*, 2002 13 619–622.

Khizroev S, Litvinov D, 'Focused-ion-beam-based rapid prototyping of nanoscale Magnetic devices', *Nanotechnology*, 2004 15 R7-R15.

Korda P, Spalding G C, Dufresne E R, Grier D G, 'Nanofabrication with holographic optical tweezers', *Review of scientific instruments*, 2002 73(4) 1956–7.

Kumar A, Whitesides G M, 'Features of gold having micrometer to centimeter dimensions can be formed through a combination of stamping with an elastomeric stamp and an alkanethiol "ink" followed by chemical etching', *Appt. Phys. Lett*., 1993 (14) 2002–4.

Li Y, Galatsis K, Wlodarski W, Ghantasala M, Russo S, Gorman J, Santucci S, Passacantando M, *J. Vac. Sci. Technol*., 2001, A19(3), 904–909.

Martin C R, 'Membrane based synthesis of Nanomaterials', *Chem. Mater*., 1996 81739–46.

Martin C R, *Acc. Chem, Res*., 1995 28 61.

Martin C R, *Science*, 1994 266 1961.

Merukulov V I, Melechko A V, Guillorn M A, Lowndes D H, Simpson M L, *Appl. Phys Lett*., 2001 79 2970.

Meyyappan M, Delzeit L, Cassell A, Hash D, *Plasma sources Sci. Technol*., 2003 12 205–216.

Michel B, Bernard A, Bietsch A, Delamache A, Geissler M, Juncker D, Kind H, Renault J-P, Rothuizen H, Schmid H, Schmidt-Winkel H, Stutz R, Wolf H, 'Printing meets lithography: Soft approaches to high-resolution patterning', *IBM J. of Research and Development*, 2001 45(5) 697–719.

Milanovic V, Doherty L, 'A simple process for lateral single crystal silicon nanowires, 2002', *Proc. Of IMECE'02*, ASME international mechanical engineering congress and exposition, New Orleans, Louisiana, November 17–22 2002.

Moore G, 'Cramming more components onto integrated circuits', *Electronics*, 1965 38(8) 1.

Nakasuji M, Suzuki S and Shimizu H, *Rev. Sci. Instr*., 1993 64 446.

Paetzel R, Albrecht H S, Lokai P, Zschocke W, Schmidt T, Bragin I, Schroeder T, Reusch C, Spratte S, personal communication, excimer systems for super high NA 193 nm lithography.

Penner R M, 'Mesocopic metal particles and wires by electrodeposition', *J. Phys. Chem. B*, 2002 106 3339–53.

Pierre A C, 1998, *Introduction to Sol-Gel Processing*, Kluwer Academic Publishers, USA.

Said R A, 'Microfabrication by localized electrochemical deposition: experimental investigation and theoretical modeling', *Nanotechnology*, 2003 14 523–531.

Satake N, Jeong G H, Hirata T, Hatakeyama R, Ishida H, Tohji K, Motomiya K, *Physica*, 2002 B-323 290.

Schmid H, Michel B, 'Siloxane polymers for high resolution, high accuracy soft lithography', Macromolecules, 2000 33 3042.

Solak H H, He D, Li W, Cerrina F, 'Nanolithography using extreme ultraviolet lithography interferometry: 19 nm lines and spaces', *J. Vac. Sci. Technol*. 1999 B17(6) 3052–3057.

Su M, Zheng B, Liu J, *Chemical physics letters*, 2000 322 321–326.

Tennant, D M 1998 'Limits of conventional lithography', *Nanotechnology*, ed. Gregory Timp, AIP press, New York, 192.

Trinchi A, Kaciulis S, Pandolfi L, Ghantasala M K, Li Y, Wlodarski W, Viticoli S, Comini E, Sberveglieri G, *Sensors and Actuators B* (in press).

Tsai C L, Chen C F, Wu L K, *App. Phys. Lett*., 2002 81 721.

Tsakadze E, Xu S, Long J, Ostrikov K, *Bull. Am. Phys. Soc*. 2003 48, 13.

Walter E C, Murray B J, Favier F, Kaltenpoth G, Grunze M, Penner R M, 'Noble and Coinage metal nanowires by electrochemical step edge decoration', *J. Phys.Chem*., 2002 B-106 11407–411.

Wang Z L, Mohamed M B, Link S, El-Sayed M A, *Surface Science*, 1999 440 L809.

Web site on Sol-Gel technology: www.chemat.com.

Web site reference: www.ceramicindustry.com.

Weller R B, Ryele R B, Newton R B, McMahon M D, 'A technique for producing ordered arrays of metallic nanoclusters by electroless deposition in focused ion beam Patterns', *IEEE Transactions on Nanotechnology*, 2003 2(3) 154–157.

Xia Y, McClelland J J, Gupta R, Qin D, Zhao X-M, *et al. Advanced Materials*, 1997, 147–149.

Xia Y, Whitesides G M, 'Soft Lithography', *Angew. Chem. Int. Ed*., 1998a 37 550–575.

Xia Y, Whitesides G M, 'Soft Lithography', *Annual Reviews on Material Sciences*,1998b 28 153–84.

Xia Y, Yang P, Sun Y, Wu Y, Mayers B, Gates B, Yin Y, Kim F, Yan H, 'One-dimensional nanostructures: synthesis, characterization and applications', *Advanced Materials*, 2003 15(5) 353.

Zach M P, Kwok H, Ng, Penner R M, 'Molybdenum Nanowires by electrodeposition', *Science*, 2000 290 2120–2123.

Zhao X M, Xia Y, Whitesides G M, *Adv. Materials*, 1996 8 837–840.

Index