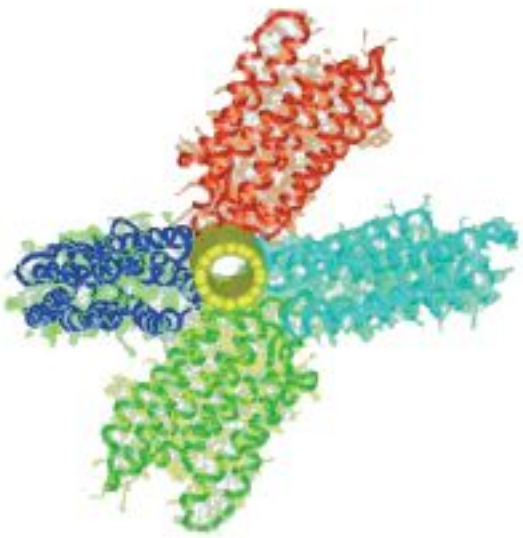
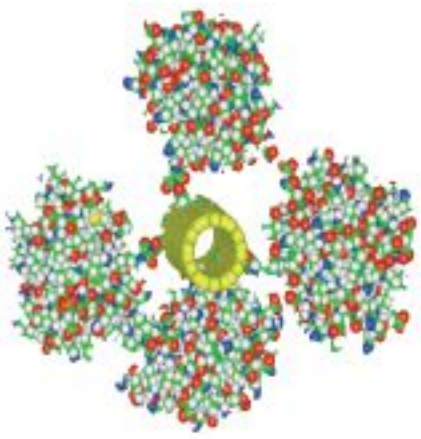
Bionanotechnology

Proteins to Nanodevices

Edited by

V. Renugopalakrishnan and Randolph V. Lewis





Springer

BIONANOTECHNOLOGY

VR: To Bhanu, Priyanka, Varun, Suraj, Sandeep

BIONANOTECHNOLOGY

Proteins to Nanodevices

Edited by

V. Renugopalakrishnan
Children's Hospital, Harvard Medical School,
Harvard University, Boston, MA, USA

and

Randolph V. Lewis
University of Wyoming,
Laramie, WY, USA

Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN-10 1-4020-4219-1 (HB)
ISBN-13 978-1-4020-4219-5 (HB)
ISBN-10 1-4020-4375-9 (e-book)
ISBN-13 978-1-4020-4375-8 (e-book)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

www.springer.com

Printed on acid-free paper

Printed in the Netherlands.

Contents

Preface

Bionanotechnology represents an intriguing facet of nanotechnology, as it strives to harness the molecules and processes of biological systems to design functional nanoscale devices. It takes advantage of the features acquired by living organisms in the course of evolution for technological purposes. It finds illustration and inspiration in biology to advance toward more and more complex artificial systems, emerging from the biological ones. It also signifies to nanotechnology in general, that such high complexity can be achieved, as does exist in biology, thus giving confidence to its practitioners.

It is of much interest to bring together in the present volume important contributions to the field of bionanotechnology, so as to provide a stepping-stone for further progress. Exciting perspectives open up, indeed!

Jean-Marie LEHN

PROTEINS TO NANODEVICES

V. Renugopalakrishnan[1], and Randolph V. Lewis[2]
[1]*Harvard University, Cambridge, MA, USA*
[2]*University of Wyoming, Laramie, WY, USA*

This edited volume includes a series of lectures delivered at the 47th Biophysical Society annual meeting in San Antonio, TX, USA.

Proteins are naturally occurring nanosytems, optimized by the process of evolution. Biotechnology of protein-based nanostructures offers vast opportunities to reengineer and combine them with either themselves by the process of self-assembly to create supramolecular aggregates or with other nanomaterials for technological applications. Bionanotechnology encompasses the study, creation, and illumination of the connections between structural molecular biology and molecular nanotechnology since the development of nanomachinery might be guided by studying the structure and function of the natural nanomachines found in living cells. Investigating the topology and communication processes of bionanotechnology is the key integrative technology of the 21st century and aims to use the knowledge, gathered from the natural construction of cellular systems, for the advancement of science and engineering cell parts and can lead to invention of novel biological devices with enormous applications. Although microscale to nanoscale research offers an excellent space for development of futuristic technologies, a number of challenges need to be overcome. Due to paucity of a dedicated literature on protein-based nanodevices, we bring you this book that combines collective research of scientists probing into this fascinating universe of bionanotechnology. The publication is written with an aim of surveying engineering design principles of biomolecular nanodevices, prototype nanodevices based on redox proteins, bacteriorhodopsins, and natural fibers, and also offers a snapshot of the future developments in the field.

Our thanks to Dr Peter Butler and his team at Springer, Dordrecht, The Netherlands for the fruition of this edited volume and the Army Research Office for supporting the Biophysical Society workshop. Editors wish to express their grateful thanks to Dr Pawan K Dhar for his valuable contribution culminating in the successful fruition of this volume. Indirect

financial support has been received from the National Science Foundation, Harvard Medical School, Wallace H Coulter Foundation, and Florida International University for which VR expresses his gratitude. Author express their appreciation to Sukanya Ratna and Dhaneesh Kumar of SPI Publisher Services for their excellent cooperation and patience in making this publication a success.

BIONANOTECHNOLOGY: PROTEINS TO NANODEVICES

Preface

(By Dr. R. A. Mashelkar, DGSIR and Secretary, DSIR)

Nanobiotechnology is an emerging area of scientific and technological opportunity. The terms "Nanobiotechnology" and "Bionanotechnology," used interchangeably, apply the tools and processes of nano/microfabrication to build devices for studying or manipulating biosystems. The field is progressing incredibly fast worldwide.

Global attention to this emerging field was triggered by an epoch making talk delivered by the 1965 Physics Nobel laureate, Robert Feynman entitled, *There is plenty of room at the bottom* on December 29, 1959 at the annual meeting of the American Physical Society at the California Institute of Technology (Caltech), followed by a series of articles and books by K. Eric Drexler, chiefly the Engines of Creation.

More than a decade ago, I gave the 9th P.V. Danckwerts memorial lecture in London on 9th June 1994. It was titled as 'Seamless Chemical Engineering Science: The Emerging Paradigm'.

I had then said 'Nanobiotechnology is emerging as an exciting field. Fabricating organized bimolecular structures is a new challenge that is being met by brining in diverse disciplines. Genetic engineering will tailor protein's makeup and also its function. Organic chemistry will provide new materials for anchoring and preserving the altered proteins. Electrical engineering will offer ways to detect signals from the internal working of proteins. We can finally turn a small assembly of biomolecules into a custom designed molecular plants.

Another exciting area is where proteins are being turned into nanoscale chemical processing plants. Proteins can be used as reaction vessels for controlling the particle size of other materials. They can cage several

compounds. If the chemical affinity of the protein can be changed, then the protein can be engineered to fit the mineralization product that one wants inside it. Such caged particles will pass through the human body in an unrestricted way. They will be useful in diagnosing and treating diseases. It is clear that chemical engineering scientists can share the grand challenges of nanotechnology by joining hands with others working in protein engineering, organometallic chemistry, semiconductor technology, etc'.

The challenge then posed seems too have been taken up by the scientific community and rapid strides have been made in this field.

According to industry sources, the estimated market for nanobiotechnology in 2003 was INR 42 billion (USD 970 M), however, it is expected to grow rapidly to reach over INR 135 billion (USD 3120 M) in 2008, reflecting growth at an annual rate of 28% worldwide. The current estimated worldwide market breakdown is US at 65% Europe at 20%, Japan at 10% and the rest of the world at 5%.

Nanomaterials, such as carbon Buckyballs, dendrimers, and metal nanoparticles, are being introduced in the market for pharmaceutical and diagnostic use both in vitro and in vivo. Currently available products fall into three segments of nanobiotechnology: drug delivery, imaging agents and biosensors.

Internationally, the role of nanobiotechnology in food, health, ecological and livelihood security has been well recognized toward meeting the UN Millennium Development Goals in the area of health and sustainable food security. In view of this, it has been advocated that a major initiative on "Nanobiotechnology, Food and Health Security" may be designed and launched. Such a National Challenge Program should bring together all appropriate R&D institutions in the country (Universities, ICAR, CSIR, ICMR, DBT and non-governmental institutions). The National Challenge Program should have well-defined goals and built-in monitoring mechanisms. It should also provide for public-private partnerships such as envisaged in New Millennium Indian Technology Leadership Initiative (NMITLI).

Chapter 1

DESIGN PRINCIPLES FOR SELF-ASSEMBLING DEVICES FROM MACROMOLECULES

F. Raymond Salemme
Imiplex Corp, Yardley PA

Abstract: Proteins are the natural building blocks for functional, self-assembling nanostructures. This short review cites the application of protein engineering technology to the formation of engineered functional polymers, as well as systems that self assemble as organized 2D molecular lattices on surfaces. Some consideration is given to approaches that could lead to self assembling structures of great complexity.
Key words: Bionanotechnology, protein structure, protein engineering, molecular lattices, biomaterials, streptavidin, molecular electronics

1. INTRODUCTION

One vision of nanotechnology anticipates the precision assembly of complex, large-scale systems incorporating individual devices engineered at the molecular scale. This chapter describes some initial studies and outlines some design principles that explore the assembly of macroscopic structures composed of proteins ordered at the molecular level. Biological systems form complex, self-assembling structures from amino acid polymers (proteins) encoded by deoxyribonucleic acid (DNA). Individual proteins containing thousands of atoms spontaneously fold to form unique three-dimensional (3D) structures replicated with atomic precision. Naturally occurring proteins exemplify many practical, structural, and signal transduction systems, including chemomechanical, electromechanical, optomechanical, and optoelectronic mechanisms, and frequently form cooperative hierarchical assemblies of great structural and functional complexity. The combination of recombinant DNA technology, enabling the synthesis of virtually any polypeptide sequence or functional domain fusion,

V. Renugopalakrishnan et al. (eds.), Bionanotechnology, 1-8.

together with lessons learned from sequentially assembled biological structures like virus particles and regular macromolecular assemblies like protein crystals, provide the basis for designing novel assemblies from engineered biological macromolecules. The successful development of these new material applications requires close integration of computer design, protein synthesis, biophysical measurements, and structural characterization using both X-ray diffraction and electron microscopy. Early results are exemplified by studies of engineered structural fibers based on architectural motifs found in viral spike structures. Examples of designs and assembly strategies for 2D "semiconductor-like" architectures based on biological macromolecular motifs are also presented.

2. DEVELOPMENT OF ORDERED POLYMERIC SYSTEMS BASED ON ADENOVIRUS TAILSPIKE PROTEIN ARCHITECTURE

During the early 1990s, an effort was launched through cooperation between the Dupont Central Research and Development Department and Dupont Polymer Products Department to investigate the use of proteins to form ordered polymeric systems.[1] The anticipated advantages of this approach included the ability, by using the methods of recombinant DNA technology, to produce high molecular weight polymers (MW$\sim$100,000 Da) with precisely defined amino acid sequences and composition. By creating polymers composed of repetitive blocks of amino acid sequences corresponding to known protein structural domains, it was envisioned that it would be possible to create polymeric structures that would spontaneously fold and assemble into 3D structures that were ordered at both the molecular and macroscopic levels. Initial objectives focused on producing materials that would emulate the unique mechanical properties of natural biological fibers like elastin and spider silk. At another level, it was believed that the molecular organization present in biological structures could be exploited to produce materials with, for example, novel nonlinear optical properties.

Efforts focused on the generation of repetitive consensus polymer sequences derived from the adenovirus spike protein. The adenovirus spike is a long stiff structure that emerges from the apex of the virus icosahedral capsid structure. Electron microscope and X-ray scattering studies of the adenovirus spike showed that it formed a long stiff shaft organized as a cross-β sheet structure. Sequence and composition analysis showed that the spike was composed of three identical polypeptide chains, each of which incorporated several approximate repeats of a homologous 15-residue amino acid sequence. At Dupont, constructs were developed that expressed

repetitive consensus sequences based on the natural sequence, in the hope that they would spontaneously assemble into trimeric, microfibrillar shafts whose intrinsic stiffness would facilitate spinning into macroscopic fibers in which the individual trimeric shaft axes were all aligned along the direction of the fiber axis. Three sequences were developed, which differed slightly with regard to repeating block sequences and extent of block polymerization, to produce polymer chains ranging from MW ~20,000 to ~100,000 Da.

The synthetic work was complemented by molecular modeling and energy minimization studies that envisioned that the shaft structure was organized as a twisted triangular beam in which each face of the beam was composed of a twisted antiparallel β-sheet with geometric and twist properties similar to that observed in many globular protein structures (Fig. 1).

In fact, the ''designed'' polypeptide chains were experimentally observed to form long fibrous shafts in electron microscope imaging and light scattering studies and to produce liquid crystalline mesophases. The liquid crystalline mesophases could be spun into macroscopic fibers. X-ray diffraction studies of spun fibers showed features characteristic of cross-β structure, although the features were not predominant along the fiber axis direction as expected if the

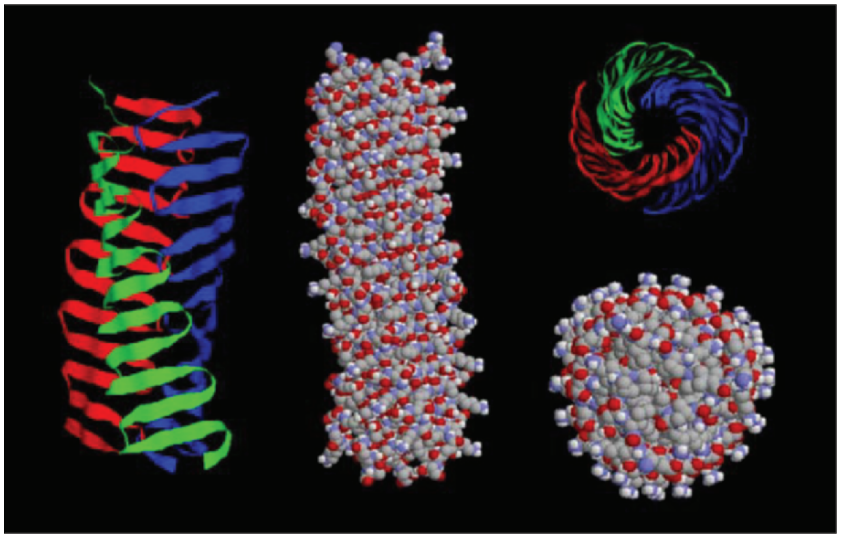


Figure 1. (A) A backbone model of a synthetic cross-β beam structure composed of three polypeptide chains in which each face of the beam is composed of a continuous, twisted, antiparallel β-sheet. (B) End views of the backbone and a space-filling model of the energy-minimized structure, illustrating the solid packing of the interior structure by hydrophobic amino acid residues of the repeating polypeptide sequence (Asparagine–Alanine–Leucine–Arginine–Isoleucine–Lysine–Glycine–Serine–Glycine–Leucine—Aspartic Acid–Phenyalanine —Aspartic Acid–Asparagine)$_n$.

shafts were closely aligned along the fiber axis. It was believed that additional experiments to improve spinning solvents and proper conditions could potentially produce more highly aligned polymeric materials, although this objective was not actively pursued at that time owing to changing priorities within Dupont Central Research.

Subsequently, the X-ray crystal structure of the adenovirus shaft was reported at near atomic resolution,[2] and it turned out to have major differences, but some similarities, with the structure that had been modeled. The model structure envisioned that each face of the trimeric shaft was composed of a single polypeptide chain, organized as repeating β-hairpins to form a contiguously hydrogen sheet whose H-bonds lay along the axis of the shaft. The actual virus structure is composed of polypeptide chains that form repetitive β-bends, which alternate "herringbone-fashion" between two adjacent faces of the trimeric beam structure. In addition, the hairpin bends in each chain twist so that the interactions between strands are primarily due to stacking between the sheets of the adjacent chains, rather than to interchain hydrogen-bonding (Fig. 2).

The true structural organization of the synthetic structures is further clouded by more recent studies[3] that investigated the propensity for both randomized and homopolymer amino-acid sequences to form amyloid structures. Both the electron microscope images of amyloid fibers and their dif-

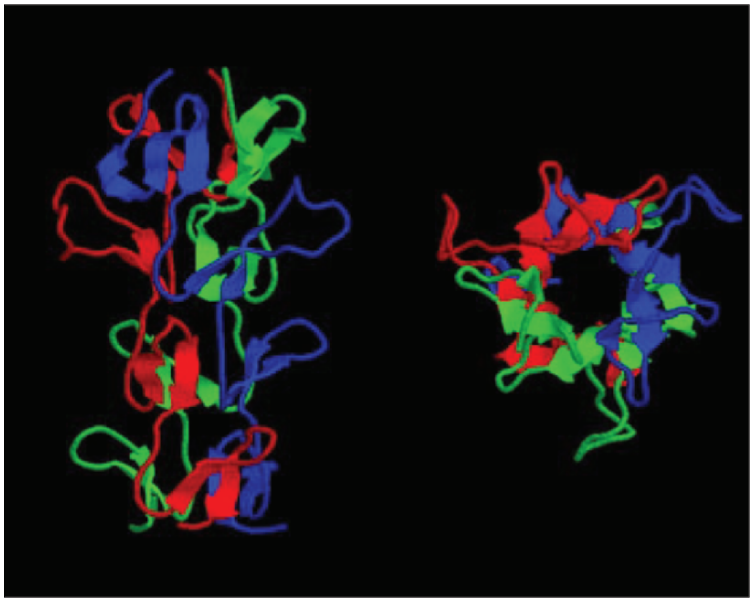


Figure 2. (A) View of a backbone model of a section of the adenovirus spike as determined from X-ray crystallography. (B) An end view.

fraction patterns appear very similar to the corresponding data derived from the engineered synthetic adenovirus sequences. Obviously, much remains to be done to further characterize these structures, both as a means of manipulating them as new structural materials and potentially to suggest a means of ameliorating amyloid diseases.

3. TWO-DIMENSIONAL LATTICES ON SURFACES

Although many practical sensor applications are already based on the immobilization of proteins on semiconductor substrates, longer range objectives envision the self-assembly of complicated molecular circuitry on semiconductor substrates that could potentially emulate the function of present day metal-oxide-semiconductor field-effect transistor (MOSFET) devices but at substantially higher device densities. Key issues involved in the realization of such ideas involve the design and functional optimization of individual molecular components, followed by their controlled hierarchical organization into large-scale functional assemblies. The strategies for producing devices with molecular scale components combine aspects of the ''top–down'' approach used for conventional MOSFET device design with a ''bottom-up'' approach to self-assembly common to biological systems.

3.1 Two-Dimensional Lattices Based on Streptavidin

Early conceptual ideas[4] envisioned the formation of 2D molecular structures that could self-organize on self-assembling monolayer (SAM) surfaces. It is desirable in many biotechnology applications to link biomolecules irreversibly through a highly specific protein–ligand interaction. The high-affinity interaction between streptavidin and its small-molecule ligand biotin,[5] which can readily be linked covalently to a variety of other biomolecules, has been widely used for biotechnology applications. In an early hypothetical example of a self-assembling 2D lattice, it was envisioned that array structures could be formed by linking ''strut molecules'' constructed, for example, from DNA segments terminated with covalently linked biotin groups to ''nodes'' formed by the symmetric, tetrameric streptavidin molecule.[4] A remarkable paper by Ringler and Schulz[6] actually realized such 2D lattice structures on SAMs by linking streptavidin and an engineered fourfold symmetric protein to form molecular lattices with predetermined dimensions and symmetry. As shown in Fig. 3, the lattice is composed of two molecular components, streptavidin, a tetramer with D2 symmetry that incorporates four biotin binding sites, and RhuA aldolase, a fourfold symmetric tetramer that has been engineered using recombinant DNA technology to introduce surface cysteine sulfhydryl groups

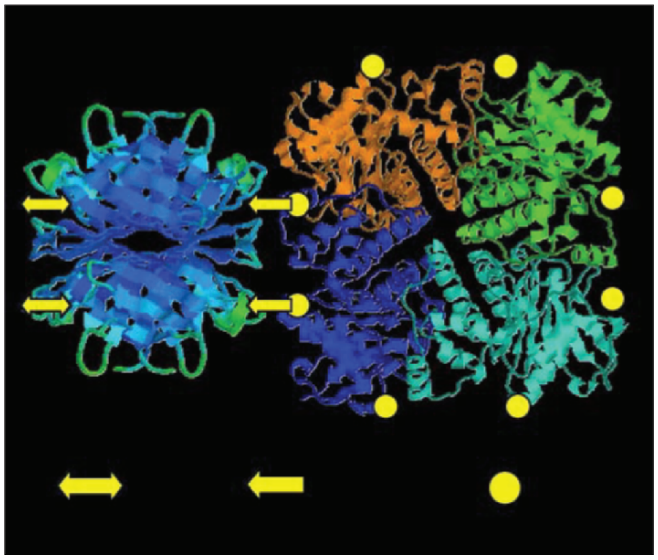


Figure 3. Molecular components of an engineered lattice formed of protein molecules arrayed on a 2D surface. The two molecular components are streptavidin, a tetramer with D2 symmetry that incorporates four biotin binding sites, and RhuA aldolase, a fourfold symmetric tetramer that has been engineered using recombinant DNA technology and then chemically modified to allow the covalent chemical attachment of biotin molecules. Circles represent sites of cysteine SH groups introduced into the tetrameric protein aldolase through site-directed mutagenesis, Arrows represents biotin groups that covalently react with the cystiene SH groups. Double arrows represent "di-biotin" crosslinking molecules.

that were subsequently functionalized through the covalent attachment of biotin molecules.

The work of Ringler and Schulz demonstrates the potential for forming controlled 2D molecular assemblies of protein molecules on surfaces. High-density lattices composed of protein molecules with switchable magnetic or photosensitive prosthetic groups might for example be components of high-density memory arrays or sensors.

3.2 Self-Assembly of Hierarchical Organized Structures

The work of Ringler and Schulz[6] demonstrates the potential for self-assembly of regular 2D lattices with defined geometrical properties constructed of protein molecules on surfaces (see Fig. 4). However, substantially more complicated self-assembling structures can be envisioned that could be constructed using a combination of molecular epitaxy and gated molecular assembly principles. There have been great advances in the use of atomic force microscopy to manipulate individual atoms on silicon or other semiconductor substrate surfaces.[7] Although it is probably unrealistic to imagine that devices of

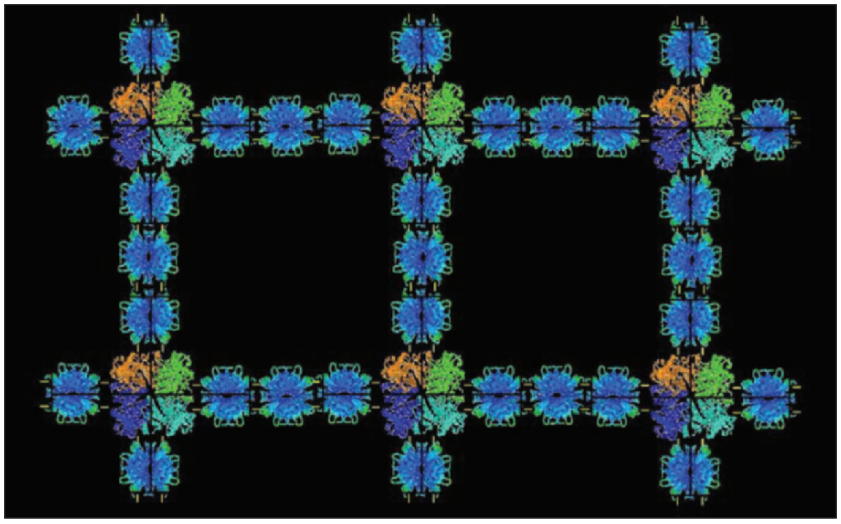


Figure 4. A schematic of the engineered 2D lattice formed by coordination of streptavidin and the engineered and modified form of RhuA aldolase, as described by Ringler and Schultz.

substantial complexity could actually be built using atom-by-atom construction principles, it is certainly feasible to produce a regular 2D lattice of individual atoms on a silicon surface. A regular lattice of gold atoms, for example, can be used as chemical immobilization sites for a 2D lattice of protein molecules linked through surface cysteine sulfhydryl groups. Additional different proteins can be added to immobilized proteins to eventually produce organized structures of potentially arbitrary complexity. In contrast to the 2D lattices that self assemble on SAM surfaces, whose formation is facilitated by the free diffusion of the assembling molecules on the SAM surface, the latter strategy would introduce fixed sites on the silicon surface to provide nucleation points for a sequence of hierarchical molecular assembly steps. The wealth of available X-ray crystal structures of proteins provides an extensive toolbox of molecular components that are able to undergo or perform chemomechanical, electromechanical, optomechanical, and optoelectronic processes. In addition, structural biology provides many examples of structures whose intermolecular interactions can be precisely controlled both in space and time.

While control of intermolecular interaction geometry is fundamental to any design process, lessons derived from the assembly of complex biological assemblies like viruses, suggest that it will be equally important to achieve temporal control over successive stages in the assembly of more complicated molecular devices. Again, structural biology provides numerous examples of how structural assembly processes can be "gated," usually as a consequence of either a precise biochemical modification of one of the molecular components or through the introduction of a small-molecule ligand that induces a

conformational change in one of the molecular components so that it can interact stably with another.

The approach outlined in the preceding paragragh envisions a process in which the interactions of discrete molecular components are precisely controlled at every level and stage of device assembly. However, for many device applications, large-scale assemblies could potentially be spontaneously organized by other means. One interesting possibility involves systems based on finite automata.[4,8] For example, studies of interactions important for the habit development in protein crystals[9] and simulations of extended lattice structures formed by finite automata nucleated at a single point,[4,8] suggest that molecular systems incorporating components whose interactions are governed by simple interaction and symmetry rules, can spontaneously assemble to form highly complex, yet finite structures.

4. CONCLUSION

Although many outstanding issues will require resolution if useful functional devices are ultimately built from macromolecules, many of the required tools for simulation, modeling, protein engineering, structure determination, biophysics, and semiconductor technology are available to begin investigating the potential of such devices.

REFERENCES

1. O'Brien, J. P., Hoess, R. H., Gardner, K. H., Locke, R. L., Wasserman, Z. R., Weber, P. C., and Salemme, F. R., 1994, Design, synthesis, and fabrication of a novel, self-assembling fibrillar protein, in *Silk Polymers : Materials Science and Biotechnology* (D. Kaplan, W. W. Adams, B. Farmer, and C. Viney, eds.), ACS Symposium Series 55, pp. 104–117.
2. Van Raaij, M. J., Mitraki, A., Lavigne, G., and Cusack, S., 1992, A triple beta-spiral in the adenovirus fibre shaft reveals a new structural motif for a fibrous protein, *Nature* **401:**935–936.
3. Dobson, C. M., 2001, The structural basis of protein folding and its links with human disease, *Phil. Trans. Roy. Soc. Lond. B*. **356:**133–145.
4. Sligar, S. L., and Salemme, F. R., 1992, Protein engineering for molecular electronics, *Curr. Opin. Struct. Bio*. **2:**587–592.
5. Weber, P. C., Ohlendorf, D. H., Wendoloski, J. J., and Salemme, F. R., 1989, Structural origins of high affinity biotin binding to streptavidin, *Science* **243:**85–88.
6. Ringler, P., and Schulz, G., 2003, Self-assembly of proteins into designed networks, *Science* **302:**106–109.
7. Strocio, J. A., and Eigler, D. M., 1991, Atomic and molecular manipulation with the scanning tunnelling microscope, *Science* **254:**1319–1326.
8. Wolfram, S., 2002, *A New Kind of Science*, Wolfram Media.
9. Salemme, F. R., Genieser, L., Finzel, B. C., Hilmer, R. M., and Wendoloski, J. J., 1988, Molecular factors stabilizing protein crystals, *J. Cryst. Growth* **90:**273–282.

Chapter 2

METALLOPROTEIN-BASED ELECTRONIC NANODEVICES

Ross Rinaldi, Giuseppe Maruccio, Adriana Biasco, Pier Paolo Pompa, Alessandro Bramanti, Valentina Arima, Paolo Visconti, Stefano D'amico, *Eliana D'Amonne, and Roberto Cingolani
**National Nanotechnology Laboratory of CNR-INFM, Strada per Arnesano Km 5, 73100 Lecce (Italy)*

Abstract: A key challenge of the current research in nanoelectronics is the realization of biomolecular devices. The use of electron-transfer proteins, such as the blue copper protein, azurin (Az), is particularly attractive because of its natural redox properties and self-assembly capability. This chapter discusses results of fabrication, characterization, and modeling of devices based on this redox protein. The prototype of biomolecular devices operate in the solid state and in air. The charge transfer process in protein devices can be engineered by using proteins with different redox centers (metal atoms) and by controlling their orientation in the solid state through different immobilization methods. A biomolecular electron rectifier has been demonstrated by interconnecting two gold nanoelectrodes with an Az monolayer immobilized on SiO_2. The device exhibits a clear rectifying behavior with discrete current steps in the positive wing of the current—voltage (*I–V*) curve, which is ascribed to resonant tunneling through the redox active center. On the basis of these results, an Az-based three-terminal device has been designed. The three-terminal device exhibits an ambipolar behavior as a function of the gate bias, thus opening the way to the implementation of a new generation of logic architectures. This peculiar characteristic allows the implementation of a fully integrated nanoscopic logic gate.

Key words: biomolecular devices, nanotechnology, three terminal electronic devices, proteins

V. Renugopalakrishnan et al. (eds.), Bionanotechnology, 9-24.

1. INTRODUCTION

In the last decade there have been dramatic advances toward the realization of molecular scale devices and integrated computers at the molecular scale.[1]. First pioneering experiments were performed demonstrating that individual molecules can serve as nanorectifiers[2] and switches,[3,4] 1000 times smaller than those on conventional microchips. The assembly of tiny computer logic circuits built from such molecular scale devices has been demonstrated.[5]

Researchers are working to join biology and nanotechnology, fusing useful biomolecules to chemically synthesized nanoclusters in arrangements, which do everything from emitting light to storing tiny bits of magnetic data. The result is a merger that attempts to blend biology's ability to assemble complex structures with nanoscientists' capacity to build useful devices. One of the biggest drivers behind nanotechnology's enthusiasm for biological systems revolves around the organism's impressive ability to manufacture complex molecules such as DNA and proteins with atomic precision.

One of the fundamental goals of bioelectronics is the realization of nanoscale devices in which a few or a single biomolecule can be used to transfer and process an electronic signal. The biomolecules have particular functionality that can be exploited for the implementation of electronic devices. In order to design and realize such devices, several steps are required: (i) choice and characterization of the suitable biomolecular system, (ii) immobilization of the molecule onto an electronic substrate, (iii) interconnection to contacts, (iv) device fabrication processing, (v) read-out, recognition of molecular information, and (vi) processing of the information. The nanobiotechnology community has started a big effort along these lines to realize biomolecular devices for information technology.[6] Among biomolecules, proteins have a fundamental role in biological processes. The combination of molecular biology (for engineering proteins with the desired, either functional and/or yield or self-assembling, properties/capabilities) and nanotechnology (for device fabrication), thus becomes the tool to realize a new class of nanoelectronic elements.[7] Different nanotechnological strategies have been selected to implement the biomolecular devices, following a bottom-up or a top–down approach depending on the biomolecule and its functionality.

Thanks to their functional characteristics, metalloproteins appear to be good candidates for biomolecular nanoelectronics. Among them, blue copper proteins[8] and, in particular Az, appear to be the best candidates of choice due to some specific structural properties and intrinsic functionality in biological environments. They can bind gold via a disulfide site present on the protein surface, and their natural electron transfer activity can be exploited for

the realization of molecular switches, whose conduction state can be controlled by tuning their redox state through an external voltage source (gate). In order to realize a real biomolecular device in the solid state operating in air, a comprehensive study of the redox, electronic, and electrical properties of the metalloproteins linked to an inorganic substrate under nonphysiological environments has been conducted.[9] Besides all those devices based on proteins and enzymes, which belong to the family of biosensors, having typical size of hundreds of microns or more, a novel example of protein-based active electronic device has been demonstrated by the group of Rinaldi and coworkers.[9,10] This chapter describes the implementation techniques and the basic principle of operations of such class of novel protein-based nanodevices.

2. DEVICE FABRICATION

2.1 Protein Monolayers

In the field of biomolecular electronics, the function of proteins adsorbed at solid interfaces is of fundamental importance. Moreover, since electronic devices fabrication aims at producing a device working in air, another critical treatment is the drying procedure that can influence the protein activity. Therefore, the optimization of the chemisorption process of a protein at a surface has to be carefully developed, and the protein functionality has to be checked after each step.

For the implementation of protein monolayers, surface functionalization by means of silanes, which enables the adsorption process, has been developed.

Azurin[8] from *Pseudomonas aeruginosa* is a small (molecular mass 14.6 kDa, Fig. 1) and soluble metalloprotein involved in the respiratory phosphorylation of its hosting bacterium. The protein folds into an eight-stranded Greek key-barrel motif with only a helix present (Fig. 1).

Given to this secondary structure and also to the presence of the disulfide bond, it can be considered as a very stable protein. Structural and electronic studies have shown that the Az capability to function as a one-electron carrier in the biological environment is due to the equilibrium between the two stable oxidation states of the Cu ion, Cu^{+} (reduced) and Cu^{2+} (oxidized), and the structural stability of the active site. The redox active center of Az contains a copper ion liganded to 5 amino acid (aa) atoms in a peculiar ligand-field symmetry, which endows the protein with unusual spectroscopic and electrochemical properties, such as an intense electron absorption band at 628 nm, a small hyperfine splitting in the electron paramagnetic spectrum, and an

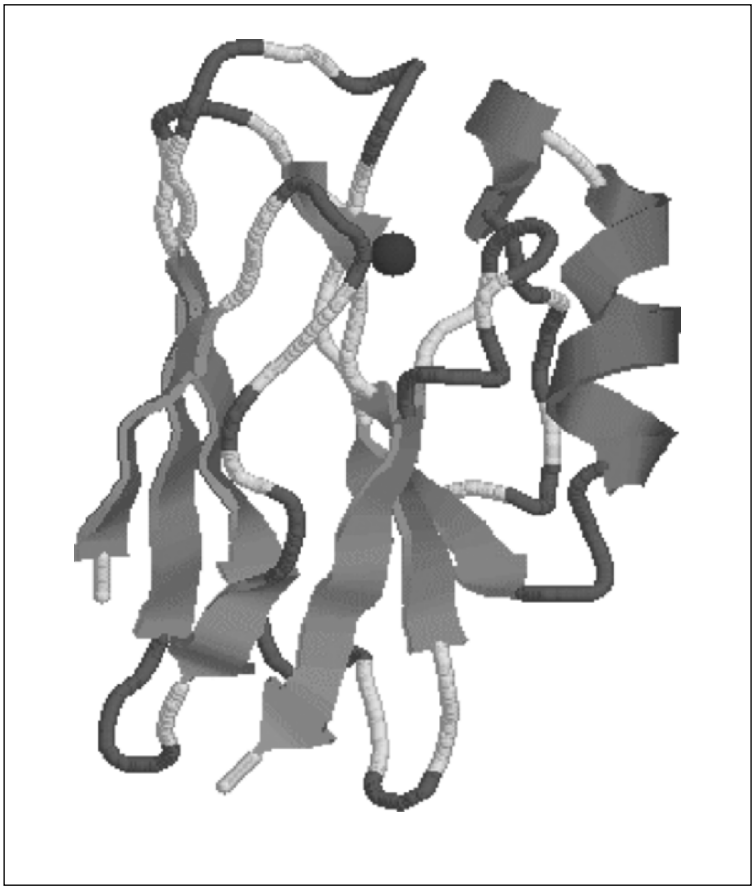


Figure 1. Schematic structure of Azurin representing eight β-barrel strands (yellow) and just one α-helix (pink).

unusually large equilibrium potential (+116 mV vs SCE) in comparison to the Cu (II/I) aqua couple (−89 mV vs SCE).

Commercial natural Az from *Pseudomonas aeruginosa* (Sigma) was used without further purification after having checked that the ratio OD_{628}/OD_{280} (OD = optical density measured at nm) was in accordance with the literature values (0.53–0.58).

Two different immobilization procedures were developed for Az on Si/SiO_2, exploiting different structural features of the protein and resulting in: (1) random orientation (three-step procedure) and/or (2) oriented immobilization (two-step procedure) in the protein monolayers. Following the three-step procedure (1), the immobilization of the proteins is achieved

through the formation of imido bonds between the exposed amino groups on the Az surface and the carbonyl heads exposed on the functionalized SiO_2 substrate. This procedure is based on a three-step chemical reaction performed first, by incubating the Si/SiO_2 substrates with gold electrodes in 3-aminopropyltriethoxysilane (3-APTS) [diluted to 6.6% (V/V) in $CHCl_3$ immediately prior to use] for 2 min and then by rinsing in $CHCl_3$, in order to remove 3-APTS molecules not linked to the surface. The second step consisted in exposing the sample already reacted with silanes that was exposed to glutharic dialdheyde (GD) (diluted in H_2O to a final concentration of 4×10^{-4} M) for 10 min and successively followed by a thorough washing in ultrapure H_2O. Finally, the precoated substrates were exposed to Az solution for 10 min and rinsed in NH_4Ac to get rid of physisorbed molecules. For this step a working solution of 10^{-4} M Az in 50 mM NH_4Ac (Sigma) buffer, pH 4.6, was prepared. The buffer was degased with N_2 flow prior to use. Milli-Q grade water (resistivity 18.2 MS/m) was used throughout all the experiments. In the three-step supramolecular reaction, the immobilization involves the chemical bonding of the exposed amino groups. Since the Az molecule contains 13 surface-exposed amino groups (found in the Lys residues and in the N-terminus), random orientation of the proteins is obtained in the solid state by this immobilization method.

In contrast, a highly ordered phase is achieved following procedure (2), which exploits a different two-step immobilization protocol acting on the unique disulfide bond found in the protein between Cys3 and Cys26. Such a procedure not only provides a special orientation of the protein molecules, but also facilitates the chemisorption process by avoiding the GD exposure. In this case, the SiO_2 substrates were cleaved and washed with organic solvents and then incubated with H_2SO_4 and bidistilled deionized water for 1 h. After washing with abundant water, the substrates were incubated for 2 min with a 2% ethanol solution of mercaptopropyltrimethoxysilane (3-MPTS) and then washed in absolute ethanol. After this step, the sample was dried under nitrogen (N_2) stream. A drop of the protein solution (0.6 g/ml) was deposited onto the silane layer, and the substrate is incubated for 5 min at room temperature and continuously shaken in order to avoid stagnant conditions. Then it was rinsed for 5 min with the buffer and finally with deionized water. The samples were dried under N_2 stream for noncontact atomic force microscopy (NC-AFM) imaging. A typical AFM image of Az monolayer is reported in Fig. 2.

Protein immobilization took place via the reaction of the free thiol groups of 3-MPTS with the surface disulfide bridge of Az, giving rise to substrate/overlayer disulfide bonds. By using the two-step method for the immobilization, the orientation of the molecule is expected to be well defined, as confirmed by electrochemical (cyclic voltammetry) and scanning force microscopy

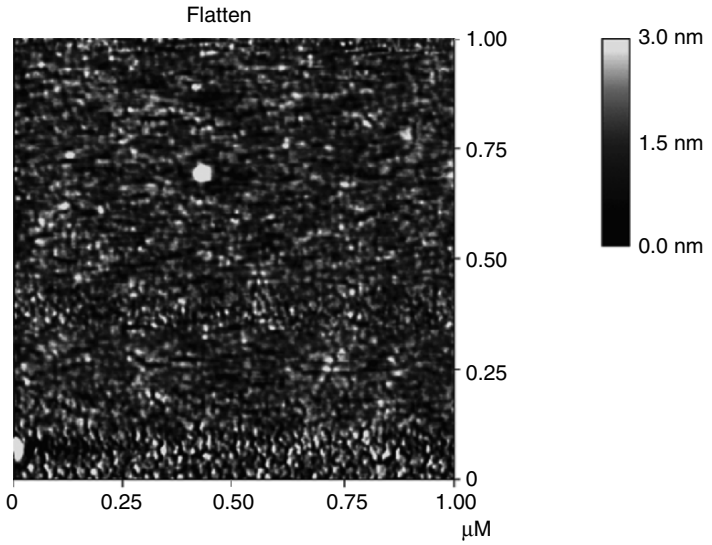


Figure 2. Non contact atomic force microscopy image of a self assembled azurin monolayer, realized by means of the two step immobilization procedure.

(SFM) measurements.[11] Samples covered just by the first molecular layers (namely 3-MPTS, 3-APTS, and 3-APTS+GD) were also produced for reference.

The last procedure is specific for all those proteins with S atom that can covalently bind the substrate. In the specific case of Az, it exploits the presence of the disulfide bond Cys-3–Cys-26, present on the protein external surface.

Synthetic wild-type Az and Az with zinc as a metal ion and without the metal (APO form) have been obtained using the recombinant DNA technique. The gene coding for the protein of interest has been introduced in a plasmid employed to transform *E. coli* cells. The two-step immobilization procedure was used to produce monolayers with synthetic proteins having the same orientation with respect to the substrate surface.

Intrinsic fluorescence experiments performed on the immobilized proteins showed no protein denaturation after the chemisorption and draying procedures.

Other blue copper metalloproteins, for example the Plastocyanin, can be used to implement the electronic devices described in the next section. In this case, however, mutants (e.g., PCSS and PCSH) have to be produced to achieve a stable and oriented immobilization of plastocyanin onto gold substrate. In the first mutant (PCSS), a disulphide bridge can be inserted within the protein. The residues Ile-21 and Glu-25 can be mutated to Cys by structurally conservative mutagenesis. In the second mutant (PCSH), usually a

residue tail (threonine, cysteine, and glycine) at the C-terminal position is inserted. The single exposed thiol group should provide a further possibility for plastocyanin immobilization.

2.2 Implementation of the Two- and Three-Terminal Protein-Based Nanodevices

Nanoelectrodes were fabricated by electron beam lithography and lift-off on a Si/SiO_2 substrate. The geometry of the device comprises two planar Au/Cr contacts separated by a gap in the range 50–100 nm. An Az monolayer was then immobilized by means of surface functionalization and chemisorption in the gap between the electrodes of the planar circuit. Both commercial natural Az and synthetic purified Az were used to implement the devices, following the methods described in Section 2.1. In the case of three-terminal devices the contact separation defined the gate width and a silver electrode was added on the back of the Si substrate as the gate to control the source–drain conduction. A schematic diagram of the device geometry and implementation is reported in Fig. 3. All the fabricated devices were tested at room temperature and ambient pressure. Prior to protein deposition, a negative control on the empty devices was performed to check the effective insulation between the source (S), drain (D), and gate (G) terminals along the different current pathways. All these tests revealed drain–source (I_{DS}) and source–gate (I_{SG}) currents lower than 20 pA and typical open circuit resistance larger than 100 G.

3. NANOELECTRONICS WITH PROTEIN

3.1 Two-Terminal Protein Devices

The achievement of oriented immobilization is extremely crucial for electronic applications in which the charge transport benefits the long-range order of the transporting material. Orientation can in principle affect conduction in two ways: (i) it allows increased protein coverage, thus favoring electron transfer among neighboring molecules and (ii) it enhances, for a given coverage, the intermolecular electron transfer (due to the fact that the positions of the Cu sites are approximately coplanar, thus offering more favorable pathways for conduction).

In addition, the molecular electrostatic potential (MEP) of the protein in solution is known to be important in protein interaction properties at

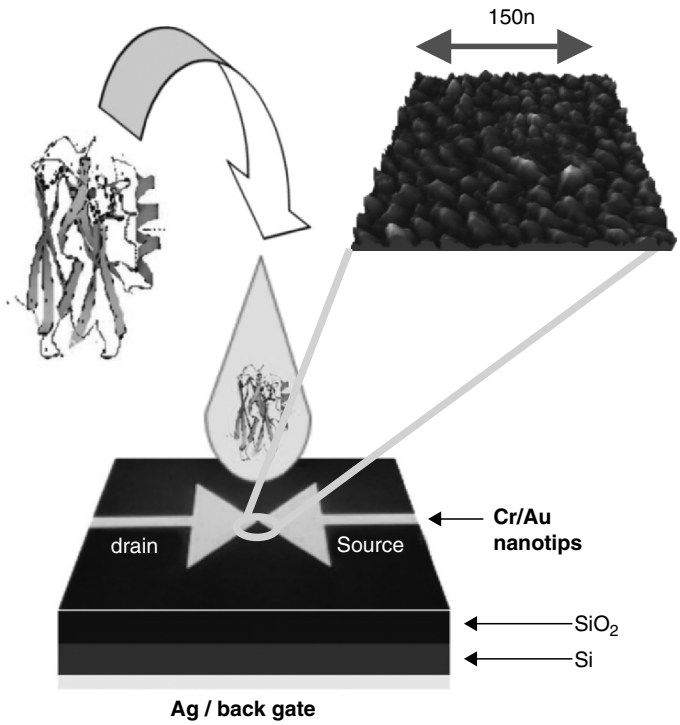


Figure 3. A schematic diagram of the device geometry and implementation procedure.

medium- and long-range distances in solution and plays a fundamental role in the recognition processes of biomolecules. It is therefore expected that electrostatics will influence both the deposition kinetics of the proteins in solution and their self-assembly. The charge distribution on the Az surface gives origin to a strong intrinsic dipole (150 Debye), suggesting that two-terminal circuits, interconnecting solid state films of immobilized Az molecules, should have an intrinsic polarity which depends on the value and the orientation of the molecular dipoles. Both oriented and nonoriented Az films should exhibit a macroscopic polarization. However, the total dipole is enhanced by depositing oriented self-assembled films with parallel dipoles. If the dipole distribution is preserved in the device after drying, it is expected to induce a macroscopic electric field favoring conduction.

The comparison between the *I–V* curves measured in the oriented and randomly oriented Az layers is reported in Fig. 4. The continuous and dotted lines represent the downward and upward sweeps, respectively. Three important effects can be deduced from this comparison. Both curves are asymmetric, with a strong rectifying behavior. The value of the current measured under

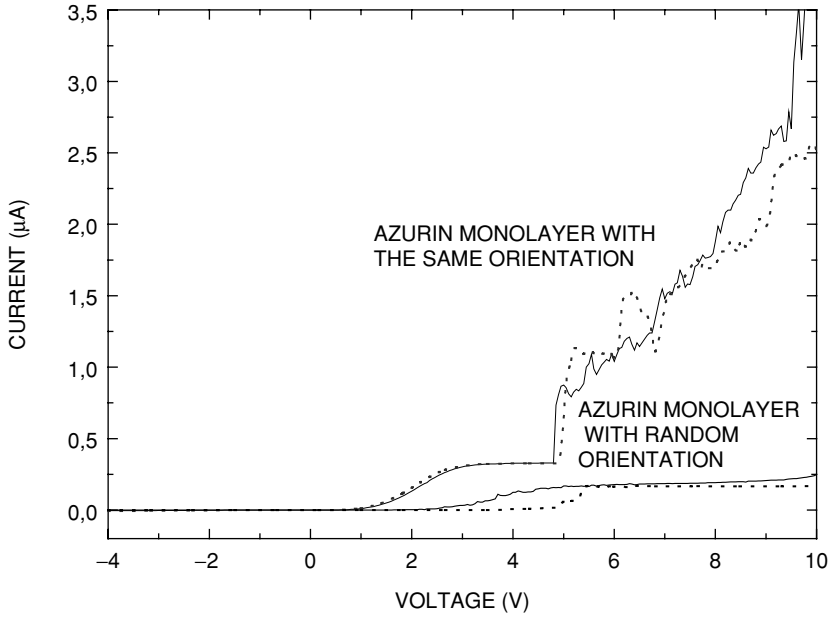


Figure 4. Current voltage curves of two terminal devices interconnecting proteins films with the same orientation and with random orientation. The gap between the gold electrodes was 70 nm.

forward bias between the nanoelectrodes suggests that the electron transfer mechanisms in the protein, between the Cu site and the edges, are quite effective in determining the conduction processes. The difference between the positive and negative wings of the current curve of both samples may be attributed to the presence of the dipole in the Az molecules, which sets the polarization of the planar devices.[9]

The current flowing through the device with the oriented layer is about ten times larger than that flowing through the device with the nonoriented layer. In fact, the regular orientation of the Az molecules, determined by the unique sticking site on the protein (the Cys3–Cys26 bridge) exploited for the layer formation, drives the self-assembly on the substrate, resulting in a distribution of parallel dipoles. This induces a macroscopic electric field favoring conduction. In the sample with random orientation of the proteins, although the electrostatic long-range interaction may favor some dipole alignment on the surface, a complete parallel orientation of the molecular dipoles cannot be achieved as a consequence of the many possible sticking sites on the protein surface used in the immobilization procedure.

In the oriented protein layer under forward bias, the current is step-like with a smooth exponential rise in the region between 1.9 and 2.3 V and a steep rise around 4.9 V (Fig.4). The step centered around 2.1 V corresponds to the energy required by the protein molecule to reduce the Cu atom by means of the electronic transition involving the S(Cys)Cu charge transfer. The step around 4.9 V corresponds to the energy required to perform resonant tunneling via the redox levels of Az. Such a process occurs through a coherent two-step tunneling in which the electrons go from the negative to the positive electrode via the molecular redox level. This is consistent with the electrochemical STM measurements and *in situ* cyclic voltammetry curves, showing a maximum at −4.96 eV (measured with respect to the vacuum level). This step is also observable in the *I–V* curve of the sample with randomly oriented proteins, though less pronounced and with a small hysteresis.

In order to further elucidate the role of the metal ion and the effects of purity of the protein sample on the device performances, three different types of high purity of engineered Azs were used for implementing oriented layer devices as shown in Fig. 5.[10]

(i) A highly purified synthetic Az, referred to as "recombinant Az," in sample (A) of Fig.5. These proteins are identical to the natural proteins (Fig. 13a), though with a higher degree of purity

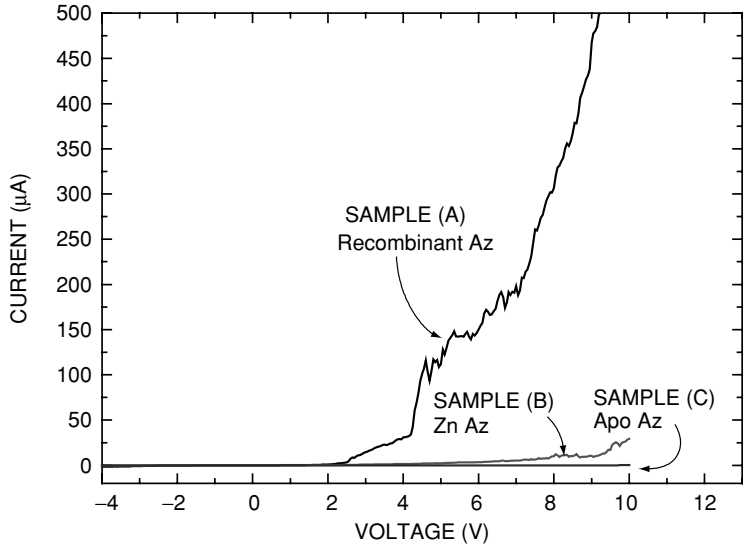


Figure 5. Current voltage curves of (A) recombinant Az (black), (B) Zn-Az (red), and (C) APO-Az (blue). The gap between the gold electrodes was 70nm.

(ii) A modified synthetic Az with the Cu atom replaced by a Zn atom (Zn–Az), in sample (B) of Fig. 5
(iii) A modified Az without metal atom (called APO-Az), in sample (C) of Fig. 5.

Sample (A) shows the same rectification and step-like character observed in the devices fabricated with oriented commercial (nonpurified) Az layers, reported in Fig. 4. However, the current flowing through device (A) with recombinant proteins is two orders of magnitude larger. This suggests that highly purified proteins are best-suited for molecular electronic applications. Reasonably, a high degree of purification favors a closer packing of the proteins onto the device substrate and avoids passivation phenomena of the substrate. Sample (B) shows a less-pronounced rectifying behavior, without the redox-induced steps characteristic of the Cu atom, and a current intensity about one order of magnitude lower than sample A. This is ascribed to the electronic properties of Zn that are different from those of Cu. Zn has in fact only one stable redox state, Zn^{2+}, which is redox inert, thus preventing Zn–Az molecules from being efficient redox species–mediated electron carriers. The importance of a metal atom in the protein structure is demonstrated by a device (C) that was built using APO-proteins obtained from recombinant Az. In the absence of metal redox center in the protein structure, the biodevice does not show any measurable conduction. This indicates that the metal in the protein is responsible for the electron transfer as reflected in the transport characteristics of Az biomolecular diodes.

3.2 Three-Terminal Protein Devices

The natural electron transfer activity of the Az can be exploited for the realization of molecular switches, whose conduction state can be controlled by tuning their redox state through an external voltage source (gate).

Fig. 6a shows the I–V characteristics of a protein device for $V_{DS} > 0$ as a function of V_G in the range between 0 and 4 V (first active region). As a general feature, the I_{DS} remains low (<20 pA) up to a V_{DS} bias voltage of about 2.3 V and then starts to increase reaching 100 pA intensity at 6 V. A clear modulation effect is visible as a function of V_G. The dependence of the I_{DS} current as a function of the V_G intensity is reported in Fig. 7a for a fixed value of $V_{DS}(V_{DS} = 4.5\,V)$. The current increases up to a maximum value of 200 pA at $V_G = 1.1\,V$, then it decreases to the 100 pA level at $V_G = 2\,V$, and finally falls down to the open circuit value at the higher applied gate voltages.

Similar set of measurements were performed under positive and negative V_{DS} bias with positive (Figs. 6a and 6c, respectively) and negative (Figs. 6b

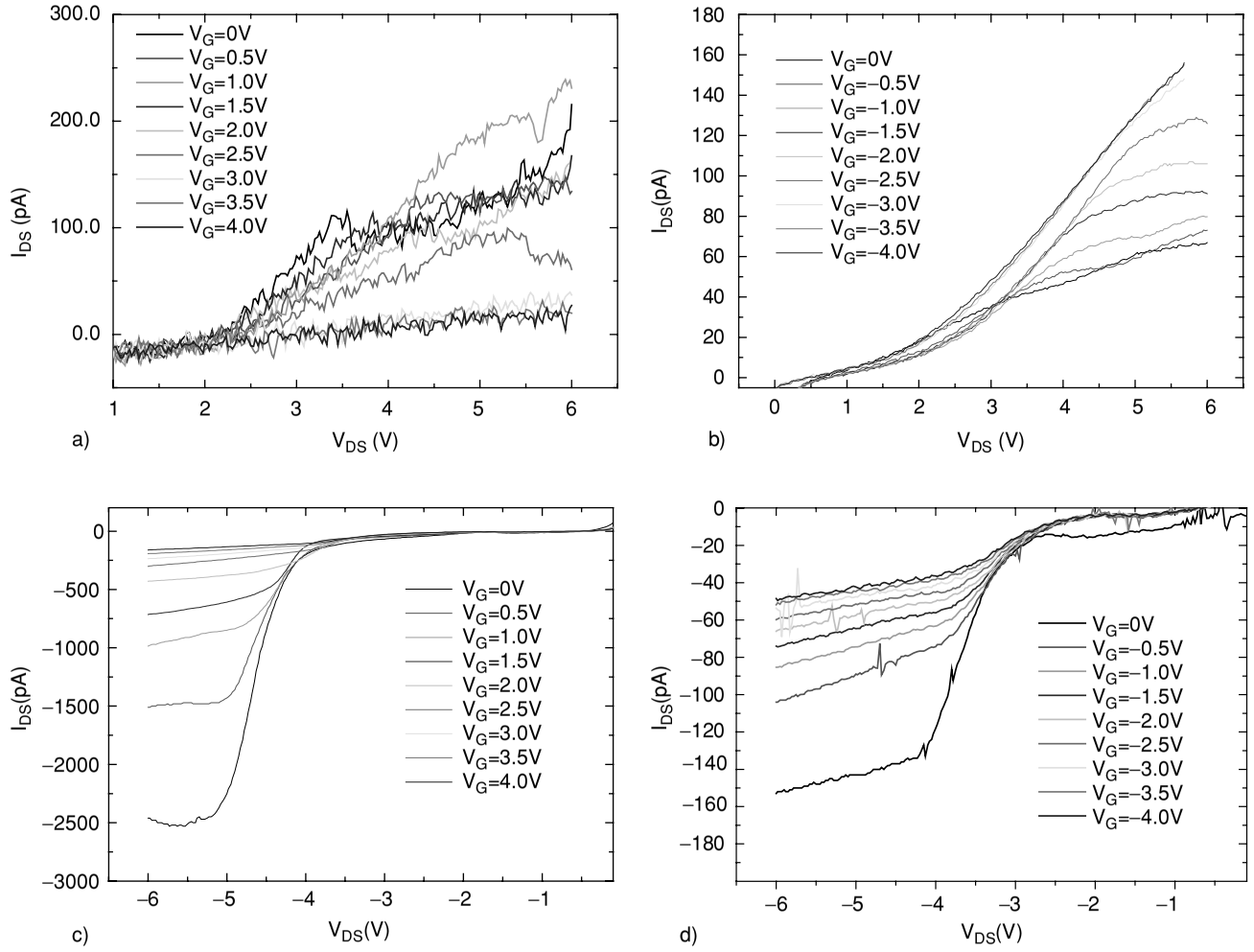


Figure 6. I-V characteristics of the protein transistor in the four different active regions : a) $V_{DS}>0$, $V_G>0$; b) $V_{DS}>0$, $V_G>0$; c) $V_{DS}<0$, $V_G>0$; d) $V_{DS}<0$, $V_G>0$.

and 6d, respectively) gate bias. Four different active regions resulted: $V_{DS} > 0$, $V_G > 0$ (16a); $V_{DS} > 0$, $V_G < 0$ (16b); $V_{DS} < 0$, $V_G > 0$ (6c); $V_{DS} < 0$, $V_G < 0$ (6d). In Fig. 6b, by decreasing the gate potential the current increases up to 160 pA at $V_{DS} = 6\,V$. The corresponding dependence of the I_{DS} current as a function of V_G is reported in Fig. 7b for a V_{DS} value of 5 V. In this case, we did not observe a peaked dependence, like in the case of Fig. 7a, but a smooth increase and a saturation at 11 5pA in the range between -2.5 and -4 V. In the last two situations, i.e., $V_{DS} < 0$, $V_G > 0$, (Fig. 7c) and $V_{DS} < 0$, $V_G < 0$ (Fig. 7d), we observed a decrease and an increase (in absolute values) of the I_{DS} negative currents in the two cases, respectively, with increasing the gate voltage. As a general trend, the current increases linearly for V_{DS} voltages higher than 3 V and then saturates in the range between 4.5 and 6 V. In this voltage range the operation of the protein transistor resembles that of an inorganic MOS-FET in the saturation region, with constant I_{DS} current value for each curve. In the case of the Az monolayer deposited in the channel of our FET, both superexchange charge transfer inside the single Az units and sequential hopping between adjacent proteins are involved in the

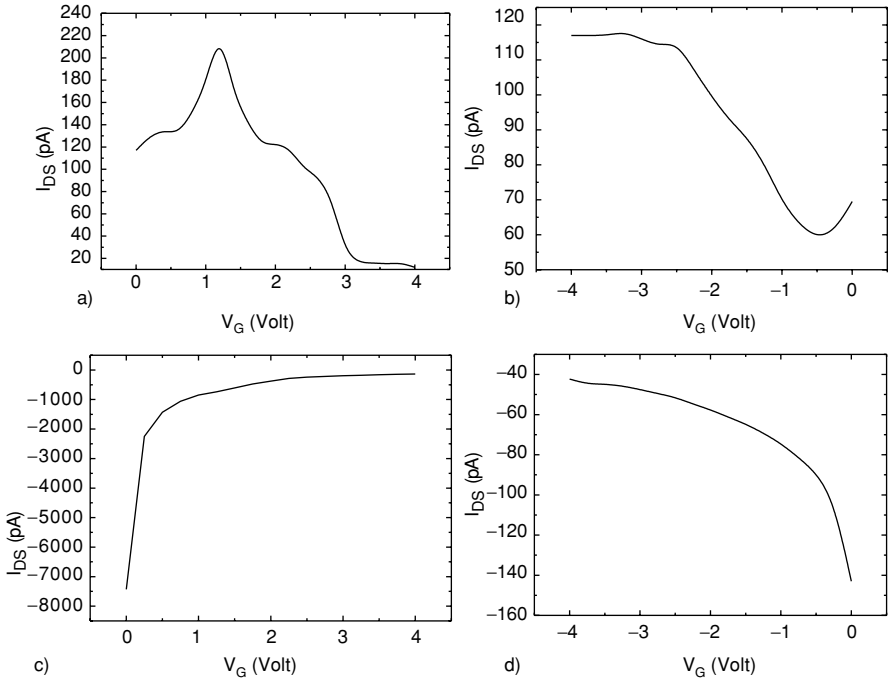


Figure 7. I_{DS} current intensities as a function of V_G at fixed V_{DS} for the four active regions reported in figure 6: a) $V_{DS} = 4.5V$; b) $V_{DS} = -5V$; c) $V_{DS} = 5V$; d) $V_{DS} = -5V$.

conduction process. The existence of a strong internal dipole due to the surface charge distribution on the protein can affect the planar conduction in the gate channel depending on whether the dipoles of the proteins in the layer are aligned or not along the source–drain axis. In the case of the characteristics of the device fabricated with commercial Az, a similar effect is observed for positive gate voltages values (Figs. 6a and c). In the case of the I–V measurements in Fig. 6c, current intensities are higher than the corresponding values in Fig. 6a, at least by a factor 10. For negative gate voltages, the measured current intensities in the 6b and d diagrams are similar.

The change of the gate voltage acts in different ways depending on the cross correlation of the signs of the bias and the internal dipole. As a general trend, the I_{DS} current decreases with the increase in the absolute value of the applied V_G. However, a different trend is usually observed in the case of $V_{DS} > 0$, $V_G > 0$ (Fig. 6a) in which a peaked dependence is observed in most of the realized devices. The maximum of the current is always found around $V_G = 1$ V. This is a peculiar behavior of the protein device that is not found in any inorganic counterpart and could be of interest for the realization of electronic devices with novel functionalities. In the case of three-terminal devices realized with synthetic and purified Az, the current increases by a factor 10 to 100 and the V_G resonance becomes sharper. The physical origin of this trend can be found in the effect of the gate voltage on the molecular levels of the protein, and in turn on the conduction pathways inside the molecule. The protein is chemisorbed onto the SiO_2 surface by the formation of a thiolate onto the silanized surface. In this way the protein sits on the surface with the axis connecting the thiolate with the Cu ion almost perpendicular to the surface. The vertical field that propagates along this direction (gate field) modifies the electronic properties of the molecule and the oxidation state of the Cu ion, since it is strongly coupled to the thiolate by the electron transfer pathway inside the protein.[12] This resonance effect has been already observed in electrochemical STM experiments[13] and standard STM experiments performed on Az molecules chemisorbed on Au(111) substrates at room temperature and ambient pressure. It has been demonstrated that the STM dsetection of the protein at the surface strongly depends on the applied bias between the tip and the gold substrate.

4. CONCLUSION

We have demonstrated the first implementation of protein-based electronic nanoscale devices working in air and at room temperature by using Az metalloproteins. Both rectifying (diode-like) and amplifying (transistor-

like) devices have been implemented. We have demonstrated the first protein transistor working in air and at room temperature by using Az metalloproteins. Though this field is rather young and very open, the results obtained so far are promising and deserve further studies to determine the actual potentiality of these biodevices and elucidate a number of important issues such as the reproducibility and ageing of the nanodevices.

ACKNOWLEDGEMENTS

We are thankful for the invaluable support and exciting collaboration by various colleagues. We would like to thank Elisa Molinari, Rosa Di Felice, Francesca De Rienzo, Paolo Facci at S^3 – INFM research center in Modena (Italy), Gerard Canters and Martin Verbeet at Leiden University (NL).

Financial support by NNL-INFM and by EC through SAMBA project is gratefully acknowledged.

REFERENCES

1. Joachim, C., Gimzewski, J. K., and Aviram, A., 2000, *Nature* **408:**541.
2. Metzger, R. M., Chen, B., Hopfner, U., Lakshmikantham, M. V., Vuillaume, D., Kawai, T., Wu, X., Tachibana, H., Hughes, T. V., Sakurai, H., Baldwin, J. W., Hosch, C., Cava, M. P., Brehmer, L., and Ashwell, G. J., 1997, *J. Am. Chem. Soc*. **119:**10455–10466.
3. Reed, M. A., Chen, J., Rawlett, A. M., Price, D. W., and Tour, J. M., 2001, *Appl. Phys. Lett*. **78:**3735.
4. Roth, S., and Joachim, C., 1997, *Atomic and Molecular Wires*, Kluwer, Dordrecht, Germany.
5. Kwok, K. S., and Ellenbogen, J., 2002, *Materials Today* Vol 5 (2) feb. 2002 pp. 28–37.
6. Ben-Jacob, E., Hermon, Z., and Caspi, S., 1999, *Phys. Lett. A* **263:**199.
7. Lee, I., Lee, J. W., and Greenbaum, E., 1997, *Phys. Rev. Lett*. **79:**3294.
8. Adman, E. T., 1985, *Topics in Molecular and Structural Biology: Metalloproteins* (P. M. Harrison, ed.), Chemie Verlag, Weinheim.
9. Rinaldi, R., Biasco, A., Maruccio, G., Cingolani, R., Alliata, D., Andolfi, L., Facci, P., De Rienzo, F., Di Felice, R., and Molinari, E., 2002, Solid-state molecular rectifier based on self-assembled metalloproteins, *Adv. Mater*. **20:**1453.
10. Rinaldi, R., Biasco, A., Maruccio, G., Arima, V., Visconti, P., Cingolani, R., Facci, P., De Rienzo, F., DiFelice, R., Molinari, E., Ph Verbeet, M., and Canters, G. W., 2003, *Appl. Phys. Lett*. **82:**472.
11. Facci, P., Alliata, D., Andolfi, L., Shnyder, B., and Koetz, R., 2002, *Surf. Sci*. **504:**282.
12. Kuznetsov, A. M., and Ulstrup, J., 1999, *Electron Transfer in Chemistry and Biology. An Introduction to the Theory*; Wiley, Chichester, UK.
13. Facci, P., Alliata, D., Cannistraro, S., 2001, *Ultramicroscopy* **89:**291.

Chapter 3

MECHANICAL CONSEQUENCES OF BIOMOLECULAR GRADIENTS IN BYSSAL THREADS

J. Herbert Waite, James C. Weaver, and Eleonora Vaccaro
Marine Science Institute, University of California, Santa Barbara, CA 93106 USA

Abstract: The hard tissues of living organisms pose serious challenges to architectural integrity. In particular, when hard tissues directly abut softer ones, there is high potential for contact deformation. Mussel byssal threads are an ideal model system for studying how molecules are assembled to mediate the joining of two mechanically divergent tissues at their common interface. The challenge is overcome by the exquisite use of gradients in the block copolymers that comprise the thread.
Key words: Mussel, *Mytilus*, byssal threads, molecular gradient, stiffness gradient

1. INTRODUCTION

Nature has invented fibers with an infinite diversity of structures, properties, and chemistries. Because of their economic importance, the most familiar animal fibers are undoubtedly the protein-based fibers of vertebrate wool/hair and the silks of arthropods, such as the silkworm and spiders, but there are numerous other less known examples. Like wool and silk, the anchoring fibers of marine mussels, for example, once supported a thriving textile industry in the Mediterranean region.[1] Byssal fibers or threads are quite unique in that, unlike other fibers, they are fabricated to be mechanically and chemically distinct at their termini. In the *Mytilus* species, for example, the proximal end is one tenth as stiff and twice as extensible as the distal portion.[2] This material difference is not the result of a sharply imposed transition along the length of the fiber, but rather, appears in a gradual, graded fashion from one end to the other.[3] The following article is devoted

V. Renugopalakrishnan et al. (eds.), Bionanotechnology, 25-38.

to describe this gradient and, more specifically, to attempt to model the effect of this chemical gradient on the mechanical properties of byssal threads.

2. PROPERTIES OF BYSSAL THREADS

2.1 Ultrastructure

The first hint of compositional gradients in byssal threads was revealed by light and later electron microscopy on the byssal threads of *Mytilus edulis* and *M. galloprovincialis*, two closely related species. In these species, thread length ranges from 2 to 6 cm, whereas diameter ranges from 50 to 100 μm.

Examination of the proximal portions both by light and scanning electron microscopy reveal a distinct corrugated or wrinkled morphology, whereas distal surfaces are deeply grooved in the direction of the thread axis[4] (Fig. 1).

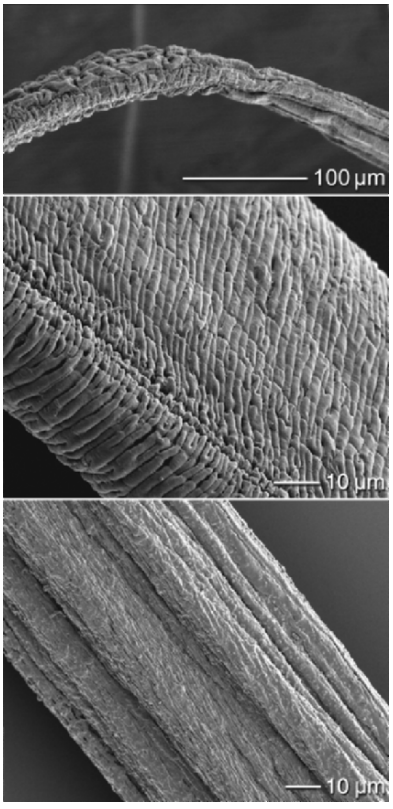


Figure 1. SEM of the transitional region (top) of a byssal thread (50 μm diameter) of *M. galloprovincialis* with enlargements of the distal (right) and proximal (left) portions.

The proximal and distal regions of a mytilid byssal thread, but not the transitional area joining them, have been scrutinized more than any other byssal structure by transmission electron microscopy.[4–6] The distal portion consists of densely fused fibrils (diameter 7–9 nm), all oriented parallel to the thread axis and interspersed with a fractional void volume of about 0.3. The voids have an average diameter of 0.1 μm and a length of 1 μm. In transverse sections, the fibrils appear to be organized into hexagonal and pentagonal arrays with 20 nm center-to-center distances.[7] The proximal portion, in contrast, contains densely packed fibrils near the outer cuticle. Solitary undulating fibrils (diameter 7–9 nm) or small clusters, thereof dominate the remainder of the core. The solitary fibers or fiber clusters are separated by substantial matrix that consists of granular or microfibrillar material. The undulations of solitary fibers have an average pitch and amplitude of about 70 and 100 nm, respectively and linearize as the thread is stretched.[6,7] The large spaces or voids that are distinct in the distal portion are completely absent in the thread's proximal region.

2.2 Mechanics

The gross and microscopic morphological differences between the proximal and distal ends of each byssal thread (Fig. 1) were the first suggestion that there may in fact be unique mechanical properties associated with these two distinct regions. These differences have now been established by several investigations.[8–10] The distal moiety is stronger, stiffer, and superior at damping (70% hysteresis), while the proximal portion is softer and weaker, with a lower but still significant hysteresis at 40% (Fig. 2). This hysteresis is evidence of energy dissipated in cyclic stress–strain tests. Toughness and energy dissipation, both are crucial properties for shock adsorbing materials such as byssal holdfasts. Other protein structures range in hysteresis from 10% in

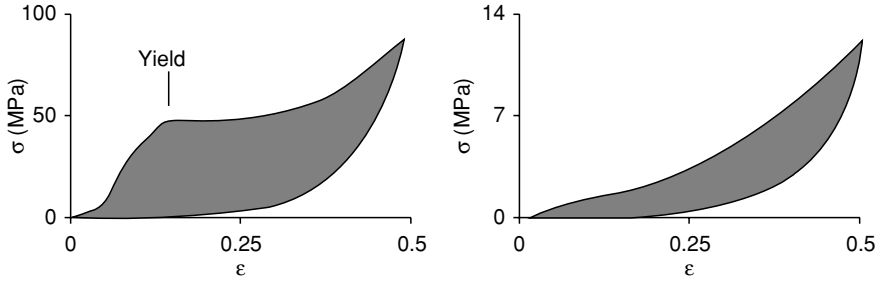


Figure 2. Cyclic extension of the distal (left) and proximal (right) portions of *M. galloprovincialis* byssal threads. Strain rate was 0.1 mm/s.

elastin and tendon[2] to a high of >80% for titin-derived passive tension in skeletal muscle.[11] Major ampullate (MA) and viscid spider silks both have a hysteresis of 65%, which is appropriate given their role in prey capture.[2] High hysteresis of byssal threads is especially important to resist drag and lift forces experienced by mussels during repetitive wave shock experienced in their natural habitat.

The mechanical behavior of byssal threads, in reality, is considerably more complex than presented earlier. Vaccaro *et al.*[12] and Carrington and Gosline[13] demonstrated that when strained beyond its yield point, the distal portion exhibits a dramatic stress softening, i.e., the initial modulus of the second cycle was reduced to about 20% of the modulus in the first cycle. Hysteresis was reduced as well; complete recovery to the modulus of the first cycle was slow (e.g., >24 h), but significant partial recovery occurred within an hour. Ironically, stress softening in individual threads may enhance the toughness in the entire byssus because it distributes recurring loads to all threads. As illustrated in Fig. 3, a mussel has a roughly radial distribution of threads, which when subjected to uniaxial loading can result in stress concentration in the leading thread (#*I*). When thread #*I* was extended beyond its yield point by an applied force, such as a wave, and this was followed by a second force from the same direction, thread #*I* offered significantly resistance to displacement due to prior stress softening (500–80 MPa). Consequently, thread #*I* can be extended with little resistance to loading, and thread #*II*, which still has high initial stiffness, will be recruited into action.

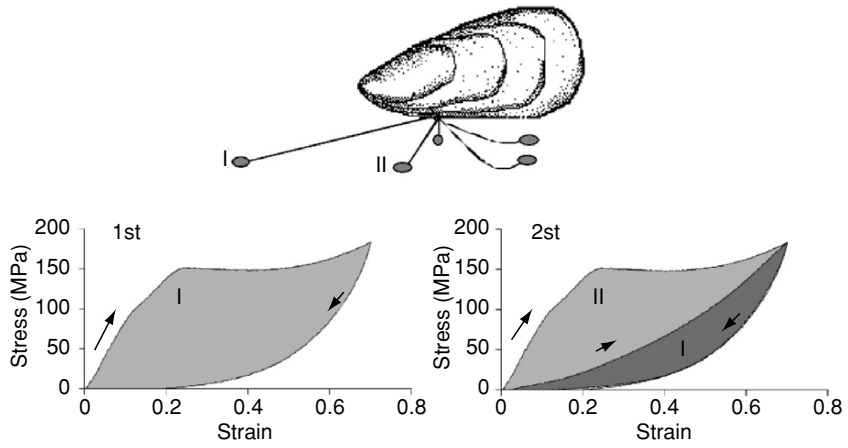


Figure 3. Two consecutive stress–strain cycles of distal byssal threads (I and II) of *M. galloprovincialis* beyond the yield point. Note that when I stress softens II picks up the load.

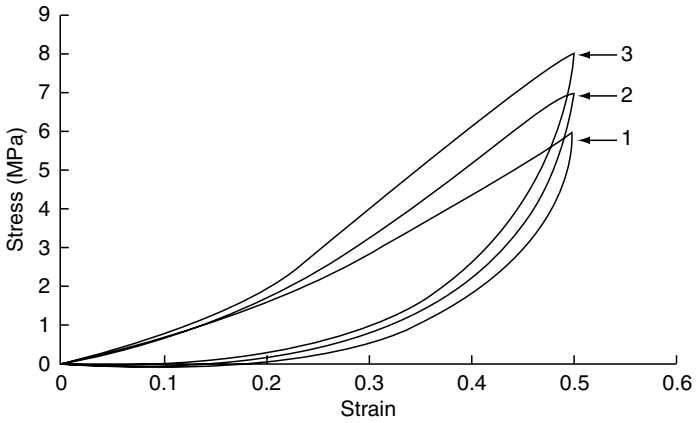


Figure 4. Three consecutive stress–strain cycles (1, 2, and 3) of a proximal byssal thread from *M. galloprovincialis*. Stress maximum occurs at about 15 MPa and is reversible. Data adapted from Sun *et al.* (2001).

Presumably, the process of stress softening followed by new recruitment to loading continues until as many threads as possible contribute to load sharing. Notably, the proximal portion also exhibits stress-dependent changes in stiffness with cyclic loading but in the opposite direction! In this case,[14] there is strain stiffening from an initial modulus of 35 MPa to an asymptotic leveling off at about 60 MPa (Fig. 4), an increase of about 40%. Maximal increase in stress at 0.5 strain is even higher at 150%.

2.3 Biochemistry

Mascolo[15] was the first to report biochemical differences along the length of *Mytilus* byssal threads. After serially cutting these threads into 2-mm segments, the sections were hydrolyzed and amino acid compositions determined. Glycine showed the greatest change from a high of 330 res/1000 at the far-distal end to 240 res/1000 at the far-proximal end; 4-hydroxyproline showed a similar trend, although the magnitude of change was not as great. This approach was recently coupled to a molecular comparison of byssal threads from *M. edulis* and *M. galloprovincialis*, two closely related mussel congeners with similar protein compositions but rather distinct gradient profiles.[16] Taken together, the Gly and Hyp gradients suggest collagens. Qin and Waite[17] identified gradient-specific collagens by isolating two pepsin-resistant collagen fragments; Col-P (from the proximal region) and Col-D (from the distal region), from serial byssal thread sections and demonstrating that the distribution of these along the thread axis was graded. Graded

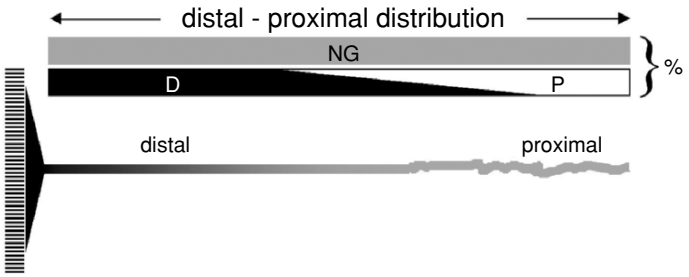


Figure 5. Gradient distribution of preCols-D and preCol-P relative to position along the length of a *M. edulis* byssal thread. There are equally likely additional, yet undiscovered, gradients in the thread matrix proteins. Nominal thread length is 3–6 cm.

distributions of two preCols (D and P) were also evident from Western blots of protein extracted from successive serial sections of byssal threads and mussel feet using specific polyclonal antibodies for the pepsin-resistant collagenous domains of preCol-D and preCol-P[17] as well as *in situ* hybridization with preCol gene-specific probes.[18,19] Because of the routine use of four to six tissue sections for discontinuous analysis, these results do not specify the precise shape but rather the general trend of the gradients. The distribution of preCol-D tapers off in a distal to proximal direction (Fig. 5), whereas preCol-P decreases in a proximal to distal direction, reaching extinction before mid-thread in *Mytilus edulis*. preCol-NG ("no gradient"), in contrast, is uniformly distributed. Evidently, these gradients are produced prior to secretion because they are already detectable in expression libraries prepared from serial foot sections screened with gene-specific probes.[20] In foot tissue, the relative abundance of each preCol mRNA can be precisely titrated using the rRNA EukA as an internal reference.[18,19]

Another gradient involving two variants of preCol-P has been reported to be present in which the proximal region fuses with the stem.[21] Moreover, gradient distributions in some of the matrix proteins of the proximal and distal portions are also suggested by an investigation of ptmp-1, a von Willebrand A domain-type protein that binds collagen with subnanomolar affinity.[22]

3. STRUCTURAL MODELS

As stated previously, byssal threads are unique from most other animal fibers in that they have an order of magnitude differential in mechanical moduli at each end. Thus, at least in terms of stiffness, each byssal thread resembles the fusion of a nylon string with a rubber band. While the combination of two such materials may offer some significant adaptive advantages to

byssal threads, the interface between the nylon and rubber is predicted to exhibit high-residual stresses due to the resulting modulus mismatch. The fact that failure does not occur at this interface with any more frequency than in any other thread portion suggests that the molecular gradient described previously may moderate and/or dissipate the stresses experienced.

The major challenge in modeling the mechanical consequences of the molecular D to P gradient, as well as the uniform distribution of NG in byssal threads is the lack of specific information regarding assembly. High-resolution structural imaging by transmission electron microscopy reveals some general properties of preCol packing in the thread, but important details, such as molecular stagger, head-to-tail orientation, and disposition of disulfide bonds, are not yet known. These have, however, not prevented attempts at modeling, and the ostensible success of some of these at predicting mechanical modulus are noteworthy.

Given their conceptual and experimental simplicity, the Reuss and Voigt models require only two parameters: Young's modulus (E) and volume fraction (V) as given by the Reuss and Voigt equations for serial or parallel arrays, respectively.

$$E_c = E_1 V_1 + E_2 V_2 + \ldots + E_n V_n, \tag{1}$$

$$1/E_c = V_1/E_1 + V_2/E_2 + \ldots + V_n/E_n, \tag{2}$$

where each E denotes a viscoelastic modulus, i.e., in this case, spring plus dashpot, and is adopted from the bulk modulus of a biomaterial homologous to the various component domains in preCol-D, preCol-P, and preCol-NG[23] as detailed later. To begin with, an individual preCol consists of a serial array of domains, i.e., the histidine-rich termini, the flanking domains, the central collagen domain, and the acidic patch. The present treatment ignores the contribution of the acidic patch to the modulus (it is small and common to all transcripts). Similarly, the modulus of the histidine-rich domains is ignored because these cross-linking nodes are also common to all transcripts. Young's moduli for the collagen and flanking domains (silk, elastin, and glycine-rich regions) were obtained from the literature or estimated experimentally (Table 1).

Table 1. The Young's moduli of protein fibers related to various preCol domain motifs. These moduli were applied to the calculation of composite preCol moduli

PreCol domain	Related fiber	Young's modulus (MPa)	Reference
Collagen	Tendon	1500	24
P flanks	Elastin	2	24
D flanks	Dragline silk	10,000	24
NG flanks	Periostracum	150	3

It should be noted that equating the modulus of any given preCol domain with a homologous bulk material is an unproven assumption in that molecules in isolation can behave differently than those packed in bulk.

It is not known whether the blocks or domains in preCol interact in a manner similar to their interaction in the comparable bulk material. Also, the contribution of matrix is unclear. In view of the considerable sequence homology, it seems to be a reasonable assumption at this stage. Taking preCol-D for an example (Fig. 6), there are three relevant block domains with corresponding moduli: the two silk-like flanking blocks together represent about half of the preCol with the remainder consisting of collagen. The composite Reuss equivalent is shown below it. Similar application of the Reuss equation to preCol-P and preCol-NG enables estimation of the composite moduli.

A slightly more complex problem is encountered in modeling how the preCols are related to one another in a three-dimensional assembly. The only available information on this subject is based on ultrastructural studies[7,25] by transmission electron microscopy or deduced from the known dimensions of tropocollagen.[26] What emerges from these data is that preCols are associated as six packs (8–9 nm in diameter when sectioned at the collagen domain), each with a narrow and a flared end, the latter occurring on the N-terminal side and induced by the collagen "kink." It is tempting to further deduce that flared ends are joined to other flared ends, but there are as of now, no data available to substantiate this claim. In fact, the absence of any further information about higher structure compels us to propose two assembly models that are inspired primarily by the demonstration of stochastic gradients in preCol distribution. In *model A*, the uniformly distributed preCol NG alternates axially with D or P, whereas in *model B*, it remains segregated from preCol-D and preCol-P (Fig. 7). The mechanical modeling assumes that the bundles self-assemble to form *continuous* anisotropic fibers in the thread and that there are no covalent cross-links between the fiber bundles. Matrix proteins also appear to be graded along the thread length to the extent that

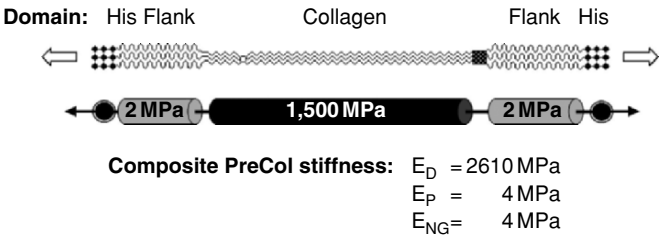


Figure 6. A series spring analog for preCol-P (top). Calculated series (Reuss) composite moduli for all three preCols (below).

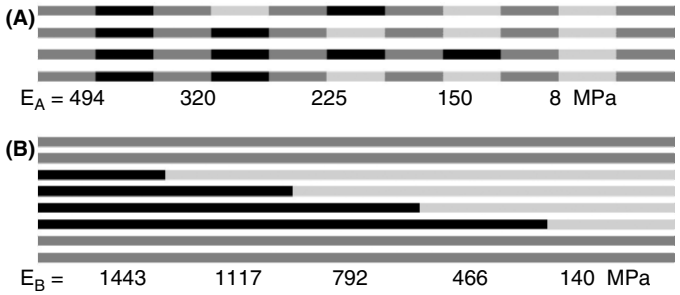


Figure 7. Possible relationship between preCols in the transitional region according to two assembly models A and B. PreCol-D (black), preCol-NG (dark gray), and preCol-P (light gray). Elastic modulus calculated according to the Voigt equation appears below each increment.

at the proximal end they represent more than 30% of the composition. These, however, require additional examination.

Clearly, both models reflect molecular gradients that resemble the graded decrease in stiffness in the transitional region between the distal and proximal portions of the thread. *Model A* has a calculated distal modulus of about 500 MPa; this is very close to the observed distal modulus, namely 500–600 MPa. *Model B*, on the other hand, has a 10-fold differential in modulus between the proximal and distal ends, which also fits the observed 10-fold decrease in the modulus of *Mytilus* byssal threads quite well. In contrast, the differential in *model A* is 60-fold. In each of these calculations, the volume fraction is assumed to be filled with preCols. According to Eqs. (1) and (2), however, the volume fraction must reflect the actual fraction of each domain or preCol type in the composite. An estimation of filled volume fraction was attempted by image analysis of scanned electron micrographs of transverse and longitudinal sections of distal and proximal byssal thread regions.[2,7] In the distal portion, preCol density is moderated by apparent voids, whereas in proximal portions, there is an increase in the proportion of matrix. Estimated volume fraction actually filled by preCol fibers is ca. 40%.[3] With this adjustment, *model A*'s gradient becomes too low when compared to observed stiffnesses, whereas *model B* is quite similar.

It is in byssal stress softening/hardening and recovery, however, that *model B* best fits the observed mechanical behavior. The reason may be the following: in a composite fiber subjected to tension, the stiffer elements break first, thus allowing load transfer to the softer elements. *Model A* only has one kind of microfiber paradigm, that is, one consisting of alternating NG and D or P preCols. When yield occurs in the distal portion, there is nothing to transfer the load to; the break is clean. In contrast, *model B* has stiffer preCol-D containing microfibrils; when these yield, the load is transferred to the

softer intact preCol-NG and further extension occurs. Upon load removal, the material may recover to the extent that it returns to initial length and any broken bonds are repaired.

Self-healing or repair is another respect in which the distal and proximal portions of the thread differ. The distal region softens before recovering, whereas the proximal hardens. Both recover at a rate that is dependent on the degree of extension beyond yield. Using threads from *M. californianus*, Carrington[12] has shown that a strain of 0.35 requires about 6 min for 50% recovery in modulus or strain energy, whereas an extension to 0.70 takes about an hour for the same recovery. Apparently, recovery is significantly slower in *M. edulis* threads. Recovery of the distal portion can be perturbed following treatment with chelators and at various pH. Incubating threads in 100 mM ethylenediamine tetraacetic acid (EDTA) for 24 h, for example, removes roughly half the transition metal content of the threads and reduces the initial modulus by about 50%.[22] Return of EDTA treated threads to natural seawater leads to complete recovery within minutes. In addition, byssal threads incubated in 0.1 M citrate–phosphate solutions at pH of 3, 4, 5, 6, 7, and 8 for 1 h prior to testing for Young's modulus. The results showed a roughly sigmoidal increase in modulus between pH 6 and 7 with a midpoint at about 6.5 (Fig. 8).

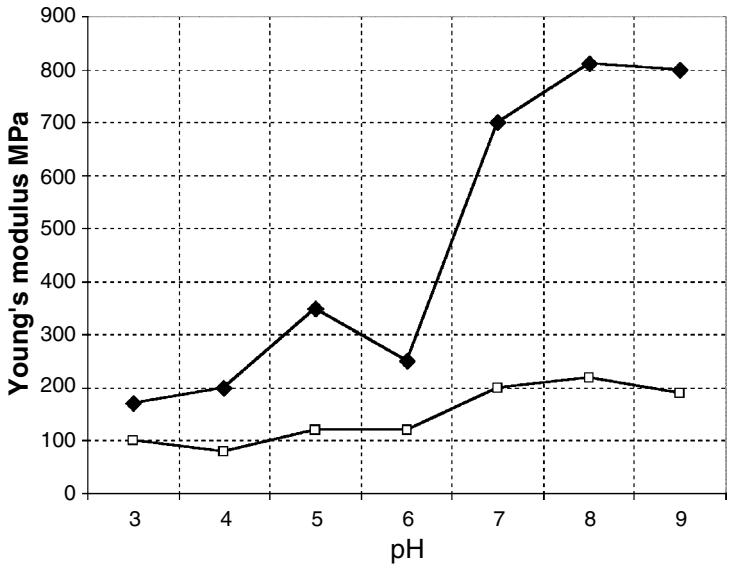


Figure 8. A plot of the relationship between the Young's modulus of the first stress–strain cycle of the distal thread portion and pH (diamonds). Lower trace (boxes) is for the second cycle. pH buffers were prepared from universal phosphate/citrate series.

This is also the midpoint for the ionization of the δ_1 imidazolate nitrogen of histidine.[27] We propose that preCols assemble end-to-end, forming linear polymers in a manner mediated by complexed transition metals such as copper, zinc, and even iron. Histidine, which is concentrated in the N- and C-terminal HIS domains of the preCols, may be the major ligand. In the charged, protonated form, histidine is unable to bind the metal; unprotonated histidine, however, is a strong ligand for several transition metals, including copper (II) and zinc (II). It is by multiple-ligand formation that metals could potentially cross-link the preCols.

There have been some significant technological advances for investigating the specific macromolecular changes that accompany yield and subsequent recovery in tendon collagen.[28] These advances couple synchrotron fiber X-ray diffraction or solid-state nuclear magnetic resonance with mechanical extension. Attempts to apply this technology to byssal threads have failed because these biofibers have insufficient structural order to allow analysis of macromolecular changes at this time.

4. CONCLUSION

A manufactured material with divergent mechanical properties at its ends is referred to as a ''functional gradient material.'' According to this definition, byssal threads definitely belong to this class of materials. Although the fabrication of ceramics with functional gradients has become quite widespread, the design of gradients in organic polymers is still largely a matter exploratory. Byssal threads thus represent a bioinspired paradigm to this end. Mussels have evolved several innovative strategies to construct these gradients: locally controlled synthesis and secretion of preCol mixtures, self-assembling polymer liquid crystals with alternating hard and soft blocks (preCols), a cross-linking strategy that relies on pH, and the addition of transition metals.

ACKNOWLEDGEMENTS

Thanks are due to NIH and NASA URETI for research support.

REFERENCES

1. Yonge, C. M., and Thompson, T. E., 1976, *Living Marine Molluscs*, Collins, London, pp. 171–172.
2. Gosline, J., Lillie, M., Carrington, E., Guerette, P., Ortlepp, C., and Savage, K., 2002, *Phil. Trans. Roy. Soc. Lond.* **357:**121–132.

3. Waite, J. H., Vaccaro, E., Sun, C. J., and Lucas, J. M., 2002, Elastomeric gradients: A hedge against stress concentration in marine holdfasts, *Phil. Trans. Roy. Soc. Lond.* **B357:**143–153.
4. Bairati, A., 1991, The byssus of the mussel *Mytilus* from the molecules to the organ: Functional performance resides in the ultrastructural assembly, in *Form and Function in Zoology* (G. Lanzavecchia and R.Valvassori, eds.), Modena, Mucchi, pp. 163–177.
5. Benedict, C. V., and Waite, J. H., 1986, Ultrastructure and composition of the byssus of *Mytilus edulis* L, *J. Morphol.* **189:**261–270.
6. Vitellaro-Zuccarello, L., De Biasi, S., and Bairati, A., 1983, Ultrastructure of the byssal apparatus of a mussel. V. Localization of the collagenic and elastic components in the threads, *Tissue. Cell* **15:**547–554.
7. Bairati, A., and Vitellaro Zuccarello, L., 1976, The ultrastructure of the byssal apparatus of *Mytilus galloprovincialis*. IV. Observations by transmission electron microscopy, *Cell Tissue Res.* **166:**219–234.
8. Bell, E. C., and Gosline, J. M., 1996, Mechanical design of mussel byssus: Material yield enhances attachment strength, *J. Exp. Biol.* **199:**1005–1017.
9. Smeathers, J. E., and Vincent, J. F. V., 1979, Mechanical properties of mussel byssus threads, *J. Moll. Stud.* **49:**219–230.
10. Price, H. A., 1981, Byssus thread strength in the mussel, *Mytilus edulis*, *J. Zool. Lond.* **194:**245–255.
11. Wang, K., McCarter, R., Wright, J., Beverly, J., and Ramirez-Mitchell, R., 1993, Viscoelasticity of the sarcomere matrix of skeletal muscles, *Biophys. J.* **64:**1161–1177.
12. Vaccaro, E., and Waite, J. H., 2001, Yield and post-yield behavior of mussel thread: A self healing biomolecular material, *Biomacromolecules* **2:**906–911.
13. Carrington, E., and Gosline, J. D., 2004, Mechanical design of mussel byssus: Load cycle and strain rate dependence, *Am. Malacol. Bull.*, **18**: 135–142.
14. Sun, C. J., Vaccaro, E., and Waite, J. H., 2001, Oxidative stress and the mechanical properties of naturally occurring chimeric collagen-containing fibers, *Biophys. J.* **81:**3590–3595.
15. Mascolo, J. M., and Waite, J. H., 1986, Protein gradients in byssal threads of some marine bivalve molluscs, *J. Exp. Zool.* **240:**1–7.
16. Lucas, J., Vaccaro, E., and Waite, J. H., 2002, A molecular, morphometric and mechanical comparison of the structural elements of byssus from *Mytilus edulis* and *M. galloprovincialis, J. Exp. Biol.* **205:**1807–1817.
17. Qin, X. X., and Waite, J. H., 1995, Exotic collagen gradients in the byssus of the mussel *Mytilus edulis, J. Exp. Biol.* **198:**633–644.
18. Waite, J. H., Qin, X. X., and Coyne, K. J., 1998, The peculiar collagens of mussel byssus, *Matrix Biol.* **17:**93–106.
19. Qin, X. X., and Waite, J. H., 1998, A collagenous precursor that may mediate block copolymer gradients in mussel byssal threads, *Proc. Natl. Acad. Sci. USA* **95:**10517–10522.
20. Qin, X. X., Coyne, K. J., and Waite, J. H., 1997, Tough tendons. Mussel byssus has collagen with silk-like domains, *J. Biol. Chem.* **272:**32623–32627.
21. Coyne, K. J., and Waite, J. H., 2000, In search of molecular dovetails in byssus: From the threads to the stem, *J. Exp. Biol.* **203:**1425–1431.
22. Sun, C. J., Lucas, J. M., and Waite, J. H., 2002, Collagen binding matrix proteins from elastomeric extraorganismic byssal fibers, *Biomacromolecules* **3:**1240–1248.
23. Wainwright, S. A., Biggs, W. D., Currey, J. D., and Gosline, J. M., 1976, *Mechanical Design in Organisms*, Princeton University Press.

24. 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:**3292–3305.
25. Vitellaro-Zuccarello, L., 1980, The collagen gland of *Mytilus galloprovincialis:* An ultrastructural and cytochemical study on secretory granules, *J. Ultrastruct. Res.* **73:**135–147.
26. Nimni, M., 1980, The molecular organization of collagen and its role in determining the biophysical properties of the connective tissues, *Biorheology* **17:**51–82.
27. Creighton, T. E., 1992, *Proteins: Structures and Molecular Properties*, 2nd edn., Freeman, New York.
28. Fratzl, P., Misof, K., Zizak, I., Rapp, G., Amenitsch, H., and Bernstorff, S., 1997, Fibrillar structure and mechanical properties of collagen, *J. Struct. Biol.* **122:**119–122.

Chapter 4

BACTERIORHODOPSIN-BASED 3D OPTICAL MEMORY

Bangwei Xi[1], Kevin J. Wise[2], Jeffrey A. Stuart[1], and Robert R. Birge[2]
[1]*W.M. Keck Center for Molecular Electronics and Department of Chemistry, Syracuse University, 111 College Place, Syracuse, NY 13244-4100, USA*
[2]*Department of Chemistry and of Molecular and Cell Biology, University of Connecticut, 55 North Eagleville Road, Storrs, CT 06269, USA*

Abstract: As semiconductor devices hit their physical limits, molecular electronics offers the potential for alternative technologies and architectures for the computer industry. Bacteriorhodopsin (BR), a photoactive protein, has long found applications as the active element of optical and photovoltaic devices. The protein is characterized by a unique photocyclic response to light, remarkable stability, cyclicity on the order of 10^6, and ease of production and processing. By taking advantage of a branch off the main photocycle, BR can be used as the active medium in optical three-dimensional (3D) memories. A sequential two-photon process is employed to drive the protein into a long-lived branched photoproduct (the Q-state). Parallel read and write processes can be realized without disturbing the data outside of the irradiated volume, making the memory largely nondestructive. In this review we will discuss the architecture of the branched-photocycle memory, our current progress in prototype development, and the challenges the technology currently faces. In addition, we will examines methods by which the protein can be optimized for wider application in biomolecular electronics.

1. INTRODUCTION

With the unprecedented development of the computer industry over the past 30 years, silicon-based technologies will probably reach their practical limit within the next few decades, i.e., improvements in feature size and device speed will no longer be possible. According to Moore's law, semiconductor device size will probably approach the molecular domain around 2030.[1] Thus,

V. Renugopalakrishnan et al. (eds.), Bionanotechnology, 39-60.

new technologies become increasingly important for both practical and economic considerations, and researchers are faced with the challenge of devising new technologies to circumvent conventional architectures. Molecular electronics, and specifically biomolecular electronics, has the potential to offer unique solutions to problems in an array of areas, including memory storage, artificial intelligence, optical processing, micro-electronics, and other applications where size and functionality require parallel optimization, i.e., reduction in size coupled with increased functionality.[2–5]

Molecular electronics can be loosely defined as the utilization of small molecules or nanoparticles as active components of electronic, optoelectronic, and sensing devices, including the encoding, manipulation, and retrieval of data at the molecular scale. The science of molecular electronics is the exploration of molecular assemblies for electronic and optoelectronic applications. These systems must be both scaleable and capable of integration into larger devices in an efficient and practical manner. An input to the system, either in the form of an applied current/voltage or photon absorption, must result in a defined and predictable response. At the level of small individual molecules, complex functionalities can only be obtained by the combinations of individual component molecules, i.e., nano-circuitry, logic gates, etc. However, interrogation becomes an issue at this scale; how can the state of an individual molecule be uniquely determined? Without extraordinarily complex experiments designed to examine a single, isolated molecule, this task becomes problematic at best. It is far easier to interrogate populations, but at the expense of device sophistication or signal complexity. Herein lies but one of the challenges of molecular electronics. This is not to say that applications involving populations of molecules cannot be effective, only that there is a limit in device complexity and signal inputs and outputs. Liquid crystal displays (LCD) provide an excellent example of a technology based on molecular electronics that has proven to be an extremely successful commercial application.

Biomolecular electronics, a sub-discipline of molecular electronics, has the potential to offer solutions to problems that conventional semiconductor technologies cannot address. In short, molecules of biological origin comprise a set of properties not easily achieved through semiconductors. These advantages derive in a large part from the natural selection process (nature has solved problems similar to those encountered in carrying out computer logic, switching, and data manipulative functions) and the ability of self-assembly.[6] Biologically based materials can conduct current, produce photovoltaic signals, behave as molecular machines, and exhibit photochromism, just to name a few. In fact, proteins and other biomolecules pose a unique solution to the problem of size versus function, in that they represent a class of extremely sophisticated molecules that perform specific functions under what can be a fairly broad range of conditions. And it is the goal of biomolecular electronics

to harness these biologically derived functions for non-biological applications. As such, biomolecular electronics represents a largely unexplored route toward new architectures for electronic switching, information storage, communications, optoelectronics, and a multitude of other device applications.

Photochromic proteins are better suited than most proteins for device architectures from a number of standpoints. The presence of an internal chromophore provides a constant probe of protein structure and function; in addition, the chromophore facilitates light to chemical energy transduction, thereby providing a convenient means of interrogation. A common property among many photochromic proteins is the presence of multiple states that are accessed upon exposure to light—examples include photosynthetic reaction centers, photoactive yellow protein, phytochrome, and bacteriorhodopsin. These states often form the basis for the protein's potential in biomolecular electronic architectures. Many photoactive proteins are also capable of exhibiting photovoltaic effects, thereby allowing multiple modes of state interrogation. Of the proteins listed above, BR has historically been the most actively investigated as the functional element in a wide variety of applications, including spatial light modulators, optical random access memories, holographic interferometers, artificial retinas, holographic associative processors, and 3D optical memories. Here, we will focus on the architecture of the BR-based 3D optical memory currently being developed at the W. M. Keck Center for Molecular Electronics at Syracuse University. Progress will be discussed with respect to prototype development, as well as the challenges that must be overcome before the technology is commercially viable.

2. BACTERIORHODOPSIN

Bacteriorhodopsin (Fig. 1) is a 26 kDa photosynthetic protein isolated from the archae-bacterium *Halobacterium salinarum*.[7] The biophysical and biochemical properties of BR have been studied intensively in recent years, making it one of the most thoroughly described proteins in the literature (for recent reviews see Refs. [8–15]). Upon the absorption of light, BR undergoes a complex photocycle, which is coupled to proton translocation across the cell membrane, thereby acidifying the extracellular medium. The all-*trans* retinal chromophore responsible for the absorption of light is bound through a protonated Schiff base to Lys-216. The photocycle (Figs. 1b and 2) can be represented by bR $\leftrightarrow$ K $\leftrightarrow$ L $\leftrightarrow$ M $\leftrightarrow$ N $\leftrightarrow$ O $\rightarrow$ bR, where bR represents the resting state of the protein (i.e., prior to light absorption), K represents the primary photoproduct, and the remaining states represent thermally driven intermediates that are involved in proton translocation.[16–19] The photocycle is initiated with a light-driven all-*trans* to 13-*cis* isomerization of the chromophore during the bR $\rightarrow$ K

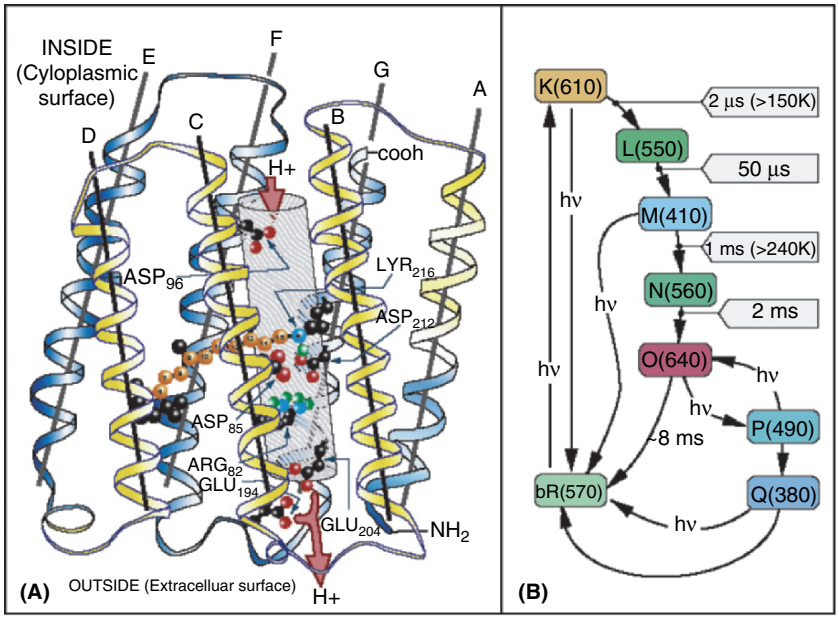


Figure 1. (A) Schematic of bacteriorhodopsin illustrating certain key amino acids and the purported path of the proton pump. The all-*trans* retinal chromophore traverses the binding pocket roughly perpendicular to the membrane normal and the alpha helices. (B) The bacteriorhodopsin photocycle, including the branched photocycle originating at the O-state. Absorption maxima in nanometer are shown in parenthesis for each intermediate.

transition. Proton release occurs during the L → M conversion, which includes the deprotonation of the Schiff base mediated by ASP_{85} and the release of the proton into the medium. During the M → N transition, the Schiff base is reprotonated by the internal proton donor ASP_{96}, followed by reprotonation of ASP_{96} from the medium during the N → O transition. The final step (O → bR) involves the transfer of proton from ASP_{85} to an acceptor group. Those intermediates and the bR resting state make up the main photocycle. In the primary photochemical event, the chromophore is excited to produce a Franck–Condon excited state. This process produces a large shift in electron density, which is equivalent to moving an electron ~2.5 Å down the polyene chain toward the nitrogen atom, and leads to a repulsive electrostatic interaction between the chromophore and two nearby negative aspartic acid residues (ASP_{85} and ASP_{212}). This interaction is the driving force for the photoisomerization about the 13-*cis* double bond to generate the primary photochemistry,[3,20] which is a subpicosecond phenomenon with a very small barrier for isomerization.[21] The energy stored in the primary event is around 48 kJ/mol.[22,23] The photoisomerization also generates a photovoltaic signal due to

the motion of the protonated Schiff base away from the negative counterions. This signal is sufficient to activate charge-sensitive semiconductor detectors.

There are several properties that make bacteriorhodopsin an appealing candidate as an active element in device applications. However, two characteristics are of primary importance: the unique manner in which BR responds to light and the protein's remarkable stability. *H. salinarum* evolved to survive in environments characterized by extreme conditions in temperature, light flux, and ionic strength—this organism is nearly ubiquitous in salt marshes world-wide and other environments where the concentration of NaCl can exceed 4 M. The temperatures in these environments routinely rise to 60°C or better, resulting in nearly unsurpassed thermal stability as compared to most proteins. Furthermore, the organism is often found at or near the surface where the actinic light flux is the greatest—the organism's ability to cope physiologically under continuous heavy illumination may be due in part to BR and its arrangement in the cell membrane. BR occurs in the membrane as a 2D hexagonal array of trimers and is often described as semi-crystalline in nature. Areas of the cell membrane where this arrangement is found are referred to as the purple membrane (PM) and consist of roughly 75% protein and 25% lipid. Purple membrane patches are isolated from the organism most easily by osmotic lysis, digestion with DNAse I, and differential centrifugation. The resulting membrane preparation is stable, and for most applications further purification is not needed (i.e., the solublized protein is seldom isolated).

Other properties that make this protein an optimal candidate for biomolecular electronics include high cyclicity (the number of times it can be photochemically cycled prior to statistical degradation, measured at $>10^6$), ease of production and isolation in large quantities, and the ability to easily modify the protein and process it into different forms. Furthermore, manipulation of the various photocycle states can be done completely with light. Bacteriorhodopsin has proven to be an exception to the common sense notion that biological materials are too fragile for use in computer architecture and other devices applications.[4]

2.1 The Bacteriorhodopsin Branched Photocycle

The photocycle illustrated in Fig. 2 includes a side-cycle off of the O-state, referred to as the branched photocycle. First described by Popp and coworkers in 1993,[24] and later by Gillespie *et al.*,[25] the branched-photocycle consists of at least two states, denoted as P and Q. Upon exposure of the O-state to red light, a small percentage of the excited protein population will be driven into P-state, the first intermediate of the branched photocycle.[24] The P-state is characterized by a 9-*cis* chromophore, which appears not to be stable in the binding site due to steric interactions between the C9 and C13 methyl groups of

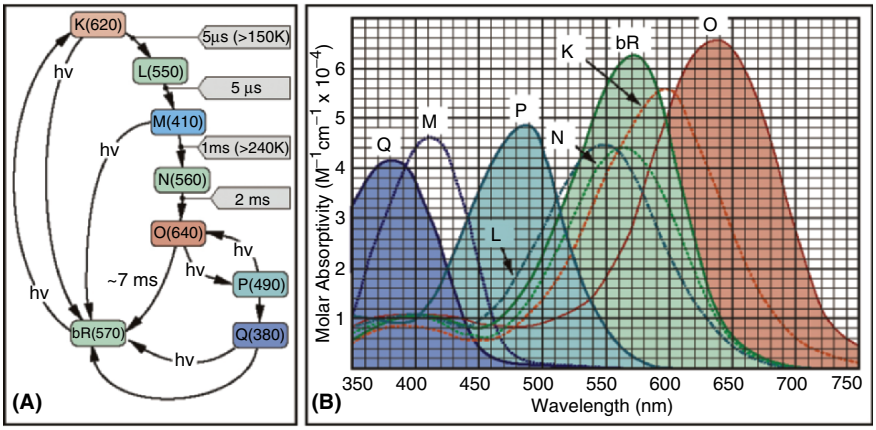


Figure 2. A simplified model of the light-adapted BR photocycle (A) and the electronic absorption spectra of selected intermediates in the photocycle (B). The height of the symbols in (A) is representative of the relative free energy of the intermediates, and the key photochemical transformations relevant to device applications are shown. Note that not all of the intermediates are shown, and that there are two species of M (M1 and M2), but only one is shown for convenience. (M1 and M2 have virtually identical absorption spectra.) The branching reactions involving P and Q that are utilized for storing data in the branched-photocycle volumetric memory are shown in the lower right-hand corner of diagram (A).

the chromophore and nearby amino acid residues. Driven by these steric constraints (i.e., to relieve the strain within the binding site), the Schiff base bond that connects the chromophore to Lys 216 hydrolyzes, resulting in a further blue shifted thermally stable product denoted as the Q-state. The Q-state has an unbound 9-*cis* chromophore trapped within the binding site. The result of this unique combination is a very stable state with an exceptional lifetime, on the order of several years. Exposure to blue light drives it directly back to the bR resting state. To date, a stable long-lived photochromic state in BR has remained elusive, thereby limiting its viability for device applications. The Q-state, however, has become a particularly valuable candidate for both memory storage and holography.

3. THE BRANCHED-PHOTOCYCLE 3D OPTICAL MEMORY ARCHITECTURE

One of the conceivable future directions for memory storage is to make the transition from two to three dimensions. Conventional semiconductor-based technologies face several imposing challenges upon attempting to bridge that gap, including three-dimensional access of data and heat dissipation. There are

pseudo-three-dimensional architectures available, but they largely consist of parallel versions of current technologies (i.e., stacked disk architectures with multiple heads that allow parallel access to multiple discs). The only truly three-dimensional designs are optical, which have the advantage of optically accessing data in a volumetric homogeneous medium. Making the transition to three dimensions has several conceivable advantages, including a substantial increase in the potential storage density and the ability to operate in parallel, thereby introducing a large gain in data throughput. The various active optical elements of these media that have been explored include organics, such as spiropyrans[26] and fulgides,[27] inorganic materials, such as lithium niobate,[28–30] and finally, the photoactive protein bacteriorhodopsin.[2,3,31–35] Several criteria must be satisfied when considering volumetric (non-holographic) optical memory operation: (1) the active component of the optical memory medium must be capable of reversible photoconversion between two stable, long-lasting states, (2) previously stored data must not be disturbed in any way, i.e., the long-term storage state should not respond in a permanent manner to wavelengths of light used in the writing or reading process, (3) the active component must be capable of undergoing a large number of routine write–read–erase operations (high cyclicity), (4) the matrix material that encapsulates the active component must be optically transparent with low light-scattering potential, and (5) the matrix material must be stable over a long time.[34]

The architecture of the BR-based three-dimensional memory is derived from the branched photocycle of bacteriorhodopsin, first described by Popp *et al.*,[24] and more recently by Gillespie *et al.*[25] In the branched-photocycle architecture, the resting state (bR) is assigned as a binary 0, and both the P- and Q-states are collectively assigned to a binary 1. The architecture is outlined schematically in Fig. 3. A key to the architecture is that writing and reading require two separate photon absorption events: the first for exciting the protein, in effect priming it for binary conversion, and the second for actual binary conversion (after the appropriate temporal separation). Only this combination will result in forcing the protein into the branched photocycle—in this way, the protein acts as a latched AND gate. Figure 4 illustrates implementation of the architecture in three dimensions. To initiate the writing process, the paging laser is fired to activate the BR photocycle, and after 2 ms, the writing laser is used to drive the protein into the branched photocycle.

The writing operation is initiated in the memory by a process referred to as paging. A thin slice of the cube is selected optically by a 630 nm laser (paging laser) beam to create the page. Around 2 ms later (depending on temperature), when protein in the paged region has reached the O-state, another laser (690 nm data laser) is fired perpendicularly to the paging laser. Using a spatial light modulator, data is encoded into the beam before it reaches the memory medium and intersects with the previously selected page. Upon intersecting the

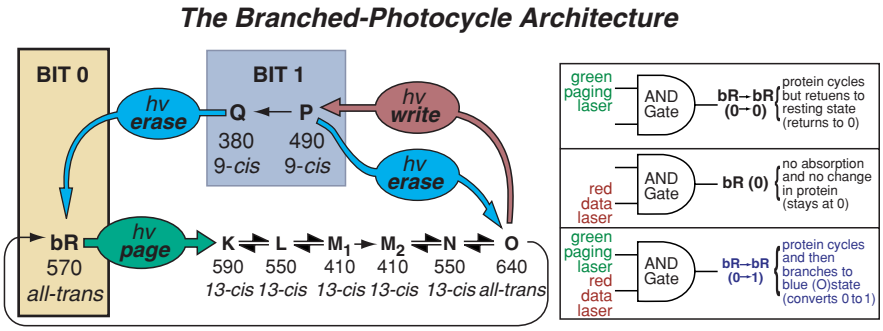


Figure 3. The native photocycle includes the resting state (bR) and the intermediates K, L, M, N, and O. The last intermediate in the photocycle, O, will convert to a blue-shifted photoproduct, P, if it absorbs red light. The P-state relaxes to form Q, a long-lived (>5 years) intermediate. Let bR represent bit 0 and P and Q represent bit 1. Data can be written into any location within a solid polymer matrix containing the protein by using a combination of paging (the green arrow) followed 2 m later by orthogonal writing (the red arrow). The protein operates like an optical AND gate, that is, data are written if, and only if, both of the input conditions are satisfied (see diagram above right).

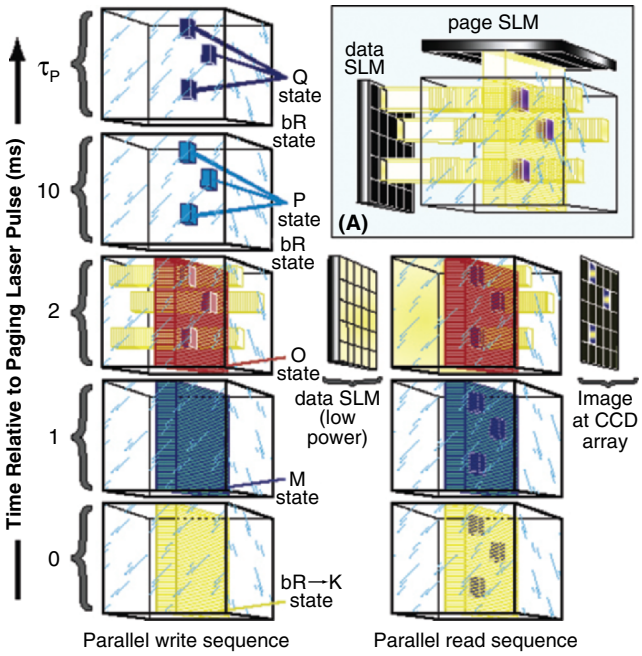


Figure 4. Parallel write sequence (left) and parallel read sequence (right) for the hypothetical storage and retrieval of three binary ones (22 binary zeros) in a page of memory using the branched-photocycle architecture (see text). The inset details the concept of orthogonal accessing, a method by which any volume in a 3D medium can be accessed.

selected page, only the protein in the doubly irradiated regions defined by the spatial light modulator will be driven into the branched photocycle, thereby driving the protein from a binary 0 to a binary 1. All BR outside both the doubly irradiated areas and the paged region will remain unchanged. The reading process is very similar to writing, in that it is also initiated by a paging laser, and after the bulk of the protein reaches the O-state, the data laser is fired. This time, however, no spatial information is encoded in the beam by the SLM, which is used instead to reduce beam intensity. The result is that in the reading process, the data laser will "illuminate" the page by introducing contrast between the binary zeros and ones. Because the absorbance maxima of the P and Q-states are blue shifted with respect to the wavelength of the data laser, previously written data will be, in effect, transparent to the data beam. The resulting image is then recorded by a charge-coupled device (CCD)—a differential read is done by comparing to a data image recorded prior to firing the data laser. An erasing operation can be accomplished by driving the 9-*cis* O-state back to the all-*trans* resting state with a blue laser (380–480 nm). Lasers in this spectral region are at this point prohibitively expensive, so only global erasures are possible at this time; until such lasers are within reach, the memory system is operated as a write once read many (WORM) device.

The advantages to the bacteriorhodopsin-based three-dimensional optical memory are several fold and include storage density (data resolution is limited by the resolution of the SLM, the diffraction limit, and the light scattering potential of the polymer matrix), parallel operation, and resistance to the damaging effects of radiation. The latter is of strong interest to space and satellite applications, where lead is needed to shield semiconductor-based storage media. Storage density in volumetric matrices has the potential to achieve improvements on the order of three orders of magnitude over conventional two disc-based architectures,[35] although the limitations detailed previously lower the expectations to about a 300-fold improvement. The ability to act in parallel is a tremendous advantage for optical data throughput; although the simplistic example in Fig. 4 shows three binary ones simultaneously written, the 5×5 SLM in the illustration also enables, in effect, the simultaneously creation of 22 binary zeros, to total 25 pieces of data written in one routine operation. And as mentioned above, data resolution is limited at least in part by the resolution of the spatial light modulator; typical resolutions on the order of 1048×1048 are available, which would make possible parallel data throughput on the order of a megabit per routine operation.

3.1 Prototype Development

The basic optical layout for a BR-branched-photocycle prototype is shown in Fig. 5. Several prototypes have been developed with the goal of

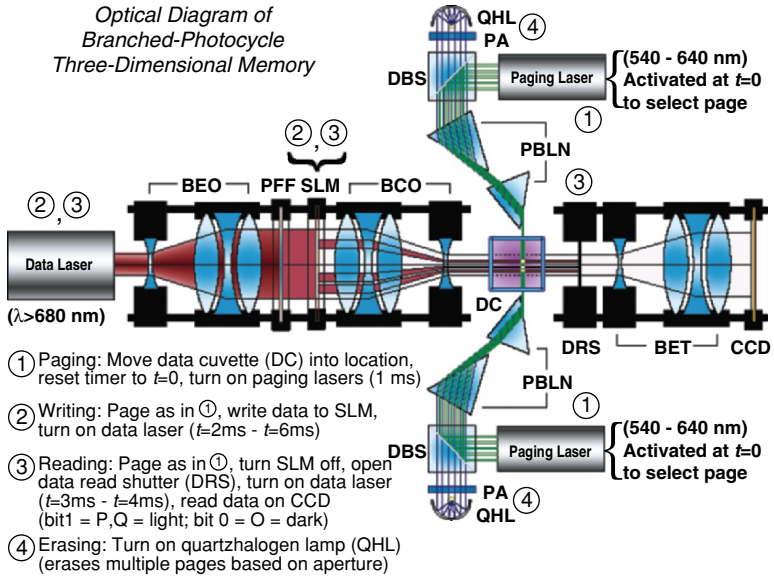


Figure 5. Optical layout schematic illustrating the key components of the 3D optical memory based on bacteriorhodopsin.

demonstrating proof-of-principle and as in-house tools for materials development. A pair of prototypes was constructed purely for evaluation of genetically and chemically modified proteins; as will be discussed below, limitations in the wild-type protein's branched-photocycle efficiency preclude its use in a commercially viable device, necessitating efforts to produce genetic and chemical bacteriorhodopsin variants better suited to memory storage. Because of the aforementioned limitations, the current series of prototypes is incapable of parallel operation and writes or reads only one bit at a time. The latest prototype was developed in collaboration with Critical Link Inc., of Syracuse NY, a custom electronics engineering and prototyping firm. The unit is capable of storing 3000 bits in the BR-based media, with lasers that are coupled with fiber optics to the media (see Fig. 6). Our efforts in protein optimization are outlined below.

4. OPTIMIZING THE BRANCHED PHOTOCYCLE

Perhaps the largest remaining challenge facing the branched-photocycle memory is the efficiency with which the branched-photocycle can be accessed. The quantum efficiency of the O to P transition is estimated to be as low as 10^{-3} to 10^{-4}, although in-house observations have yielded values on the order

Figure 6. Two views of the current branched-photocycle 3D optical memory based on BR. Top: Overall system view, measuring $\sim 31\,\text{cm} \times 38\,\text{cm} \times 23\,\text{cm}$. The memory media is inserted through an opening in the unit's top. Bottom: Close-up view of the optical platform, with a memory gel in the foreground. The gel is inserted into a central holder, mounted on an x–y–z translation stage. The lasers are fiber optically coupled to the holder.

of a few percent. The larger issue is one of unwanted photochemistry, where photoconversion is achieved along the axes of irradiation during paging and writing operations, i.e., from single photon rather than sequential dual photon processes. The source of this single photon route into the P and Q-states is thought to be due to a well-described direct path from the deionized form of BR, the blue membrane.[36] Illumination of the blue membrane with red light results in the formation of the purported pink membrane, which is considered to be synonymous with the P-state.[37–42] Because the blue membrane is always present in some equilibrium concentration, the direct route to the branched photocycle can never be completely suppressed in the wild-type protein (see Fig. 7).

Bacteriorhodopsin has long been touted as a strong candidate for biomolecular electronic applications, for memory storage as described above, for holographic applications, artificial retinas, and spatial light modulators, as well as other applications. A particularly exciting and new set of applications currently being developed involves the use of BR as a photochromic ink for use in security and verification (see the contribution by Markus Wolperdinger in these proceedings). Yet to date, only one commercial device has resulted, the holographic interferometer developed by Hampp and colleagues of Munich Innovative Biomaterials.[43] No devices have been marketed based on the wild-type protein—modification, either by chemical or genetic means, seems

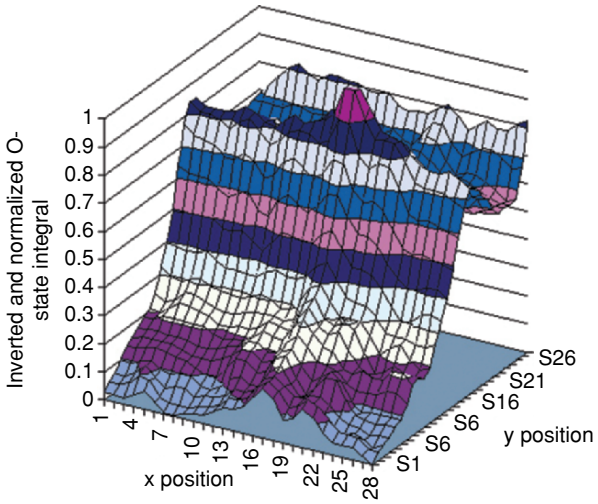


Figure 7. Contour plot illustrating the existence of a bit written in three dimensions, using the BR mutant F208N. Unwanted photochemistry can clearly be seen along both axes of irradiation. Such unwanted photochemistry ultimately degrades both bit resolution and storage density.

to be necessary to produce commercially viable technologies. Although a powerful technique, chemical modification has limitations and will probably not be successful alone (see Refs. [3,33,44,45] for a review of some of these techniques). Genetic engineering, however, is an exceptionally powerful technique enabling control at the level of individual amino acid residues.

As applied to the branched-photocycle memory, two genetic approaches toward optimization are conceivable: optimization of the O-state magnitude and optimization of the O to P/Q transition quantum efficiency—both routes toward greater efficiency are essential for a faster and more efficient BR-based optical memory. The former approach, optimizing the magnitude of the O-state, is based on a simple mass-transfer premise—the more O-state that is produced, the more that can be driven into the branched photocycle. The latter approach, optimizing the efficiency of the O to P transition, poses a more difficult challenge, as the transition is not well understood on a mechanistic basis. The first attempt to optimize the production of the O-state was made through site-directed mutagenesis, by targeting specific amino acids that are known to affect its rise and decay. And enough is known about the photocycle and the O-state that specific site-directed mutants were possible. Several such mutants were constructed and evaluated. Although the results were promising, it became clear that this approach alone would not ultimately be successful; despite considerable understanding of protein structure and function (especially in the case of BR), science still does not have an adequate ability to predict *de novo* the effects of most individual amino acid substitutions, especially when they are far removed from the protein's active site. Therefore, site-directed mutagenesis is not a viable route toward a specific goal in situations that lack sufficient mechanistic understanding. Fortunately, an alternative approach exists that bypasses these limitations; random mutagenesis and directed evolution provide techniques by which a targeted goal can be achieved without specific knowledge of the changes being made to the protein.

Random mutagenesis and directed evolution are, in fact, based on the premise that we cannot predict the nature of many genetic mutations—in essence, through the addition of some sort of random mutagenic factor, mutations are introduced to the genetic code with unpredictable consequences, resulting in a library of genetically altered bacterial strains. The mutant proteins ultimately isolated from those strains are screened and evaluated with respect to the properties of interest (the targeted goals); the strain producing the most efficient mutant protein can then be selected to act as parent to the next generation of random genetic progeny. This process can be repeated as many times as necessary, thereby building upon successive improvements in an incremental fashion, until the desired goal is met. A strength of this approach is that the researcher is not biased by a perceived

understanding of protein structure–function relationships, i.e., mutations are allowed that would undoubtedly be overlooked in a more systematic approach. This procedure is conceptually illustrated in Fig. 8 and more specifically in Fig. 9, which includes the results of our first round of random mutagenesis. An effective screening protocol is key to the success of both random mutagenesis and directed evolution; the ideal screening protocol can identify potential candidate proteins *in vivo*, thereby eliminating the need for time consuming and tedious cell culture and protein isolation procedures. As illustrated in Fig. 10, screening for both absorption maxima and O-state kinetics can be done with little difficulty *in vivo*. To date, all of the studies are focused on O-state optimization—the alternative approach of targeting the O to P quantum efficiency is far more complicated and poses problems with respect to an efficient screening process.

A technique that essentially combines site-directed and random mutagenesis techniques is called semi-random mutagenesis. Semi-random mutagenesis involves mutation of a local region of approximately 15 amino acids. This method is less specific than site-directed mutagenesis but provides the opportunity to partition the gene into 17 manageable segments. Details on how

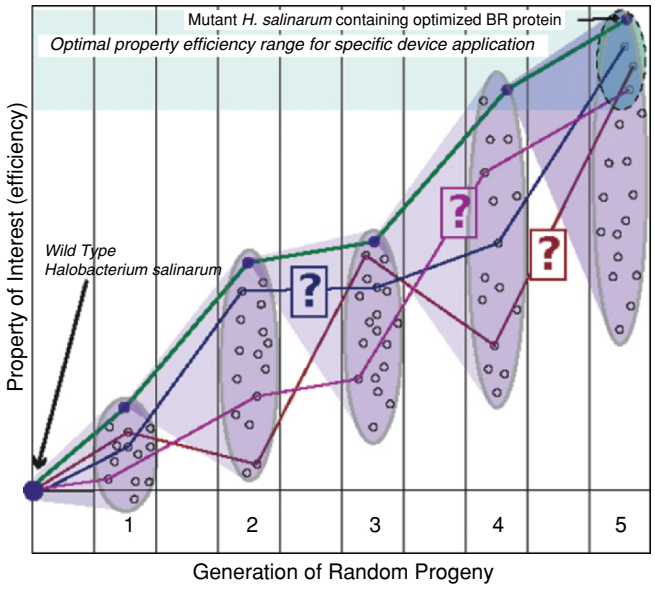


Figure 8. Schematic illustration of directed evolution of a bacterial protein toward a pre-defined goal. The bold green line indicates the route selected by the researcher, where, based on screening, the strain containing the most efficient mutant protein of each generation was chosen to parent the next. However, as shown by the question marks, other synergistic routes may be possible but never discovered.

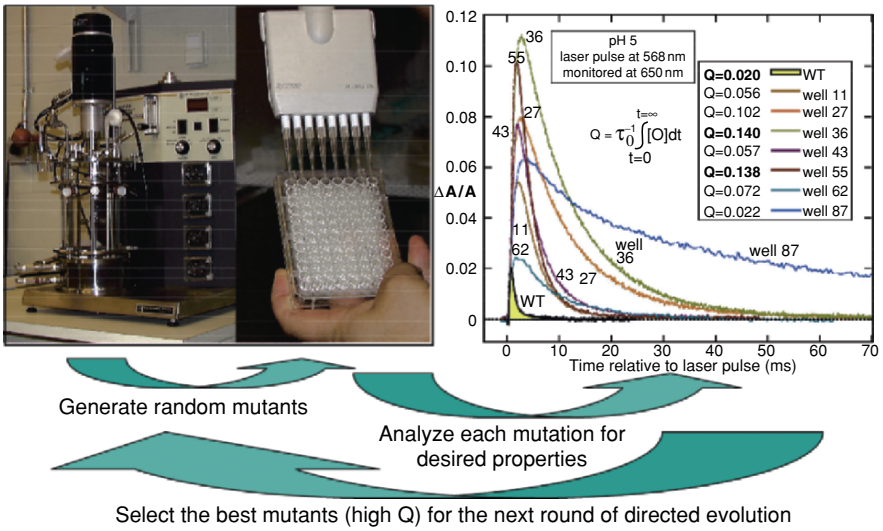


Figure 9. Another illustration of directed evolution as applied to selection of the O-state lifetime from the experimental point-of-view. Natural evolution occurs as a function of selective pressures placed on a species by its environment. In this case, selective pressure is applied by the researcher, in the form of the Q factor, a quality factor that assigns significance to mutant proteins with an O-state characterized by a fast rise time, high yield, and a fast decay. The mutant that has the best Q factor is selected to be the parent for the next round of random genetic progeny. Individual strains are brought up and screened in 96-well plates, prior to selection for large-scale fermenter culture. O-state mutant kinetics derived from a number of mutant strains are shown on the right along with their associated Q values.

these mutants are created can be found in Ref. [46]. This medium-throughput method is used to generate and screen small amounts of protein in 96-well plate format.

There are two approaches to directed evolution, denoted as Type I and Type II. Both are based on the concept illustrated in Fig. 8, the difference lying in the chosen screening method. Type I focuses on evaluating small amounts of purified protein with respect to enzymatic activity, kinetics, and other standard benchmarks. Mutant strains must be cultured in sufficient quantity to isolate the protein of interest. Type II has the potential to be much more efficient, by evaluating the protein *in vivo*—however, not all proteins are amenable to this approach, which depends on the *in vivo* protein concentration and the sensitivity of the screening method. As such, photoactive proteins are better candidates for Type II directed evolution because the protein can be interrogated and monitored with light, making possible a much faster and more efficient screening process. Both Type I and Type II rely on iterative

screening/selection to identify proteins with optimized properties, the best of which are used as a template for the next round of mutation. Type II methodologies could be implemented using a modified flow cytometer to screen and select individual cells. These techniques are currently being developed and are discussed at greater length in Ref. [46]. Just as for the semi-random techniques discussed above, the most serious limitation thus far is the selection and screening process, which can be both time-consuming and challenging depending upon the nature of the enhanced property. Progress along these lines is being made by researchers at the W. M. Keck Center for Molecular Electronics at Syracuse University, where the ability to detect various photocycle intermediates *in vivo* has already been demonstrated (M, O, see Fig. 10). After just one round of directed evolution, a number of new variants are exhibiting improved properties (Fig. 11). These new mutants are currently being analyzed with respect to O-state production and branched-photocycle efficiency.

5. FUTURE DIRECTIONS

The library of proteins available for biomolecular electronics is currently limited to just a few, including bacteriorhodopsin, photosynthetic reaction centers, photoactive yellow protein, and a few other, predominantly photoactive, proteins. Industrial scale utilization of proteins is largely limited to enzymes employed for their catalytic activity in diverse applications ranging from chemical processing and synthesis to stain degradation agents in laundry detergents. The field of biomolecular electronics is small by comparison, and has not as of yet demonstrated its full potential. However, at the time of this writing, there are at least four companies dedicated to photoactive protein-based devices, with several larger companies diversifying along similar lines. It is clear that the next 5 to 10 years will be pivotal for this field.

The bacterial rhodopsins have garnered a great deal of interest as active elements in device applications because of their unique response to photoactivation and the potential for multiple methods of state interrogation, including optical and photovoltaic. Of this group, bacteriorhodopsin has proven to be the most successful because of an extremely well-defined response to light (resulting in both optical and voltaic signals), remarkable stability, and ease of isolation, modification, and materials processing. As mentioned above, a number of devices have been devised that employ BR, and several have been successfully built. Included are holographic interferometer systems,[43] artificial retinas,[47,48] and optical memory systems (2D and 3D holographic, binary photonic).[3,6,32,34,49–51] The application described above represents the

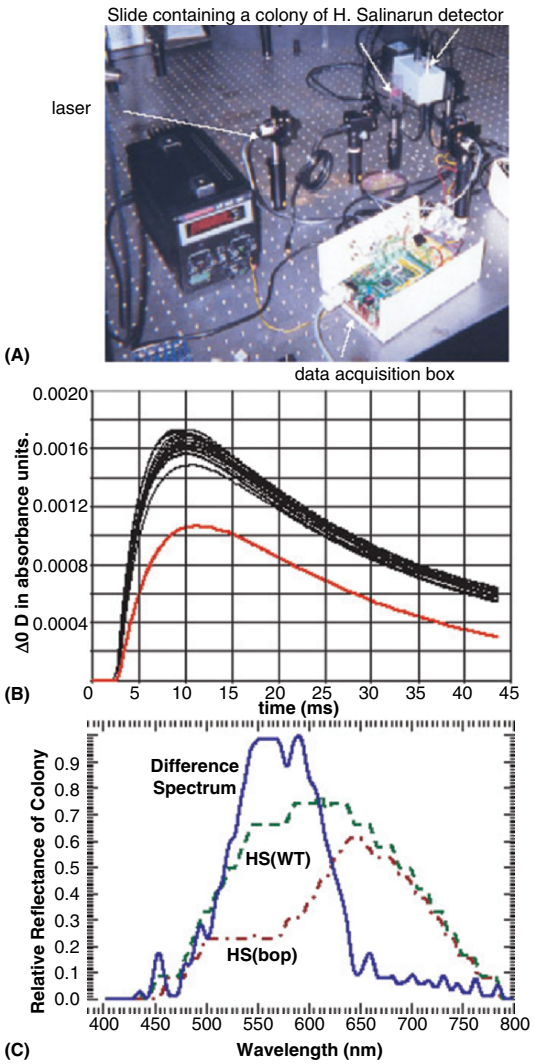


Figure 10. (A) Existing screening apparatus, which determines the absorbance change versus time at 690 nm in live colonies of *H. salinarum*. (B) O-state screening data are shown for 18 consecutive pump-probe cycles, each separated by 5 min. The sample is then subjected to a constant illumination at 635 nm for 2 h. The lower curve, shown in red, is a final pump-probe cycle after illumination. The reduced amount of O-state photoproduct is due to the formation of P-state and Q-state. (C) Reflectance data measured in wild-type (green) and bop- (red) colonies of *H. salinarum*. Bop- cells do not produce BR and appear white, while wild-type cells appear purple. A difference spectrum shows a peak at 568 nm, corresponding to the chromophore peak of BR. This method will be used to detect differences in emission spectra between the BR mutants. Measurements were taken with an Ocean Optics fiber optic emission spectrometer Model SD-2000.

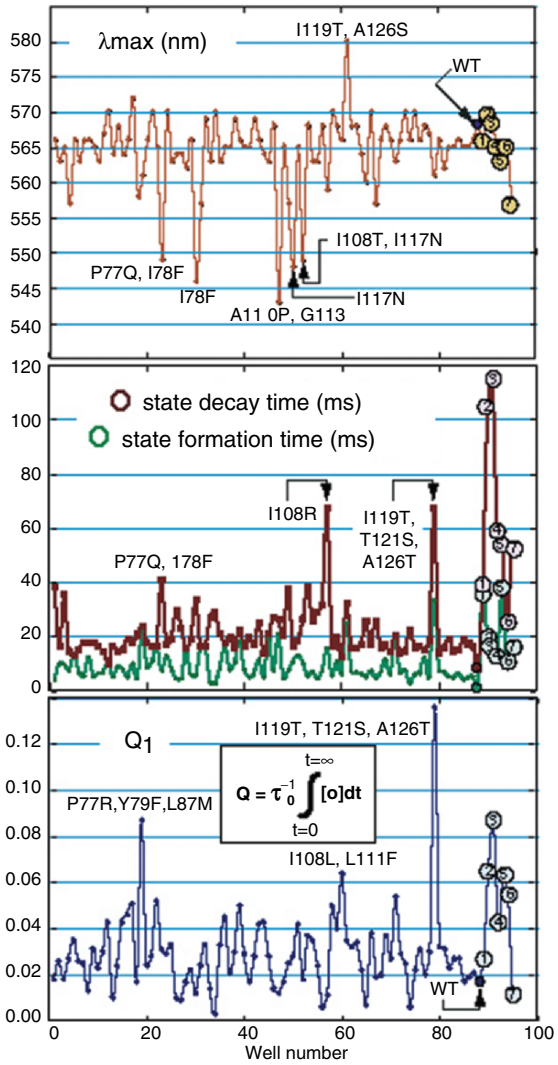


Figure 11. Spectrokinetic screening of semi-random mutants for branched-photocycle memory applications. Semi-random mutant proteins were constructed and analyzed in 96-well plates. Wild-type protein and 7 site-directed mutants (labeled 1–7 and positioned to the right of wild-type) are shown for comparison. 1 = E194Q; 2 = E204Q; 3 = F208N, 4 = E194C, 5 = E204C, 6 = E194C/E204C; 7 = E194Q/E204Q. Some of the semi-random mutant extrema are labeled. The top panel shows the λ_{max} of mutant proteins and WT. The middle panel shows the O-state intermediate (as measured at 650 nm) decay and formation times for the mutant proteins and WT. Formation times are shown in green, decay times in red. The bottom panel shows the Q value for each mutant (see Fig. 10). Note that some of the semi-random mutants show higher values in some categories than do the site-directed mutants designed from *a priori* knowledge of the protein.

current state-of-the-art in bacteriorhodopsin-based memory storage. The largest remaining hurdles still concern materials development and involve optimizing bacteriorhodopsin for use in specific applications. However, never before have techniques such as random and semi-random mutagenesis and directed evolution been available to researchers in this field—the landscape of potential new materials based on bacteriorhodopsin is nearly unlimited. Limitations in the wild-type protein no longer pose serious problems—indeed the potential now exists to custom engineer a wide variety of proteins for a diverse range of applications.

ACKNOWLEDGEMENTS

This work was supported by the Air Force Research Laboratories in Rome, NY, the National Science Foundation, and by the W. M. Keck Foundation.

REFERENCES

1. Birge, R. R., 1994, Introduction to molecular and biomolecular electronics, *Adv. Chem.* **240:**1–14.
2. Birge, R. R., 1995, Protein-based computers, *Sci. Am.* **272:**90–95.
3. Birge, R. R., Gillespie, N. B., Izaguirre, E. W., Kusnetzow, A., Lawrence, A. F., Singh, D., Song, Q. W., Schmidt, E., Stuart, J. A., Seetharaman, S., and Wise, K. J., 1999, Biomolecular electronics: Protein-based associative processors and volumetric memories, *J. Phys. Chem. B.* **103:**10746–10766.
4. Birge, R. R., Parsons, B., Song, Q. W., and Tallent, J. R., 1997, Protein-based three-dimensional memories and associative processors, in *Molecular Electronics* (M.A. Ratner and J. Jortner, eds.), Blackweel Science Ltd., Oxford, pp. 439–471.
5. Vought, B. W., and Birge, R. R., eds., 1999, Molecular electronics and hybrid computers, in *Wiley Encyclopedia of Electrical and Electronics Engineering*. (J. G. Webster, ed.), vol. 13, Wiley-Interscience, New York, pp. 477–490.
6. Birge, R., 1992, Protein-based optical computing and memories, *Computer* 56–67.
7. Oesterhelt, D., and Stoeckenius, W., 1971, Rhodopsin-like protein from the purple membrane of Halobacterium halobium, *Nat. New Biol. (Lond.)* **233:**149–152.
8. Birge, R. R., 1981, Photophysics of light transduction in rhodopsin and bacteriorhodopsin, *Ann. Rev. Biophys. Bioeng.* **10:**315–354.
9. Birge, R. R., 1990, Nature of the primary photochemical events in rhodopsin and bacteriorhodopsin, *Biochim. Biophys. Acta* **1016:**293–327.
10. Ebrey, T. G., 1993, Light energy transduction in bacteriorhodopsin, in *Thermodynamics of Membrane Receptors and Channels* (M. B. Jackson, ed.), CRC Press, Boca Raton, pp. 353–387.
11. El-Sayed, M. A., 1992, On the molecular mechanisms of the solar to electric energy conversion by the other photosynthetic system in nature, bacteriorhodopsin, *Accts. Chem. Res.* **25**(7):279–286.

12. Lanyi, J., and Luecke, H., 2001, Bacteriorhodopsin, *Curr. Opin. Struc. Biol*. **11:**415–419.
13. Lanyi, J. K., 1992, Proton transfer and energy coupling in the bacteriorhodopsin photocycle, *J. Bioenerg. Biomembr*. **24:**169.
14. Lanyi, J. K., 1999, Bacteriorhodopsin, *Int. Rev. Cytol*. **187:**161–202.
15. Stuart, J. A., and Birge, R. R., 1996, Characterization of the primary photochemical events in bacteriorhodopsin and rhodopsin, in *Biomembranes* (A. G. Lee, ed.), JAI Press, London, pp. 33–140.
16. Lanyi, J., 1993, Proton translocation mechanism and energetics in the light-driven pump bacteriorhodopsin, *Biochim. Biophys. Acta* **1183**(2):241–261.
17. Lanyi, J., and Varo, G., 1995, The photocycles of bacteriorhodopsin, *Isr. J. Chem*. **35:**365–385.
18. Lozier, R. H., Bogomolni, R. A., and Stoeckenius, W. 1975, Bacteriorhodopsin: A light-driven proton pump in *Halobacterium halobium*, *Biophys. J*. **15:**955–962.
19. Mathies, R., Lin, S., Ames, J., and Pollard, T., 1991, From femtoseconds to biology: Mechanism of bacteriorhodopsin's light-driven proton pump, *Annu. Rev. Biophys. Biophys. Chem*. **20:**491–518.
20. Birge, R. R., 1990, Photophysics and molecular electronic applications of the rhodopsins, *Annu. Rev. Phys. Chem*. **41:**683–733.
21. Hasson, K. C., Gai, F., and Anfinrud, P. A., 1996, The photoisomerization of retinal in bacteriorhodopsin: Experimental evidence for a three-state model, *Proc. Natl. Acad. Sci*. **93:**15124–15129.
22. Birge, R. R., and Cooper, T. M., 1983, Energy storage in the primary step of the photocycle of bacteriorhodopsin, *Biophys. J*. **42:**61–69.
23. Birge, R. R., Cooper, T. M., Lawrence, A. F., Masthay, M. B., Zhang, C. F., and Zidovetzki, R., 1991, Revised assignment of energy storage in the primary photochemical event in bacteriorhodopsin, *J. Am. Chem. Soc*. **113:**4327–4328.
24. Popp, A., Wolperdinger, M., Hampp, N., Bräuchle, C., and Oesterhelt, D., 1993, Photochemical conversion of the o-intermediate to 9-*cis*-retinal-containing products in bacteriorhodopsin films, *Biophys. J*. **65:**1449–1459.
25. Gillespie, N. B., Wise, K. J., Ren, L., Stuart, J. A., Marcy, D. L., Hillebrecht, J., Li, Q., Ramos, L., Jordan, K., Fyvie, S., and Birge, R. R., 2002, Characterization of the branched-photocycle intermediates P and Q of bacteriorhodopsin, *J. Phys. Chem. B*. **106**(51):13352–13361.
26. Rentzepis, P. M., 1993, *Three-Dimensional Optical Memory*, The Regents of the University of California, US.
27. Liao, N., Gong, M., Xu, D., Qi, G., and Zhang, K., 2001, Single-beam two-photon three-dimensional optical storage in a pyrryl-substituted fulgide photochromic material, *Chin. Sci. Bull*. **46**(22):1856–1859.
28. Mok, F. H., 1993, Angle-multiplexed storage of 5000 holograms in lithium niobate, *Opt. Soc. Am*. **18**(11):915–917.
29. Paek, E. G., and Psaltis, D., 1987, Optical associative memory using Fourier transform holograms, *Opt. Eng*. **26:**428–433.
30. Psaltis, D., and Pu, A., 1996, Holographic 3-D disks, *Proc. IEEE Nonvol. Mem. Tech. (INVMTC)* pp. 34–39.
31. Birge, R. R., 1994, Three-dimensional optical memories, *Am. Sci*. **82:**349–355.
32. Birge, R. R., and Govender, D. S. K., 1993, *Three-Dimensional Optical Memory*, Syracuse University, US.
33. Stuart, J. A., Marcy, D. L., and Birge, R. R., 2000, Bacteriorhodopsin-based three-dimensional optical memory, in *Bioelectronic Applications of Photochromic Pigments*, IOS Press, Szeged, Hungary.

34. Stuart, J. A., Marcy, D. L., Wise, K. J, and Birge, R., 2002, Volumetric optical memory based on bacteriorhodopsin, *Synth. Met*. **127:**3–15.
35. Stuart, J. A., Tallent, J. R., Tan, E. H. L., and Birge, R. R., 1996, Protein-based volumetric memory, *Proc. IEEE Nonvol. Mem. Tech. (INVMTC)* **6:**45–51.
36. Tallent, J. R., Stuart, J. A., Song, Q. W., Schmidt, E. J., Martin, C. H., and Birge, R. R., 1998, Photochemistry in dried polymer films incorporating the deionized blue membrane form of bacteriorhodopsin, *Biophys. J*. **75:**1619–1634.
37. Chang, C.-H., Liu, S. Y., Jonas, R., and Govindjee, R., 1987, The pink membrane: The stable photoproduct of deionized purple membrane, *Biophys. J*. **52:**617–623.
38. Chung-Ho, C., Jonas, R., Ebrey, T. G., Hong, M., and Eisenstein, L., 1987, Protonation changes in the interconversions of the pink membrane, blue membrane and purple membrane, in *Biophysical Studies of Retinal Proteins* (T. G. Ebrey, H. Frauenfelder, B. Honig, and K. Nakanishi, eds.), University of Illinois Press, pp. 156–166.
39. Gross, R. B., Todorov, A. T., and Birge, R. R., 1995, The wavelength-dependent refractive index change associated with the blue to pink membrane photochemical conversion in bacteriorhodopsin, in *Applications of Photonic Technology* (G. A. Lampropoulos, ed.), Plenum Press, New York, pp. 115–121.
40. Liu, S. Y., and Ebrey, T. G., 1987, The quantum efficiency for the interphotoconversion of the blue and pink forms of purple membrane, *Photochem. Photobiol*. **46**(2): 263–267.
41. Pande, C., Callender, R. H., Chang, C. H., and Ebrey, T. G., 1986, Resonance raman study of the pink membrane photochemically prepared from the deionized blue membrane of *H. Halobium*, *Biophys. J*. **50:**545–549.
42. Tallent, J., Song, Q. W., Li, Z., Stuart, J., and Birge, R. R., 1996, Effective photochromic nonlinearity of dried blue-membrane bacteriorhodopsin films, *Optics Lett*. **21:**1339–1341.
43. Hampp, N., and Juchem, T., 2000, Fringemaker—The first technical system based on bacteriorhodopsin, in *Bioelectronic Applications of Photochromic Pigments*, IOS Press, Szeged, Hungary.
44. Gross, R. B., 1995, *The Development, Characterization and Applications of Polymeric Films Containing Bacteriorhodopsin*, Syracuse University.
45. Gross, R. B., Izgi, K. C., and Birge, R. R., 1992, Holographic thin films, spatial light modulators and optical associative memories based on bacteriorhodopsin, *Proc. SPIE* **1662:**186–196.
46. Wise, K. J., Gillespie, N. B., Stuart, J. A., Krebs, M. P., and Birge, R. R., 2002, Optimization of bacteriorhodopsin for bioelectronic devices, *Trends Biotechnol*. **20**(9):387–394.
47. Chen, Z., and Birge, R. R., 1993, Protein based artificial retinas, *Trends Biotechnol*. **11:**292–300.
48. Martin, C. H., Chen, Z. P., and Birge, R. R., eds., 1997, Towards a bacteriorhodopsin-silicon neuromorphic photosensor, *Proc. Pacific Symp. Biocomputing* (R. B. Altman, A. K. Dunker, L. Hunter, and T. E. Klein, eds.), World Scientific, Maui, pp. 268–279.
49. Birge, B., Fleitz, P., Gross, R., Izgi, J., Lawrence, A., Stuart, J., and Tallent, J., 1990, Spatial light modulators and optical associative memories based on bacteriorhodopsin, in *Materials Research Society*, Boston, MA.
50. Birge, R. R., 1992, Protein based optical computing and optical memories, *IEEE Comput*. **25:**56–67.
51. Birge, R. R., Chen, Z., Govender, D., Gross, R. B., Hom, S. B., Izgi, K. C., Stuart, J. A., and Vought, B. W., 1995, Biomolecular photonics based on bacteriorhodopsin, in *CRC Handbook of Organic Photochemistry Photobiolology*, pp. 1568–1587.

Chapter 5

SPIDER SILK PRODUCTION

Randolph V. Lewis
Dept. of Molecular Biology, Univ. of Wyoming 1000E. University, Dept. 3944, Laramie, WY 82071-3944

Abstract: Spider silks have been noted for their extraordinary mechanical properties for centuries. Recent research has provided information that helps explain these properties on a molecular basis. This article describes the biological and mechanical properties of spider silks. Three different spider silks are discussed with particular emphasis on the protein sequences and how they compare. The structures of key amino acid sequences are used to indicate how they generate the mechanical properties.
Key words: Spider silk, fiber, major ampullate, minor ampullate, flagelliform, biophysical, protein structure

1. INTRODUCTION

Human thinking in material and design frequently draws on nature for inspiration. Applications from architecture to ceramics frequently seek to utilize the ideas that natural evolution has developed and tested for millions of years. This is particularly true of spider silk. This material has held human fascination for many years. There are stories regarding spiders and their ability to spin silks in Greek mythology as well as the Koran. To scientists, spider silk is remarkable not only for its exceptional material properties but also for its production, processing, and evolution. The ability to produce a synthetic spider silk remains something of a holy grail of material science. Yet, an understanding of the composition and design of spider silk suggests that it may be possible not only to mimic spider silk but also to alter its composition and thus its material properties. Scientific studies have revealed the composition of spider silk to be almost exclusively protein.[1] This has raised questions

V. Renugopalakrishnan et al. (eds.), Bionanotechnology, 61-78.

and revealed surprising findings regarding the development and evolution of many different types of spider silk.

2. BIOLOGICAL ASPECTS OF SPIDER SILK PRODUCTION

Various species of spiders have developed many different applications for silk. These include constructing egg cases, lining burrows, ensnaring or swathing prey, and creating custom habitats. Most species are capable of producing multiple types of silk, each suited to a particular purpose. A distinct gland contained within the abdomen of the spider produces each type of silk. Normally each type of gland occurs in pairs with bilateral symmetry.

Each type of silk gland has a unique morphology, but the overall organization of these glands is very similar between species. Most of what is known about spider silks was learned from the study of silk produced by the major ampullate gland of large, orb-web weaving species such as members of genus *Nephila*. This particular gland produces dragline silk that is a component of an orb web, as well as the lifeline of the spider when falling from high places. More basal taxa of spiders do not produce orb webs but have other uses for the silks that they produce. Frequently, these silks are considered to be orthologous but only by the morphology of the gland. Clearly, they are not related by application or gene sequence.

Transmission electron micrograph analysis has shown that silk proteins are produced in specialized columnar epithelial cells in a portion of the gland commonly referred to as the "tail."[2] In the major ampullate glands, there appears to be two different types of these cells[3] that may correlate with the production of two distinct types of silk proteins in this gland. According to Bell and Peakall, there was no evidence for Golgi in these cells; instead, protein molecules are seen to aggregate in granules in the cytoplasm of the cell, which are then secreted into the lumen of the gland.

After synthesis and secretion into the lumen of the gland, the silk is stored in the body of the gland in an aqueous solution. In this part of the gland, the silk is in a liquid crystalline state, which may inhibit polymerization.[4] A duct leads from the distal end of the silk gland to the spinnerets at the posterior end of the spider. These ducts are at least five times longer than the distance from the body of the gland to the spinneret.[2] This excess length strongly indicates that important processing is taking place as the silk material travels down the duct. As silk is drawn down the duct leading to the spinnerets, it becomes increasingly dehydrated, slightly acidified, and increasingly birefringent.[5] This indicates that as the silk exits the spider, it has already been modified in a manner that alters the shape and orientation of the molecules. Other than the influx of

protons, there is a little evidence for substances being introduced to, or removed from, the spinning dope as the material passes through the duct and out of the spider. Unlike many artificial materials, spider silk emerges from the spinneret in a finished state and requires no further processing such as postspin drawing.

3. MECHANICAL PROPERTIES OF SPIDER SILKS

Insect-produced silks have long been of use, most frequently for the textiles. Spider silks have received a great deal of attention for their exceptional mechanical properties, although for reasons discussed later, these silks have not been as amenable to human utilization. Spiders that spin aerial webs have long been subjected to strong evolutionary pressures to create a panel of materials that are ideally suited to this demanding application. A web must be made out of a minimal amount of material on order to be both economical to produce and as undetectable as possible to the prey. A web must also be able to absorb and dissipate a great deal of kinetic energy from the impact of a flying insect. Due to the effects of these various evolutionary pressures, some silks, such as dragline, have evolved to have very high-tensile strengths, while others, like flagelliform, have evolved to exhibit extreme elasticity. For an excellent review of the elastic properties of spider silks and other biological materials, see Ref. 6, Table 1.

Many early studies have shown that dragline silk, in particular, exhibited very high tensile strength, as well as some degree of elasticity.[7–9] However, these studies were based on an implied mean diameter, which was calculated based on the known length, mass, and density of a particular silk sample. Later microscopic analysis demonstrated high variability in this diameter. In

Table 1. Mechanical properties of high-performance materials

Material	Tensile strength (GPa)	Tensile elasticity (%)	Young's modulus (GPa)	Density (kg/L)
S-Glass	2.53	2.9	87	2.5
Kevlar 29	2.27	2.8	83-100	1.44
Kevlar 49	2.27	1.8	124	1.44
High-strength carbon fiber	1.1	1.1	230	1.8
Major ampullate silk	4.6	35	22-34	1.36
Minor ampulate	0.96	5		1.19
Flagelliform	1	200		
Silkworm silk	1.3	12	16	

some silk fibers, the minimum diameter could be as little as 50% of the average diameter. If minimum diameters are used for these calculations, the calculated tensile strength of the silk can be as much as four times greater. These variations in fiber diameter have also been supported by observations using scanning electron microscopy.[10]

4. PROTEIN SEQUENCES AND EVOLUTIONARY IMPLICATIONS

Among the earliest studies of the chemical makeup of spider silks showed that these were made almost exclusively of protein.[1] Even recent studies have shown that all of the silks that are used to make an orb web are composed of protein and very little else. One study has shown the presence of some unusual branched lipids in one type of silk.[11] However, silks are, for the most part, protein-based materials. In major ampullate silks from a variety of different species, alanine, glycine, proline, and glutamine make up over 80% of the constituent amino acids (see Fig. 1).[12] Levels of proline vary widely between different species, but amounts of other amino acids remain relatively constant.

Genomic and cDNA work have shed a great deal of light on the evolution of the silk genes, as well as the structure/function relationships that make spider silk such a remarkable material. For araneoid species that spin aerial webs, genes have been sequenced encoding the proteins that form major ampullate, minor ampullate, and flagelliform silks. Many silk genes have also been sequenced from more basal taxa, features of which will be discussed later.

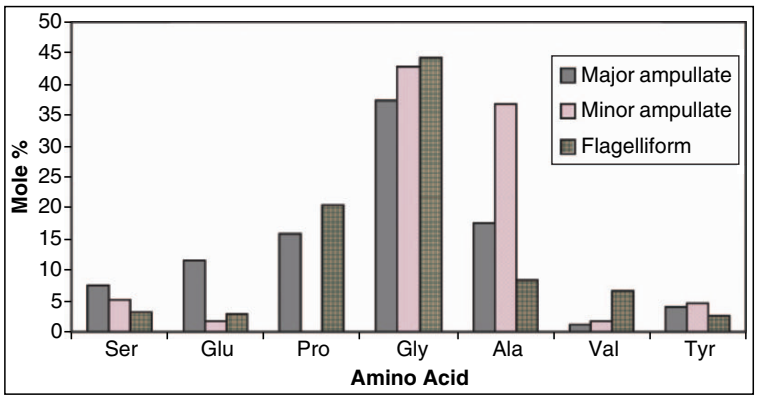


Figure 1. Amino acid composition of three types of silk (from Andersen, 1970).

4.1 Major Ampullate

Major ampullate or dragline silk has been the most extensively studied of the spider silks. In 1990, this silk became the first to have a silk protein gene sequenced.[13] A cDNA clone was isolated utilizing a degenerate nucleotide probe based on a short peptide isolated from acid hydrolyzed dragline silk from *Nephila clavipes*. This gene encoded a series of similar but nonidentical repeats. Each of these repeat units is approximately 30 amino acids (aa) in length and composed of a number of identifiable motifs. These motifs include a series of up to seven consecutive alanine residues alternating with a tandem repeats of Gly-Gly-X. The X amino acid is limited to Tyrosine (Y), Leucine (L), or Glutamine (Q). In a series of these GGX motifs, the X residues frequently follow the above order, resulting in a larger repeat of GGYGGLGGQ. Between repeat units, there are wide variations in the number of GGX repeats, but virtually no other substitutions for the X residue. This gene sequence bears no significant homology to other known genes, indicating that it is a truly unique structure.

The amino acid composition of the protein encoded by this cDNA nearly matched that already known for the dragline silk fiber, with one glaring exception. It was known that dragline silk contained nearly 13% proline; however, no proline was encoded by this known gene. The cDNA library was therefore reprobed and a second cDNA, named MaSp2 was identified.[14] Like MaSp1, this gene encoded a repetitive protein, where each repeat unit was of a different length but also composed of similar short amino acid motifs (see Table 2). The motifs of MaSp2 include polyalanine (like MaSp1). The other motif that composed MaSp2 was a proline-containing motif of GPGQQGPGGY. Furthermore, like MaSp1, each repeat unit varies in the number of motifs, but truncated motifs as well as amino acid substitutions are not found.

Table 2. MaSp1 consensus repeats

Nep.c.	GGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAA
Nep.m.	GGAGQGGYGGLGSQGAGRGGYGGQGAGAAAAAA
Nep.s.	GGAGQGGYGGLGGQGAGAAAAAA
Tet.k.	GGLGGGQGAGQGGQQGAGQGGYGSGLGGAGQGASAAAAAAAA
Tet.v.	GGLGGGQGGYGSGLGGAGQGGQQGAGQGAAAAAASAAA
Lat.g.	GGAGQGGYGQGGQGGAGAAAAAAAAA
Arg.a.	GGQGGXGGYGGLGSQGAGQGYGSGLGGQGGAGQGGAAAAAAAAAA
Arg.t.	GGQGGQGGYGGLGXQGAGQGYGAGSGGQGGXGQGGAAAAAAAA
Ara. d.	GQGGQGGQGGLGSQGAGGAGQGGYGAGQGGAAAAAAAA

Table 3. MaSp2 consensus repeats

Nep.c.	GPGQQGPGGYGPGQQGPGGYGPGQQGPSGPGSAAAAAAAA
Nep.m.	GPGQQGPGGYGPGQQGPGGYGPGQQGPSGPGSAAAAAAA
Nep.s.	GPGQQGPGXYGPSGPGSAAAAA
Lat. g.	GPGGYGPGPGXQQGYGPGGSGAAAAAAAA
Arg.a.	GGYGPGAGQQGPGSQGPGSGGQQGPGGXGPYGPSAAAAAAAA
Arg.t.	GGYGPGAGQQGPGSQGPGSGGQQGPGGQGPYGPSAAAAAAAA
Gas.m.	GGYGPGSGQQGPGQQGPGSGGQQGPGGQGPYGPGAAAAAAAA
Ara.b.	GGYGPGSGQQGPGQQGPGQQGPYGPGASAAAAAA
Ara.d.	GGYGPGSGQQGPGQQGPGGQGPYGPGASAAAAAA
Nep.m.	GRGPGGYGPGQQGPGGPGAAAAAA
Arg.t.	GPGGQGPGQQGPGGYGPSGPGGASAAAAAAAA
Ara.d.	GPGGYGPGSQGPSGPGAYGPGGPGSSAAAAAAAAS

As will be discussed in more detail in Section 4.2, both MaSp1 and MaSp2 encode a 100 aa carboxy-terminal sequence, which is very distinct from the repetitive region that comprises the majority of the gene. However, although MaSp1 and MaSp2 only share polyalanine as the common motif, the C-term regions are very highly conserved.

Northern blotting has shown that both genes encode large transcripts of 12.5 kb for MaSp1 and 10.5 kb for MaSp2. Genomic analysis for these genes indicates a lack of any large introns throughout most of the gene.

4.2 Minor Ampullate Silk Proteins

Genes encoding minor ampullate silk were the next spider silk genes to be sequenced.[15,16] Minor ampullate silk also appears to be composed of two distinct proteins, named MiSp1 and MiSp2, both of which have similar architecture to MaSp1 and MaSp2 (see Table 4). The main portion of both genes encode many repeats composed of certain amino acid motifs, which is followed by a 100 aa, nonrepetitive C-term region very similar to the carboxy terminus of MaSp1 and MaSp2. There were also some critical differences between MiSp and MaSp genes. The first is interruption of the repetitive regions by a conserved 130 aa nonrepetitive "spacer," which is composed of larger, hydrophilic

Table 4. Minor ampullate silk consensus repeats

Nephila clavipes MiSp1
GAGGAGGYGRGAGAGAGAAAGAGAGAGGYGGQGGYGAGAGA-GAAAAAGA
Nephila clavipes MiSp2
GGYGRGVGAGAGAGAAAGXGAGAGGYGGQGGYGAGXGAAAAGAG

amino acids and shows no similarity to any part of any other characterized gene. The function of this spacer is currently unknown, but it is hypothesized to play a role in fiber formation, possibly as a site of protein–protein interaction to aid protein accretion for formation of a solid fiber.

The amino acid motifs that compose MiSps were also of interest. MiSp contains GGX motifs and short polyalanine segments similar to MaSp1. But, MiSp also contains larger blocks of alternating glycine and alanine. The significance of this poly-GA motif will be discussed in greater detail later.

Northern blotting revealed the MiSp1 and MiSp2 genes to also be responsible for relatively large transcripts of 9.5 and 7.5 kb, respectively. Restriction analysis, followed by Southern blotting, suggests that the gene is nearly the same size as its transcript, suggesting a lack of large introns throughout the coding region.

4.3 Flagelliform Silk Proteins

The most recent silk gene to be sequenced was that from the flagelliform gland.[17] This gland produces the silk that composes the highly elastic silk of the capture spiral of a silk web. Separate cDNA clones identified for portions of this large gene contain the 5′ end of this gene, including the 5′ UTR, initiation codon, and signal peptide, C-terminus and numerous iterations of a pentapeptide motif. This information reveals a similar, yet distinct architecture from the other types of silk. Flagelliform silk contains only a single repeated motif, GPGGX, which is very similar to the proline-containing repeat from MaSp2. A unique and highly conserved 34 aa spacer interrupts repetitive regions of this silk. Although flagelliform silk does contain a nonrepetitive C-term region, there is no sequence similarity to the C-term regions of either of the ampullate type silks.

Genomic sequence data has also been obtained for this type of silk and demonstrates some unusual findings.[18] This data indicates a pattern of repeating introns and exons, with the introns being more highly conserved than the exons (see Table 5). Furthermore, when flagelliform orthologs are aligned, sequential introns within a single gene are more similar than are corresponding introns between orthologous genes. This suggests that there must be significant homogenizing forces acting on introns.

Table 5. Flagelliform consensus repeat sequences

Nep. c	$[GPGGX]41[GGX]_7TIIEDLDITIDGADGPITISEELTISGAGGS[GPGGX_n]_{26}$
Nep.m	$[GPGGX]_{36}[GGX]_7TVIEDLDITIDGADGPITISEELTIGGAGAGGS[GPGGX_n]_{19}$
Arg.t	$[GPGGX_n]_6GPVTVDVDVSVGGAPGG[GPGGX_n]_5[GGX]_6[GPGGX_n]_7$

4.4 Silk Evolution, Convergence, and Divergence

Silk genes for major and minor ampullate and flagelliform silks have been sequenced from a variety of spider species.[19] Orb-weaving species demonstrate a great deal of gene sequence conservation, having maintained the same motifs and gene architecture over more than 125 million years of separation. One interesting aspect of these other similar silks is that differences in motifs are propagated throughout the gene, suggesting that powerful and likely novel mechanisms are acting to maintain silk gene sequences.

When silks from more distantly related, nonorb web-weaving species are compared, the results are very different. These silk genes show very few of the recognizable motifs that are so highly conserved among orb-weaving species. Some elements of the architecture are conserved, however. Poly-alanine or poly gly–ala occur occasionally, but other similar patterns also appear, such as poly glutamine–alanine. Even these more basic silks still show a repetitive nature, but the consensus repeat is frequently much longer. This strongly indicates that silks from more basal nonorb-weaving species are subject to the same homogenizing evolutionary forces as are silks from more highly evolved species but that the sequence requirements are very different. This may be reflective of the selection for spiders producing silks with optimal material properties when their lifestyle is dependant upon an effective web for prey capture. Species that do not produce orb webs are less dependent upon the mechanical properties of their silk for their continued survival. These species frequently take a more active role in pursuing and capturing prey.

Some species of spider produce silks that appear to be more similar to silks produced by silkworm species than other spiders. Two examples include the silk gene of the jumping spider *Phiddipus* which is clearly very similar to the silk of the silkworm *Bombyx*. The second is the spider *Zorocrates* which is clearly very similar to the silk of the oak silkworm, *Antheria*. Two hypotheses may explain this interesting phenomenon: (1) convergent evolution or (2) horizontal gene transfer. Horizontal gene transfer seems unlikely because such a hypothesis must explain both the transfer and utilization of the genetic material. Convergent evolution would imply that these organisms have independently arrived at a nearly identical solution for creating a high performance silk. This theory would imply that there are a very limited number of protein sequences that can result in the necessary conformations to confer the properties that a silk fiber must have to be an evolutionary benefit to the organism.

Convergent evolution is also bolstered by the fact that the motifs that comprise spider silk are found in many other protein applications. Poly-alanine, the role of which will be discussed later, is also found in mollusk

shell proteins,[20] and mussel byssus threads,[21] both of these are structural proteins which are very strong and are extracellular in application. GGX, a component of MaSp1, is also found in mussel byssus threads. The proline containing motif, GPGXX, is believed to be responsible for the elasticity observed in some types of spider silk, as will be discussed later. This motif also occurs in other elastic proteins, such as wheat glutin and elastin,[22] and mussel bysus threads.[21] These organisms are unrelated through all but the most ancient ties. This demonstrates that multiple organisms have independently arrived at a small set of motifs for creating structures with certain mechanical properties.

5. BIOPHYSICAL STUDIES

Almost all biophysical studies to date have been performed on major ampullate fibers from *N. clavipes*. This type of silk fiber has a relatively large diameter and is easily obtained in larger quantities than any other type of silk.

The earliest biophysical studies performed on spider silk were X-ray diffraction. The only apparent spacing corresponded with that of antiparallel beta sheets. However, this technique demonstrated that much of the protein in the fiber was actually unordered.

Wide angle X-ray diffraction (WAXD) indicates that the gross morphology of silk fibers is small, rigid crystals embedded in a far less rigid amorphous matrix.[23] This is known as a two-phase model of the structure of spider silk. This methodology suggests that there may actually be a third phase of structured but unorganized regions.

The next technique to be applied to spider silk fibers was Fourier-transform infrared (FTIR) spectroscopy. Data from this type of analysis supported the existence of beta sheet structure within the fiber, which was not altered when the fiber was stretched or supercontracted. However, there was also emergence of a signal indicating additional helical structure when the fiber was stretched. This signal disappeared when tension was released from the fiber. These helical regions appeared to arise from regions that were unordered when the fiber was in a relaxed state. Parallel polarized spectra indicated that the proteins forming the recognizable structures were oriented parallel to the long axis of the fiber.[24]

Techniques of solid-state nuclear magnetic resonance (NMR) spectroscopy are more recently being applied to analyzing the molecular structures of spider silk fibers. These studies confirm the beta sheet structure present in silk fibers.[25,26] Studies have demonstrated that the chemical shifts of α, β, and

carbonyl carbons of amino acids are sensitive to the confirmation of the peptide backbone. Therefore, these chemical shifts are indicative of the secondary structure of the peptide. Studies using silk from spiders fed ^{13}C alanine indicate that the beta sheet regions are composed strictly of alanine. These studies detected no change in the structure of the alanine residues upon stretching or supercontraction.

This leads to a model of silk that entails a largely amorphous matrix of protein molecules which are oriented parallel to the long axis of the fiber. These fibers are bound together through interspersed rigid crystals of poly-alanine motifs. This type of structure is supported by computer modeling of such a network that accurately reproduces the shape of stress–strain curves that silk fibers demonstrate.[27]

6. PROTEIN STRUCTURE FUNCTION RELATIONSHIPS

The occurrence of alanine in apparently rigid beta sheet structures corresponds very well with the utilization of alanine almost exclusively in poly-alanine blocks in the protein sequence. It is widely believed that these beta-structured regions are composed almost exclusively of poly-alanine regions, which aggregate to form the rigid crystals within the otherwise amorphous silk. Furthermore, these poly-alanine crystals are believed to be a site of cross linking between multiple, parallel protein molecules. Since poly-alanine stretches occur many times per molecule (at least 100 times), this allows tight interaction with many different molecules. This tight binding between protein molecules is believed to be responsible for the tensile strength of silk fibers. Alanine residues in a beta sheet form have an ability to form very tightly packed aggregates through interactions very similar to a zipper, with protruding methyl groups occupying the void near the alpha carbon of a residue on a neighboring chain. Such an aggregate would be very rigid, have no void space, and be impenetrable to water, since it is stabilized by hydrogen bonds in one dimension and hydrophobic interactions in the other. Molecular dynamics simulations support such a model through analysis of similar peptides that occur in prion proteins and seem to be responsible for the fibril formation seen in the pathologies associated with prion proteins. This type of analysis suggests that in free solution, poly-alanine adopts an extended helical conformation, but when enough peptide molecules are brought together, they adopt a beta sheet confirmation and aggregate into a rigid structure identical to that described above. This aggregate can also nucleate condensation of other peptides, which build on the rigid crystalline seed.[28]

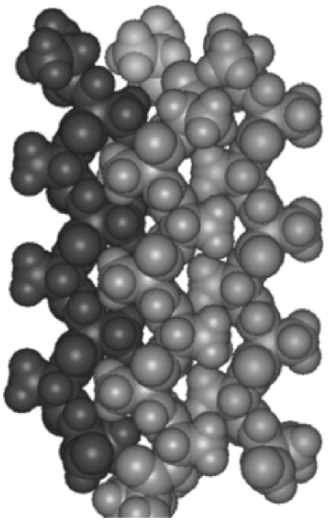


Figure 2. Three poly-alanine beta sheets.

The presence of the crystalline beta sheet regions is also supported by the properties of the corresponding structure found in minor ampullate silk (see Fig. 2). Instead of poly-alanine, minor ampullate silk has stretches of alternating glycine–alanine residues which are likely to form corresponding structures. The major difference in this case would be an absence of methyl side chains on one side of the protein strand. This results in introduction of small void spaces into the crystals, as well as weaker interactions due to reduced hydrophobic interactions and steric interactions due to the ''zippering'' of adjacent molecules. Correspondingly, minor ampullate silk exhibits substantially less-tensile strength than does major ampullate silk. Elasticity was originally attributed to the GGX regions that were first seen in MaSp1. These regions were thought to be able to form an extended helical form upon extension, but otherwise unordered in relaxed silk.[26] This agrees to the predictions that as much as 85% of the energy absorbed by the fiber upon impact is utilized to reduce molecular entropy.[29] This is also in agreement with the appearance of an alpha helical region upon stretching as is detected by FTIR spectroscopy.[24] The NMR studies of this motif also support an extended 3_1 helix.[30]

When flagelliform silk was first sequenced, the above assumptions were challenged. This type of silk is highly elastic and the protein sequence is composed almost exclusively of repeats of GPGXX that is very similar to one of the motifs found in MaSp2 (see Table 3). This motif draws attention when protein sequences for multiple silks are compared with the mechanical

properties of the silks they encode. Minor ampullate silk genes encode none of these motifs and the resultant fiber demonstrates no elasticity. On a percentage basis, flagelliform silk genes encode seven times more of this motif than do major ampullate silk genes, and the resultant flagelliform fiber exhibits seven times more elasticity than does major ampullate silk. Very similar motifs are also found in other elastic structural proteins such as wheat gluten, mussel byssus threads, and elastin. In silk, as well as other occurrences of this motif, it is hypothesized to form a series of linked type II beta turns, which would result in a type of flattened molecular coil spring.[31] Proline is known to form the core of tight, hairpin type turns because of the ability of this amino acid to adopt a *cis* conformation.

There are a number of different hypothesized structures for the proline-containing silk motifs,[32] but the common element of these predictions is the presence of a stack of beta turns, resulting in a structure resembling a coil spring (see Fig. 3). The function of this structure is yet to be understood. Some propose that the turns suspend the intervening sequence in a conformationally free state. This structure could also be seen to function more akin to a

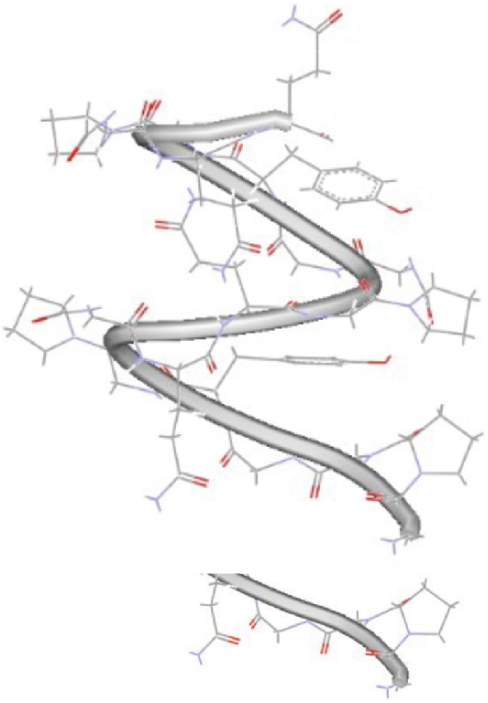


Figure 3. Hypothesized structure of proline-containing motif.

macroscopic spring in which energy is stored in strain induced on chemical bonds. If this motif is truly responsible for silk elasticity, its role is likely to be a mixture of enthalpic and entropic energy storage. If this forms a rigid molecular spring, it should be detectable in X-ray diffraction studies, but the presence of beta sheets may be masking it.

7. EXPRESSION OF SYNTHETIC SPIDER SILK PROTEINS

Spiders are aggressive and antisocial creatures and are not amenable to farming or growing at high density. Furthermore, spider silk fibers are very fine, with diameters often less than 10 microns. This means that even though a large spider may produce over 100 m of silk, this is actually less than a milligram of material. Silks reeled forcibly from a spider frequently show inconsistency in their mechanical properties,[33] and some types of silk are very difficult to reel from a spider at all. Silk collected from a web may exhibit more consistency for mechanical properties but is limited in quantity and prone to contamination from other silks and detritus. An alternative ''spider-free'' method of producing silks would be ideal for material for study, as well as future commercial application of this remarkable material. Furthermore, genetic manipulation of the silk-encoding genes opens up the possibility of spider silk as a designer material, but there are virtually no tools developed for genetic manipulation in spiders.

Microbial systems for protein production are standard for laboratory applications as they produce large amounts of material quickly and inexpensively. However, efforts to express spider cDNAs in *E. coli* resulted in no detectable expression. It was hypothesized that unusual secondary structure and codon bias found within the spider silk cDNA were responsible for the lack of expression in a microbial host. For this reason, the first synthetic spider silk genes were created. Design of synthetic spider silk genes is simplified by the repetitive nature of true spider silk genes. This allows the design of a single consensus repeat, which can then be concatenated using a number of different strategies to create large tandem arrays of this consensus repeat. Careful consideration must be used during the design of the consensus repeat to avoid formation of hairpins and other unusual secondary structures that can be found in repetitive genes. Codon preferences of the host must also be considered in the design of synthetic genes.

Strategies for creation of the concatemers of spider silk repeats fall into two broad categories: condensation or iterative. Condensation strategies entail ligation of a pool of monomeric DNA molecules. It is necessary to design these monomers so that they only assemble in a head-to-tail fashion.

One means to accomplish this is to design the monomers with nonpalindromic restriction sites at the ends. Such a strategy was employed by Capello and McGrath,[34] as well as Fukushima and Nakajima.[35] A variation on this idea was employed by Prince *et al.* [36] This group designed silk monomers with compatible but non-regenerating restriction enzyme sites at each end. A pool of these monomers were ligated and then digested with both restriction enzymes. Head-to-head and tail-to-tail ligated products would regenerate a restriction site, but correctly assembled head-to-tail constructs are not cleaved. Using this strategy, genes encoding proteins as large as 684 aa long were created. However, even single isolates demonstrated DNA "laddering," with each construct separated by the size of a single monomer unit. This indicates that the construct was unstable and subject to internal deletions.

Iterative strategies generally involve construction of a synthetic silk gene through controlled addition of silk monomer units or larger constructs to a core structure. This strategy is employed by Lewis[37] using ideas described by Kempe *et al*.[38] Advantages of this strategy include control over the number of repeats in a construct, as well as the ability to introduce other monomers into the construct. This allows for greater control and variability in the construction of larger synthetic silk genes. This strategy has been utilized to construct genes encoding silk proteins as large as 110 kDa. Constructs made in this manner appear to be stable, with no apparent DNA laddering.

Another iterative strategy was utilized by a group at DuPont to produce a synthetic MaSp2.[39] Four different overlapping oligonucleotides encoding four different native repeat sequences were used to construct a synthetic spider silk gene. Iterative polymerization was utilized to create constructs of 8 or 16 iterations of the monomer unit. This strategy resulted in a much larger construct, resulting in proteins up to 1904 aa in size. This same group also utilized this same strategy to produce a MaSp1 synthetic gene for expression in yeast.[40] In *Pichia pastoris*, the host organism, a DNA ladder effect was noted, indicating internal deletion or duplication events. However, the ladder pattern was stable, even after more than 100 doublings of the host.

MaSp1-based genes have also been used in a similar fashion to explore the mechanics of beta sheet formation.[41] In this study, MaSp1-based genes were designed to incorporate amino acids that function as trigger sites to either block or allow beta sheet formation. This strategy utilized digested monomer cassettes that were ligated into a linearized second plasmid. This strategy was repeated until constructs of the desired size were constructed. This strategy resulted in much smaller constructs (156 aa proteins), but stability was reported to be excellent.

Other organisms have also been explored as hosts for production of synthetic spider silk protein. These systems include mammalian cell culture,[42] as well as plants such as tobacco and potato.[43]

After creation of the synthetic spider silk genes and transformation into the intended host organism, purification and processing of the silk protein are necessary to produce a true synthetic silk fiber. Frequently, silk protein is isolated from homogenized cell lysate using an affinity tag, which is encoded by nucleotide sequences appended to either end of the silk-encoding gene. Popular tags include glutathione S-transferase (GST) and poly-histidine both of which have a high affinity for substances that can be immobilized on solid resins, allowing binding of soluble silk proteins to solid matrices for easy purification.

The mechanics of spinning silk are reasonably well understood, especially because of parallels that occur in the process between spiders and silkworms. For a review of the current view of this process, see Ref. [33]. Duplication of this process for artificial silk production has proven to be more difficult. One paper describes successfully spinning a solid fiber out of recombinantly produced silk proteins.[42] For this experiment, a high-concentration spinning dope was produced by dissolving synthetic MaSp2 in water. Forcing this solution through a small aperture into a methanol bath produces fibers. The resulting fibers are not as strong as true dragline silk. This observation may be accounted for by two limitations of the current design: (1) natural dragline silk is composed of two proteins, not just one and (2) the synthetic MaSp2 proteins are smaller than the natural MaSp2. Production of strong fibers using a system with the above limitations is a clear indicator that this type of system has great promise, especially, once larger synthetic silk proteins are produced. Once synthetic MaSp1 is added to the mixture, there will also be interest as to whether the produced fibers show the same molecular structures as natural silk, namely, whether rigid crystalline regions of poly-alanine beta sheets can be detected.

8. OUTLOOK AND PERSPECTIVES

In addition to being a remarkable material, spider silk provides a unique opportunity to examine the relationship between molecular structures and material properties. There are currently hypothesized structures and roles for two of the elements found in silk-encoding genes. Verifying that material properties of silk fibers are conferred by these motifs, as well as establishing the structures of these regions will open the door for true designer fibers. With this type of understanding, it becomes possible to design a material to have specific properties by engineering a gene to include varying proportions of these motifs. The stiffness of the material is also hypothesized to be conferred by the sequence of the proline-containing repeat. Further study will be

necessary to reveal the roles of other silk motifs such as GGX found in MaSp1 and the spacers found in minor ampullate and flagelliform silk.

There is currently no sequence information on the genes that encode other silks found in araneoid spiders such as tubuliform, aciniform, and piriform silks. Tubuliform and aciniform silks are used for creating egg sacs, which may be the most basal application for spider silk. These types of silks may be more closely related to the most ancestral silks, and, by this same reasoning, may be more homologous to silks from more basal taxa of spider. The evolution of the different types of spider silks may yield some surprises, similar to the cases mentioned earlier.

The unusual architecture, predicted instability, and extreme bias toward a select set of amino acids suggest that there may be fundamental differences in the manner in which these genes are handled in a spider. For example, silk-producing insects have been shown to possess alternate pathways for producing glycine, a key component of almost all types of silk. Studies have also shown that tRNA pools vary, with populations of charged tRNAs for silk constituents rising significantly during silk production. Understanding how the silk-producing cells in a silk gland handle such a large single open reading frame, as well as maintaining a supply of charged tRNA will assist in efforts to produce these silk proteins in other host organisms.

Regardless of our level of understanding, spider silk remains a remarkable material, both for its material properties, as well as the host of human applications that it presents.

REFERENCES

1. Fischer, E., 1907, About spider silk, *Hoppe-Seyler's Z. Physiol. Chem.* **53:**440–450.
2. Bell, A. L., and Peakall, D. B., 1969, Changes in the fine structure during silk protein productions in the ampullate gland of the spider *Araneus sericatus*, *J. Cell Biol.* **42:**284–295.
3. Kovoor, J., 1972, Etude histochimique et cytologique des glandes sericigenes de quelques Argiopidae, *Ann. Sci. Natl. Zool. Biol. Anim.* **14:**1–40.
4. Willcox, P. J., Gido, S. P., Muller, W., and Kaplan, D. L., 1996, Evidence for a cholesteric liquid crystalline phase in natural spinning processes, *Macromolecules* **29:**5109–5110.
5. Work, R. W., 1977, Dimensions, birefringence and force-elongation behavior, *Textile Res. J.* **47:**650–662.
6. Gosline, J. M., Denny, M. W., and DeMont, M. E., 1984, Spider Silk as Rubber, *Nature* **309:**551–552.
7. Brooks, A. E., Steinkraus, H. B., Nelson, S. R., and Lewis, R. V., 2005, An Investigation of the Divergence of Major Ampullate Silk Fibres from *Nephila clavipes* and *Argiope aurantia*. *Biomacromolecules* **6:**3095.
8. Gosline, J. M., DeMont, M. E., and Denny, M.W., 1986, The structure and properties of spider silk, *Endeavor* **10:**37–43.

9. Stauffer, S. L., Coguill, S. L., and Lewis, R. V., 1994, Comparison of physical properties of three silks from *Nephila clavipes* and *Araneus gemmoides*, *J. Arachnol*. **22:**5–11.
10. Theil, B. L., Kunkel, D. D., and Viney, C., 1994, Physical and chemical microstructure of spider dragline: A study by analytical transmission electron microscopy, *Biopolymers* **34:**1089–1097.
11. Schulz, S., 2001, Composition of the silk lipids of the spider *Nephila clavipes*, *Lipids* **36:**637–647.
12. Andersen, S. O., 1970, Amino acid composition of spider silks, *Comp. Biochem. Physiol*. **35:**705–711.
13. Xu, M., and Lewis, R.V., 1990, Structure of a protein superfiber: Spider dragline silk, *Proc. Natl. Acad. Sci. USA,* **87:**7120–7124.
14. Hinman, M. B., and Lewis, R. V., 1992, Isolation of a clone encoding a second dragline silk fibroin, J. *Biol. Chem*. **267:**19320–19324.
15. Colgin, M., 1996, *Spider Minor Ampullate Silk Proteins*, Ph.D. Thesis, University of Wyoming.
16. Colgin, M., and Lewis, R. V., 1998, Spider minor ampullate silk proteins contain new repetitive sequences and highly conserved non-silk-like "Spacer Regions", *Protein Sci*. **7:**667–672.
17. Hayashi, C., and Lewis, R. V., 1998, Spider flagelliform silk proteins, *J. Mol. Biol*. **275:**773–784.
18. Hayashi, C. Y., and Lewis, R. V., 2000, Molecular architecture and evolution of a modular spider silk protein gene, *Science* **287:**1477–1479.
19. Gatesy, J., Hayashi, C. Y., Motriuk, D., and Lewis, R. V., 2001, Extreme diversity, conservation, and convergence of spider silk fibroin sequences, *Science* **291:**2603–2605.
20. Sudo, S., Fujikawa, T., Nagakura, T., Ohkubo, T., Sakaguchi, K., Tanaka, M., Nakashima, K., Takahashi, T., 1997, Structures of mollusc shell framework proteins, *Nature* **387:**563.
21. Qin, X., Coyne, K., and Waite, J., 1977, Tough tendons. Mussel byssus has collagen with silk-like domains, *J. Biol. Chem*. **272:**32623–29.
22. Tatham, A., Shewry, P., and Miflin, B., 1984, Wheat gluten elasticity: A similar molecular basis to elastin? *FEBS Lett*. **177:**205–212.
23. Theil, B.L., Kunkel, D.D., and Viney, C. (1994). Physical and chemical microstructure of spider dragline: a study by analytical transmission electron microscopy. *Biopolymers* **34:**1089–1097.
24. Dong, Z., Lewis, R. V., and Middaugh, C. R., 1991, The molecular basis of spider silk elasticity, *Arch. Biochem. Biophys*. **284:**53–58.
25. Simmons, A., Michal, C. A., and Jelinski, L. W., 1996, Molecular orientation and two-component nature of the crystalline fraction of spider dragline silk, *Science* **271:**84–87.
26. Parkhe, A. J., Seeley, S. K., Gardner, K., Thompson L., and Lewis, R. V., 1997, Structural studies of spider silk proteins in the fiber, *J. Mol. Recognit*. **10:**1–6.
27. Termonia, Y., 1994, Molecular modeling of spider silk elasticity, *Macromolecules* **27:**7378–7381.
28. Ma, B., and Nussinov, R., 2002, Molecular dynamics simulations of alanine rich beta sheet oligomers: Insight into amyloid formation, *Protein Sci*. **11:**2335–2350.
29. Gosline, J. M., Denny, M. W., and DeMont, M. E., 1984, Spider silk as rubber, *Nature* **309:**551–552.
30. Kummerlen, J., van Beek, J. D., Vollrath, F., and Meier, B. H., 1996, Local structure in spider dragline silk investigated by two-dimensional spin-diffusion NMR, *Macromolecules* **29:**2920–2928.

31. 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. Macromol*. **24:**271–275.
32. Becker, N., Oroudjev, E., Mutz, S., Cleveland, J. P., Hansma, P. K., Hayashi, C. Y., Makarov, D. E., and Hansma, H. G., 2003, Molecular nanosprings in spider capture-silk threads. *Nat. Mater*. **2**(4):278–283.
33. Vollrath, F., and Knight, D. P., 2001, Liquid crystalline spinning of spider silk, *Nature* **410:**541–548.
34. Capello, J., and McGrath, K. P., 1994, Spinning of protein polymer fibers, in *Silk Polymers: Materials Science and Biotechnology* (D. Kaplan, W. W. Adams, B. Farmer, and C. Viney, eds.), pp. 311–327.
35. Fukushima, Y., and Nakajima, H., 1997, Secondary structural studies of biosynthetic polypeptides with a repeating sequence of glycine-rich sequence of spider dragline silk, *Chem. Lett*. 1087–1088 Volume 26.
36. Prince, J. T., McGrath, K. P., DiGirolamo, C. M., and Kaplan, D. L., 1995, Construction, cloning, and expression of synthetic genes encoding spider dragline silk, *Biochemistry* **34:**10879–10885.
37. Lewis, R. V., Hinman, M., Kothakota, S., and Fournier, M. J. (1996) Expression and purification of a spider silk protein: a new strategy for producing repetitive proteins. *Protein Exp. Purif*., **7**, 400–406.
38. Kempe, T., Kent, S. B. H., Chow, F., Peterson, S. M., Sundquist, W. I., L'Italien, J. L., Harbrecht, D., Plunkett, D., and DeLorbe, W., 1985, Multiple-copy genes: Production and modification of monomeric peptides from large multimeric fusion proteins, *Gene* **39:**239–245.
39. Fahnestock, S. R., and Irwin, S. L., 1997, Synthetic spider dragline silk proteins and their production in *E. coli*, *Appl. Microbiol. Biotechnol*. **47:**23–32.
40. Fahnestock, S. R., and Bedzyk, L. A., 1997, Production of synthetic spider dragline silk protein in *Pichia pastoris*, *Appl. Microbiol. Biotechnol*. **47:**33–39.
41. Winkler, S., Szela, S., Avtges, P., Valuzzi, R., Kirschner, D.A., and Kaplan, D. (1999) Designing recombinant spider silk proteins to control assembly. *Int. J. Biol. Macromol*., **24**, 265–270.
42. Lazaris, A., Arcidiacono, S., Huang, Y., Zhou, J. F., Duguay, F., Chretien, N., Welsh, E. A., Soares, J. W., and Karatzas, C. N., 2002, Spider silk fibers spun from soluble recombinant silk produced in mammalian cells, *Science* **295**(5554):472–476.
43. Scheller, J., Guhrs, K.-H., Grosse, F., and Conrad, U., 2001, Production of spider silk proteins in tobacco and potato, *Nat. Biotechnol*. **19:**573–577.

Chapter 6

A PROJECTION DISPLAY BASED ON A BACTERIORHODOPSIN THIN FILM

Lars Lindvold[1] and H. Lausen[2]
[1]*Stensborg Inc, Frederiksborgvej 399/Niels Bohr, DK 4000 Roskilde, Denmark*
[2]*Laser Inteface, Fruebjergvej 3, DK 2100 Copenhagen Ø, Denmark*

Abstract: A novel laser-projection display utilizing an optically addressed spatial light modulator (OASLM) is presented. The OASLM is based on the photochromic protein, bacteriorhodopsin (BR), immobilized in a thin gelatin film (80 μm). The photochromism of this material facilitates the light-induced switching of the optical absorption properties of the film. In this manner, the film can dynamically be switched between a purple-colored state and a yellow-colored state within 50 μs. Furthermore, this material is capable of more than 10^8 read–erase cycles without showing any sign of fatigue. This monochrome system (purple graphics on a yellow background) is very suitable for radar displays, due to the high contrast of the projected images. Other possible applications include public information displays in airports and railway stations. The advantages of this projection system can be summarized as: (1) flicker-free graphics due to the temporal response of the OASLM causing no eye fatigue, (2) due to the molecular nature of the photochromic process, the spatial resolution is only limited by the scanners (no pixels) and projection lens, (3) as a result of the dynamic process in the OASLM, the system can be used at different ambient light levels by simply altering the intensity of the write laser and the read projection lamp simultaneously, and (4) highly efficient use of the power from the projection lamp because it actively pumps the photochromic process.

Key words: bacteriorhodopsin thin film; coherent-to-incoherent converter; optically addressed spatial light modulator; photochromism; projection display

V. Renugopalakrishnan et al. (eds.), Bionanotechnology, 79-96.

1. INTRODUCTION

The proliferation of information technology has increased the demand for systems that permit graphics from a PC to be projected and enlarged onto a screen. Such systems are generally referred to as projection display systems. The majority of these systems utilize an electrically addressed liquid crystal display (LCD) as the spatial light modulator (SLM). The major drawback of this device is the use of two polarizers to read out the image, which alone introduces an absorption loss of at least 75% and imposes a severe limit on the power density of the projection lamp. Research results, however, have demonstrated a significant improvement in the thermal stability of polarizers based on dichroic dyes.[1]

The use of polarization filters in projection displays can be avoided if an optically addressed spatial light modulator (OASLM) based on the photochromic effect is used. One material that is particularly well-suited for this purpose is bacteriorhodopsin (BR). This material is a naturally occurring light-sensitive protein with strong photochromic properties.

2. PHOTOCHROMISM IN BACTERIORHODOPSIN

Optical applications, such as light modulators and optical storage based on organic photochromic materials have been under investigation since 1970. One of the major problems of conventional organic photochromic materials has been unwanted side reactions that have caused fatigue in the read–erase cycles. A material without such a deficiency is the BR, a naturally occurring photochromic protein. The BR is found in the cell membrane of bacteria, known as *Halobacterium Salinarium*, which live in salterns and salt lakes. In these organisms, it serves as a light-harvesting system for generating adenosine triphosphate (ATP)—the fuel of most living organisms, under anaerobic conditions. Patches of this membrane-bound protein can be extracted from bacteria without the loss of its photochromic properties. These patches are often referred to as purple membranes (PM). The PMs can be suspended in water-soluble polymers like gelatin and polyvinyl alcohol (PVA) and cast into solid thin films on a glass substrate.

The photochromic cycle (read–erase cycles of more than 10^8 has been reported) of the PM is quite unique; the ground state has a broad absorption band centered at 568 nm and a metastable state centered at 412 nm with a thermal relaxation time of 10 ms. The metastable state offers two potential advantages. First, it can be stimulated either by electrical fields or photons to decay into the ground state in 200 ns. Second, the thermal relaxation time of

this state may be extended by five orders of magnitude by suitable chemical treatment of the PM, without compromising the photostimulated decay of 200 ns. Furthermore, this process also increases the actinic sensitivity of the material by several orders of magnitude.

Although there still seem to be some uncertainties regarding the more subtle details of the photocycle in the PM, a model that describes the main events in this process has been suggested.[2] The first step is the absorption of photons by the BR, ranging from 480 to 650 nm with a quantum efficiency of about 0.64. This excites the BR into an intermediate state with a lifetime of approximately 0.5 ps. From this state, it decays through a series of intermediates with different lifetimes and absorption peaks in the visible spectrum. Two states in this photocycle are pertinent to OASLMs: the ground state B (sometimes referred to as bR) and the metastable M-state. Due to the complex nature of the entire photocycle, a simplified model of the photochromic reaction, like the one in Fig. 1, is often used to illustrate the salient points of the photocycle. The light-absorbing unit in the BR is the retinal (Fig. 2), the precursor of vitamin A. This molecule, however, does not absorb in the visible spectrum; it has a broad absorption band at about 380 nm and should absorb at about 370 nm in its protein-linked form.

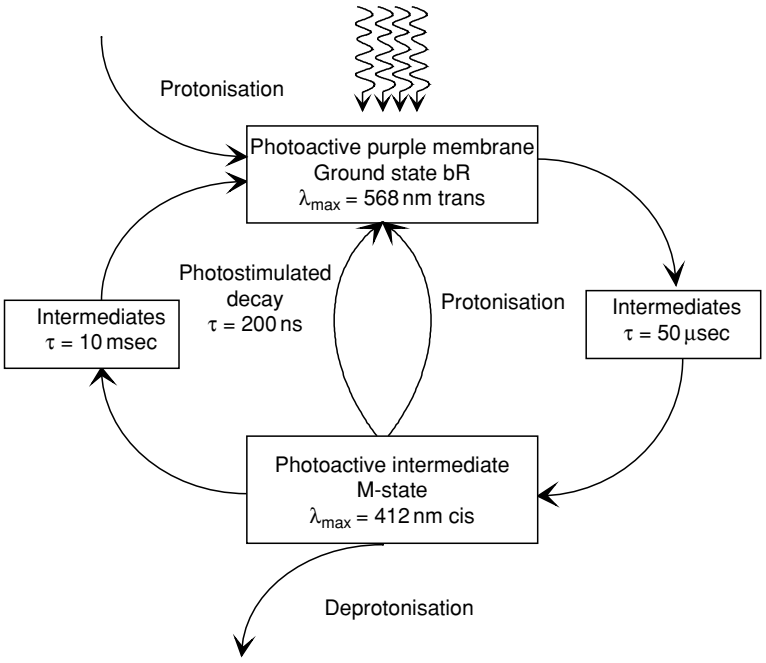


Figure 1. The photocycle of the BR adapted and modified from Ref. [3].

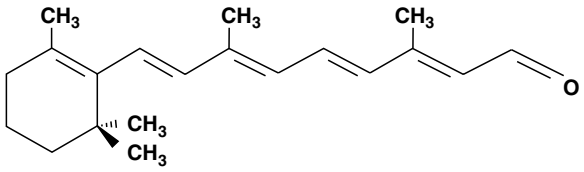


Figure 2. All-*trans* retinal, the chromophore in the BR.

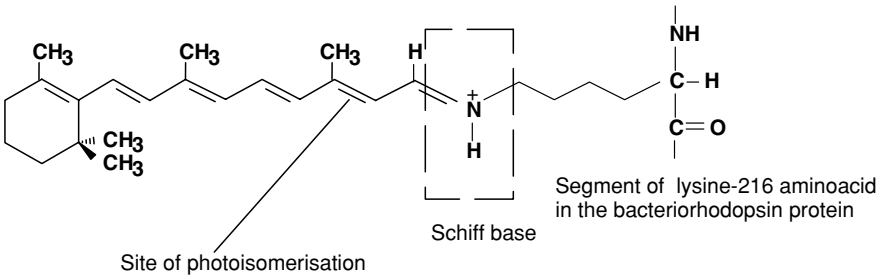


Figure 3. The protonated Schiff base bond between retinal and the protein opsin that constitutes the BR. Deprotonization of the Schiff base shifts the absorption maximum from 568 to 412 nm.

The major absorption band for the PM, however, peaks around 568 nm. This strong bathochromic shift is explained by the particular conditions under which the retinal is complexed to the transmembrane protein opsin through a so-called protonated Schiff base (Fig. 3).

This compound now has different spectroscopic properties than the all-*trans* retinal. Its position inside the lipid bilayer membrane, which constitutes the cell membrane, provides a hydrophobic pocket that shields the Schiff base from the aqueous environment of the bacterial suspension, which would otherwise destroy the Schiff base linkage by hydrolysis.[3] This is also the reason why most detergents and organic solvents tend to bleach the PM irreversibly.[4,5]

The formation of the Schiff base does not entirely explain the remarkable spectroscopic properties of the BR. Another crucial factor that determines the spectroscopic properties of the BR is the presence of point charges in the lipid-bilayer membrane. Especially divalent anions like Ca^{++} and Mg^{++} have been found to influence the proton pumping process facilitated by the photochromic reaction in the BR.[6,7] Deionization of the membrane actually shifts the absorption peak of the BR from 568 to 603 nm. As a result the PM, or rather blue membrane, loses its ability to act as a light-induced proton pump.

Apart from the conspicuous shift in absorption wavelength between the B-state and M-state, which is caused by deprotonization of the Schiff base in

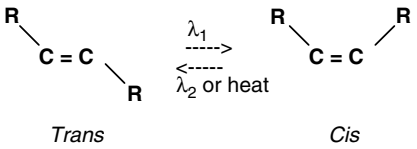


Figure 4. The conformation of the two photoactive states of the BR. Only the twist around the carbon double bond is shown for clarity.

the B-state, the B to M transition also involves a *trans* to *cis* isomerization in the retinal chromophore as shown in Fig. 4. This type of reaction is commonly known as a photoisomerization. The main feature of this type of reaction is the photoinduced excitation of a double bond, in this case, a carbon double bond, −CC−. The excited double bond loses its stiffness that permits the molecule to alter its conformation. As a result of this isomerization, an induced dipole moment is induced in the chromophore of the PM. The *cis*-state is, however, transient and follows either of the following decay pathways: (1) thermal relaxation or (2) photostimulated relaxation.

Usually, the change in dipole moment is accompanied by a change in the absorption coefficient of the two species, $\lambda_{max}(cis) > \lambda_{max}$ (*trans*). This effect is not observed in the BR, and it may be concluded that the major contribution to the photochromic effect in this material is due to the protonization/deprotonization of the Schiff base. The photoinduced anisotropy of the material is a direct result of this photoisomerization process. Furthermore, the charge translocation, i.e., the proton pumping (hydrogen ions), following the change from B-state to M-state facilitates conformational changes of the entire bR molecule. Such a reaction will perturb the induced dipole in the chromophore and, depending on the pH value; the response may be augmented or inhibited by this reaction.

3. FABRICATION OF THIN FILMS WITH BACTERIORHODOPSIN

Bacteriorhodopsin is an ideal material to immobilize in thin polymer films. Due to its high-absorption coefficient (Table 1) and molecular weight ($M_W = 26{,}000$), thin films with a thickness typically ranging from 40to100 μm will have an optical density, OD, at 568 nm, between 2 and 6.

Several methods and matrix materials are available for fabrication of thin films containing BR. Thin-film fabrication procedures for BR films differ somewhat in complexity, depending on whether alignment of the molecules in

Table 1. The optical properties of bacteriorhodopsin

	B-state	**M-state**
Peak absorption wavelength	568 nm	412 nm
Extinction coefficient ε	63,000/mol /cm	34,000/mol/cm
FWHM absorption	100 nm	100 nm
Lifetime	Infinite	10–100,000 ms

the film is required. Aligned PMs are mostly of interest in those cases in which either nonlinear phenomena like frequency doubling[8] or photovoltaic properties[9] are to be studied. Another method for immobilization of the BR is the use of the sol–gel process to form an optically transparent and porous silica matrix with the BR. The method, however, is rather time consuming (3 weeks) and elaborate compared to the film fabrication methods discussed here.[10]

The simplest method for obtaining a BR film is to spread an aqueous PM suspension on a leveled glass plate and evaporate the water in a desiccator. A BR film fabricated in this manner is very fragile and exhibits an excessive light scattering due to the aggregation of the PMs.

The best way to form a BR film, which is robust and with a high-optical quality, is to embed it in a water-soluble polymer. Fortunately, a couple of commercially available water-soluble polymers readily form dry polymer films, viz., PVA and gelatin.

In principle, any coating method for polymer films can be used. The high cost of the PM suspensions (£400/100 mg), however rules out most of the familiar coating techniques. Methods, such as spin-coating, dip-coating, and spray-coating, as well as doctor blading cause excessive waste of material. Two known methods to produce polymer films of high-optical quality are gravity settling and gel casting. Most workers in the field prefer the former method. This method is relatively simple in instrumentation and yields dry films of high-optical quality. The principle of gravity coating is quite simple; the substrate is placed on a leveled plate, and a coating solution is applied to the substrate until it covers the surface. Surface tension holds back the liquid, and solvent evaporation subsequently causes the formation of a dry thin film. The polymer content in the coating solution as well as the volume applied to the substrate controls the thickness of the film.

3.1 Gelatin As a Matrix Material for Bacteriorhodopsin

Gelatin is a well-known matrix material for the ubiquitous holographic recording materials: silver halide and dichromated gelatin. Gelatin is a biopolymer that is fabricated by extraction of collagen from animal hides

and bones, typically of bovine or porcine origin. Collagen is a so-called structure protein that is the "glue" of bone and skin cells. By treating this protein with either strong base or acid, a partial hydrolysis of the collagen takes place and yields gelatin through several intermediate washing and extraction steps.

The main constituent of gelatin is a polypeptide chain with various amino acids tied to the chain, predominantly alanine, hydroxyproline, proline, and glycine. This high content of charged molecules makes gelatin very hydrophilic that also accounts for its excellent solubility and ability to swell in water. Furthermore, this mixture of different amino acids facilitates a tremendous buffer capacity in the gelatin. This property is most fortunate; like most biochemical processes the reaction kinetics of the BR is dependent on the pH of the matrix. Gelatin, therefore, constitutes an intrinsic buffer for the BR—a property only shared by other polymer, viz., methylcellulose. Gelatin, however, differs due to its ability to form a gel spontaneously. This property makes it difficult to handle gelatin solutions with gelatin content above 10% w/v as they tend to gelate almost immediately at room temperature. This property also limits the thickness of the gelatin film to about 150 μm.

Details of the method for immobilizing the BR in gelatin films can be found in Ref. [14]. A brief summary of the most salient steps in this procedure is presented in this section. The procedure for making a gelatin coating solution consists of the following steps:

- Gelatin powder is suspended in water; during this process the gelatin swells like a sponge in water. When the solution looks like wallpaper glue, the gelatin has been properly conditioned. This takes 15–20 min depending on the gelatin concentration. It is important that this procedure is followed. If not, the gelatin will not dissolve during the next step.
- The suspension is then heated to 40–50°C, and the gelatin is completely dissolved in water within 5–10 min.
- The suspension is filtered through a 0.45-μm filter to remove particulates (dust etc.).
- The gelatin solution is thoroughly mixed with the PM suspension.[11]

The coating solution can now readily be coated onto a glass substrate that has been cleaned[12] to improve spreading and adhesion during the coating process. In order to achieve the best coating quality, the leveled substrate should be at 10–15°C and the coating solution lukewarm at 30–35°C. Drying the gel in a laminar flow bench ensures that no dust contaminates the surface. It also facilitates an even evaporation of water from the gel. This process takes about 10 h for completion.

3.2 Methods of Extending the M-State Lifetime in Gelatin Films

The description of the photochemical processes taking place in the PM presented in Section 2 gives a qualitative explanation of the photochromism in the BR. This, however, is not adequate to describe how the M-state lifetime affects the ratio of molecules in the ground state B to the excited metastable state M. Furthermore, the actinic sensitivity of the BR film cannot be inferred from such a qualitative description. We, therefore, have to resort to rate equations to give a more accurate description of the photoinduced reaction in the BR.

The reaction kinetics of the BR photocycle is generally very complicated due to its large number of intermediates. It is, however, possible to simplify the reaction if only the ground state B and the long-lived metastable M-state are taken into account. In that case, the system constitutes a unimolecular reaction of the type:

$$\mathrm{B}_{trans} \underset{k_2}{\overset{\mathrm{k}_1}{\rightleftarrows}} \mathrm{M}_{cis}.$$

Assuming a unimolecular reaction, a rate equation relating the concentration of B molecules, $[B]$, which are converted into the M-state, $[M]$, can be written as:

$$\frac{\mathrm{d}[B]}{\mathrm{d}t} = -\sigma_{\mathrm{B}}\Phi_{\mathrm{B}\to\mathrm{M}}I_1[\mathrm{B}] + (\tau^{-1}_{\mathrm{M}\to\mathrm{B}} + \sigma_{\mathrm{M}}\Phi_{\mathrm{M}\to\mathrm{B}}I_2)[M], \tag{1}$$

where σ_{B} and σ_{M} are the absorption cross sections of the B-state and M-state at the write and read wavelengths;

$\Phi_{\mathrm{B}\to\mathrm{M}}$ and $\Phi_{\mathrm{M}\to\mathrm{B}}$ are the quantum efficiencies of the photoreactions;

$\tau^{-1}_{\mathrm{M}\to\mathrm{B}}$ are the decay time of the thermal relaxation of the M-state to B-state;

and I_1 and I_2 are the intensities of the write and read beams.

The solution to Eq. (1) has previously been presented in Ref. [10]. The ratio of molecules in the ground state $[B_0]$ to the number of molecules in the metastable M-state $[M]$, assuming steady state, is stated in Eq. (2) as:

$$\frac{[M]}{[B_0]} = \frac{1}{1 + \frac{\tau^{-1}_{\mathrm{M}\to\mathrm{B}}}{k_{\mathrm{B}\to\mathrm{M}}}}, \tag{2}$$

where $k_{\mathrm{B}\to\mathrm{M}}$ is the photoinduced rate constant for exciting ground state B into the metastable M-state.

This rate constant is defined by Eq. (3) as follows:

$$k_{B\to M}(\lambda_1, I_1) = \frac{2.3026\varepsilon_B\lambda_1\Phi_{B\to M}I_1}{N_A hc}, \quad (3)$$

where ε_B is the molar extinction coefficient at wavelength λ_1 of the B-state;
λ_1 is the wavelength of the write light;
N_A is the Avogadro constant, 6.02×10^{23}/mol;
h is the Planck constant, 6.63×10^{-34} Js;
and c is the speed of light 3.00×10^8 m/s.

In order to model the effects of extending the lifetime of the M-state, a plot simulating the steady-state behavior of the M-state is presented in Fig. 5.

The curve shows that the ratio of conversion of M-state to initial B-state strongly depends on the lifetime of the M-state. It can also be inferred from the curve that the sensitivity of the bR film increases simultaneously with the increase of this time constant. The equations can furthermore be used to estimate the saturation level for given pump intensity, wavelength, or M-state lifetime.

The results presented in the previous section indicate that the extension of the M-state lifetime offers a convenient way of increasing the sensitometric properties of BR films. A number of methods have been developed, primarily as a means to facilitate spectroscopic studies of the metastable M-state. A summary of these methods can be found in a couple of studies.[13,14] It is beyond the scope of this chapter to describe these methods in detail, but the most common methods will briefly be outlined here. These are as follows:

(1) Increase the pH of the polymer film to the range 9–10.
(2) Use genetically modified BR.
(3) Exchange the retinal chromophore in the BR with retinal analogs.

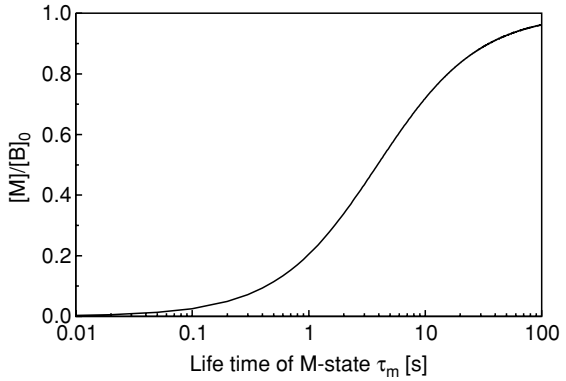


Figure 5. The diagram shows the ratio of conversion of M to initial B-state as a function of the lifetime of the M-state. The assumed values for the various parameters were taken to be: $\varepsilon B = 20{,}000$/cm/mol (at 633 nm), $I_1 = 25\,mW/cm^2$ and 0.64.

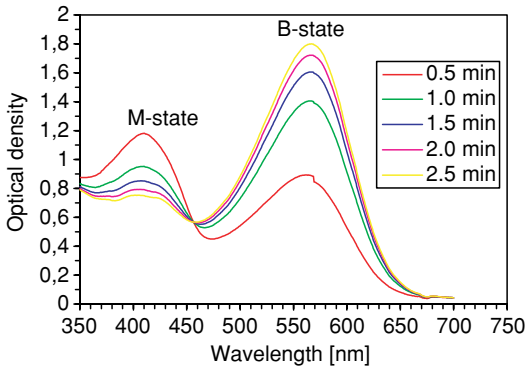


Figure 6. The decay of the M-state, recorded on a Cary 50 spectrophotometer. The sample was flooded with light from a halogen spotlight, and spectra were then recorded time lapsed with an interval of 0.5 min. The film had been chemically modified with Tris–buffer at pH 9.3. The BR was the wild type immobilized in a gelatin film with a thickness of 80 μm. The lifetime of the M-state was determined from this scan to be 47 s.

3.2.1 pH control of the film

The effect of increasing the pH of the gelatin film can be seen in Fig. 6. The intrinsic lifetime of the M-state, 10 ms, has been extended to about 47 s. The actinic sensitivity of the film has been enhanced from 100 to 10 mW/cm^2. An increase of the M-state lifetime of 10 min by increasing the pH of a PVA film containing BR to 10.3 is reported in the literature.[15]

The extension of the M-state lifetime furthermore permits to control the "afterglow" of the OASLM in projection display applications.

3.2.2 Genetically modified bacteriorhodopsin

One of the promising methods for extending the lifetime of the M-state is site-directed mutagenisis. Using this method, certain amino acids in the wild type BR, bR_{WT} can be replaced yielding a variant known as D96N.[16] The M-state of this variant is typically of the order of 1 s at pH = 7. The effect of increasing the pH value to 9 of a film containing this mutated variant is quite spectacular. The sensitivity of such a film is typically 1 mW/cm^2, which makes it possible to excite a film doped with D96N by an incandescent lamp. A variant L93T has been created in a joint effort between University of Oregon, Wayne State University, and Bend Research Inc.[17] This material has metastable intermediate, the O-state,[18] absorbing at 610 nm. This makes it possible to make projection displays with He–Ne lasers or high-power laser diodes.

3.2.3 Bacteriorhodopsin with retinal analogs

Exchanging the retinal chromophore of the BR with retinal analogs has been particularly perfected by a group of Russian scientists.[19] This method makes it possible to shift the absorption peak of the B-state significantly in the visible spectrum, typically from 500 to 630 nm. The absorption of the M-state is only altered slightly by this procedure, whereas the lifetime can be changed from a couple of seconds to several minutes. The drawback of the method however is that the durability of the PM encountered in the two previous methods is lost.

3.3 Commercial Suppliers of Bacteriorhodopsin Films

Bacteriorhodopsin is available in the following states, viz. (1) aqueous PM suspensions, (2) lyophilized (freeze-dried), and (3) polymer thin films mounted with a clear aperture of 0.5–2 in.. Both wild type and genetically modified versions are available. The optical density of the films ranges from 0.5 to 6 at 568 nm.

Bacteriorhodopsin is commercially available from at least three companies, viz. (1) Fluka (Europe), (2) Sigma-Aldrich Worldwide, and (3) Munich Innovative Biomaterials <www.mib-biotech.de/> (Germany).[20]

4. SPATIAL LIGHT MODULATORS BASED ON BACTERIORHODOPSIN

A cornerstone in most projection-display systems is the availability of devices capable of controlling light with light. Such devices are also known as SLMs. Such a device requires an optical material with generic properties that can be photostimulated, i.e., absorption, refractive index, or polarization. The permutations of these modulation schemes are legion, and the number of devices fabricated based on these mechanisms is vast.[21] It is, however, fair to state that the workhorses in optically addressed SLMs are liquid crystal (LC) devices. The LC–SLM devices are almost ideal for projection displays in terms of resolution, speed spectral, and actinic sensitivity. The major problem is the highly complex fabrication of the device. The LC–SLMs require external means to change the optical properties of the incident light. This results in devices that require complicated fabrication technologies, e.g., photoconductive layers, alignment layers, electric circuitry, etc.

4.1 Photochromic OASLM Using Bacteriorhodopsin

Due to the photochromic properties of gelatin film doped with the BR, a simple single layer OASLM can be fabricated. This OASLM permits direct optical manipulation of absorption, polarization, and phase of the incoming light. The spatial resolution of a gelatin film has been previously determined by interferometric means[22] and is shown in Fig. 7.

It is evident that the limitations in a projection display system based on an OASLM, using BR, are the scanning system and the projection lens. In order to demonstrate the applicability of the BR as an OASLM for projection displays, a system capable of projecting graphics from a PC onto a screen has been constructed.

The outline of this system is depicted in Fig. 8. The OASLM is operated in the dynamic mode i.e., light from the halogen lamp optically pumps the bR film from the ground B-state into the metastable M-state. The blue light from the krypton laser photostimulates the M-state back into the ground state. Using this pumping scheme, the limiting factor in the speed of the OASLM becomes the 50 μs rise time of the B-state to M-state. This is sufficient to match a 25 Hz frame rate.

The krypton laser is used due to its match with the peak absorption wavelength of the M-state viz., 413 nm. Other possible laser lines for this purpose are listed in Table 2. Using a laser line off the peak absorption of the M-state, however, inevitably requires a higher power, but experiments have shown that the deep blue lines of the He–Cd and Argon lasers can be used as well. The typical output power required for the experiment is 40–60 mW.

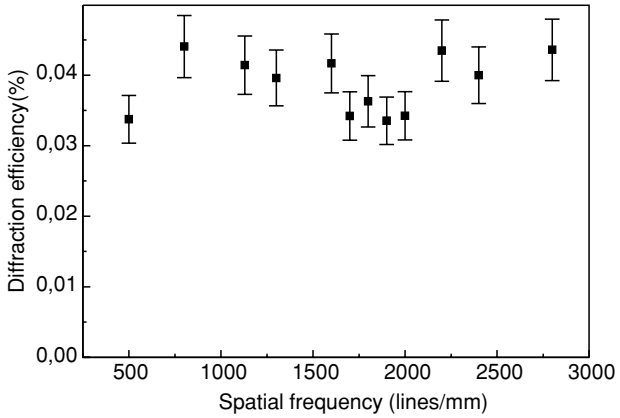


Figure 7. The spatial frequency response of the BR immobilized in a gelatin thin film.

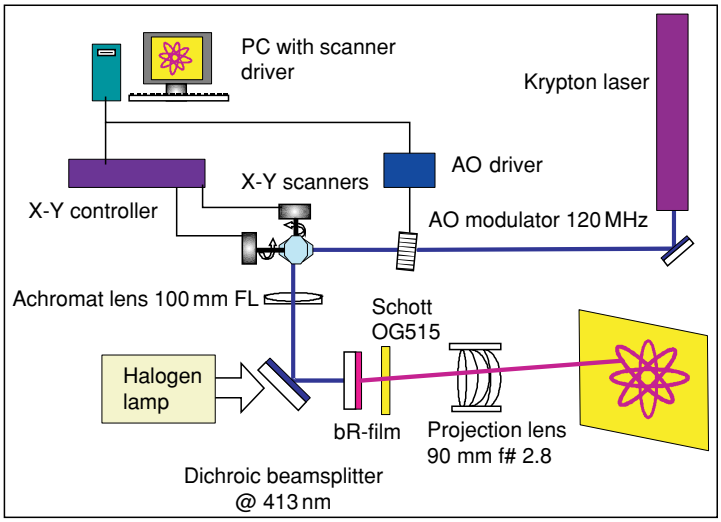


Figure 8. Outline of the experimental setup for demonstrating the laser projection display using a BR-thin film (80 μm) as an OASLM. Both wild type bRWT and genetically modified D96N material are used in the setup. The graphics from the PC are transferred to the controller that controls the X–Y scanners. These scanners deflect the laser beam mechanically. A lens focuses the laser beam on the film. The dichroic beam splitter is a color selective mirror that only reflects the laser light. The Schott OG515 filter was used to block the blue laser light (413 nm). The read-out light from the halogen lamp is used to project the laser written image on the BR-film via the lens.

Table 2. Suitable lasers for M-state photostimulation

Wavelength (nm)	Laser type
405	Diode laser
410–430	Solid state
441.5	Helium–Cadmium
454	Argon ion
458	Argon ion, solid state

4.2 OASLM Based on the Photoinduced Anisotropy of Bacteriorhodopsin

So far, only the photochromic properties of the BR are utilized in the projection display. It has been previously demonstrated that an OASLM using the photoinduced anisotropy in the BR can be fabricated[23] for incoherent-to-coherent conversion. The photoinduced anisotropy is facilitated by the

cis–trans photoisomerization of the retinal chromophore in the BR as described in Section 2.

Initial experiments in our laboratory indicate that this mode of operation also can be used for projection displays. The mode makes it possible to generate white graphics on a black background or vice versa. This mode of operation would be particularly relevant under low ambient light level applications like radar displays on a ship at night.

The difficulty of using a bR OASLM as a polarization-sensitive device is the stringent requirement on the optical quality of the film used for embedding the BR. Slight stress-induced birefringence and nonuniformities in the film will decrease the contrast ratio of the OASLM significantly.

Furthermore, it requires the use of polarizing filters that inevitably results in a loss of optical power.

5. THE DISPLAY SOFTWARE

The active bR OASLM has a high-spatial resolution 2800 /mm. This makes it possible to generate truly smooth vector graphics, without the steps and edges that characterize a pixel-based material. The system resolution, therefore, is defined by the laser modulation system and optics, rather than by the material.

The present demonstrator is controlled by a PC software system, LaserXi,[24] which generates vector line graphics by way of an acousto optic (AO) modulator and an XY laser scanner. The system generates animated sequences by processing digitized frames in real time. A small number of basic frames are subjected to build-in processes such as morphing, zooming, fading, rotation, drawing, etc. The resulting stream of coordinates is sent to the scanning system, constituting ''virtual'' frames. The update rate is 30–50 frames/s.

The idea is to separate movement information from graphic information, contrary to the classical concept of a live image as linear stream of frames. The software structure can be seen as a two-dimensional patch board with vertical coordinate streams and horizontal process selections. The horizontal layers constitute the time levels on which the processes operate. Processes and graphics can be edited independently, and there are tremendous savings in disk and memory space.

Processes may be time-linear or following a graphic curve, e.g., acceleration, cosine, etc. The system can be expanded by a live interface, modifying movements and graphics ''on the fly'' from sensor signals or other systems.

The software can be changed to run in ''Instrument'' mode: the basic frames now constitute image elements, each with its own processing attached to it. Some of the elements may be scales and numbers; others may be indicators with a follow-scale-path function and a signal interface. This will make it possible to create a projection instrument display that can be easily

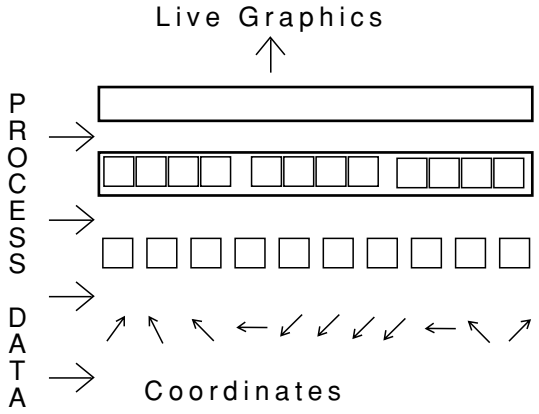


Figure 9. Structure of LaserXi^{TM} laser scanning software.

modified. Since the program and files take up very little space, the software is ideal for low-cost embedded-PC systems.

Laser-based raster scanning ''TV'' systems have been invented, using AO deflectors or combined AO and Galvo deflectors. It should be possible to integrate such systems in the BR display, if the color/laser modulation technology is developed.

A third alternative is a different type of hybrid scanning system, combining random access galvo and AO scanning. This concept works as ''live etching'' with overall vector modes and local, individually programmable, raster modes. It would be suitable for very large, high-intensity graphic displays with text, drawings, and local high-resolution images.

The internal processing speed per vector is typically 150 ns on a 1000 MHz PC. This means that only the output board and modulation system will be the limiting factor for this type of image generation.

6. CONCLUSION

The advantages of using BR-thin films for OASLM can be summarized as follows:

- The photochromic reaction in the BR shows no sign of fatigue even after 10^8 write and erase cycles
- High photosensitivity due to large photochemical quantum efficiency
- Spectral range covering a major portion of the visible region (400–650 nm)

- Long-term stability to thermal and photochemical degradation
- Commercial availability
- A spatial frequency response up to 5000 lines/mm

The advantages of using an OALSM based on the bR in this projection system can be summarized as:

- Flicker-free graphics due to the temporal response of the OASLM causing no eye fatigue
- The spatial resolution is only limited by the scanners (no pixels) and projection lens due to the molecular nature of the photochromic process
- As a result of the dynamic process in the OASLM, the system can be used at different ambient light levels by simply altering the intensity of the write laser and the read projection lamp simultaneously
- Highly efficient use of the power from the projection lamp because it is actively pumping the photochromic process

Although the projection system presented previously is projecting onto a screen, there is no problem in using it in a rear-projection configuration.

This monochrome system (purple graphics on a yellow background) is very suitable for radar displays, due to the high contrast of the projected images. Other applications include public information displays in airports and railway stations.

True color displays are under consideration using the OASLM as a phase modulator with a subsequent phase filtering process in order to convert the phase image into an amplitude image. Another possible method for this purpose would be a device based on the surface plasmon-resonance effect.

ACKNOWLEDGEMENTS

This work is supported by the European Communities as a part of the POPAM project no. 6863 under the auspices of the ESPRIT Basic Research Program.

REFERENCES

1. Mohri, S., and Matsuo, T., 1995, Highly durable dyed polarizer for use in LCD projection, *Proc. SPIE* **2407:**62–72.
2. Birge, R. R., 1990, Nature of the primary photochemical events in rhodopsin and bacteriorhodopsin, *Biochem. Biophys. Acta*. **1016:**293–327.

3. Stoeckenius, W., 1976, The PM of salt-loving bacteria, *Sci. Am*. **234**(6):38–46.
4. This statement does not hold true in all cases. Certain organic solvents like dimethyl sulphoxide can reversibly bleach the PM. This effect can be used for fundamental studies pertaining to the photochemical reactions in the PM.
5. Oesterhelt, D., Meentzen, M., and Schuhmann, L., 1973, Reversible dissociation of the purple complex in bacteriorhodopsin and indentification of 13-*cis* and all-*trans*-retinal and its chromophores, *Eur. J. Biochem*. **40:**453–463.
6. Ariki, M., and Lanyi, J. K., Characterisation of metal ion-binding sites in bacteriorhodopsin, *J. Biol. Chem*. **261:**8167–8174.
7. Chang, C. H., Chen, J.-G., Grovindjee, R., and Ebrey, T., 1985, Cation binding by bacteriorhodopsin, *Proc. Natl. Acad. Sci. USA* **82:**396–400.
8. Huang, J. Y., Chen, Z., and Lewis, A., 1989, Second-harmonic generation in purple membrane-poly(vinyl alcohol) films: Probing the dipolar characteristics of the bacteriorhodopsin chromophore in bR_{570} and M_{412}, *J. Phys. Chem*. **93:**3314–3320.
9. Vsevolodov, N. N., and Dyukova, T. V., 1994, Retinal-protein complexes as optoelectronic components, *Trends Biotechnol*. **12:**81–88.
10. Wu, S., Ellerby, L. M., Cohan, J. S., Dunn, B., El-Sayed, M. A., Selverstone, J. Valentine, and Zink, I., 1993, Bacteriorhodopsin encapsulated in transparent sol-gel glass: A new biomaterial, *Chem. Mater*. **5:**115–120.
11. The suspension is sonicated by inserting a needle vibrating at ultrasonic frequencies in the suspension. This process breaks up any aggregates of PM patches that might cause light scattering.
12. It is important that substrate (glass) has been cleaned in an ultrasonic bath with alkaline lab detergent to remove any grease from the surface. Subsequent rinsing in distilled water to remove detergent is necessary to give a clean surface.
13. Lindvold, L. R., Imam, H., and Ramanujam, P. S., The sensitometric properties of chemically modified bacteriorhodopsin films, *Proc. SPIE* **2429:**22–33.
14. Birge, R. R., 1990, Photophysics and molecular electronic applications of the rhodopsins, *Ann. Rev. Phys. Chem*. **41:**683–733.
15. Gross, R. B., Can Izgi, K., and Birge, R. R., 1992 Holographic thin films, spatial light modulators and optical associative memories based on bacteriorhodopsin, *Proc. SPIE* **1662:**186–196.
16. Bräuchle, C., Hampp, N., and Oesterhelt, D., 1991 Optical applications of bacteriorhodopsin and its mutated variants, *Adv. Mater*. **3:**420–428.
17. Downie, J. D., and Smithey, D. T., 1996, Red-shifted photochromic behaviour of a bacteriorhodopsin film made from the L93T genetic variant, *Opt. Lett*. **21:**680–862.
18. Popp, A., Wolperdinger, M., Hampp, N., Bräuchle, C., and Oesterhelt, D., 1993, Photochemical conversion of the O-intermediate to 9-*cis*-retinal containing products in bacteriorhodopsin film, *Biophys. J.* **65:**1449–1459.
19. Vsevolodov, N. N., and Dyukova, T. V., 1994, Retinal-protein complexes as optoelectronic components, *Trends Biotechnol*. **12:**81–88.
20. This company provides both suspensions of wild type bR as well as genetically modified D96N. Furthermore films in different sizes and optical densities can be obtained from this company.
21 Fisher, A. D., 1990, Spatial light modulators: functional capabilities, applications, and devices *Intl. J. Optoelec*. **5:**125.
22. Lindvold, L., Imam, H., and Ramanujam, P. S., 1995, Spatial frequency response and transient anisotropy in bacteriorhodopsin thin films, *Opt. Rev*. **2**(1):32–38.
23. Imam, H., Lindvold, L. R., and Ramanujam, P. S., 1995, Photoanisotropic incoherent to coherent converter using a bacteriorhodopsin thin film, *Opt. Lett*. **20:**225–227.
24. Developed by Laser interface.

Chapter 7

THE ROLE OF COLLAGEN IN ENERGY STORAGE AND DISSIPATION IN EXTRACELLULAR MATRIX

Frederick H. Silver, Joseph W. Freeman, Istvan Horvath, and Gino Bradica
Department of Pathology and Laboratory Medicine, UMDNJ – Robert Wood Johnson Medical School, 675 Hoes Lane, Piscataway, NJ USA 08854

Abstract: Gravity plays a central role in vertebrate development and evolution. Gravitational forces present on earth increase the muscle forces required to support locomotion and other bodily movements. Increased muscle forces lead to increased energy stored in tendons and other extracellular matrices (ECMs) during movement that is eventually dissipated as heat by muscle cells. The role of collagen molecular and fibrillar stretching and slippage in energy storage and dissipation in ECMs has been studied in tendon, articular cartilage, and blood vessel wall. The results of these studies suggest that collagen is a block copolymer of alternating flexible and rigid regions with energy storage occurring primarily in the flexible regions devoid of the aromatic amino acids, proline, and hydroxyproline. Energy storage in collagen involves increase in steric energy associated with stretching of pairs of charged amino acids that are present in flexible regions of both the molecule and fibril; these regions coincide with the positively stained bands seen in the D period when collagen fibrils are stained with heavy metals and observed in the electron microscope. Energy storage is also associated with changes in steric energy of backbone carbonyl oxygen groups and stretching of the triple helical backbone. Energy stored in collagen fibrils of the ECM is transduced by the attached cells found in the ECM into changes in cell division and protein synthesis.

The purpose of this chapter is to relate the structure of collagen fibrils in extracellular matrix (ECM) to the molecular and fibrillar bases of energy storage and dissipation. Besides this the relationship between energy storage and mechanochemical transduction by cells found in the ECM is explored and a model of how energy storage may lead to changes in cell division, gene expression, and protein synthesis in the ECM is presented.

Key words: collagen, energy storage and dissipation, mechanochemical transduction, cell-extracellular matrix interactions

V. Renugopalakrishnan et al. (eds.), Bionanotechnology, 97-116.

1. EXTRACELLULAR MATRICES

Extracellular matrices are the primary structural materials found in vertebrates that serve to maintain tissue shape (skin), aid in locomotion (bone), transmit and absorb mechanical loads (tendon and ligament), prevent premature mechanical failure (tendon, ligament, skin, and blood vessel wall), partition cells and tissues into functional units (fascia), act as scaffolds that define tissue and organ architecture (organ parenchyma), act as storage and dissipative devices for elastic energy (tendon and blood vessel wall) and as the substrate for cell adhesion, growth, and differentiation of a variety of cell types.[1–6]

Gravity plays a central role in the development and evolution of ECMs. Gravitational forces acting on mammalian tissues cause the net muscle forces required for locomotion to be higher on earth than on a body subjected to a microgravitational field. Thus, in the presence of a gravitational field, muscular forces required for locomotion or other activities are increased. As the body mass increases during development, increased muscular forces are required to propel large vertebrates; this requires that the musculoskeleton must be able to adapt by increasing the size of its functional units. This also requires that the increased energy stored in both the musculoskeleton and the cardiovascular system must be dissipated by the ECM components in order to prevent disruption of locomotion and other bodily movements. This argument suggests that inherent to biological systems subjected to gravitational forces is the need to adapt to the level of external force through anabolic and catabolic processes. For example, astronauts exposed to microgravity experience muscle atropy and bone resorption as a result of the decreased force required for locomotion or other bodily movements; this results in downregulation of transduction of mechanical energy and catabolism of muscle and bone.[7,8] The purpose of this chapter is to analyze the role of collagen in energy storage and dissipation by an ECM at the molecular and fibrillar levels. Energy storage and dissipation are also important in regulating cellular behavior in ECMs.[9]

Mechanochemical transduction is the phrase that is used to describe the biological processes by which external forces, such as gravity, influence the biochemical and genetic responses of cells and tissues. Specifically, these responses include stimulation of cell proliferation or cell apoptosis (death) and synthesis or catabolism of components of the ECM.[10,11] This chapter briefly explores the literature on mechanochemical transduction processes in an attempt to identify how the balance between external mechanical forces acting on vertebrate tissues and internal forces acting at the cellular level influence energy storage and dissipation by an ECM.

1.1 Gravity and Cellular Responses

In order for vertebrates to achieve locomotion and move efficiently, they must be able to develop muscular forces, store elastic energy in tendons, and transfer this energy to the attached joints. In addition, to support locomotion, the cardiovascular system must be able to provide the required oxygen perfusion to musculoskeletal tissues, which requires auxiliary pumping efficiency that occurs through energy storage and dissipation by blood vessel walls.[12,13]

Results of studies on the mechanophysiology of energy storage, transmission, and dissipation by the ECM suggest that although the muscle generates force, work is stored as energy in stretching the attached tendon.[1–3] Locomotion and joint movement are the result of the coordinated development of muscular forces that are transferred to bone by tendon. Although muscle can generate force, tendon stores most of the muscular force as work in a spring-like fashion.[3] Tendon not only transmits the developed muscular force to the bone but also acts as an energy storage device.[14,15,6] Elastic energy is stored through reversible stretching of collagen molecules and fibrils in a variety of ECMs.[16,6,10]

In a similar manner to tendon, elastic energy is stored in articular cartilage during musculoskeletal movement. Some of the energy transmitted from the muscle–tendon junction to the attached joint is stored during compression of articular cartilage, which lines the surface of the long bones.[6,10] A recent study concluded that articular cartilage is stretched over the joint surface and acts like a drumhead to distribute the stresses across the cartilage surface (see Fig. 1).[10] During compressive loading of the joint, some of this energy is stored elastically in the collagen fibrils.[15,10] The collagen fibrils in the surface zone of articular cartilage act to store this energy during tensile deformation that occurs as the compressive load is increased on the cartilage surface (see Fig. 1).[15,10]

In vessel wall, energy storage in collagen fibrils occurs during radial expansion; during blood flow both elastic and collagen fibers bear the loads that occur as a result of the pumping action of the heart (see Fig. 2).[12,17] Some of the collagen fibers in vessel wall appear to be in series with smooth muscle cells, suggesting that energy storage and dissipation in vessel wall are closely linked with mechanochemical transduction by cells in the wall.[10,11] The following section concentrates on describing how collagen is involved in energy storage and dissipation in the musculoskeletal system during locomotion.

2. ROLE OF COLLAGEN IN ENERGY STORAGE AND DISSIPATION IN MUSCULOSKELETAL ECM

Elastic energy storage in tendons, and in the legs and feet of many animals is an important mechanism that saves substantial quantities of muscular

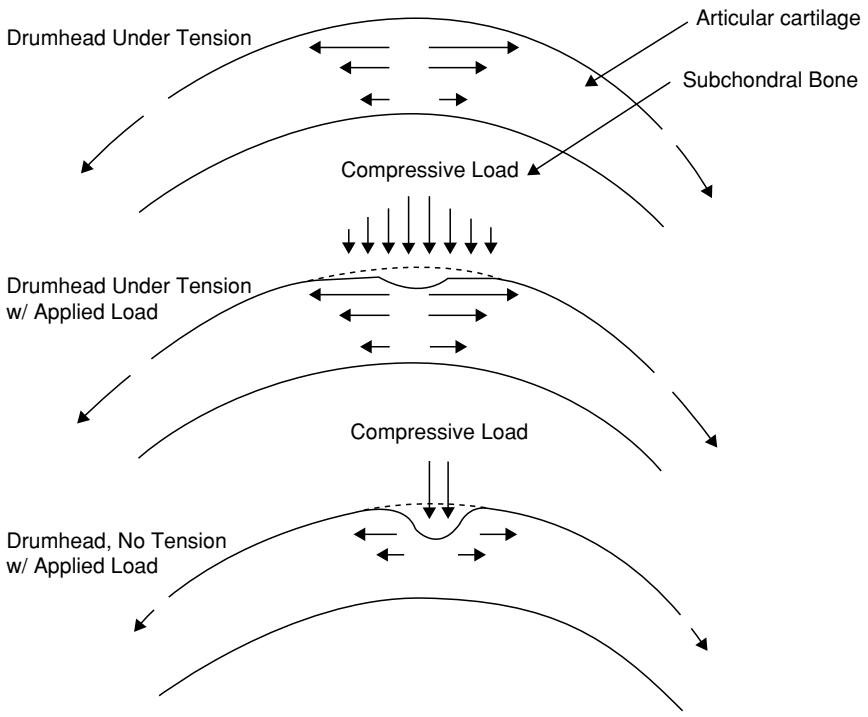


Figure 1. Diagram illustrating the tension present in the superficial zone of articular cartilage. Normal articular cartilage uncurls when it is removed from the subchondral bone. This is due to the presence of tension (top diagram) in the superficial zone that exists to pull the cartilage taught across the surface. The presence of tension in the superficial zone makes articular cartilage behave like a drumhead, allowing compressive forces applied to the surface at specific points to be distributed (middle diagram) across the surface to lower the local stresses. In osteoarthritic cartilage (bottom), loss of the tension leads to concentration of the compressive load, which promotes cartilage degradation.

energy during locomotion.[1,2] During normal gait, potential energy is stored as strain energy in the muscles and tendons that are stretched upon impact with the ground.[1,2] Elastic recoil, primarily by the tendons, converts most of the stored energy back to kinetic energy. Elastic energy storage in tendons has been studied in several animal models. In the pig, the digital flexor tendons are involved in the elastic storage of strain energy[4]; the amount of elastic energy stored in the digital flexor tendons decreases with age after the animal reaches maturity. In the turkey, direct measurement of force and fiber length in the lateral gastrocnemius muscle reveals that the active muscle produces high force but little work, while the tendon produces much of the work through elastic deformation and recovery.[3] A recent study in horses suggests that the

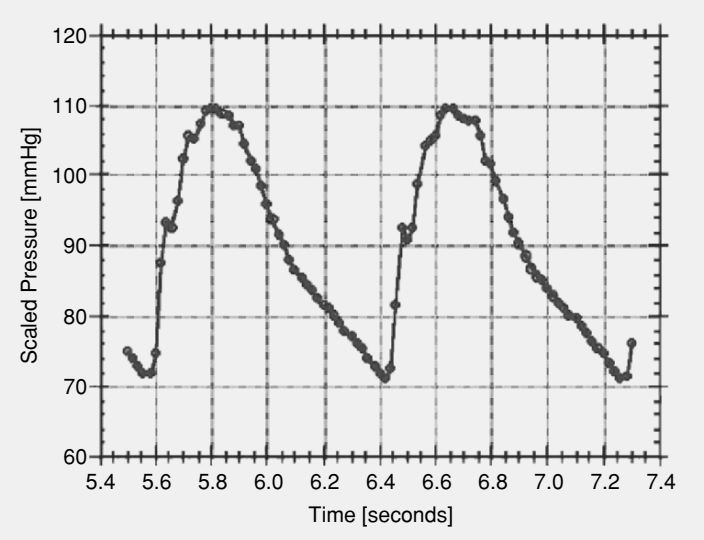


Figure 2. Pressure–time curves for human carotid artery. This waveform shows the loading (rising slope) and unloading (decaying slope) of the carotid artery during the beginning and end of systole. The initial loading portion of the curve that represents energy storage, reflects loading of elastic fibers (initial rapid slope increase) and then loading of collagen fibers in series with smooth muscle (second less steep slope increase). The unloading portion of the curve reflects unloading of elastic fibers (sharp slope decline) and then the slow decay in slope associate with unloading of collagen fibers. Note the time dependence (viscoelasticity) of the unloading of collagen fibers during which energy is dissipated.

muscle–tendon unit not only stores energy during extension, but it also dissipates the energy after extension is complete.[18]

In accomplishing either energy storage or force transmission, tendon must cyclically undergo reversible deformation without undergoing irreversible structural changes. Structurally, this tissue is made up of dense fibrous connective tissue composed of aligned collagen fibers separated by rows of cells (see Fig. 3). Tendon contains regions with specialized structures and functions such as connections to muscle and bone. This chapter considers only the region that makes up the mid-substance of tendon to simplify the analysis. In the following section, the composition and structure of tendon are considered.

3. STRUCTURE AND DEVELOPMENT OF DENSE REGULAR CONNECTIVE TISSUE

Dense regular fibrous connective tissue, also termed ECM, is made up of cells, aligned collagen and elastic fibers, proteoglycans, and water. Although

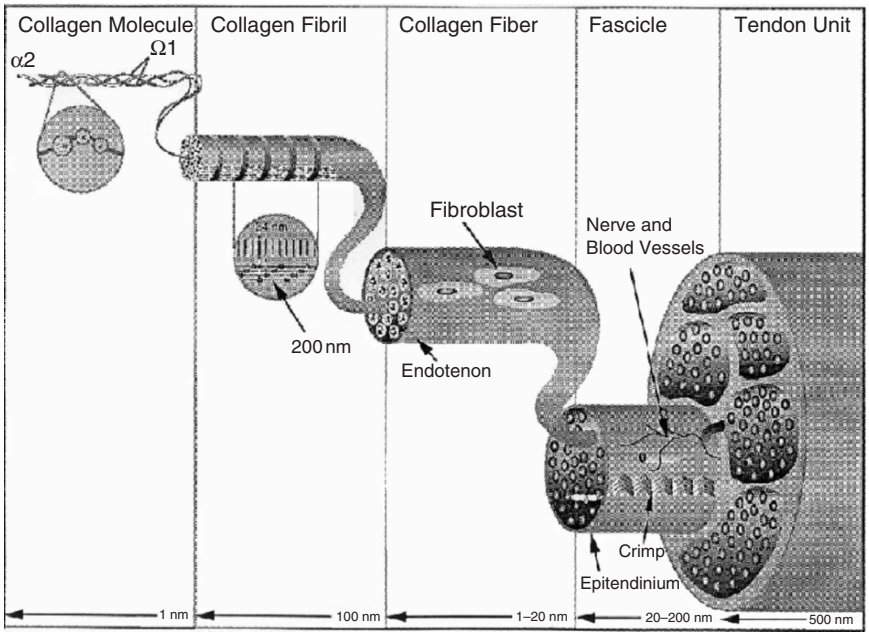


Figure 3. Structural hierarchy in tendon. This diagram illustrates the relationship between collagen molecules, fibrils, fibers, fascicles, and tendon units. Although the diagram does not show fibril subunits, collagen fibrils appear to be self-assembled from intermediates that are integrated within the fibril. The structural hierarchy in ligament and joint capsule is similar to that of tendon although other components, such as elastic tissue, are also present in the latter types of connective tissue.

elastic fibers are present in ligament and other tissues, in dense regular connective tissue that is associated with tendon, elastic fibers are only minor components. Collagen is the most abundant protein in dense fibrous connective tissue and forms essential mechanical building blocks in the musculoskeletal system. It can be found in both fibril and nonfibril forming forms. The fibril-forming collagens provide the structural framework of tissues; they include types I, II, III, V, and XI collagen (see Refs. [5,10,11]). All fibril-forming collagens self-assemble into cross striated fibrils with a characteristic 67 nm repeat made up of alternating hole and overlap regions (see Fig. 4); they all share a triple helical region that is roughly a 1000-amino acid (aa) residues long with a length of about 300 nm.[5] These collagens are synthesized within cells in a precursor form termed procollagen, which has amino and carboxyl-terminal nonhelical ends that are about 15.0 and 10.0 nm long, respectively (see Fig. 5). In addition, fibril-associated collagens with interrupted triple helical sequences (FACIT) collagens, such as type XII collagen,

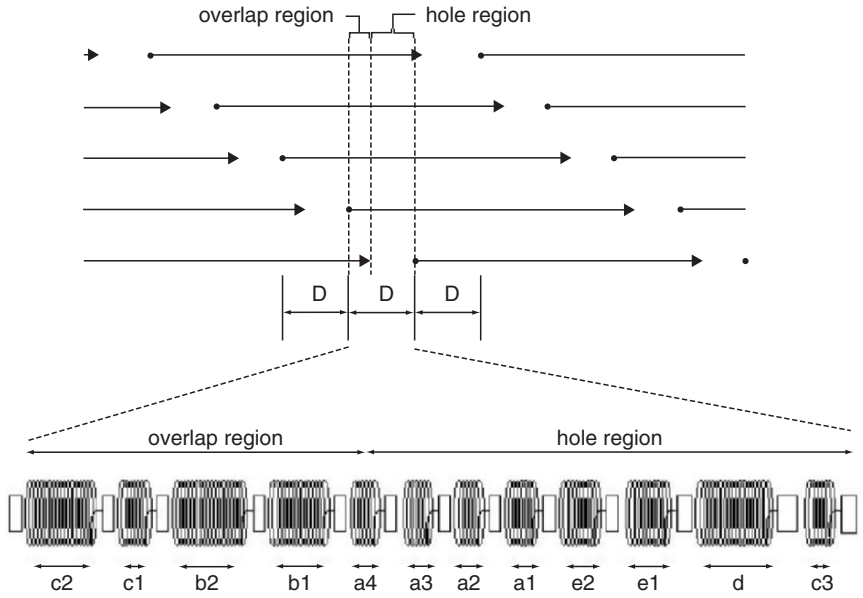


Figure 4. Packing pattern of collagen molecules in collagen fibrils. In the quarter-stagger packing pattern, collagen molecules, 4.4D long, are staggered with respect to their neighbors. In tendon, the collagen molecules are shifted with respect to each other by a distance D equal to 67 nm. When the collagen molecules are stained with metal ions and then viewed in the electron microscope, a series of light and dark bands are observed across the axis of the fibril and designated b2, b1, a4, a3, a2, a1, e2, e1, d, c2, and c1. The distance D is made up of a hole region of about 0.6D and an overlap region of about 0.4D. The D period is the characteristic fingerprint of fibrous collagen. Note the collagen molecules form head-to-tail cross-links in tendon.

are found on the surface of collagen fibrils and may connect fibrillar collagens to other components of the ECM.[19]

Tendon is composed predominantly of type I collagen, although it contains small amounts of types III and V.[20,21,22,5] The type III collagen content has been reported to be 5% for tendon.[23] The high content of type I collagen in tendon not only leads to mechanical stability but also promotes elastic energy storage.

The mechanical stability of dense fibrous connective tissue and type I collagen in tendon is related to the rod-like structure of type I collagen and its inherent flexibility.[24,6] The rod-like behavior of type I collagen was first established based on measurement of the translational diffusion coefficient;[25,26,27] however, later measurements indicated that the type I collagen molecule had numerous bends and was not completely rigid.[28] Types II and III collagen molecules have slightly higher translational diffusion coefficients

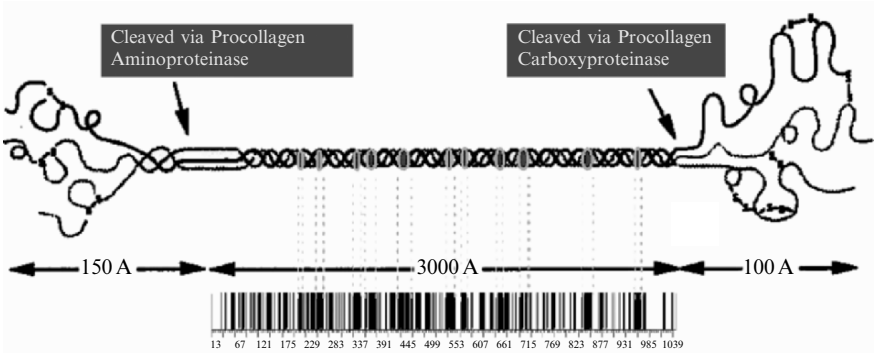


Figure 5. Structure of collagen molecule. This diagram illustrates the structure of procollagen type I, the precursor of collagen molecules found in tendon. The procollagen molecule shown consists aminopropeptides (large left hand portion of molecule), a triple helical region (3000 Å long), and a carboxylic propeptide (right hand end of molecule). The amino (N-) and carboxylic (C-) propeptides are cleaved by specific proteinases during collagen self-assembly and aid in regulating linear and lateral fibril growth during self-assembly of collagen fibrils. The circles in the triple helix represent major sequences devoid of proline and hydroxyproline that are the likely sites of folds where flexibility is introduced into a normally rigid helix. The striated pattern shown below the helical portion of the molecule is a diagrammatic representation of the flexible (dark bands) and rigid regions (light bands) found in the triple helix.

and slightly shorter end-to-end distances,[28] suggesting that the later collagen molecules are more flexible than type I collagen. Results of recent modeling studies suggest that the types I, II, and III collagen molecules are made up of alternating rigid and flexible domains that are conserved within the collagen fibril; types II and III collagen molecules appear more flexible than type I.[6] The flexible domains to a first approximation coincide with the positively stained bands in the collagen D period observed in the electron microscope after staining with heavy metals.[6,16]

The molecular basis for the flexibility of type I collagen comes from sequences that lack the amino acids proline and hydroxyproline. These sequences are the sites where bends can occur in the triple helix. Five sites have been identified by Hofmann *et al.*[29] based on electron microscopic images; these sites occur at characteristic distances of 30–45, 90–105, 150–157.5, 210–217.5, and 270–277.5 nm from the C-terminus of the molecule. However, results of recent modeling studies suggest that additional flexible sites are present.[6] Results of another modeling study suggest that sequences without the amino acid residues proline and hydroxyproline are able to form internal loops[30] that give these regions more flexibility than other regions of the triple helix. Stereochemical maps constructed for dipeptides containing amino and aromatic imino acid residues suggest that in the absence of proline and

hydroxyproline the number of available conformations of the molecule and consequently the flexibility are increased.[31,5,6] A flexibility profile as a function of axial displacement for the type I collagen triple helix is shown in Fig. 5. This diagram suggests that the collagen triple helix can be considered a composite of regions with varying degrees of stiffness; regions of the molecule devoid of proline and hydroxyproline appear to have the highest flexibility, whereas regions with the sequence Gly–Pro–Hyp or sequences containing lysine or hydroxylysine are very rigid. This variation in molecular flexibility affects collagen self-assembly as well as the resulting energy storage and dissipative properties of collagen.

Energy storage during stretching of collagen is postulated to be associated with increases in: (a) steric energy of van der Waals and electrostatic interactions that occur when pairs of oppositely charged amino acid side chains on collagen are stretched, (b) stretching of electrostatic interactions between amino acid side chains and the carbonyl oxygen found in the backbone, and (c) stretching of the molecular backbone[10,11] (see Fig. 6). Modeling studies on a five-molecule wide subfibrillar structure composed of quarter-staggered collagen molecules suggest that the most flexible regions are bands a3, a4, a2, b1, b2, and d; other regions between bands b1 and b2 and between bands c1 and c2 are also flexible (see Fig. 4). These areas are consistent with the flexible sites identified by Hofmann *et al.*[29] and with Ramachandran's plots based on peptide sequences observed in collagen.[31]

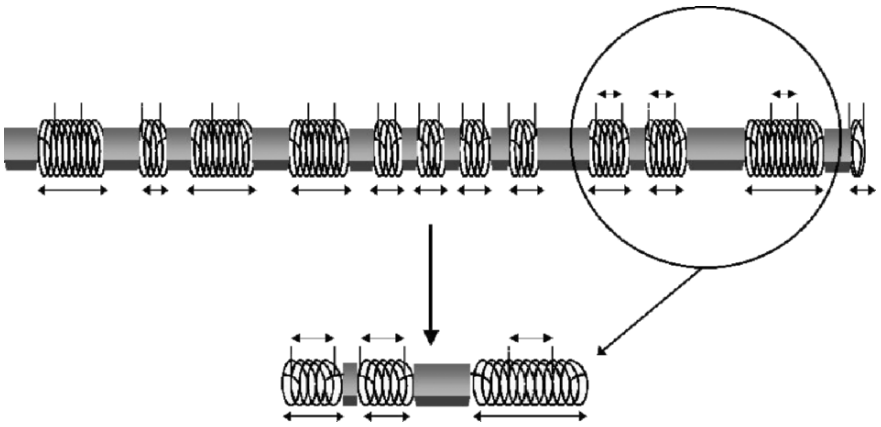


Figure 6. Diagrammatic representation of energy storage in collagen fibril. When the collagen fibril is stretched to tension, part of the energy is stored as potential energy in the flexible regions contained in the positive staining bands. This diagram illustrates how external mechanical loading is stored as steric potential energy by stretching the spring (helix) backbone, which also leads to increased distances between pairs of oppositely charged atoms.

4. MECHANICAL PROPERTIES OF ECM

Much of our understanding of the relationship between the structure and mechanical properties of dense regular connective tissue comes from analysis of the mechanical properties of developing tendon. The properties of developing tendons rapidly change just prior to the onset of locomotion. McBride *et al*.[32,33] reported that the ultimate tensile strength (UTS) of developing chick extensor tendons increases from about 2 MPa (day 14 embryonic) to 60 MPa 2 days after birth. This rapid increase in UTS is not associated with changes in fibril diameter but with increases in collagen fibril lengths,[33] which is related to the viscoelastic properties of tendons.[14] The viscoelastic properties of ECM have been studied using incremental stress–strain tests.[14,16,15]

Application of incremental strains to tendon, followed by measuring the initial and equilibrium stresses, yields information on the molecular and fibrillar bases for energy storage and dissipation in dense regular connective tissue (Fig. 7). From the equilibrium stresses obtained at different strains, one can plot an elastic stress–strain curve, while from the difference between the total and elastic stress, one constructs a "viscous" stress–strain curve.[14,16,15] The slope of the elastic stress–strain curve is proportional to the elastic modulus of the collagen molecule and fibril, while the area under the elastic

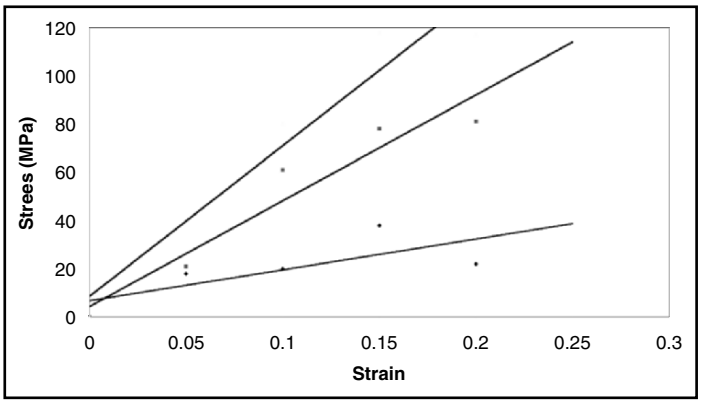


Figure 7. Incremental stress–strain curves for rat-tail tendon. This figure is a plot of stress versus strain obtained from incremental stress–strain testing of rat-tail tendon in tension at room temperature and at strain rate of 10% per min. The top curve represents the initial stress at increasing strains obtained at the end of the strain increment prior to relaxation of the specimen. The middle curve represents the elastic stress as a function of strain obtained by plotting the stress at equilibrium after relaxation has occurred. The bottom curve is the viscous stress–strain curve, which is obtained by subtracting the middle curve from the top curve. The energy stored is equal to the area under the elastic stress–strain curve (middle curve), while the energy dissipated is proportional to the area under the viscous stress–strain curve (bottom curve).

stress–strain curve is related to the elastic energy stored during stretching. Using hydrodynamic theory, the viscous stress is proportional to the fibril length and the area under the viscous stress–strain curve is proportional to the energy dissipated.[14,16,15]

On a molecular basis, the slope of the elastic stress–strain curve (elastic modulus) can be related to the stretching of collagen triple-helices within cross-linked collagen fibrils.[14,15] Results of molecular modeling studies suggest that stretching first occurs in the flexible domains of the collagen molecule (see Figs. 4 and 6). Studies on self-assembled type I collagen fibers show that in the absence of cross-links, the elastic slope is reduced[14] suggesting that cross-links are important in mechanical coupling between collagen molecules.

Results of modeling studies on self-assembled type I collagen fibers suggest that elastic energy storage measured as the area under the elastic stress–strain curve is closely related to the changes in steric energy calculated using a molecular model. This model assumes that energy storage occurs as a result of stretching pairs of oppositely charged amino acids present in the flexible regions of the collagen fibril, stretching of the triple-helix backbone, and stretching of hydrogen bonds involving backbone carbonyl oxygen (Fig. 8).

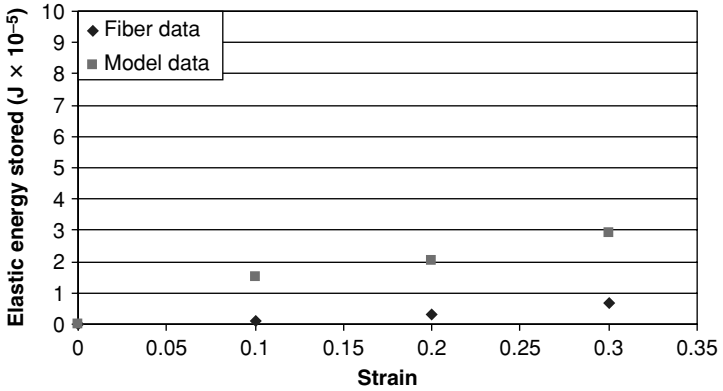


Figure 8. Elastic energy storage in collagen fibrils. Energy versus strain calculated using a SYBYL molecular modeling program for type I collagen fibrils thereon determined from experimental fiber measurements. Fiber data was obtained by measuring the area under experimental elastic stress–strain curves of self-assembled type I collagen fibrils. This data was obtained using an Instron mechanical testing device operating at a strain rate of 10% per min. Model data was calculated using energy minimization techniques and a SYBYL molecular modeling program. A five stranded collagen fibril was constructed, and the energy was minimized in the absence of water molecules. Minimized energy calculations were then conducted on the microfibril after strains of 1–3%, and the difference in energy of stretched minus unstretched microfibrils was plotted versus strain after the energy data corrected for the fiber cross-sectional area. Additional calculations using the SYBYL indicated that in the presence of structural water molecules the difference between the experimental and theoretical energy is reduced.

Elastic energy storage is reported to be impaired in osteoarthritic cartilage (see Fig. 9) and associated with decreased collagen fibril lengths; the observation that energy storage and dissipation are both decreased in osteoarthritis suggests that the mechanisms of energy storage and dissipation are related in collagen[15] (see Fig. 10).

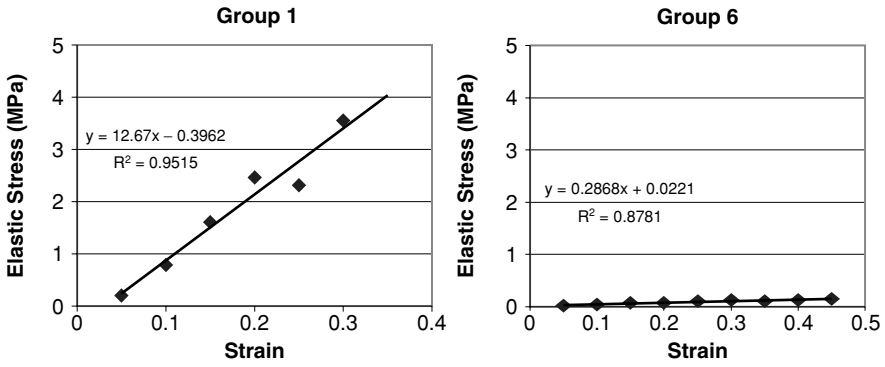


Figure 9. Impaired elastic energy storage associated with progression of osteoarthritis. Plots of elastic stress versus strain for quasi-normal group 1 and osteoarthritic articular cartilage (group 6) as a function of the degree of fibrillation and fissure formation (increased group number). The area under the elastic stress–strain curve is equal to the elastic energy stored. Note the rapid loss in the ability of osteoarthritic cartilage to store elastic energy as indicated by the smaller area under the elastic stress–strain curve.

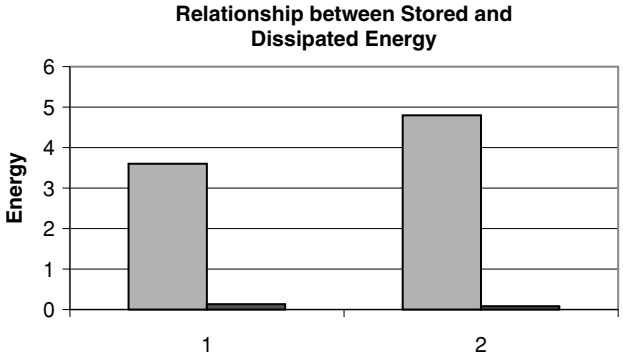


Figure 10. Loss of energy storage and dissipative properties of osteoarthritic cartilage. Elastic energy stored (1) and dissipated (2) for normal (left bar) and osteoarthritic (right bar) articular cartilage. Energy stored and dissipated were calculated as the areas under the elastic (stored) and viscous (dissipated) stress–strain curves. Note the large decrease in energy stored and dissipated for osteoarthritic cartilage.

Energy storage in collagen fibrils during tensile stretching of the flexible domains can be transferred into changes in cell membrane shape and cytoskeletal structure via conformation and other physical effects as illustrated in Fig. 11. Key to these structural changes in the cell are changes in conformation of the collagen triple helix and fibril as manifested by changes in the collagen positively stained banding pattern.

5. EFFECTS OF PHYSICAL FORCES ON CELL–ECM INTERACTIONS: INTRACELLULAR SIGNAL TRANSDUCTION MECHANISMS

During evolution, mechanical as well as other stimuli, such as light, pH, and temperature are, critical environmental parameters that are sensed by cells. It is postulated that intracellular signal transduction pathways were originally developed by cells to respond to these basic stimuli; this response is preserved and thought to develop further during evolution.[34] The signaling events initiated by mechanical, as well as other stimuli include generation of secondary messengers (molecules that are found within the cell cytoplasm, which activate other messengers that move from the cytoplasm to the cell nucleus), changes in the phosphorylation status of intracellular proteins (proteins within the secondary messenger pathways are activated and deactivated by adding or removing a phosphate group), amplification through enzymatic cascades (the signal is amplified by going through a series of steps), and transmission via a complicated network of signaling molecules (changes in cell nuclear events occur only after a series of molecules are activated) (see Fig. 11).[34]

The phosphorylation and dephosphorylation of molecules, such as enzymes found within the cell, many of which are attached to the cell membrane or in proximity, is one of the ways in which these signaling pathways are activated. Protein phosphorylation is also involved in control of cell proliferation and differentiation[35] with phosphorylation being controlled genetically by the synthesis and regulation of molecules termed protein kinases (PKs). External physical forces are transduced by the activation of different PKs. The effect of PK activation is regulated through the activation of secondary messengers that transmit signals within a cell. An example of a PK is focal adhesion kinase (FAK), which is phosphorylated after integrin binding to a substrate (see Fig. 12); phosphorylation of the FAK is integrin mediated and leads to activation of an intracellular signaling pathway termed mitogen-activated protein kinase (MAPK).[36] The MAPKs constitute a family of kinases that regulate cellular activities ranging from gene expression, mitosis, movement, metabolism, and programmed cell death (apoptosis).[37] Substrates

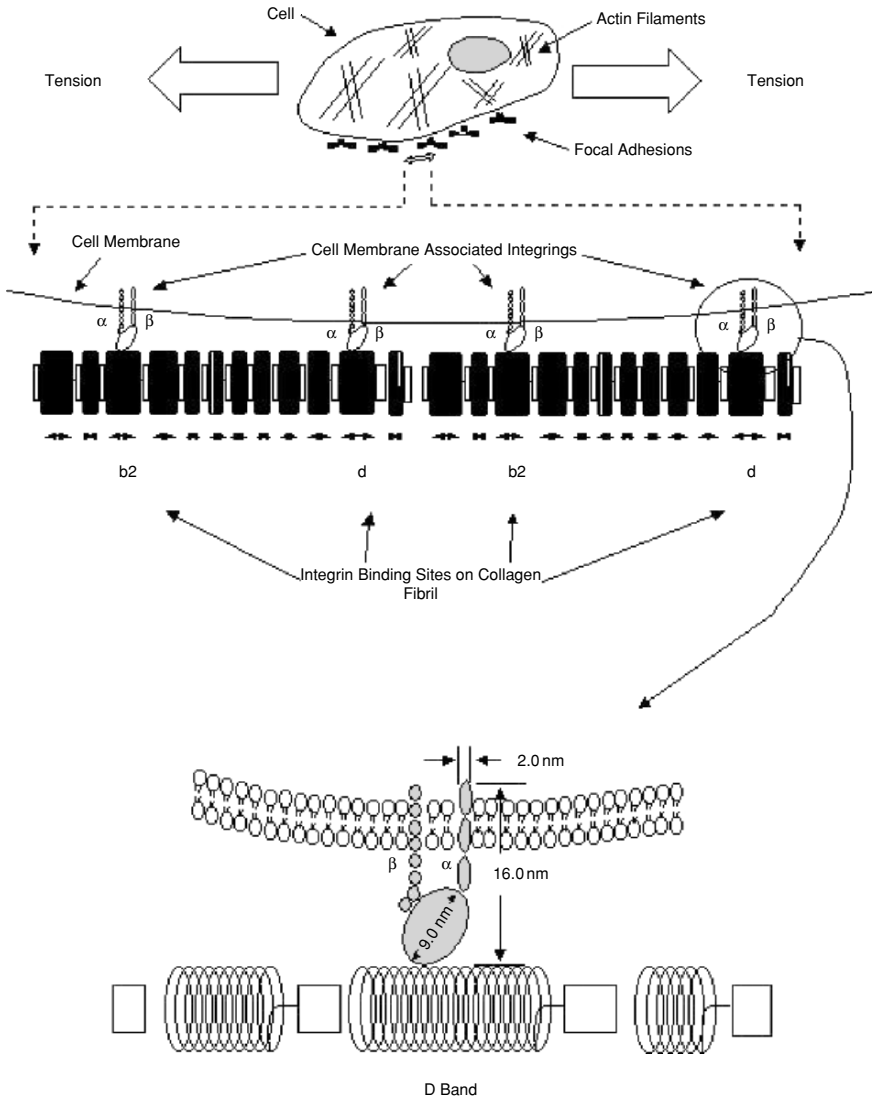


Figure 11. Diagram of energy transfer from collagen fibril to cell cytoskeleton. This diagram illustrates how a cell binds to a collagen fibril at the b2 and d bands via integrin α and β chains. The collagen fibril is represented by a series of springs (flexible regions) and solid cylinders (rigid regions) that transfer tensile loads and elastic energy to the cell cytoskeleton via stretch induced conformational changes. An expanded view is shown of the d band (see bottom) illustrating how energy transfer could occur between collagen fibrils and cell surface integrins.

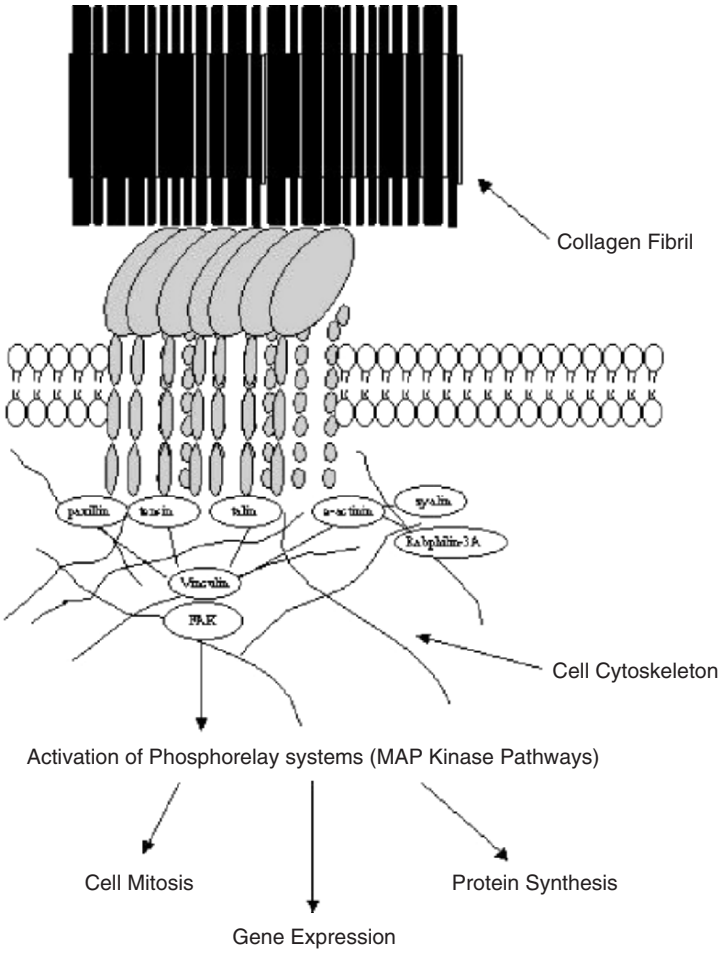


Figure 12. Diagram illustrating how tensile energy transfer leads to synthesis of new tissue in an ECM. This diagram illustrates how tensile stretching of the collagen fibril shown in Fig. 11 could cause energy transfer through clusters of integrin subunits in a focal adhesion complex. This energy transfer to the attached integrin subunits may lead to physical changes occurring in the cell cytoskeleton (actin filaments and other cytoskeletal molecules shown) that affect activation of the FAK and stimulation of secondary messengers. Secondary messengers activate several different pathways that are part of the phosphorelay system. This leads to the activation of MAPK, up-regulation of protein synthesis and cell mitotic activity, and changes in gene expression.

that are phosphorylated include other protein kinases, transcription factors, and cytoskeletal proteins. Down-regulation of the MAPK occurs through the activity of protein phosphatases that remove the phosphates that were transferred to the protein substrate by the MAPK.

At least two mechanisms exist by which external loads and energy storage in the ECM can affect gene expression of resident cells in the same; these include cell–ECM interactions and cell–cell interactions.[6,10] Integrin adhesion receptors that connect ECM components and cytoskeletal elements have been implicated in mediating signal transduction through the cell membrane in both directions.[38] Integrin adhesion receptors are heterodimers of two different subunits termed α and β.[38] They contain a large ECM domain responsible for binding to substrates, a single transmembrane domain, and a cytoplasmic domain that in most cases consists of 20–70 aa residues.[39] They mediate signal transduction through the cell membrane in both directions; binding of ligands to integrins transmits signals into the cell and results in cytoskeletal reorganization, gene expression, and cellular differentiation (outside-in signaling). On the other side, signals within the cell can also propagate through integrins and regulate integrin–ligand binding affinity and cell adhesion (inside-out signaling).[39]

Mechanochemical transduction is postulated to occur by several other mechanisms besides direct stretching of protein-cell surface integrin binding sites that occur on all eukaryotic cells as reviewed recently.[6,10] Stress induced conformational changes in the fibrillar structure of collagen may alter integrin structure and lead to activation of several secondary messenger pathways within the cell. Activation of these pathways leads to altered regulation of genes that synthesize and catabolize ECM proteins as well as to alterations in cell division. The second mechanism by which mechanochemical transduction occurs, involves deformation of gap junctions that contain calcium sensitive stretch receptors that in some tissues form dendritic patterns. Once activated, these channels trigger secondary messenger activation through pathways similar to those involved in integrin dependent activation and allow cell-to-cell communications between cells with similar and different phenotypes.[6,10] Another mechanism by which mechanochemical transduction occurs is through activation of ion channels in the cell membrane. Mechanical forces have been shown to alter cell membrane ion channel permeability associated with Ca^{2+} and other ion fluxes.[6,10] In addition, the application of mechanical forces to cells leads to the activation of tyrosine kinase growth factor receptors even in the absence of growth factor binding. These are some of the mechanisms that have evolved in vertebrates by which cells respond to changes in external forces.

6. SUMMARY

Energy storage in dense regular connective tissue occurs by stretching of flexible regions in collagen molecules and fibrils that are cross-linked into a three-dimensional network. External forces applied to dense regular

connective tissue not only cause storage and dissipation of energy within collagen molecules and fibrils, but also lead to the transduction of energy into cellular changes including changes in cell division and gene expression through cell–ECM and cell–cell interactions. Cell–ECM interactions involve cell surface integrins and focal adhesion complexes that activate secondary messengers attached to the cell membrane. Focal adhesion complexes once formed affect several pathways including the MAPK pathway. Cell–cell interactions occur by activating cadherin dependent cell junction stretch receptors that lead to release of intercellular calcium and activation of secondary messenger pathways. Mechanical stretching of cells also leads to alterations in cell membrane ion permeability that affects cell function. All of these processes affect the balance between external loading, energy storage, and cell generated contractile forces that ultimately lead to changes in composition and mechanical properties of the ECM. Thus, energy storage within collagen fibrils found in ECMs not only has implications in biological control mechanisms involved in growth and development of multicellular organisms, but it also provides a mechanism to study how mechanical energy is converted into chemical energy within the cell.

ACKNOWLEDGEMENTS

The authors would like to thank the graduate students Gurinder P. Seehra and Ali Ibrahimi for their assistance in preparing this chapter.

REFERENCES

1. Alexander, R. M., 1983, *Animal Mechanics*, 2nd edn., Blackwell Scientific, Oxford, UK.
2. Alexander, R. M., 1984, Elastic energy stores in running vertebrates, *A. Zool.* **24:**85–94.
3. Roberts, T. J., Marsh, R. L., Weyand, P. G., and Taylor, C. R., 1997, Muscular force in running turkeys: The economy of minimizing work, *Science* **275:**113–115.
4. Shadwick, R. E., 1990, Elastic energy storage in tendons: Mechanical differences related to function and age, *J. Appl. Physiol.* **68:**1033–1040.
5. Silver, F. H., and Christiansen, D. L., 1999, *Biomaterials Science and Biocompatibility*, Chapter 6, Springer, New York.
6. Silver, F. H., Horvath, I., and Foran, D. J., 2002, Mechanical implications of the domain structure of fibril forming collagens: Comparison of the molecular and fibrillar flexibilities of the α-chains found in types I, II, and III collagen, *J. Theor. Biol.* **216:**243–254.
7. Collet, P., Uebelhart, D., Vico, L., Moro, L., Hartmann, D., and Roth, M., 1997, Effects of 1- and 6-month spaceflight on bone mass area and biochemistry in two humans, *Bone* **20:**547–551.

8. Vandenburgh, H. H., 1992, Mechanical forces and their second messengers in stimulating cell growth in vitro, *Am. J. Physiol*. **262:**R350–R355.
9. Chiquet, M., 1999, Regulation of extracellular matrix gene expression by mechanical stress, *Matrix Biol*. **18:**417–425.
10. Silver, F. H., and Bradica, G., 2002, Mechanobiology of cartilage: How do internal and external stresses affect mechanochemical transduction and elastic energy storage? *Biomechan. Model Mechanobiol*. **1:**219–238.
11. Silver, F. H., Siperko, L. M., and Seehra, G. P., 2003, Mechanobiology of force transduction in dermal tissue, *Skin Res. Technol*. **9:**3–23.
12. Kalath, S., Tsipouras, P., and Silver, F. H., 1986, Non-invasive assessment of aortic mechanical properties, *Annals Biomed. Eng*. **14:**513–524.
13. Silver, F. H., Christiansen, D. L., and Buntin, C. M., 1989, Mechanical properties of the aorta: A review, *Crit. Rev. Biomed. Engg*. **17:**323–258, 1989.
14. Silver, F. H., Christiansen, D. L., Snowhill, P. B., and Chen, Y., 2000, Role of storage on changes in the mechanical properties of tendon and self-assembled collagen fibers, *Connect. Tissue Res*. **42:**155–164.
15. Silver, F. H., Bradica, G., and Tria, A., 2001, Viscoelastic behavior of osteoarthritic cartilage, *Connect. Tiss. Res*. **42:**223–233.
16. Silver, F. H., Freeman, J. W., Horvath, I., and Landis, W. J., 2001, Molecular basis for elastic energy storage in mineralized tendon, *Biomacromolecules* **2:**750–756.
17. Roach, M. R., and Burton, A. C., 1957, The reason for the shape of the distensibility curves of arteries, *Can. J. Biochem. Physiol*. **35:**681–690.
18. Wilson, A. M., McGuigan, M. P., Su, A., and van den Bogert, A. J., 2001, Horses damp the spring in their step, *Nature* **414:**895–899.
19. Keene, D. R., Lunstrum, G. P., Morris, N. P., Stoddard, D. W., and Burgeson, R. E., 1991, Two type XII-like collagens localize to the surface of banded collagen fibrils, *J. Cell Biol*. **113:**971–978.
20. Bland, Y. S., and Ashhurst, D. E., 2001, The hip joint: The fibrillar collagens associated with development and ageing in the rabbit, *J. Anat*. **198:**17–27.
21. Kaltsas, D. S., 1983, Comparative study of the properties of the shoulder joint capsule with those of other joint capsules, *Clin. Orthop*. **173:**20–26.
22. Niyibizi, C., Visconti, C. S., Kavalkovich, K., and Woo, S. L.-Y., 1994, Collagens in an adult bovine medial collateral ligament: Immunofluorescence localization by confocal microscopy reveals that type XIV collagen predominates at the ligament–bone junction, *Matrix Biol*. **14:**743–751.
23. Riechert, K., Labs, K., Lindenhayn, K., and Sinha, P., 2001, Semiquantitative analysis of types I and III collagen form tendons and ligaments in a rabbit model, *J. Orthop. Sci*. **6:**68–74.
24. Engle, J., 1997, Versatile collagens in invertebrates, *Science* **277:**1785–1786.
25. Fletcher, G. C., 1976, Dynamic light scattering from collagen solutions. I. Translational diffusion coefficient and aggregation effects, *Biopolymers* **15:**2201–2217.
26. Silver, F. H., Langley, K. H., and Trelstad, R. L., 1979, Type I collagen fibrillogenesis: Initiation via a reversible linear growth step, *Biopolymers* **18:**2523–2535.
27. Thomas, J. C., and Fletcher, G. C., 1979, Dynamic light scattering from collagen solutions. II. Photon correlation study of the depolarized light, *Biopolymers* **18:**1333–1352.
28. Silver, F. H., and Birk, D. E., 1984, Molecular structure of collagen in solution: Comparison of types I, II, III, and V, *Int. J. Biol. Macromol*. **6:**125–132.
29. Hofmann, H., Voss, T., Kuhn, K., and Engle, J., 1984, Localization of flexible sites in thread-like molecules from electron micrographs: Comparison of interstitial, basement membrane and intima collagens, *J. Mol. Biol*. **172:**325–343.

30. Paterlini, M. G., Nemethy, G., and Scheraga, H. A., 1995, The energy of formation of internal loops in triple-helical collagen polypeptides, *Biopolymers* **35:**607–619.
31. Silver, F. H., 1987, Self-assembly of connective tissue macromolecules, in *Biological Materials: Structure, Mechanical Properties, and Modeling of Soft Tissues*, chapter 5, NYU Press, New York, pp. 150–153.
32. McBride, D. J., Hahn, R., and Silver, F. H., 1985, Morphological characterization of tendon development during chick embryogenesis: Measurement of birefringence retardation, *Int. J. Biol. Macromol.* **7:**71–76.
33. McBride, D. J., Trelstad, R. L., and Silver, F. H., 1988, Structural and mechanical assessment of developing chick tendon, *Int. J. Biol. Macromol.* **10:**194–200.
34. Liu, M., Tanswell, K., and Post, M., 1999, Mechanical force-induced signal transduction in lung cells, *Am. J. Physiol. Lung Cell. Mol. Physiol.* **277:**L677–L683.
35. Lord, J. M., Bunce, C. M., Brown, G., 1988, The role of protein phosphorylation in control of cell growth and differentiation, *Br. J. Cancer* **58:**549–555.
36. Riveline, D., Zamir, E., Balaban, N. O., Schwarz, U. S., Ishizaki, T., Narumiya, S., Kam, Z., Geiger, B., and Bershadsky, A. D., 2001, Focal contacts as mechanosensors: Externally applied local mechanical force induces growth of focal contacts by an mDia1-dependent and ROCK-independent mechanism, *J. Cell Biol.* **153:**1175–1186.
37. Johnson, G. L., and Lapadat, R., 2002, Mitogen-activated protein kinase pathways mediated by ERK, JNK, and p38 protein kinases, *Science* **298:**1911–1912.
38. Liu, S., Calderwood, D. A., and Ginsberg, M. H., 2000, Integrin cytoplasmic domain-binding proteins, *J. Cell Sci.* **113:**3563–3571.
39. Hynes, R. O., 1992, Integrins: Versatility, modulation, and signaling in cell adhesion, *Cell* **69:**11–25.

Chapter 8

ENHANCEMENT OF PROTEIN THERMAL STABILITY: TOWARD THE DESIGN OF ROBUST PROTEINS FOR BIONANOTECHNOLOGICAL APPLICATIONS

V. Renugopalakrishnan[1], X. Wei[2], G. Narasimhan[2],
C. S. Verma[3], Pingzuo Li[1,4], A. Anumanthan[5]
[1]*Children's Hospital, Harvard Medical School, Harvard University, Boston, MA 02115, USA*
[2]*School of Computer Science, Florida International University, University Park, Miami, FL 33199, USA*
[3]*Bioinformatics Institute, 30 Biopolis Street, 07-01 Matrix, Singapore 138671, Singapore*
[4]*Shanghai Research Center for Biotechnology, Chinese Academy of Sciences, Shanghai 200233, Peoples Republic of China*
[5]*Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA*

Abstract: Protein thermal stability is still an unsolved problem in protein biochemistry. Although a large number of investigations were carried out in the past two decades in order to understand the factors that contribute to the thermal stability of thermophilic proteins, we do not have a complete understanding of the stabilization strategies adopted by thermostable proteins. These investigations have adopted two major approaches. The first approach was to compare the structures and sequences of homologous proteins from hyperthermophiles, thermophiles, and mesophiles, which has resulted in the identification of various differentiated features related to the thermal stability such as an increased number of salt bridges, better hydrogen bonding, higher internal packing, stronger intersubunit associations, and so on. The second major approach involved protein engineering methods. There have also been attempts to combine both of these approaches to study thermal stability. In this review, we discuss a novel approach to achieve enhanced thermal stability using rational site-directed mutagenesis to increase the heat capacity (ΔC_p) by relying on bioinformatics databases and by developing a thermodynamic force field for proteins. Rubredoxin was used as a test platform, and mutants with T_m as high as 200°C were developed.

V. Renugopalakrishnan et al. (eds.), Bionanotechnology, 117-140.

Key words: bionanotechnology; hyperthermophilic rubredoxin; protein heat capacity; protein thermal stability; protein thermodynamics; rational site-directed mutagenesis

1. INTRODUCTION

Extensive studies of mesophilic, thermophilic, and hyperthermophilic proteins have contributed to a deeper and better understanding of the factors that influence protein thermal stability. This chapter defines *hyperthermophiles* as those organisms that grow optimally at temperatures between 80 and 110°C.[1] As of 2001, more than 70 hyperthermophilic species have been cataloged and described in the literature.[2,3] With the accumulated data on thermophiles and hyperthermophiles, it is possible to derive a number of factors that contribute to the stability of these organisms and the proteins contained in them. Some of these factors include: (1) amino acid composition and their intrinsic propensities, (2) disulfide bridges, (3) hydrophobic interactions, (4) aromatic interactions, (5) hydrogen bonds, and (6) ion pairs. From the studies of thermophilic and hyperthermophilic proteins, it is strongly believed that the thermal stability of a protein depends on the enhancement of the heat capacity of the protein, which in turn is dependent on its structure.

Proteins are naturally occurring nanosystems endowed with diverse properties. Advances in molecular biology have made it possible to manipulate proteins and selectively modify their properties through site-specific mutagenesis. Random site-specific mutagenesis has been the method of choice used in the studies reported in the literature. The authors have worked on a rational site-specific mutagenesis approach that rests on the foundation of optimizing specific physical properties by relying upon critical information from bioinformatics databases. The focus is on the thermal properties of proteins.

2. MECHANISMS OF PROTEIN THERMOSTABILIZATION

The hydrophobic effect is considered to be the major driving force of protein folding.[4] Hydrophobicity drives the protein to a collapsed structure from which the native structure is defined by the contribution of all types of forces (e.g., H bonds, ion pairs, and van der Waals interactions). Dill[4] reviewed the evidence supporting this theory as follows: (1) nonpolar solvents denature proteins, (2) hydrophobic residues are typically sequestered into a core, where they are largely devoid of contact with water, (3) hydrophobic

residues in the protein core are more strongly conserved than any other type of residues (replacements of core hydrophobic residues are generally more disruptive than other types of substitutions), and (4) protein unfolding involves a large increase in heat capacity.[4] The sequence, structure, and mutagenesis information accumulated in the last 20 years confirm that hydrophobicity is a major force in protein stability. Given the central role of the hydrophobic effect in protein folding, the assumption that the hydrophobic effect is also the major force responsible for protein stability seems to be justified. The following are two critical observations about the thermal stability imparted by the conserved protein core of mesophilic and their hyperthermophilic homologs: (1) hydrophobic interactions and residues involved in secondary structures in the core are better conserved than the ones on the surface and (2) numerous stabilizing substitutions are found in solvent-exposed areas. The high level of similarity encountered in the core of mesophilic and hyperthermophilic protein homologs suggests that even mesophilic proteins are packed almost as efficiently as possible and there is not much room left for stabilization inside the protein core. Stabilizing interactions in hyperthermophilic proteins are often found in the less conserved areas of the protein. As illustrated in the following section, factors, such as surface ion pairs, decrease in solvent-exposed hydrophobic surface, and anchoring of "loose ends" (i.e., the N and C termini and loops) to the protein surface seem to be instrumental in hyperthermophilic protein thermostability. Sufficient experimental evidence (e.g., sequence, mutagenesis, structure, and thermodynamics) has been accumulated on hyperthermophilic proteins to conclude that no single mechanism is responsible for the remarkable stability of hyperthermophilic proteins. Increased thermostability is caused, instead, by a small number of highly specific mutations that often do not obey any obvious rules.

3. ENGINEERING PROTEIN THERMOSTABILITY

A simple rule of thumb suggests that most reaction rates double with each 10°C increase in temperature. According to this rule, one would expect hyperthermophilic enzymes to have specific activities between 50 and 100 times higher than mesophilic enzymes. What is observed instead is that hyperthermophilic and mesophilic enzymes have approximately the same activities and catalytic efficiencies in their respective physiological conditions. The fact that hyperthermophilic enzymes are not as catalytically optimized as their mesophilic homologs was attributed to the principle that there had to be a trade-off between thermostability and activity, that is, a protein could not be both hyperstable and catalytically optimized. This principle came from the

observation of natural proteins. However, numerous protein engineering studies performed in the last 10 years suggest that protein stability can be enhanced without deleterious effects on activity and stability and activity can be increased simultaneously.[5–7] The fact that mesophilic enzymes are not optimized in terms of stability and hyperthermophilic enzymes are not catalytically optimized is probably only a reflection of the absence of selection pressure for these characteristics. Organisms need to have proteins that they are able to degrade in order to rapidly adapt to changes in environmental stimuli. Hence, mesophilic as well as hyperthermophilic enzymes are only marginally stable under their respective physiological conditions. Unless their substrate is highly unstable, there is no selection pressure in nature for hyperthermophilic enzymes to be highly active. A protein engineer should thus be able to increase the thermal stability of an enzyme without negatively affecting its catalytic properties. The limit to this thermostability increase is defined only by the maximum temperature at which proteins are stable and is believed to be around $\sim$200°C.

4. MECHANISM OF INACTIVATION AND CHOICE OF THERMOSTABILIZATION STRATEGY

Two types of protein stability, thermodynamic and global, are of interest from an applied perspective. Increasing the thermodynamic thermostability is a major challenge especially when an enzyme is used under denaturing conditions. Industrial applications require active enzymes rather than enzymes that are in a reversibly inactivated state. For other enzymes, such as diagnostics enzymes, it is often the long-term stability that needs to be improved.[8] Depending on an enzyme's first inactivation step (i.e., chemical inactivation or unfolding), stabilizing the native, active conformation should involve either substituting temperature-sensitive residues with residues that are chemically more stable or increasing the enzyme's resistance to unfolding, respectively. A number of attempts at stabilizing (or destabilizing) proteins by site-directed mutagenesis have failed because they did not target protein areas that were critical for the unfolding process.[9–11] On the other hand, mutations in regions where the unfolding limits the protein denaturation process can provide extensive stabilization. Good illustrations can be found in the stability studies of thermolysin-like protease from *Bacillus stearothermophilus*. This enzyme is irreversibly inactivated by autolysis that is made possible by partial unfolding of local surface areas. Stabilizing mutations were all located on the surface, around one flexible loop located in the β-pleated N-terminal domain.[12–14] The association of eight mutations in the same area resulted in a 340-fold kinetic

stabilization of *B. stearothermophilus* thermolysin-like protease at 100°C and did not affect the catalytic activity at 37°C. In this enzyme, since the target for autolysis is a loop belonging to the N-terminal domain, any attempt to stabilize the enzyme introducing mutations at the C-terminal domain would have failed. While improving thermodynamic, thermostability can have a beneficial effect on the long-term stability; other strategies can also be used to increase long-term stability. In this case, the focus for stabilization is decelerating the irreversible inactivation process that usually follows reversible unfolding. The strategies that can be used include: (1) eliminating protein diffusion to block aggregation and other biomolecular processes (the most effective approach so far is immobilization), (2) replacing temperature-sensitive residues by residues that are chemically more stable (see Ref. 8 and the references cited therein), and (3) stabilizing the reversibly unfolded state by introducing more hydrophilic residues on the enzyme surface or by adding low-molecular weight compounds to the solute (i.e., inorganic and organic salts, organic co-solvents, or classical denaturants). Hen egg white lysozyme slowly deamidates once reversibly unfolded.[15] Tomizawa[16] replaced Gly residues by Ala in the potentially deamidable Asn–Gly sequences of lysozyme. These mutations generally increased the rate of reversible unfolding, but they decreased the rate of irreversible inactivation and, as a result, stabilized the enzyme against irreversible inactivation. This study is a good illustration of the fact that resistance against irreversible inactivation is not synonymous with thermodynamic thermostability.

5. STRATEGIES FOR ENGINEERING PROTEIN STABILITY

We examine here the principles of random mutagenesis first and then proceed to outline the rational site-specific mutagenesis approach, as developed in our laboratory.

5.1 Random Mutagenesis

The ultimate goal of protein engineering is to design proteins to perform specific functions.[17] However, since the exact rules governing protein folding, molecular recognition, and their precise relationship to functions remain unresolved, rational design of proteins continues to be a challenge. Modification of proteins is carried out either via random site-directed mutagenesis or by combinatorial methods that combine random mutagenesis with phage display depending on the desired properties for the candidate proteins. These

desired properties include: improved bioactivity, photochemical properties, thermostability, and many more. Factors important for the stability of proteins, such as stabilization of α-helices and reducing the number of conformations in the unfolded state,[18,19] are taken into consideration.

There are three widely used strategies for random mutagenesis outlined in Fig. 1: (1) the first one is the oligonucleotide-directed mutagenesis. Oligonucleotides can be synthesized to contain mixtures of nucleotides at specific codons, including all the bases in the first two positions and only two bases in the third position; with each combination allowing 20 possible amino acids. (2) The second one is the error-prone polymerase chain reaction (PCR) that uses a deoxyribonucleic acid (DNA) polymerase to replicate the target gene. (3) The third method is through DNA shuffling, whereby the gene is first spliced into pieces and then regenerated using a DNA polymerase.[20] Since the pieces of DNA get mixed, mutations from separate copies of the gene can be combined. The appropriate method of mutagenesis depends on the choice of physical

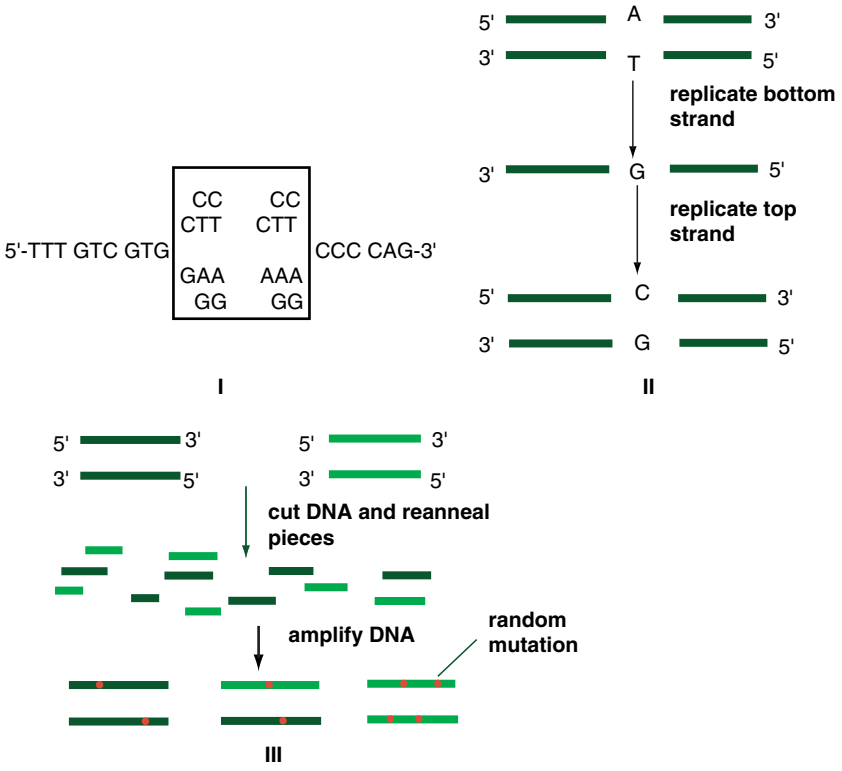


Figure 1. Random mutagenesis.

properties that are chosen for optimization as mutagenesis techniques differ in the number and dispersion of mutations introduced into a gene.

Most of the large numbers of random mutants produced using the above methods do not possess the desired properties. To efficiently screen the mutants, an *in vitro* phage display linking the protein library to DNA was introduced.[21] In other words, each member of the protein library is connected with its gene so that the objective mutants can be identified and readily amplified. Specific enzyme inhibitors, increased enzymatic activity, identifying novel peptide ligands, and agonists of receptor molecules are examples of the successful use of this method. By sorting these libraries produced by random mutation to select for a predicted function, the small number of active proteins can be separated from the millions of inactive variants.

5.2 Rational Site-Directed Mutagenesis

Rational site-directed mutagenesis rests on the principle of optimizing a chosen physical property by analyzing the dependence of that property on specific candidate amino acids that are to be substituted by selecting known or putative sites in the protein under investigation.

We will illustrate this by selecting specific thermodynamic properties of a protein. Proteins in their native state prefer to occur in a state of minimum free energy, ΔG^0, by carefully balancing the intramolecular and the intermolecular energies (such as solvation).

6. FREE ENERGY OF PROTEINS

The minimum free energy, ΔG^0, can be described as follows:

$$\Delta G^0 = \Delta G^0_{\text{intrinsic}} + \Delta G^0_{\text{extrinsic}}, \tag{1}$$

where $\Delta G^0_{\text{intrinsic}}$ is given by the Gibbs–Helmoltz equation:

$$\Delta G^0_{\text{intrinsic}} = \Delta H^0_{\text{intrinsic}} - T \cdot \Delta S_{\text{intrinsic}} \tag{2}$$

and can be partitioned into individual components arising from constituent amino acid residues as:

$$\Delta G^0_{\text{intrinsic}} = \sum_{i=1}^{n} \Delta G^0_{\text{intrinsic}}, \tag{3}$$

where n is number of residues.

One of the fundamental goals in protein engineering is to generate thermally stable protein structures with large free energies of unfolding, ΔF_{u}, and

high midpoint melting temperatures, T_m. The temperature dependence, ΔF_u can be analyzed by examination of the temperature dependence of ΔH and ΔS. Assuming that the heat capacity, ΔC_p, of a protein is independent of temperature (at least in the range 20–200°C), we have,

$$\Delta H = \Delta H(T_r) + \Delta C_p(T - T_r), \tag{4}$$

$$\Delta S = \Delta S(T_r^{'}) + \Delta C_p \ln\left(\frac{T}{T_r^{'}}\right), \tag{5}$$

where T_r and $T_r^{'}$ are appropriate reference temperatures. Note that:

$$\Delta H(T_r) = \Delta H \text{ of the protein at the temperature} T_r, \tag{6}$$

$$\Delta S(T_r^{'}) = \Delta S \text{of the protein at the temperature} T_r^{'}. \tag{7}$$

The convergence temperatures[22,23] were chosen here because the reference ΔH for different proteins upon derivation converged to the same value ΔH^* at some temperature T_h^*. A similar convergence was also observed for the entropy change per residue, ΔS^*. In physical terms, the convergence temperatures are now believed to be the temperatures at which the polar contribution to ΔH^0 and ΔS^0 reaches zero, respectively.

Therefore,

$$\Delta H = N_{ref} H^* + \Delta C_p(T - T_h^*), \tag{8}$$

$$\Delta S = N_{ref} \Delta S^* + \Delta C_p \ln\left(\frac{T}{T_s^*}\right), \tag{9}$$

where N_{ref} is the number of amino acids in the protein. Therefore, the free energy of unfolding, ΔF_u, can be expressed as[23]

$$\Delta F_u = N_{ref}(\Delta H^* - T\Delta S^*) + \Delta C_p\left[(T - T_h^*) - T \cdot \ln\left(\frac{T}{T_s^*}\right)\right]. \tag{10}$$

The above equation is quite useful when comparing the ΔF between a wild-type protein and its mutants with different values of ΔC_p.

$$\Delta\Delta F_u = \Delta\Delta C_p\left[(T - T_h^*) - T \cdot \ln\left(\frac{T}{T_s^*}\right)\right]. \tag{11}$$

From the above equations, it is evident that for a protein with N_{ref} residues, changes in ΔC_p can result in changes in ΔF_u that can be quantified at all the temperatures. For T_h^* and T_s^* of 100.5 ± 6°C and 112 ± 1°C, ΔH^* and ΔS^* are 1.35 ± 0.11 kcal/(mol.res) and 4.30 ± 0.1 cal/K/(mol.res), respectively.

Therefore,

$$\Delta F_u = N_{res}(1.35 \pm 0.0043T) + \Delta C_p\left[(T - 373.6) - T \cdot \ln\left(\frac{T}{385.1}\right)\right]. \quad (12)$$

ΔT_m can be directly estimated from the changes in ΔC_p as follows:

$$\Delta T_m = (3.6 \pm 2.5)\Delta\Delta C_p. \quad (13)$$

7. CALCULATION OF ΔC_P OF PROTEINS

For a protein, the quantity ΔC_p can be expressed as a function of the change of buried apolar and polar surface areas upon unfolding[24]:

$$\Delta C_p = 0.45\Delta ASA_{apolar} - 0.26\Delta ASA_{polar}, \quad (14)$$

where ΔC_p is in cal/mol/K. ΔASA_{apolar} and ΔASA_{polar} are the changes in solvent accessible surface area (unit $Å^2$) upon protein denaturation for apolar and polar atoms, respectively. The constants are taken from model-compound studies[25] and have been shown to be reasonably accurate in calculating ΔF_u.[26] Equation (14) shows that increasing buried apolar surface area in the folded state will increase ΔC_p, while increasing buried polar surface area in the folded state will decrease ΔC_p. The ΔC_p values of apolar and polar residues are listed in Table 1.

8. EXPERIMENTAL DETERMINATION OF ΔC_P

The quantity ΔC_p can be experimentally obtained from differential scanning calorimetric studies (DSC). The DSC measures the heat capacity of a protein solution as a continuous function of temperature. In a typical DSC

Table 1. ΔC_p values for Amino Acid Residues

Amino acid	ΔC_p (cal/K/(mol.res))[†]
Ala	30.15
Ile	63.00
Leu	61.65
Met	41.47
Phe	78.75
Trp	80.73
Tyr	53.62
Tyr[‡]	70.00
Val	52.65

[†]The ΔC_p values for totally buried hydrophobic amino acids were calculated from reported solvent accessible surface area.
[‡]This value is used in special cases when the hydroxyl group of Tyr is solvent exposed.

measurement, the heat capacity of the sample cell containing the protein solution is measured in relation to that of the reference cell containing only the solvent or buffer. The measured heat capacity C_p (cal /K) can be written as:

$$C_p = m_p C_{p,p} + m_b C_{p,b} + C_{p,\,ref}, \tag{15}$$

where $C_{p,p}$ and $C_{p,b}$ are the heat capacity of the protein and solvent/buffer, while m_b and m_s are the masses of protein and solvent/buffer, respectively. $C_{p,\,ref}$ includes the heat capacity of the buffer solution in the reference cell. Since the volume of the sample cell is constant, the volume of solvent (or buffer) and consequently its mass depends on the mass of protein present in the cell, and the above equation can be written as,

$$C_p = m_p C_{p,p} + m_p \cdot C_{p,b}\rho_b(V^0 - V_p \cdot m_p) + C_{p,\,ref}, \tag{16}$$

$$= C_{p,b} \cdot \rho_b \cdot V^o + (C_{p,\,p} - V_p \cdot \rho_b \cdot C_{p,b}) \cdot m_p + C_{p,\,ref} \tag{17}$$

where ρ_b is the density of the solution, V^o is the volume of the cell, and V_p is the partial specific volume of the protein. Generally, $V_{pz}\rho_b$ and $C_{p,\,b}$ are obtained from standard tables. From the above equation, the slope of a plot of C_p versus the mass of protein in the cell ($\delta C_p/\delta m_p$) provides all the information necessary to calculate the heat capacity, C_p.

ΔC_p can be used to calculate T_m:

$$T_m = T_1^* \exp\left[\frac{-\Delta S(T_1)}{\Delta C_p}\right]. \tag{18}$$

9. DEFINITION OF T_{MS}

The temperature of maximal stability of a protein, denoted by T_{ms}, is defined as the midpoint temperature of transition, T_m, during heat denaturation of a protein. Proteins are characterized by two denaturation temperatures (T_d) at which $\Delta G_0 = 0$ and T_w as determined by DSC studies. A plot of $\Delta G_0 = f(T)$ is usually referred to as the stability curve.

The values of T_{ms} and ΔC_p for a number of proteins are listed in Table 2 (adapted with permission from Ganesh *et al.*[26] and shown in Fig. 2).

10. HYPERTHERMOPHILIC PROTEINS

Hyperthermophilic organisms and the proteins found in them serve as an excellent test bed for the design of thermally stable proteins[27]. Nature has optimized these proteins by the process of evolution, thus providing us with a

Table 2. Experimental and calculated ΔC_p and T_m for selected proteins

Protein	PDB code	N_{s-s}	N_{res}	ΔA_{ap}[a]	ΔC_p[b]	ΔC_p^b	T_{mso}	T_{msc}
Ovomucoid III	1cho	3	56	2712	10.5	10.5	268	257
RNase A	9rsa	4	124	7170	10.6	12.6	256[d]	275
CSP[c]	1csp	0	67	3736	10.7	11.5	280	266
Protein G[c]	1pgb	0	56	2834	11.1	10.5	269	257
CI2[c]	2ci2	0	65	3409	11.1	11.4	262	265
Parvalbumin	5cpv	0	108	6393	12.0	13.5	277	280
Hen Lysozyme	6lyz	4	129	7701	12.0	12.8	272[d]	276
HH myoglobin[c]	1ymb	0	153	9757	12.2	15.3	291	291
Hu lysozyme	1lzl	4	130	8119	12.2	12.8	276	276
Interleukin 1β	5ilb	0	153	10175	12.4	14.4	285	286
Chymotrypsin	4cha	5	241	15383	12.5	14.4	281[d]	286
Iso-1-cytochrome c[e]	1ycc	1	108	6220	12.7	14.4	271	286
Barnase	1rnb	0	110	6166	12.8	13.6	255	281
Sac 7[d]	1sap	0	66	3446	13.0	11.5	296	265
RNase T1	9rnt	2	104	5847	13.0	12.6	259	274
Trypsin	1tkl	6	223	14161	13.8	14.1	281	284
α-Lactalbumin	1alc	4	123	7404	14.6	12.6	291	274
CAB[c]	2cab	0	260	16850	14.6	15.3	290	291
S. nuclease	2sns	0	149	8360	14.8	14.3	289	286
Thioredoxin	2trx	1	108	6701	15.4	13.1	298[d]	278
Papain	9pap	3	212	13776	15.6	14.5	290	286
T4 Lysozyme	1l63	0	164	10024	15.7	14.5	281	287
Cytochrome c[e]	2pcb	1	104	5978	16.1	14.3	293	285
Barstar	1bta	0	89	5770	16.4	12.8	299	276
Hpr[c]	2hpr	0	87	5121	16.7	12.7	290	275
MBP[c]	1omp	0	370	25283	17.6	15.7	306	293
myoglobin[c] SW	5mbn	0	153	9964	18.1	15.3	305	291
PGK[c]	3pgk	0	415	26466	18.1	15.8	301	294

The ΔC_p values in column 7 are calculated from the amino acid sequence using Eq. 4. The observed temperatures of maximal stability (T_{mso}) are obtained from calorimetric data. The calculated temperatures of maximum stability (T_{mso}) are obtained from the calculated values of ΔC_p in column 7 using Eq. 4 with T_1=358 K and $\Delta S(T_1) = 3.15$ cal (mol residue)$^{-1}$ K^{-1}

[a] Areas calculated as described in the text using the algorithm of Connolly [8].

[b] Observed and calculated values of ΔC_p in units of cal (mol residue)$^{-1}$ K^{-1}. Values for ΔC_p (obs.) are obtained from the following sources: for PDB codes 1cho, 1pgb, 1lzl, 5ilb, 1ycc, 1rnb, 9rnt, 1163, 2pcb, 5mbn, 1ymb, 6lyz, 4cha, 2trx, 2ci2 from [9]; for PDB codes 2cab, 5cpv, 1tld, 9pap from [2]; for PDB codes 1alc and 3pgk from [4]; for PDB codes 2sns, 1sap, 1bta, 2hpr, 1csp, 1omp and 9rsa from [26,27,18,12,13,28,29].

[c] Protein G:.IgG binding domain of protein G; C12: chymotrypsin inhibitor 2; HH: horse heart; CSP: cold shock protein; CAB: carbonic anhydrase; SW: sperm whale; PGK: phosphoglycerate kinase; Hu: human; HPr: histidine-containing phosphocarrier protein; CAB: carbonic anhydrase; SW: sperm whale; PGK: phosphoglycerate kinase; Hu: human; HPr: histidine-containing phosphocarrier protein; MBP: maltose binding protein.

[d] See Table 2.

[e] The indicated S-S linkage for the cytochromes is the thioether linkage between the vinyl groups in the home prosthetic group to the sulfur atoms in the cysteine side chains in the protein.

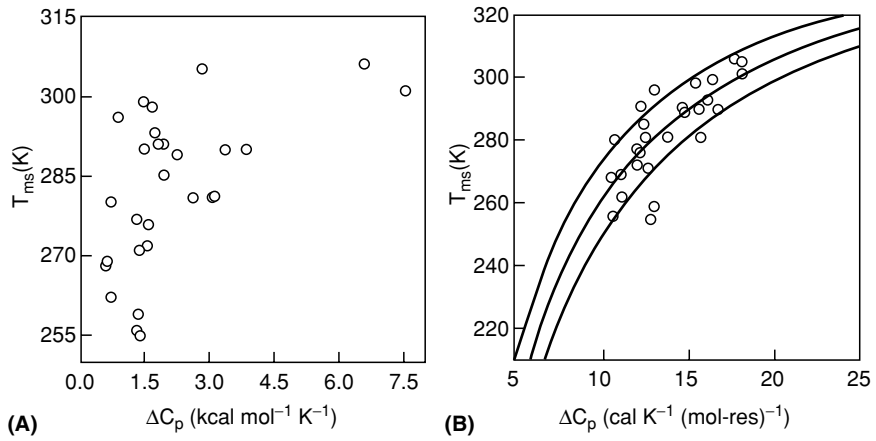


Figure 2. T_m of proteins (whose X-ray structures are known) as a function of ΔC_p. (Adapted with permission from Ganesh et al.[26])

wealth of principles for the design of thermally stable proteins. Table 3 shows a compendium of hyperthermophilic organisms and selected proteins contained in these organisms.

There are currently three proposed models to explain the higher denaturation temperatures of thermophilic proteins. The free energy profile of a mesophilic protein and the proposed models for converting it into a thermophilic protein are shown in Fig. 3. This model was based on the observation that the thermophilic protein shows only a modest increase in ΔG_u at room temperature in comparison to their mesophilic counterparts. However, increase in ΔG_u at room temperature as high as 20 kcal/mol for thermophilic proteins have been reported,[28] which lend support to the "raised" model. There is not much experimental support for the "shifted" model because decreases in ΔG_u at room temperature have not been observed.

According to the theory presented in this chapter, any one of the three models alone cannot explain the protein thermal stability. First of all, trying to judge the dependence of free energy on temperature by looking at ΔG_u values reported (in support of the "raised" model in Fig. 2) by assuming the same ΔC_p value for both the thermophilic and mesophilic proteins.[29] The above assumption is incorrect as ΔC_p difference is an important factor in determining protein stability.

A quantitative way to look at thermal stability is by taking the derivative of Eq. (10) with respect to temperature:

Table 3. Sample list of hyperthermophilic organisms with select proteins present in them

Organism	Temperature (°C)	Aa	Dr	Fe	Hi	Pk	Ru	RN
Archaea								
Acidianus ambivalens (*Desulfurolobus ambivalens*)	80			•				
Acidianus infernos	90			•				
Aeropyrum prernix	90–95					•		•
Archaeoglobus fulgidus	83	•		•		•	•	•
Methanococcus jannaschii	83	•		•	•	•	•	•
Methanopyrus kandleri$_{AV19}$	98	•		•	•	•		•
Methanothermus fervidus	83			•	•	•		
Pyrobaculum aerophilum	100	•		•		•		
Pyrobaculum islandicum	100			•				
Pyrococcus abyssi	97	•		•	•	•	•	•
Pyrococcus furiosus	98	•		•	•	•	•	•
Pyrococcus horikoshii$_{OT3}$	9	•			•	•		•
Pyrococcus woesei	100–103			•		•		
Pyrodictium abyssi	97					•		
Sulfolobus tokodaii	80			•		•		
Thermococcus litoralis	85			•				
Thermococcus profundus	80			•				
Bacteria								
Aquifex aeolicus	95			•		•		•
Thermotoga maritima	80	•	•	•		•	•	•

A dot (•) indicates that the particular protein is present in that organism. The absence of a dot indicates that the presence of that protein in that organism is not known. The proteins selected and their corresponding abbreviations are as follows: Aspartate Aminotransferase (Aa); Dihydrofolate reductase (Dr); Ferredoxin (Fe); Histone (Hi); Phosphoglycerate kinase (Pk); Rubredoxin (Ru); and R Nase H (RN).

$$\frac{\mathrm{d}(\Delta G)}{\mathrm{d}T} = -N_{\mathrm{res}} \times \Delta S^{*} + \Delta C_{\mathrm{p}} - \Delta C_{\mathrm{p}} \times \ln\left(\frac{T}{T_{\mathrm{s}}^{*}}\right). \tag{19}$$

Since this is the slope of the ΔG_{u} versus T plot at any temperature T, at $T_{\max}$ (where ΔG_{u} reaches maximum), $\mathrm{d}(\Delta G)/\mathrm{d}T = 0$.

Therefore,

$$T_{\max} = T_{\mathrm{s}}^{*}.\exp\left(\frac{-\Delta S^{*} \cdot N_{\mathrm{res}}}{\Delta C_{\mathrm{p}}}\right) = T_{\mathrm{s}}^{*} \exp\left(\frac{-\Delta S^{*}}{\mathrm{residue}\Delta C_{\mathrm{p}}}\right). \tag{20}$$

Because the errors in the values of T_{s}^{*} and ΔS^{*} are relatively small, $T_{\max}$ therefore depends mainly on the value of ΔC_{p} and the nature and number of residues. Because of this conclusion, a thermophilic or hyperthermophilic protein and its mesophilic counterparts differ essentially on ΔC_{p} and consequently manifest different values of T_{m}.

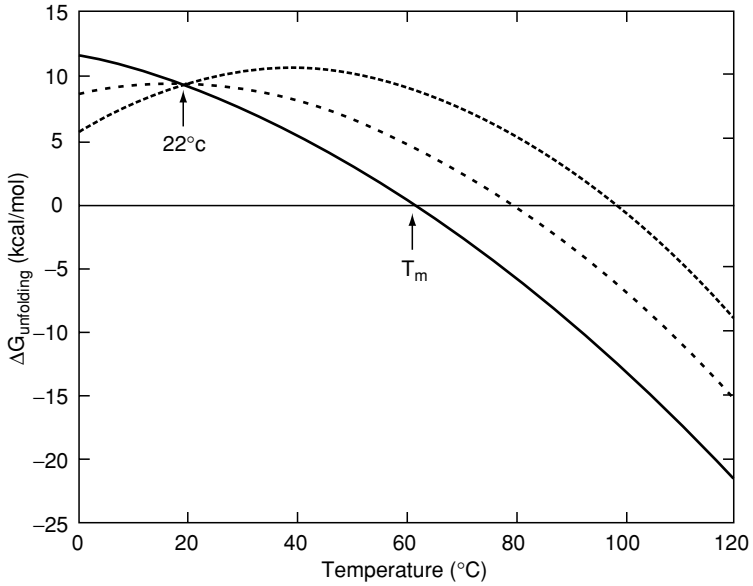


Figure 3. Dependence of ΔF_u on temperature for a hypothetical 100 residue protein at three different values of ΔC_p (10 and 20 cal/K/mol.res) are usually lower and upper limit of ΔC_p for most proteins. (Adapted with permission from Ref. 38).

11. CONFORMATIONAL ENTROPY, ΔS_{CONF}

The conformational entropy change, ΔS_{conf}, arises from the reorganization of hydrophobic amino acid side chains inside the protein and can be calculated by taking the difference of the entropy of the folded and unfolded states from the Boltzman equation:

$$\Delta S_{conf} = S_f - S_u = -R[\ln W^* - \ln W], \tag{21}$$

where S_f is entropy of the folded state, S_u is entropy of the unfolded state, W is the number of rotamers in unfolded state, and W^* is the effective number of rotamers in the folded state as given by:

$$W^* = \exp\left(-\sum_i^n p_i . \ln p_i\right), \tag{22}$$

where p_i is the fractional population of each rotamer state, i, in the folded state and R is the gas constant.

The ΔS_{conf} change for the hydrophobic core of the entire protein is defined as the mean of the individual conformational entropies for all core residues. The quantity ΔS_{conf} is an important component of protein stability.

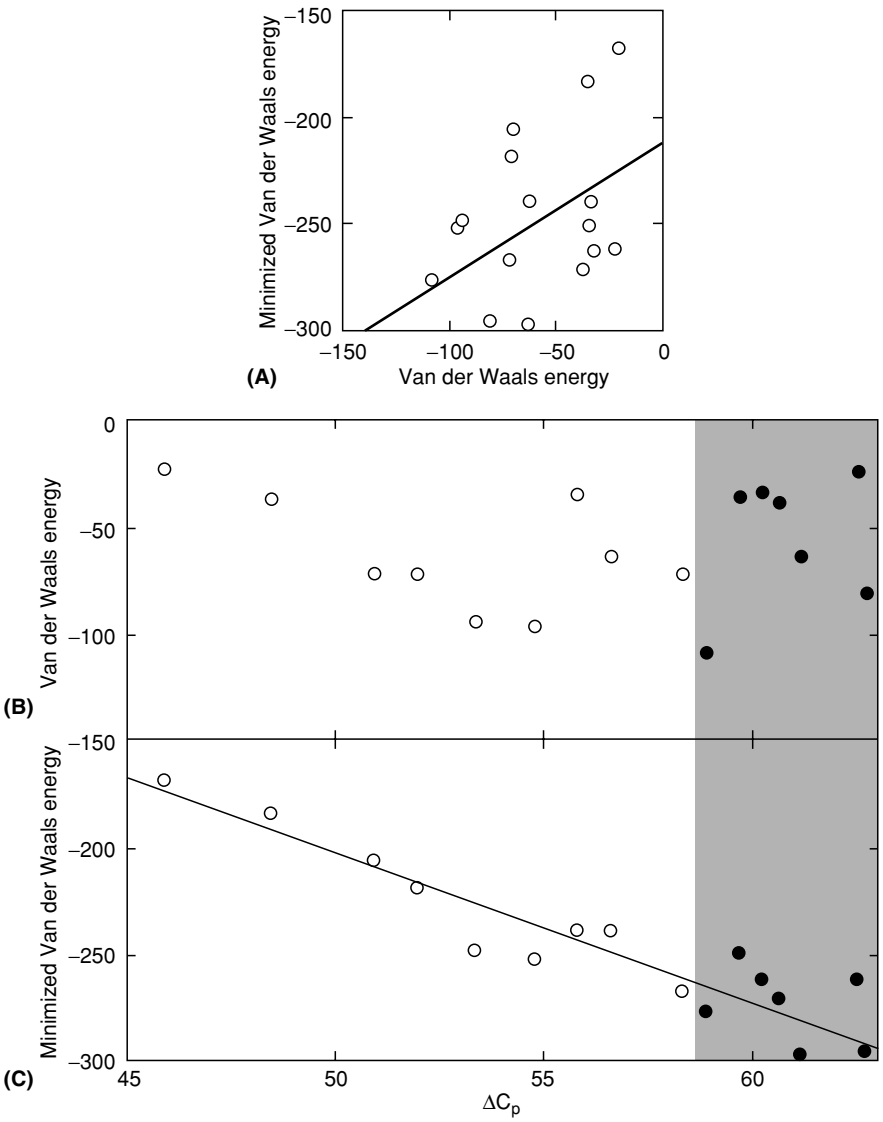


Figure 4. (A) Plot of van der Waals energy after a 10-iteration energy minimization vs van der Waals energy before minimization for random myoglobin structures sampled during a CORE run. (B) van der Waals energy plotted as a function of ΔC_p for these same myoglobin structures showing no obvious correlation. The shaded region represents the range of ΔC_p values for predicted myoglobin sequences: black circles represent predicted myoglobin sequences and open circles represent sequences with zero bumps sampled, but not predicted. (C) van der Waals energy after the 10-iteration energy minimization vs ΔC_p showing good correlation between these parameters. (Adapted with permission from Ref. 38).

Figure 4 shows the dependence of ΔC_p on E_{vdw} for several myoglobin mutants. It shows a weak correlation between E_{vdw} calculated using discrete rotamers library and ΔC_p. The correlation improves after 10 iterations of minimization. Protein sequences with high ΔC_p values and side-chains compatible with the backbone of the protein, calculated by using a rotamer library are stable proteins, are selected.

12. THERMODYNAMIC FORCE FIELD FOR PROTEIN

The design of a thermally stable protein must include the following criteria:

(a) The side chains of the amino acid substituted to enhance T_m must be sterically compatible within the backbone structure.
(b) The mutated sequence must have a ΔC_p value that can be achieved by burying more apolar surface area or exposing more polar surface area in the folded protein.
(c) The hydrophobic core of the predicted protein must have low conformational entropy of folding to ensure a unique and rigid internal architecture.

13. COMPENSATION AND CONVERGENCE TEMPERATURES

The convergence of thermodynamic quantities at some temperature will occur when there are two dominant interactions, such as apolar and polar, which independently contribute to the thermodynamic properties. If only apolar contributions are modified by mutation of core residues without significant effect on polar contribution or backbone structure, convergence should not be expected to occur.

It is expected, therefore, that within a series of proteins in which the backbone structure is maintained and only hydrophobic core residues are altered, the values of ΔH^*, ΔS^*, T_H^*, and T_s^* will remain constant. Under these conditions, increases in ΔC_p will result in increases in T_m and ΔG_u above a fixed temperature.

Let us assume that for a specific protein, the apolar contribution to ΔH^o and ΔS^o are equal to 0 at temperatures T^+ and T_s^+, respectively. Let us also assume ΔH^+ and ΔS^+, respectively. Equation (10), described previously, can be used to calculate ΔG_u for any protein, where ΔC_p is assumed to be

temperature independent. Let us also assume that this specific protein has a compensation temperature of T_c. Based on the enthalpy–entropy compensation relationship, at T_c, we have:

$$\Delta\Delta H = T \cdot \Delta\Delta S, \tag{23}$$

and

$$(T_c - T_h^+) - T_c \cdot \ln(T_c - T_s^+) = 0. \tag{24}$$

T_c is shared among a native protein and its mutants with approximately the same solution structure, as shown empirically. Based on the above analysis, we deduce T_c as the actual crossing point on the ΔG_u versus T plot for a group of proteins, with T_h^+ and T_s^+ as the temperature at which the apolar contribution to ΔH^o and ΔS^o reaches zero, respectively. The origin of T_c is believed to be the hydrophobic effect or the hydration of the apolar groups. Therefore, the convergence temperatures (T_h^+ and T_s^+) depend on the composition, structure, and the environment of each individual protein.

Enthalpy–entropy compensation is a common property of weak intermolecular interactions. Based on the relationship between compensation and convergence temperatures, it is therefore rational to conclude that T_c is the temperature at which the apolar and polar hydration effects reach certain equilibrium. The value of T_c depends on the primary structure of a protein and hence on its three-dimensional (3D) structure.

Enthalpy–entropy compensation is a general feature of many chemical reactions and processes in biological systems. The slope of ΔH_0 versus ΔS_0 plot is called the compensation temperature, T_c, and the values for T_c falls in the range of 260–315°K. ΔG_0 stays around a constant value. To engineer a protein at a given temperature, T_c, ΔS will also decrease by an equivalent amount. The important points we learn in both components, ΔH_{conf} and ΔS_{conf}, must be optimized simultaneously.

14. COMPUTER ALGORITHM FOR DETERMINING THE CORRECT SUBSTITUTIONS OF THE SELECTED RESIDUES

We have developed a program named ''rational site-directed mutagenesis selection algorithm'' (RSDMSA) (US Patent Disclosures[30] and Renugopalakrishnan[31]), which was improvised from an earlier program (CORE) developed by Jiang *et al.*[32] The RSDMSA predicts protein hydrophobic core sequences that can fold into a target backbone structure. Basically, RSDMSA rejects or eliminates unfavorable mutations starting from an input backbone.

The sequences are selected via simulated annealing and Monte Carlo runs to yield proteins with high thermal stability. A typical prediction run in RSDMSA starts by randomizing the sequence of hydrophobic core of a selected protein to ensure that no bias is introduced at the start of a run. After this, a simulated annealing run driven by the Metropolis algorithm is initiated. Following the last step, a single mutation of a core residue chosen based on the specific property and nature of the protein is allowed. Residues are allowed to mutate to Ala, Ile, Met, Phe,Tyr, Trp, or Val. RSDMSA then initiates a nested simulated annealing run to determine the best rotamer configuration of all core residues. The simulated annealing run reveals the number of unfavorable van der Waals interactions by calculating the number of hard sphere bumps. If the number of bumps for the simulated annealing run is greater than zero, the sequence is rejected. If the number of bumps is zero, a second simulated annealing on the same sequence is initiated followed by a low-temperature Monte Carlo run that yields two parameters as follows: (i) Mobility of each amino acid in the form of conformational entropy, which is averaged to obtain global conformational entropy for the whole protein (ΔS_{conf}) as shown below

$$\Delta S = \sum_{1}^{n} \Delta S_{\text{conf}}^{\text{residue}}. \tag{25}$$

(ii) Heat capacity (ΔC_{P}) of the hydrophobic core of the protein calculated based on Table 1.

These two parameters, plus the number of bumps, are used to calculate the ''scores'' for this particular sequence. This score then drives the main sequence simulated annealing run. After 10 sequences sampled at each temperature, the Metropolis temperature is gradually decreased during the simulated annealing process. This gradual decrease in the Metropolis temperature ensures that local minima are avoided by slowly lowering the probability that the sequences are accepted with scores higher than the previously accepted sequence. The simulated annealing run is automatically terminated when the number of accepted sequences is consistently zero. The temperature at which this occurs, or an arbitrarily low temperature, T, is used for a final sequence Monte Carlo run initiated with the sequence determined from the simulated annealing run. This Monte Carlo run is conducted to sample the sequence space around the simulated annealing sequence. A large value of T allows the program to sample from sequences with wider range of scores around the score for the simulated annealing sequence. Typically, these runs generate a family of 100–1000 sequences of proteins with good thermal stability depending on the value of T and the number of core residues mutated.

15. DESIGN OF HYPERTHERMOPHILIC RUBREDOXIN

Rubredoxin from the hyperthermophilic archaebacterium, *Pyrococcus furiosus*, an organism that grows optimally at 100°C and above, provides a test platform for thermal stabilization of proteins.[33] The thermal denaturation of rubredoxin has been previously shown to exhibit unusual kinetics characterized by nearly constant rate of change and a reaction time independent of the starting amount of the protein. Wampler and Bradley[34] have described a kinetic model that explains these data and a variety of other thermal denaturation data for various rubredoxins, as well as other proteins.

Rubredoxins are probably the simplest members of the ubiquitous and huge family of redox metalloenzymes[33] and consist of a relatively short polypeptide chain (~53 aa) endowed with a prosthetic group in the form of a ferrous/ferric ion tetrahedrically bound to four cysteine (Cys) residues. Even though their exact metabolic role in anaerobic cells has not yet been fully clarified, their structural features are fairly well known from the primary structures of 19 bacterial rubredoxins reported to date. These rubredoxins are quite similar with over 20% of the residues strictly conserved. Among the conserved residues, there are four Cys residues from the active site and five aromatic residues that constitute the hydrophobic core of the proteins. However, these rubredoxins are from a variety of mesophiles and thermophiles, displaying a wide range of thermal stability. In particular, no rationale has been found on the basis of the sequence and/or structure information, for the fact that the rubredoxin from *Pyrococcus furiosus* (RUBR_PYRFU) that lives normally at 90°C has a half-time of thermal denaturation of 400 h at 92°C as compared to 6 h of *Clostridium pasteurianum* (RUBR_CLOPA) rubredoxin that is most similar to it in terms of 3D structure.

The RSDMSA program was applied to rubredoxin to select mutants with high thermal stability. We have carried out extensive optimization of rubredoxin thermodynamics by increasing its total ΔC_p, by selectively substituting key residues shown in Fig. 5 with more hydrophobic residues with large ΔC_p values and consequently, increasing T_m to 200°C. The underlying rationale for the selection of residues A9, A13, A31, A42, A43, and A51 for site-directed mutagenesis rests on the hypothesis that the thermostability is conferred by interaction between β-sheet segments 1 and 2 (Renugopalakrishnan *et al.*, unpublished). Three designed mutants of rubredoxin (A9F, A13W, A43I; A9G, A13D, A31E, A42G, A51L; A9Y, A31L, A42F, A51I) with T_m in the vicinity of 200°C are shown in Fig. 6.

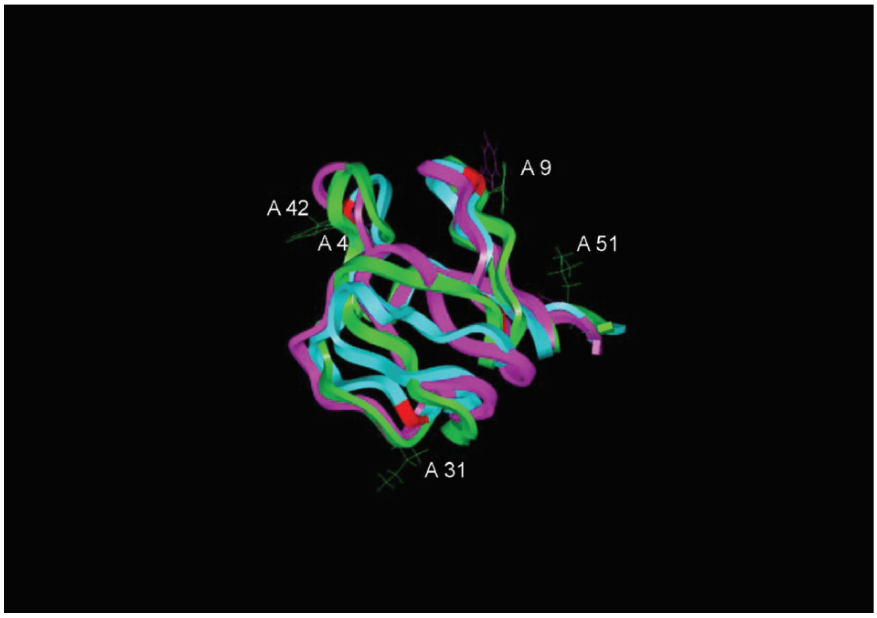


Figure 5. Three-dimensional structure of wild-type rubredoxin from *Pyrococcus furiosus*.

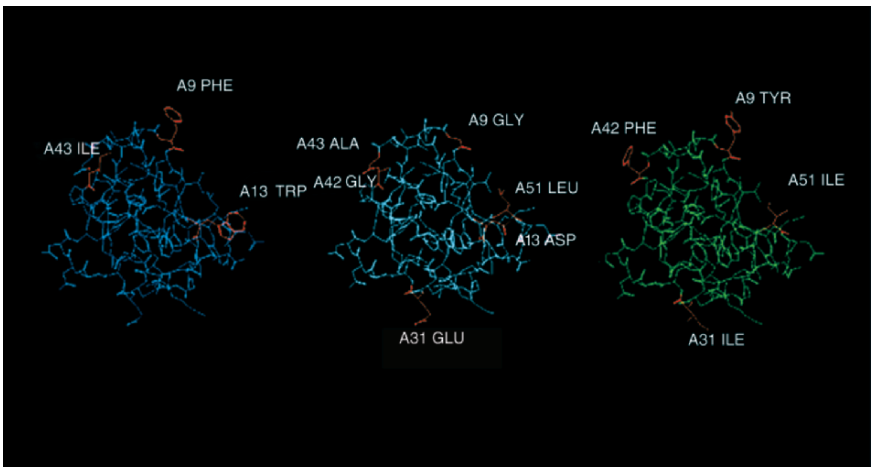


Figure 6. Three-dimensional structure of three mutants of rubredoxin. T_m in the range of 200°C.

16. CONCLUSION

Design of robust proteins for bionanotechnological applications is receiving wide attention.[35,36] Protein thermodynamic stability, ΔG, is determined by a multitude of intramolecular and intermolecular interactions. The side chain interactions are important contributors to the thermodynamic stability, and the exact placement of side chains in the sequence is therefore of great importance. Once the sequence is determined, all other secondary and tertiary structural features become spontaneously dependent on the sequence, and hence it is the sequence that determines ΔG.[37] Therefore in our opinion, based on accumulated data, protein thermal stability is a direct consequence of the sequence itself. Berezovsky and Shakhnovich[37] concluded that basically all strategies for thermal stabilization of protein, from which rational guidelines for design of thermally stable proteins are derived, consist of two strategies: (1) general, nonspecific, structure-based, with contribution from all types of stabilizing interactions and (2) specific, sequence-based, with use of only dominating factor for adaptation to extreme conditions. The exponentially increasing number of fully sequenced genomes will be an invaluable help in deciphering as to which sequence variations among homologous proteins are related to stability and which ones are simply a result of evolution.

ACKNOWLEDGEMENTS

Early and significant conceptual developments of the research was initiated at Children's Hospital, Harvard Medical School and was funded by a grant from the Office of Naval Research (ONR), and partially from an NIH grant to the Laboratory for Skeletal Disorders and Rehabilitation, Department of Orthopedic Surgery, Children's Hospital and Harvard Medical School. Authors express thanks for copyright permission to the Editor, FEBS Letters, Elsevier, Ganesh, C. *et al*., FEBS Lett, 1999 454:31–36 [Ref. 26], to reproduce Fig. 2 and Table 2; and Editor, Protein Science for Fig. 4, Jiang et al., Protein Science, 2000, Vol. 9:403–416 [Ref. 38]. Research of GN was supported in part by NIH Grant P01 DA15027-01. M Prabhakaran's help in obtaining Figs. 5 and 6 are acknowledged. BII is an A-STAR institute CVS is adjunct faculty at the Dept. of Biological Sciences, National University of Singapore. Authors thank Ms. S.L.Lakka for her help with Table 2. Authors express their thanks to American Scientific Publishers for permission to reproduce Fig. 6, Ref. 36.

REFERENCES

1. Vieille, C., Burdette, D. S., and Zeikus, J. G., 1996, Thermozymes, *Biotechnol. Annu. Rev*. **2:**1–83.
2. Hicks, P. M., and Kelly, R. M., 1999, Extremely thermophilic microorganisms, in *Encyclopedia of Bioprocess Technology* (M. C. Flickinger and S. W. Drew, eds.), Wiley, New York, pp. 2536–2552.
3. Stetter, K. O., 1996, Hyperthermophilic prokaryotes, *FEMS Microbiol. Rev*. **18:**149–158.
4. Dill, K. A., 1990, Dominant forces in protein folding, *Biochemistry* **29:**7133–7155.
5. Giver, L., Gershenson, A., Zhao, H., and Miyazaki, K., 1998, Directed evolution of a thermostable esterase, *Proc. Natl. Acad. Sci. USA* **95:**12809–12813.
6. Van den Burg, B., Vriend, G., Veltman, O. R., Venema, G., and Eijsiak, V. G. H., 1998, Engineering an enzyme to resist boiling, *Proc. Natl. Acad. Sci. USA* **95:**2056–2060.
7. Zhao, H., and Arnold, F. H., 1999, Directed evolution converts subtilisin E into a functional equivalent of thermitase, *Protein Eng*. **12:**47–53.
8. Mozhaev, V. V., 1993, Mechanism-based strategies for protein thermostabilization, *Trends Biotechnol*. **11:**88–95.
9. Kawamura, S., Kakuta, Y., Tanaka, I., Hikichi, K., Kuhara, S., Yamasaki, N., and Kimura, M., 1996, Glycine-15 in the bend between two alpha-helices can explain the thermostability of DNA binding protein HU from *Bacillus stearothermophilus*, *Biochemistry* **35:**1195–1200.
10. Kawamura, S., Tanaka, I., Yamasaki, N., Kimura, M., and Eijsiak, V. G. H., 1997, Contribution of a salt bridge to the thermostability of DNA binding protein HU from *Bacillus stearothermophilus* determined by site-directed mutagenesis, *J. Biochem. (Tokyo)*, **121:**448–455.
11. Tomschy, A., Bohm, G., and Jaenicke, R., 1994, The effect of ion pairs on the thermal stability of D-glyceraldehyde 3-phosphate dehydrogenase from the hyperthermophilic bacterium Thermotoga maritime, *Protein Eng*. **7:**1471–1478.
12. Hardy, F., Vriend, G., Veltman, O. R., van der Vinne, B., Venema, G., and Eijsiak, V. G. H., 1993, Stabilization of Bacillus stearothermophilus neutral protease by introduction of pralines, *FEBS Lett*. **317:**89–92.
13. Mansfeld, J., Vriend, G., Dijkstra, W., Veltman, O. R., van der Burg, B., Venema, G., Ulbrich-Hofman, R., and Eijsiak, V. G. H., 1997, Extreme stabilization of a thermolysin-like protease by an engineered disulfide bond, *J. Biol. Chem*. **272:**11152–11156.
14. Veltman, O. R. et al., 1996, Analysis of structural determinants of the stability of thermolysin-like proteases by molecular modelling and site-directed mutagenesis, *Protein Eng*. **9:**1181–1189.
15. Ahern, T. J., and Klibanov, A. M., 1985, The mechanisms of irreversible enzyme inactivation at 100 C, *Science* **228:**1280–1284.
16. Tomizawa, H., Yamada, H., Hashimoto, Y., and Imato, T., 1995, Stabilization of lysozyme against irreversible inactivation by alterations of the Asp-Gly sequences, *Protein Eng*. **8:**1023–1028.
17. Dahiyat, B. I., and Mayo, S. L., 1997, De novo protein design: Fully automated sequence selection, *Science* **278:**82–87.
18. Cohen, C., and Parry, D. A., 1990, Alpha-helical coiled coils and bundles: How to design an alpha-helical protein, *Proteins* **7:**1–15.
19. Naik, R. R., Kirkpatrick, S. M., and Stone, M. O., 2001, The thermostability of an alpha-helical coiled-coil protein and its potential use in sensor applications, *Biosens. Bioelectron*. **16:**1051–1057.

20. Crameri, A., Raillard, Sun-Al, Bermudez, E., and Stimmer W. P. C., 1998, DNA shuffling of a family of genes from diverse species accelerates directed evolution, *Nature* **391:**288–291.
21. Dennis, M. S., Herzka, A., and Lazarus, R. A., 1995, Potent and selective Kunitz domain inhibitors of plasma kallikrein designed by phage display, *J. Biol. Chem*. **270:**25411–25417.
22. Privalov, P. L., and Khechinashvili, N. N., 1974, A thermodynamic approach to the problem of stabilization of globular protein structure: A calorimetric study, *J. Mol. Biol*. **86:**665–684.
23. Murphy, K. P., Privalov, P. L., and Gill, S. J., 1990, Common features of protein unfolding and dissolution of hydrophobic compounds, *Science* **247:**559–561.
24. Murphy, K. P., and Freire, E., 1992, Thermodynamics of structural stability and cooperative folding behavior in proteins, *Adv. Protein Chem*. **43:**313–361.
25. Murphy, K. P., and Gill, S. J., 1991, Solid model compounds and the thermodynamics of protein unfolding, *J. Mol. Biol*. **222:**699–709.
26. Ganesh, C., Narayanan, E., Srivatsava, S., Ramakrishnan, C., and Varadarajan, R., 1999, Prediction of the maximal stability temperature of monomeric globular proteins solely from amino acid sequence, *FEBS Lett*. **454:**31–36.
27. Hollien, J., and Marqusee, S., 1999, A thermodynamic comparison of mesophilic and thermophilic ribonucleases H, *Biochemistry* **38:**3831–3836.
28. Grattinger, M., Dankesreiter, A., Schurig, H., and Jaenicke, R., 1998, Recombinant phosphoglycerate kinase from the hyperthermophilic bacterium *Thermotoga maritima*: Catalytic, spectral and thermodynamic properties, *J. Mol. Biol*. **280:**525–533.
29. Beadle, B. M., Baase, W. A., Wilson, D. E., Gilkes, N. R., and Shoichet, B. K., 1999, Comparing the thermodynamic stabilities of a related thermophilic and mesophilic enzyme, *Biochemistry* **38:**2570–2576.
30. Renugopalakrishnan, V., 2004, *US Provisional Patents*.
31. Renugopalakrishnan, V., Strzelczyk, A., Pingzou, Li, Mokhnatyuk, A. A., Gursahani, S. H., Nagaraju, M., Prabhakaran, M., Arjomandi, H., and Lakka, S. L., 2003, Retro-engineering bacteriorhodopsins: Design of smart proteins by bionanotechnology, *Int. J. Quantum Chem*. **95:**627–631.
32. Jiang, X., Bishop, E. J., and Farid, R. S., 1997, A de novo designed protein with properties that characterize natural hyperthermophilic proteins, *J. Am. Chem. Soc*. **119:**838–839.
33. Elkins, C. J., 1995, *Thermostability Comparison of Rubredoxin from Clostridium pasteurianum and Pyrococcus furiosus by Site-Directed Mutagenesis,* Masters Thesis, University of Georgia.
34. Wampler, J. E., and Neuhaus, E. B., 1997, A model for the unusual kinetics of thermal denaturation of rubredoxin, *J. Protein Chem*. **16:**721–732.
35. Verma, C. S., and Renugopalakrishnan, V., 2004, Computer experiments in the design of bionanodevices, in *Modeling and Simulating Materials Nanoworld, Advances in Science and Technology* (P. Vincenzini and F. Zerbetto, eds.), vol. 44, Techna Group, Srl, pp. 321–328.
36. Renugopalakrishnan, V., Garduño-Juárez, R., Narasimhan, G., Verma, C. S., Wei, X., and Pingzuo, Li, Rational design of thermally stable proteins, *J. Nanosci. Nanotechnol*. **5:**1759–1767.
37. Berezovsky, I. N., and Sakhhnovich, E. I., 2004, Of sequence and structure: Strategies of protein thermostability in evolutionary perspective, Los Alamos National Laboratory, Preprint Archive, *Quant. Biol*. 1–33.
38. Jiang, X., Harid, H., Pistor, E., and Farid, R. S., 2000, A new approach to the design of uniquely folded thermally stable proteins, *Protein Sci*. **9:**403–416.

Chapter 9

DECIPHERING ENGINEERING PRINCIPLES FOR THE DESIGN OF PROTEIN-BASED NANOMACHINES

Dan W. Urry
BioTechnology Institute, University of Minnesota, Twin Cities Campus, 1479 Gortner Avenue, Suite 240, St. Paul, MN 55108-6106
Bioelastics, Inc., 2423 Vestavia Drive, Vestavia Hills, AL 35216-1333

Abstract: *De novo*-designed elastic-contractile model proteins exhibit unique phase transitions (called inverse temperature transitions) that result in hydrophobic association on raising the temperature and that are capable of performing the set of energy conversions that occur in biology. The energy conversions involve the intensive variables of the free energy of mechanical force, temperature, pressure, chemical potential, exectrochemical potential and electromagnetic radiation frequencies ranging from the ultraviolet to the acoustic. Their physical characterizations and analyses have resulted in an understanding of reversible elastic force development, even by a single dynamic chain, and of the change in Gibbs free energy for hydrophobic association, $\mathbf{\Delta G_{HA}}$, that results from virtually every variable of interest to protein structure and function. A functional component within $\mathbf{\Delta G_{HA}}$ is a water-mediated apolar-polar repulsion, $\mathbf{\Delta G_{ap}}$. These physical principles have been applied with success to gain an understanding of the function of key biological energy converting proteins, namely Complex III of the mitochondrial electron transport chain, the F_1-motor of ATP synthase of the inner mitochondrial membrane, and the myosin II motor of muscle contraction.

With the assist of atomic force microscopy (AFM) in the single-chain force-extension mode and in a dynamic force spectroscopy mode that takes advantage of the acoustic mechanical resonances exhibited by these elastic contractile model proteins, these physical processes become engineering principles for the construction of nanosensors with the potential to detect single molecular events.

V. Renugopalakrishnan et al. (eds.), Bionanotechnology, 39-60.

1. INTRODUCTION

1.1 Renowned Nanomachines of the Biological Cell Confirm Engineering Principles and Inspire Nanomachine Design

1.1.1 Definition of a machine

The *McGraw-Hill Concise Encyclopedia of Science and Technology*, 2nd Edition, 1989, defines a machine as the following: "A combination of rigid or resistant bodies having definite motions and capable of performing useful work."

Proteins and protein-based polymers constitute the "resistant bodies" of interest here. Proteins are polypeptides, and protein-based polymers are composed of repeating peptide sequences in which the repeating unit may be as small as a dipeptide or as large as hundreds of residues. The definite motions of protein arise from changing hydrophobic associations that result in deformations and relaxations of protein chain segments capable of performing mechanical work. The changing hydrophobic associations are also capable of performing a range of other forms of work, most commonly chemical work.

1.1.2 A *de novo*-design approach using elastic-contractile model proteins gives rise to engineering principles

The approach to the elucidation of engineering principles required for the design of protein-based nanomachines, as reviewed in the following section, began with a relatively simple repeating pentapeptide sequence from an extracellular matrix protein. Then, without recourse to an already known protein-based machine, *de novo* design involving stepwise introduction of compositional changes in elastic-contractile model proteins resulted in the demonstration of diverse energy-converting capacities. The approach of designing and characterizing elastic-contractile model proteins allowed recognition of engineering principles that have now been used, as demonstrated in the following section, to describe the function of key intracellular protein machines. By this means, engineering principles for design of protein-based nanomachines have been developed and substantiated.

1.1.3 Nanomachines of living cell confirm engineering principles

Driving forces for cellular machines differ from those of man-made macroscopic machines. The protein machines within the cell are of nanometer

dimensions and require an aqueous milieu in which to function. These protein nanomachines are remarkable structures capable of operating at high efficiencies. In some cases, they bear astonishing structural analogy to machines of the macroscopic world, even though composed of very different ''rigid or resistant'' materials that operate by means of very different forces. The forces that dominate in protein-based machines are not so familiar to the mechanical engineer. The forces of hydrophobic association/dissociation and those concerned with the extent of irreversible deformation of single protein chain segments replace gravitation and surface frictional loss of the macroscopic world. The interior of the living cell, therefore, provides the ultimate site for confirming model protein-derived engineering principles for the design of protein-based nanomachines.

The difficulty of determining engineering principles from cellular protein machines. The compositional complexity of the evolutionarily optimized protein nanomachine of the cell makes difficult a direct attack on elucidating the engineering principles. Rather a *de novo* design approach, in which the energies and forces can be dissected by careful control of composition and selection of experimental method, provides the opportunity, one variable at a time, to characterize the energies and forces involved in function. Elastic-contractile model proteins, capable of a phase transition unique to aqueous systems, exhibit the properties essential for revealing the sought-after engineering principles.

Engineering principles from designed elastic-contractile model proteins confirmed by crystal structure in two key functional states of three vital cellular protein machines. Once a set of engineering principles is derived using elastic-contractile model proteins, these principles can be tested by detailed consideration of cellular protein machines for which crystal structural detail is available in critical functional states. In the present case, the principles are remarkably demonstrated in Complex III of the electron transport chain of mitochondria that produces the proton gradient, in the adenosine triphosphate (ATP) synthase of the inner mitochondrial membrane that uses the proton concentration gradient to produce nearly 90% of the ATP for an organism, and in the myosin II motor of muscle contraction that uses ATP to produce motion essential to function of the organism.

1.1.4 Nanomachine designs inspired by intracellular protein constructs and properties

An elastic extracellular protein provided the starting point that led to the design of energy-converting elastic-contractile model proteins. On the other hand, the elastic intracellular protein, titin, composed of a linear array of repeating small globular units, functions within the muscle cell as a series of

shock absorbing elements that irreversibly absorb energy of deformation. When a single globular element is combined with subtending elastic-contractile model protein, the potential exists for the detection of single molecular events. As seen by single-chain force–extension studies using the atomic force microscope (AFM), a single chain of titin globular elements gives a sawtooth profile.[1] Each sawtooth represents a single unfolding globular protein element irreversibly absorbing a unique packet of deformation energy. On the other hand, the elastic-contractile model proteins give a simple monotonic curve in the AFM single-chain force–extension curve.[2] By the design of a single globular sensing element subtended by elastic–contractile sequences, a single-chain force–extension curve is obtained with a single sawtooth. As a single interaction with the globular element can change the characteristic sawtooth for that globular element, it becomes a sensing element for detection of that interaction.

1.2 *De Novo* Design of Diverse Elastic-Contractile Protein Machines

The potential for more effective design of new protein-based nanomachines arises from the previous *de novo* design of elastic-contractile model proteins capable of performing the set of energy conversions familiar in living organisms.[3] The set of energy conversions involves the intensive variables of mechanical force, pressure, chemical potential, temperature, electrochemical potential, and electromagnetic radiation covering the frequency range from the ultraviolet to the acoustic. The interconversions of energy involving the 6 intensive variables result in 15 possible pair-wise energy conversions. When chemo-chemical transduction (the conversion of one chemical energy into another), electro-electrical transduction (the use of the reduction of one redox group to change the reduction potential of another redox group), and energy conversions between different frequencies of the electromagnetic radiation spectra are included, the number of 15 expands to 18 classes of pair-wise energy conversions.

1.2.1 The initial elastic-contractile model protein

The starting protein-based polymer was found as a repeating pentapeptide sequence in the mammalian elastic protein, elastin. The repeating pentapeptide sequence is glycine–valine–glycine–valine–proline, or Gly–Val–Gly–Val–Pro, or simply GVGVP which in bovine elastin is found to repeat as many as 11 times in a single unbroken sequence.[4,5] High molecular weight protein-based polymers of this repeating sequence, $(GVGVP)_n$, were found to be soluble in water at temperatures below 25°C and phase separate on

raising the temperature above 25°C to form a more ordered viscoelastic and filamentous state. When cross-linked, for example by γ-irradiation, the resulting elastic sheets swell on lowering the temperature and contract on raising the temperature above 25°C and do so when loaded with a weight. *Thus, the cross-linked polymer is capable of thermo-mechanical transduction*.[3]

1.2.2 Changing hydrophobicity to change temperature range for contraction

Changing the polymer composition by addition of a CH_2 group per pentamer, achieved by replacing a Val residue by an isoleucine (Ile, I) residue, makes the polymer more hydrophobic, and the contraction begins at 10°C. When deleting two CH_2 groups by replacing a Val residue by an alanine (Ala, A) residue, the contraction begins on raising the temperature above 45°C. *Accordingly, the temperature at which contraction occurs, becomes lower as a polymer becomes more hydrophobic and higher as a polymer becomes less hydrophobic*.

1.2.3 Introduction of functional groups with different states of hydrophobicity

Now, it becomes possible to change the temperature of the phase separation by substitution of an aliphatic amino acid residue by an amino acid residue with a functional side chain. For example, the charged carboxylate of glutamic acid (Glu, E), or of aspartic acid (Asp, D), or the charged amino function of lysine (Lys, K) are less hydrophobic and more polar than the uncharged (more hydrophobic) state of the side chain. Accordingly, changing the pH to form the uncharged side chain lowers the temperature of the transition. When at an intermediate temperature, therefore, neutralizing the charged side chain lowers the temperature of the transition and drives contraction. *Thus, the suitably substituted cross-linked polymer becomes capable of efficient chemo-mechanical transduction*.[6] Similarly, a redox function can be bound to the protein-based polymer. Since the oxidized state can be more polar and the reduced state can be more hydrophobic, *designed polymers capable of electro-mechanical transduction result*.[7]

1.2.4 Phenomenological control of diverse energy conversions

Thus, *de novo* designs of diverse energy conversions utilizing elastic-contractile model proteins are achieved, phenomenologically, by exercising design control of the temperature at which hydrophobic association occurs. These protein-based motors were not simply derivative of then known biological machines. Subsequent extensive analysis of these phenomenological designs

gave rise to new understandings of the factors that control hydrophobic association/dissociation and the nature of entropic elasticity.[8] These new understandings, referred to as the hydrophobic and elastic consilient mechanisms, become the engineering principles for the design of diverse protein-based nanomachines.

1.2.5 Relevance to biology's protein-based machines

Coherence of phenomena between designed elastic-contractile model protein machines and biology's protein machines. When considering coherence of phenomena between the behaviors of elastic-contractile model proteins designed for energy conversion and the natural energy converting proteins of the cell, the hydrophobically associated state is identified with the contracted (rigor) state of a muscle protein or the taut, tense, T-state of a globular protein such as hemoglobin. Formation of this contracted or T-state occurs with loss of hydration. As concerns muscle contraction, clinically there are recognized acid rigor (association of protein caused by acids), calcium rigor (contraction caused by excess calcium ion), heat rigor (heating that drives hydrophobic association), and rigor mortis (contraction on loss of ATP, e.g., dephosphorylation). These muscle phenomena have demonstrable correlation with the elastic-contractile model proteins designed for energy conversion.

Another well-characterized protein is hemoglobin with its T(taut)-state and R(relaxed)-state. Increased temperature favors the T-state; binding polar oxygen favors the R-state that forms with the uptake of water, and ion-pairing decreases charge to favor the T-state. Thus, there exists a clear coherence of phenomena.

Remarkable positive cooperativity of designed elastic-contractile model proteins when compared to biology's renowned protein machines. These elastic-contractile model proteins of translational symmetry are designed as chemo-mechanical transductional model proteins that exhibit positive cooperativities with Hill coefficients as large as 8. This is a greater positive cooperativity than seen by Monod and others, when concerned with symmetries of associating globular proteins. This was not expected from the prevailing view of cooperativity. As stated by Monod,[9] "One may set aside the simple problem of fibrous proteins. Being used as scaffolding, shrouds, or halyards, they fulfill these requirements by adopting relatively simple types of translational symmetries."

In the popular hemoglobin example, the Hill coefficient found for the oxygen-binding curve is 3.2, which indicates a degree of positive cooperativity much less than that obtained with the most efficient design of elastic-contractile model proteins. Thus, protein structure with translational

symmetry is designed to exhibit positive cooperativity that is central to efficient function (e.g., energy conversion) of nanomachines and exceeds the positive cooperativity considered so remarkable in protein machines of biology.

Direct demonstrations of relevance of the hydrophobic and elastic consilient mechanisms to biology's protein machines. Even though these demonstrations of energy conversion by *de novo* designed elastic-contractile model proteins proceeded visually and convincingly with the above-noted coherence of phenomena, suggestions that biology's protein-based machines functioned similarly were not well-received by specialists in their particular protein machine. These new understandings, nonetheless, have now been demonstrated to be significant in three key classes of biological energy conversion: the formation of the proton gradient in the mitochondrion, the use of the proton gradient to produce ATP by ATP synthase, and the use of ATP to produce motion essential to life.[10] Thus, the important new engineering principles did not arise from examination of the more complex, site localized, and evolutionarily optimized energy converting machines of biology. Instead, the new engineering principles arose from the design, preparation, characterization, and data analysis of relatively simple model proteins. In doing so, the model proteins could be designed to perform the energy conversions of biology by simple changes made in an elastic repeating sequence derived from the mammalian elastic fiber—a fiber not recognized as an energy-converting machine of biology.

1.3 Hydrophobic and Elastic Consilient Mechanisms: Definitions

Definition of consilience. Consilience reflects the presence of a pervasive, basic law of nature that underlies otherwise seemingly disparate phenomena. In this case, the phenomena involve a diverse set of energy conversions with an underlying common mechanism, "a common groundwork of explanation" in the words of E. O. Wilson.[11] Accordingly, the common groundwork of explanation for diverse energy conversions, as discovered using designed elastic-contractile model proteins, is called a consilient mechanism.

The hydrophobic and elastic consilient mechanisms. Two interlinked consilient mechanisms have been identified in studies on the *de novo* designed elastic-contractile model protein machines and also observed in the energy conversion of biology's protein machines. These are the hydrophobic and elastic consilient mechanisms. They derive their consilience from a general applicability, in the former case to all amphiphilic polymers in water and in the latter case to all polymers, of whatever composition, containing chain segments with a backbone mobility that can become damped on deformation whether by extension, compression, or repulsion.

1.4 Hydrophobic Consilient Mechanism and the Inverse Temperature Transition

Hydrophobic consilient mechanism epitomized in amphiphilic protein-based polymers. While the hydrophobic consilient mechanism applies to all amphiphilic polymers, it is most readily characterized in model proteins. This is because only in protein-based polymers is it possible to have strict control of a sequence in which each position may be filled by any one of 20 different residues of the same absolute configuration. This diversity of composition and control of sequence allows for determination of maximal interaction energies and their correlation using different experimental methods.

1.4.1 The inverse temperature transition and the formation of peptide structure on raising the temperature

Origin of the designation of an inverse temperature transition (ITT). The ITT is a unique kind of phase transition in which a clear solution of polypeptide in water at lower temperature becomes cloudy on raising the temperature into the transition range. On standing, the cloudy suspension phase separates to form a denser polypeptide phase and a clear overlying equilibrium solution. The term, ITT, originates from the observation that these polypeptide compositions dissolved in water become more ordered on phase separation. The most unambiguous demonstration of this occurs when cyclic polypeptide analogs, miscible in all proportions in water below a certain temperature, crystallize on raising the temperature through the temperature range of a phase transition and redissolve on lowering the temperature.[12] The corresponding linear polymers phase-separate on raising the temperature through the transition range to form parallel-aligned twisted filaments of about 5 nm width that redissolve on lowering the temperature.[13]

Implications of the Second Law of Thermodynamics on the process of the ITT. The Second Law of Thermodynamics requires that the order of the total system, polypeptide plus water, decreases on increasing the temperature. Since the order of the polypeptide part of the system clearly increases on raising the temperature through the range of the ITT, the Second Law requires that the water part of the system undergoes a decrease in order that is greater than the increase in order of the polypeptide. The one recognized way by which this might occur is when structured water around hydrophobic groups disappears as hydrophobic association of structure formation occurs.

Evidence for structured water around hydrophobic groups. Several studies established the presence of structured water around hydrophobic groups. In 1937, Butler[14] found dissolution of hydrophobic CH_2 groups in water to be

exothermic, yet solubility became limiting due to an increase in order of water. Frank and Evans[15] based on a marked decrease in vapor pressure on dissolution of hydrophobic groups in water, spoke of "icebergs" surrounding hydrophobic groups. Stackelberg and Müller[16] determined the crystal structure of water surrounding small hydrophobic molecules to be one of a pentagonal dodecahedron, that is, water molecules arranged at the apices of a pentagon. Water surrounding hydrophobic groups is indeed well ordered.

Accordingly, this ordered water, called hydrophobic hydration, would become less ordered bulk water as the hydrophobic groups associate. This loss of hydrophobic hydration during an ITT of elastic-contractile model proteins has been confirmed using microwave dielectric relaxation.[17] Furthermore, an ITT of hydrophobic association is expected to be an endothermic transition, that is, the heat released (as shown by Butler)[14] on formation of hydrophobic hydration is recovered as heat is taken up to return the hydrophobic hydration to disordered bulk water as hydrophobic groups associate. This means that the temperature and heat of an ITT would be central to an expression for the change in Gibbs free energy of hydrophobic association.

1.4.2 Thermodynamic quantities that control hydrophobic association

Determining the Gibbs free energy of hydrophobic association, ΔG_{HA}, *by calorimetry*. By means of differential scanning calorimetry (DSC), the endothermic heat of the ITT can be determined.[18] It is shown in Section 2.1 that an expression for the change in Gibbs free energy of hydrophobic association, ΔG_{HA}, can be derived and used to calculate this important thermodynamic quantity for any variable that changes the temperature and heat of an ITT. One fundamental variable is the amino acid composition, and the approach results in a ΔG_{HA}-based hydrophobicity scale for amino acid residues.[19] Importantly, the hydrophobicity scale includes both states of functional amino acids such as the carboxyl and carboxylate states of glutamic and aspartic acid residues. In your author's view, virtually every variable that affects protein function does so by their common effect on and their responsiveness to ΔG_{HA}.

Titration data and the apolar–polar repulsive free energy of hydration, ΔG_{ap}. The formation of carboxylate during an acid-based titration can disrupt hydrophobic association and redissolve a phase-separated state. A number of different experimental methods, including microwave dielectric relaxation,[17] demonstrate that the formation of charge, as in the formation of the carboxylate anion, destroys hydrophobic hydration. Furthermore, as the composition of the model protein becomes more hydrophobic, the pKa of the carboxyl shifts to higher pH values.[20] As the hydrophobicity increases, the carboxylate finds it more difficult to achieve adequate hydration; the free

energy of the carboxylate becomes less favorable, and the result is an increase in pKa. This reflects a competition for hydration between the apolar (hydrophobic) and the polar (charged) groups. A formal expression for such competition is an apolar–polar repulsive free energy of hydration, ΔG_{ap}.[3] Physically the competition for hydration expresses as repulsion because each group uses the flexibility of the intervening polymer chain to reach out and find water less perturbed by the other.

1.4.3 Hydrophobic hydration determines heat and temperature of inverse temperature transitions

Correlation of transition heat and temperature with amount of hydrophobic hydration. As carboxylates form, the number of waters of hydrophobic hydration, N_{hh}, decrease; the temperature of the transition, T_t, increases in a concentration dependent mirroring of T_t and N_{hh}, and the heat of the transition, ΔH_t, decreases.[10] Thus, the amount of hydrophobic hydration determines the heat and temperature of the ITT. If the value of T_t for the carboxyl state is below the operating temperature, hydrophobic association prevails, but, as charge forms, the value of T_t increases above the operating temperature and hydrophobic dissociation results. *Thus, the formation of charged species disrupt hydrophobic association*.

1.4.4 The source of the free energy of hydrolysis of ATP to form ADP plus P_i

Free energy of hydrolysis due to improved hydration. Several reasons are routinely given for the free energy of hydrolysis of ATP in the formation of adenosine diphosphate (ADP) and inorganic phosphate, P_i. These are opposing resonance, charge–charge repulsion within the triphosphate tail, and limited free energy of hydration of the triphosphate tail. As reviewed and argued elsewhere,[10] the improved free energy of hydration of ADP and P_i when compared to that of ATP dominates the change in free energy on hydrolysis. As the competition for hydration between hydrophobic and charged groups dominates in ΔG_{HA} and is specifically expressed as ΔG_{ap}, and as ATP carries multiple charges and has such a thirst for hydration, the states of ATP are expected to exhibit the most profound effects on hydrophobic association.

Why is ATP the universal energy currency of Biology? On the basis of the hydrophobic consilient mechanism, ATP functions as the universal energy currency of biology due to the capacity of negative charges to disrupt hydrophobic association and from the fact that phosphates constitute the most charged and hydration limited species of biology. At the moment of hydrolysis by ATPase, the localized ADP and P_i constitute the most polar state

routinely formed in biology. At that moment, any transient occurrence of hydrophobic hydration (as when a hydrophobic association undergoes a fluctuation toward dissociation) is immediately recruited for phosphate hydration, thereby removing the driving force for hydrophobic reassociation. Dephosphorylation, therefore, provides the most dramatic means for driving hydrophobic reassociation. Herein resides the power, whereby the most charged groups commonly available to biology control protein function.

1.4.5 The power of ATP hydrolysis to drive protein machines and perform the work of the cell

Formation of more polar states disrupt hydrophobic association. From studies on elastic-contractile model proteins, the enzymatic use of ATP to phosphorylate the serine (S) residue of poly[30(GVGIP)(RGYSLG)] achieves phosphorylation to about 50% and shifts the transition temperature on a per pentamer basis, $\mathbf{T_t}$, from 18° to 860°C.[21,10] Using the plot of $\mathbf{T_t}$ versus $\mathbf{\Delta G_{HA}}$, developed in Section 2.3.1, the value of $\mathbf{\Delta G_{HA}}$ appears to be about +8 kcal/mol. Furthermore, the extent of phosphorylation indicated an equilibrium constant of essentially 1. This means that the γ-phosphate of ATP and the phosphate bound to the serine of the elastic-contractile model protein are at the same free energy, which would be about +8 kcal/mol. The apolar–polar repulsive free energy, i.e., the change in Gibbs free energy of hydrophobic association on phosphorylation, and the free energy of hydrolysis of ATP to form ADP and P_i are the same quantity. *Thus, phosphorylation appears to achieve function by its disruption of hydrophobic association*.

Formation of the most polar state can physically repulse a hydrophobic domain. As discussed in the following section for the F_1-ATPase of ATP synthase, the hydrolysis of ATP to form the most polar state of localized ADP plus P_i repulses the most hydrophobic side of the γ-rotor to drive rotation in a counterclockwise direction. On the other hand, when the γ-rotor is driven by the F_0-motor in a clockwise direction, the most hydrophobic face of the rotor is rotated in the opposite direction to develop a through-water repulsion with the most polar ADP plus P_i state. In your author's view, the resulting repulsion is sufficient to cause the most polar ADP plus P_i state to become less polar by formation of ATP.[10,22]

1.5 The Elastic Consilient Mechanism and the Nature of Near Ideal Elasticity

The perfectly reversible stress–strain curve of an ideal elastomer states that energy of deformation is completely recovered on relaxation. The area under a

stress–strain curve, which plots elastic force versus extension, gives the energy expended during deformation. If the relaxation curve exactly traces the extension curve, then all of the energy expended on deformation is recovered on relaxation. Such a perfectly reversible stress–strain curve identifies an ideal elastomer. When a protein machine develops an elastic force, optimal energy conversion requires an ideal elastomer.

The consequence of hysteretic stress–strain curves is that much energy of deformation is not recovered on relaxation. When the stress–strain curve for relaxation falls below the curve obtained on extension, hysteresis occurs. As the area under the relaxation curve becomes less than the area under the curve for extension, energy of deformation is lost to its surroundings as the chain is stretched. As the hysteresis becomes greater, the efficiency of a machine that utilizes the elastomer for energy storage during transduction becomes lower. Thus, the mechanism of elasticity becomes an important issue in the design of efficient protein-based nanomachines.

Means for overcoming hysteresis with elastic–contractile model proteins. In the AFM-based single-chain force–extension studies, the elastic-contractile model proteins,[2,23] $(GVGVP)_{nx251}$ and $(GVGIP)_{nx260}$, provide informative examples. Under dilute conditions, $(GVGVP)_{nx251}$ exhibits reversible force–extension curves as long as time at lower extensions was insufficient for association of chain with substrate. If reassociation with substrate was allowed to occur, a small hysteresis could be observed, that is, on pealing a chain away from nonspecific adsorption to the substrate, energy required for extension is irreversibly lost to substrate on forced dissociation.

For the more hydrophobic $(GVGIP)_{nx260}$ under dilute conditions, very marked hysteresis was observed. Only under high dilution (where the order of 100 attempts would be required before a chain would be picked up)[24] could reversible, non-hysteretic, force–extension curves be observed. At higher concentration, however, if a chain were pulled out of a cluster of chains, or for that matter from an amorphous globule, hysteresis would occur. In this case, loss of energy to the associated non-load bearing chains would result in hysteresis. This is one reason why your author does not believe that reversible entropic elastic force is possible when the load-bearing chain is pulled from an amorphous globule of associated chains.[25]

Even for a single chain, if time were allowed at low extension for self-association or back folding to occur, hysteresis results, that is, energy of deformation is lost to non-load bearing chains.[2,23] This demonstrates much about the requirements for a protein-based machine to perform mechanical work efficiently. Accordingly, a freestanding chain would be optimal for approaching ideal elasticity and efficient energy conversion. A freestanding chain is observed at two different extensions in the Rieske Iron Protein (RIP) of Complex III (the cytochrome bc_1 complex) of the electron transport chain.

Extension of the freestanding chain occurs on hydrophobic association at the globular end of the single chain. Once the hydrophobic association is lost due to the formation of two positive charges, on oxidation of ubiquinol at the site of hydrophobic association, the relaxing chain retracts the FeS-containing globular domain of the RIP and moves it to the heme c_1 site. Thereby elastic retraction of an extension due to hydrophobic association transports an electron from ubiquinol to cytochrome c_1.

1.5.1 Entropic (ideal) elastic force, the property of a single dynamic chain

As became evident in the immediately preceding paragraphs, if extension involves dissociation of the chain from other chains, from a surface, or from non-load bearing sections of the same chain, at the moment of dissociation, energy is lost from the chain undergoing load-bearing extension. The extension will not result in a reversible extension/relaxation, that is, will not provide for an entropic, near-ideal elastic force. Any protein-based machine that dissipates some of its input energy into hysteretic deformations of that sort would be a less efficient machine. By this reasoning, reversible (near-ideal) elasticity is a key property that is best provided by a single, freestanding dynamic chain. What then must be the origins of the entropic elastic force?

1.5.2 Damping of internal (backbone) chain dynamics on deformation

Sources of entropy in the extended load-bearing elastic chain. In the AFM-based single-chain force–extension study, entropic elastic force develops.[2,23] Accordingly, entropic elastic force, the surest way to ideal (reversible) elastic force development is attainable with a single chain. Therefore, entropic elastic force development does not require a random chain network, as taught by the classical (random chain network) theory of rubber elasticity.[26,27]

From statistical mechanics, the entropy of a molecular system can be expressed by use of the Boltzmann's relation, $S = k \ln W$, where S is the entropy, k is the Boltzmann's constant, and W is the number of *a priori* equally probable states accessible to the molecular system.[28] W is commonly obtained as the product of partition functions for the different degrees of freedom accessible to the molecular system, there being $3N$ degrees of freedom where N is the number of atoms in the molecule. There are three translational degrees of freedom, but in the force–extension experiment the single chain is gripped at each end giving no translational freedom. There are also three whole molecule rotational degrees of freedom, but again these cannot occur in the elasticity experiment. In the stress–strain experiment the only degrees of freedom are those due to internal chain motions such as vibrational modes

and torsional oscillation motions about bonds. *Therefore, the entropic elastic force must arise from the damping of internal chain dynamics on extension.*

On the basis of the harmonic oscillator partition function; however, higher frequencies, such as vibrations, contribute very little to the entropy. On the other hand, torsion angle oscillations that occur at much lower frequencies contribute much more to the chain entropy.[8] *Thus, the development of entropic elastic force, reversible elastic deformation, derives from the decrease in amplitude of torsion angle oscillations on extension.*[29]

At extension forces greater than available from damping of internal chain dynamics, yet well below the limit of chain rupture, a reversible elastic force development can occur that would contribute an internal energy component to the total force. When force development arises from hydrophobic association, the energy storage due to the sum of the internal energy and entropy components of the force would, of course, be substantially less than the free energy gained on hydrophobic association in the absence of chain extension.

1.5.3 Unique internal chain dynamics of family of elastic-contractile model proteins

The occurrence of frequency-localized mechanical resonances above the transition temperature. Repeating dynamic sequences provide the opportunity for each of these dynamic repeats, once in a regular structure, to contribute to a common frequency with the result of strong absorptions due to motion over limited frequency ranges. On raising the temperature through the range of the ITT, the elastic-contractile model proteins form regular structures in which the individual repeats fold into a common conformation and undergo motions of a common localized frequency. The result is the development of an intense relaxation frequency that has been given the name of a *mechanical resonance*. In particular, as the temperature of aqueous solutions of $(GVGVP)_{251}$ and $(GVGIP)_{320}$ is raised from below to above the temperature range of their respective ITTs, dielectric relaxation studies demonstrate the development of intense mechanical resonances near 5 MHz and 3 kHz. These resonances represent the coordinated motion of peptide dipoles within the repeat so that a net oscillating dipole moment results and sums from one repeat to the next to give an intense relaxation at the noted localized frequencies. For example, the dielectric increment for the 5 MHz relaxation of $(GVGVP)_{251}$ is 70.[30] On the other hand, the dielectric constant of water at a similar temperature is less than 80, even though the density of water molecules is several times that of the peptide moiety.

The importance of mechanical resonances to the stability of folded state. Using the harmonic oscillator partition function, the contribution to the entropy, S_i, as a function of frequency, ν_i, is written as

$$S_i = R\left[\ln(1 - e^{-h\nu_i/kT})^{-1} + \left(\frac{h\nu_i}{kT}\right)(e^{h\nu_i/kT} - 1)^{-1}\right], \quad (1)$$

where R is the gas constant, h and k are the Planck and Boltzmann's constants, and T is the temperature in °K. When the entropic contribution to the free energy, TS_i, was calculated for a temperature of 298°K (25°C), the contribution to the Gibbs free energy of the 5 MHz mechanical resonance was found to be 9 kcal/mol and that of the 3 kHz mechanical resonance was 14 kcal/mol.[8] The harmonic oscillator partition function was derived for quantized vibrational motions. The results, nonetheless, remain instructive and indicate a great stabilization of the structure that results from such low frequency mechanical resonances. Incidentally, this is thought to explain the remarkable biocompatibility of this family of model proteins.[10]

The presence of the 3 kHz mechanical resonance, being in the acoustic absorption frequency range, presents a number of applications for these polymers, including sound absorption. In the area of nanomachine type biosensors, the 3 kHz absorption would allow for the design of a relatively robust single molecule detection device that could be used in the field for detection of analytes such as explosives, nerve gasses, toxins, etc. This is because the cantilever would be fixed in space, that is, the z-direction of the AFM as well as the x, y dimensions would not be scanned. This would allow for the device to function without taking such extreme measures for vibration isolation.

1.6 Coupled Hydrophobic and Elastic Consilient Mechanisms

Coupling of hydrophobic and elastic consilient mechanisms allows insight into the *élan vital*, literally "the vital force or impulse of life."[31] In a common scenario, the impulse can be thought of as initiated by hydrophobic associations that cause elastic force development within interconnecting chain segments. The result is a sustained impulse imparted by relaxation of the stretched chain. Dephosphorylation or formation of another favorable hydrophobic domain with which to associate can initiate hydrophobic association. In specific protein-based machines, formation of the most favorable hydrophobic associations results in the stretching of interconnecting chain segments.

1.6.1 Insights developed from elastic-contractile model proteins

Entropic elastic force development during isometric contractions due to hydrophobic association. Properly designed elastic-contractile model proteins

demonstrate thermally and chemically driven contractions at fixed extension, that is, under isometric conditions. Under the former experiment, the elastic band is held at fixed extension, and the temperature is raised from below to above the temperature range of the ITT. On passing through the transition, hydrophobic associations increase within the elastic band, but since the elastic-contractile band is of fixed length, the result is the development of an elastic force, that is a thermally-driven isometric contraction.[8]

Alternatively, the elastic-contractile model protein may contain one glutamic acid (Glu, E) residue per 30 residues. The elastic-contractile band is again mounted at fixed length at physiological pH, and the pH is lowered causing neutralization by protonation of the charged carboxylate residue. This lowers the temperature of the transition from above to below the operating temperature of 37°C and drives hydrophobic association. But since the band is held at fixed length, the result is an increase in elastic force, that is, there is a chemically driven isometric contraction.[32] *Thus, hydrophobic associations stretch interconnecting chain segments!*

Under isometric conditions, hydrophobic association can occur only with the loss of hydrophobic hydration, that is, with an increase in entropy of the water. The resulting stretched interconnecting chain segments represent entropic elastic force due to a decrease in entropy, that is, due to a decrease in chain entropy as the result of a decrease in amplitude of torsion angle oscillations. *Thus, the development of entropic elastic force under isometric conditions cannot be due to proposed solvent entropy changes*.[25]

1.6.2 The role of hydrophobic and elastic consilient mechanisms in design of nanomachines

It is your authors' thesis that the hydrophobic and elastic consilient mechanisms constitute the two most dominant engineering principles in the design of protein-based nanomachines. During function of biology's protein-based nanomachines, changes in hydrophobic association often effect changes in the elastic forces sustained by specific peptide sequences. In particular, as argued in Section 3, the coupling of the hydrophobic and elastic consilient mechanisms is central to the electron transfer function of the RIP subunit of Complex III and to the development of elastic force during an isometric contraction in the myosin II motor. Most to the point of this review, nanomachines designed using protein-based constructs utilize the capacity of changes in hydrophobic association to effect changes in elastic force.

2. PRINCIPAL THERMODYNAMIC QUANTITIES CONTROLLING DIVERSE ENERGY CONVERSIONS IN MODEL PROTEINS

2.1 The Change in Gibbs Free Energy for Solubility, $\Delta G(\text{solubility}) = \Delta H - T\Delta S$

The source of hydrophobic insolubility and the equivalence of phase separation by hydrophobic association. The statement of the Gibbs free energy for solubility is $\Delta G(\text{solubility}) = \Delta H - T\Delta S$, and, of course, solubility occurs as $\Delta G(\text{solubility})$ becomes negative. As noted earlier, the original insights into the insolubility of hydrophobic groups come from the alcohol data of Butler[14] in which the addition of each CH_2 gives a favorable average heat release of $-1.3\,\text{kcal/mol-CH}_2$ for the series of normal alcohols from methanol to pentanol.

One then asks why is *n*-octanol insoluble in water? The answer is that the $(-T\Delta S)$ term for *n*-octanol dominates such that $\Delta G(\text{solubility})$ is positive. In fact, as noted by Butler for the alcohol series, the $(-T\Delta S)$ term changes by $+1.7\,\text{kcal/mol-CH}_2$. Accordingly, for methanol $\Delta G(\text{solubility})$ is a sufficiently negative value that the addition of four CH_2 groups, i.e., $4\ CH_2\ \text{groups} \times 1.7\,\text{kcal/mol-CH}_2 = +6.8\,\text{kcal/mol}$ for the $(-T\Delta S)$ term still leaves a negative value for $\Delta G(\text{solubility})$ for *n*-pentanol. By the time three more CH_2 groups are added, as for *n*-octanol, $\Delta G(\text{solubility})$ becomes substantially positive and solubility is lost.

For the insolubility of a phase separation brought about by raising the temperature, the $(-T\Delta S)$ term simply increases with increase in the value of T and overcomes the inherently negative ΔH term to result in the insolubility of phase separation. For a chemically driven phase separation, the temperature may decrease by 10% in degree kelvin, but the negative change in entropy for the transition can increase in magnitude by several 100% until the positive $(-T\Delta S)$ term dominates and solubility is lost.

Components of an endothermic ITT. Again as shown by Butler in 1937 for the methanol to *n*-pentanol series,[14] dissolution of the hydrophobic CH_2 moiety in water is an exothermic reaction with a mean value for the series of four CH_2 additions of $-1.3\,\text{kcal/mol-CH}_2$. This means that the reverse process of a phase separation by hydrophobic association in which CH_2 goes from soluble (dissolved, i.e., surrounded by water in solution) to insoluble (phase separated where water no longer surrounds CH_2 groups) is necessarily an endothermic process, as observed for the ITT.

In the polymer association process of the ITT, there is also an exothermic component due to van der Waal's contacts between associating chains. This exothermic element of the ITT has been resolved by temperature modulated differential scanning calorimetry (TMDSC)[33] and found to be less than one-third the magnitude of the total endothermic reaction of the ITT for $(GVGVP)_n$ and $(GVGIP)_n$. As long as there is a measurable net endothermic transition, the endothermic component, which defines the ITT, will be larger than the exothermic component, and the transition will be *dominated* by the effect of disruption of hydrophobic hydration.

One calculates the change in Gibbs free energy of hydrophobic association ($\mathbf{\Delta G_{HA}}$) using the temperature and net endothermic heat of the transition, which contains the exothermic aspect of van der Waal's interactions (the association of protein chains) as part of the hydrophobic association. $\mathbf{\Delta G_{HA}}$, therefore, represents a water-*dominated* endothermic process, but it also contains the smaller oppositely signed, exothermic effect of polymer association involving van der Waal's contacts.

2.2 The Change in Gibbs Free Energy for a Phase Transition

The change in Gibbs free energy during a narrow phase transition is zero. At the temperature of a phase transition, the Gibbs free energy per mole (also called the chemical potential, μ) of the molecules undergoing transition is the same in the soluble state as it is in the phase separated state, i.e., μ(soluble) = μ(insoluble). For an infinitely narrow transition, all of the heat going into the system serves to shift the molecules from soluble to insoluble and not to raise the temperature. Accordingly, for a sufficiently narrow transition the change in Gibbs free energy for the transition, ΔG_t, is zero, i.e., $\Delta G_t(\text{solubility}) = 0$, and therefore,

$$\Delta H_t = T_t \Delta S_t. \tag{2}$$

The change in Gibbs free energy for a phase transition. If we write the transition for two different experimental conditions denoted by a change, χ, in the system, and consider χ as a change from a reference state, ref, we can write, in general, for the change in Gibbs free energy for a phase transition,[19]

$$\delta\Delta G_t(\chi) = \delta\Delta H_t(\chi) - T_t(\chi)\delta\Delta S_t(\chi), \tag{3}$$

where the definitions are $\delta\Delta H_t(\chi) \equiv [\Delta H_t(\chi) - \Delta H_t(\text{ref})]$ and $\delta\Delta S_t(\chi) \equiv [\Delta S_t(\chi) - \Delta S_t(\text{ref})]$.

2.3 The Change in Gibbs Free Energy for an Inverse Temperature Transition

The width of the ITT involving high-molecular-weight polymers that fold and associate to form regular structures on raising the temperature, at the rate of 1°C/min, is quite broad with a half-width of 10°C or more. At any given temperature within the transition, the chemical potential of the soluble, hydrophobically dissociated polymer, μ_{HD}, and that of the hydrophobically associated polymer, μ_{HA}, are the same, i.e., $\mu_{HD}(T_{t1}) = \mu_{HA}(T_{t1})$, where (T_{t1}) is a specified temperature within the transition temperature range. When the heat of the transition is determined over the entire transition width, however, ΔG_t is only approximated as zero, because over the width of the transition, heat has also gone into raising the temperature of the system. Therefore, the equivalent of Eq. (2) is only approximate, i.e., $\Delta H_t \approx T_t \Delta S_t$. Specifically, for the endothermic ITT of a high-molecular weight protein-based polymer, the change in Gibbs free energy for hydrophobic association, $\Delta G_{HA}(\chi)$ is given as approximating $\delta\Delta G_t(\chi)$, such that, $-\Delta G_{HA}(\chi) \approx \delta\Delta H_t(\chi) - T_t(\chi)\delta\Delta S_t(\chi)$.

2.3.1 The change in Gibbs free energy for hydrophobic association, ΔG_{HA}

Definition of the reference state, f_X=1. When the change in thermodynamic quantity is written with regard to a specific reference state, the quantity is boldfaced. For example, χ may represent a change in amino acid composition within a pentamer, such as GVGVP replaced by GXGVP, where X is any amino acid residue or a chemical modification thereof. The experimental thermodynamic quantity, written for the pentamer, (GXGVP) in poly(GXGVP), may not be experimentally measurable because it may lie outside the experimentally accessible range. In practice, one considers the general composition, poly[f_X(GXGVP), f_V(GVGVP)], where f_X and f_V represent the mole fraction of pentamer in the polymer such that $f_X + f_V = 1$. When X stands for the phenylalanine (Phe, F) residue, the value of T_t(GFGVP) for poly(GFGVP) would be below −30°C and not directly measurable. It is possible, however, to measure T_t for poly[f_F(GFGVP), f_V(GVGVP)] for several values of f_F, say 0.1, 0.2, and 0.3, and then linearly extrapolate to $f_F = 1$. This example of $f_X = 1$ is called the reference state, and thermodynamic quantities obtained for the state of $f_X = 1$ are bold-faced, e.g., $\mathbf{T_t}$(GXGVP) and $\mathbf{\Delta H_t}$(GXGVP) and the derived quantities, $\mathbf{\Delta S_t}$(GXGVP) and $\mathbf{\Delta G_{HA}}$(GXGVP).

2.3.2 ΔG°_{HA}-based hydrophobicity scale for amino acid residues, using the heat and temperature of the inverse temperature transition

Choice of an amino acid residue as the reference state for a hydrophobicity scale. More than 40 hydrophobicity scales have been proposed for amino acid residues, but none of these has been based on the fundamentally relevant quantity, ΔG_{HA}. Many of these have taken the value of the glycine (Gly, G) residue as the zero reference state, as it is perceived to be neither polar nor apolar. The following scale follows this convention. Accordingly, $\Delta G^{\circ}_{HA}(GGGVP) = 0$, where the superscript is added to indicate that a choice of reference state has been made. This gives[19]

$$\Delta G^{\circ}_{HA}(GXGVP) \equiv [\Delta H_t(GGGVP) - \Delta H_t(GXGVP)]. \quad (4)$$

Equation (4), evaluated at $T_t(GXGVP)$, has been used in Table 1 to calculate the value of ΔG°_{HA}. For several X groups with only the value of $T_t(GXGVP)$ known, the sigmoid curve of T_t versus ΔG°_{HA} from Ref. [19] has been used to obtain ΔG°_{HA}. These values are indicated in Table 1 below by the notation (from graph).

2.3.3 ΔG°_{HA}-based hydrophobicity scale for chemical modifications of amino acid residues and prosthetic groups of proteins

Table 2 demonstrates large effects in ΔG°_{HA} attending the change in oxidative state of several biological prosthetic groups. For example, reduction of nicotinamide adenine dinucleotide (NAD), $NAD^+ \rightarrow NADH$, changes the Gibbs free energy for hydrophobic association by -7.5 kcal/mol, whereas the reduction of flavin adenine dinucleotide (FAD), $FAD \rightarrow FADH_2$, changes ΔG°_{HA} by -4.5 kcal/mol, about two-thirds that of NAD. This is in keeping with their respective roles in the electron transport chain of the inner mitochondrial membrane. The additional data of Table 2 show a number of useful modifications of amino acid side chains that may be employed in the design of nanomachines. The example of the large effect of serine phosphorylation is utilized in Section 4 in the design of a nanosensor for single molecule detection.

2.4 Apolar–Polar Repulsive Gibbs Free Energy of Hydration, ΔG_{ap}

2.4.1 Change in ΔG_{HA} on change of state of functional amino acid residues

For $(GVGVP\ GVGVP\ GEGVP\ GVGVP\ GVGVP\ GVGVP)_{36}$, the carboxyl (COOH) and carboxylate (COO^-) states of the glutamic acid (Glu, E)

Table 1. Hydrophobicity scale in terms of $\mathbf{\Delta G^{\circ}_{HA}}$, the change in Gibbs free energy of hydrophobic association for amino acid residues

Residue X	$\mathbf{T^{o}_{t}}$C	$\mathbf{\Delta G^{\circ}_{HA}}$ (GXGVP) kcal/mol-pent
W: Trp	−105	−7.00
F: Phe	−45	−6.15
Y: Tyr	−75	−5.85
H°: His°	−10($\mathbf{T_t}$)	−4.80 (from graph)
L: Leu	5	−4.05
I: Ile	10	−3.65
V: Val	26	−2.50
M: Met	15	−1.50
H^+: His^+	30($\mathbf{T_t}$)	−1.90 (from graph)
C: Cys	30($\mathbf{T_t}$)	−1.90 (from graph)
E°: Glu(COOH)	20 (2)	−1.30 (−1.50)
P: Pro	40	−1.10
A: Ala	50	−0.75
T: Thr	60	−0.60
D°: Asp(COOH)	40	−0.40
K°: Lys(NH_2)	40 (38)	−0.05 (−0.60)
N: Asn	50	−0.05
G: Gly	55	0.00
S: Ser	60	+0.55
R: Arg	60($\mathbf{T_t}$)	+0.80 (from graph)
Q: Gln	70	+0.75
Y^-: Tyr(ϕ-O^-)	140	+1.95
D^-: Asp(COO^-)	170($\mathbf{T_t}$)	≈ +3.4 (from graph)
K^+: Lys(NH_3^+)	(104)	(+2.94)
E^-: Glu(COO^-)	(218)	(+3.72)
Ser (PO_4^{2-})	860 ($\mathbf{T_t}$)	≈ +8.0 (from graph)

Data within parentheses utilized microbial preparations of poly(30)mers, e.g., (GVGVP GVGVP GXGVP GVGVP GVGVP GVGVP$_n$, with $n \approx 40$. The notation (from graph) indicates that the value of $\mathbf{T_t}$ was used to obtain $\mathbf{\Delta G^{\circ}_{HA}}$(GXGVP) from the experimental sigmoid curve of $\mathbf{T_t}$ vs $\mathbf{\Delta G^{\circ}_{HA}}$ of Ref. [19]. $\mathbf{T_t}$-values adapted from Ref. [3].

residue are represented as E°/0F and E^-/0F, respectively, in Table 1. Thus, using the values given in Table 1, i.e., −1.5 kcal/mol-(GE°GVP) for E°/0F and +3.72 kcal/mol-(GE^-GVP) for E^-/0F, one can write $\mathbf{\Delta G_{HA}}$[(GE°GVP)−(GE^-GVP)] $= -5.22\,kcal/mol$-(GEGVP). Protonation markedly decreases ΔG_{ap}, which relaxes the apolar–polar repulsion.

2.4.2 Hydrophobicity-induced changes in pKa values and Hill coefficients

An instructive set of glutamic acid (Glu, E) containing elastic-contractile model proteins, **I** through **V**, is listed below. These repeats contain one E

Table 2. ΔG°_{HA}-based hydrophobicity scale (preliminary T_t and ΔG°_{HA} values) to include chemical modifications and prosthetic groups of proteins[a]

Residue X	ΔG°_{HA}(kcal/mol)[g]	T_t at $f_X = 1$
Lys{dihydro NMeN}[b,d]	−7.0	−130
Glu{NADH}[c]	−5.5	−30°C
Lys{6-OH tetrahydro NMeN}[b,d]	−3.0	15°C
Glu{$FADH_2$}	−2.5	25°C
Glu{AMP}	+1.0	70°C
Ser{-O-SO_3H}	+1.5	80°C
Thr{-O-SO_3H}	+2.0	100°C
Glu{NAD}[c]	+2.0	120°C
Lys{NMeN, oxidized}[b,d]	+2.0	120°C
Glu{FAD}	+2.0	120°C
Tyr{-O-SO_3H}[e]	+2.5	140°C
Tyr{-O-NO_2^-}[f]	+3.5	220°C
Ser{PO_4^{2-}}[g]	+8.0	860°C

[a]Usual conditions are for 40 mg/ml polymer, 0.15 N NaCl and 0.01 M phosphate at pH 7.4.
[b]NMeN is for *N*-methyl-nicotinamide pendant on a lysyl side chain, i.e., *N*-methyl-nicotinate attached by amide linkage to the -NH_2 of Lys and the most hydrophobic reduced state is *N*-methyl-1,6-dihydronicotinamide (dihydro NMeN), and the second reduced state is *N*-methyl-6-OH 1,4,5,6-tetrahydronicotinamide or (6-OH tetrahydro-NMeN).
[c]For the oxidized and reduced nicotinamide adenine dinucleotides, the conditions were 2.5 mg/ml polymer, 0.2 M sodium bicarbonate buffer at pH 9.2.
[d]For the oxidized and reduced *N*-methyl-nicotinamide, the conditions were 5.0 mg/ml polymer, 0.1 M potassium bicarbonate buffer at pH 9.5, 0.1 M potassium chloride.
[e]The pK_a of polymer bound -O-SO_3H is 8.2.
[f]The pK_a of Tyr{-O-NO_2^-} is 7.2.
[g]Gross estimates of ΔG°_{HA} using the T_t-values in the right column in combination with the T_t versus ΔG°_{HA} values from sigmoid curve of T_t versus ΔG°_{HA}. Adapted from reference [10].

residue per 30 residues, constituted as six pentamers, and more hydrophobic phenylalanine (Phe, F) residues stepwise replace valine (Val, V) residues. In particular, model protein **I** contains one E residue and no F residues and is designated E/0F. In model proteins **II**, **III**, **IV**, and **V**, modestly hydrophobic V residues have been stepwise replaced by 2, 3, 4, and 5 very hydrophobic F residues, respectively. Obviously, increasing hydrophobicity systematically causes pKa shifts and increases the Hill coefficient, ***n***.[3]

	n	pKa
I:(GVGVP GVGVP G<u>E</u>GVP GVGVP GVGVP GVGVP)$_{36}$:E/0F,	1.5;	4.5
II:(GVGVP GVGFP G<u>E</u>GFP GVGVP GVGVP GVGVP)$_{40}$:E/2F;	1.6;	4.8
III:(GVGVP GVGVP G<u>E</u>GVP GVGVP GVGFP GFGFP)$_{39}$:E/3F;	1.9;	5.2
IV:(GVGVP GVGFP G<u>E</u>GFP GVGVP GVGFP GVGFP)$_{15}$:E/4F;	2.7;	5.6
V:(GVGVP GVGFP G<u>E</u>GFP GVGVP GVGFP GFGFP)$_{42}$:E/5F;	8,0;	6.4

The values of the pKa and ***n***, the Hill coefficient, increase as the competition for hydration between the carboxylate and the more hydrophobic F residues increases. Even in model protein **I**, the pKa is shifted from the value generally considered for an unperturbed glutamic acid residue of 3.8–4.0 due to the presence of the moderately hydrophobic Val residues.

Derivation of ΔG_{ap} *in terms of the* ΔpK_a. Starting with the definition of chemical potential for proton, $\mu_H = RT \ln a_H$, where R and T have their usual meaning of the gas constant (1.987 cal/mol-deg) and the absolute temperature in degree Kelvin, and a_H is the activity of proton. At the usual proton concentrations, proton activity can be replaced by proton concentration, $[H^+]$, and the natural logarithm can be replaced by log to the base ten, to give $\ln a_H = 2.3 \log [H^+]$. As $pH = -\log [H^+]$, μ_H becomes in general $-2.3\,RT\ pH$, and at the pH of the pKa, $\mu_H = -2.3\,RT\ pKa$. The change in chemical potential of proton, $\Delta\mu_H$, represented by a shift in pKa can be written as $\Delta\mu_H = -2.3\,RT\ \Delta pKa$. Now, since $\Delta\mu_H$ is the change in Gibbs free energy per mole, ΔG per mole, and as the shift derives from apolar–polar repulsion, we have,

$$\Delta G_{ap} = 2.3\,RT\Delta pKa. \tag{5}$$

Assuming an unperturbed pKa of 3.9 for Glu (E) and a T of 310°K in Eq. (5), for the above listed model proteins, **I**, **II**, **III**, **IV**, and **V**, the values for the apolar–polar repulsive free energy of hydration, ΔG_{ap}, is calculated to be 0.85, 1.28, 1.84, 2.41, and 3.54 kcal/mol-Glu residue, respectively.

2.4.3 Extent of positive cooperativity correlates with magnitude of hydrophobic-induced pKa shifts.

The relationship between the change in Gibbs free energy due to a change in cooperativity, $\Delta\Delta G_0$, has been approximated by Wyman as, $\Delta\Delta G_0 = RT(1 - 1/n)/\alpha(1 - \alpha)$, where α is the degree of ionization and ***n*** is the Hill coefficient. At an α of 0.5, the value at the pKa, the Wyman equation becomes $\Delta\Delta G_0 = 4\,RT(1 - 1/\boldsymbol{n})$, and taking RT to be 0.6 kcal/mol, the maximal $\Delta\Delta G_0$ would be 2.4 kcal/mol. For model proteins, **I**, **II**, **III**, **IV**, and **V** with a Hill coefficient, ***n***, of 1.5, 1.6, 1.9, 2.7, and 8.0, the values of $\Delta\Delta G_0$ are calculated to be 0.82, 0.92, 1.16, 1.55, and 2.16 kcal/mol Glu residue.

These values are to be compared to those obtained from Eq. (5) above. The values, of ΔG_{ap} for model protein **I** with a small pKa shift (0.6) and of $\Delta\Delta G_0$ with a small Hill coefficient (1.5), are essentially the same, 0.8 kcal/mol. On the other hand for model protein **V** with a pKa shift of 2.5, and a Hill coefficient of 8.0, the former gives 3.5 kcal/mol whereas the latter gives only 2.2 kcal. Thus, as pKa shifts become higher, the approximations of the Wyman equation cause

$\Delta\Delta G_0$ to become highly unsatisfactory, that is, to fall increasingly below ΔG_{ap} with the maximum value calculated by $\Delta\Delta G_0$ being 2.4 kcal/mol.

Phenomenologically, the source of positive cooperativity and pKa shift may be described similarly. The first ionization of the carboxyl is less probable because of the energy required to disrupt hydrophobic hydration in the process of achieving its own hydration. This delays the onset of ionization as the pH is raised for a carboxylate titration. As the pH is raised further, occasionally a second ionization will occur and assist the first such that both ionizations hold, but the presence of two facilitate the third in gaining its hydration and so on. Thus, there is a delayed onset followed by a rapid proliferation of ionizations, giving rise to the related phenomena of pKa shifts and positive cooperativity.

2.4.4 Correlation of independently derived values of ΔG_{HA} and ΔG_{ap}

Both $\Delta\Delta G_0$ and ΔG_{ap} values were calculated for model protein **I**:(GVGVP GVGVP GEGVP GVGVP GVGVP GVGVP)$_{36}$(GVGVP), which is abbreviated as E/0F. In this structure, one Glu (E) residue repeats every 30 residues, that is, one GEGVP repeats after every six pentamers. Taking 0.6 as the hydrophobic-induced pKa shift for Glu, both $\Delta\Delta G_0$ and ΔG_{ap} gave values of 0.85 kcal/mol-Glu (see Sections 2.4.2 and 2.4.3). For comparison with the calculation of $\Delta G_{HA}[(GE^{\circ}GVP) - (GE^{-}GVP)]$ in Section 2.4.1, which used the values in the ΔpKa-derived values and the ***n***-derived values need to be multiplied by a factor of 6, i.e., $6 \times 0.85 = 5.1$ kcal/mol. The value calculated using the change in Gibbs free energy for hydrophobic association using Eq. (4) in Section 2.3.2 gave -5.22 kcal/mol-(GEGVP). While the approximation for $\Delta\Delta G_0$ is inadequate to describe the energetics for a higher degree of competition for hydration, the thermodynamic quantities of $\mathbf{\Delta G_{HA}}$ and ΔG_{ap} appear to be quite adequate for consideration in the design of protein-based nanomachines.

Thus, two entirely different experimental approaches (differential scanning calorimetry to determine the temperature and heat of the transition and acid–base titrations to determine pKa) for the calculation of the apolar–polar repulsive Gibbs free energy of hydration result in the same statement of repulsive interaction between hydrophobic and carboxylate groups.

2.5 Calculations of the Entropic Elastic Force and Energy

An understanding of energy conversion by elastic-contractile model proteins, and equivalently by biology's protein-based nanomachines, requires insight into the relationship of energy input and output and temporary

storage of energy in the stretching of associated chain segments. Earlier, the importance of reversible or entropic elastic force development to efficient energy conversion was noted. The intention here is to make that relationship more tangible. The route is to calculate the change in entropy on extension, to use that entropy change in the calculation of entropic elastic force and to utilize force, acting through distance, to obtain energy involved in the performance of mechanical work.

2.5.1 Relation of elastic force and the energy in the performance of mechanical work

As noted earlier, the area under a force–extension curve measures the energy expended in achieving extension. During the function of natural protein nanomachines, energy derived from a favorable hydrophobic association or from an apolar–polar repulsion becomes stored in elastic deformation. It is, therefore, useful to gain insight into (1) the relationship between the elastic mechanism of the damping of internal chain dynamics on chain extension or the apolar–polar repulsion of the chain, (2) the development of elastic force, and (3) the energy stored on deformation.

2.5.2 Calculation of entropy change due to decreased amplitude of backbone torsional oscillations on extension

Calculation of the entropy change resulting from the decrease in amplitude of backbone torsion angles can utilize molecular dynamics calculations to obtain the root-mean-square (rms) amplitudes of the protein and backbone torsion angles, $\Delta\phi$ and $\Delta\psi$, respectively, for two states of extension, e.g., relaxed (r) and extended (e). Using the Boltzmann's relation of Section 1.5.2, $S = R \ln W$, where S is per mole because R is in units of kN, and the product of the rms values provides a measure of W to give entropy. Furthermore, the ratio of the products for the extended state divided by those of the relaxed state gives the change in entropy, ΔS, i.e.,

$$\Delta S = R \ln \left(\frac{\Pi_i \Delta\phi_i^e \Delta\psi_i^e}{\Pi_i \Delta\phi_i^r \Delta\psi_i^r} \right) \tag{6}$$

Calculating for $(GVGVP)_n$, the product, Π_i, signifies multiplication of all $\Delta\phi_i \Delta\psi_i$ for the $5n$ residues of the repeating pentapeptide segment under consideration, that is i runs from 1 to $5n$. With two angles for each residue this gives a $10n$-dimensional volume (space) per pentamer. Since the rms amplitudes of the torsion angle oscillations decrease on extension, the product, $\Pi_i \Delta\phi_i^e \Delta\psi_i^e$, represents a smaller $10n$-dimensional volume than represented by $\Pi_i \Delta\phi_i^r \Delta\psi_i^r$. This gives a decrease in entropy on extension. Using the CHARMM molecular

dynamics program of Karplus, the change in entropy per residue for a 130% extension, i.e., for a pentamer extension from 0.35 to 0.8 nm, is calculated to be -5.5 cal/mol-deg-pentamer.[29] Identical values result using the molecular mechanics approach (the ECEPP/2 program) of Scheraga.[8]

2.5.3 Use of entropy change in calculating entropic elastic force[2,23]

The conversion of a change in entropy, ΔS or ∂S, to an entropic elastic force, f_S, can occur by the following simple statement,

$$f_S = -\mathrm{T}\left(\frac{\partial S}{\partial L}\right)_{VT}. \tag{7}$$

As ΔS is inherently negative, i.e., -5.5 cal/mol-deg-pentamer, the entropic elastic force is positive. Avogadro's number for model protein chains per mole is 6.023×10^{23}; 4.184 J = 1 cal; the change in length, ∂L, for a single pentamer was from 0.35 to 0.80 nm, and the physiological temperature T is taken to be 310°K. As the force in Newtons has units of Joules/meter, J/m, these numbers translate into a force of 26 pN per model protein chain.

2.5.4 The area under a force–extension curve gives energy

Calculation of energy arising from an extension causing an entropic elastic force. For the simple case of a linear force–extension curve, the energy stored in a single model protein chain for the extension of 11 pentamers (55 residues) by 130% would be a change of length, ΔL, of about 5×10^{-9} m (i.e., 11 pentamers $\times$ 0.45 nm) to give a product of 13×10^{-20} J/chain.

This represents the $f\Delta L$ rectangular area, but for a linear force–extension curve, the work performed is given by the area under the curve, which would be one-half of the rectangular area, i.e., 6.5×10^{-20} J for a 55-residue chain. For a mole of such chains the stored energy of extension would be 3.9×10^4 J/mol or 9.4 kcal/mol of 55-residue chains. This would be about 0.2 kcal/mol-residue, and it represents an estimate of the *energy stored per residue* for a purely entropic deformation on extension of 55 residues of the $(\mathrm{GVGVP})_{11}$ model protein, which represents a soft chain structure.

By the coupled hydrophobic and elastic consilient mechanisms, part of the energy obtained on hydrophobic association is stored in the elastic extension of chains connecting sites of hydrophobic association. One may also assume that there would be an internal energy component to the force of about the same magnitude. Of course, the energy stored due to deformation would need to be less than the Gibbs free energy available from hydrophobic association. Therefore, the deformation energy of a single chain segment is expected to be not more than a few kcal/mol of extended chain segment.

Relevance to entropic elastic force developed in single chains by hydrophobic associations in Complex III and the myosin II motor. These considerations are applied later to the extensions attending chains in the Complex III and myosin II protein nanomachines. In the case of Complex III, hydrophobic association at the Q_O-site of the hydrophobic tip of the RIP globular domain extends a single chain of 11 residues, A66 to K77. This single chain continues on to be hydrophobically anchored in the cell membrane. In the example of the myosin II motor, it is a chain that becomes stretched as two hydrophobic associations occur, one at the actin binding site and the other at the head of the lever arm and the N-terminal domain of the cross-bridge. The myosin hydrophobic associations occur on going from an analog of the ATP bound state to a representation of the rigor state.

3. BIOLOGY'S PROTEIN-BASED NANOMACHINES CONFIRM THE HYDROPHOBIC AND ELASTIC CONSILIENT MECHANISMS

3.1 The Three Classes of Energy Conversion Within the Cell

Energy conversions that occur within the cell may be considered in terms of three primary classes: (1) the formation of a proton concentration gradient, (2) the use of the proton concentration gradient to produce ATP, and (3) the use of ATP to perform the essential work of the cell. Two of these occur within the mitochondria, known as the energy factory of the cell. Energy conversions of the first class occur in the crucial electron transport chain that resides in the inner mitochondrial membrane. These energy conversions may be classified as *electro-chemical transduction* as the energy obtained from the electron transfers of oxidation–reduction reactions is used to develop a proton concentration gradient across the inner mitochondrial membrane. The energy conversion of the second class within the mitochondrion may be classified as *chemo-chemical transduction*. It is in this class where the chemical energy of the proton concentration gradient is used to produce ATP from ADP and P_i. The third class of energy conversions involves the many uses of ATP in performing the work required to sustain the organism. A prominent use of ATP is in producing motion as in the myosin II motor of muscle contraction. This represents perhaps the most classical example of *chemo-mechanical transduction*.

A more complete treatment of these three natural nanomachines is in press as part of a book on the hydrophobic and elastic consilient mechanisms in protein-based machines and materials,[10] and a separate publication on the F_1-motor (F_1-ATPase) of ATP synthase has recently appeared.[22] Some of the analyses, however, have been extended in this report.

Even though the *input* energy and the *output* work of a natural nanomachine may not include motion, many energy conversions within their mechanism do utilize motion; intermediate steps do involve the performance of mechanical work. In the process of achieving electro-chemical transduction, Complex III of the electron transport chain utilizes movement of a key domain to achieve electron transfer, that is, there is an element of electro-mechanical transduction within the mechanism of Complex III function. Also ATP synthase, which makes nearly 90% of the cell's ATP and where the overall process is one of chemo-chemical transduction, the mechanism breaks down into two clearly physically separable energy conversions, chemo-mechanical transduction and mechano-chemical transduction. *Thus, motion is a common component of many nanomachines of the cell, and commonly the capacity to achieve motion involves an elastic deformation*.

3.2 Complex III of the Electron Transport Chain Within the Inner Mitochondrial Membrane

3.2.1 Electro-chemical transduction

The coupling of electron transfer to proton translocation across the inner mitochondrial membrane. Four complexes function to transfer electrons from reduced nicotinamide to oxygen. Complexes I, III, and IV each pump four protons across the membrane during the oxidation of each reduced nicotinamide, i.e., for every two electrons passed through each complex. Complex II supplies additional reduced ubiquinone, i.e., ubiquinol, to the lipid bilayer pool of the inner mitochondrial membrane to be used by Complex III. Crystal structure data is available for both complexes III and IV. The crystal structures available for Complex III (cytochrome bc_1), however, allow substantiation of the consilient mechanisms in this natural nanomachine in which the coupling of hydrophobic and elastic consilient mechanisms achieves coupling of electron flow to proton transport.

3.2.2 Structure of complex III

A homodimer with each monomer composed of three redox-containing protein subunits. As shown in stereo view in Fig. 1A, the crystal structure of Complex III, due to Lange and Hunte,[34] exhibits twofold symmetry with an interesting cross-over of the membrane anchor of the Rieske Iron Protein (RIP). While the anchor of RIP occurs in one monomer, the globular domain, containing the FeS center, functions in electron transfer in the other monomer. This is shown in Fig. 1B as a crisscrossing of freestanding tether

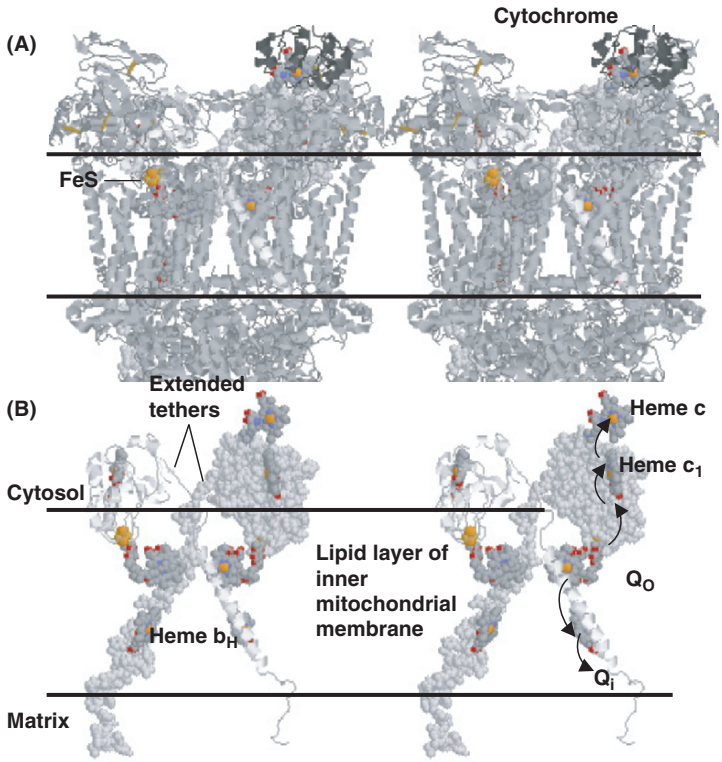


Figure 1. (A) Stereo view of homo-dimeric structure (gray) of yeast Complex III, plus one molecule of cytochrome c (dark gray at upper right), but with symmetry related single chains of the white RIP domain in space-filling representation (rising left to right) and in a ribbon representation (rising right to left).[34] The result is a crossing of the globular component by means of the single chains into the opposite monomers where they elastically move electrons from the Q_O-site to the heme c_1 site. (B) Stereo view of RIP and redox centers showing electron transfer with the Q_O-site for the ubiquinol oxidation site and Q_i-site for the ubiquinone reduction, showing extended tethers that subsequently contract to move the FeS center to the heme c_1 on positively charging Q_O-site to disrupt hydrophobic association by ΔG_{ap}. Crystal structure due to Lange and Hunte.[34] Protein Data Bank, Structure File, 1KYO. Adapted from reference 10.

segments of RIP, apparent in stereo view, one in space filling and the other in ribbon representation.

While the monomer of yeast Complex III contains 11 protein subunits, only three subunits contain redox centers. The redox-containing protein subunits are cytochrome b with two hemes, b_L and b_H, cytochrome c_1 with a single heme c_1, and the RIP with the FeS, iron sulfide, redox center that transfers one electron from ubiquinol at the Q_O-site of cytochrome b on the cytosolic or cytoplasmic (intermembrane) side of the inner mitochondrial

membrane to the heme of cytochrome c_1, also on the cytoplasmic (cytosolic) side. Figure 1B shows a cross-eye stereo view of the electron flow. It should also be noted that some of the protein of Complex III on the matrix side of the membrane has been cut off in Fig. 1A.

3.2.3 The RIP as a key protein subunit of electron transfer

The globular domain of the RIP undergoes movement.[35] The RIP is a single protein chain with a globular FeS-containing domain and a single chain extending from the globular component by means of a flexible tether to a hydrophobic α-helical segment that anchors RIP to the inner mitochondrial membrane within the opposite monomer from which RIP transfers its electron. The globular domain of RIP is shaped like a stylus with a very hydrophobic tip containing the FeS center that forms its most favorable hydrophobic association at the cup-shaped ubiquinol-occupied Q_O-site. In order to form this energetically most favorable hydrophobic association, however, the single chain tether becomes extended.

3.2.4 Mechanism as deduced from crystal structures of two key states

Two crystal states available, one with the FeS center of the RIP component of Complex III at the Q_O-site and a second with FeS at the heme c_1 site. From the crystal structure data of Zhang *et al.*[35] using Complex III from chicken in the presence and absence of the inhibitors, stigmatellin and antimycin, the FeS tip of the globular domain of RIP is found at the Q_O-site in the presence of inhibitor and at the heme c_1-site in the absence of inhibitor. These two states of the RIP subunit of Complex III are shown in Fig. 2.

Extension of the tether of RIP when the FeS center is at the Q_O-site. When the FeS tip of the globular domain of RIP is found at the Q_O-site, the tether connecting the globular domain to its membrane anchor is extended, whereas when the FeS tip of the globular domain is at the heme c_1-site, the tether connecting the globular domain to its membrane anchor is relaxed. As described in detail elsewhere using eight developed structural perspectives,[10] crystal structure data of the two states,[34,35] presented in part in Figs. 1 and 2, demonstrate the basis for the relevance of the hydrophobic and elastic consilient mechanisms. These are briefly noted in the following.

The hydrophobic consilient mechanism provides for release of the FeS tip from the Q_O-site. As depicted in Fig. 1B, hydrophobic ubiquinol, which underlies the Q_O-site, gives up one electron to the FeS center, and immediately following ubiquinol^{+1} gives up a second electron to heme b_L. This results in a single ubiquinol bearing two positive charges at the Q_O-site. By the apolar–polar repulsive Gibbs free energy of hydration, ΔG_{ap}, of the hydrophobic

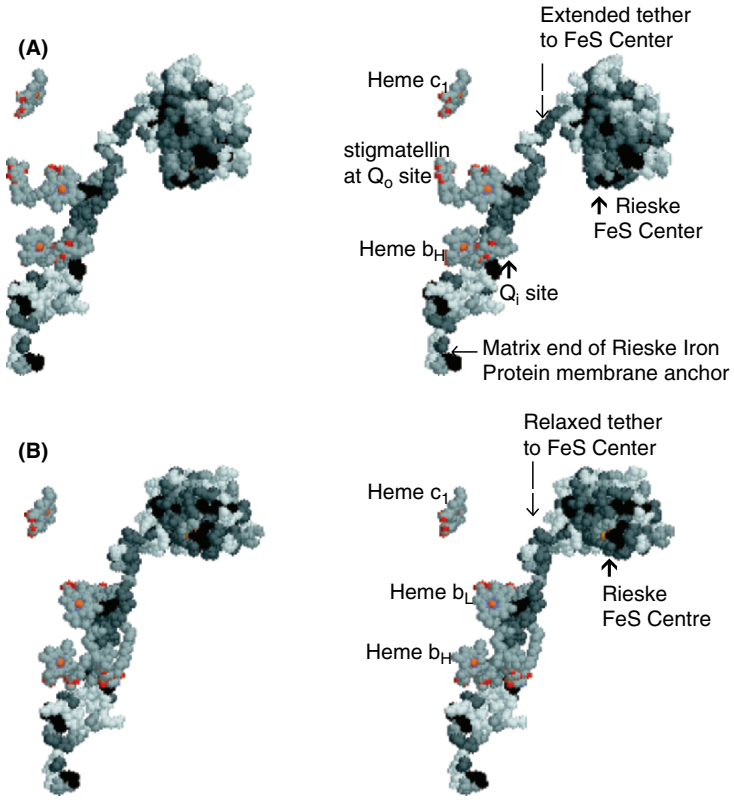


Figure 2. Stereo view of Complex III RIP of monomers for chicken[35] with the amino acid residues gray-coded to denote hydrophobicity: charged residues, white; neutrals, light gray; aliphatics, gray; and the most hydrophobic aromatics, black. The globular domain is very hydrophobic at its tip, the location of the FeS center. Shown also are the ligands. (A) With the stigmatellin and antimycin inhibitors, the FeS center of RIP is directed downward in position at the Q_O-site to be reduced by second monomer. Protein Data Bank, Structure File 3BCC. (B) Without inhibitors, the FeS center is directed out of the plane in position to reduce cytochrome c_1 of second monomer. Protein Data Bank, Structure File 1BCC. Reproduced with permission from reference 10.

consilient mechanism, the two positive charges cause $\mathbf{\Delta G_{HA}}$ to become less negative, i.e., less favorable, and, therefore, susceptible to hydrophobic dissociation, especially with the elastic retraction pulling toward dissociation. But can we estimate what energy would be stored on extension of the tether?

The elastic consilient mechanism explains the FeS tip relocation to the heme c_1-site. With the formation of two positive charges at the ubiquinol bound Q_O-site, the change in Gibbs free energy for hydrophobic association, $\mathbf{\Delta G_{HA}}$, is no longer sufficient to sustain the stretched tether, and the result is

hydrophobic dissociation. On retraction the stretched single chain tether lifts the reduced FeS-containing globular domain of RIP and brings it to rest in a secondary hydrophobic association at the heme c_1 site where it is oxidized. Thus, RIP domain movement transfers an electron from the ubiquinol at the Q_O-site to heme c_1.[35] This is mechano-electrical transduction.

A ballpark estimate of the deformation energy stored in the extended tether can be obtained using the analysis of Section 2.5.4. In particular, the extension of the A66 to K77 sequence of the tether appears to be about half of the 130% extension used in the calculation of $(GVGVP)_{11}$. The extension of 11 residues with the same degree of damping of the 11 pairs of ϕ_i and ψ_i backbone torsion angles as occurred with the elastic-contractile model protein would give an energy of deformation for 11 residues of 11×0.2 kcal/mol-residue $=$ 2.2 kcal/mol-tether. The percentage of the extension, however, is about one-half of that used in the $(GVGVP)_{11}$ calculation. Accordingly, the energy stored in the extended tether would be approximately 1 kcal/mol of tethers. This amount of entropic deformation energy and an equivalent amount due to elastic force arising out of a change in internal energy result in an energy of a reasonable magnitude of about 2 kcal/mol-tether. The change in Gibbs free energy for hydrophobic association, ΔG_{HA}, on association of the hydrophobic tip of the FeS-containing globular component of RIP with the hydrophobic Q_O-site would be several times that amount. Therefore, the energies involved in the coupling of hydrophobic and elastic consilient mechanisms are appropriate. With a less favorable $\mathbf{\Delta G_{HA}}$ after oxidation of the ubiquinol, the energy of retraction would be of a magnitude sufficient to effect relocation of the FeS center.

As the RIP domain leaves the Q_O-site, the oxidized ubiquinol, bearing two positive charges, releases two protons to the inner membrane space to become ubiquinone. Once the globular domain of RIP vacates the Q_O-site on the cytoplasmic side of the inner mitochondrial membrane with replacement by water that is not hydrophobic hydration, the oxidized ubiquinol releases its two positive charges in the form of two protons to the mitochondrial intermembrane space. *Thus, proton translocation couples to electron transport because hydrophobic association/dissociation couples to extension/relaxation*.

The second electron from the first ubiquinol and from the second ubiquinol pass from heme b_L to heme b_H and on to ubiquinone at the Q_i-site. On receipt of two electrons, one from each of two ubiquinols at the Q_O-site, ubiquinone at the Q_i-site bears two negative charges. Again by the apolar–polar repulsive Gibbs free energy of hydration, ΔG_{ap}, of the hydrophobic consilient mechanism, the two negative charges of the reduced ubiquinone open incipient aqueous channels on the matrix side of the inner mitochondrial membrane and the ubiquinone with two negative charges picks up two protons from the matrix side to become ubiquinol. With a ubiquinol from Complex II, this

completes coupling of transfer of two electrons from ubiquinol to two cytochrome c_1 molecules to the transport of four protons across the membrane.

3.3 ATP Synthase of the Inner Mitochondrial Membrane

This discussion of ATP synthase will be short. Instead of the 17 figures and many references in the more extended treatment,[10] only a single figure and two crystal structure references are considered here.[36,37] A communication in Cell Biology International provides the initial disclosure.[22] The principal reason for inclusion here is to argue for the central role of apolar–polar repulsive free energy of hydration, ΔG_{ap}, in the synthesis of ATP from ADP plus inorganic phosphate, P_i, which is relevant to the design of a nanomachine. By clockwise rotation of a rotor, the most hydrophobic face of the rotor is forced into opposition with the most polar ADP plus P_i state, and the high level of repulsion is lowered by formation of the less polar ATP state. In the ATPase function, ATP hydrolysis does exactly the reverse; formation of the most polar ADP plus Pi state repulses the most hydrophobic side of the rotor in such a way as to drive rotation of the γ-rotor in the counterclockwise direction.

3.3.1 Structure of ATP synthase

The structure of ATP synthase can be described as having the following four main protein components:[36] (1) An intramembrane component with 10 hairpin α-helical subunits arranged perpendicular to the membrane, like the tread of a wheel that functions as a rotary motor (the F_0-motor) with its rotation driven by proton return from cytosol to the matrix side of the inner mitochondrial membrane, (2) an axle (called the γ-rotor) that extends from the F_0-motor in an orientation perpendicular to the plane of the membrane and directed outward from the matrix side of the inner mitochondrial membrane, (3) an orange-shaped structure with six sections arranged as three αβ sections that form a housing with threefold symmetry and into which the γ-rotor inserts to form the F_1-motor, and (4) a stator component that spans from the base of the housing of the F_1-motor to the side of the F_0-motor where it contains the channels for proton influx and efflux and keeps the housing of the F_1-motor from rotating as the γ-rotor rotates.

3.3.2 Function of the F_0-motor of ATP synthase

The 10 hairpin α-helical subunits of the F_0-motor have hydrophobic sequences that span the lipid layer of the inner mitochondrial membrane with the exception of a single carboxyl-containing amino acid residue, e.g., an aspartic acid (Asp, D) residue approximately half way across the lipid

layer, but which when ionized can be separated from it by a channel-containing component of the stator. A proton from the cytoplasmic side of the membrane enters the channel, protonates the carboxylate, and thereby, lowers $\mathbf{\Delta G_{HA}}$ for rotating the now carboxyl-containing helical subunit into the lipid layer. As this occurs, a carboxyl rotated from the lipid layer releases its proton by means of a channel to the matrix side of the inner mitochondrial membrane. Thus, protonation of the carboxylate lowers the Gibbs free energy for hydrophobic association and favors rotation into the lipid layer, and the carboxyl-containing helical segment that rotates into the efflux channel position gives up its proton. The side of the membrane with the highest chemical potential of proton dictates the direction of rotation. Accordingly, the return of 10 protons across the lipid layer from the cytoplasmic to the matrix side of the mitochondrial membrane drives one complete clockwise rotation of the γ-rotor within the F_0-motor.

3.3.3 The F_1-motor (F_1-ATPase) of ATP synthase

The hydrophobic sidedness of the γ-rotor within the (αβ) housing of the F_1-motor.[10,22,36] If the hydrophobic consilient mechanism is to be relevant to the function of ATP synthase, the γ-rotor would need to be asymmetrically hydrophobic. The expectation, based on the apolar–polar repulsive free energy of hydration, is that the most hydrophobic side of the rotor would be forced by F_0-motor rotation to face-off through water with the most polar state, ADP plus P_i. The mechanically forced state of high elastic repulsion would be relieved by formation of ATP. The state of high elastic repulsion between rotor and housing for the β-subunit containing the ADP plus sulfate (SO_4^{-2}) analog of the ADP plus P_i state is indeed observed in the crystal structure data as a physical distancing of the γ-rotor from the housing for the β-subunit containing the ADP plus sulfate analog for phosphate.[22,36]

Shown in stereo view in Fig. 3 below is a vertical cross-section of the F_1-motor with the β-ADP-SO_4^{-2} subunit in space-filling representation directly opposed to the most hydrophobic side of the γ-rotor in ribbon representation. The very polar, doubly charged SO_4^{-2} faces-off through a water-filled cavern against the hydrophobic side of the γ-rotor. The arc of repulsion, indicated by the heavy lines, is directed against the α-helical segment of the γ-rotor that ends at the base of the central aqueous cavern. Looking at the structure in Fig. 3 in cross-eye stereo view allows recognition of the steric directing effects of the repulsion. The finer lines are due to charged amino acid side chains, a number of which became unpaired in a positively cooperative manner with the appearance of sulfate to achieve their own hydration and increase the repulsion. With the primary repulsion directed at the short helix, the repulsion drives the rotation of the γ-rotor in a counterclockwise direction, as observed when functioning as an ATPase.[38] As elaborated in detail elsewhere,[10,22] the

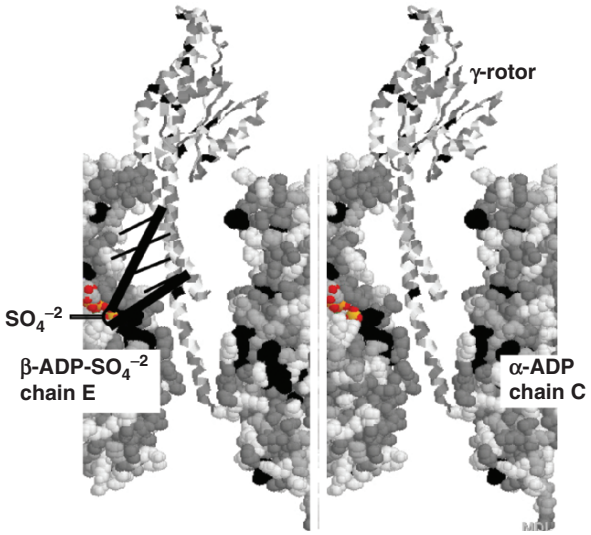


Figure 3. Cross-eye stereo view of a cross-section of the F_1-motor of ATP synthase with the γ-rotor in ribbon representation and chain E (the β_E-subunit) and chain C (an α-subunit) in space-filling representation and with neutral residues in light gray, aliphatic residues in gray, aromatic residues in black and charged residues in white. The lines indicate a repulsive interaction between the hydrophobic side of the γ-rotor and SO_4^{-2} (as the analog for hydrolyzed γ-phosphate) that is augmented by newly emerged charged groups occurring in the β_E chain. The ΔG_{ap} due to SO_4^{-2} and due to the charged groups that emerge on the appearance of sulfate applies a torque to the short α-helical portion of the double-stranded γ-rotor. The result of the repulsion would be a counterclockwise rotation when functioning as an ATPase. Developed from crystal structure data of Menz *et al.*[37] Protein Data Bank, Structure File 1H8E. Adapted from reference 22.

hydrophobic and elastic consilient mechanisms, given the general structure of ATP synthase, allow some eight predictions that appear to be borne out by analysis of the details of the crystal structure data.[37]

3.4 The Myosin II Motor of Muscle Contraction

More detailed analysis of the function of the myosin II motor of muscle contraction is given elsewhere,[10] where the approach is as for Complex III above. Hydrophobic associations develop due to decreased polar state of the protein nanomachine and provide the impulse for motion by developing elastic force in single protein chain segments. Hydrophobic associations develop (on ATP hydrolysis and loss of phosphate) between cross-bridge and the actin filament binding site and between the head of the lever arm and the underside of the N-terminal domain within the cross-bridge. These hydrophobic associations stretch interconnecting chain segments.

3.4.1 Narrow Clefts that Direct Apolar–Polar Repulsion from Nucleotide Binding Site to Disrupt Hydrophobic Associations

Figure 4 demonstrates narrow clefts in the crystal structure of the scallop muscle cross-bridge in near rigor and ATP analog states.[39] The clefts emanate in different directions from the location of the γ-phosphate. In your author's view, one cleft directs apolar–polar repulsion due to ATP, to disrupt association of cross-bridge with the actin filament. The second cleft, partially obscured from the view in Fig. 4 but more exposed by the perspective in Fig. 5, is directed toward the head of the lever arm. One possibility is that on

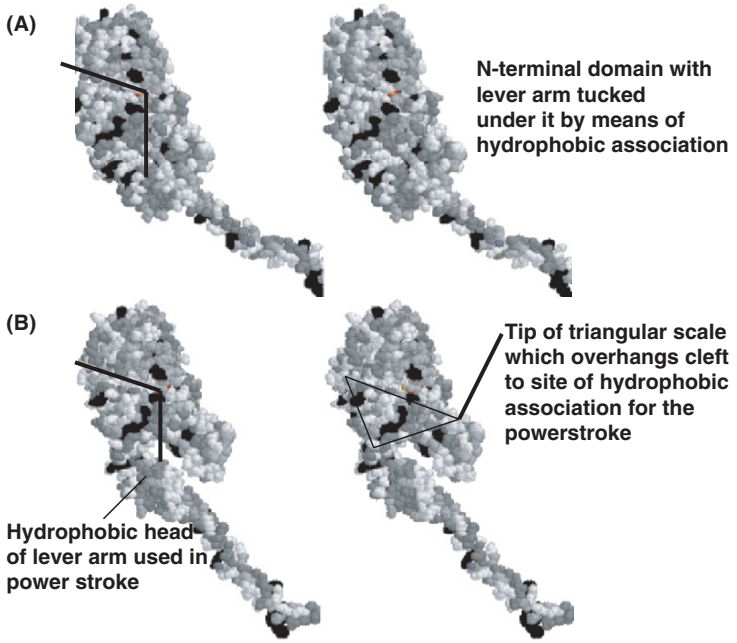


Figure 4. Stereo (cross-eye) view of the scallop muscle cross-bridge (S1) taken from the actin filament toward the actin binding site on the myosin cross-bridge. (A) The near rigor state, Protein Data Bank, Structure File 1KK7. (B) The ATP analog state, Protein Data Bank, Structure File 1KK8.[39] Two narrow clefts are indicated, approximated by the two lines meeting at the location of the γ-phosphate. The first cleft directs the apolar–polar repulsive force of the ATP to disrupt hydrophobic association of cross-bridge to actin. ATP hydrolysis to form ADP plus P_i, with a slight shift of the phosphate location, causes the phosphate to lose its line of sight by means of the first cleft to the actin-binding site and allows reattachment to the actin filament. In this slightly shifted position, the phosphate now directs its repulsion to the head of the lever arm and disrupts hydrophobic association of the head of the lever arm with the underside of the N-terminal domain. Release of the phosphate would result in the powerstroke due to hydrophobic reassociation of head of the lever arm with the N-terminal domain. From reference 10.

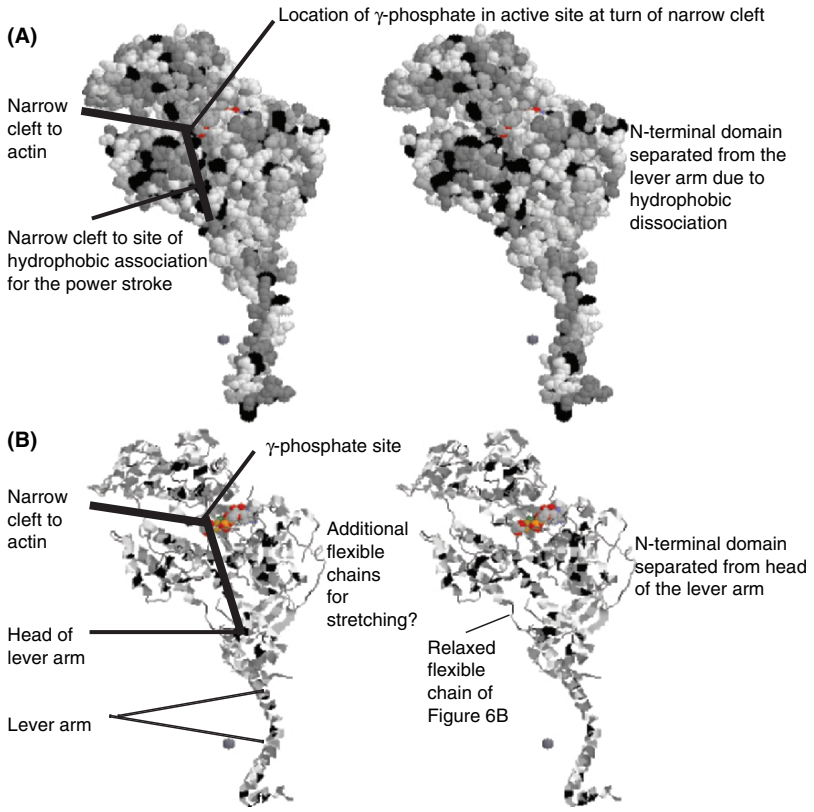


Figure 5. Stereo view (cross-eye) of scallop muscle cross-bridge (S1) for the ATP analog state to show the ATP binding site with phosphate at the junction of two clefts. One cleft runs toward the actin–cross-bridge binding site and the other toward the hydrophobic dissociation between head of lever arm and N-terminal domain. (A) Space-filling representation and (B) ribbon representation, Protein Data Bank, Structure File 1KK8 due to Himmel *et al.*[39] Adapted from reference 10.

ATP hydrolysis to form the ADP plus P_i state, phosphate shifts its location slightly, loses site by the first cleft to the actin binding site, and allows hydrophobic reassociation of the cross-bridge with the actin filament. With the shift, the phosphate now directs its apolar–polar repulsion to disrupt hydrophobic association of the head of the lever arm with the underside of the N-terminal domain.

Rayment and colleagues recognized very early the importance of clefts. A decade ago in a most insightful study, Rayment and coworkers[40] noted the presence of a narrow cleft and its structural relationship to the ATP binding site in the S1-fragment, by the following statement:

The current structural results emphasize the importance of the narrow cleft that splits the 50-kDa segment in the molecular origin of myosin based motility. These suggest further that it functions not only in sensing the presence of the γ-phosphate of ATP but is also responsible for transducing the conformational change that results in the power stroke.[40]

The apolar–polar repulsive free energy of hydration, ΔG_{ap}, developed in Sections 1 and 2 and demonstrated with Complex III and ATP synthase, provides in this Section 3, the added insight that could readily have brought Rayment and coworkers to the mechanism proposed here.

3.4.2 Extension of dynamic single chains on forming near rigor state

Another structural feature noted by the Rayment group was the presence of a number of flexible loops.[41] One such flexible loop is the focus of Fig. 6.

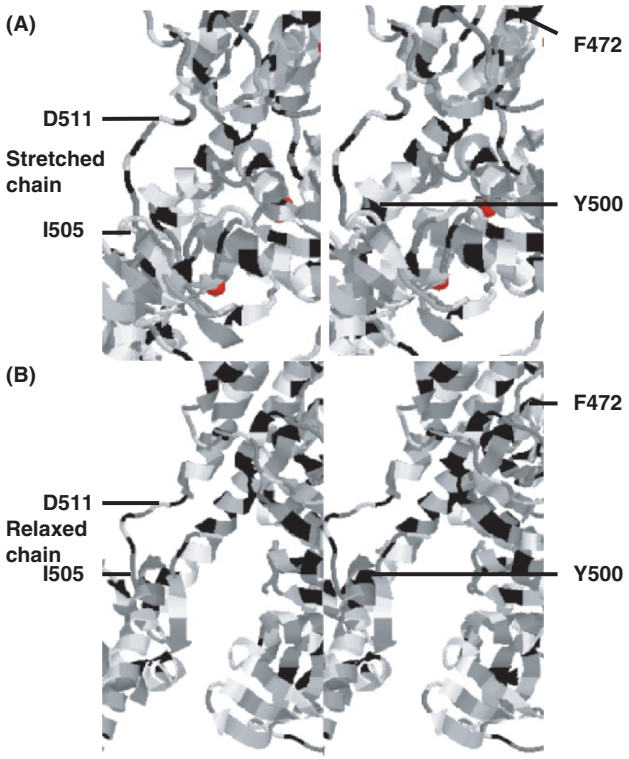


Figure 6. Expanded stereo (cross-eye) views of scallop muscle cross-bridge (S1)[39] in ribbon representation for the purpose of showing the length change in a portion of the relay loop. The ratio of the length of I505-D511 loop segment to that of the Y500-F474 segment of the relay helix indicates as much as a 40% extension on loss of phosphate. (A) Near rigor state. Protein Data Bank, Structure File 1KK7. (B) ATP state using ATP analog. Protein Data Bank, Structure File 1KK8. Adapted from reference 10.

The indicated protein chain segment, L505-D511, is relaxed in the polar state with disrupted hydrophobic association, but is extended in the near rigor state with the reformation of the above pair of noted hydrophobic associations. Using the relay helix as a measuring stick, there is about a 40% extension in the near rigor state of the L505-D511 sequence as compared to the ATP analog state. It is to be expected that a number of such flexible loops would be stretched on isometric contraction to achieve the experimental forces. Additional flexible loops are seen on the right-hand side of Fig. 5B. These loops and many more features require further examination.

3.5 Confirmation of the Hydrophobic and Elastic Consilient Mechanisms

The two fundamental hydrophobic and elastic consilient mechanisms overlay the usual molecular mechanics and dynamics treatments of structure/function to provide the primary energetic foundations that give rise to protein function. These mechanisms derive from designed elastic-contractile model proteins capable of ITTs of hydrophobic association leading to elastic contractions on raising the temperature or on lowering the temperature of the transition, for example, by decreasing charge. Importantly, they have been reasonably confirmed to be operative in the above-considered natural nanomachines. In particular, with the Complex III and myosin II motor examples, favored hydrophobic associations cause chain extension with elastic force development, and it is your author's view that the combination of these hydrophobic and elastic mechanisms results in function.

To recap, the mechanisms are the change in Gibbs free energy for hydrophobic association, $\mathbf{\Delta G_{HA}}$, and elastic force development resulting from the damping of backbone dynamics on extension of even a single chain. In analogy to the old adage "oil and vinegar do not mix," a change in $\mathbf{\Delta G_{HA}}$ arises from a change in competition for hydration between oil-like (hydrophobic) and vinegar-like (e.g., charged) groups constrained by primary, secondary, tertiary, and quaternary structure to coexist and thereby forced to interact in a protein chain. The result is termed an apolar–polar repulsive free energy hydration, ΔG_{ap}. Formation of charge, on recruiting its own hydration, disrupts hydrophobic association by destroying hydrophobic hydration that forms during a transient hydrophobic dissociation. The potential for formation of too much hydrophobic hydration as occurs on loss of a competing polar (e.g., charged) species results in hydrophobic association (insolubility).

4. DESIGNING PROTEIN-BASED NANOMACHINES USING THE HYDROPHOBIC AND ELASTIC CONSILIENT MECHANISMS

Design of a nanomachine as a biosensor for chemical detection. A nanomachine that gives a mechanical output to a chemical energy input could be called a chemo-mechanical transducer. In the examples under development below, the chemical input is phosphorylation and the mechanical output will be either the change in force or the change in the acoustic absorption intensity of a mechanical resonance. The change in force, measured by AFM in the single-chain force–extension mode, is to occur on phosphorylation-induced hydrophobic unfolding of a single globular domain. Plans are to measure the change in acoustic absorption intensity by employing a vibrating cantilever in a loss shear modulus type of determination as a function of frequency. A device for measuring changes in intensity of mechanical resonances has been designed by Vlado Hlady.[42,43]

4.1 Design of an AFM-Based Stress–Strain Nanomachine for the Detection of a Single Molecular Event

4.1.1 Single-chain force–extension curve for a near ideal elastic element

The smooth monotonic stress–strain curve for elastic model proteins. The single-chain force–extension curve for the family of elastic-contractile model proteins based on $(GVGVP)_n$ is a reversible smooth curve that exhibits a monotonic increase in force with an increase in extension.[23] This provides the appropriate background curve on which to superimpose a sensing element.

4.1.2 Single-chain force–extension curve for a series of globular elements

Series of repeating globular subunits yields a series of sawtooths superimposed on the smooth monotonic curve of an elastic chain. The intracellular megaprotein, titin, and the extracellular protein, fibronectin, contain long sequences that are protein-based polymers composed of repeating 90–100 residue globular protein domains. When a single strand of such sequences is examined by AFM, the force–extension curve is found to be a series of sawtooths, one for every (hydrophobically folded) globular domain. Each globular domain yields a characteristic sawtooth that can be analyzed as to the peak force reached and the number of residues in the globular domain. On relaxation only a smooth underlying elastic retraction is observed. Thus, a

series of sawtooths are observed that superimpose on a smooth monotonic curve of an elastic chain.

The hysteretic elastic elements of titin. It would seem that a function of titin in the sarcomere of the muscle cell would be to absorb energy in a series of packets so that the total force on chain extension never rises to the level of chain rupture. Thus, titin would appear to be a molecular shock absorber in which energy is consumed in the forced unfolding and chain extension of a single globular element. Titin dissipates the extending force into a set of energy packets too small to affect chain rupture. The particular interest here is that the characteristics of the sawtooth representation of globular protein unfolding report the energetics of the dominantly hydrophobic folding/unfolding process. To the extent that an interaction with the globular element alters the energetics of the process, it should be detectable as a change in the characteristics of the sawtooth such as a change in the peak force, or a change in the complexity of the sawtooth, or its obliteration.

4.1.3 Combining near ideal elastic element with hysteretic (sensing) elements

Design of the phosphate-sensing element. The design of the sensing globular element begins with the tenth Type III domain of fibronectin, Fn3. This 99-residue domain gives a characteristic sawtooth in the single-chain force–extension curve. The Fn3 domain contains the cell attachment sequence, GRGDSP, which will be changed to the cyclic AMP dependent protein kinase site, RGYSLG. This construct is indicated as Fn3′(GRGDSP → RGYSLG) wherein the serine (Ser, S) of RGYSLG becomes the site of phosphorylation. As Fn3 is a hydrophobically folded β-barrel, by the apolar–polar repulsive free energy of hydration of the hydrophobic consilient mechanism, the effect of phosphorylation would be to disrupt part or all of the hydrophobically folded domain.

Insertion of the globular sensing element into an elastic model protein chain. In order to have a single sawtooth on a long monotonically rising force–extension curve, the design becomes $[(GVGIP)_{10}]_{n1}$ – Fn3′(GRGDSP → RGYSLG) – $[(GVGIP)_{10}]_{n2}$. A successful gene construction, verified by base sequence analysis and agarose gels, was achieved and the resulting protein-based polymer, also verified by the NMR analysis, was expressed with $n1 = 16$ and $n2 = 6$.

Stress–strain curve for $[(GVGIP)_{10}]_{n1}$*-Fn3′(GRGDSP → RGYSLG)-*$[(GVGIP)_{10}]_{n2}$*, obtained as E. coli expressed protein*. The experimental force–extension curve, due to *E. coli* expressed $[(GVGIP)_{10}]_{16}$– Fn3′(GRGDSP → RGYSLG) – $[(GVGIP)_{10}]_6$ in Fig. 7A, exhibits a single sawtooth with the length expected for the 99-residue domain (Hugel, Seitz,

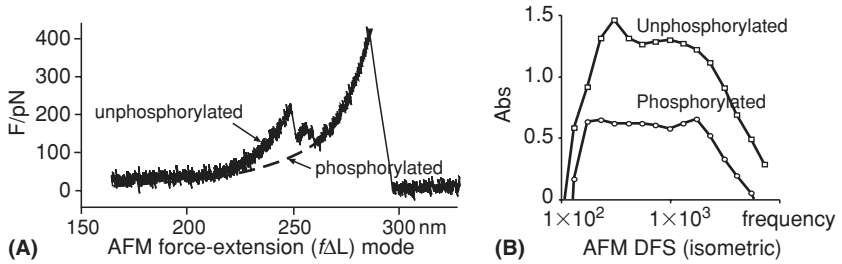


Figure 7. Proposed means whereby elastic protein-based polymers can be designed as nanomachines in combination with special adaptations of the atomic force microscope with the ultimate capacity to detect a single molecular interaction. (A) Single-molecule force–extension curve of a designed single globular protein detection element subtended by elastic contractile model protein chains. The experimental curve is without the molecular interaction of phosphorylation (Hugel, Seitz, Gaub, and Urry, unpublished data). The dashed curve would result when an interaction, such as phosphorylation, completely unfolded the hydrophobically folded globular protein domain. (B) Acoustic absorption of cross-linked (GVGIP) largely unfolded (lower curve) and hydrophobically folded (upper curve).[2,23] Such curves with the noted relationship is expected on phosphorylation when using AFM/DFS, an atomic force microscope modified for dynamic force spectroscopic measurements on a single twisted filament designed to contain one or more sites for enzymatic phosphorylation. A suitable AFM/DFS device is designed and constructed by Vladimir Hlady.

Gaub, and Urry, unpublished data). The sawtooth has some structure that could be the result of a two-stage unfolding process. The complete disruption of the hydrophobically folded barrel, as indicated by the dashed line, represents the most dramatic possible result of phosphorylation. Addition of SH-containing cysteinyl residues at the ends of the protein chain will provide stable attachment of the construct to the cantilever tip and the substrate. This will allow repeated force–extension curves with and without phosphorylation and determination of the stochastic nature of the sawtooth. Once the nature and extension at which the force change occurs are known, the chemical event can be monitored at a single extension without having to change length.

Experimental data showing phosphorylation to shift the temperature of the ITT. The effect of phosphorylation of $[(GVGIP)_{10}]_{16} - Fn3'(GRGDSP \rightarrow RGYSLG) - [(GVGIP)_{10}]_6$ on the temperature profiles for determining the onset of the inverse temperature is significant and interesting (Pattanaik, Hayes, Parker, and Urry, unpublished data). Either enzymatic phosphorylation goes to 50% or there is an initial micellization using one end (perhaps the $n = 16$ end) followed by complete aggregation and phase separation at a

higher temperature. The aggregation profile itself reports the occurrence of the singular molecular event of phosphorylation. In any event, the effect of the single bound phosphate is profound, and there is a good reason to expect that the sawtooth of the force–extension curve will be markedly altered by phosphorylation.

4.2 Use of the 3 kHz Mechanical Resonances in the Design of an AFM-Based Nanomachine for Detection of Interactions at Fixed Length

4.2.1 Formation of a triple-stranded filament

The 3 kHz mechanical resonance, measurable as acoustic absorption and a low frequency dielectric relaxation, builds as the temperature progresses into the range of the ITT. As the ITT progresses the elastic-contractile model protein hydrophobically folds and assembles. The fundamental repeating structural unit is thought to be a triple-stranded twisted filament, as schematically represented in Fig. 8C. Accordingly, in order to utilize the 3 kHz mechanical resonance in a nanomachine-based biosensor, the formation of the fundamental structural unit is required in the context of the AFM device of cantilever and substrate. Figure 8 outlines the approach, as described in the figure legend.

4.2.2 Anticipated change in the acoustic absorption profile

With the fundamental filament assembled as in Fig. 8 and with the device designed by Vladimir Hlady,[42,43] the absorption profile in the acoustic range is expected to resemble that of the acoustic absorption profile obtained for the bulk sample as previously reported.[2,23] As proposed in Fig. 7B, the binding of a polar analyte, such as phosphate, would raise the temperature for the onset of the ITT and thereby lower the intensity of the acoustic absorption.

The site for analyte binding can be positioned along the twisted filament, for example as the RGYSLG sequence, where the result of Fig. 7B would be most directly relevant, or the analyte binding site can be part of a globular element. Interestingly, the globular sensing element, Fn3′(GRGDSP → RGYSLG), is of a size and shape, based on the crystal structure of the 10th Type III domain of fibronectin,[44] that it would fit into the twisted filament in such a way as to be almost indistinguishable from native filament, as shown in Chapter 9 of Ref. [10].

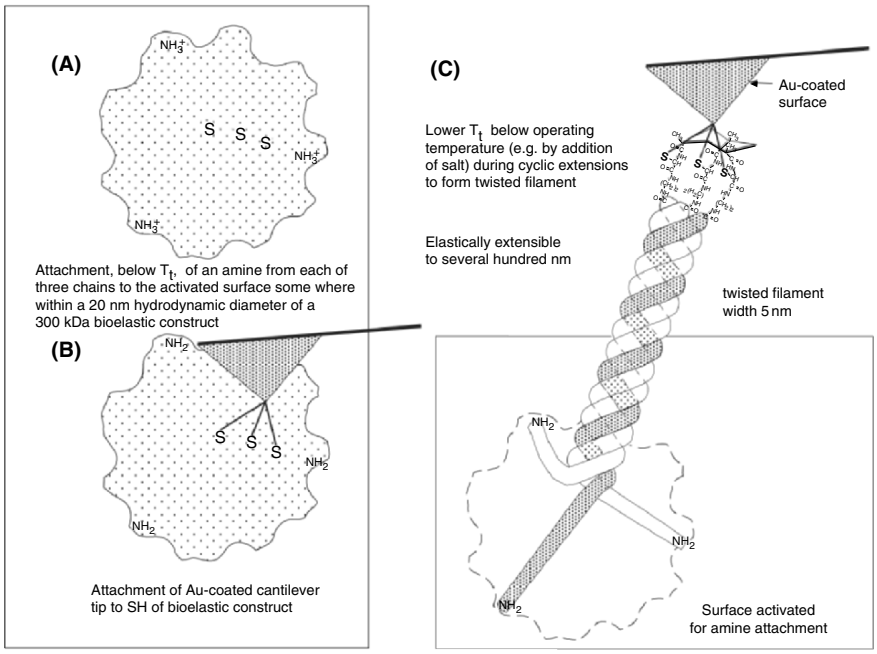


Figure 8. Twisted filament formation for AFM dynamic force spectroscopy. (A) Representation, below onset of the ITT, of the hydrodynamic sphere of three 100 kDa strands of an elastic-contractile model protein bound together at one terminus with sulfhydryls and attached by amino functions at the other terminus. (B) The gold-coated cantilever tip picks up the SH-containing termini. (C) Cyclic extension/relaxations form a twisted filament. At fixed partial extension, the twisted filament is expected to exhibit the dephosphorylated curve of Fig. 7B. On phosphorylation, the intensity of the peak is expected to drop due to hydrophobic dissociation.

4.3 An Additional Opportunity in the Deciphering of Engineering Principles for the Design of Protein-Based Nanomachines

Prediction of an elastic force due to apolar–polar repulsion within a single chain. At a temperature below the onset temperature for hydrophobic association, a substantially hydrophobic protein-based polymer, constrained by apolar–polar repulsive free energy of hydration, could be expected to exhibit characteristics of a higher single chain elastic modulus for the ionized state than for the uncharged state. On the other hand, at a temperature above the onset temperature for hydrophobic association, the situation would be reversed.

ACKNOWLEDGEMENTS

The author gratefully acknowledges the support of the Office of Naval Research (ONR) by means of Grant No. N00014-98-1-0656 and Contract No. N00014-00-C-0404.

REFERENCES

1. Gaub, H. E., and Fernandez, J. M., 1998, The molecular elasticity of individual proteins studied by AFM-related techniques, *AvH-Magazin* **71:**11–18.
2. Urry, D. W., Hugel, T., Seitz, M., Gaub, H., Sheiba, L., Dea, J., Xu, J., and Parker, T., 2002, Elastin: A representative ideal protein elastomer, *Phil. Trans. Roy. Soc. Lond. B* **357:**169–184.
3. Urry, D. W., 1997, Physical chemistry of biological free energy transduction as demonstrated by elastic protein-based polymers, *J. Phys. Chem. B* **101:**11007–11028.
4. Sandberg, L. B., Leslie, J. G., Leach, C. T., Torres, V. L., Smith, A. R., and Smith, D. W., 1985, Elastin covalent structure as determined by solid state amino acid sequencing, *Pathol. Biol*. **33:**266–274.
5. Yeh, H., Ornstein-Goldstein, N., Indik, Z., Sheppard, P., Anderson, N., Rosenbloom, J. C., Cicila, G., Yoon, K., and Rosenbloom, J., 1987, Sequence variation of bovine elastin mRNA due to alternative splicing, *J. Collagen Rel. Res*. **7:**235–247.
6. Urry, D. W., Haynes, B., Zhang, H., Harris, R. D., and Prasad, K. U., 1988, Mechanochemical coupling in synthetic polypeptides by modulation of an inverse temperature transition, *Proc. Natl. Acad. Sci. USA* **85:**3407–3411.
7. Urry, D. W., Hayes, L. C., and Gowda, D. C., 1994, Electromechanical transduction: Reduction-driven hydrophobic folding demonstrated in a model protein to perform mechanical work, *Biochem. Biophys. Res. Comm*. **204:**230–237.
8. Urry, D. W., and Parker, T. M., 2002, Mechanics of elastin: Molecular mechanism of biological elasticity and its relevance to contraction, *J. Muscle Res. Cell Motility* **23**(5–6); Special Issue: *Mechanics of Elastic Biomolecules* (H. Granzier, M. Kellermayer Jr., W. Linke, eds).
9. Monod, J., 1968, On symmetry and function in biological systems, in *Nobel Symposium 11: Symmetry and Function of Biological Systems at the Macromolecular Level* (A. Engstrom, and B. Strandberg, eds.), Almqvist and Wiksell Forlag AB, Stockholm, p. 1527.
10. Urry, D. W., 2006, *What Sustains Life? Consilient Mechanisms for Protein-based Machines and Materials*, Springer-Verlag, LLC, New York (in press), ISBN: 0817639101.
11. Wilson, E. O., 1998, *Consilience, The Unity of Knowledge*, Alfred E. Knopf, New York, p. 8.
12. Urry, D. W., Long, M. M., and Sugano, H., 1978, Cyclic analog of elastin polyhexapeptide exhibits an inverse temperature transition leading to crystallization, *J. Biol. Chem*., **253:**6301–6302.
13. Urry, D. W., 1992, Free energy transduction in polypeptides and proteins based on inverse temperature transitions, *Prog. Biophys. Molec. Biol*. **57:**23–57.

14. Butler, J. A. V., 1937, The energy and entropy of hydration of organic compounds, *Trans. Faraday Soc*. **33:**229–238.
15. Frank, H. S., and Evans, M. E., 1945, Free volume and entropy in condensed systems: III Entropy in binary liquid mixtures; partial molal entropy in dilute solutions; structure and thermodynamics in aqueous electrolytes, *J. Chem. Phys*. **13:**507–532.
16. Stackelberg, M. V., and Müller, H. R., 1951, Zur Struktur der Gashydrate, *Naturwissenschaften* **38:**456; Stackelberg, M. V., and Müller, H. R., 1954, Feste Gashydrate II: Struktur und Raumchemie, *Zeitschrift für Elektochemie* **54:**25–39 (now *Berichte der Bunsengesellschaft für physicalische Chemie*).
17. Urry, D. W., Peng, S.-Q., Xu, J., and McPherson, D. T., 1997, Characterization of waters of hydrophobic hydration by microwave dielectric relaxation, *J. Am. Chem. Soc*. **119:**1161–1162.
18. Luan, C.-H., and Urry, D. W., 1999, Elastic, plastic, and hydrogel protein-based polymers, in *Polymer Data Handbook* (J. E. Mark, ed.), Oxford University Press, New York, p. 78–89.
19. Urry, D. W., 2004, The change in Gibbs free energy for hydrophobic association: Derivation and evaluation by means of inverse temperature transitions, *Chem. Phys. Letters*, 399, 177–183.
20. Urry, D. W., Gowda, D. C., Peng, S.-Q., and Parker, T. M., 1995, Non-linear hydrophobic-induced pKa shifts: Implications for efficiency of conversion to chemical energy, *Chem. Phys. Lett*. **239:**67–74.
21. Pattanaik, A., Gowda, D. C., and Urry, D. W., 1991, Phosphorylation and dephosphorylation modulation of an inverse temperature transition, *Biochem. Biophys. Res. Comm*. **178:**539–545.
22. Urry, D. W., 2006, Function of the F_1-motor (F_1-ATPase) of ATP synthase by apolar-polar repulsion through internal interfacial water, *Cell Biology International*, **30**/1, 44–55.
23. Urry, D. W., Hugel, T., Seitz, M., Gaub, H., Sheiba, L., Dea, J., Xu, J., Hayes, L., Prochazka, F., and Parker, T., 2003, Ideal protein elasticity: The elastin model, in *Elastomeric Proteins: Structures, Biomechanical Properties and Biological Roles* (P. R. Shewry, A. S. Tatham, and A. J. Bailey, eds.), Chapter 4, The Royal Society, Cambridge University Press, Cambridge, UK, pp. 54–93.
24. Hugel, T., personal communication.
25. Alonso, L. B., Bennion, B. J. and Daggett, V., 2001, Hydrophobic hydration is an important source of elasticity in elastin-based polymers, *J. Am. Chem. Soc*. **123:**11991–11998.
26. Flory, P. J., 1968, Molecular interpretation of rubber elasticity, *Rubber Chem. Tech*. **41:**G41–G48.
27. Hoeve, C. A. J., and Flory, P. J., 1974, Elastic properties of elastin, *Biopolymers* **13:** 677–686.
28. Eyring, H., Henderson, D., Stover, B. J., and Eyring, E. M., 1964, *Statistical Mechanics and Dynamics*, Wiley, New York.
29. Chang, D. K., and Urry, D. W., 1989, Polypentapeptide of elastin: Damping of internal chain dynamics on extension, *J. Computational Chem*. **10:**850–855.
30. Henze, R., and Urry, D. W., 1985, Dielectric relaxation studies demonstrate a peptide librational mode in the polypentapeptide of elastin, *J. Am. Chem. Soc*. **107:**2991–2993.
31. Webster's Dictionary defines vital as to be "concerned with or necessary for the maintenance of life" and "performing an essential role in the living body." Britannica 2002 expresses the sense of *élan vital* as "the vital force or impulse of life," as a creative

principle due to Henri Bergson indicating "to be immanent in all organisms and responsible for evolution."

32. Urry, D. W., Haynes, B., Zhang, H., Harris, R. D., and Prasad, K. U., 1988, Mechanochemical coupling in synthetic polypeptides by modulation of an inverse temperature transition, *Proc. Natl. Acad. Sci. USA* **85:**3407–3411.
33. Rodríquez-Cabello, J. C., Reguera, J., Alonso, M., Parker, T. M., McPherson, D. T., and Urry, D. W., 2004, Endothermic and exothermic components of an inverse temperature transition for hydrophobic association by TMDSC, *Chem. Phys. Lett*. **388:**127–131.
34. Lange, C., and Hunte, C., 2002, Crystal structure of the yeast cytochrome bc_1 complex with its bound substrate cytochrome c., *Proc. Nat. Acad. Sci. USA* **99:**2800–2805. Protein Data Bank, Structure File 1KYO.
35. Zhang, Z., Huang, L., Shulmeister, V. M., Chi, Y. I., Kim, K. K., Hung, L. W., Crofts, A. R., Berry, E. A., and Kim, S. H., 1998, Electron transfer by domain movement in cytochrome bc_1., *Nature* **392:**677–684. Protein Data Bank, Structure Files 1BCC and 3BBC.
36. Abrahams, J. P., Leslie, A. G. W., Lutter, R., and Walker, J. E., 1994, Structure at 2.8 Å of F_1-ATPase from bovine heart mitochondria, *Nature* (*London*) **370:**621–628. Protein Data Bank, Structure File 1BMF.
37. Menz, R. I., Walker, J. E., and Leslie, A. G. W, 2001, Structure of bovine mitochondrial F_1-ATPase with nucleotide bound to all three catalytic sites: Implications for mechanism of rotary catalysis, *Cell* **106:**331–341. Protein Data Bank, Structure File 1H8E.
38. Kinosita, K., Yasuda, R., and Noji, H., 2000, F_1-ATPase: A highly efficient rotary ATP machine, *Essays in Biochemistry, Molecular Motors*, **35:**3–18, G. Banting, and S. J. Higgins, eds. Portland Press.
39. Himmel, D. M., Gourinath, S., L Reshetnikova, Shen, Y., Gyorgyi, A. G., and Cohen, C., 2002, Crystallographic findings on the internally uncoupled and near-rigor states of myosin: Further insights into the mechanics of the motor, *Proc. Nat. Acad. Sci. USA* **99:**12645–12650 . Protein Data Bank, Structure Files 1KK7 and 1KK8.
40. Fisher, A. J., Smith, C. A., Thoden, J., Smith, R., Sutoh, K., Holden, H. M., and Rayment, I., 1995, Structural studies of myosin: Nucleotide complexes: A revised model for the molecular basis of muscle contraction, *Biophys. J*. **68:**19s–28s.
41. Rayment, I., Smith, C., and Yount, R. G., 1996, The active site of myosin, *Annu. Rev. Physiol*. **58:**671–702.
42. Stroup, E. W., Pungor, A., and Hlady, V., 1997, A constant compliance force modulation technique for SFM imaging of polymer surface elasticity, *Ultramicroscopy* **66:**237–249.
43. Hlady, V., Pungor, A., and Stroup, E. W., Method for mapping mechanical property of a material using a scanning force microscopy, US Patent 5,700,953, issued 12/23/97.
44. Leahy, D. J., Hendrickson, W. A., Aukhil, I., and Erickson, H. P., 1992. Structure of a fibronectin type III domain from tenascin phased by MAD analysis of the selenomethionyl protein, *Science* **258:**987–991.

Chapter 10

BIOPROCESSING OF SILK PROTEINS-CONTROLLING ASSEMBLY

Hyoung-Joon Jin, Jaehyung Park, Regina Valluzi, Ung-Jin Kim, Peggy Cebe, and David L. Kaplan
Departments of Chemical & Biological Engineering; Biomedical Engineering; Physics; Tufts University; 4 Colby Street; Medford, Massachusetts 02155 USA

Abstract: A model for silk processing in silkworms and spiders is proposed and based on the unique domain structure in the sequences of silk proteins. A hypothesis for the silk-spinning process begins with chain folding at lower concentrations of protein, proceeds through formation of micelles and then micellar aggregates (globules) through water loss and increasing protein concentration, and finally leads to fiber formation due to the physical shear process during fiber spinning. Many aspects of this process can be mimicked *in vitro*, and the all-aqueous environment used is instructive as a model for polymer processing in general.
Key words: Silk, self-assembly, fibers, bioprocessing

1. INTRODUCTION TO FIBROUS PROTEINS

A number of groups have been studying various aspects and length scales of fibrous protein structure with the common goal of a mechanistic basis for fibrous protein structure and the associated processes, allowing us to direct these processes and the materials structures generated, as well as to generalize them to the more general class of complex blocky polymers. This general grouping of proteins is characterized by the presence of highly repetitive primary sequences,[1–12] which lead to long stretches of regular secondary structures (unlike highly mixed secondary structures seen in globular proteins). Thus, direct helix–helix interactions lead to the formation of complex long-range ordered materials, such as fibers,[13–22] basement membranes,[23,24]

V. Renugopalakrishnan et al. (eds.), Bionanotechnology, 61-78.

and related structures, "encoded" into the protein structure at the primary sequence and secondary structural levels.[2,16,18,23,25–42] Fibrous proteins form mechanically robust structures,[1,4,15,43–50] while also providing recognition sites in regular patterns for cell binding, for mineralization (e.g., in bone[51,52]) and interactions with other extracellular matrix components such as glycosaminoglycans and other matrix-associated proteins. The multiscale order achieved with this group of proteins provides a remarkably versatile and useful template for the exploration of a wide range of questions in materials science and engineering. The highly repetitive primary sequence allows salient features of the sequences of these large proteins to be abbreviated as short synthetic genetic variants[2,28–30,34,53–58]—a technique useful in simplifying the complex behavior of these proteins to an intelligible level while retaining biological relevance. These shorter genetic variants can be polymerized (multimerized) into longer genes to explore sequence and size relationships in systematic ways.[2,59–61]

The route from polymer biosynthesis to complex material structure is of fundamental and applied interest in the fields of materials science and engineering, biology, and biophysics. This process is influenced both by the content of the polymer (e.g., protein primary sequence and chain length) and the surrounding environment.[3,4,15,37,47,53,55,62–65] These processes are instructive and essential in living systems as they form the basis for the development of tissue structure and architecture, cell responses, metabolic compartmentalization, and many related aspects of anatomy and physiology. Furthermore, an understanding of the process of tissue development has direct implications with respect to mutations and disease states, as well as possible interventions in this process.

Fibrous proteins provide the "materials" of nature, whether as fibers (orb webs of spiders), tissue structures (collagens in tissues), hair (keratin), or ceramics (sea shell organic–inorganic composites) and so on. These proteins provide readily available templates for evolutionary tailoring of structure and function. Multigene families exist for the silks and collagens and have been partially characterized. A great deal of insight into protein structure and function is anticipated as an understanding is developed of the structure–function relationships of these fibrous proteins, pattern recognition, relationships between hydrophobicity and structure formation, and many related structural issues.

Fibrous proteins can also be considered nature's equivalent of block copolymers. While the majority of synthetic block copolymer studies have focused on microphase separated patterns in *amorphous* block copolymers, recent advances in synthesis have allowed incorporation of chiral, liquid crystal forming, or crystallizable blocks into synthetic polymers: molecular structures that are far more analogous to fibrous proteins. Using the fibrous proteins as experimentally accessible model systems, we can begin to emulate the complex polymer systems that will be synthetically possible with simpler and well-controlled (template-based protein synthesis) strategies permitting

reasonable questions to be answered. Within this context we can either use native proteins (e.g., silkworm silk, spider dragline silk), model peptides representing conserved regions of the repetitive sequence, or genetically engineered variants of the sequence by using synthetic gene technology to modify in selective ways the chemistry in the sequence. These powerful tools available through molecular biology and protein engineering facilitate direct insight in the role of sequence chemistry, block sizes and distributions, and chain length on materials structure and function.

2. CURRENT STATE OF KNOWLEDGE

The majority of studies on the ''polymer physics'' of fibrous proteins, their physical chemistry, and their processability or process engineering (biological or *in vitro*) have focused on insect silks. An additional substantial literature exists on long-range ordered materials formed by collagen. These protein types have several features in common that are also shared by many fibrous proteins, making them good examples for the ordering processes and process considerations common to fibrous proteins. These common features include a ''blocky'' sequence structure, with long subsequences of amino acids defining large hydrophobic and hydrophilic ''blocks.''[1–12] The tendency to form long stretches of homogeneous secondary structure is an outgrowth of the patterns built into the sequence.[3,4,9,29,30,34,47,66–82] The sequence blocks in these proteins are often significantly hydrophobic or hydrophilic.[83] Thus, in the flexible solution state form, we would expect many of the features associated with flexible amphiphiles, including surfactancy, micelle formation, and possible micelle-based lyotropic liquid crystalline phase behavior.[22,28,30,46,47,49,62,84–91] In their final state and at the last biological ''processing'' stages leading to the final protein material, there is significant evidence that silks, collagens, and possibly other fibrous proteins exist in ''helicoidal'' structures[26,27,30,31,35,44,47,91–100] resembling liquid crystals.[16,26,27,30,31,35,47,53,91–95,99–101] The term ''pseudomorph'' has been coined to describe this state.[92,93] There is also evidence for true liquid state chiral liquid crystals as helicoids precursors in both silks and collagens. In both these proteins, liquid crystalline textures are observed in solutions of molecules—before fibrils (collagen) or crystalline domains (silks) have formed.[19,27,29,35,47,53,91,94,96,102–104] Thus a role for precursive chiral liquid crystalline order is suggested as a way to promote long-range order as aggregated structures (fibrils, crystalline domains, etc.) grow in these helicoidal multiscale ordered biological materials. When the amphiphilic nature of many of these molecules is considered along with the final helicoidal or multiscale ordered structure observed in many materials formed by the molecules, it is evident that there is a predisposition toward long-range ordered structures driven by the

"hydrophobic effect" (or through amphiphilic interactions forming micelles or similar structures) that can either enhance or compete with the liquid crystalline ordering from chiral and other steric intermolecular interactions.

Little is known about collagen self-assembly prior to fibril formation in the context of liquid crystallinity and long-range material or fluid state order.[16,27,29,53,93,103] The prevailing view holds that the triple helix is formed early, perhaps prior to leaving the cell, resulting in a rigid amphiphilic molecular unit and little possibility for *flexible* micelle formation. Amphiphilic interactions may still play a role in materials order. In helicoids of collagen fibers, the helicoid structure at micron length scales suggests a chiral liquid crystal,[16,26,92,93] but the internal structure of the collagen fibers implies smectic ordering typical of both amphiphile-based structures (lamellar phases) and "rodlike" liquid crystals (smectic phases).[16,53] Thus the two types of liquid crystallinity may act in concert to enhance the formation of structure, but too little is known about collagen self-assembly mechanisms and processes at these length scales to say much more at this point. A number of intriguing studies are progressing in this area. These include studies of sequence–secondary structure–liquid crystallinity relationships in model collagens,[29,53,54,105] descriptions of the liquid crystalline behavior of native collagen and tropocollagen molecules and collagen-like features in materials prepared *ex-vivo*,[27,103,106] and molecular structure studies at high resolution using single crystals of collagen-like triple helical peptides that elucidate sequence-secondary structure relationships in greater detail.[107,108]

In our ongoing studies, we see stabilization of chiral smectic phases in studies using collagen-like model molecules with enhanced amphiphilicity—an instance of an amphiphilic interaction enhancing the smectic phase formation expected for low polydispersity rigid rods such as the model triple helices used.[29,53,105,109] In these studies, simple cholesteric or nematic liquid crystallinity is rarely observed. Lyotropic liquid crystalline phases based on amphiphilic interactions between unfolded (non-triple-helical), flexible (not rigid rods) collagen molecules are observed, and the formation of well-ordered amphiphile phases greatly inhibits chiral liquid crystallinity.

In contrast, the processing and self-assembly pathways of silks are far better elucidated at all stages of the process. Development of a tentative model for the early stages of structural organization in silk processing can thus provide a useful intellectual roadmap to organize data on different aspects of silk structure occurring at different length scales and stages in the spinning process. Such a model, while tentative and by no means exhaustive, provides a basis for comparisons between fibrous proteins, by defining stages in the multiscale self-assembly process in which different molecular/materials feature interactions are dominant. As this model is revised by additional data, we will be better able to "connect the dots" and knit together the variety of

studies on silks and other fibrous proteins into a framework for portable knowledge and understanding.

In the literature on silks, there are several crystal structures reported: Silk I, a ''precursor'' structure observed when silk solutions or silk glands are dried unperturbed;[70,110–113] Silk II, the antiparallel β-sheet structure observed in spun fibers;[69–71,73,74,112,114,115] and Silk III, a structure stabilized at interfaces that optimizes the surfactancy of the silk repetitive motif.[33,34,69–71,73,74,87,88,112,114–117] A number of studies exists connecting secondary structure and crystalline polymorph to process environment. Studies of silk films prepared on different surfaces indicate that secondary structure, crystal structure, and morphology are surface sensitive and can be templated by surfaces with differing chemistries. Hydrophobic solid surfaces favor silk I-like structures whereas hydrophilic surfaces favor silk II-like structures.[118,119] These results are confirmed with liquid–liquid interface experiments, where silk III, a structure with crystal diffraction pattern more similar to silk I than silk II, is observed at hydrophobic fluid/aqueous silk solution interfaces.[31,32,34] At these interfaces, swollen, highly concentrated surface excess layers of silks form with a helicoid structure resembling a cholesteric or chiral smectic liquid crystal.[31] Synthetic moderate weight (4000–8000 Da) oligopeptides with silk-like repetitive sequences and enhanced amphiphilicity (from charged or strongly hydrophilic end-blocks) reproduce the same general types of results observed for silk proteins prepared at surfaces with different chemistries,[28] and a surface-induced transition from a soluble β-strand to silk I and silk III upon adsorption to a hydrophobic surface has been demonstrated.[34] The silk I structure is also observed in dried air-blown foams of silk solution.[42] All of these results and observations indicate a high degree of surfactancy for the silk molecule and point to amphiphilic interactions as a major determinant of structure at even the smallest length scales. Other model proteins vary in the size and distribution of silk-like and hydrophilic blocks. Preliminary data suggest that different folded supersecondary structures, driven by hydrophilic and hydrophobic interactions with the aqueous environment, are favored by different patterns of blocks. Supersecondary structure, as well as secondary structure, may be important in silk spinning, because the ''loosely'' folded versions of these structures can provide well-segregated regions of hydrophilic and hydrophobic chemistry to aid micelle formation. A more tightly folded, hydrogen bonded supersecondary structure can act as a rigid asymmetric structure, allowing chiral liquid crystalline textures and helicoid materials to form.

In view of the observed surfactancy of silks, it is not at all unreasonable to expect other types of amphiphile behavior in silk solutions at different concentrations. Several studies in the recent literature report results suggestive of micelles or other types of amphiphilic structures in silk from silk

glands. ''Globules'' have been observed in silks with dimensions too large to be single molecule phenomena, suggesting an aggregated structure such as a micelle.[120–122] On smaller length scales disk-shaped structures have been observed and correlated to liquid crystalline phases in spider silks, suggesting organized micelles of a single molecule or a small number of molecules.[46] These observations suggest true lyotropic liquid crystallinity in which micelles of flexible molecules grow and alter their curvature and shape in response to changes in concentration and solubility. Typically, an amphiphilic flexible surfactant can form small spherical micelles, larger cylinders, aligned cylinders, and lamellar layered structures as the surfactant concentration is increased. Bilayered and multiwalled structures are also a possibility. Given the observed surfactancy of silk proteins and the presence of micelle-like structures in solution, lyotropic liquid crystalline phase behavior suggests itself as a mechanism for organizing silk into high concentration domains, biasing the orientation at the molecular level through orientation of micelles, and for possibly enhancing water removal in the gland by providing clear channels for fluid to flow.

New biosynthetic variants of spider silks may be useful in probing amphiphilic interactions and phase behavior *in vitro*.[55,57,62,123] Surface and interface induced long-range ordering has been observed for amphiphilic tri-block silk-like peptides, with material patterns very reminiscent of textures observed in the silk gland, developed when peptides are prepared at an air–water interface.[28,34] The detailed structure of the pattern is controlled by solubility of the peptides—in this case screening of charges on the hydrophilic end-blocks, and by secondary structure stabilization and orientation at the interface, both of which are designed into the molecules at the sequence level. Promising model proteins have been developed incorporating redox (methionine oxidation) and phosphorylation (enzymatic phosphorylation of serine) triggers that allow their global solubility and local amphiphilicity to be chemically altered.[55,57,62,123] Studies of these synthetic silk analogues indicate that triggering changes secondary structure, solubility, and crystallinity.[55] Additional changes in long-range pattern are observed with solubility changes induced by triggering of these silks, suggesting that a solubility dependent mechanism from liquid crystalline organization may be present and sensitive to triggering. These biosynthetic variants also feature different hydrophobic/hydrophilic blocky structures and may eventually help elucidate relationships, interactions, and possible transitions from lyotropic micelle driven fluid order to the chiral rigid rod interactions needed to produce helicoid materials and materials aligned at the molecular level.

There is a clear need for both studies on the early stages of the silk-spinning process, in which micellar structures are expected and for an intellectual framework to help us understand and connect silk secondary structure, supersecondary structure, micellar and surfactancy driven structures, chiral ''rigid rod'' liquid crystalline materials morphologies, and the multiscale

ordered structure of silk fibers. We have decided to focus on surfactancy in silks and micelle formation and development as a first step in building a hypothetical framework for the silk-spinning process.

3. A MODEL FOR SILK PROCESSING

We have been studying reconstituted silkworm fibroin using various strategies to modify and process the solutions to fabricate materials and understand solution properties relevant to silk spinning. Among these strategies, there have been blending of silk fibroin and poly(ethylene oxide) (PEO) to obtain a solution viscosity suitable for electrospinning nanoscale diameter fibers and studies on film formation and structure in which predictable nano- and microscale surface and bulk morphologies (unpublished data) have been obtained. A hypothesis for the process of silk spinning by insects and spiders is based on insight gained from these approaches (Jin *et al*., 2003d). We can break down the overall process (natural or mimicked) of silk processing the glands of silkworm and spider in the following sections.

3.1 Silk Biosynthesis and Formation of Aqueous Solutions

Fibroin proteins are synthesized in epithelial cells lining the gland and then secreted into the lumen of the gland.

3.2 Biopolymer Folding (Hypothetical)

Portions of the silk fibroin molecule adopt a silk I conformation as the concentration of protein in the lumen of the gland increases. We have observed features suggesting transient chain folding with hydrophobic hydration thermodynamics driving the process. It is also apparent that the rate at which water is removed from the silk solution has a direct impact on the formation of micelles versus the formation of a gel in which slower rates of concentrating the fibroin facilitate micelle formation, consistent with the slow dynamics expected as large molecules order into larger structures (micelles) in dynamic equilibrium with the surroundings. We suspect that micelles represent a soft organized precursive state that "sorts out" the regions of the molecule that will become β-sheets in the fiber, while preventing interactions that will inhibit an ordered β-sheet fiber. These relationships remain to be clarified with experimental studies.

3.3 Micelle Coalescence or Curing

As chain folding of fibroin progresses to a metastable folded structure (the asymmetric rigid object needed for chiral liquid crystalline phases such as the

cholesteric textures observed in silk glands), intermicellar interactions lead to the formation of "globules" and alignable micellar structures such as cylindrical micelles. These larger globules may be either aggregates of smaller micelles or multilayered structures. The end result is the separation of locally planar domains of concentrated fibroin that is segregated from much of the surrounding water but still swollen and mobile.

3.4 Final Curing with Polymers

At the final stage of processing, water-absorbing polymers (sericin in the natural process, PEO in our mimicked process) draw additional water out of the globules to facilitate the curing process. This stage occurs just before spinning.

There is clearly a transition from a complex micellar solution to a solution state structure more like a rigid rod liquid crystal with intercalated water at high concentrations and late stages in the spinning process, and polycrystallinity and molecular alignment are notable features of the final fibers. Studies on shearing of micelles may provide insight into the features of silk solutions as this transition is approached.

There remains little detail in the literature on silk biosynthesis and the initial stages in silk processing. It is unknown if the silk, when secreted into the gland from the cells lining the gland is in a dimer (disulfide liked) or monomer state. It is also unclear what content of calcium or other ions are present. Regenerated silk solutions can be prepared to reproduce the chemical makeup of silk solutions at various stages in the silk-spinning process. These regenerated solutions are prepared by cleaning and degumming the cocoon fibroin and then dissolving the clean fibroin in LiBr or LiSCN salts. These solutions can be dialyzed against pure water or other salt solutions to produce a silk solution with the desired properties.

From the data available, it is uncertain whether the silk fibroin biopolymers "fold" to adopt a supersecondary structure early in the spinning process, especially considering the absence of knowledge regarding the components of the initial solution and their concentrations. Figure 1 illustrates the biopolymer folding step. It is possible that folding into supersecondary structures occurs after micelles have begun to form, but this is unlikely. A conformational bias toward a supersecondary structure can occur when the protein amphiphiles are in the monomeric form, whereas micelle structures require a critical concentration of protein amphiphile (the critical micelle concentration) to be stable. Regenerated silkworm silk fibroin can be used to mimic this process *in vitro*, and these studies suggest that surfactancy dominates the initial ordering interactions as hydrophobic and hydrophilic sequence

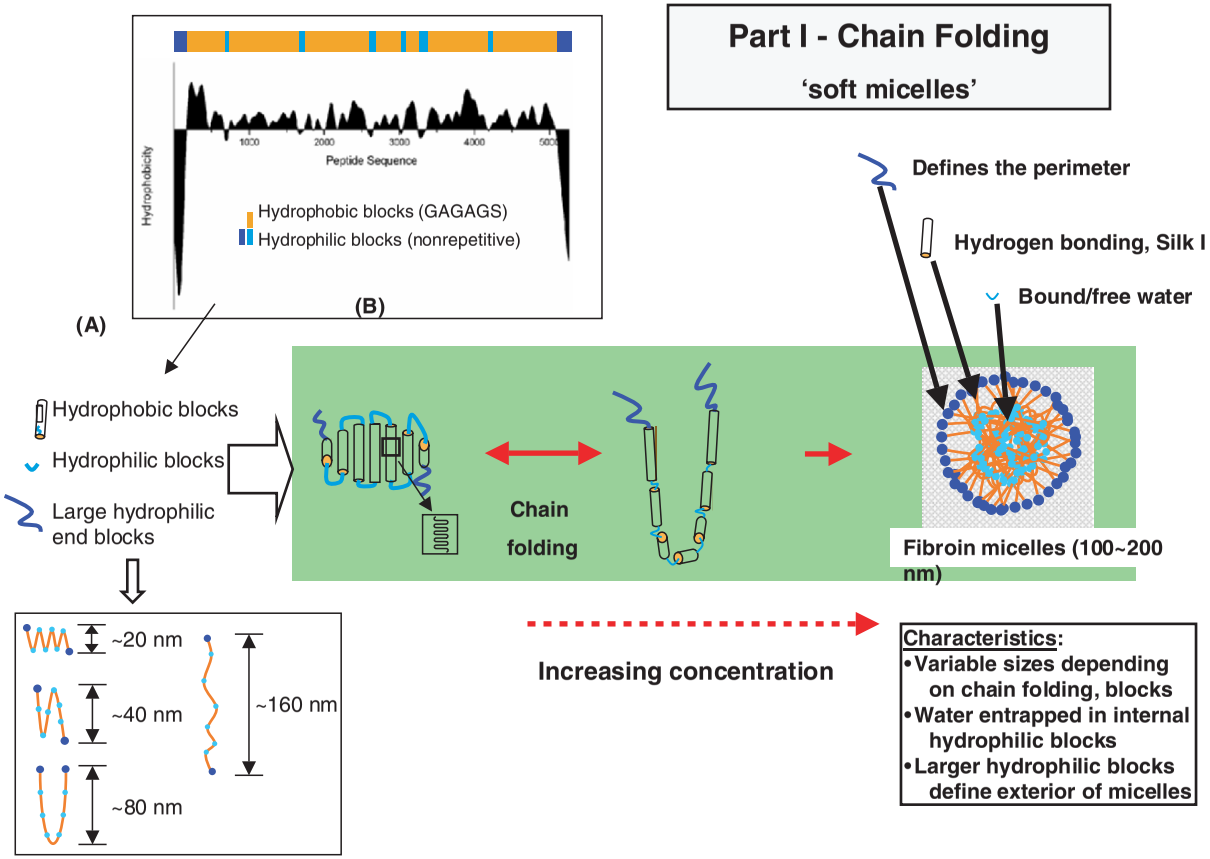


Figure 1. Transient chain folding allows the formation of silk biopolymer micelles and then miscellar aggregates (globules) incorporating entrapped water. Size and properties of micelle walls are a consequence of the chain folded structure present and the transient nature of the folds may help regulate, control of water, and rate of increase in protein concentration. Silk fibroin sequence data are from Zhou et al., 2000.

domains separate into different localized regions in solution, resulting in loosely "folded" chains with some defined secondary and supersecondary structure. A hydrophobic plot of the sequence illustrates the nature of the hydrophobic and hydrophilic segments in this repetitive protein, suggesting types of likely supersecondary transient folded structures present prior to micelle formation. Formation of secondary (silk I conformation) and supersecondary structures results in spatially separate solution regions that are rich in hydrophilic or hydrophobic monomers—local patterning of the solution state hydrophilicity occurs through folding of the chain into a loose structure rather like a "self-micelle." As water is extracted from the solution, a critical concentration is reached where micelles form until at concentrations greater than 20% fibroin, in the storage region of the gland, micellar structures are prevalent and water content reaches a steady state. Here the protein is likely to have a local secondary structure (and stretches of secondary structure) similar to the silk I crystalline form, which is believed to represent a minimum in torsional energy for proteins with small side chains. It is instructive to keep in mind that this is not a process that can occur easily unless the protein concentration increases gradually; otherwise premature gelation due to hydrophobic physical cross-links occurs, precluding micellar structures to form a gel. Significant water is buried internally in the micelles that may in fact be bilayered and multiwalled structures with hydrophilic interior regions. From interaction strength alone, it is likely that most of the large hydrophilic N- and C-termini of the fibroin interact with the external aqueous environment. These chain "ends" are significantly larger than the internal blocks in the protein. At this stage, we presume that some level of intermicellar interactions may occur due to transient hydrophobic interactions in the dynamic micellar structures. Because the silk can still be collected in liquid form at this stage, it is unlikely that strong long-lived interactions, such as cysteine bonds or β-sheet nucleation, occur at this point—interactions of this type would freeze the micellar structures into gelatinous insoluble globules.

Once micelles form and incorporate most of the fibroin in solution, a number of different processes can come into play. Continued "curing" or "coalescence" of the fibroin micelles, as further hydrophobic interactions take place as water is actively transported out of the lumen of the gland, may occur. Due to the dynamic nature of these solution state structures, evolution of micellar structures into tubes and lamellar layers is possible, as is the formation of large globules by adding more layers onto a multiwalled "onion" structured micelle. The evolution of the micellar structures may also result in lyotropic liquid crystallinity, as soft micellar structures with an asymmetric shape align in solution. While there is only weak evidence for this idea, it is somewhat appealing intellectually because it provides a mechanism for spontaneous alignment of the oriented molecules in the walls of the micelle

by aligning the large dynamic micellar superstructure. This would allow significant ordering in the silk solutions and regions of high-fibroin concentration that remain plasticized and mobile due to the surrounding water.

The addition of sericin at the end of the gland extracts water from the fibroin micelles into the sericin phase due to the hydrophilic nature of this polymer (with up to 30% serine content).[124–128] Sericin coats the fibroin while also lubricating the passage of the fibroin through the later stages of the gland during spinning. The end result is a core fibroin structure surrounded by a skin coating of sericin (Fig. 2).[36,129–132] Despite the fact that the two proteins appear miscible in solution, sericin occurs in silk fibers as a separate phase on the outside of the fiber, possibly due to the high-molecular weight and slow diffusion of the sericin protein. One should note that there is a miscibility gap between fibroin and sericin, where the two proteins are miscible in dilute solution in water and immiscible in solids (and by extrapolation, in concentrated solution),[127,130,131] making mixing unlikely at the late stage in spinning in which sericin is added.

At these high concentrations, there must be a point at which the micellar organization of the solutions is replaced by homogeneous solutions of fibroin containing a structure that can act as a liquid crystalline rigid rod or plate with intercalated water plasticizing the liquid crystalline phase to promote fluidity. ''Cholesteric'' helicoids have been observed in sectioned silk glands using transmission electron microscopy (TEM),[35] but the associated electron diffraction patterns suggest significantly more order than is developed in a cholesteric phase. Silk fibers have chiral features reminiscent of hexatic structures. Micellar structures neither are observed associated with late stages of the spinning process, nor are they expected from comparison with other lytotropic liquid crystals in which very low concentrations of solvent result in a homogeneous phase of solvent dissolved in amphiphile. Thus while amphiphilic interactions may contribute to ordering at this stage in spinning, it is more likely that the supersecondary structures, developed in dilute solution and oriented and helped to order in a dimensionally constrained micelle wall, have developed into pseudo-rigid hydrogen-bonded (but metastable) folded rods or plates, generating the chiral liquid crystalline features observed. There is very little knowledge about the middle stages of the silk-spinning process, and this possibility is advanced as a sensible hypothesis tenuously supported by the limited available data.

During the final spinning step, physical shear plays an important role in many aspects of the process. Our current view is that shear aligns the helicoidal liquid in the late stage concentrated solution and ''pulls out'' some of the supersecondary structures in the aligned liquid crystal. For a secondary structural change to occur at low shear, the secondary structure must be among the ''softest'' or most easily sheared objects in the multiscale ordered

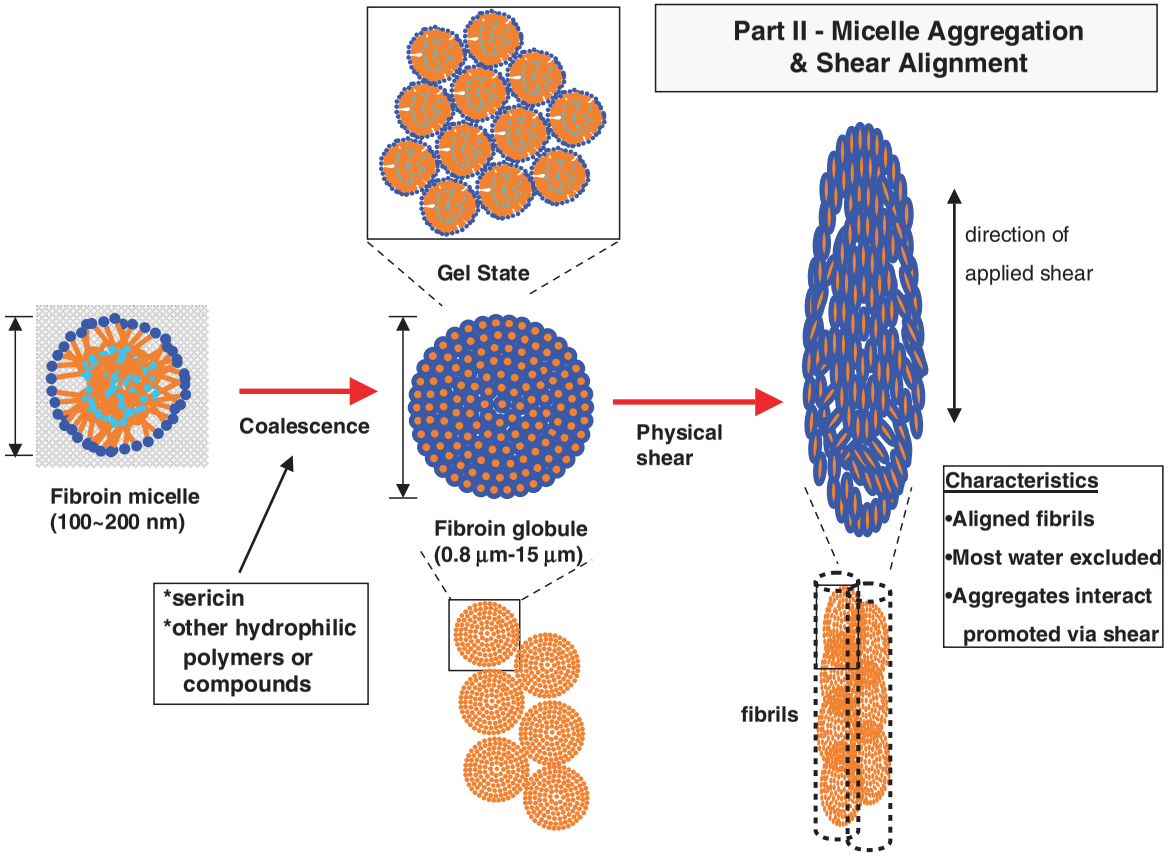


Figure 2. Final stages of globule curing and shear induced fibril formation where PEO and sericin may play a role at the later stages of this process in further removal of water.

fluid. Thus micelles, if present, would be expected to shear first, leaving the sturdier component, supersecondary and secondary structures, intact. Thus a long-range ordered fluid phase that can *resist* shear is proposed such as a cholesteric or smectic liquid crystalline phase. The latter is more likely in view of the order observed in electron diffraction from silk gland sections, the amphiphilic nature of the silk molecule, and the proposed structural evolution pathway that would promote segregation of fibroin into layers. Similarly the supersecondary structure formed must be far less stable than a globular protein fold, perhaps existing as a ''bias'' toward folding at a particular set of locations or a weakly stabilized loop or turn. The secondary structures present—the silk I conformation—must also be poorly stabilized suggesting a conformation involving incomplete or highly strained interchain hydrogen bonds.

When the ''micellar'' liquid is subjected to shear, fibrils form. The shear processing induces significant interglobule interactions. We can recapitulate this process *in vitro* by shearing the globules between a cover slip and a glass slide (unpublished data). When the micellar or globular liquids are sheared, only very weak orientation is observed on the molecular level at shear rates similar to those occurring in natural silk spinning. Furthermore, the weakly oriented structure is a variant of the silk I crystalline polymorph and not the silk II polymorph observed in spun fibers. Both of these observations support the idea that it is the *micelles* that orient and not the molecules and crystalline domains and that most of the shear stress is dissipated in the destruction and alignment of micelles, leaving the secondary structure of the fibroin molecule intact. The fact that weak molecular orientation (and the silk I precursor conformation) is observed in WAXD studies suggests that alignment of micelles during the spinning process may help provide some degree of molecular orientation without nucleating the β-sheet structures leading to insolubility and intractability in silks.

In our biomimetic approach, PEO is used to extract water from concentrated silk solutions instead of sericin and also serves as a processing aid. In these biomimetic approaches, there is no requirement for exogenous divalent cations or other small-molecular weight species. Other than control of the water and chain folding, the process seems to be functional in the absence of cofactors. The change in pH from the beginning to end of the process likely plays some role in the nature of the interglobule interactions, mediated through changes in amino acid side chain ionization; however, experimentation is needed to validate this possibility. Cations, other small-molecular weight species, posttranslational modifications in the form of glycosylation, or phosphorylation are also feasible, and changes in disulfide bond formation may all play a role in the natural process. These options will provide rich area of study to understand how these factors may modulate the fundamental processing paradigm described here.

4. ANALOGIES BETWEEN SILKWORM SILK PROCESSING AND SPIDER SILK PROCESSING

A similar process can be hypothesized for spider silks, but the detailed features of the silk and solution structure are expected to be different due to the different size, blocky pattern, and repetitive motifs present in spider silks, as well as due to the absence of a sericin helper protein. One should note that silkworm silk fibers also have a much larger diameter than spider silks providing a possible rationale for the role of sericin as a water-extracting agent in silkworms but not in spiders. Spider dragline silks have a large number of short-hydrophobic and hydrophilic blocks in their main repetitive sequence in which the contrast in hydrophobicity is also stronger than in silkworm silk. There is also no clear motif that would promote secondary structure polymorphism through surfactancy, but the size, definition, and arrangement of blocks suggest that supersecondary folded structures, such as *cross β-sheets* and *greek key* motifs, may form more readily in solution. It is difficult to predict how these changes in local structure would translate into micelle curvature, but a large number of folds might enhance the ability of spider silk molecules to form single walled (very small) micellar structures containing one or a small number of molecules by providing a large number of hydrophilic fold points per molecule. Dragline silks are spun at higher shear rates than silkworm silks, suggesting that more stable supersecondary and secondary structures can exist in the gland and still be processed by shear.

In the absence of sericin (the case for the spider), the "curing" process would be slightly different. Either less curing is required to begin with since the water content would be lower to begin with, and this can be controlled to a large extent based on the rate of silk concentration increase, or other molecules replacing the role of the sericin. There are a number of reports indicating significant coatings of spider silk fibers by hydrophilic sugars, neurotransmitters, and other small molecules, thus contributing in a way similar to sericin. These molecules may be playing a dual role in drawing water out of the globules during the later stages of processing, while also providing biological functionality related to prey capture.

An interesting unanswered question within the context of these insights into silk processing is the role of fibroin molecular weight. Within limits, it is not clear that chain end defects due to lower-molecular weight protein chains would have a major negative impact on fiber properties. It is problematic that the N-terminus of the major ampulate protein in *N. clavipes* dragline silk is not known. However, based on extrapolation from other spider silk sequence data (e.g., flagelliform silk), combined with the data from the silkworm, extrapolation of our observations from the silkworm to the spider seems

reasonable. By analogy, we would expect a large hydrophilic domain to be present at the N-terminus of the ampullate gland silk protein from the spider dragline silk. The presence of the cysteine in the larger hydrophilic domain at the C-terminus of silkworm and spider silks may not be a requirement for the process. Since intermicelle interactions can be driven by physical shear, the cysteines may facilitate the process, enhance overall stability of these interactions, and may also be important in the environmental stability of the proteins when incorporated into orb webs or as cocoons. However, sufficient stabilization of the structures can occur in the absence of such disulfide bonds based on the apparent absence of cysteines in some terminal domains of spider silks (R. Lewis, personal communication), yet stable structures are still formed.

5. SILK PROCESSING AS A MODEL FOR SYNTHETIC POLYMER PROCESSING

The proposed mechanistic model for silk protein processing into fibers suggests some intriguing options and opportunities in the field of synthetic polymer science and engineering, particularly with respect to utilizing aqueous processing environments and the design of polymer blocks to optimize specific interactions at specific stages in the process. A key feature of the process appears to be an amphiphilic sequence design with a hierarchy of amphiphilic patterns in the sequence allowing multiple scales of interaction to be addressed in differing environments. Many of these proteins can be approximated as tri-block polymer designs with large hydrophilic end blocks and a series of internal hydrophilic and hydrophobic blocks. Molecular complexity appears to be a key to self-assembly of complex materials.

In fibroin, initial steps in chain folding lead to hydrogen bonding and van der Waals interactions that are responsible for the formation of key structural transitions in silks: the metastable silk I and the very stable (almost irreversible) silk II β-sheet crystalline state. The chemistry and pattern of the amino acids is clearly important in the overall process, and patterned synthetic polymers may be induced to form useful transient folded structures as well as to aid in processing.

6. CONCLUSION

By analogy to the role of genetics and environment on protein function in human biology, we can equate silk processing described herein. Clearly sequence in the form of amino acid chemistry is essential to silk processing,

defining the sizes and distributions of the hydrophilic and hydrophobic domains, as well as their potential interactions with water and among themselves (intra- and inter-). It is also apparent that the rate and extent of buildup in protein concentration, as well as the spinning process itself provide the environmental inputs to the process. With the processing model as presented, the synergy between these two features (genetics vs environment) relative to their contributions to structure, morphology, and ultimately function (e.g., mechanical properties, biological interactions) can be more explicitly explored. This opportunity will help clarify evolutionary considerations with respect to hydrophobic domain size in different silks, selection and sequence, sizes of the fibroin chain, and extent to which conservation in chain length is important, and relationships between sequence and function.

ACKNOWLEDGEMENTS

We thank the NSF-DMR, NSF-BES, NIH-NIDR, and the AFOSR for support at various phases of the research described here.

REFERENCES

1. Lucas, J. M.; Vaccaro, E.; Waite, J. H. *J Exp Biol* **2002**, *205*, 1807–1817.
2. Asakura, T.; Kato, H.; Yao, J.; Kishore, R.; Shirai, M. *Polym. J.* **2002**, *34*, 936–943.
3. Zhou, C. Z.; Confalonieri, F.; Jacquet, M.; Perasso, R.; Li, Z. G.; Janin, J. *Proteins* **2001**, *44*, 119–122.
4. Hayashi, C. Y.; Lewis, R. V. *Bioessays* **2001**, *23*, 750–756.
5. Datta, A.; Ghosh, A. K.; Kundu, S. C. *Insect Biochem Molec* **2001**, *31*, 1013–1018.
6. Tatham, A. S.; Shewry, P. R. *Trends Biochem Sci* **2000**, *25*, 567–571.
7. Nousiainen, M.; Rafn, K.; Skou, L.; Roepstorff, P.; Andersen, S. O. *Comp Biochem Phys B* **1998**, *119*, 189–199.
8. Qin, X. X.; Waite, J. H. *J Exp Biol* **1995**, *198*, 633–644.
9. Andreva, N. S.; Esipova, N. G.; Millionova, M. I.; Rogulenkova, V. N.; Shibnev, V. A. In *Conformation of Biopolymers*, 1 ed.; Ramachandran, G. N., Ed.; Academic Press: New York, New York, 1967; Vol. 2, p. 785.
10. Mello, C. M.; Senecal, K.; Yeung, B.; Vouros, P.; Kaplan, D. *Silk Polymers* **1994**, *544*, 67–79.
11. Case, S. T.; Powers, J.; Hamilton, R.; Burton, M. J. *Silk Polymers* **1994**, *544*, 80–90.
12. Rothfus, J. A. *J Agr Food Chem* **1996**, *44*, 3143–3152.
13. Becker, M. A.; Mahoney, D. V.; Lenhert, P. G.; Eby, R. K.; Kaplan, D.; Adams, W. W. In *Silk Polymers*, 1994; Vol. 544, pp. 185–195.
14. Cappello, J.; Mcgrath, K. P. *Silk Polymers* **1994**, *544*, 311–327.
15. Hayashi, C. Y.; Shipley, N. H.; Lewis, R. V. *Int. J. Biol. Macromol.* **1999**, *24*, 271–275.
16. Hukins, D. W. L.; Woodheadgalloway, J. *Molecular Crystals and Liquid Crystals* **1977**, *41*, 33–39.

17. Kaplan, D. L.; Fossey, S. A.; Viney, C.; Muller, W. S. In *Hierarchically Structured Materials*; Aksay, I. A.; Baer, E.; Sarikaya, M.; Tirrell, D. A., Eds., 1992; Vol. 255, pp. 19–29.
18. Magoshi, J.; Magoshi, Y.; Nakamura, S. *Journal of Applied Polymer Science: Applied Polymer Symposium* **1985**, *41*, 187–204.
19. Magoshi, J.; Magoshi, Y.; Nakamura, S. *Silk Polymers* **1994**, *544*, 292–310.
20. Ratho, T.; Patel, A. *J. Polym. Sci. Pol. Chem*. **1975**, *13*, 2345–2352.
21. Thiel, B. L.; Kunkel, D. D.; Viney, C. *Biopolymers* **1994**, *34*, 1089–1097.
22. Trabbic, K. A.; Thiel, B. L.; Gillespie, D. B.; Viney, C.; Yager, P. *Biophys. J*. **1994**, *66*, A294–A294.
23. Gathercole, L. J.; Barnard, K.; Atkins, E. D. T. *Int. J. Biol. Macromol*. **1989**, *11*, 335–338.
24. Gathercole, L. J.; Keller, A.; Shah, J. S. *J Microsc* **1974**, *102*, 95–105.
25. Beck, K.; Brodsky, B. *J Struct Biol* **1998**, *122*, 17–29.
26. Giraud-Guille, M. M. *Microsc Res Tech* **1994**, *27*, 420–428.
27. Martin, R.; Farjanel, J.; Eichenberger, D.; Colige, A.; Kessler, E.; Hulmes, D. J. S.; Giraud-Guille, M. M. *J. Mol. Biol*. **2000**, *301*, 11–17.
28. Valluzzi, R.; Kaplan, D.; Probst, W.; Zellmann, E.; Jacksch, H. *Soft Mater*. **2003**, *In Press*.
29. Valluzzi, R.; Kaplan, D. L. *Macromolecules* **2003**, *In Press*.
30. Valluzzi, R.; Kaplan, D. L. *Biopolymers* **2000**, *53*, 350–362.
31. Valluzzi, R.; He, S. J.; Gido, S. P.; Kaplan, D. *Int. J. Biol. Macromol*. **1999**, *24*, 227–236.
32. Valluzzi, R.; Gido, S. P.; Muller, W.; Kaplan, D. L. *Int. J. Biol. Macromol*. **1999**, *24*, 237–242.
33. Valluzzi, R.; Gido, S.; Zhang, W.; Muller, W.; Kaplan, D. *Macromolecules* **1996**, *29*, 8606–8614.
34. Wilson, D.; Valluzzi, R.; Kaplan, D. *Biophys. J*. **2000**, *78*, 2690–2701.
35. Willcox, P. J.; Gido, S. P.; Muller, W.; Kaplan, D. L. *Macromolecules* **1996**, *29*, 5106–5110.
36. Tsukada, M.; Nagura, M.; Ishikawa, H. In *Proceedings of the 7th International Wool Textile Research Conference*: Tokyo, 1985; Vol. 1, pp 383–388.
37. Thiel, B. L.; Guess, K. B.; Viney, C. *Biopolymers* **1997**, *41*, 703–719.
38. Lotz, B.; Cesari, C. *Biochimie* **1979**, *61*, 205.
39. Lotz, B.; Gonthiervassal, A.; Brack, A.; Magoshi, J. *J. Mol. Biol*. **1982**, *156*, 345–357.
40. Kamien, R. D. *Molecular Crystals and Liquid Crystals* **2001**, *358*, 97–101.
41. Harris, A. B.; Kamien, R. D.; Lubensky, T. C. *Rev. Mod. Phys*. **1999**, *71*, 1745–1757.
42. He, S. J.; Valluzzi, R.; Gido, S. P. *Int. J. Biol. Macromol*. **1999**, *24*, 187–195.
43. Cunniff, P. M.; Fossey, S. A.; Auerbach, M. A.; Song, J. W.; Kaplan, D. L.; Adams, W. W.; Eby, R. K.; Mahoney, D.; Vezie, D. L. *Poly. Adv. Technol*. **1994a**, *5*, 401–410.
44. Gaill, F.; Herbage, D.; Lepescheux, L. *Matrix* **1991**, *11*, 197–205.
45. Hinman, M. B.; Stauffer, S. L.; Lewis, R. V. In *Silk Polymers*; D. Kaplan, W. W. A., B. Farmer, C. Viney, Ed.; American Chemical Society: Washington D. C., 1994; Vol. 544, p. 222.
46. Knight, D.; Vollrath, F. *Tissue Cell* **1999**, *31*, 617–620.
47. Knight, D. P.; Vollrath, F. *Philos. Trans. R. Soc. Lond. Ser. B-Biol. Sci*. **2002**, *357*, 155–163.
48. Parry, D. A. D. *Biophys. Chem*. **1988**, *29*, 195–209.
49. Viney, C. *Supramol. Sci*. **1997**, *4*, 75–81.
50. Yang, Z. T.; Liivak, O.; Seidel, A.; LaVerde, G.; Zax, D. B.; Jelinski, L. W. *J. Am. Chem. Soc*. **2000**, *122*, 9019–9025.

51. Levi, Y.; Albeck, S.; Brack, A.; Weiner, S.; Addadi, L. *Chem-Eur J* **1998**, *4*, 389–396.
52. Litvin, A. L.; A.Samuelson, L.; Charych, D. H.; Spevak, W.; Kaplan, D. L. *J. Phys. Chem*. **1995**, *99*, 12065–12068.
53. Valluzzi, R.; Cebe, P.; Kaplan, D. L. *Science* **2003**, *In preparation*.
54. Valluzzi, R.; Kaplan, D. L.; Cebe, P.; Guertin, R. *Abstr. Pap. Am. Chem. Soc*. **2002**, *224*, 527-POLY.
55. Valluzzi, R.; Szela, S.; Avtges, P.; Kirschner, D.; Kaplan, D. *J. Phys. Chem. B* **1999**, *103*, 11382–11392.
56. Winkler, S.; Szela, S.; Avtges, P.; Valluzzi, R.; Kerschner, D.; Kaplan, D. L. *Int. J. Biol. Mol*. **1999**, *24*, 265.
57. Szela, S.; Avtges, P.; Valluzzi, R.; Winkler, S.; Wilson, D.; Kirschner, D.; Kaplan, D. L. *Biomacromolecules* **2000**, *1*, 534–542.
58. Cebe, P.; Georgiev, G.; Valluzzi, R.; Kaplan, D. *Abstr. Pap. Am. Chem. Soc*. **2002**, *224*, 26-BTEC.
59. Mita, K.; Ichimura, S.; James, T. C. *J Mol Evol* **1994**, *38*, 583–592.
60. Gillespie, D. B.; Thiel, B. L.; Trabbic, K. A.; Viney, C.; Yager, P. *Macromolecules* **1994**, *27*, 6177–6182.
61. Goeden-Wood, N. L.; Conticello, V. P.; Muller, S. J.; Keasling, J. D. *Biomacromolecules* **2002**, *3*, 874–879.
62. Valluzzi, R.; Winkler, S.; Wilson, D.; Kaplan, D. L. *Philos. Trans. R. Soc. Lond. Ser. B-Biol. Sci*. **2002**, *357*, 165–167.
63. Winkler, S.; Szela, S.; Avtges, P.; Valluzzi, R.; Kirschner, D. A.; Kaplan, D. *Int. J. Biol. Macromol*. **1999**, *24*, 265–270.
64. Thiel, B. L.; Viney, C. *J. Microsc.-Oxf*. **1997**, *185*, 179–187.
65. Kaplan, D.; Adams, W. W.; Farmer, B.; Viney, C. In *Silk Polymers*; AMER CHEMICAL SOC: Washington, 1994; Vol. 544, pp. 2–16.
66. Kenney, J. M.; Knight, D.; Wise, M. J.; Vollrath, F. *Eur J Biochem* **2002**, *269*, 4159–4163.
67. Rathore, O.; Winningham, M. J.; Sogah, D. Y. *J. Polym. Sci. Pol. Chem*. **2000**, *38*, 352–366.
68. Nemethy, G.; Scheraga, H. A. *Biopolymers* **1965**, *3*, 155-&.
69. Rudall, K. M., Ed. *Protein Ribbons and Sheets*, 1956; Vol. 5.
70. Elliott, A.; Malcolm, B. R. *Biochimica Et Biophysica Acta* **1956**, *21*, 466–471.
71. Kratky, O. *Transactions of the Faraday Society* **1956**, *52*, 558–570.
72. Lucas, F.; Shaw, J. T. B.; Smith, S. G. *Nature* **1956**, *178*, 861–861.
73. Kendrew, J. C.; Perutz, M. F. *Annu. Rev. Biochem*. **1957**, *26*, 327–372.
74. Lucas, F.; Shaw, J. T. B.; Smith, S. G. *Adv. Protein Chem*. **1958**, *13*, 107–242.
75. Traub, W.; Yonath, A.; Segal, D. M. *Nature* **1969**, *221*, 914–917.
76. Venkatac.Cm; Ramachan.Gn. *Annu. Rev. Biochem*. **1969**, *38*, 45-&.
77. Doyle, B. B.; Traub, W.; Lorenzi, G. P.; Brown, F. R.; Blout, E. R. *J. Mol. Biol*. **1970**, *51*, 47-&.
78. Doyle, B. B.; Traub, W.; Lorenzi, G. P.; Blout, E. R. *Biochemistry* **1971**, *10*, 3052-&.
79. Mattice, W. L.; Mandelke. L. *Biochemistry* **1971**, *10*, 1926-&.
80. Ananthan. Vs. *J. Sci. Ind. Res*. **1972**, *31*, 593–612.
81. Anderson, J. M.; Chen, H. H.; Walton, A. G.; Rippon, W. B. *J. Mol. Biol*. **1972**, *67*, 459-&.
82. Brack, A.; Caille, A.; Spach, G. *Mon. Chem*. **1972**, *103*, 1604–1612.
83. Brack, A.; Caille, A. *Int. J. Pept. Protein Res*. **1978**, *11*, 128–139.
84. Lekkerkerker, H. N. W.; Vroege, G. J. *Philos T Roy Soc A* **1993**, *344*, 419–440.
85. Muller, W. S.; Samuelson, L. A.; Fossey, S. A.; Kaplan, D. L. *Langmuir* **1993**, *9*, 1857–1861.

86. Valluzzi, R.; Atvges, P.; Winkler, S.; Szela, S.; He, S. J.; Gido, S. P.; Kaplan, D. L. *Abstr. Pap. Am. Chem. Soc.* **1998**, *216*, 077-BTEC.
87. Gido, S. P.; Valluzzi, R.; He, S. J. *Abstr. Pap. Am. Chem. Soc.* **1998**, *216*, 466-PHYS.
88. Valluzzi, R.; He, S. J.; Gido, S. P.; Kaplan, D. *IJBM* **1999**, *In Press*.
89. Viney, C.; Huber, A. E.; Dunaway, D. L.; Kerkham, K.; Case, S. T. In *Silk Polymers*; D. Kaplan, W. W. A., B. Farmer, C. Viney, Ed.; American Chemical Society: Washington D. C., 1994; Vol. 544, p. 120.
90. Muller, W. S.; Samuelson, L. A.; Fossey, S. A.; Kaplan, D. In *Silk Polymers: Materials Science and Biotechnology*; Kaplan, D.; Adams, W. W.; Farmer, B.; Viney, C., Eds.; American Chemical Society: Washington, D. C., 1994; pp 342–352.
91. Kerkam, K.; Viney, C.; Kaplan, D.; Lombardi, S. *Nature* **1991**, *349*, 596–598.
92. Giraud-Guille, M. M. *Molecular Crystals and Liquid Crystals* **1987**, *153*, 15–30.
93. Bouligan. Y. *Tissue Cell* **1972**, *4*, 189-&.
94. Iizuka, E. *Adv Biophys* **1988**, *24*, 1–56.
95. Magoshi, J.; Nakamura, S. *Acs Sym Ser* **1992**, *489*, 231–248.
96. Knight, D. P.; Feng, D.; Stewart, M.; King, E. *Philos. Trans. R. Soc. Lond. Ser. B-Biol. Sci.* **1993**, *341*, 419–436.
97. Neville, A. C. *Biology of Fibrous Composites: Development Beyond the Cell Membrane*; Cambridge University Press: Cambridge, U.K., 1993.
98. Borgstrom, J.; Quist, P. O.; Piculell, L. *Macromolecules* **1996**, *29*, 5926–5933.
99. Giraud-Guille, M. M. *Int Rev Cytol* **1996**, *166*, 59–101.
100. Das, P.; Roy, J.; Chakrabarti, N.; Basu, S.; Das, U. *J Chem Phys* **2002**, *116*, 9028–9035.
101. Hwang, J. J.; Iyer, S. N.; Li, L. S.; Claussen, R.; Harrington, D. A.; Stupp, S. I. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 9662–9667.
102. Vollrath, F.; Knight, D. P. *Nature* **2001**, *410*, 541–548.
103. Giraudguille, M. M. *J. Mol. Biol.* **1992**, *224*, 861–873.
104. Besseau, L.; Giraudguille, M. M. *J. Mol. Biol.* **1995**, *251*, 197–202.
105. Valluzzi, R.; Bini, E.; Haas, T.; Cebe, P.; Kaplan, D. L. In *NASA MIcrogravity and Materials Science Conference*; NASA: Hunsville, Alabama, 2002; In Press.
106. Giraudguille, M. M. *Biol. Cell* **1989**, *67*, 97–101.
107. Bella, J.; Brodsky, B.; Berman, H. M. *Structure* **1995**, *3*, 893–906.
108. Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. M. *Science* **1994**, *266*, 75–81.
109. Valluzzi, R.; Jin, H. J.; Haas, T., Ed.: Medford, 2002.
110. Kratky, O.; Schauenstein, E.; Sekora, A. *Nature* **1950**, *165*, 527–528.
111. Kratky, O.; Schauenstein, E.; Sekora, A. *Nature* **1950**, *165*, 319–320.
112. Kratky, O.; Schauenstein, E.; Sekora, A.; Lapp, I.; Berndt, W. *Nature* **1951**, *167*, 861–862.
113. Kratky, O. *Transactions of the Faraday Society* **1954**, *50*, 1013–1014.
114. Kratky, O.; Schauenstein, E. *Discussions of the Faraday Society* **1951**, 171-&.
115. Kratky, O.; Schauenstein, E.; Sekora, A. *Nature* **1950**, *166*, 1031–1032.
116. Valluzzi, R.; Gido, S. P. *Biopolymers* **1997**, *42*, 705–717.
117. Valluzzi, R.; Guertin, R. P.; Haas, T. E.; D., P.; Kriss, F.; NI, C.-Y. *Chem Materials* **2003**, *In Preparation*.
118. Kuzuhara, A.; Asakura, T.; Tomoda, R.; Matsunaga, T. *J. Biotechnol.* **1987**, *5*, 199–207.
119. Ishida, M.; Asakura, T.; Yokoi, M.; Saito, H. *Macromolecules* **1990**, *23*, 88–94.
120. Li, G. X.; Yu, T. Y. *Makromol Chem-Rapid* **1989**, *10*, 387–389.
121. Inoue, S.; Tanaka, K.; Arisaka, F.; Kimura, S.; Ohtomo, K.; Mizuno, S. *Journal of Biological Chemistry* **2000**, *275*, 40517–40528.

122. Kerkam, K.; Viney, C.; Kaplan, D. L.; Lombardi, S. J. *Nature* **1991**, *349*, 596–598.
123. Winkler, S.; Szela, S.; Avtges, P.; Valluzzi, R.; Kirschner, D.; Kaplan, D. L. *Int. J. Biol. Macro*. **1999**, *24*, 265–270.
124. Sprague, K. U. *Biochemistry* **1975**, *14*, 925–931.
125. Gamo, T.; Inokuchi, T.; Laufer, H. *Insect Biochemnistry* **1977**, *7*, 285–295.
126. Okamoto, H.; Ishiwaka, E.; Suzuki, Y. *Journal of Biological Chemistry* **1982**, *257*, 15192–15199.
127. Tsukada, M.; Komoto, T.; Hirabayashi, K. *Sen-i Gakkaishi* **1983**, *39*, T227–232.
128. Freddi, G.; Svilokos, A. B.; Ishikawa, H.; Tsukada, M. *J. Appl. Polym. Sci*. **1993**, *48*, 99–106.
129. Mercer, E. H.; Meyer, F. H. *Textile Research Journal* **1953**, *23*, 243–246.
130. Hirabayashi, K.; Tsukada, M.; Ishikawa, H.; Yasumura, S. *Sen-i Gakkaishi* **1974**, *30*, T459–464.
131. Kataoka, K. *Kobunshi Robunshu, Eng. Ed*. **1977**, *6*, 1–9.
132. Hirabayashi, K.; Arai, M. *Journal of Sericultural Science Japan* **1988**, *58*, 81–82.

Chapter 11

NANOBIOTECHNOLOGY ENABLES NEW OPPORTUNITIES IN MATERIAL SCIENCES: BACTERIORHODOPSIN AS A FIRST EXAMPLE

Norbert Hampp
University of Marburg, Faculty of Chemistry & Material Science Center D-35032 Marburg, Germany

Abstract: In Nanobiotechnology the combination of rational design of biomolecular functions and biotechnological syntheses may lead to a new class of materials not only mimicing biomolecules but going beyond all design and synthetic capabilities existing today. Nature supplies the blueprints and functional modules which may be used as building blocks for the design of new functions leading to new applications, even far out from the established fields for biomolecules like diagnostics, medicine and food industry. An example is described here. Using bacteriorhodopsin as a test bed the development from a biomolecule into a functional biomolecular material for technical uses, here security applications, is reviewed.
Key words: Bacteriorhodopsin, security features, photochromism, data storage

1. STATE-OF-THE-ART IN MATERIAL SCIENCES IS THE DESIGN OF MATERIAL PROPERTIES

The design of material functions is the challenge of this century. Nanoscaled structures, which react upon excitation in a very specific way, characterize this ultimate goal in material science. Some examples from nature are molecular motors, like the F1-ATPase, which rotates as soon as the proper energy form is available, or the photoreaction center that captures and transforms light energy. Semiconductor technology and supramolecular chemistry are both heading to generate artificial functional nanoscaled structures. The borders between disciplines like biochemistry, macromolecular chemistry, and semiconductor research disappear on the nanoscale.

V. Renugopalakrishnan et al. (eds.), Bionanotechnology, 209-216.

2. NATURE'S BLUEPRINTS—RESOURCES FOR THE FUTURE

Today, functional nanostructures are the domain of biology. Every little cell is stuffed with a huge variety of fascinating efficient molecular systems. Methods like X-ray diffraction and nuclear magnetic resonance (NMR) have taught us much about the structures of these molecules and supported by molecular dynamics modeling, we are beginning to understand how they function. May be the relation to material science is not immediately obvious. Of course, medical uses of these insights in new and more powerful therapies potentially helping to gain victory over cancer and other diseases are the first associations that come into mind. But from a materials perspective, we recognize that, e.g., the smallest motor, the strongest material, the most selective sensor, the best solar energy converter, all were discovered in the biosphere. These efficient "nanomachines" are made of amino acids as biological building blocks, which commonly are not expected to be that powerful as far as their physical properties are considered. All these "nanoscaled treasures" have been there for millions of years; however, our scientific tools were not developed far enough to recognize their performance until recently. This has changed now. Headlines come into mind: molecular motors,[1] DNA-computing,[2] self-reproducing molecules,[3] neuron-semiconductor junctions,[4] single molecule sensors,[5] not to mention the analysis of the photoreaction center,[6] and the ribosomal protein synthesis machine.[7] The dramatically increased knowledge about the function of biomolecules led to the birth of a new research direction named nanobiotechnology.

What is the philosophy behind? The only way for the design and realization of functions on the nanoscale is the use of biological systems so far. The production of such materials is accomplished indirectly. An organism is genetically altered in order to produce the compound of interest and then amplified by conventional biotechnology. The desired product is isolated and may be further modified for final use or integration into technical products. As this seems to be an efficient and powerful method, a dramatic increase of the economic importance of nanobiotechnology is expected. Because it is impossible to give an overview on the whole research area in such a short article, one example is selected—the molecule bacteriorhodopsin (BR).

3. BACTERIORHODOPSIN—A LIGHT ENERGY CONVERTER

Since several years we develop technical applications of the BR (Fig. 1). The natural function of the BR is that of a light-energy converter. It is the key

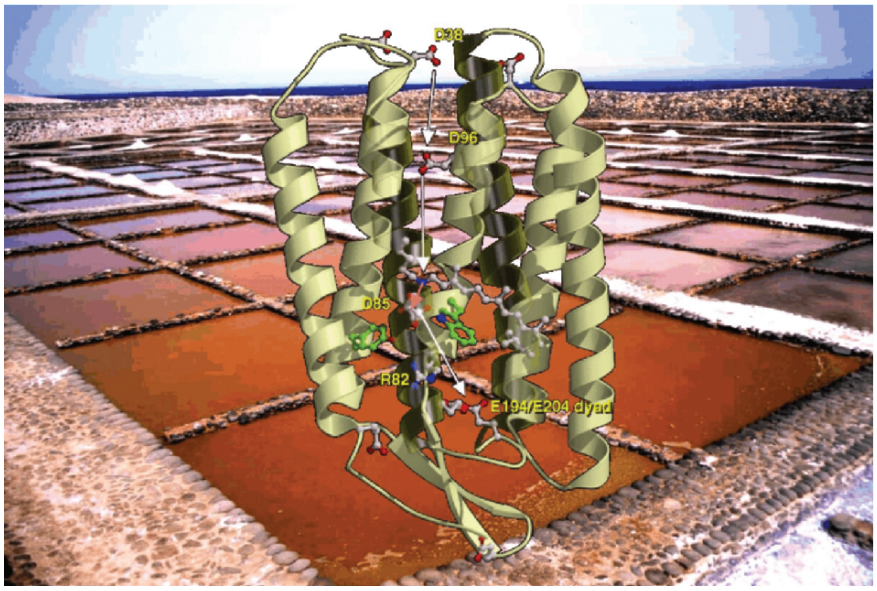


Figure 1. Salt lakes on the Canary Islands. The purple color is caused by halobacteria, in particular by the BR produced in their cell membrane. A model of its 3D structure and the proton transport pathway is shown.

molecule in the halobacterial photosynthetic apparatus that is extremely simple as it comprises only two molecules: a BR and a membrane-bound ATPase.[8] In the halobacterial cell, the BR molecule acts as a light-driven proton pump, which upon capture of a photon, transports a proton from the inner of the cell to the outer medium. Thereby, it converts light energy into chemical energy by a complex molecular process that is well understood. The transport of the proton in the wildtype takes about 10 ms. A transient photovoltage of about 250 mV per molecular layer creates a photochromic color change. Each of these features may be used in different technical applications. And there is something that makes the BR somewhat different—it occurs in the halobacterial cell membrane in the form of a two-dimensional (2D) crystalline lattice, the so-called purple membrane. In this form the BR is astonishingly stable, which makes its technical use much easier and thus a first choice for such an attempt.[9]

4. MUTATED BACTERIORHODOPSINS FOR SECURITY APPLICATIONS AND DATA STORAGE

Technical applications of the BR are investigated by several groups in the world. We focus on the photochromic properties of the material. Constantly,

new applications have been developed. The first was a holographic correlator, an optical system for the comparison of images.[9] For that purpose only a few milligrams of the material embedded into tiny polymeric films was required. Years later, much larger BR-films with a significantly improved light sensitivity led to the development of a holographic system for nondestructive testing.[9] New biotechnological procedures have been developed, which allow to produce the BR in large quantities at reasonable costs, and this paved the way to new applications. The BR-based inks have been developed, which can be applied to almost every material (e.g., paper and plastics) by conventional printing technologies. The resulting structures combine three different functions on the molecular level. First, it is a photochromic element. On exposure to light it changes its color (Fig. 2). This property can be visually inspected without any technical equipment. In addition, the material may be used as an optical data storage, which enables to record considerable amounts of digital data (Fig. 3). Of course, a terminal is needed to write and read data, but the application is very flexible. The data written into the BR-material are non-erasable, which further prevents counterfeiting attempts. Finally, each of the BR-molecules carries a molecular tag that identifies a certain production lot. In case there is some misuse detected, the material can be traced back to its origin. This combination of properties makes the BR a unique material for security applications. For decorative purposes, the material can be modified to have other initial colors, e.g., blue, red, or green. The 1000-fold increased light-sensitivity, the data storage capabilities, and the molecular traceability tags are all designed into the material by genetic engineering of the wildtype form (Fig. 4).

How far has this developed? The fundamental problems have been solved, and most of the lab work has been done. MIB, a start-up company focusing on the BR-based products, has been established. They implemented in a first step a production in the range of 1–2 kg/year. In parallel field-tests with ID cards using this new technology have been started with industrial partners. For each ID card, about 1–5 mg is required depending on the size and data capacity of the BR-feature. The current production capacity is sufficient for a few million samples a year but has to be further scaled up in the near future in order to reach technically relevant quantities.

5. BACTERIORHODOPSIN—THE FIRST EXAMPLE OF A NANOBIOTECHNOLOGY PRODUCT

The example given shortly summarizes the identification and generation of a bionanotechnology product. The first step is to identify a biomaterial that has a technically interesting property. Then the analysis of the structure–function

Figure 2. ID card sample in which the symbol "D" is printed with a BR-based ink. Top—initial color after storage; bottom—color after flashing with light.

relation is a must before any modifications can be designed. This process is currently the bottleneck. Tuning of the physical properties/functions of a biomolecule cannot be done by conventional methods, e.g., statistical mutagenesis and selection procedures. The reason why this approach fails is that there is no possibility to establish a selection system for physical properties. The only way, at least today, is rational design of molecular modifications. The genetic alteration of organisms is well established. For the biotechnological production, the transfer of the desired gene into one of the standard organisms like *E. coli* is desired. However, this is not possible in the case of the BR. Because the 2D crystalline lattice of the purple membrane, which stabilizes

Figure 3. Optical data storage in an ID card sample made from BR-inks. In the microscope and on the screen, the stored seal of the University of Marburg is seen.

Figure 4. From the clone to the product. Clones of *H. salinarium* on the petri dish and isolated BR in the Erlenmeyer flask. Lyophilized BR and variants with different colors are shown in the sample vials. An ID card, as shown in Fig. 2, made using BR-based inks is positioned in the middle.

the BR to a great extent, is only formed by halobacteria that have to be used also for the production. This means cultivation in concentrated salt solutions. Everybody can imagine that several technical problems had to be solved before such a production could be set up reliably. The isolated material is then integrated into specially developed inks, which also require a special processing in order to preserve the properties of the embedded protein.

6. SUMMARY

Is the example given unique or representative for a wide variety of new materials to appear? We should be optimistic because the same approach

nature uses during evolution. Humans carry three different rhodopsins in their eyes, which show sensitivities to blue, green, and red wavelengths. They are different in their physical properties. Each of the rhodopsins comprises a light-antenna pigment and an amino acid cage in which the antenna pigment is located inside. All three rhodopsins have an identical light-antenna pigment, retinal. The differences in the physical properties of the rhodopsins are caused by changes in the amino acids interacting with the retinal molecule. The tuning of the physical properties in this way is very effective. Not three different antenna pigments need to be synthesized but just one. To generate three different amino acid cages for the same pigment is a more flexible approach—a so-called platform technology. Using a reduced set of chemical structures, the 20 amino acids, in principle every molecular function may be designed. Until mathematical models have been developed, which allow a rational design of a molecular function from scratch, we have to use functional biomolecules as blueprints and adapt their physical functions to the needs of technical applications—thereby taking benefit from millions of years of evolution distilled into each of the biomolecular system—the most developed nanomachines we have on earth today.

ACKNOWLEDGEMENTS

This work was supported by the Federal Ministry of Education and Research through grant 13N8196.

REFERENCES

1. Junge, W., 1999, ATP synthase and other motor proteins, *Proc. Natl. Acad. Sci. USA* **96:**4735–4737.
2. Adleman, L. M., 1994, Molecular computation of solutions to combinatorial problems, *Science* **266:**1021–1024.
3. Lee, D. H., Granja, J. R., Martinez, J. A., Severin, K., and Ghadiri, M. R., 1996, A self-replicating peptide, *Nature* **382:**525–528.
4. Fromherz, P., 2002, Electrical interfacing of nerve cells, *Chem. Phys. Chem*. **3:**276–284.
5. Bayley, H., and Cremer, P. S., 2001, Stochastic sensors inspired by nature, *Nature* **413:**226–230.
6. Deisenhofer, J., Michel, H., and Huber, R., 1985, The structural basis of photosynthetic light reactions in bacteria, *Trends Biochem. Sci*. **10:**243–283.
7. Harms, J., Schluenzen, F., Zarivach, R., Bashan, A., Gat, S., Agmon, I., Bartels, H., Franceschi, F., and Yonath, A., 2001, High resolution structure of the large ribosomal subunit from a mesophilic eubacterium, *Cell*, **107:**679–688.
8. Oesterhelt, D., and Stoeckenius, W., 1971, Rhodopsin-like protein from the purple membrane of *Halobacterium halobium*, *Nature New Biol*. **233:**149–152.
9. Hampp, N., 2000, Bacteriorhodopsin as a photochromic retinal protein for optical memories, *Chem. Rev*. **100:**1755–1776.

Chapter 12

SYNTHETIC PHOTOREFRACTIVE AND PHOTOCHROMIC MATERIALS AND THEIR COMPARISON WITH BACTERIORHODOPSIN MUTANTS FOR OPTICAL INFORMATION PROCESSING

A. Miniewicz[1]* and V. Renugopalakrishnan[2]
[1]*Institute of Physical and Theoretical Chemistry, Wroclaw University of Technology, Wybrzeze Wyspianskiego 27, 50–370 Wroclaw, Poland*
*Address correspondence to: Prof. A. Miniewicz *andrzej.miniewicz@pwr.wroc.pl*
[2]*Children's Hospital, Harvard Medical School, Boston, MA 02215, USA*

Key words: bacteriorhodopsin, photorefractive polymers, photochromic materials, optical information storage, optical information processing

Abbreviations: bacteriorhodopsin (bR), liquid crystal (LC), polymer-dispersed liquid crystal (PDLC), computer-generated hologram (CGH), optically addressed liquid crystal spatial light modulator (OA LC SLM).

Abstract: Encoding, manipulation and retrieval of optical information in series or in parallel is of pivotal importance in optical computing and photonics. Strong and fast interaction of light with matter is therefore highly desirable for producing reversible holographic media for information storage and/or processing. Widespread, well developed inorganic materials may loose their leading position to organics whose rising potential for photonic devices is envisioned e.g. liquid crystals for display applications. In this chapter we describe the main features of photorefractivity and photochromism in polymers, liquid crystals and bacteriorhodopsin (bR) and its mutants designed by rational site-directed mutagenesis. Our goal was to provide a glimpse of complexity of phenomena occuring in optical media leading to conversion of spatial light intensity or polarization patterns into respective chemical and/or physical changes.

We have discussed details of classic photorefractive effect which is the main effect allowing for volumetric optical storage in inorganic crystals. By chemical synthesis and sophisticated material engineering a whole class of efficient polymeric photorefractive materials have been reported in the

V. Renugopalakrishnan et al. (eds.), Bionanotechnology, 217-256.

literature. Despite extensive research on photorefractive polymers it is increasingly felt that they fall short of ideal properties of bR and its mutants which are promising candidates as ideal material for information storage. Next we describe the emerging field of dynamic holographic materials basing on low-molecular mass liquid crystals. The photorefractive effect observed in this class of materials differs from the classic one, nevertheless gives new options for spatial light modulation and parallel optical processing. We have discussed details of the physics underlying such interesting real-time holographic properties of liquid crystals and their limitations.

Further we briefly describe photochromic effect and organic photochromic materials currently studied with emphasis to azo-functionalized polymers in which photoisomerization of chromophores is responsible for the light-induced anisotropy. Again this group of materials brings some novel features i.e. relief grating formation as well as exhibit drawbacks. Following that we move toward photochromic properties and branched photocycle of bR. Later in the chapter we have focused attention on optimization of bR properties by rational site-directed mutagenesis. A combination of theoretical analysis of the energy level diagram of wild-type bacteriorhodopsin and genetic engineering methods lead to the design of thermally stable bR mutants with perfect holographic properties.

The goal of the chapter is also to explore the application of photorefractive polymers, liquid crystals and bR mutants for real-time holography and high-density holographic memories. These demanding applications need optimization of physical properties of materials, stability, processability and low cost which are discussed at the end of this chapter. Comparison of bR with other synthetic organic materials is not straightforward due to fact that each group of materials is optimized toward precisely defined application. However, superior capabilities of bR family over synthetic materials seems to be unsurpassable. Biological macromolecules yield themselves to extensive modification of their properties by biotechnology and the future developments in the design of biological macromolecules with tailored properties will require a multi-disciplinary approach utilizing the arsenal of methods from the rapidly developing field of bionanotechnology.

1. INTRODUCTION

The growing demands of generation, accumulation, access and archiving of electronic data require novel concepts for processing and storage of information. Conventional electronic and magnetic surface-storage techniques have reached their physical limitations. Progress in nonlinear optics and material science hold much promise in ushering in materials and methods for storing vastly enhanced data storage and processing. For instance, it seems

that volumetric approaches such as optical holography which offers the storage density of 1 Tbit/cm^3 on a single CD-sized disk would be the choice of next generation memory devices. Photonic applications like optical signal processing and optical storage require optical materials with fast and strong response to incident light. Many materials e.g. among them photorefractive crystals, polymers, photopolymers and biomaterials, manifest interesting linear optical and non-linear optical properties can be exploited for processing, transport or storing information with the help of photons—the fastest carriers. Classic photorefractive properties[1] (e.g. mapping of light intensity pattern into refractive index pattern of the material) can now be achieved in specially designed and treated polymeric systems.[2] However, other approaches are equally interesting namely those exploiting molecular optical bistability – photochromism[3] (e.g. photochromic proteins[4]) or the emerging branch of photonic band gap materials[5,6] which offer numerous novel phenomena that can be used for light processing.

Organic polymeric systems supplemented or doped either by photochromic moieties or combined with liquid crystalline systems are emerging as promising materials for advanced information and communication technology. Polymers—because of their structural flexibility, easy chemical modification and low-cost fabrication, liquid crystals—because they combine optical anisotropic properties (birefringence) with the property of being soft, which means that the anisotropy can easily be influenced by external fields and finally photochromic systems which by nature reversibly change their absorptive and refractive properties upon photon absorption. We will briefly discuss chosen mechanisms and materials used for optical signal processing and storage and compare them with bacteriorhodopsin (bR), a protein produced by halobacteria.[7] bR crafted by nature shows astonishing stability toward chemical and thermal degradation and its photosensitivity and cyclicity to illumination is far beyond that of synthetic materials. Moreover bR can be genetically modified to a large extent forming a platform for the whole class of materials.[8–10] Several different types of utilisation of bR have been demonstrated, most basing on its transient or dynamic holographic properties.[9] These include: reversible holographic media, spatial light modulators, holographic optical correlators and pattern recognition systems, nonlinear optical filters, associative memories, dynamic time-average interferometers, neural type logic gates, multilevel logic gates, thin film memories, volumetric memories, optical computing, photon counters and photovoltaic converters.[10]

In the second part of this chapter the emphasis will be placed on the technical applications of bR, liquid crystalline systems and photorefractive and/or photochromic polymers as transient holographic materials.

2. EFFECTS EXPLOITED FOR OPTICAL INFORMATION PROCESSING IN ORGANIC MATERIALS

In order to optically record a bit of information in a material a physical and/or chemical change of this material influencing the complex refractive index should occur. Therefore a suitable description of such a process can be done with the use of formalism of nonlinear optics for dielectric nonmagnetic materials. Magnetic materials must be treated separately. Within the dipolar approximation formalism one can describe the light–matter interaction using expansion of a polarization vector $\vec{P}$ of a medium into series of electric field $\vec{E}$ of interacting light waves:[11]

$$P_i = \varepsilon_0(\chi_{ij}^{(1)}E_j + \chi_{ijk}^{(2)}E_jE_k + \chi_{ijkl}^{(3)}E_jE_kE_l + \cdots) \tag{1}$$

where tensors $\chi_{ij}^{(1)}$, $\chi_{ijk}^{(2)}$, $\chi_{ijkl}^{(3)}$ known as optical susceptibilities of first, second and third order describe the macroscopic material properties. The complex $\chi_{ij}^{(1)}$ linear susceptibility which is related to the dielectric constant and refractive index by

$$n^2 = \varepsilon/\varepsilon_0 = 1 + \chi_{ij} \tag{2}$$

describes both changes of material real refractive index n as well as its absorption coefficient α. Any permanent or temporary light induced changes influence the susceptibility tensors $\chi_{ij}^{(1)}$, $\chi_{ijk}^{(2)}$, $\chi_{ijkl}^{(3)}$ of a material and determine the light - matter interactions. Generally organic materials (molecular crystals and polymers) are composed of weakly interacting molecules described by their local field $\vec{F}$ dependent dipolar moments:

$$\mu_i = \mu_{0i} + \alpha_{ij}F_j + \beta_{ijk}F_jF_k + \gamma_{ijkl}F_jF_kF_l + \cdots \tag{3}$$

where μ_0 stands for permanent molecular dipole and α_{ij}, β_{ijk}, γ_{ijkl} are tensor elements of polarizability, first and second hyperpolarizability, respectively. All known optical effects used for information storage or processing such as linear absorption, two-photon absorption, linear electro-optic effect, dc Kerr effect, optical Kerr effect, light intensity I induced refractive index change $n(I) = n + n_2I$, optical phase conjugation, photorefraction, photochromism, etc. could be described within this formalism.

Two major categories of organic materials suitable for reversible optical write-read processes can be distinguished: photorefractive and photochromic ones (cf. Table 1).

Photorefraction refers to spatial modulation of the index of refraction n under non-uniform illumination via space-charge field formation and electro-optic

Table 1. Organic photorefractive and photochromic systems

Photorefractive materials[a]	**Photochromic materials**[b]
Polymer composites	Bacteriorhodopsin mutants
Organic amorphous glasses	Chromophore doped polymers
Fully functionalised materials	Chromophore functionalised polymers
Polymer dispersed liquid crystals PDLC	Liquid crystalline photochromic systems
Hybrid LC-photoconducting polymer	Surface command LC systems
Doped liquid crystals	Chirooptical molecular switches

[a]O. Ostroverkhova, W.E. Moerner, "Organic photorefractives: Mechanisms, Materials, and Application", Chemical Reviews, vol. 104, 3267–3314 (2004)
[b]Chemical Reviews, vol. 100(5), 2000 "Photochromism: Memories and Switches"

nonlinearity. Photochromism is defined as a reversible change induced by light irradiation, between two states of molecule having different absorption spectra $\alpha(\lambda)$.

We describe below more precisely the physics underlying the two above mentioned effects in order to discuss the differences in material performance with respect to optical information processing. We will limit ourselves to chosen materials only representing these phenomena with the emphasis to those studied by us.

2.1. Photorefractive effect

Classic photorefractive effect refers to a spatial modulation of refractive index in a material due to the light-induced redistribution of charges.[1] Since its first discovery, in inorganic non-centrosymmetric crystal of $LiNbO_3$ in 1966,[12] photorefractive effect has been recognized to be promising for various holographic applications due to the large and reversible modulation of refractive index. At present the photorefractivity term is used widely to the light induced change of refractive index, therefore we will recall below the fundamentals of the classic photorefractive effect.

Let us consider a photorefractive material illuminated with spatially modulated light intensity due to plane waves interference $I(x)$:[13]

$$I(x) = I_0\left(1 + m\cos\frac{2\pi x}{\Lambda}\right) \tag{4}$$

where $I_0 = I_1 + I_2$ is the mean intensity, fringe contrast is $m = \frac{2(I_1 I_2)^{1/2}}{I_1 + I_2}$ and Λ is the grating period. The four main physical processes (charge photogeneration, carrier diffusion, space charge formation and refractive index modulation $n(x)$ via Pockels effect are involved in the classic photorefractive phenomenon. It is

necessary to briefly describe them for understanding the complexity of other nonclassic photorefractive effects.

(i) Photogeneration of mobile charge

Absorption of a photon of a suitable energy at x leads in a photoconductive material to creation of a hole-electron pair. The photogeneration rate $G(x)$ is proportional to the light intensity as well as to the density of photosensitive centers. Photogeneration quantum efficiency in organic materials is strongly electric field-dependent and well described by the Marcus model.[14] Usually a constant electric field $\vec{E}_a$ is applied along the x-direction. Photoinduced charge is often provided by a molecule that absorbs light and injecting hole into the material, which generally is more mobile than electron. Most widely used generators are donor-acceptor charge transfer complexes, for example carbazole and trinitrofluoreneone (TNF) or fullerene C_{60}.

(ii) Diffusion and transport

Spatial distribution of electron and hole densities undergoes diffusion. Charge carrier transport in organics mainly occurs via hopping between neighboring molecules. The transport process in polymers is well described within the Baessler model[15] taking into account statistical distribution of hoping sites. For efficient photorefractive effect it is advantageous that mobilities μ of charge carriers of opposite sign differ significantly and charge carrier trapping is effective. In amorphous polymers the effective carrier mobility $\mu^{eff}(E)$ is electric field dependent and given by:[15]

$$\mu^{eff}(E) \propto \exp[-(E_A(E) - \beta_T E^{1/2}/k_B T] \tag{5}$$

where E_A is an activation energy and β_T is a constant positive below glass transition temperature T_g (β_T might be negative above T_g), k_B is a Boltzmann constant and T is temperature. The other physical process giving rise to charge transport is a drift in an external electric field and photovoltaic effect. Usually the drift of charges is a dominant mechanism in photorefractive polymers.

(iii) Space charge

Charge carriers created by light are retrapped in another place thus leading to appearance of distribution of space charge $\rho_{SC}(x)$. The presence of trapping sites is important for preventing carriers participation in transport for some period of time. The lifetime of the carrier in the trap is determined by

the trap depth as compared to thermal energies. Non-uniform distribution of charge results in appearance of space charge field $E_{SC}(x)$, which can be determined from the equality of drift and diffusion currents (in one charge carrier (hole) approximation):

$$J(x) = e\mu_h n_h(x) E_{SC}(x) - k_B T \mu_h \frac{dn_h}{dx} = 0 \tag{6}$$

where μ_h is a hole mobility, $n_h(x)$ is the density of mobile holes. Within this approximation Poisson's equation of electrostatics describes the space charge field:

$$E_{SC}(x) = \frac{k_B T}{e} \frac{1}{n_h(x)} \frac{dn_h}{dx}. \tag{7}$$

(iv) Refractive index modulation via Pockels effect (the effect is nonzero only for noncentrosymmetric materials)

For material showing linear electro-optic properties (the Pockels effect is described by a third rank tensor r_{ijk}, linked directly with $\chi^{(2)}_{ijk}$ cf. Eq. 1) the internal electric filed modifies the material's refractive index n:[1]

$$\Delta n(x) = -\frac{1}{2} n^3 r_{eff} E_{SC}(x) \tag{8}$$

where r_{eff} is an effective Pockels coefficient (projection of respective tensor elements on given direction). A field-dependent refractive index can also occur by a quadratic or Kerr orientational effect. An electric field dependent refractive index in organics is usually provided by exploiting molecular functionalities such as high ground-state dipole moment and large linear polarizability anisotropy (molecular birefringence). Molecules called NLO chromophores are added to polymer matrix at high concentration.

The scheme of light induced refractive index modulation is shown in Fig. 1a) together with typical two-beam coupling (2BC) experiment Fig. 1b) facilitating the study of photorefractive properties of polymers.

In the small intensity modulation limit:

$$\Delta n(x) = \Delta n_{\max} \sin(2\pi x/\Lambda) \tag{9}$$

where $\Delta n_{\max} = \frac{1}{2} n^3 r_{eff} \frac{2\pi m k_B T}{e\Lambda}$ with Λ being a period of the interference fringes and m - fringe contrast. Sinusoidally varying index modulation is a diffraction grating or hologram that can diffract light. If the sample is much thicker than the grating period a volume hologram[13] is formed. Readout of the grating occurs only when the Bragg condition is satisfied for the readout beam angle and wavelength. If material thickness d is smaller than Λ the Raman-Nath

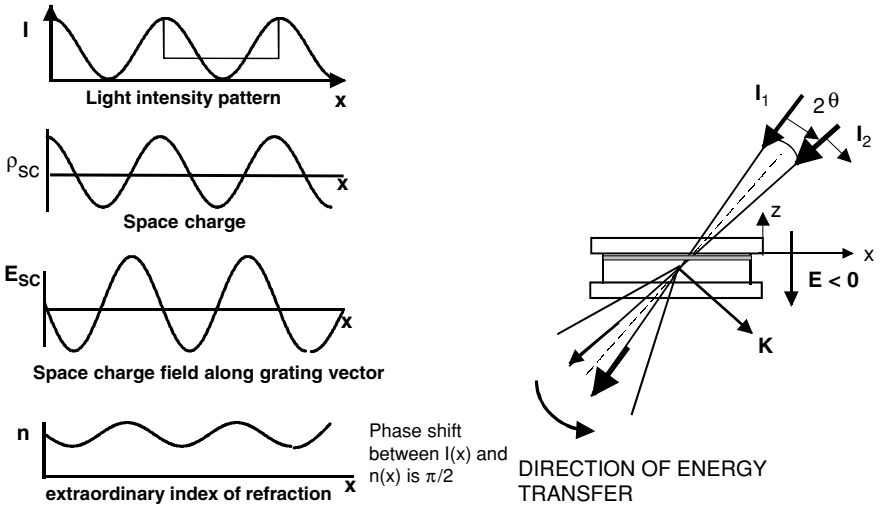


Figure 1. Model of classic (charge carrier diffusion controlled) photorefractive effect in inorganic NLO materials (left). Refractive index modulation maps the light intensity distribution irradiating the material except of phase shift $\varphi_p = \pi/2$ between them. Two-beam coupling experiment (right).

scattering regime[13] occurs and there is almost no limitation for hologram readout. Calculating a Klein parameter $Q = 2\pi d\lambda/(\Lambda^2 n)$ one can distinguish between Raman–Nath (surface or thin) hologram for $Q \leq 1$ and Bragg (volume or thick) hologram for $Q > 10$. Amplitude of the grating Δn can be measured observing light diffraction efficiency defined as $\eta = I_{diff}/I_{in}$ i.e. the ratio of light intensity diffracted into first order of diffraction to incident light intensity. In the case of thick and thin holograms the simple relations hold between diffraction efficiency and grating amplitude Δn:[13]

$$\eta_B = \exp(-\alpha_0 d)\sin^2\left(\frac{\pi\Delta n d}{\lambda}\right) \quad \text{and} \quad \eta_{R-N} = \exp(-\alpha_0 d)J_1^2\left(\frac{\pi\Delta n d}{\lambda}\right) \quad (10)$$

where d is the sample thickness, α_0 its average absorption coefficient and $J_1(\Delta\phi)$ is the first order Bessel function of phase difference $\Delta\phi$ seen by light beam.

In classic photorefractive crystals (eg. $LiNbO_3$) for $\Lambda = 1\,\mu m$, $m = 1$ and $T = 300\,K$ the space charge electric field reaches $E_{max} = 1.6 \times 10^5\,V/m$. Typically $\Delta n_{max}(\sim 1 \times 10^{-4})$ is proportional to contrast factor m and r_{eff} and to $1/\Lambda$. The resulting photorefractive grating $\Delta n(x)$ is spatially phase-shifted with respect to the intensity interference pattern, which indicates the non-local nature of the process (when $\varphi_p \neq 0$). The nonzero phase shift φ_p enables transfer of energy from beam to beam in a two-beam coupling (2BC) process

providing optical amplification[1] (cf. Fig. 1b). This is a unique process, one of the most important of photorefractive materials, which leads to many fascinating applications. Usually a weaker power beam gains some power from stronger one and optical amplification effect is observed.[16] The optical gain g experienced by I_1 beam in the presence of the another beam I_{12} is characterized by an exponential gain coefficient Γ which, for a nearly co-directional two-beam coupling, is given by:[16,17]

$$\Gamma = \frac{4\pi}{\lambda}\frac{\Delta n}{m}\sin\varphi_p \quad \text{and} \quad g = \frac{I_{12}}{I_1} = \frac{1+m}{1+me^{-\Gamma d}}e^{-\alpha_0 d} \tag{11}$$

where φ_p is a phase mismatch between the refractive index $n(x)$ grating and the intensity pattern $I(x)$. It is clear that a phase shift of $\varphi_p = \pi/2$ leads to optimum energy transfer. If, by some reason, the refractive index grating is local with respect to the fringe pattern, i.e. $\varphi_p = 0$ then $\Gamma = 0$ and no energy transfer is possible.

Linear electro-optic effect is one of the second order nonlinear optical effects related with the susceptibility $\chi^{(2)}(-\omega;\omega,-0)$ which is non-zero only in noncentrosymmetric materials. Only few polymers are noncentrosymmetric by nature. However, the invention of electric field poling techniques,[18] enforcing molecular dipoles embedded in polymer matrix to show polar ordering and also electro-optic effect, enabled the discovery and widespread utilisation of photorefractive effect in polymers.

In polymeric photorefractive materials the non- centrosymmetricity is induced by an external electric field E_z due to chromophore (dipole) alignment at elevated temperatures and freezing of this order at room temperature. The ordering is never permanent and its retention depends mainly on polymer glass transition temperature T_g. The effective Pockels coefficient at $T < T_g$ decreases in time due to relaxation of dipole arrangements. In polymers with low $T_g \cong T_a$ (T_a is an ambient temperature) the dominant effect of refractive index modulation is connected with light induced birefringence. This effect results from molecular dipole reorientation due to mutual action of externally applied electric field and space charge field ($\vec{E} = \vec{E}_{sc} + \vec{E}_a$).[19] The induced birefringence Δn_{BR} and electrooptic Δn_{EO} contributions to the grating amplitude are given by:[2,19]

$$\begin{aligned} \Delta n_{BR} &\propto \frac{1}{2n}\Delta\chi^{(1)} \propto \frac{1}{2n}\frac{2Nf_\infty\Delta\alpha}{45}\left(\frac{\mu_g}{k_BT}\right)^2 E^2 \\ \Delta n_{EO} &\propto \frac{1}{2n}\Delta\chi^{(2)}E \propto \frac{1}{2n}\frac{2Nf_0f_\infty^2}{5}\frac{\mu_g\beta}{k_BT}E^2 \end{aligned} \tag{12}$$

where N is the density of guest molecules (dipoles) in polymer matrix, f_0, f_∞ are the local field factors at zero and optical frequencies, μ_g a ground state

dipole moment, β first hyperpolarizability and $\Delta\alpha$ is a molecular polarizability anisotropy. Externally applied electric field is therefore vital for obtaining high index modulation via molecular orientation effect. The best photorefractive polymers are obtained when NLO chromophores meet the so called chromophore figure-of-merit (FOM) defined as:[2]

$$FOM = \frac{1}{M}\left[9\mu_g\beta + \frac{2\mu_g^2\Delta\alpha}{k_B T}\right] \tag{13}$$

where M is the molar mass. Chromophores can be designed by quantum chemical calculations and subsequent synthesis to maximise FOM, however many other parameters of chromophore-matrix system must be considered (like viscosity, the chromophore size, the presence of plasticizing agents, temperature relative to T_g and other factors) to obtain high optical quality and well performing material.

2.2. Photorefractive materials

Photorefractive polymers are obtained by mixture of different type functional molecules in polymer matrix (i.e. host-guest approach) or via synthesis of fully functionalised complicated polymeric structures. In the latter approach functional groups can be attached to the main chain as side groups, can be mounted in the main chain or make a 3D structures. Generally materials of host-guest type are characterized by lower T_g temperature than fully functionalised structures. Nature of phenomena occurring in polymers with low and high T_g is different. In 1991, Ducharme et al.[20] reported the first polymeric photorefractive material based on the nonlinear optical polymer, bisphenol A diglycidyl diether 4-nitro-1,2-phenyl-enediamine (bisA-NPDA) doped with a hole transporting diethylamino-benzaldehyde diphenylhydrazone. The diffraction efficiency η of the first polymeric photorefractive material was as poor as $5 \times 10^{-5}\%$ and no net gain has been observed. However, during the last decade, the performance of photorefractive polymers has rapidly improved. Through gradual improvements of the charge transport and electro-optic characteristics of the polymers ($r_{eff} \approx 50\,\mathrm{pm/V}$), the photorefractive gain of these materials now exceed those of the inorganic crystals currently available. Recently, polymeric materials with high exponential gain coefficient Γ of over $200\,\mathrm{cm}^{-1}$ and the response time of several ms have been reported.[21–23] One of the drawbacks of polymeric materials, which seriously limits their practical applications, is the necessity of usage of high external electric field in the range of $50–100\,\mathrm{V/\mu m}$ (i.e. a bulk sample of 1 mm thickness will require the impracticably high voltage up to ca. 100 kV). Such a high electric field limits the active layer thickness to several tens of microns,

and as a consequence, the Bragg selectivity of polymeric material is reduced. The most common composition of the best photorefractive polymer is ~40–60 wt% of a photoconductive polymer (poly(N-vinylcarbazole, PVK) to provide sufficient density of transport sites, ~25–35 wt% of NLO chromophore (e.g. dyes, LCs, etc.) to ensure a sufficient electro-optic response, ~15–30 wt% of a plasticizer (e.g. ethyl carbazole, ECZ) to facilitate chromophore orientation by lowering T_g of the composite, and finally, a small amount (usually below 1 wt%) of sensitizer (e.g trinitrofluorenone, TNF) to assist in charge photogeneration.[2] Good results were also observed in polysiloxane type polymers with carbazole side groups and DMNPAA as a dopant. Polysiloxane has lower T_g than PVK and there is no need of a plasticizer. The best performing photorefractive materials reach gain coefficients $\sim 400\,cm^{-1}$, internal diffraction efficiency ~100% (PSX/stilbene A/TNF) and photorefractive response time ~30 ms (PSX/DB-IP-DC/TNF).[2] Examples of many optimised photorefractive polymers are given in the excellent review of Ostroverkhova and Moerner.[2] However, one must realize that there are trade-offs between high gain coefficients and/or diffraction efficiencies and response speed, so materials are usually optimised for a certain precisely defined application.

Temporal dynamics of refractive index grating recording depends on light intensity used for recording and exposure time, grating erasure can be performed using spatially uniform illumination. Because chromophore orientation is usually a slow process the dynamics is either limited by charge carrier mobilities or by reorientation times of chromophores. For volumetric storage applications the two-photon grating recording techniques are elaborated in order to avoid partial erasure of earlier written holograms during readout process.

The new emerging organic photorefractive materials are various liquid crystalline systems.[24–26] The photorefractivity of nematic liquid crystals is derived from the reorientation of the birefringent rod-like molecules in the photoinduced space charge field. In fact, it has been reported that 80% of the observed photorefractivity in the best photorefractive polymers was due to the orientational (or quadratic) electro-optical effect.[19] Since liquid crystals are composed entirely of the birefringent chromophores, as opposed to a fraction within polymer composites, liquid crystals represent an excellent opportunity to improve the photorefractive gain performance of organic materials. The electric fields of the order of $1\,V/\mu m$ are sufficient for molecular reorientation in LC's.[27] Low-molecular-weight liquid crystals are particularly advantageous in the dynamic holography because the mobility of mesogenes is high compared to that of other LC materials.[28–31] However, it is difficult to form the holographic grating with narrow fringe spacing (up to 200 lines/m) and permanent stability within this group of materials. In photorefractive polymer-dispersed liquid crystals (PDLCs),[32–36] the mobility of mesogenes is suppressed with polymers as binder. Droplets of LC are confined

in polymer cavities. Best performing PDLCs exhibit photorefractive internal diffraction efficiency of 100% in 105 μm films at electric fields of only 8 V/μm and net gain coefficients of 37 cm^{-1}.[37] The weak points of PDLC's include high losses due to scattering and slow, of the order of minutes, photorefractive dynamics.

Photorefractive effect in calamitic liquid crystals deserves slightly modified theoretical description as the material is usually centrosymmetric and there in no electro-optic contribution due to Pockels effect. The space charge field resulting from sinusoidal light intensity modulation in the dye-doped LC (dye is serving as a sensitizer) was found by Rudenko and Sukhov[25] to depend on local difference in liquid crystal conductivity:

$$E_{SC} = \frac{mk_B T}{e} K \nu \frac{\sigma - \sigma_d}{\sigma} \sin(qx) \tag{14}$$

$$\text{and} \quad \nu = \frac{D^+ - D^-}{D^+ + D^-} \tag{15}$$

where D^+, D^- are the diffusion constants of cation and anion, $K = 2\pi/\Lambda$ is the grating wavevector and σ and σ_d are the total and dark conductivity, respectively. Alternative mechanisms for space charge creation in LC's have been formulated since that time. The so called Carr-Helfrich effect is due to conductivity anisotropy:[38]

$$E_{SC}^{\Delta\sigma} = -\left(\frac{\Delta\sigma \sin\theta \cos\theta}{\sigma_\parallel \sin^2\theta + \sigma_\perp \cos^2\theta}\right) E_a \tag{16}$$

where $\sigma_\parallel$ and $\sigma_\perp$ are conductivities along and perpendicular to the LC director (a preferred local orientation of long molecular axes) with Δσ determining the conductivity anisotropy, θ—director reorientation angle and E_a—the applied electric field. Similar effect arises from the substantial dielectric anisotropy of LC's $\Delta\varepsilon = \varepsilon_\parallel - \varepsilon_\perp$.[39]

$$E_{SC}^{\Delta\varepsilon} = -\left(\frac{\Delta\varepsilon \sin\theta \cos\theta}{\varepsilon_\parallel \sin^2\theta + \varepsilon_\perp \cos^2\theta}\right) E_a. \tag{17}$$

Ferroelectric liquid crystals (FLCs), known from fast switching times of the order of μs, seldom appear in the literature of photorefractive effect. It is well known that aligned FLCs have a net polarization (P_s) in the smectic C* phase due to surface stabilization. Having C_2 symmetry these materials allow for observation of electronic electro-optic effect. In FLC's due to presence of coupling of P_s and the space-charge field a strong orientational effect is expected. The first observation of orientational photorefractive effect in FLCs was reported by Wiederrecht et al.[40] in a smectic C* mixture

characterized by birefringence $\Delta n = 0.14$, spontaneous polarization $P_s = -6.6\,nC/cm^2$ and doped with suitable molecules to produce mobile ions upon light illumination.

Photorefractive-like properties of nematics are well-recognized and it is widely accepted that light could impose molecular reorientation through different mechanisms: like thermal, dye-induced or surface induced reorientations.[41–45] Here we present a designed and fabricated by us LC cell able to perform tasks of photorefractive medium.[46,47] The pure nematic liquid crystal mixture is confined in a planar cell (10 μm in thickness) with transparent ITO electrodes and orienting coatings. One of the coatings was made of a 100 nm thick layer of photoconductive poly(N-vinyl carbazole) PVK doped with 2,4,7-trinitrofluorenone (TNF) or 2,4,7-trinitro-9-fluorenilidene-malononitrile (TNFDM) to induce its sensitivity in the visible. A schematic structure of the hybrid photoconducting polymer—liquid crystal system developed by us is shown in Fig. 2.

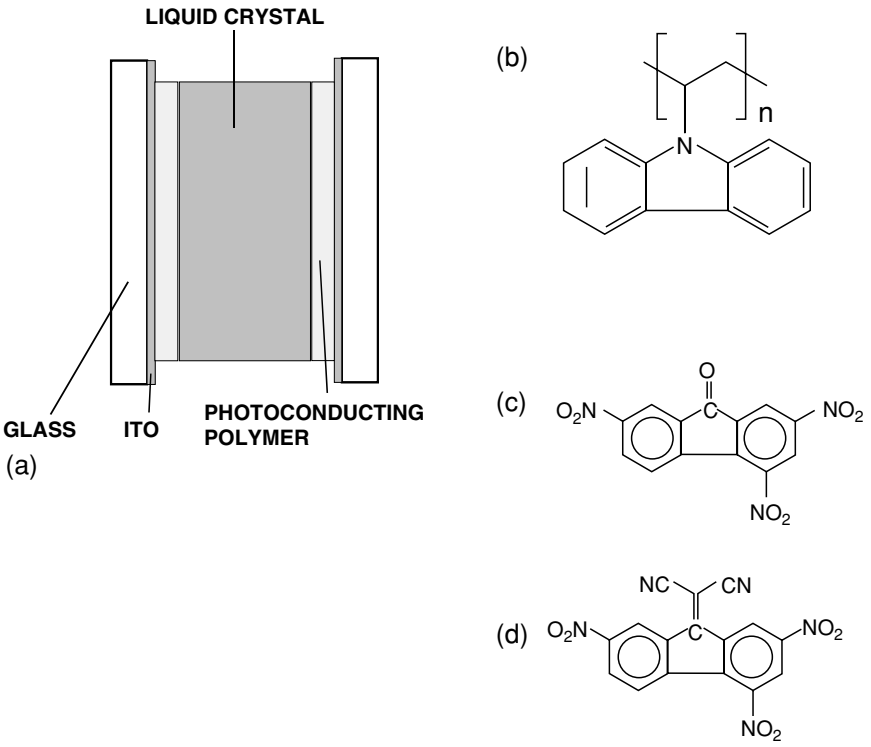


Figure 2. A schematic structure of the photorefractive liquid crystal panel. The external dc voltage is applied to the panel via the ITO conducting layers (a). Chemical formulae of PVK (b), TNF (c) and TNFDM (d).

The cell contains a multicomponent nematic liquid crystal mixture E7 (Merck KGaA, Darmstadt, Germany) characterized by $\Delta\varepsilon = +13.8$, $\Delta n = 0.2253$ at 589 nm ($n_e = 1.7464$, $n_o = 1.5211$) and viscosity $\gamma = 39\,\mathrm{mm^2 s^{-1}}$. The dc voltage (0–20 V) is sufficient to change the director orientation from planar to nearly homeotropic. Spatially modulated light ($\lambda = 514$ nm or 632.8 nm) $I(x,y)$ incident on the biased LC cell induces charge carrier photogeneration process with photoconductive sensitivity depending on light intensity and a biasing electric field within a polymeric layer E_p: $\sigma/I = e\alpha\mu(E_p)\tau\Phi(E_p)/h\nu$, where e is the electron charge, α the absorption coefficient, $h\nu$ the photon energy, τ the life-time of the photocarrier and $\mu(E_p)$ the carrier mobility. The efficiency of photogeneration $\Phi(E_p)$ according to the Onsager/Marcus theory[14] is a function of local electric field E_p. Under constant light illumination the charge density distribution is established which is converted into respective changes of surface potential $E_{sc}(x)$. In the process also ionic transport through the LC layer is important, as it tends to compensate an excess charge. The electric field distribution within LC layer changes extraordinary effective index of refraction $\Delta n_e(x,y,z)$ in its bulk:[27]

$$n_e^{eff} = \frac{1}{d}\int_0^d \frac{n_e n_o dz}{\sqrt{n_e^2 \sin^2\varphi(z) + n_o^2 \cos^2\varphi(z)}} \tag{18}$$

where d is the thickness of the LC layer, n_e and n_o are extraordinary and ordinary indexes of refraction and $\varphi(z)$ is the angle between the nematic director $\hat{n} = (\cos\varphi(z),\ 0,\ \sin\varphi(z))$ and the x-axis. When intensity grating $I(x) = I_0(1 + m\cos(qx))$ is projected onto LC panel refractive index grating will set-up and the extraordinary polarized laser light will be diffracted on it. Sinusoidal grating of the form $n_e(x) = n_{0,e} + \Delta n_e \cos(qx - \phi_P)$ constitutes the phase hologram. In such hybrid photoconducting polymer - liquid crystal system we measured nearly 30% diffraction efficiency at $2\,\mathrm{V/\mu m}$, net exponential gain $\sim 3700\,\mathrm{cm^{-1}}$ and response speed of the order of 30 ms.[48] Spatial resolution was up to 200 line pairs/mm.

2.3. Photochromic effect

Photochromism is defined as a light-induced reversible transformation of a single chemical species between two states having distinguishably different absorption spectra.[49,50]

$$A(\lambda_1) \underset{h\nu_2,\Delta}{\overset{h\nu_1}{\Leftrightarrow}} B(\lambda_2)$$

Usually, B has at least one absorption band appearing at longer wavelength than the absorption bands of A. The activating photon energies ($h\nu_1$ and $h\nu_2$)

induce the forward and backward molecular transformations and the absorption spectra changes are accompanied usually by thermally (Δ) induced change in one direction (cf. Fig. 3). The single chemical species A may be a molecule or an ion whereas the product B may in some cases represent more than one species.

For holographic applications, the spectral dependence of the light induced absorption coefficient changes $\Delta\alpha$ and refractive index changes Δn are important. The molar refraction and the molar absorption coefficients are the real and imaginary parts of a single complex value. The spectral (v and s denote frequency in eq. 19) dependence between both values is described by the Kramers–Krönig relations for the case of an undisturbed chromophore:

$$\begin{aligned}\chi'(\mathrm{v}) &= \frac{2}{\pi}\int_0^\infty \frac{s\chi''(s)}{s^2 - v^2}\,ds \\ \chi''(v) &= \frac{2}{\pi}\int_0^\infty \frac{v\chi''(s)}{v^2 - s^2}\,ds\end{aligned} \tag{19}$$

where linear susceptibility is given by $\chi^{(1)} = \chi' + j\chi''$ corresponding to a complex permittivity $\varepsilon = \varepsilon_0(1 + \chi^{(1)})$. Equation relating the refractive index n and the absorption coefficient α to the real and imaginary parts of the susceptibility χ' and χ'' is the following:

$$n - j\frac{\alpha c_0}{2\omega} = (1 + \chi' + j\chi'')^{1/2} \tag{20}$$

where c_0 is a speed of light in vacuum and $\omega = 2\pi v$. In solid-state materials, due to matrix effects the Kramers–Krönig relations might not hold.

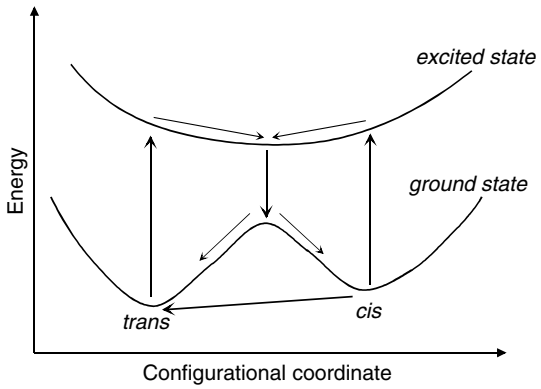


Figure 3. Schematic view of simplified energetic diagram of photochromic molecule and light induced conformational transitions of *trans–cis* type.

The most prevalent photochromic systems are based on unimolecular reactions though bimolecular reactions, such as reversible photocycloadditions of electron transfer processes of the type ($A + B \rightarrow P$) are equally possible. Reversibility is the main criterion for photochromism. The photochromic transformation and the observed spectral changes or changes in physical or chemical behaviour are related to the modification of geometry of the system and electronic density. The medium in which the photochromic compound is incorporated (e.g. polymeric matrices) strongly influence and/or control the kinetics of the thermal back reaction and sometimes the forward photochemical reaction, too. The major photochromic processes are *trans–cis* photoisomerisation, proton transfer (tautomersim), ring formation and cleavage, cycloadditon reactions, electrocyclic reactions, dissociation processes, zwitter-ion formation, etc... (cf. Fig. 4 for examples).[49,50] Classes of photochromes may be differentiated according to the chemical transformation that they undergo:[50] -*E*-*Z* isomerisation of double bonds (e.g., stilbenes, azo compounds, polymethines, indigoids, etc.), electron transfer (e.g. Fe(II) thiazines, chlorophyll cell), proton transfer (tautomerism) (e.g. metal dithizonates, N-salicylidene-anilines), photodissociation (e.g. triarylmethanes), pericyclic reactions: intramolecular electrocyclic reaction (e.g. spiropyrans,[51] spirooxazines,[51] fulgides, etc.)

The most intensively studied photochromic systems are spiropyrans and related compounds[51], azobenzenes[52], viologens and fulgides[53]. For volumetric optical memories the most actively studied are spiropyrans and the photochromic bR.

Particularly widely studied are azo-functionalized polymers.[54–59] In this group of materials the azobenzene derivatives are covalently attached to the backbone as pending groups. Azobenzenes when attached as side groups to main polymer chains pertain their photochromic properties. The presence of

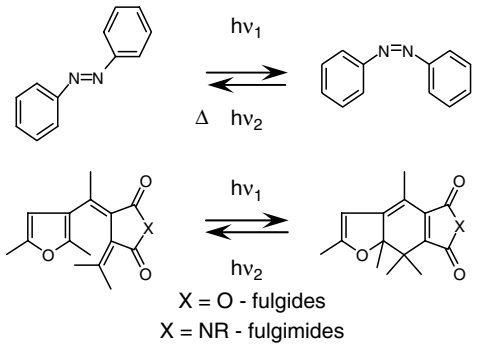


Figure 4. Photochromic reactions of *trans-cis* photoisomerisation in azobenzenes and tautomerism in fulgides and fulgimides.[52,53]

N=N bonds ensures the existence of two photoisomers *trans* and *cis*. These isomers differ in shape (*trans*- is elongated rod-like molecule whereas *cis* is bend and can be treated as being isotropic). Switching between two isomers is possible via absorption of photon of suitable energy. Absorption crosssection for *trans*-isomer fulfils relation $\sigma_{\parallel} > \sigma_{\perp}$. Dipole moment μ_1 and molecular hyperpolarisability β_{111} (cf. Fig. 5) are much higher for *trans*-isomer than for *cis*-isomer. Molecules after prolonged irradiation tend to stay with their long molecular axes of *trans* form perpendicularly to the ***E***-vector of light. Multiple *trans-cis* photoisomerisation processes which lead to molecular alignment can be understood in terms of sequential molecular reorientations.

Examples of azo-functionalised polymers in which photoinduced birefringence anisotropy $\Delta n = n_{\parallel} - n_{\perp}$ and dichroism $\Delta\alpha = \alpha_{\parallel} - \alpha_{\perp}$ is observed are given in Fig. 6.

Molecular reorientation is in fact a result of complex process consisting of angular hole burning (AHB)[59–61] in which those molecules which are nearly parallel to the light polarization direction are excited to *cis** and next they relax either to ground *trans* or *cis* state. Next the angular diffusion process takes place both for *trans* and *cis*, populations. Molecular alignment state can

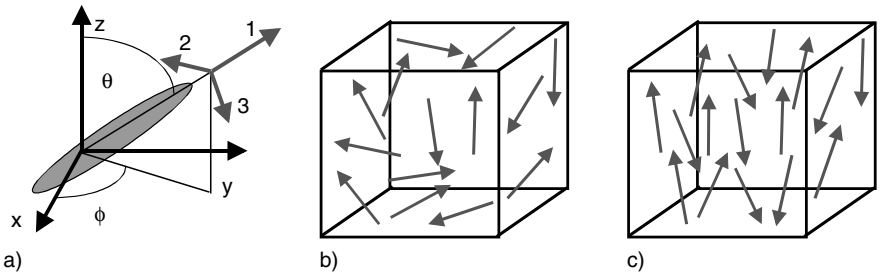


Figure 5. Scheme of photochromic trans molecule in coordination systems (a). Random distribution $f(\Omega)$ of chromophores in a polymer matrix before illumination by light (b) and after illumination by a linearly polarized $\boldsymbol{E} \parallel y$ light (c).

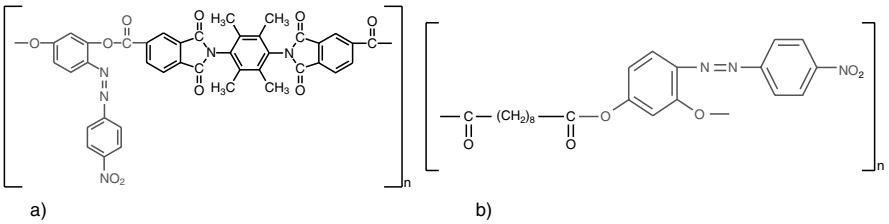


Figure 6. Examples of azo-functionalised photochromic polymers: polyesterimide and polyester.

be frozen in those polymer matrices that are sufficiently rigid to not allow for fast molecular reorientation after removal of irradiation. Slow angular redistribution having place in polymers having high glass transition temperature T_g results in appearance of axial molecular alignment (cf. Fig. 5). This mechanism has been detailed by Fischer et al.[59] and is schematically depicted in Fig. 7. In the case of azo-compounds attached to polymer chains frequently the mass transport occurs resulting in formation of amplitude and phase gratings in those materials including surface relief gratings (SRG)—a novel mechanism of nanostructuring, photonic device fabrication and information storage.[62,63]

However, several factors such as molecular relaxation and electrostatic interactions reduce the photoanisotropy. Depending on matrix properties and pending azogroups various angular distribution functions for *trans*-molecules might be obtained in relation to light illumination conditions, including axial, polar and chiral molecular arrangements. This allows for polarization sensitive holographic recording.[58]

When, after grating recording with the use of two interfering laser beams, as shown in Fig. 8, sample will be moved along grating wavevector K in the presence of optical field the correlated oscillations of diffraction power will be observed. As follows form the theory of moving gratings[64] the sum $\Sigma = I_L + I_R$ and difference $\Delta = I_L - I_R$ of intensities of two zero order diffraction beams allow to separate contributions to diffraction efficiency

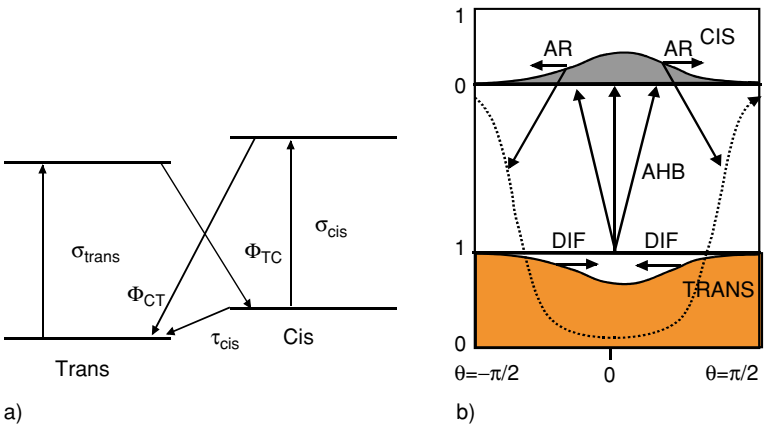


Figure 7. Scheme of energy levels for photoisomerization process of trans-cis-trans (a) mechanism of axial ordering of trans molecules after irradiation with linearly polarized light (b). AHB—angular hole burning, AR—angular redistribution of cis-molecules, DIF—diffusion, thermal relaxation.

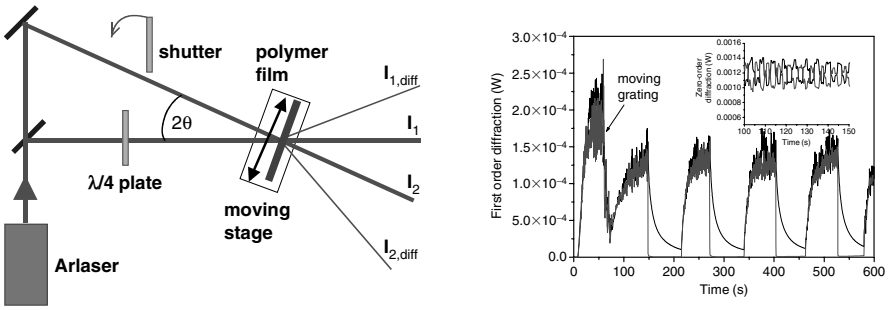


Figure 8. Scheme of two-wave mixing set-up for study of dynamics of polarization (s-p) grating recording in azo-functionalised polymers (a) and temporal evolution of subsequent recording and erasure of the diffraction grating supplemented (inset) with results of moving grating technique experiment (b).

coming from refractive index Δn (phase) grating and absorption coefficient $\Delta\alpha$ (amplitude) grating:

$$\begin{aligned} I_L &= I_0 \exp[-\alpha_0 d/\cos(\theta/2)] \cdot (1 - 2A\cos\phi_\alpha - 2P\sin\phi_n), \\ I_R &= I_0 \exp[-\alpha_0 d/\cos(\theta/2)] \cdot (1 - 2A\cos\phi_\alpha + 2P\sin\phi_n), \end{aligned} \tag{21}$$

where $A = \Delta\alpha d/[4\cos(\theta/2)]$, and $P = \pi\Delta n/[\lambda\cos(\theta/2)]$.

Typically diffraction efficiencies η in azo-funcionalized polymers are $\sim 0.1-4\%$, recording times are of the order 100 ms to hours and light sensitivity is relatively poor. An advantage is that no biasing voltage is needed for hologram recording and both transient and permanent but reversible recording is possible. The ability of azo-functionalised polymers to form surface relief gratings utilized in nanotechnologies and other applications[65] is advantageous with respect to photorefractive materials.

2.4. Transient holographic properties of bacteriorhodopsin

bR is a transmembrane protein containing seven transmembrane α-helices with a retinal co-factor (Fig. 9) attached to Lys_{216}. It forms a two-dimensional crystalline lattice in purple membrane of *Halobacterium salinarium*.[4,9,10]

On light exposure, bR transports protons from halobacterial cell to the outer medium, and thereby converts light energy into chemical energy. The molecular processes occurring in bR after photon absorption are complex and involves not only the reversible photochromic and conformational changes but also reversible protonation and deprotonation of Asp residues. The photocycle of bR (quantum yield of 0.67) is initiated by the absorption of a photon by the retinal which is linked to the protein by a protonated Schiff

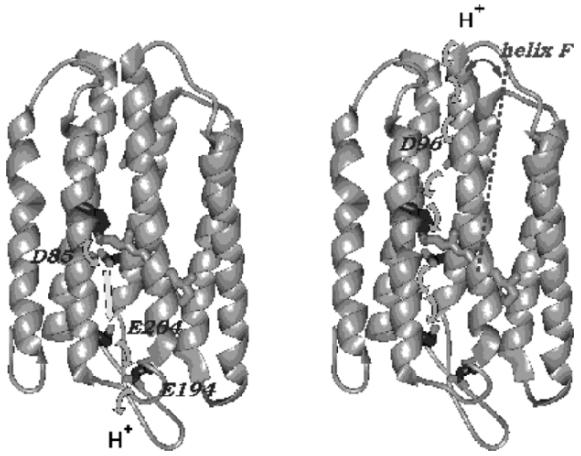


Figure 9. Structure of bR. A proton H^+ is pumped through membrane after initiation by a photon of a suitable energy.

base at Lys_{216}. This causes rapid rearrangement of the electronic structure of the extended conjugated retinal π-electron system which ultimately results in *trans* ⇔ *cis* isomerization at the C13=C14 double bond (K intermediate). This isomerization in turn reduces the proton affinity of the charged Schiff base nitrogen which loses its proton to the initial acceptor group, Asp_{85}, probably via a water molecule (M intermediate). The major steps of the bR photocycle are shown in Table 2.[66]

bR has been optimized by evolution, the shelf-life is very long in comparison to synthetic photochromes. Isolated purple membranes are stable for many years when exposed to air and sunlight which makes it attractive for applications. In the dark, bR contains a mixture of 13-*cis* and *all-trans* retinal (D- and B-state). They are converted by light to a pure *all-trans* population of light-adapted bR (B-state). The photochromic shift observed in bR between B

Table 2. Photocycle of bacteriorhodopsin with characteristic life times, τ, of various electronic states

τ	**500fs**	**10ps**	**2 μs**	**40 μs**	**7ms**	**4ms**	
BR_{570}	J_{625}	K_{600}	L_{550}	M_{412}	N_{520}	O_{640}	BR_{570}
T NH^+		twC NH^+	C NH^+	C N	C NH^+	twT NH^+	T NH^+

T = *all-trans*; C = 13=14 *cis*; twC = twisted 13=14 *cis*; NH^+ = protonated Schiff base; N = deprotonated Schiff base. (H. Luecke, H-T. Richter & J. K. Lanyi (1998) *Science* 280, 1934–1937).

and M state is substantial (570 nm ↔ 412 nm). The lifetime of the photointermediate M is only 7 ms in wild-type bR. By changing pH-value of environment, using less protic solvents such as glycerol, and rational site-directed mutations in bR, the M-lifetime can be enhanced and modulated.

Optimization of bR by rational site-directed mutagenesis

Wild type bR possesses ideal characteristics required for an ideal optical storage medium. Nevertheless re-optimization of bR by rational site-directed mutagenesis[68–71] has resulted in bR mutants with significantly enhanced physical and optical properties that make them potentially ideal candidates for an optical binary storage material. Their naturally evolved properties, along with their genetically engineered variants, make them superior to any magnetic material used for binary data storage in hard-drives. Birge et al.[7] and Hampp[8] have discussed the application of bR in 3D holographic volumetric memory. For example, the bR mutants with less than 3 nanometers in diameter have demonstrated a long-term stability with a shelf life of 10 years at room temperature. In comparison, magnetic grains used in the best hard-drives today become highly unstable if the average grain size is reduced below 3 nm due to magnetic domain instabilities[72]. Among other advantages of bR medium is its unprecedented recyclability and durability including resistance to microbial degradation. For example, bR media can be rewritten more than 10^9 times[73]. Finally, the bR media have demonstrated a faster time response, as compared to the magnetic disks, picosecond range versus nanosecond range, respectively. The faster time response makes the bR media superiors with respect to the data transfer rate. However, before the protein-based data storage can be finally implemented, suitable methods for immobilizing bR mutants on platforms (Tatke et al.[74]) must be elaborated.

Rational site-directed mutagenesis

Rational site-directed mutagenesis rests on the principle of optimizing a chosen physical property by analyzing the dependence of that property on specific candidate amino acids that are to be substituted by selecting known or putative sites in the protein under investigation. On absorption of light, bR undergoes a cascade of structural changes during its photochemical cycle. The cis-trans isomerization, the heart of its photochemistry, triggers an avalanche of structural changes in the bR secondary structure which consists of at least three major steps: (a) the release of proton into the extracellular medium (b) the uptake of a proton from the cytoplasmic medium and (c) the thermal re-isomerization of retinal to the starting all-*trans* configuration. These structural changes are reflected as changes in the visible region of the spectrum as observed by the sequential formation and decay of the optical intermediates, K, L, M, N, and O, see Fig. 10 below.[75,76]

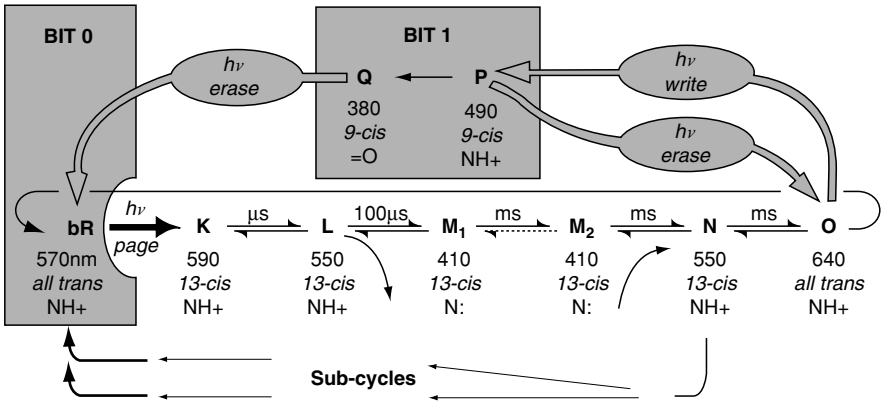


Figure 10. Diagrammatic representation of bR photocycle.

Extensive mutagenesis studies have established that the presence of a carboxylate group at Asp residue 96 is important for rapid reprotonation of the Schiff base[77–81]. The protein conformational changes in the bacteriorhodopsin photocycle have been discussed by Subramaniam et al.[82] Reprotonation of the Schiff base by Asp_{96} is reflected by the spectroscopically detectable conversion of the M intermediate to the N intermediate. Asp_{96} is reprotonated from the cytoplasmic medium, initiating the formation of O-intermediate. The mutation of Asp_{96} to Asn_{96} therefore influences the above process and confers long resident time or half-life for the M intermediate. Secondary structure of [D96-N] bR obtained from x-ray crystallography[83] is shown in Fig. 11 below.

We have taken a different approach in the design of high melting mutants of bR. While other laboratories relied on stabilizing the intermediates by temperature, pH, and chemical optimization, we have taken the view that the design of mutants manifesting high quantum yield and thermal robustness, can be better accomplished by a combination of theoretical analysis of the energy level diagram of wild-type bacteriorhodopsin and genetic engineering methods. The band gap between the energy levels, E, and its dependence on amino acid substitution are important determinants in the design of photostable and photorobust mutants of bR. Detailed QM/MM [Hartree–Fock (HF) combined with molecular mechanics force field using AMBER force field] on bR from *Halobium salinarium* were undertaken to develop an energy-level diagram for bR to ascertain the detailed energetics of the excited of bR and then investigated their dependence on amino acid substitutions of key residues, e.g., Arg_{82}, Asp_{85}, and Asp_{212}

The photochemical cycle of bR consists of several intermediates, which can be identified by their spectroscopic properties, mainly their absorption

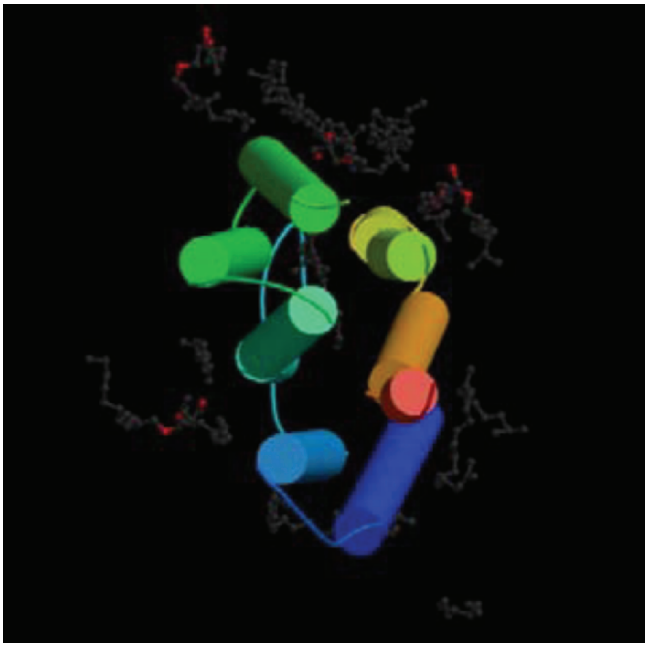


Figure 11. 3D Structure of Bacteriorhodopsin [D96 N].

maxima. Metastable, blue-shifted M state has been the principal focus in most holographic applications of bR because it has the longest lifetime. Numerous attempts have been made by chemical modification techniques to extend this lifetime. Time-resolved Raman and Fourier transform infrared spectroscopic studies in our laboratory (Renugopalakrishnan, unpublished) and other laboratories have demonstrated that the photointermediate O state shows a branched photocycle in which it converts itself into P and Q states. These two photointermediates may not be important to the wild-type bR but assume significance when it comes to the design of bR mutants with long-lived intermediates, especially in their use as bit-by-bit digital storage media, which is our principal goal in utilizing bR in the design of high performance storage devices.

One of the mutants so designed, bR192, that satisfies the criterion of high quantum efficiency for the branched photochemistry ($O \rightarrow P$) and further satisfying the criterion of thermal robustness involves multiple mutations. The design of bR192 represents a significant breakthrough resulting from a combined approach of theoretical energy levels and their dependence on amino acid substitutions and principles of thermal stabilization. bR192, may the most advanced and useful bR mutant for commercial application in our

hands because the Q and O states are robust for this mutant and thereby increasing the shelf life of the bR mutant significantly. The combined approach of theoretical principles based on a hybrid quantum mechanical approach to elucidate the energy-level diagram of bR, their dependence on amino acid substitutions, and the principles developed in our laboratory for calculation of ΔT_m from the heat capacity ΔCp (of individual amino acid residues) and the primary structure of protein are novel methods in protein engineering that were utilized in bR a test case to re-engineer bR. Detailed mathematical analysis of the decay kinetics (details are unpublished) and a detailed comparison with other mutants is in progress in our laboratories (Renugopalakrishnan et al., to be published). We used a computer-aided rational design concept and homology of sequences followed by experimental genetic manipulation of the respective cDNAs of bacteriorhodopsin to express the mutants initially in *Escherichia coli* and more recently in *P. pastoris*[84].

Eight residues in w-bR (D85, W86, L93, D96, D115, W182, W189, D212), see Fig. 12 below, : derived from numerous studies reported in the literature

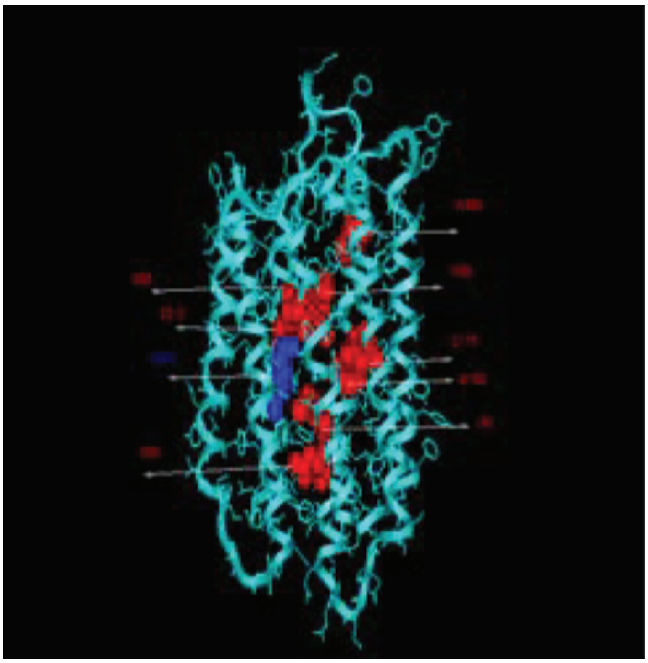


Figure 12. Candidate residues where site-directed mutations were induced reproduced with permission from ref. 68, Wiley & Sons, New York, NY.

were systematically mutated to enhance T_m and lifetimes, τ, of M and O photointermediates. Arg and Tyr residues were the chosen residues whenever possible as the choice for the above eight residues in the rational site-directed mutations because these residues occur much more frequently in thermophilic proteins. The overall α-helicity of bR was retained in the designed mutants. With the aid of a computer routine developed in our laboratory we generated hundreds of possible mutations that can be induced in the selected eight residues in w-bR and selected some of the mutants to test them for their information storage capacity. Two criteria were used in the selection of bR mutants: (1) their thermal stability and (2) slower decay of the photocycle and extension of half-life of the M and O intermediates. We have calculated the optical properties of the designed mutants and compared them with reported values in the literature.[85] We estimated the holographic efficiency, the corresponding change in refractive index, Δn, from the change in absorption in function of frequency, using the Kramers–Krönig relations[85]. The calculations were performed for each of the mutants. The results of the calculations are shown in Table 3.

The holographic measurements were performed on the thin film of the mutants encapsulated in polymethylmethacrylate Langmuir–Blodgett film. For reasons of protecting the proprietary rights of bR192, for which provisional patents have been filed at present, the precise mutations induced in bacteriorhodopsin are not reported here. One of the mutants, bR192, manifests a relative diffraction efficiency of 4.8, a large change in refractive index, Δn, of 4.0×10^{-4} at 645 nm in comparison to other mutants reported in the literature.[85]

Table 3. Summary of optical properties for selected bR mutants

Mutant	**Ground State Absorption (nm)**	**Max. Δn in 10^{-4}**	**Relative Diffraction Efficiency (%)**
bR192	4.0@645nm	4.8	
bR184	540	3.7@560nm	3.8
bR D85N	610[a]	2.4@680 nm	1.0
br D85N/A103C	600[a]	1.9@680nm	0.6
bR D85N/D96N	575[a]	2.2@660nm	0.8
bR D85N/V49A	610[a]	4.0@670nm	2.8

[a]Millerd, J. E., Rohrbacher, A., Brock, N., Chau C-K., Smith, P. and Needleman, R. Opt Lett 1999, 24, 1355.

3. HOLOGRAPHIC OPTICAL PROCESSING, STORAGE AND OTHER IMPLEMENTATIONS

Holography is an image-recording process in which both the phase and the amplitude of the optical field are recorded in a photosensitive medium. The interference of laser coherent image bearing beam and a reference beam overlap within a photosensitive medium and form an interference pattern which is recorded. Complicated diffraction grating-like pattern stored in the material diffracts reference beam during hologram retrieval and reconstructed real or imaginary object wave appear. The scheme of a simple holographic recording set-up is shown in Fig. 13. The attractive feature is that, despite the hologram is usually recorded on a flat surface, it contains information about entire three-dimensional wavefield scattered form the object. Holograms can also be recorded in a volume of a photosensitive material instead of a thin film (3-D holographic process). The phenomenon called Bragg selectivity is used to holographically store many different images within the same volume. Therefore the so called volumetric optical storage technique allows for achieving the theoretically highest ($10^{14}\,\mathrm{bit/cm^3} \sim 1/\lambda^3$) information packing density among the all known today recording techniques. Higher information packing densities are only offered within near field optical approach (SNOM) employing nanometer scale moving optical fibre tip but at the cost of much slower access time to the data. Any particular image stored in the material can be independently retrieved using the reference beam with the exact incidence angle (wavelength or spatial code) that was used during recording. Hologram multiplexing techniques (angular, wavelength, shift or phase- code) allow to

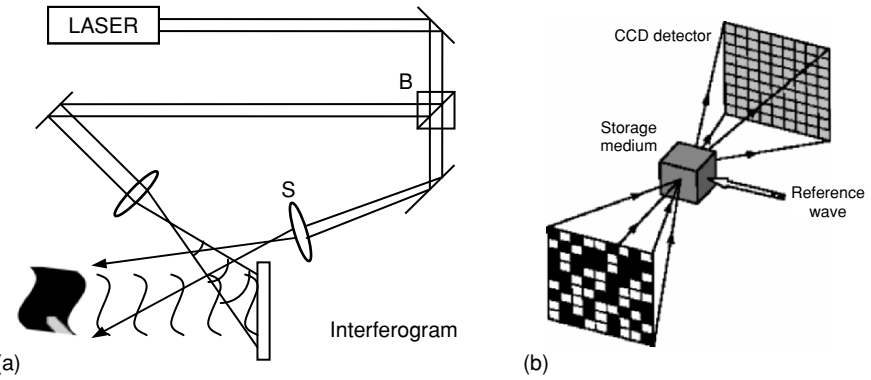


Figure 13. Off-axis holographic recording of a 3-D object in a 2-D photosensitive medium (a) and an idea of holographic binary data storage system using SLM for data input and matrix detector for data output (reconstructed hologram).

store about 20000 holograms in 1 cm^3 crystal. System architecture for typical holographic optical data storage or processing system must comprise: (i) a coherent source or collection of sources to provide object, reference and reconstruction waves and erasure; (ii) spatial light modulator for preparing binary or multilevel data pages; (iii) a storage medium allowing for holograms writing accompanied with suitable optics for routing and imaging wavefields; (iv) a detector array and subsequent electronics for data read-out, signal processing and error correction.

High fidelity reconstruction is achieved already with diffraction efficiency $\eta \approx 10^{-4}$. Holographic techniques have demonstrated the ability to transfer 1 Gbits/s (60 times faster than DVD's)[86]. Therefore, holography is expected to be the most promising candidate for massive storage of high-density information with an extremely high data transmission rate. Recently the optical holographic storage and holographic light processing become more feasible due to improvement of transmissive electrically addressed liquid crystalline spatial light modulators (EA LC SLMs) enabling transfer of pages of data onto laser beams in the form of phase binary or grey scale encoding of information. Ferroelectric LC SLM's can provide as many as 2000 binary pages per second each containing 1024×1204 pixels. The optically retrieved binary pages can be read by matched detector arrays (high density CCD camera chips) decoded and electrically transmitted to output devices.

High-performance holographic data storage requires high-power, single spatial-mode lasers. Currently used are red or green diode pumped solid state lasers ($\sim$50 mW). However, the short-wavelength blue lasers are the ultimate light sources for holographic storage because the theoretical limit on the data density is 1 bit/λ^3, where λ is the wavelength of the light used in the writing/readout process. The most important requirements for optical data storage devices are: high storage capacity, fast access time, low error rate (10^{-7}), long data stability, compactness and low manufacturing costs. The gridlock for the wide usage of holographic volumetric optical data storage lies in the lack of suitable high optical quality and durability materials. One of the best known material for permanent optical storage is developed at Aprilis photopolymer using light induced cationic ring opening polymerization (CROP).[87] What is important this polymer epoxy-modified siloxy silane is of excellent optical quality and undergoes minimal volume changes during recording. Architecture of the future holographic compact disk drive is shown schematically in Fig. 14. However, there is a tremendous progress in development of photosensitive materials enabling the dynamic recording of holograms. Self- or light-assisted erasure of just recorded holograms is a basis of the so-called real-time holography. The real-time holography is by definition any repetitive dynamic holographic recording/erasure cycle. The holograms are stored temporarily only and can be erased either by light or by spontaneous relaxation

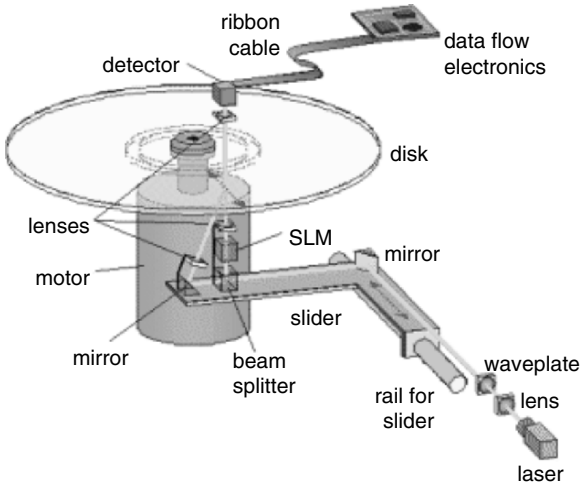


Figure 14. Holographic compact memory disk. Spatial light modulator (SLM) provides binary data pages that are superimposed onto object laser beam. Interferogram is formed in the overlapping area of the object and reference beams. Matrix detector is used for hologram retrieval.

processes. Future commercial visualization systems as well as many scientific implementations (e.g. holographic interferometry[88] or holographic optical tweezers[89]) will require holographic projection i.e. a three-dimensional (3-D) image exhibiting all the effects of perspective and depth of focus[90].

The dynamic (or real-time) holography requires the fully reversible photosensitive material lacking any post-processing.

3.1. Optical retrieval of binary holograms

Computer-generated hologram (CGH) is a vital element in the optical system and it generally relates to a desired pattern or replay field via a Fourier transform. We constructed and characterized the holographic projection system utilizing the optically addressed liquid crystal spatial light modulator (OA LC SLM) as a holographic medium.[91] For that purpose we used a commercially available computer controlled multimedia projector with Liquid Crystal Display (LCD) displaying computer generated holograms onto OA LC SLM. This proof-of-principle demonstration of excellent transient holographic properties of LCs is based on projection with incoherent white light of digital binary holograms and their reconstruction with collimated laser beam. The extraordinary polarized laser light diffracts on the holograms and the first Fraunhofer diffraction order forms the real reconstructed image. The optical

set-up is shown in Fig. 15.[92] Spatially modulated light passing through Nikon camera lens system forms a sharp image of CGH on liquid crystal modulator. Light from He-Ne laser illuminates the phase pattern and a clear holographic image is formed at a distance of about 4.7 m. Computed interferogram of the input image which feeds the multimedia projector can be changed at video rate.

The physical sizes of the binary holograms illuminated onto the OA LC SLM through camera lenses were typically less than $1.5 \times 1.5\,\mathrm{cm}^2$. Calculated from diffraction pattern virtual pixel size on OA LC SLM was $\Delta x = 22.2\,\mu m$ and $\Delta y = 30.9\,\mu m$. Image was fine-tuned by an adjustment of SLM position in the optical setup, however performance of the system was dependent on characteristics of the key-element: the photosensitive OA LC SLM. In Fig. 16 we show a target (4a), its hologram (4b) calculated by DBS method, numerically retrieved image (4c) and the image retrieved optically (4d). Optically reconstructed image contains the dominant 0-order light spot (undiffracted reconstruction wave).

Short holographic films can be composed by changing holograms up to frame rate of 30 Hz. No long-term changes of the performance of the OA LC SLM under permanent usage (8-hour video-rate experiment corresponding to 0.864 million completed exposure and erasure cycles) were observed.[92] The size of the object seen at screen depends on the total size of hologram projected on OA LC SLM and can be enlarged by decreasing the size of the projected hologram. Spatial resolution of tested LC panel reached 250 lp/mm. Using higher resolution (i.e. 512×512 pixels) four phase holograms the more image details could be seen. In Fig. 17 we present optically reconstructed

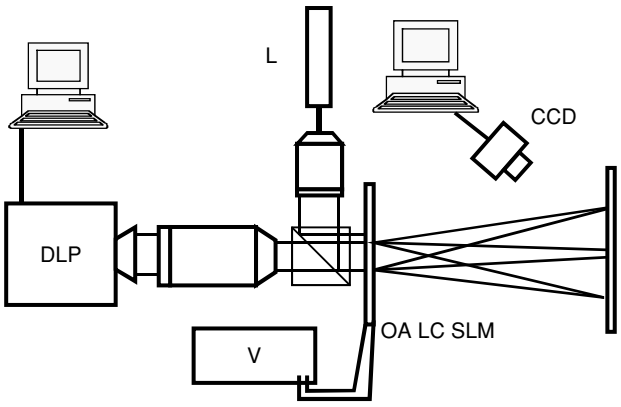


Figure 15. Digital binary hologram projection system with OA LC SLM. DLP—digital light projector, L—laser, V—dc voltage supply, CCD—digital camera.

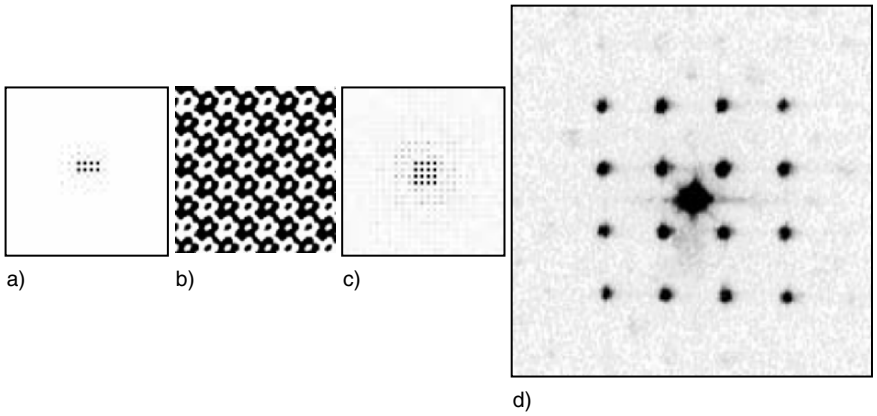


Figure 16. An example of optically reconstructed image (d) from a digital binary hologram (b) generated on a 50×50 pixel matrix for a target (a). The hologram was projected on OA LC SLM via multimedia projector. The symmetry of binary hologram makes that numerical as well as optical reconstruction gives doubled image.

Figure 17. Optically reconstructed image (a) of J.-B. Fourier portrait (b) from a 512×512 pixels four-phase hologram (c).

image of a portrait of J.-B. Fourier, kindly supplied to us by HoloEye Photonics AG, in the form of a four phase hologram.

For the best performance OA LC SLM's required the dc voltage in the range (20–30 V). LC panel response and recovery times to the changing light scenario depend on many factors and for nematics are typically: time-on $\approx 10 - 30$ ms and time-off $\approx 50 - 200$ ms. These limits originate from the photoconductive properties of PVK:TNF polymer and ion mobilities in LC. Motion picture holography is a challenge for many research laboratories.[93–96] The sequential reconstruction of the holograms creates a holographic video. The dynamical change of holograms projected on LC panel at rate of 25 images per second was feasible.

4. CRITICAL COMPARISON OF BACTERIORHODOPSIN PROPERTIES WITH OTHER PHOTOREFRACTIVE MATERIALS

In this chapter we outlined two groups of phenomena as well as two groups of synthetic organic materials designed for optical information processing: photorefractive polymers and liquid crystals and photochromic polymers. We also described the properties of mutated bacteriorhodopsin family. Each group of materials can be tailored for various applications in the field of photonics, optical data storage and processing. Many laboratories and research group have already shown fascinating applications of bacteriorhodopsin and various applications of polymers and liquid crystals. However, it should not be surprising that bR crafted by evolutionary process and further re-optimized by biotechnology shows superior transient holographic properties than any manufactured or synthesized material or group of materials up to date. This is evident when one compares, the described in this chapter, simplicity of the mechanisms underlying photochromic properties in azo-functionalised polymers or photorefractive properties of polymers or liquid crystals in opposition to complex mechanism of photochromic transitions in bacteriorhodopsin. What is advantageous in bR and cannot be obtained in other materials is the structure—a two-dimensional crystalline lattice in purple membrane. Each photosensitive site is surrounded by adapted seven peptide helixes that facilitate the light initiated reaction without changing the surrounding. Therefore bR shows extremely high spatial resolution $\sim$5000 lines/mm. The writing speed of bR can be remarkably high (40 μs for bR $\rightarrow$ M) but using other states and different modified BR forms even shorter switching times were reported (eg. BR $\rightarrow$ J transition: 500 fs). However, the main problem with the utilization of bR as a temporary holographic material is a relatively long recovery time of the M $\rightarrow$ bR transition which can take several hundreds of ms to seconds. This can be overcome by the use of intense blue light that drives M state back to the bR within 50–100 μs. Special structure of bR allows for implementations of polarization holography and polarization based architectures in optical dynamic memories, associative memories, Joint Fourier Transform correlators and others.[98–99] On the other hand liquid crystals combined with polymers promise the highest dynamic refractive index changes a key-factor for dynamic optical data processing.

Nonlinear spectroscopic properties of bR, large two-photon absorption ($\delta = 290 \times 10^{-50}$ cm^{4}s molecules^{-1} photon^{-1}) that permits the efficient activation of the primary event using long wavelength lasers[100,101], large second- and third-order hyperpolarizabilities[102] arising from large change in dipole

moment, $\Delta\mu = 13.5 \pm 0.8$ D that accompanies excitation into the lowest-lying strongly allowed state, make it a promising candidate for all-optical switching vis-à-vis organic molecules and polymers.

We would like to conclude this chapter with a direct comparison of important optical and holographic properties of bR and other artificial organic materials designed for similar purposes. They are summarized in Table 4 and averaged for photorefractive polymers, photorefractive liquid crystals

Table 4. Optical and holographic properties of BR in the visible part of the spectrum versus best photochromic and photorefractive materials for dynamic holographic optical information processing

Property	**Bacteriorhodopsin**	**Photorefractive polymers and LCs**	**Photochromic Polymers**
Resolution (optical)	≥5000 lines/mm	<400 lines/mm <250 lines/mm	<2000 lines/mm
Maximal bleaching ratio	95%	not adequate	90%
Index of refraction	1.47	1.4–1.6 1.4–1.8	1.4–1.6
Refraction index change	0.001–0.01 (depends on OD)	0.01–0.01 0.001–0.2	0.001–0.01
Diffraction efficiency	1–3%, max 7%	100% (int.) 34% (int.)	6.7% max
Light sensitivity	0.1–20 mJ/cm^2	1–30 mJ/cm^2	
Polarization recording	possible	impossible impossible	
Reversibility	≥10^7 cycles	10^4 cycles 10^8 cycles	10^5–10^6 cycles
Shelf life	years	months years	years
Film thickness	10–500 μm, typ.20–40 μm	50–200 μm 2–50 μm	10–1000 μm
Rise and decay times	ms – s	ms – min 10 ms – s	ms – hours
Aperture	unlimited	unlimited ∼15 × 15 cm^2	unlimited
2BC gain coeff. Γ	not reported	10–400 cm^{-1} <4000 cm^{-1}	not reported
2-photon writing	possible	possible possible	possible
Bias electric field	not required	∼50–100 V/μm 0.1–5 V/μm	not required

and photochromic polymers. Some specific materials can exhibit superior properties than those reported in the Table 4.

Analysis of the data contained in Table 4 should be done along the preferred characteristics of the ideal holographic storage medium. One must consider the following features:

1. *Mechanism of optical recording*—should result in production of large dynamic range of refractive index change (e.g., $\Delta n_{\max} \approx 10^{-3} - 10^{-2}$), preferably optically erasable. Large absorption coefficient changes are much less effective though they lead to refractive index change through Kramers-Kronig relation. The ultimate physical limit on the storage density of a medium comes from its finite dynamic range as each of the stored holograms uses a portion of a $\Delta n_{\max}$.
2. *Sensitivity*—material should be responsive to visible light (laser diodes are widely available and cheap). Desired holographic writing sensitivity should be of the order of $10^{-2}\,\mathrm{cm}^3/\mathrm{J}$.
3. *Optical quality*—thick slabs with large surface area should be easily fabricated, showing negligible light scattering. Aggregation, crystallization, material diffusion or shrinkage should be absent. The angular multiplexing of holograms requires thick media. The attainable storage density is proportional to the medium thickness.
4. *Self-processing*—storage medium should be directly addressed by light without need of any kind of developing (e.g., chemical, thermal, magnetic, UV, IR, etc.) both before and after recording.
5. *Stability*—material should retain recorded data indefinitely over a wide range of ambient (temperature, humidity, etc.) conditions. Number of write-read-erase cycles should exceed 10^6. Particularly important is lack of any photochemical side-reactions.
6. *Volatility*—simple ways of fixing of recorded holograms, disabling their partial erasure by subsequent recording and read-out.
7. *Cost*—holographic material for massive optical storage must be available, easily manufacturable and cheap.

The nature of the physical recording process determines most of the properties of the storage medium. In bR the recording mechanism is of photochromic origin therefore a modulation of absorption occurs as the primary event. This fact makes that the material is optically erasable and hence suitable for rewritable memory. However, this also lead to volatility. In order to achieve data persistence the complex methods must be engineered (e.g. two-photon gated recording, electrical or thermal fixing). The clear advantage of photorefractive crystals is their optical quality enabling the achievement of minimum signal-to-noise ratio partially determined by

medium dynamic range, thickness, diffraction efficiency and light scattering. Photorefractive and photochromic polymers are characterized by insufficient dynamic range, typically small thickness and high level of light scattering. The necessity of the use of high voltage in photorefractive polymers prevents them from their widespread use. Contrary to that photochromic polymers and bR can operate without any voltage applied. Photochromes also enable use of light polarization properties to store information by polarization sensitive type of data recording. This offers a vast area of novel optical data storage and processing architectures. On the other side liquid crystals offer the highest dynamic range but are limited to the smallest thickness and therefore are useless for applications of high density data storage. However they properties make them attractive candidates for dynamic holographic applications (e.g. holographic interferometry for strain, vibration and flow analysis, imaging through distorting media, holographic optical tweezers, three-dimensional displays).

5. CONCLUDING REMARKS

Traditional holography based on photographic silver-halide emulsions. However, they fail to meet many requirements of contemporary holographic storage and novel more suitable materials have been elaborated. None of the material for holographic storage considered so far has been able to fulfill all the demands. Instead an arsenal of possible materials, each with a unique set of strengths and weaknesses has emerged: photopolymer films, photorefractive crystals, photochromic films, mutated bacteriorhodopsins. In striking contrast to bR mutants there is no single material among photorefractive polymers, liquid crystals, glasses or photochromic polymers that could be used as broadly as bR and its mutants for many demanding applications. But there is hope that all mentioned above materials will find their niches in the field of future optical technologies.

It should be also noted that apart discussed in this chapter organics there is a vast number of inorganic materials useful in the domain of optical storage and light processing including nanomaterials, quantum-well structures and photonic crystals.

6. ACKNOWLEDGEMENTS

Early and significant conceptual developments of the research was initiated at Children's Hospital, Harvard Medical School and was funded by a

grant from the Office of Naval Research (ONR) and partially from a NIH grant to the Laboratory for Skeletal Disorders and Rehabilitation, Dept. of Orthopaedic Surgery, Children's Hospital and Harvard Medical School. V. Renugopalakrishnan expresses his thanks to National Science Foundation for the award of a MRI grant ECS-0421255. V. Renugopalakrishnan would like to thank Florida International University, Wallace H Coulter Foundation, Office of Naval Research (ONR), and Harvard Medical School for supporting this project. A. Miniewicz is grateful to Wroclaw University of Technology for support. We would like to thank *J. Physical Chemistry* (Ref. 66) and International *J. Quantum Chemistry* (Ref. 68) for copyright permission to reproduce Figs. 10 and 11, respectively. Fig. 12 was taken from PDB data base.

REFERENCES

1. Günter, P. and Huignard, J.-P. (eds.), 1988 *Photorefractive Materials and their Applications*, Vols. 1 and 2, Springer Verlag, Berlin.
2. Ostroverkhova, O. and Moerner, W. E. 2004 "Organic Photorefractives: Mechanisms, Materials, and Applications" *Chem. Revs*. 104: 3267–3314
3. Irie, M. (ed.) 2000 "Photochromism: Memories and Switches", *Chem. Revs*. 100: 1683–1890
4. Oesterhelt, D., Brauchle, C. and Hampp, N. 1991 "Bacteriorhodopsin: a biological material for information processing", *Quart. Rev. Biophys*. 24: 425–478
5. Yablonovitch, E. 1987 "Inhibited spontaneous emission in solid state physics and electronics", *Phys. Rev. Lett*. 58: 2059–2062
6. Joannopoulous, J. D., Meade, R. D. and Winn, J. N. 1995 *Photonic Crystals*, Princeton U. Press, N.J.
7. Birge, R. R. 1990 "Nature of the primary photochemical events in rhodopsin and bacteriorhodopsin", *Biochim. Biophys. Acta*, 1016: 293–327
8. Hampp, N. and Zeisel, D. 1994 "Mutated bacteriorhodopsins versatile media in optical image processing" *IEEE Engineering in Medicine and Biology*, 13: 67–111
9. Birge, R. R., 1990 "Photophysics and Molecular Electronic Applications of the Rhodopsins", *Anu. Rev. Phys. Chem*. 41: 683–733
10. Hampp, N. 2000 "Bacteriorhodopsin as a photochromic retinal protein for optical memories", *Chem. Rev*. 100: 1755–1776
11. Prasad, P. N. and Williams D. J. 1991 *Introduction to Nonlinear Optical Effects in Molecules and Polymers*, John Wiley and Sons, N. Y.
12. Ashkin, A., Boyd, G. D., Dziedzic, J. M., Smith, R. G., Ballman, A. A., Levinstein, J. J. and Nassau K., 1966 "Optically-induced refractive-index inhomogenities in $LiNbO_3$ and $LiTaO_3$" *Appl. Phys. Lett*. 9: 72–74
13. Eichler, H. J., Günter, P. and Pohl, D. 1986 *Laser-Induced Dynamic Gratings*, Springer-Verlag, Berlin/Heidelberg
14. Marcus, R. A. 1964 "Chemical and Electrochemical Electron-Transfer Theory", *Annu. Rev. Phys. Chem*. 15: 155–168

15. Binh, N.T., Gailberger, M. and H. Baessler, 1992 ''Photoconduction in poly(3-alkylthiophene) I. Charge carrier generation'', *Synthetic Metals*, 47: 77–86
16. Meerholz, K., Volodin, B.L., Sandalphon, Kippelen, B. and Peyghambarian, N., 1994 ''A photorefractive polymer with high optical gain and diffraction efficiency near 100%'', *Nature* 371: 497–499
17. Kajzar, F., Bartkiewicz, S. and Miniewicz, A. 1999 ''Optical amplification with high gain in hybrid polymer-liquid crystal structures'', *Appl. Phys. Lett*. 74: 2924–2926
18. Dalton, L. R., Harper, A. W., Wu, B., Ghosn, R., Laquindanum, J., Liang, Z., Hubbel, A. and Xu C., 1995 ''Polymeric electro-optic modulators: materials, synthesis and processing'', *Adv. Mater*. 7: 519–539
19. Moerner, W.E., Silence, S.M., Hache, F. and G.C. Bjorklund, 1994 ''Orientationally enhanced photorefractive effect in polymers'', *J. Opt. Soc. Am*. B11: 320–330
20. Ducharme, S., Scott, J. C., Twieg, R. J. and Moerner, W. E. 1991 ''Observation of the photorefractive effect in a polymer'' *Phys. Rev. Lett*., 66: 1846–1849
21. Liphardt, M., Goonesekera, A., Jones, B.E., Ducharme, S., Takacs, J.M. and Zhang, L. 1994 ''High-performance photorefractive polymers'', *Science*, 263: 367–369
22. Grunnet-Jepsen, A., Thompson, C.L. Twieg, R.J. and Moerner, W.E. 1998 ''Amplified scatering in a high-gain photorefractive polymer'', *J. Opt. Soc. Am*. B15: 901–904
23. Grunnet-Jepsen, A., Thompson, C.L. and Moerner, W.E. 1998 ''Systematics of two-wave mixing in a photorefractive polymer'', *J. Opt. Soc. Am*. B15: 905–913
24. Simoni, F. 1997 *Nonlinear optical properties of liquid crystals and polymer dispersed liquid crystals*, World Scientific, New Jersey
25. Rudenko, E.V. and Sukhov, A.V. 1994 ''Optically induced space charge field in nematics and orientational nonlinearity'', *JETP (in Russian)*, 105: 1621–1634
26. Wiederrecht, G.P., Yoon, B.A., Wasielewski, M.R. 1995 ''High photorefractive gain in nematic liquid crystals doped with electron donor and acceptor molecules'', *Science*, 270: 1794–1797
27. Khoo, I.-C. 1995 *Liquid Crystals*. *Physical Properties and Nonlinear Optical Phenomena*, Wiley, New York
28. Khoo, I.-C. 1996 ''Orientational photorefractive effects in nematic liquid crystals films'', *IEEE J. Quant. Electron*. 32: 525–534
29. Bartkiewicz, S., Miniewicz, A., Januszko, A., and Parka, J. 1996 ''Dye-doped liquid crystal composite for real-time holography'', *Pure and Applied Optics*, 5: 799–809
30. Komorowska, K., Miniewicz, A. and Parka, J., 2000 ''Holographic grating recording in large area photoconducting liquid crystal panels'', *Synth. Metals*, 109: 189–193
31. Pagliusi, P., Macdonald, R., Busch, S., Cipparone, G. and Kreuzer, M. 2001 ''Nonlocal dynamic gratings and energy transfer by optical two-beam coupling in a nematic liquid crystal owing to highly sensitive photoelectric reorientation'', *J. Opt. Soc. Am. B* 18: 1632–1638
32. Date, M., Takeuchi, Y. Kato, K., 1999 ''Droplet size effect on the memory-mode operating temperature of smectic-A holographic polymer–dispersed liquid crystal'' *J. Phs. D. Appl. Phys*. 32: 3164–3168
33. Han, J. W., Kang, T. J., Park, G., 2000 ''Effects of composition, curing-time, and temperature on the electrooptical characteristics of polymer-dispersed liquid crystal films'', *Journal of the Korean Physical Society*, 36: 156–163
34. Golemme, A., Volodin, B. L., Kippelen, E., Peyghambarian, N. 1997 ''Photorefractive polymer-dispersed liquid crystals'', *Opt. Lett*., 22: 1226–1229
35. Ono, H. and Kawatsuki, N. 1997 ''Orientational photorefractive effects observed in polymer-dispersed liquid crystals'', *Opt. Lett*., 22: 1144–1146

36. Cipparone, G., Mazulla, A., and Simoni, F. 1997 ''Writing and erasure of holographic gratings in dye doped PDLC'', *Mol. Cryst. Liq. Cryst*., 299: 329–336
37. Golemme, A., Kippelen, E., Peyghambarian, N. 1998 ''Highly efficient photorefractive polymer-dispersed liquid crystal'', *Appl. Phys. Lett*., 73: 2408–2410
38. Helfrich, W. 1969 ''Conduction-induced alignment of nematic liquid crystals: basic model and stability considerations'', *J. Chem. Phys*., 51: 4092–4105
39. Simoni, F., Francescangeli, O. 1999 ''Effects of light on molecular orientation of liquid crystals'', *J. Phys. Condens. Mater*, 11: R439–R487
40. Wiederrecht, G.P., Yoon, B.A. and Wasielewski, M.R., 2000 ''Photorefractivity in ferroelectric liquid crystal composites containing electron donor and acceptor molecules'', *Adv. Mater*., 12(20): 1533–1536
41. Khoo, I.C., Shih, M. Y., Shishido, A., Chen, P. H., and Wood, M. V. 2001 ''Liquid crystal photorefractivity – towards supra-optical nonlinearity'', *Opt. Mater*. 18: 85–90
42. Jánossy, I. and Lloyd, A. D., 1991 ''Low-power optical reorientation in dyed nematics'', *Mol. Cryst. Liq. Cryst*. 203: 77–84
43. Miniewicz, A., Bartkiewicz, S., Turalski, W. and Januszko, A. 1997 ''Dye-doped liquid crystals for real-time holography'', in R. W. Munn, A. Miniewicz and B. Kuchta (eds.), *Electrical and Related Properties of Organic Solids*, NATO ASI Series, 3/24, Kluwer Academic Publishers, Dordrecht, 323–337
44. Pagliusi, P., and Cipparone G., 2002 ''Surface-induced photorefractive-like effects in pure liquid crystals'', *Appl. Phys. Lett*. 80: 168–170
45. Ono, H., and Kawatsuki, N. 1997 ''Orientational holographic grating observed in liquid crystals sandwiched with photoconductive polymer films'', *Appl. Phys. Lett*. 71: 1162–1164
46. Komorowska, K., Miniewicz, A., Parka, J., Kajzar, F., 2002 ''Self-induced nonlinear Zernike filter realized with otically addressed liquid crystal spatial light modulator'', *J. Appl. Phys*., 92: 5635–5641
47. Kajzar, F., Bartkiewicz, S. and Miniewicz, A. 1999 ''Optical amplification with high gain in hybrid polymer-liquid crystal structures'', *Appl. Phys. Lett*., 74: 2924–2926
48. Bartkiewicz, S., Matczyszyn, K., Miniewicz, A. and Kajzar, F., 2001 ''High gain of light in photoconducting polymer – nematic liquid crystal hybrid structures'', *Opt. Commun*., 187: 257–261
49. Dürr, H., Bouas-Laurent, H. (eds), 1990 *Photochromism: Molecules and Systems*, Elsevier, Amsterdam
50. Crano, J. C., Guglilmetti (eds), 1999 *Organic Photochromic and termochromic Compounds,* Kluwer Academic / Plenum Publishers, New York
51. Berkovic, G., Krongauz, V., Weiss, V., 2000 ''Spiropyrans and Spirooxazines for memories and switches'', *Chem. Rev*. 100(5): 1741–1754
52. Delaire, J.A., Nakatani, K. 2000 ''Linear and Nonlinear Optical Properties of Photochromic Molecules and Materials'', *Chem. Rev*. 100(5): 1817–1746
53. Yokoyama, Y., 2000 ''Fulgides for Memories and Switches'', *Chem. Rev*. 100(5): 1717–1740
54. Meng, X., Natahnson, A. Barett, C., and Rochon, P. 1996 ''Azo Polymers for Reversible Optical Storage. Cooperative Motion of Polar Side Groups in Amorphous Polymers'', *Macromolecules*, 29: 946–952
55. Blinov, L. M., Kozlovsky, M. V. and Cipparone, G. 1999 ''Photochromism and Holographic Grating Recording on a Chiral Side-Chain Liquid Crystalline Copolymer Containing Azobenzene Chromophores'', *Chem. Phys*., 245: 473–481

56. Jiang, X. L., Li, L., Kumar, J., Kim, D. Y., and Tripathy, S. K. 1988 ''Unusual polarization dependent optical erasure of surface relief gratings on azobenzene polymer films'', *Appl. Phys. Lett*. 72: 2502–2504
57. Natansohn, A., Rochon, P., 2002 ''Photoinduced Motion in Azo-Containing Polymers'', *Chem. Rev*., 102: 4139–4175
58. Naydenova, I., Nikolova, L., Todorov, T., Holme, N.C.R., Ramanujam, P. S., Hvilsted, S. 1998 ''Diffraction from polarization holographic gratings with surface relief in side-chain azobenzene polyesters'', *J. Opt. Soc. Am. B*, 15:1257–1265
59. Fischer, M., El Osmanb, A., Blanche, P.-A., Dumont, M., 2000 ''Photoinduced dichroism as a tool for understanding orientational mobility of photoisomerizable dyes in amorphous matrices'', *Synthetic Metals*, 115: 139–144
60. Dumont, M. 1996 ''A general model for optically induced molecular order in amorphous materials via photoisomerisation'', *Nonl. Opt*. 15: 69–72
61. Huang, T., Wagner, K.H., 1996 ''Diffraction analysis of photoanisotropic holography: an anisotropic saturation model'', *J. Opt. Soc. Am. B*, 13: 282–299
62. Natansohn, A., and Rochon, P. 1999 ''Photoinduced motions in azobenzene-based amorphous polymers: possible photonic devices'', *Adv. Mater*. 11: 1387–1396
63. Yager, K. G., Barrett, C. J., 2001 ''All-optical patterning of azo polymer films'', *Current Opinion in Solid State and Materials Science* 5: 487–494
64. Sutter, K., Günter, P., 1990 ''Photorefractive grating in the organic crystal 2- cyclooctyloamino-5-nitropyridine doped with 7,7,8,8-tetracyanoquino-dimethane'' *JOSA B* 7: 2274–2279
65. Rocha, L., Dumarcher, V., Denis, C., Raimond, P., Fiorini, C., Nunzi, J., 2001 ''Laser emission in periodically modulated polymer films'', *J. Appl. Phys*. 89: 3067–3071
66. Birge, R. R., Gillespie, N. B., Izaguirre, E. W., Kusnetzow, A., Lawrence, A. F., Singh, D., Song, Q. W., Schmidt, E., Stuart J. A., Seerharaman, S., and Wise K. J., 1999 ''Biomolecular Electronics: Protein-Based Associative Processors and Volumetric Memories'', *J. Phys. Chem*. 103: 10746–10766
67. Wise, K. J., Gillepsie, N. B., Stuart, J.A., Krebs, M. R., and Birge, R.R. 2002 ''Optimization of bacteriorhodopsin for bioelectronic devices''. *Trends in Biotechnology* 20: 387–394
68. Renugopalakrishnan, V., Strzelczyk, A., Li, Pingzou, Mokhnatyuk, A. A.. Gursahani, S. H., Nagaraju, M. Prabhakaran, M. Arjomandi, H., Lakka, S. L. 2003 ''Retro-engineering bacteriorhodopsins: Design of smart proteins by bionanotechnology''. *International Journal of Quantum Chemistry* 95: 627–631
69. Renugopalakrishnan, V., Strzelczyk, A., Mokhnatyuk, A. A., Pingzou Li, Senthil, D., Monfort, J., Duzgunes, N., Balschi, J. A., Ram Mohan, Samet, M., and Grimes, D. E. 2002 ''BioFold's high performance bacteriorhodopsin 192-based storage devices'', *Insight* 11:14–19
70. Renugopalakrishnan V., Lindvold L.R., Pingzuo Li, Arjomandi, H., Litvinov, D., Chomko, R., and Khizroev, S., unpublished.
71. Renugopalakrishnan, V. and Pingzuo Li, ''Design of a Bacteriorhodopsin, bR 192, with seven point mutations and its expression'' in P. Pastoris: ''Stabilization of the M and O photo intermediates with enhanced lifetimes'', US Provisional Patent (pending).
72. Moser, A., Takano, K., Margulies, D. T., Albrecht, M., Sonobe, Y., Ikeda, Y., Sun, S. and Fullerton, E. E., 2002, ''Magnetic recording: advancing into the future'' *J. Phys. D: Appl. Phys*. 35: R157–R167
73. (Lindvold, unpublished).

74. Tatke, S. S., Renugopalakrishnan, V., and Prabhakaran, M. 2004, ''Interfacing biological macromolecules with carbon nanotubes and silicon surfaces: a computer modelling and dynamic simulation study'', *Nanotechnology* 15: S684–S690
75. T. G. Ebrey, ''Light energy transduction in bacteriorhodopsin. In Thermodynamics of Membrane Receptors and Channels'', 1993 Ed. M. Jackson, CRC Press, Boca Raton, FL, USA, pp. 353–387
76. Xi, B., Wise, K. J., Stuart, J. A., and Birge, R. R. 2005 ''Bacteriorhodopsin-based 3D optical memory'', in V. Renugopalakrishnan (eds.): ''Bionanotechnology'', Springer Verlag, 39–59.
77. Gerwert, K., Hess, B., Soppa, D., Osterhelt, D. 1989 ''Role of aspartate-96 in proton translocation by bacteriorhodopsin'' *Proc. Natl. Acad. Sci*. USA 86: 4943–4947
78. Marinetti, T., Subramaniam, S., Mogi, T., Marti, T., Khorana, H. G., 1989 ''Replacement of asparatic residues 85, 96, 115, or 112 affects the quantum yield and kinetics of proton release and uptake by bacteriorhodopsin'' *Proc. Natl. Acad. Sci. USA* 86: 529–533
79. Tittor, J., Soell, C., Osterhelt, D., Butt, J. J., Bamberg, E. 1989 ''A defective proton pump, point-mutated bacteriorhodopsin Asp96—Asn is fully reactivated by azide'' *EMBO J*. 8: 3477–3482
80. Otto, H., Mareti, T., Holz, M., Mogi, T., Lindau, M., Khorana, H. G., Heyn, M. P., 1989 ''Aspartic acid-96 is the internal proton donor in the reprotonation of the Schiff base of bacteriorhodopsin'' *Proc. Natl. Acad. Sci. USA* 86: 9228–9232
81. Miller, A., Osterhelt, D., 1990 ''Kinetic optimization of bacteriorhodopsin by asparatic acid 96 as an internal proton donor'' *Biochim. Biophys. Acta* 1020: 57–64
82. Subramaniam, S., Lindahl, M., Bullough, P., Faruqi, A. R., Tittor, J., Osterhelt, D., Brown, L., Lanyi, J., Henderson, R., 1989 *J. Mol. Biol*. 287: 145–161
83. Luecke, H., Schobert, B., Richter, H.-T., Cartailler, J.-P Lanyi, J. K. 1999 ''Bacteriorhodopsin during ion transport at 2 Angstrom resolution'', *Science* 286: 255–260
84. Pingzuo Li, Anumanthan A., Xiu-Gao, Gong, Nagaraju, **M.,** Duzgunes, N., Suzura, V., Renugopalakrishnan V.. ''Review on Expression of Proteins in *Pichia Pastoris*'', 2004, submitted *Applied Biochem. and Biotech*. Submitted.
85. Peda'el, A., Daisy, R., Horowitz, M., Fischer, B. 1998 ''Beam coupling induced transparency in a bacteriorhodopsin-based saturable absorber'' *Opt. Lett*. 23: 1173–1175
86. http://www.optware.co.jp
87. http://www.aprilisinc.com
88. Powell, R. L., and Stetson, K. A., 1965 ''Interferometric vibration analysis by wavefront reconstructions'', *J. Opt. Soc. Am*. 55: 1593–1598
89. Dufresne, E., Spalding, G., Dearing, M., Sheets, S., and Grier, D. 2001 ''Computer generated holographic optical tweezer arrays'', *Rev. Sci. Instrum*., 72: 1810–1816
90. Huebschman, M. L., Munjuluri, B., and Garner, H. R., 2003 ''Dynamic holographic 3-D image projection'', *Optic Express*, 11: 437–445
91. Miniewicz, A., Gniewek, A., Parka, J., 2002 ''Liquid crystals for photonic applications'', *Opt. Mater*. 21: 605–610
92. Miniewicz, A., Mysliwiec, J., Gryga, L., and Kajzar, F., 2004 ''Retrieval of computer generated holograms projected onto liquid crystal - photoconducting polymer system'', *Proc. of SPIE*, 5351: 134–143
93. Wick, D. V., Martinez, T., Wilkes, J. M. and Gruneisen, M. T., 1998 ''Development of a novel liquid crystal spatial light modulator for real-time holography'', *Proc. SPIE*, 3475: 42–48

94. Ito, T., Shimobaba, T., Godo, H., Hoiruchi, M. 2002 ''Holographic reconstruction with a 10μm pixel-pitch reflective liquid-crystal display by use of a light-emitting diode reference light'', *Opt. Lett*., 27: 1406–1409
95. Hast, T., Schonleber, M., Tiziani, H. J., 1997 ''Computer-generated holograms from 3D-objects written on twisted-nematic liquid crystal displays'', *Opt. Commun*., 140: 299–304
96. Sutkowski, M. and Kujawinska, M., 2000 ''Application of liquid crystals (LC) for optoelectronic reconstruction of digitally stored holograms'', *Optics & Lasers in Engineering*, 33: 191–203
97. Okada-Shudo, Y., Yamaguchi, I., Tomioka, H., and Sasabe, H. 1996 ''Real-time image processing using polarization discrimination of bacteriorhodopsin'', *Synth. Metals*, 81: 147–149
98. Joseph, J., Aranda, F. J. Rao, D. V., Akkara, J. A., and Nakashima, M. 1996 ''Optical Fourier processing using photoinduced dichroism in a bacteriorhodopsin film'', *Opt. Lett*., 21: 147–150
99. Birge, R. R., 1992 ''Protein-based optical computing and memories'', *Computer*, 11: 56–67
100. Birge, R. R. and Zhang, C. F. 1990 ''Resonant optical rectification on bacteriorhodopson'' *J. Chem. Phys*. 92: 7178–7195
101. Birge, R. R., Gross, R. B., Masthay, M. B., Smart, J. A., Talent, J. R. and Zhang, C. F. 1992 ''Protein-based Two-Photon Three Dimensional Memories'' *Mol. Cryst. Liq. Cryst. Sci. Technol. Sect. B. Nonlinear Opt*. 3: 133–147
102. Huang, J. Y., Chen, Z. and Lewis, A. 1989 ''Second-harmonic generation in purple membrane-poly(vinylalcohol) films: probing the dipolar characteristics of the bacteriorhodopsin chromophore in bR 570 and M 412'' *J. Phys. Chem*. 93: 3314–3320

Chapter 13

SUBMONOLAYER MEASUREMENTS OF ADSORBED PROTEINS IN MICROFLUIDIC CHANNELS

David Henry[2], Keith Lenghaus[1,2], Kerry A. Wilson[1,2], Melissa Hirsch-Kuchma[1], Jerry Jenkins[3], Shankar Sundaram[3] and James J. Hickman[1,2]
[1]*University of Central Florida, 12424 Research Parkway, Suite 400, Orlando, FL 32826, USA*
[2]*Bioengineering Department, Clemson University, Clemson, SC 29634, USA*
[3]*CFD Research Corporation, Huntsville, AL 35805, USA*

Key words: Microfluidics, Protein Adsorption, CFD, XPS, Simulation, Flow
Abstract: Microelectromechanical systems (MEMS), and soon nanoelectromechanical systems (NEMS), are projected to have a major impact in systems for toxin detection, proteomics and biomedical applications. This chapter describes experimental systems that were developed in our labs, as well as those being developed elsewhere, that are being used to understand protein deposition at sub-monolayer coverages in these devices. Modeling tools, used in conjunction with experimental results, can also be utilized for predicting the interactions of proteins with these microsystems, under static and flow conditions, with the goal of creating biocompatible MEMS devices that can be easily integrated with biologically based assay systems. The long-term combination and integration of biological and electronic components requires a thorough understanding of surface/biomolecule interactions in these devices. The use of surface modification techniques has allowed the tailoring of the interface between biological/nonbiological materials, independent of the bulk composition of the nonbiological material, to enable reproducible experimental design. Our results to date have indicated we can measure protein adsorption down to less that 1% of a monolayer and that static results are very different than those observed under flow conditions, and this has been supported by the modeling results. Thus, data obtained under static conditions cannot be used as a predictor for behavior under flow conditions, which should be the predominant situation in most biological assay systems.

V. Renugopalakrishnan et al. (eds.), Bionanotechnology, 257-274.

1. INTRODUCTION

Microelectromechanical systems (MEMS), and soon nanoelectromechanical systems (NEMS), are projected to have a major impact on biological system analysis for toxin detection, proteomics, and biomedical applications. Many existing biological assays and analysis techniques can achieve very low detection limits in biomedical and biotechnological devices. Anticipated improvements in performance will push these limits even closer to ultimate sensitivities. As detection limits fall, there has been a corresponding trend toward handling smaller and smaller volumes of material, resulting in an increase in the surface to volume ratio of the analysis systems. Under these conditions, even apparently minor nonspecific adsorption (perhaps less than 1% of the total available surface) could severely or completely deplete the analyte stream of the molecule of interest before the sample reaches the detector. Thus, there is a pressing need for tools that can resolve surface interactions with submonolayer precision, both experimentally and by utilizing simulation tools. Progress toward this goal is presented in this monograph, as well as other options that are available to determine low levels (submonolayer) of surface interactions with biological macromolecules.

This monograph also demonstrates that simulation and modeling tools, combined with experimental findings, can be used to predict protein interactions with surfaces in MEMS devices, even at submonolayer coverages. Taken as a whole, this research indicates that existing systems can be vulnerable to failure by sample depletion. The work presented in this chapter focuses on protein/surface behavior, but investigators are encouraged to see this system as analogous to many other situations in which small amounts of surface active materials are being handled under microfluidic conditions.

2. CONSEQUENCES OF NONSPECIFIC ADSORPTION OF BIOLOGICAL MACROMOLECULES IN MEMS AND NEMS DEVICES

As a consequence of the ongoing improvements in instrument sensitivity, routine detection of substances at the single molecule limit is being postulated. Along with the movement toward lower and lower detection limits, device size, and thus, sample volumes are likewise becoming smaller and smaller, increasing the surface area to volume ratio in analysis systems. As a consequence, those who develop or anticipate using equipment with these increased capabilities must face the possibility of false negative or false positive errors

due to the adsorption or desorption of small numbers of molecules in the devices. As the relative amount of available surface goes up, the opportunities for adverse interactions with the wall increase. Since sensors and other microfluidic devices are getting smaller and projecting much lower detection limits, producers and end users of these new devices have to worry about, and deal with, protein adsorption at much less than a monolayer of coverage and possibly less than 1% coverage of the total surface area.

On the basis of previous points, adsorption of less than a monolayer could lead to depletion of a molecule of interest from the analyte stream in many devices being envisioned before it reached the detector. Figure 1 shows the intersection of detection limits and available sensor surface area in terms of the amount of surface required to deplete the sample at a given detection limit for some selected devices. The trend for all assays is to use the least amount of material possible, and it can be readily shown that slightly suboptimal surfaces can severely degrade a sample stream. Capillary electrophoresis, for instance, is used to analyze (principally) protein, in quantities between 10^{-12} and 10^{-15} moles in devices ranging from 10^6 to 10^8 μm^2 (the lower size limit covers CE in MEMS microdevices). The shaded area does not represent all this range but draws attention to the fact that a 10^6 μm^2 device surface will completely deplete a 10^{-15} mole sample at only ~5% surface coverage, and a 10^8 μm^2 device can remove 10^{-12} moles at ~30% coverage. The other three shaded areas indicate the operating size and surface coverage of a tunable

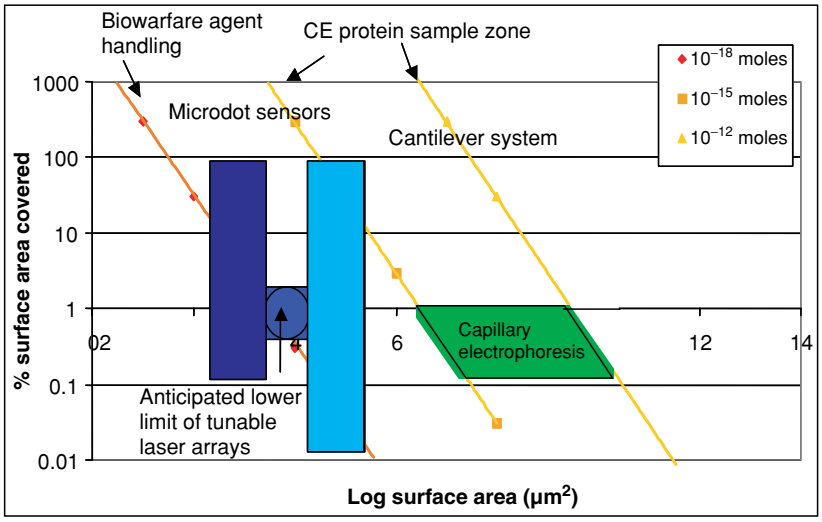


Figure 1. Detection limits, surface area, and depletion. This diagram shows the intersections between surface area, surface coverage, and the amount of material that is introduced into a given system.

laser array, a cantilever detector from University of California, Berkeley[1] and various microsensors under development in the $100 - 1000\ \mu m^2$ range.

Imagine a clean, freshly prepared surface, prepared under the best, most ideal conditions imaginable. Is it homogeneous? With very few exceptions, the answer to this will be no. No matter how fastidious the surface scientist, some heterogeneities will be found: steps in an imperfectly cleaved crystal plane, dislocations, and other flaws. These regions have measurably different adsorption properties from the bulk surface.[2,3] By analogy, the microfluidic surfaces that encounter biological molecules are not going to be homogeneous either, nor can their adsorption properties be validly expressed by means of an average overall behavior. Viewed from the perspective of a fresh surface encountering a stream containing biological molecules, many different sites for adsorption will be present, with different intrinsic rates of adsorption, and these differences are of particular concern when relatively low numbers of molecules interacting with adsorption sites is considered.

Adsorption can be especially complex in those systems in which the molecule/wall interaction is transient, i.e., where there is flow in the system. Adsorption is generally assumed to be occurring under thermodynamic control, that is, the system can, and does, rapidly reach a state that is more or less steady, and this corresponds to the local thermodynamic equilibrium. Although outside the scope of this chapter, the possibility of thermodynamically favorable changes in the conformation of a biomolecule, especially proteins, means that the point at which true thermodynamic equilibrium occurs is particularly difficult to pin down or even define. Given a range of adsorption site energies, with a range of adsorption behavior, and under flow conditions, the more energetically favorable sites will still be able to rapidly adsorb and hold molecules. Less favorable sites that might have been capable of initiating adsorption under static conditions will be unable to do so, or less effective at doing so, under flow conditions. To the experimenter, this would appear as the rate of adsorption being a function of flow. In gas adsorption, similar problems in measuring the true available surface area are faced, as a consequence of a heterogeneous surface, although for different reasons.[4,5]

3. TECHNIQUES FOR QUANTIFICATION OF PROTEIN ADSORPTION AT SUBMONOLAYER ADSORPTION

The following material outlines some of the options available to study protein adsorption at fractional, submonolayer coverage. No single technique is best, and this is by no means a complete list, but each of these is capable of

providing reliable information about different aspects of protein adsorption under microfluidic conditions. Various options, such as protein assays,[6–8] mass spectroscopy,[9] Raman spectroscopy,[10,11] X-ray photoelectron spectroscopy,[12] hollow cathode analysis,[13,14] BCA assays,[15] SDS-polyacrylamide gel electrophoresis (PAGE),[16] enzyme mass balance assays,[17] and standard Ag staining[18] methods are techniques that have been used by various researchers, including in our laboratory, to study protein adsorption. In addition, results from our laboratory that combine some of the previously mentioned techniques with computational fluid dynamics (CFD) to analyze and model the protein/surface interactions in microfluidic channels are presented. Figure 2 compares the sensitivity of these methods to those developed in our laboratory.

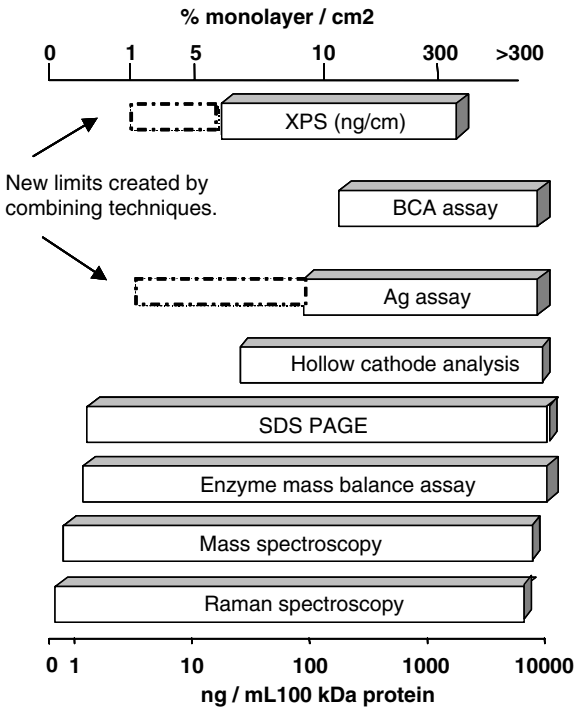


Figure 2. Comparison of the sensitivities of various protein detection techniques. The ability to resolve protein adsorption from 0.1% of a monolayer to, in some cases, over 100% coverage, and multilayer formation has been demonstrated as shown in this monograph. This range of coverage strongly overlaps with the operating limits of microdevices under development. This is, however, by no means an exhaustive description of all applicable techniques.

3.1 Studying Adsorption by Mass Balance

It is possible to measure the total amount of protein lost inside a microfluidic channel by comparing what comes out of the channel with what is known to have been put in, and such measurements can be surprisingly sensitive.[17] Although the detection limits of chemical based total protein assays, such as BCA, are, at the time of writing, still relatively insensitive, assays based on measuring the activity of enzymatic proteins offer comparable sensitivity to those obtainable with radioactively labeled proteins, without the added concerns of the labeling process altering the adsorption properties of the protein in some fundamental way.[6]

The popularity of enzyme linked immunosorbent assays (ELISAs) means that several different enzymes and their assays are readily available off the shelf. We have published results obtained from alkaline phosphatase, shown in Fig. 3, and glucose oxidase[17] and have shown that we can readily resolve changes in adsorbed material in an area as small as 2×10^{-8} m^2. By comparison, a microchannel with a length of 10 mm and a square cross section with a 50 μm diameter has a total available surface of 2×10^{-6} m^2. Resolution equivalent to 1% (or better, depending on the enzyme and the assay) of a monolayer is thus possible.

3.2 Mass Spectroscopy

Assuming that a microchannel surface can be exposed and placed in a vacuum chamber, then various mass spectroscopy techniques can provide an indication of how much adsorbed material is present.[19] Time of flight second-

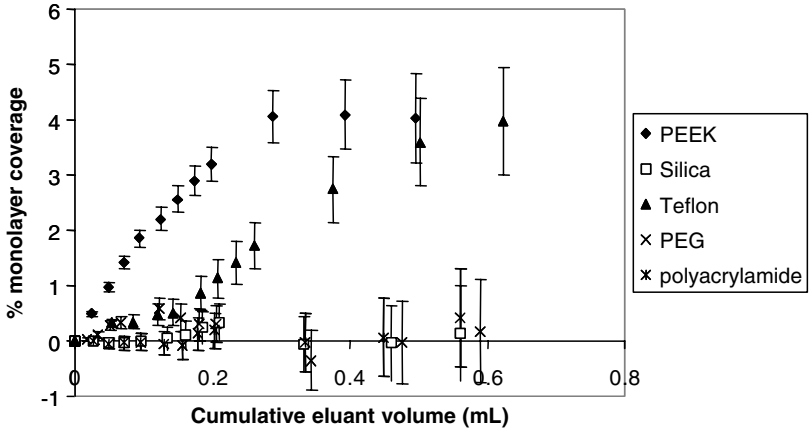


Figure 3. Alkaline phosphatase adsorption onto five capillaries of different composition. Enzyme concentration = 30 ng/mL, flow rate = 0.1 mL/h.

ary ion mass spectroscopy (ToF-SIMS) is attracting a wide amount of attention for this application. The method involves bombarding the surface with ions, which in turn ablate off material present on the surface; in the case of adsorbed proteins, a wide array of ionized organic species are generated, which can be collected and analyzed by the ToF mass spectroscopy system.

The choice of ionic species to generate the secondary ions is important. Some ions are much more effective at generating secondary species than other,[20,21] and the investigation of different ion types is ongoing. Polyatomic species, such as SF_5^+, offer much better sputter rates than monatomic species when dealing with organic films.[9] The spatial resolution of ToF-SIMS can be very good, with spot sizes down to 50 nm with some ionic species, but such a small spot size may result in an insufficient material being ablated, however, depending on the amount of protein adsorbed.

The type and amount of ionic species generated provide a fingerprint of the protein adsorbed on the surface, and research with multivariate analysis techniques have shown that simple binary mixtures of proteins can be distinguished using ToF-SIMS. By appropriately weighting certain characteristic ions in the mass spectrum, the presence of two different proteins on the same surface can be resolved and their relative amounts estimated.[21–23]

Protein adsorption has also been studied using matrix-assisted laser desorption/ionization (MALDI) mass spectroscopic techniques.[24,25] A protein-bearing surface is coated with the matrix material, which is then laser ablated as normal. In this case, only relatively weakly adsorbed proteins are removed and detected, and this method, in theory, can distinguish weakly bound from strongly bound material. It may provide an indication of relative binding strengths of different proteins in an adsorbed mixture, or the changes in binding, which occur on different surface types.

A recent, and potentially exciting, development allows for the generation of ions from a surface under ambient conditions. Desorption electrospray ionization (DESI) avoids the need to place a sample into a vacuum chamber.[26] Instead, a stream of ions is generated by electrospray ionization, which in turn ablate ionic species off a surface, even under normal pressures. These desorbed ions are then allowed into a mass spectrometer fitted with an atmospheric inlet. Detection limits of 10–50 pg of lysozyme are claimed in the text of this article. Some spatial resolution is described as well, although resolution at this point is limited.

3.3 Raman Spectroscopy

Raman spectroscopy (RS) measures the vibrational energies of molecules by the inelastic scattering of an incident laser sourced photon and can be a powerful method for elucidating the structure of proteins in solution without

the need for additional labeling.[27] The technique provides information on bond energies, which is complementary to that obtained with infrared (IR) spectroscopy, and like IR, RS can distinguish the vibrational modes of protein amide bonds with sufficient resolution to provide information about the secondary structure of proteins. However, the sensitivity of conventional RS is limited by the relatively low photon scattering cross-section of the typical atomic constituents of proteins. There is also the possibility of sample damage and strong fluorescence background from extracts.

Many of these issues have been overcome using a method known as surface-enhanced Raman scattering (SERS), although the exact mechanism of this enhancement is not known to date. A biomolecule is physically adsorbed to colloid or nanoparticle suspensions of certain transition metals, particularly Ag, and using this method the detection of single molecules of protein has been reported.[28] SERS also maintains the ability to distinguish alpha helix, beta sheet, and random coil protein conformations, allowing the conformation of the adsorbed protein to be investigated. Additionally, since only atoms close to the metal surface contribute to the SERS effect, those vibrational modes that show significant SERS activity can be concluded to be close to the surface, allowing protein orientation to be assessed. Podstawka and coworkers have determined unique spectra for individual amino acids adsorbed to silver colloid surfaces and have further shown that this method can be expanded to the entire protein adsorbed to the metal with structural information elucidated.[29–31]

To date, however, no data have been presented on SERS studies of proteins adsorbed in microchannels. Typically, SERS studies have focused on suspensions of proteins and colloids/nanoparticles rather than studies of preadsorbed layers of protein on a microchannel wall. This method, however, could easily yield useful information in this area as the spot size of the incident light source can be focused below 2 μm, and the typical dimensions of a microchannel are 50 μm or greater. Connaster and coworkers have demonstrated SERS detection of biomolecules on silver colloids immobilized in polydimethylsiloxane (PDMS) microchannels.[10] This approach does not probe the intrinsic affinity of a biomolecule for a microchannel wall but does show the applicability of SERS detection in microchannels. Furthermore, studies by Cao and coworkers on SERS detection on protein microarrays could be directly applicable to the study of proteins adsorbed in microchannels. These studies have shown that Raman-dye labeled nanoparticles with conjugated antibodies can be used to selectively detect physically adsorbed proteins in microarray configurations.[32] It is at least conceivable that this technique could be applied to physically adsorbed proteins in microchannels. By using various antibody labels for specific protein epitopes, it could also be possible to obtain information on surface density, as well as the adsorbed conformation of the proteins or biological macromolecules in microfluidic devices.

3.4 X-Ray Photoelectron Spectroscopy (XPS)

Applying XPS to the detection and quantification of adsorbed proteins on a surface has allowed surface concentrations in the range of 10 ng/cm^2 (3% monolayer equivalent, assuming a 100-kDa protein with a footprint of 50 nm^2) to be detected.[9] Several variations on the method have been employed to monitor the amount of adsorbed protein and the distribution within the layer. Usually, the percentage of nitrogen or the nitrogen/carbon atomic ratio is used to quantify the amount of protein on a surface. However, when a surface, such as Al, Si, or a transition metal, is used to adsorb the protein, then the attenuation of the substrate's XPS signal can be monitored to determine protein adsorption. With gold substrates, the Au 4f signal has been used to monitor the protein layer's thickness, while on silicon, the Si 2p signal has been monitored.[33,34] As adsorption occurs, the Au or Si signal decreases, owing to the short mean free path of photoemitted electrons, and the thickness of the layer can then be calculated by comparing the intensities of the signals of the covered surface with the bare surface (Briggs and Seah). This information can be especially useful when used in conjunction with ellipsometry. By providing information regarding the range of surface coverage and the thickness of the layer, XPS allows the amount of protein adsorbed to be determined. This method is limited, however, to adsorbed layers of less than 5–8 nm in thickness, as beyond this the attenuation of the substrate signal is too high to allow the layer thickness to be readily measured.[35]

The amount of protein adsorbed depends on the type of surface and its topography. X-ray photoelectron spectroscopy allows different types of surfaces to be evaluated for their resistance to protein adsorption for possible use in medical devices or as biosensors.[36,37] It is also possible to measure the rate of protein deposition via XPS; adsorption kinetic studies of human serum albumin (HSA) on TiO_2 surfaces have been performed by monitoring the increase in carbon and nitrogen intensities and the decrease in titanium and oxygen intensities.[37] If the substrate gives interfering XPS signals, such as with an amine monolayer, the amide carbon binding energy can be monitored in order to determine the protein coverage.[9]

Static ToF-SIMS has been used in conjunction with XPS and radiolabeling to characterize protein surfaces.[38,39] As mentioned previously, a ToF-SIMS can distinguish between different adsorbed proteins, which XPS cannot. XPS has also been used to study proteins adsorbed on organic polymers. Halogenated proteins were used as markers to aid in the analysis of the spectra. In addition, a quartz crystal microbalance (QCM) was also used in conjunction with the halogen counts from the XPS spectra to quantitate the amount of protein adsorbed.[40]

Angle-dependent XPS, in which the variation in peak intensity with emission angle can be determined, can provide additional information owing to the short electron mean free path in which the change in peak intensity with angle is related to how deep the emitting atom is located from the surface. This method has been used to determine how a metalloprotein was bound to an alkanethiol SAM.[41] In this case, the orientation of azurin upon adsorption to the surface via sulfur binding to gold was monitored by the sulfur peaks at different locations on the molecule because the disulfide group coordinated to the gold surface is nearly opposite to the sulfur atoms coordinated to the copper center; thus, sulfur signals from the protein can be monitored at the surface and in the upper layer of the surface simultaneously.

Combined with other methods, such as atomic force microscopy, surface plasmon resonance (SPR),[42] and scanning electron microscopy (SEM) or scanning tunneling microscopy (STM),[41,43,40] additional information regarding a protein layer can be obtained. X-ray photoelectron spectroscopy and atomic force microscopy (AFM) used together provide information about the degree of coverage of the adsorbed protein, the thickness of the layer, and the conformation of the protein.[44,33,35,37] A comparison study of protein adsorption on hydrophobic and hydrophilic functionalized surfaces on both smooth and rough substrata using XPS and AFM demonstrated that the adsorbed amount is dependent on the surface chemistry, but the supramolecular organization of the layer is dependent on the surface chemistry and the topography of the underlying substratum.[45]

4. SIMULATION AND MODELING OF PROTEIN ADSORPTION

CFDRC, of Huntsville, AL, has developed a fluidics-resolved model that evaluates protein adsorption, accounting for both the fluidic transport and the biochemical kinetics in complex microfluidic systems.[46] Traditionally, two different methods have been used to extract kinetics from adsorption data. The simplest approach involves neglecting nonideal (convective–diffusive) effects by assuming well-mixed protein in contact with an idealized surface. Coupled with pseudo-second order kinetics, this approach leads to the classical analytical expression for the Langmuir adsorption isotherm,[47] which is regression fitted to experimental data to calculate the kinetic constants. However, this method can lead to large errors in situations in which transport is the rate-limiting step. In an incrementally better approach (favored by many), the convective-diffusive transport of the protein to the surface under fully developed laminar flow conditions, coupled with surface adsorption, is described using a

lumped-parameter (so-called two compartment) model. Although better than the well-mixed approach, this method still introduces several drawbacks. The most significant of these are: (a) introduction of an additional fitting parameter k_m (transport coefficient), which can introduce errors in the extracted values and (b) nonapplicability to flow situations other than fully developed (idealized situations). In summary, neither method yields the true or intrinsic kinetics but rather a value that is influenced partly by the conditions of the experiment.

In contrast, using a Computational Fluid Dynamics (CFD) model to analyze the protein/surface interactions provides a convenient methodology for simultaneously characterizing and studying convective–diffusive transport and adsorption kinetics within a single framework. The model explicitly considers fluidic (convective–diffusive) transport of protein to the surface coupled with a model of protein adsorption on a substrate (Fig. 4). It considers a monolayer formed by a reversible adsorption process that yields a Langmuir-type adsorption isotherm. The salient features of this approach are the following:

(a) This is capable of *ab initio* accounting for convection, diffusion, and electromigration of multiple protein species toward a biomaterial surface in any arbitrary three-dimensional volume.
(b) This accounts for finite-rate binding kinetics, with the reversible adsorption process modeled as a second-order reaction and the desorption process modeled as a first-order reaction.
(c) This allows for simulation of localized variations in surface composition through variations in the density of binding sites and/or variations in the binding rate constants at different locations along the surface.

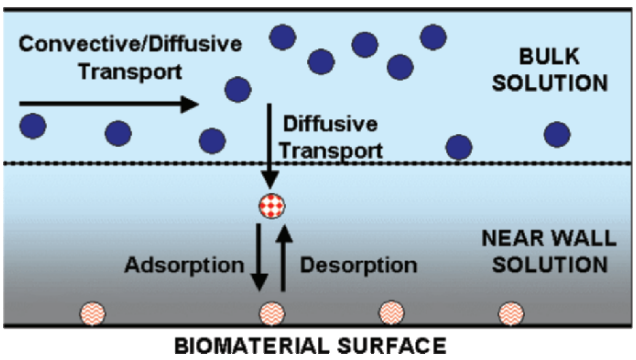


Figure 4. The convective/diffusive transport and adsorption model employed in the CFD-ACE+ simulation tool.

5. COMBINATION OF EXPERIMENT AND SIMULATION TO PREDICT PROTEIN ADSORPTION

Unfortunately, mass balance analysis alone does not provide much information about the spatial distribution of adsorbed protein in microfluidic devices. If the microchannel is uniform along its length, it is a reasonable assumption that protein adsorption is not preferentially occurring at any particular area, but if the channel has bends, constrictions, or other features, then it is entirely possible that preferential adsorption could occur. It may also, depending on the application, be important to allow for the variation in protein concentration and hence adsorption along the channel. The surface of the microchannel close to the channel inlet will be exposed, at early times of the experiment at least, to higher protein concentrations than the end of the channel. In this case, computer simulations and finite element analysis may allow the adsorption process to be modeled, based on parameters extracted from a best fit to the available adsorption data.[46]

To compare the simulation results to an experimental system, a method to map proteins in a microfluidic channel was adapted from a standard Ag assay technique in combination with XPS. To start, using a modification of a kinetic silver staining technique utilized in microtiter plate wells[48] that reduce silver nitrate to metallic silver in the presence of a protein, a procedure was developed to quantitate proteins adsorbed to self-assembled monolayers (SAMs) on the nanogram scale. The amount of protein adsorbed to the surface can be correlated to the time the silver development reaction takes by monitoring the optical density. The shorter the lag time of the reaction, the more protein is present on the surface. Figure 5 shows data for coverslips coated with the silane N-1[3-(trimethoxysilyl)propyl]diethylenetriamine (DETA) that were soaked in various concentrations (ng/mL) of bovine serum albumin (BSA) and stained with a modified silver staining solution. The solution used was silver nitrate 0.5%, ammonium nitrate 0.5%, tungstosilicic acid 0.5%, formaldehyde 0.75%, and a sodium carbonate/bicarbonate solution adjusted to a pH of 10.4.

These data show that for concentrations of 25–400 ng/mL of protein, there is a good correlation of the lag time to the amount of adsorbed protein. Combined with XPS, silver staining of proteins can give an enhanced signal and another method for quantitation of the protein. The silver XPS signal is approximately 13 times more sensitive than nitrogen and about 20 times more sensitive than carbon—the elements most commonly used for protein detection in XPS. A representative silver enhanced protein XPS spectrum can be seen in Fig. 6 for quantification of adsorbed proteins.

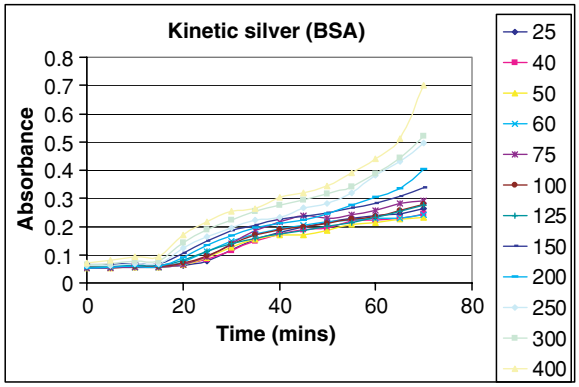


Figure 5. Optical density measurements of kinetic silver staining lag time correlating BSA protein solution concentration (ng/mL) on DETA coated coverslips.

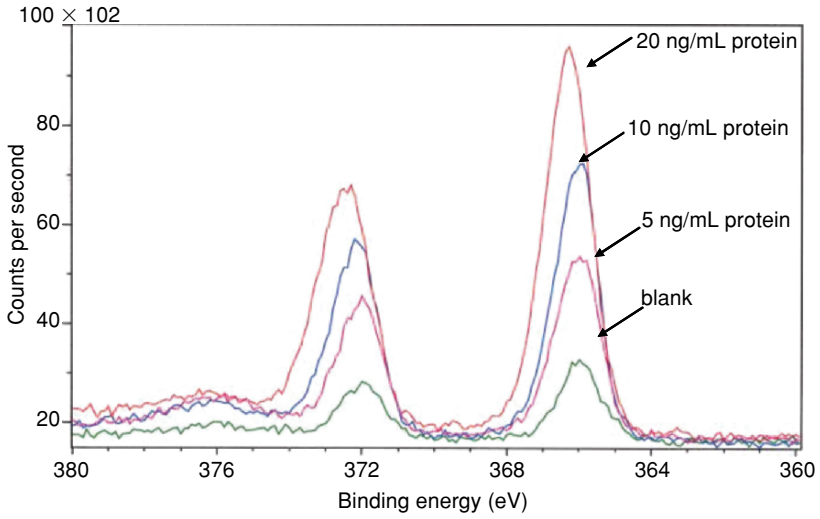


Figure 6. Quantitative XPS spectrum of silver enhanced protein at concentrations ranging from 5 to 20 ng/mL.

5.1 Experimental Verification of Simulation Results

This silver staining technique, described previously, was implemented to visualize the effect of flow through microfluidic channels coated with 1H, 1H, 2H, 2H tridecafluorotrichlorosilane (13F). This technique was able to show that the silver assay could resolve *spatial* differences in protein deposition with excellent resolution.

In this case, protein solution was passed through a 200 μm wide serpentine channel etched into a silicon wafer that was clamped to a matching, unetched silicon wafer (flat lid), and the flat lid exposed to the silver staining solution after the protein flow experiment (Fig. 7). The inlet concentration of protein was set to 300 ng/mL at a flow rate of 0.10 mL/h. The binding behavior was then simulated using our model in the channel geometries used in the experiments.

A comparison of the silver staining assay and the predicted protein binding from the CFD-ACE+ simulation tool (Fig. 8) showed good qualitative agreement. Specifically, we note the correspondence between the experiment and simulation in the location of dead zones, where little protein is being deposited, and in the overall shape of the deposited regions. Observing that the darkness of the silver stain is related to the amount of protein deposited, the modeling is seen to be in good semiquantitative agreement with experiments. These images clearly show the presence of the protein, and further comparison with the predictions obtained from finite element analysis provided by CFDRC showed many of the same features.

6. CONCLUSION

There are many techniques that can be utilized to study protein adsorption in MEMS, and ultimately in NEMS devices, and some have been demonstrated in this monograph. More importantly, it is the ***combination*** of simulation tools with experiments that will prove to be the method needed to design these systems. This combination will be especially important in the future because at the moment only simple protein adsorption is being

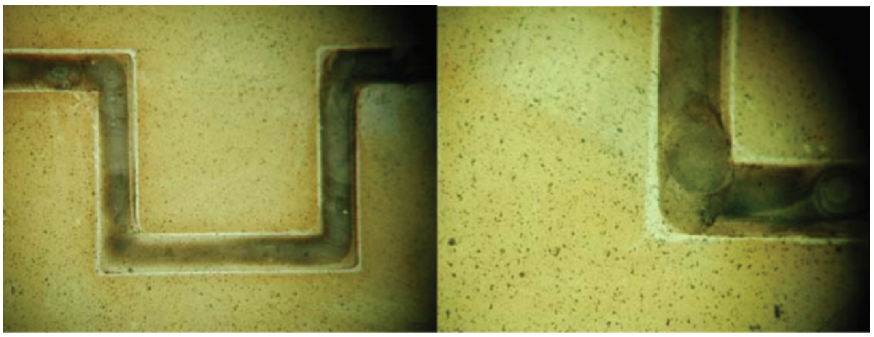


Figure 7. Silver deposited on the surface of the flat wafer used to seal the microfluidic system clearly shows that the protein has been successfully retained within the channels. These are 200 μm features, indicating that the resolution allows it to be applied to microfluidic systems.

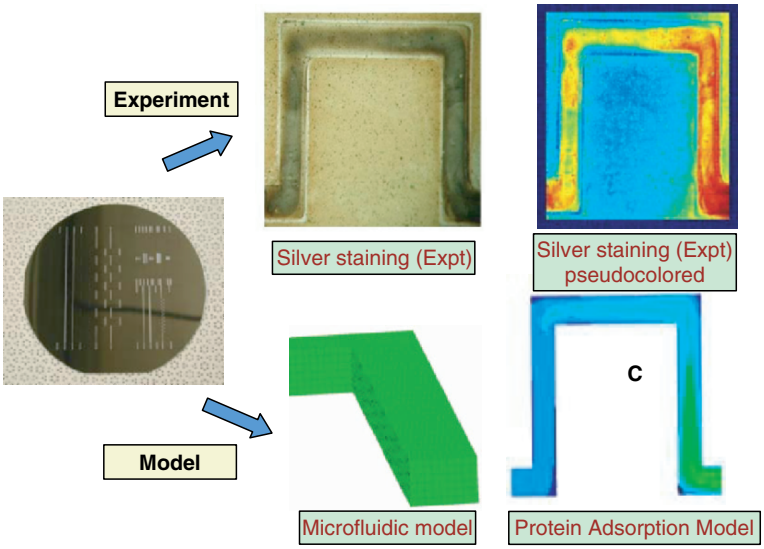


Figure 8. Comparison between the features observed using the silver staining technique on microchannels etched onto a silicon wafer with those predicted by the CFDRC simulation shows qualitative agreement.

considered, but protein adsorption can be an extremely complex dynamic that will need to be understood in order to create MEMS devices for *in vivo* applications. For example, protein adsorption often displays cooperative effects, called fouling in some circumstances, which refers to a behavior in which protein adsorption occurs much more rapidly in the presence of already adsorbed material.[49,50] If fouling is present, preferential adsorption at a particular feature can lead to rapid build-up of adsorbed material at this point due to the self-reinforcing nature of the process. In extreme cases, this may cause complete device failure. Thus, techniques that can resolve the spatial distribution of adsorbed protein provide important information as well and will need to be developed for many of the futuristic applications in biomedical and biotechnological systems currently envisioned.

ACKNOWLEDGEMENTS

We would like to acknowledge the MTO Office at DARPA, grant number F 30602-01-2-6541, and NSF, grant number ECS-0003227, for funding this project. The initial experiments for this chapter were performed at Clemson University as indicated by the dual affiliations for K. Wilson, K. Lenghaus, and J. Hickman.

REFERENCES

1. Chakraborty, A., Majumdar, A., Sundaram, S., and Dravid, V., 2004, Simulation of Biomolecular Nanomechanical Systems (BioNEMS), in: *Proceedings of the SIMBIOSYS PI Meeting*, Palm Springs, CA.
2. Rudzinski, W., 1997, *Equilibria and Dynamics of Gas Adsorption on Heterogeneous Solid Surfaces* (Rudzinski, W., Steele, W.A., and Zgrablich, G. eds.), vol. 104, Elsevier, Amsterdam.
3. Tovbin, Y. K., 1997, *Equilibria and dynamics of gas adsorption on heterogeneous solid surfaces* (S. W. Rudzinski, W. A. Zgrablich, and G. eds.), vol. 104, Elsevier, Amsterdam.
4. Lenghaus, K., Qiao, G., Solomon, D. H., Gomez, C., Rodriguez-Reinoso, F., and Sepulveda-Escribano, A., 2002, Controlling carbon microporosity: The structure of carbons obtained from different phenolic resin precursors, *Carbon* **40**(5):743–749.
5. Vallardares, F., Reinoso, R., and Zgrablich, G., 1998, Characterization of active carbons: The influence of the method in determination of the pore size distribution, *Carbon* **36**(10):1491–1499.
6. Butler, J. E., 1981, The amplified ELISA: Principles of and applications for the comparative quantitation of class and subclass antibodies and the distribution of antibodies and antigens in biochemical separates, in *Methods in Enzymology*, Academic Press, New York, pp. 482–523.
7. Butler, J. E., Ni, L., Brown, W. R., Joshi, K. S., Chang, J., Rosenberg, B. and Voss, E. W. Jr., 1993, The immunochemistry of sandwich ELISAs-VI greater than 90% of monoclonal and 75% of polyclonal anti-fluorescyl capture antibodies (CAbs) are denatured by passive adsorption, *Mol. Immun*. **30**(13):1165–1175.
8. Butler, J. E., Navarro, P., and Sun, J., 1997, Adsorption-induced antigenic changes and their significance in ELISA and immuological disorders, *Immun. Invest*. **26**(1–2):39–54.
9. Wagner, M. S., 2002, Limits of detection for time of flight secondary ion mass spectrometry (ToF-SIMS) and X-ray photoelectron spectroscopy (XPS): Detection of low amounts of adsorbed protein, *J. Biomater. Sci. Polym. Ed*. **13**(4):407–428.
10. Connaster, R. M., Riddle, L. A., and Sepaniak, M. J., 2004, Metal-polymer nanocomposites for integrated microfluidic separations and surface enhanced Raman spectroscopic detection, *J. Sep. Sci*. **27**(17–18):1545–1550.
11. Koo, T. W., Chan, S., and Berlin, A. A., 2005, Single-molecule detection of biomolecules by surface-enhanced coherent anti-stokes Raman scattering, *Opt. Lett*. **30**(9):1024–1026.
12. Ratner, B. D., Castner, D. G., Horbett, T. A., Lenk, T. J., Lewis, K. B., and Rapoza, R. J., 1990, Biomolecules and surfaces, *J. Vacuum Sci. Technol. A: Vacuum, Surf. Films* **8**(3):2306–2317.
13. Jin, F., Lenghaus, K., Hickman, J., and Marcus, R. K., 2003, Total protein determinations by particle beam/hollow cathode optical emission spectroscopy, *Anal. Chem*. **75**(18):4801–4810.
14. Jin, F., Hickman, J., Lenghaus, K., and Marcus, R. K., 2004, Total protein determinations by particle beam/hollow cathode optical emission spectroscopy (PB/HC-OES) system Part 2. Improved sensitivity through use of silica micro-particles as carrier agents for enhanced particle transport, *J. Anal. Atom. Spectrom*. **19**(9):1199–1205.
15. Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J., and Klenk, D. C., 1985, Measurement of protein using bicinchoninic acid, *Anal. Biochem*. **150**(1):76–85.

16. Hames, B. D., and Rickwood, D., 1990, *Gel Electrophoresis of Proteins: A Practical Approach*, 2nd edn., IRL Press, New York.
17. Lenghaus, K., Dale, J. W., Henderson, J. C., Henry D. C., Loghin, E. R., and Hickman, J., 2003, Enzymes as ultrasensitive probes for protein adsorption in flow systems, *Langmuir* **19**(15):5971–5974.
18. Gottlieb, M., and Chavko, M., 1987, Silver staining of native and denatured eucaryotic DNA in agarose gels, *Anal. Biochem*., **165**(1):33–37.
19. Pacholski, M., and Winograd, N., 1999, Imagine with mass spectrometry, *Chem. Rev*. **99**(10):2977–3006.
20. Wagner, M. S., 2005, Molecular depth profiling of multilayer polymer films using time-of-flight secondary ion mass spectrometry, *Anal. Chem*. **77**(3):911–922.
21. Wagner, M. S., Horbett, T. A., and Castner, D. G., 2003, Characterizing multicomponent adsorbed protein films using electron spectroscopy for chemical analysis, time-of-flight secondary ion mass spectrometry, and radiolabeling: capabilities and limitations, *Biomaterials* **24**(11):1897–1908.
22. Wagner, M. S., Tyler, B. J., and Castner, D. G., 2002, Interpretation of static time-of-flight secondary ion mass spectra of adsorbed protein films by multivariate pattern recognition, *Anal. Chem*. **74**(8):1824–1835.
23. Wagner, M. S., Shen, M., Horbett, T. A., and Castner, D. G., 2003, Quantitative analysis of binary adsorbed protein films by time-of-flight secondary mass spectrometry, *J. Biomed. Mater. Res. A* **64**(1):1–11.
24. Griesser, H. J., Kingshott, P., McArthur, S. L., McLean, K. M., Kinsel, G. R., and Timmons, R. B., 2004, Surface-MALDI mass spectrometry in biomaterials research, *Biomaterials* **25**(20):4861–4875.
25. Kingshott, P., St. John, H. A. W., and Griesser, H. J., 1999, Direct detection of proteins adsorbed on synthetic materials by matrix-assisted laser desorption ionization-mass spectrometry, *Anal. Biochem*. **273**(2):156–162.
26. Takats, Z., Wiseman, J.M., Gologan, B., and Cooks, R.G., C., 2004, Mass spectrometry sampling under ambient conditions with desorption electrospray ionization, *Science* **306**(5695):471–473.
27. Sane, S. U., Cramer, S. M., and Przybycien, T. M., 1999, A holistic approach to protein secondary structure characterization using Amide I Band Raman Spectroscopy, *Anal. Biochem*. **269**(2):255–272.
28. Delfino, I., Bizzarri, A. R., and Cannistraro, S., 2005, Single-molecule detection of yeast cytochrome c by surface-enhanced Raman spectroscopy, *Biophys. Chem*. **113**(1):41–51.
29. Podstawka, E., Ozaki, Y., and Proniewicz, L. M., 2004a, Adsorption of S-S containing proteins on a colloidal silver surface studied by surface-enhanced Raman spectroscopy, *Appl. Spectrosc*. **58**(10):1147–1156.
30. Podstawka, E., Ozaki, Y., and Proniewicz, L. M., 2004b, Part I: Surface-enhanced Raman spectroscopy investigation of amino acids and their homodipeptides adsorbed on colloidal silver, *Appl. Spectrosc*. **58**(5):570–580.
31. Podstawka, E., Ozaki, Y., and Proniewicz, L. M., 2004c, Part II: surface-enhanced Raman spectroscopy investigation of methionine containing heterodipeptides adsorbed on colloidal silver, *Appl. Spectrosc*. **58**(5), 581–590.
32. Cao, Y. C., Jin, R., Nam, J. M., Thaxton, C. S., and C. A., M., 2003, Raman-dye labeled nanoparticle probes for proteins, *J. Am. Chem. Soc*. **125**(48):14676–14677.
33. Coen, M. C., Lehmann, R., Groning, P., Bielmann, M., Galli, C., and Schlapbach, L., 2001, Adsorption and bioactivity of protein A on silicon surfaces studied by AFM and XPS, *J. Colloid Interf. Sci*. **233**(2):180–189.

34. Schnyder, B., Kotz, R., Alliata, D., and Facci, P., 2002, Comparison of the self-chemisorption of azurin on gold and on functionalized oxide surfaces, *Surf. Interface Anal*. **34:**40–44.
35. Dufrene, Y. F., 1999, Probing the organization of adsorbed protein layers: Complementarity of atomic force microscopy, X-ray photoelectron spectroscopy and radiolabeling, *Appl. Surf. Sci*. **144–145**:638–643.
36. Lasseter, T., Clare, B., Abbott, N., and Hamers, R., 2004, Covalently modified silicon and diamond surfaces: Resistance to nonspecific protein adsorption and optimization for biosensing, *J. Am. Chem. Soc*. **126:**10220–10221.
37. Sousa, S. R., Moradas-Ferreria, P., Saramago, B., Viseu Melo, L., and Barbosa, M., 2004, Human Serum albumin adsorption on TiO_2 from single protein solutions and from plasma, *Langmuir* **20:**9745–9754.
38. Lhoest, J.-B., 1998, Fibronectin adsorption, confirmation, and orientation on polystyrene substrates studied by radiolabeling, XPS, and ToF SIMS, *J. Biomed. Mater. Res*. **41:**95–103.
39. Tidwell, C. D., 2001, Static time of flight secondary ion mass spectrometry and X-ray photoelectron spectroscopy characterization of adsorbed albumin and fibronectin films, *Surf. Interface Anal*. **31:**724–733.
40. Glidle, A., 2003, Analysis of protein adsorption and binding at biosensor polymer interfaces using X-ray photoelectron spectroscopy and scanning electrochemical microscopy, *Anal. Chem*. **75**:2559–2570.
41. Chi, Q., Zhang, J., Nielson, J., Friis, E., Chorkendorff, I., Canters, G., Andersn, J., and Ulstrup, J., 2000, Molecular monolayers and interfacial electron transfer of *Pseudomonas aeruginosa* Azurin on Au(111), *J. Am. Chem. Soc*. **122**:4047–4055.
42. Petoral Jr., R., and Uvdal, K., 2002, Arg-Cys and Arg-cysteamine adsorbed on gold and the G-protein-adsorbate interaction, *Colloids Surf. B*. **25**:335–346.
43. Dodero, G., De Michieli, L., Cavalleri, O., Rolandi, R., Oliveri, L., Dacca, A., and Parodi, R., 2000, L-Cysteine chemisorption on gold: An XPS and STM study, *Colloids Surfaces A Physicochemical and Engineering Aspects* **175:**121–128.
44. Browne, M. M., 2004, Protein adsorption onto polystyrene surfaces studied by XPS and AFM, *Surf. Sci*. **553:**155–167.
45. Denis, F., Harnap, P., Sutherland, D., Gold, J., Mustin, C., Rouxhet, P., Dufrene and Yves, F., 2002, Protein adsorption on model surfaces with controlled nanotopography and chemistry, *Langmuir* **18:**819–828.
46. Jenkins, J., Prabhakarpandian, B., Lenghaus, K., Hickman, J., and Sundaram, S., 2004, Fluidics-resolved estimation of protein adsorption kinetics in a biomicrofluidic system, *Anal. Biochem*. **331**(2):207–215.
47. Gregg, S. J., and Sing, K. S. W., 1967, *Adsorption, Surface Area and Porosity*, Academic Press, London.
48. Root, D. D., and Wang, K., 1993, Kinetic silver staining and quantification of proteins adsorbed to microtiter plates, *Anal. Biochem*. **209:**354–359.
49. Robert H. Perry, Donald W. Green editors., 1997, *Perry's Chemical Engineer's Handbook*. The McGraw-Hill Companies.
50. Hook, F., Rodahl, M., Kasemo, B. and Brzenzinski, P., 1998, Structural changes in hemoglobin during adsorption to solid surfaces: Effects of pH, ionic strength, and ligand binding, *Pro. Natl. Acad. Sci*. **95**(21):12271–12276.
51. Briggs, M.P., and Seah, M.P., 1992, Practical Surface Analysis by Auger and X-ray Photoelectron Spectroscopy, Second Edition, John Wiley and Sons, New York.

Chapter 14

PROGRAMS OF THE EUROPEAN COMMISSION ON ORGANIC MATERIALS FOR THE ELECTRONICS INDUSTRY

Dr Anne F. de Baas
European Commission, DG RTD G3, Office CDMA 0/170, Brussels, Belgium

1. INTRODUCTION

The European Commission (EC) develops and follows up on political initiatives to realize a European Research Area. To do so, the Commission issues Framework Programs for research and technological development. There are two Directorates General (DG) each dealing with a subset of the research areas: DG Research and DG Information Society Technologies (INFSO).

The DG Research deals a.o. with the thematic areas: nanotechnology, materials, and production (NMP), which aims to promote industrial breakthroughs, based on scientific and technical excellence. The topic of organic materials for the electronic industry falls within this NMP program. The DG INFSO deals with information society technology (IST).

An overview of projects currently funded by the EC in the field of organic materials for the electronics industry is given in Section 2. The content of future calls for proposals (2005) in the area will also be discussed. Further details on these programs, calls, and projects can be found on the EC's Website www.cordis.lu.

2. ONGOING PROJECTS IN THE NMP PROGRAM

Three projects will be highlighted, of which the first, a network of excellence, is the most relevant for organics electronics. Networks of excellence have the goal to integrate the participants into an organization, which is more

V. Renugopalakrishnan et al. (eds.), Bionanotechnology, 275-282.

efficient than its parts. A lasting new organization is to be developed. This organization will plan the research and give access to personnel, equipment, and facilities. Execution of the research is partly on the EC budget, but it is envisaged that the new organization will become self-sufficient. These networks should also integrate education and skills development with research activities to increase European knowledge. In addition, it is expected that breakthrough research activities should help to foster dialog with society and generate enthusiasm for science. In this way, a better use of still-fragmented distribution of resources will be made possible.

The other two projects mentioned are of another type called specific targeted research projects (STREPs). This type of project does long-term, advanced research and addresses specific parts of a technology or value chain.

2.1 Nanofun-Poly

The network of excellence Nanofun-Poly started in June 2004. The main objective of Nanofun-Poly is to install the European Center for multifunctional nanostructured polymers and polymer matrix nanocomposites (ECNP). This new organization should strengthen the scientific, technological, and training excellence in all the disciplines that contribute to the development of polymer-based nanostructured polymers and nanocomposites. This objective will be reached through a transdisciplinary partnership of 120 scientists combining excellence in different scientific areas. Applications that will benefit from Nanofun-Poly can be found in optoelectronics, telecommunications, packaging, agriculture, building construction, automotive, and aerospace. Besides integrating the excellence in scientific expertise, Nanofun-Poly will integrate the considerable resources in advanced instrumentation available to its partnership. On the research side, the project will do road mapping, strategy development, and plan research in the area. The vision is that the new organization will produce better research projects with even better results.

The network will allow continuous contact with large companies and SMEs through the constitution of an Industrial Council with 28 members and training activities.

The coordinator if this Network of Excellence is Italian Consortium for Science and Technology of Materials (INSTM) and information on the other 29 partners can be found on the EC Website.

2.2 NAIMO

The STREP project NAIMO develops materials, processes manufacturing capabilities to transform a plastic film substrate into a multifunctional composite with designed electronic, optical, sensing, and magnetic capabilities by

a series of solution-based additive manufacturing steps. This solution-based manufacturing approach will enable control of structure on a nanometre scale. The NAIMO builds on the important advances made by the partners in the use of molecular materials including conjugated polymers for low-cost large-area electronic and optical applications.

2.3 NaPa

The STREP project nanopatterning methods (NaPa) integrate new patterning methods in order to arrive at low-cost scalable processes and tools. They cover the needs of nanopatterning from CMOS back-end processes through photonics to biotechnology. To achieve this, research in three technology strands is carried out: nl, nanoimprint lithography; soft lithography; and self-assembly and MEMS-based nanopatterning.

3. COMING CALLS IN THE NMP PROGRAM

The NMP Program targets new, high knowledge-content materials, providing new functionalities and improved performance, which will be critical drivers of innovation in technologies, devices, and systems. The reader is invited to read the complete Work Program 2005 on Web: www.cordis.lu. The two topics most relevant for "Proteins to Devices," are mentioned in this chapter. Both topics invite for the STREP-type of project. Other topics, like material characterization and computational modeling could also involve organic electronics. Further information can be found on the EC Website.

3.1 Materials by Design: Multifunctional Organic Materials—STREP

The flexibility in the design and processing of organic solids (usually polymers) allows them to meet the requirements of many technologically significant applications. Multifunctional organic materials are used in displays, electronic circuits, solar cells, chemical sensors and actuators, lasers, storage media, and electronic paper, as well as for insulation and packaging in electronics.

The expected STREPs should aim at highly innovative long-term research for the development of new multifunctional organic materials for electronics, including modeling and experimentation, and considering their processing and potential applications as well. The objective is to increase the capability

of industry to have materials with characteristics needed for the intended applications. Molecular and nanoelectronics are excluded.

3.2 Materials for Solid-State Ionics—STREP

Understanding ionic transport in solids, and in particular in nanostructured materials, is the key to many technological applications in wide temperature ranges, for instance, in solid-state batteries, conducting membranes, electrochemical gas sensors, and electroceramic devices.

The STREPs should focus on highly innovative multidisciplinary research to develop electrodes and electrolytes providing improved electron transport and reactivity, in particular for microbatteries in mobile microsystems. Developments in the field of fuel cells are excluded.

4. ONGOING PROJECTS IN THE IST PROGRAM

Current projects on organics electronics in the IST program are all concerned with the development of light-emitting diode OLEDs. Here seven projects are shortly mentioned; further information can be found on the EC Website.

4.1 FLEXled

The FLEXled project aims at demonstrating reliable flexible displays for nomadic consumer electronics applications, based on a flexible emissive technology: polymer light-emitting diodes (PLED). Passive matrix displays (up to 4 in. 320×240 pixel RGB color) on flexible substrates will be made.

4.2 Photoled

The project develops novel photopatternable organic light-emitting diode (p-OLED), RGB emitter materials, and associated fabrication processes as the basis for making full color p-OLED-on-silicon displays. This unique p-OLED technology is based on a photopatternable liquid crystalline polymer blue emitter doped with red and/or green fluorescent dyes to generate the RGB emission. The developed p-OLED materials will also provide the basis for the development of a new display with a unique stereoscopic 3D effect. The target for 2D p-OLED materials is 15 lm/W and the target for 3D p-OLED materials is 7 lm/W (both measured for green emission).

4.3 Contact

Contact addresses the production of organic electronic arrays: transistors, solar cells, sensors, and light-emitting diodes. These develop a process for the fabrication through gravure printing of the various components of liquid crystal displays (LCD). The project formulates suitable materials and substrates and modifies test printers. The process should provide a significant enhancement of thin-film electronics in terms of cost and throughput. Two test displays will be realized: namely an optical triple-cell LCD and an all-plastic LCD.

4.4 FlexiDis

This project ''Flexible Displays,'' or FlexiDis, has the realization of flexible active-matrix displays as its primary goal. This will include the materials and processing technology, physical studies of the mechanical properties, and materials behavior of multilayer structures during processing, under flexing, and during service life, and novel fabrication tools and substrate handling procedures to make and introduce reliable, flexible, active-matrix displays to the market. The demonstrator vehicles will work with two different display technologies: full-color organic light-emitting displays (OLED) on bendable metal and/or plastic foils with inorganic thin-film driving transistors and monochrome electrophoretic (EP) displays on rollable plastic substrates with organic thin-film driving transistors.

4.5 Improve

The aim of this project is to improve lightweight near-to-the-eye displays and tiled stereoscopic large size displays. The improvements on the hardware level consist in developing a unique stereoscopic head mounted display (HMD) using emerging display technology such as LCOS and OLEDs. For tiled stereoscopic large screen displays, improved calibration techniques will be developed to ease and accelerate their use. On the software level, improvements comprise the fidelity of the content to be displayed (rendering quality), the interfacing between the user and the displays through innovative 2D/3D interaction techniques for mixed realities, and advanced tracking systems. The achievements of IMPROVE are integrated into a collaborative mixed-reality product development environment, showcased and evaluated in two application scenarios: collaborative product design in the car industry and architectural design.

4.6 Rolled

The projects' goal is the development of a cost-effective, volume-scale, roll-to-roll manufacturing technology for the realization of flexible OLED devices with arbitrary size and shape pixels and displays on the Web. To reach the cost level of OLEDs required for packaging and printing, the production of the devices must be transferred to the low-cost roll-to-roll processes integrated in the conventional printing and packaging processes used in industry. This requires selection and modification of suitable materials, effective patterning technologies, development of the manufacturing processes, and improved encapsulation technologies.

4.7 Shift

The objective of the project is the development of smart, highly integrated, mechanically flexible electronic systems for a wide variety of applications. ''Smart'' means that the flexible multilayer laminate has embedded components, and that the different flex layers in the multilayer structure can have different functions, meaning that it might be necessary to combine layers of different base material in the laminate. Compactness of the resulting circuit will be boosted in two ways: by using the third dimension for electronic component integration (not only on front and back side, but potentially on every conductive layer) and by drastically increasing the wiring density through the introduction of new flex manufacturing and lamination techniques.

5. COMING CALLS IN THE IST PROGRAM

The IST Program wants to stimulate research and development of information communication technologies, so as to master the technologies that will drive future innovation and growth. From the Work Program 2005, four topics relevant for organic electronics are selected and discussed in later sections. For the full Work Program text the reader is referred to the EC Website.

5.1 Nanoelectronics

The technical goals are to reduce the transistor size, deep into the nanoscale, to radically transform the process technologies through the integration of a large number of new materials and to master the design technologies for achieving competitive systems-on-chip and systems-in-package with increasing functionality, performance, and complexity.

5.2 Technologies and Devices for Micro/Nanoscale Integration

To push the limits of integrated micro- or nanosystems through research on a family of mixed technologies (combining for instance micronanotechnology, ICT, and biotechnology) and integration technologies for very high density or for integrating micro- or nanodevices in various materials and into large surfaces. Validation and demonstration of maturing silicon-based and polymer-based technologies, manufacturing, and design issues are also targeted.

5.3 Photonic Components

To develop advanced materials, solid-state sources and micro- and nanoscale photonic devices, and to integrate photonic functions in micro-/nanoelectronics components (''photonic system on a chip''), biophotonic functional components and subassemblies are explicitly mentioned, and the call invites for Integrated Projects in this area.

5.4 Micro/Nano-Based Subsystems

To validate integrated micro-/nanosystems technology for new products and services in key application fields such as miniaturized autonomous robotic systems, mass storage systems, and visualization systems. Micro-/nano-based integrated medical systems are also targeted to explore many opportunities offered by combining bio-, nano-, and information-related technologies. Part of this topic addresses very large area displays and highly integrated display solutions. These tasks are to be addressed through IPs and STREPs.

6. CONCLUSION

The EC Programs offer many possibilities for advanced research in the area of organic electronics. The reader is advised to consult the EC Website on details of ongoing projects and future calls for proposals.

REFERENCES

The EC Website on www.cordis.lu, where more information on the research Programs and Projects can be found.

Chapter 15

BIODIVERSITY: AN ARCHIVE OF OPPORTUNITY FOR NANODEVICES

Paul Galvin[1], Dennis A. Dempsey[2], Alan O'Riordan[1], Gareth Redmond[1], Eric Moore[1], W. Shannon Dillmore[1], and Marin Gheorghe[1]
[1]*Tyndall National Institute, Lee Maltings, University College Cork, Ireland*
[2]*Analog Devices B.V., Raheen Business Park, Limerick, Ireland*

Abstract: Since the invention of the transistor in 1947 and the subsequent invention of the integrated circuit, the feature density of microelectronic integrated circuits and devices has increased exponentially (Moore's Law). In complementary metal-oxide semiconductor technology, the minimum feature size in volume production has decreased from 2 μm in 1980 to 0.1 μm in 2002. However, new fabrication strategies are required as conventional top-down fabrication techniques continue to encounter many obstacles and challenges associated with the approach of their fundamental size limits. Interest has therefore focused on molecular scale materials, including biomolecules, to provide the building blocks for next generation functional nanoscale devices. Since nature is an obvious archive of nanofabrication and self-assembly solutions, the opportunity exists to exploit bioinspired and biohybrid solutions to fabrication of nanodevices. This approach has the potential to enable the formation of complex component architectures and offers a number of advantages for fabrication of future nanoscale devices, such as, spatial control at the nanometre length-scale, parallel self-assembly mechanisms and the tendency towards self-correction and defect minimisation. Examples based on templated assembly of nanoparticles using DNA, bacterial S-Layers and other peptides, provide convincing evidence of the feasibility of nanoscale fabrication based on biomimetics. Therefore, as the industry considers its future beyond the perceived limits of Moore's Law, new bioinspired processes and biohybrid devices are envisaged, based on unconventional ''bottom-up'' self-assembly routes, which, while compatible with silicon processing techniques, provide alternatives to replace or augment high resolution lithography based methods.

Key words: bionanotechnology, self-assembly, Moore's Law, nanodevices, nanofabrication, biodiversity

V. Renugopalakrishnan et al. (eds.), Bionanotechnology, 283-296.

1. INTRODUCTION

The transition from microfabrication to nanofabrication has presented many new challenges to industry, as Moore's law reaches the physical limits of ''top–down'' fabrication tools. In order to continue to reduce feature sizes, nanofabrication needs to incorporate radically different approaches based on ''bottom-up'' assembly approaches. While there are numerous examples of ''self-assembly'' in chemistry, it is clear that many challenges remain to be resolved to enable self-assembly methods to be incorporated in routine nanodevice fabrication. There is now a growing realization that many of these challenges have been resolved in nature, and it is possible to use nature as a model to find solutions for nanofabrication. In this context, biodiversity provides a library of biomolecules that should be considered as an archive of nanofabrication techniques.

The diversity of habitats within the marine environment alone ranges from hydrothermal deep-sea vents with organisms operating at extremes of temperature and pressure, to intertidal areas with dramatic changes of salinity and sometimes prolonged shortages of oxygen, and finally to Antarctic regions where some organisms complete their life cycle in extremes of cold. By studying and understanding organisms living in these environments (known as extremophiles) and comprehending the complex solutions they have been applied to enable them to survive and thrive in their challenging environments, innovative solutions to existing nanofabrication challenges can be anticipated. In addition, novel solutions and novel nanodevices can be expected through isolation and exploitation of functional proteins from these organisms. This chapter discusses the potential for innovative bioinspired solutions for fabrication of three-dimensional (3D) structures.

2. CHALLENGES FOR BIOINSPIRED SEMICONDUCTOR DEVICES

Silicon, a group IV semiconductor, is the dominant integrated circuit (IC) technology and has led the growth of electronics into a world leading industry today. Many years have passed since Gordon Moore penned his 1965 prediction[1] of continued exponential IC growth, with the number of IC transistors doubling every year. Moore views the 1959 invention of the planar transistor and planar wafer fabrication by Jean Horni as the origin of this law. Many generations of engineering development have abided by this law, shrinking the process geometries and continuing integration growth exponentially. How-

ever, as IC device dimensions approach the nanometer (atomic scale), it is clear that the fundamental limits of physics will, ultimately, limit Moore's law.

Foreseeing these ''limits'', the semiconductor industry is working on leading edge technologies[2] to go beyond these limits and embrace the new challenges, as it has done for many years. An alternative perspective is to say that Moore's law is about human ingenuity and inventiveness, a core engineering skill set across many engineering disciplines. The industry has been challenged to abide by ''Moore's Law'' and has succeeded in overcoming the challenges of many generations. As the industry now struggles with new integration challenges/boundaries, a paradigm shift seems to be required in the medium term. If the rewards are worthwhile, history tells us that a solution, a path of progress will be found to meet the goal of continued growth. The semiconductor industry has remained substartially unchallenged by competitive technology in pursuit of systems to complete tasks in a desired, controlled manner. Will the solid-state-circuits industry live long after its 50-year window, where many other industries have not? In this regard, the potential to learn new nanofabrication approaches from nature is clear and remains as a potential goldmine awaiting exploitation.

Jean Horni's invention of the now institutionalized silicon planar wafer fabrication process was a breakthrough, which has been used to great effect for over 30 years to integrate circuits in both lateral dimensions, with ongoing wafer fabrication development enabling higher levels of integration, and the industry is content with success, foregoing the third dimension. More recent developments have expanded the capability of integration by increased number of metalization layers to ease the communication bottlenecks. Stacking of die and other lateral passive devices are also used to successfully increase the level of integration, using this third dimension somewhat. Progress on the third dimension is much slower than on the planar dimensions, and the relative sloth is attributable to the fact that all such development efforts are constrained by a requirement for compatibility with further development of the core planar wafer fabrication process. MOS gate leakage is a well understood technology limit 1,2 with leakage increasing approx 2.5 times per nm gate dielectric thickness reduction. In terms of network design and interconnect, integration of circuits lingers behind the biological world where all three dimensions are used without prejudice, allowing a much superior communications network. Increased usage of the third dimension appears to be an area where ICs have a lot of potential for further development.

2.1 Geometry Shrinking Limits

At the end of 2004, leading edge research was at the 5 nm scale. In this sub-20 nm scale, atomic level statistical fluctuations are not completely

understood as quantum effects come into play and the conventional electron and hole conduction model loses validity. Both quasi-ballistic operation and enhanced velocity saturation are required at the 20 nm technology node. A lot remains to be understood at a device level, and this understanding will be required before circuitry is built with these devices. To this end, research is actively underway with field-effect transistors (FinFETs) at this geometry. This compares with the diameter of a deoxyribonucleic acid (DNA) structure at 0.5–2.0 nm and spacing between silicon atoms in the crystal lattice of tenths of a nanometer. Lithography is a critical part of the wafer fabrication process, defining the features in the planar dimensions. Research is ongoing, and the application of deep ultraviolet (DUV) for defining features <10 nm is demonstrated and remains a prospect in this range. In this respect, future devices are likely to combine both bottom-up (e.g., self-assembly) and top–down (e.g., lithographic techniques).

2.2 Self-Assembly

What lesson, if any, can we learn from nature that builds organisms, with devices of much higher complexity and tissues that can self-heal, even after significant injuries? The IC world uses software and hardware encoded (circuitry) procedures to control the function(s) to yield the desired response. The biological world takes a different and very successful approach using "self-assembly." The core cells are the building blocks of the system, which work together in their respective functions to deliver a final system. Biological systems have the well-proven capability to replace cells when required and cater for discrepancies in the system. Integrated circuits use a self-control version of this concept in a number of ways:

- Built in self-test (BIST)
- Self-calibration, in a variety of forms
- Control loop behavior–self-correcting nonidealities
- Usage of built-in redundancy in hardware and signal protocols

The latter are all very specific, controlled, and very contained solutions to very specific requirements that do not match the complexity of capabilities in the natural world. We are, as yet, many steps away from truly understanding the complexity of a biological system. The even greater challenge will be integration of bioinspired nanofabrication techniques and development of biohybrid technologies compatible with IC design, a very different technology.

"Self-assembly" remains a goal on the horizon of the integrated circuits industry. Self-healing, as nature does, is not even under consideration, being far beyond where the industry currently stands. An analogous practice in electronics is the use of redundancy such that faults/errors, e.g., in a large

DRAM, can be detected, and the relevant unit(s) substituted with new unit(s) in a manner invisible to the user. This is, notably, a practice not prevalent in nature.

2.3 Backward Compatibility

The Semiconductor industry[2] focuses on a number of integration paths, all of which are compatible with current IC manufacturing processes. New technological developments need to be compatible with and/or capable of interfacing Metal Oxide Semiconductor (MOS) technology in order to avoid direct competition with the significant power of the semiconductor industry. This may be a great challenge for bioengineered solutions designed to be integrated with electronic systems.

3. INTEGRATION OF PROTEINS IN NANODEVICES

Surface functionalization facilitates the interface between nanomaterials and biomolecules. When considering such an interface three criteria need to be considered:

- Nonspecific binding of biomolecules to nanomaterials
- Inhibition of nonspecific biomolecule interaction to nanomaterials
- Control of specific selective binding of biomolecules to nanomaterials

Proteins consist of long chains of amino acids known as polypeptides which differ in chain length, amino acid sequence, and consequently their folded shapes. Although proteins are produced as long chains of amino acids, the strands fold to give secondary structures that comprise four primary elements; α-helices, β-pleated sheets, turns, and loops. A third structure level is also possible and is referred to as the tertiary structure. This occurs when secondary structures wrap or twist in formation, and they are held together by a combination of forces, such as hydrophobicity, weak van der Waals forces, electrostatic interactions, hydrogen bonding, and also by through bonds such as disulphide bridges. Quaternary structures are associated with chemically separate protein chains that bind with one another, and they are held in place by the same forces as described for the tertiary structure. The process by which these final two structures of proteins are broken down is known as denaturing.

There are a number of ways by which biomolecules can be immobilized to nanomaterials.

Passive Adsorption is perhaps the simplest way in which biomolecules are allowed to contact with an activated solid support for a given time and under specific conditions, during which the biological agents adsorb onto the surface of the nanomaterial. Physisorption of the biomolecule occurs from electrostatic interaction, hydrogen bonding, hydrophobic interactions, and/or a combination of these.

Cross-linking is perhaps one of the best-known methods of immobilizing enzymes. Commonly used cross-linking bifunctional agents for this application are glutaraldehyde and hexamethylene. The linker attaches itself to the amino groups of the enzyme and also to the solid support. A stable polymer is formed when these bifunctional reagents bound several enzymes together and produce a thin film on the solid support. The major problem associated with this method is the decrease in activity of the enzyme, while the advantage is increased enzyme loading due to the spatial arrangement of product. Swelling of the bulky layer formed can also be a problem, as large analytes may not diffuse through.

Covalent bonding involves direct covalent attachment of the biomolecule to the surface of the nanomaterial. While similar to the above-mentioned cross-linking method, it differs in that it creates a monolayer of protein on the surface as opposed to a random orientation of protein. The most commonly used chemical for covalently attaching antibodies to the surface of transducers is 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC). The advantage that this system offers is that the layer formed is very thin and, therefore, the diffusivity is excellent.

Physical Entrapment is a method in which the biomolecules are trapped in a gel or a polymer. They are still close enough to the surface to enable signal detection, and there is no chemical change to the immobilized species. The principle involves trapping the biomolecules in a lattice that allows smaller molecules, such as the substrate, to diffuse through and interact with them. However, the larger the antigen the less likely that this immobilization method will work as the antigen may have difficulty diffusing through the lattice. Examples of commonly used gels or polymers for physical entrapment are nylon, starch, and polyacrylamide.

4. BIOHYBRID AND BIOINSPIRED NANODEVICES

Many proteins have distinct conformations that make them excellent candidates for a bottom-up strategy of synthesizing nanoscale materials with defined structures. Many successful demonstrations of this concept

exist in the literature. A notable example was the construction of a mesh with nanopores using a tetrameric aldolase protein linked to streptavidin hubs. Biotin, a ligand that binds streptavidin nearly irreversibly, was conjugated to the aldolase using a helical spacer. Ringler and Schulz[3] were able to control the size of the pores by varying the length of this spacer. Sone and Stupp[4] were able to template the growth of CdS fibers using peptide-amphiphile molecules. A classic example by Sleytr and coworkers[5] demonstrates the synthesis of superlattice structures using S-layer bacterial wall proteins.

4.1 Bacterial Surface Layers (S-Layers) as Biomolecular Templates

Crystalline bacterial cell surface layers (S-layers) possess pores identical in size, morphology, and functional groups in well-defined lattice positions and can recrystallize and self-assemble into SAMs in suspension or on solid surfaces. S-layers are generally 5–10 nm thick with unit cell dimensions in the range 3–30 nm with pore dimensions 2–8 nm diameter, depending on the strain,[5] see Fig. 1. The periodicity of the pores has enabled S-layers to be used as molecular lithography nanomasks and biomolecular assembly templates. Metalized S-layer composites are successfully employed as shadow masks for selective nanoscale etching of Si, GaAs, glass, metal, and graphite substrates,[6–10] for example, fabrication of an array of holes in silicon as a

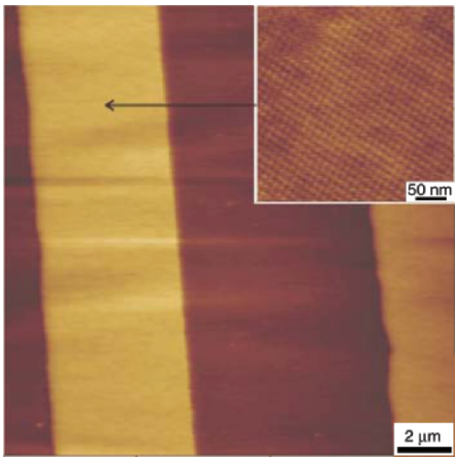


Figure 1. Contact mode AFM image of 6-μm wide S-layer protein tracks patterned and recrystallized on a silicon wafer (Z-range 50 nm). Inset shows the square lattice symmetry of the SbpA S-layer protein (Z-range 1.2 nm). Images acquired under 100 mM NaCl.[5] © American Chemical Society.

template for growth of 10 nm hexagonal magnetic dots has been demonstrated[11]. Synthesis of 5 nm CdS nanoparticle arrays nucleated in S-layer pores are reported[12]. S-layers are used as biomolecular templates for fabrication of metal nanocluster arrays, site-specific organization of gold nanoparticles, and molecular arrays of streptavidin (as a potential route to fabrication of quantum dot arrays).[13–17] Thus, the wide range of possible S-layer conformation allows for many possible superlattice structures.

Templated self-assembly based on the molecular recognition properties of biological molecules provides a facile route for fabrication of nanoscale components and component architectures. In such a bottom-to-top approach, the assembly information is programmed or encoded into the components, which then self-assemble according to that information. In addition, many well-developed techniques now exist for synthesis, manipulation, and structural tailoring of a broad range of biological molecules. To this end, research has focused on application of biological scaffolds for templated self-assembly of functional nanostructures. Some of the recent highlights in this area are as follows.

4.1.1 Nanofabrication templated by individual strands of DNA

This is used to template the assembly of 1D gold nanoparticle arrays with close packed structures. Here, coulombic interactions between cationic ligand-stabilized gold nanoparticles and anionic DNA were employed to drive the site selective localization and self-assembly of the nanoparticles at the DNA surface.[18] Parallel 1D and crossed 2D platinum nanowire arrays were also fabricated by first using molecular combed DNA to direct the assembly of positive metal ions at the DNA surface, followed by chemical reduction of the cations to yield 30 nm diameter nanowires.[19] Electrostatic binding of ligand stabilized nanoparticles to a DNA backbone to form extended linear chain-like structures, ribbon-like structures composed of parallel nanoparticle chains, and branched structures was demonstrated.[20] DNA templated fabrication of well-defined networks of oligonucleotide duplexes was also reported. Chemically modified DNA component tiles, comprising 4×4 DNA strands, were successfully self-assembled into two lattice structures, nanoribbons and nanogrids, which were then employed to template the assembly of protein arrays, and highly conductive silver nanowires were demonstrated (see Fig. 2).[21]

Braun and coworkers have recently demonstrated sequence specific molecular lithography on individual DNA molecules.[22] In this bottom-up approach, RecA proteins were polymerized onto single stranded probe DNA molecules forming nucleoprotein filaments. These filaments were then mixed with double stranded DNA substrate molecules leading to nucleoprotein–substrate binding at sequence specific homologous probe-substrate locations. Such homologously combined nucleoprotein filaments were employed as site

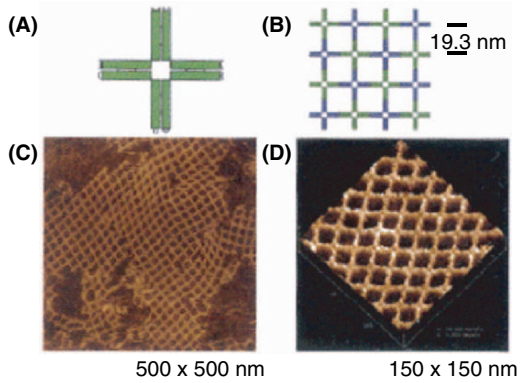


Figure 2. Self-assembly of 2D nanogrids with corrugated design. (A) Schematic of the 4 × 4 DNA strand component tile. (B) Schematic of the self-assembled 2D nanogrids. (C) AFM image of self-assembled 2D lattices (nanogrids) displaying a square aspect ratio. (D) Surface plot of a magnified region from (C).[21] © American Association for Advancement of Science.

selective resists by, e.g., preventing metalization of the DNA substrate molecules in the protected areas. The position and size of the insulating gap on the DNA nanowires could be tailored by altering the probes sequence.

In an extension of this method, the Braun group have fabricated carbon nanotube field-effect transistors using streptavidin-functionalized semiconducting single wall carbon nanotubes selectively localized at biotin modified nucleoprotein filaments bound at DNA substrate molecules.[23] Following localization of the nanotube, the remaining ends of the DNA molecules were metalized to form extended conductive wires that electrically contacted the carbon nanotubes (see Fig. 3). Electrical characterization of a number of devices assembled in this manner indicated that approximately one-third of the devices demonstrated FET behavior with partial or full gating.

4.1.2 Viruses as templates for growth and self-assembly of inorganic materials

The structural and recognition properties of a number of virus capsids, including M13 bacteriophage, tobacco mosaic virus (TMV), cowpea chlorotic mottle virus (CCMV), and cowpea mosaic virus (CPMV), have been exploited to nucleate the growth and self-assembly of many different inorganic materials.[24] A critical factor in developing a virus directed self-assembly template is identification of appropriate compatible biological-inorganic material systems. To this end, peptide combinatorial libraries may be employed to identify

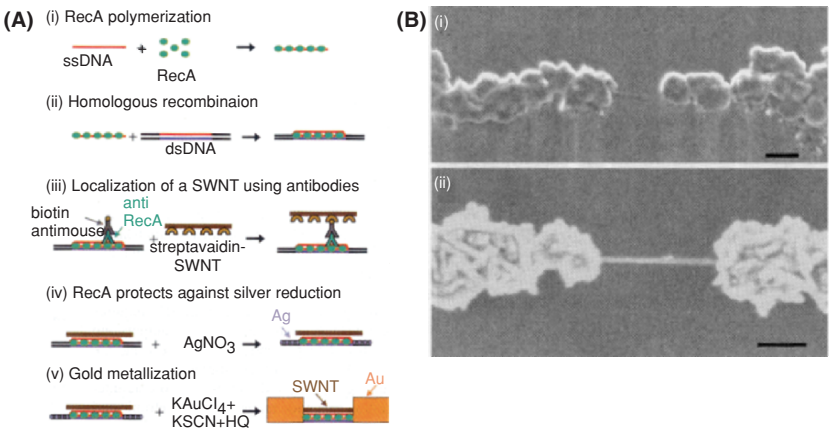


Figure 3. (A) Schematic detailing self-assembly of DNA templated carbon nanotube FETs. (B) Scanning electron micrograph of DNA templated carbon nanotube FETs comprising (i) an individual nanotube and (ii) a rope of nanotubes. Scale bars, 100 nm.[23] © American Association for Advancement of Science.

viral peptide sites with specific interactions to inorganic target materials. Using this approach, Belcher and coworkers[25,26] have isolated, cloned, and amplified peptide sequences from the M13 bacteriophage virus with binding specificity to ZnS and CdS crystal surfaces, respectively, and demonstrated directed self-assembly of quantum dots with controlled nanoparticle size and crystal structures for both.

By genetically modifying an M13 virus to display the selected peptides in the helical major protein coat, pVIII, ZnS hybrid viral-semiconductor quantum dot nanowires are successfully fabricated.[27] The technique is extended to the templated self-assembly of CoPt and FePt magnetic nanowires.[28] Mann and coworkers[29] have used TMV templates for controlled deposition and organization of Pt, Au, and Ag nanoparticles (see Fig. 4). Cage architectures based on CCMV were also employed in the controlled synthesis of 5 nm diameter ferrimagnetic iron oxide maghemite nanoparticles.[30,31]

4.1.3 Protein templated self-assembly

Microtubule components of the cytoskeletons of eukaryotic cells are protein filaments with outer diameters of 25 nm and lengths of several micrometers. Due to their large geometrical aspect ratios, they are well suited to be used as biomolecular templates for assembly of magnetic and electrically conductive nanowires. Three-dimensional metalization of microtubules is demonstrated using electroless plating of nickel initiated by palladium cata-

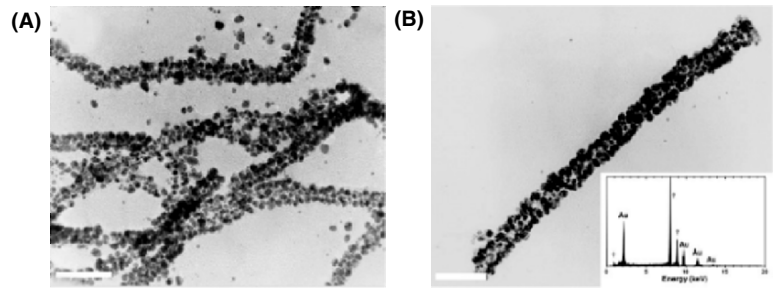


Figure 4. Tunneling electron micrographs of gold nanoparticles in wire-like structures formed on the surface of a TMV virus. (A) Low magnification image showing multiple TMV rods. Scale bar, 100 nm. (B) Single TMV rod with dense coating of gold nanoparticles. Scale bar, 50 nm. Inset shows the corresponding EDXA spectrum.[29] © American Chemical Society.

lysts adsorbed at the microtubules surface.[32] Lindquist *et al*.[33] have demonstrated fabrication of conducting nanowires using self-assembled amyloid fiber templates. Genetically modified protein fibers (from *Saccharomyces cerevisiae*) that presented reactive, surface-accessible cysteine residues are used to covalently link to colloidal gold particles. Additional metal, deposited by chemical enhancement of the colloidal gold, was deposited by reductive deposition to form biotemplated metal wires that exhibited low resistance and ohmic behavior when characterized (see Fig. 5).

Yu and coworkers[34] have also reported synthesis of shape-controlled silver nanoparticles using nanotube templates.[34] A specific peptide sequence, AG4 (identified using peptide combinatorial libraries), with binding specificity for Ag nanocrystals was isolated, sequenced, and incorporated onto template nanotube surfaces. Biomineralization of Ag ions, using these templates, enabled the isotropic growth of hexagonal shaped Ag nanocrystals.

Magnetotactic bacteria (MTB) orient and migrate along the geomagnetic field towards favorable habitats, a behavior known as magnetotaxis. Since the first report of MTB by Blakemore,[35] subsequent studies have shown that MTB are a morphologically diverse and cosmopolitan group of aquatic microorganisms inhabiting freshwater and marine environments ranging from aerobic to anoxic.[36,37] Blakemore noticed that some of the bacteria observed under a microscope always moved to the same side of the slide. On holding a magnet near the slide, the bacteria would move towards the north end of the magnet. These bacteria are able to do this because they make tiny, iron-containing, magnetic particles. The bacteria arrange these tiny magnets in a line to make one long magnet, and they use this magnet as a compass to align themselves to the earth's geomagnetic field. Based on this property, an opportunity exists using genetic recombination to combine this

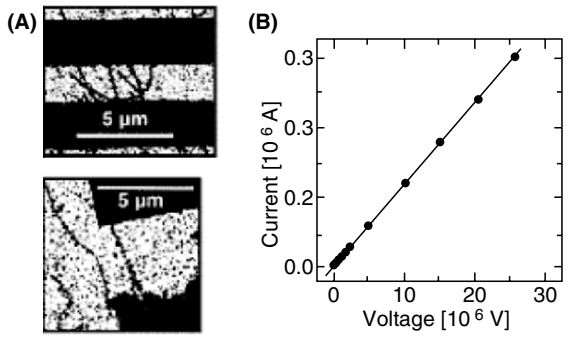


Figure 5. (A) Gold-coated protein fibers bridging the gap between two electrodes. (B) Corresponding IV curve, demonstrating ohmic conductivity with low resistance.[33] © National Academy of Sciences.

magnetotactic characteristic with expression of specific proteins, to enable magnetic field programmable deposition of specific proteins in nanodevices.

5. CONCLUSION

Faced with the current challenge of sustaining Moore's Law in the semiconductor industry as "top–down" fabrication methods based on lithography approach, the fundamental limits of physics, research on fabrication solutions provided by nature is likely to lead to the development of bioinspired and biohybrid nanodevices. The potential exists for the development of alternative fabrication processes to lithography, capable of yielding 3D nanodevices using a combination of biomolecules and bioinspired processes. While research to date has demonstrated that functional biohybrid nanodevices can be fabricated using bioinspired "bottom-up" approaches, a vast array of useful biomolecules and bioprocesses await discovery.

Despite the ever-increasing understanding of the functionality and interaction of different cell types within existing model organisms (i.e., including humans), the complexity of many advanced organisms together with their specific environmental requirements can be limiting. Therefore, more primitive organisms with a greater diversity of environmental tolerances would provide more opportunities for identification and exploitation of proteins useful for novel nanodevices. The diversity envisaged is especially evident in organisms known as "extremophiles" that can survive and even thrive in what are generally considered as harsh environmental conditions. Exploring these organisms in detail, therefore, may yield proteins with unique properties exploitable for innovative nanodevices.

REFERENCES

1. Moore, G. E., 1965, Cramming more onto integrated circuits, *Electronic Magazine* (35th anniversary issue), April 19, 1965.
2. International Technology Roadmap for Semiconductors, http://public.itrs.net
3. Ringler, P., and Schulz, G. E., 2002, *Chem. BioChem*. **463:**27.
4. Hartgerink, J. D., Beniash, E., and Stupp, S. I., 2002, *Proc. Natl. Acad. Sci*. **99**(8):5133–5138.
5. Gyorvary, E. S., O'Riordan, A., Quinn, A. J., Redmond, G, Pum, D., and Sleytr, U. B., 2003, *Nano Lett*. **3**:315.
6. Douglas, K., Clark, N. A., and Rothschild, K. J., 1986, *Appl. Phys. Lett*. **48**:676.
7. Douglas, K., Devaud, G., and Clark, N. A., 1992, *Science* **257:**642.
8. Moore, J. T., Beale, P. D., Winningham, T. A., and Douglas K., 1997, *Appl. Phys. Lett*. **71**:1264.
9. Winningham, T. A., Whipple, S., and Douglas, K., 2001, *J. Vac. Sci. Technol. B* **19**:1796.
10. Douglas, K., 1996, *Biomimetic Materials Chemistry* (S. Mann, ed.), VCH Inc, Weinhein, Germany p. 117.
11. Malkinski, L., Camley, R. E., Celinski, Z., Winningham, T. A., Whipple, S. G., Douglas K., 2003, *J. Appl. Phys*. **93:**7325.
12. Shenton, W., Pum, D., Sleytr, U. B., and Mann, S., 1997, *Nature* **389:**585.
13. Winningham, T. A., Gillis, H. P., Choutov, D. A., Martin, K. P., Moore, J. T., and Douglas, K., 1998, *Surf. Sci*. **406:**221.
14. Mertig, M., Kirsch, R., Pompe, W., and Engelhardt, H., 1999, *Eur. Phys. J. D* **9:**45.
15. Hall, S. R., Ahenton, W., Englehardt, H., and Mann, S., *Chem. Physchem*. **2:**184.
16. Moll, D., Huber, C., Schlegel, B., Pum, D., Sleytr, U. B., and Sara, M., 2002, *Procs. Natl. Acad. Sci., USA* **99:**14646.
17. Sleytr, U. B., Bayley, H. Sara, M. Breitwieser, A. Kupcu, S. Maderm C. Weigert, S. Unger, F. M., Messner, P., Jahn-Schmid, B., Schuster, B., Pum, D., Douglas, K., Clark, N. A., Moore, J. T., Winningham, T. A., Levy, S., Frithsen, I., Pankovc, J., *et al*., 1997, *FEEMS Microbiol. Rev*. **20:**151.
18. Yonezawa, T., Onoue, S., and Kunitake, T., 1999, *Koobunshi Ronbunshu* **56:**855.
19. Zhaoxiang, D., and Chengde, M., 2003, *Nano Lett*. **3:**1545.
20. Warner, M. G., and Hutchinson, S. E., 2003, *Nat. Mater*. **2:**272.
21. Yan, H., Park, S. H., Finkelstein, G., Reif, J. H., and LaBean, T. H., 2003, *Science* **301:**1882.
22. Kerren, K., Krueger, M., Gilad R., Ben-Yospeh, G., Sivan, U., and Braun, E., 2002, *Science* **297:**72.
23. Keren, K., Berman, R. S., Buchstab, E., Sivan, U., and Braun, E., 2003, *Science* **302:**1380.
24. Flynn, C. E., Lee, S-W., Peelle, B. R., and Belcher, A. M., *Acta Materialia* **51:**5867.
25. Lee, S.-W., Mao, C., Flynn, C. E., and Belcher, A. M., 2002, *Science* **196:**892.
26. Flynn, C. E., Mao, C., Hayhurst, A., Williams, J. E., Georgiou, G., Iverson, B., and Belcher, A. M., 2003, *J. Mater. Chem*. **13:**2414.
27. Mao, M., Flynn C. E., Hayhurst, A., Sweeney, R., Qi, J., Georgiou, G., Iverson, B., and Belcher, A. M., 2003, *Procs. Natl. Acad. Sci., USA* **100:**6946.
28. Mao, C., Solis, D. J., Reiss, B. D., Kottmann, S. T., Sweeney, R. Y., Hayhurst, A., Georgiou, G., Iverson, B., and Belcher, A. M., 2003, *Science* **303:**213.
29. Dujardin, E., Peet, C., Stubbs, G., Culver, J. N., and Mann, S., 2003, *Nano Lett*. **3:**413.

30. Douglas, T., Strable, E., Willits, D., Aitouchen, A., Libera, M., and Young, M., 2002, *Adv. Mater*. **14:**415.
31. Allen, M., Willits, D., Mosolf, J., Young, M., and Douglas, T., 2002, *Adv. Mater*. **14:**1562.
32. Kirsch, R., Mertig, M., Pompe, W., Wahl, R., Sadowski, G., Bohm, K. J., and Unger, E., 1997, *Thin Solid Films* **305:**248.
33. Scheibel, T., Parthasarathy, B., Swaicki, G., Lin, X-M., Jaeger, H., and Lindquist, L., 2003, *Proc. Natl. Acad. Sci., USA* **100:**4527.
34. Yu, L., Banerjee, I. A., and Matsui, H., 2003, *J. Am. Chem. Soc*. **125:**14837.
35. Blakemore, R., 1975, *Science* **190:**377–379.
36. Fassbinder, J. W. E., Stanjek, H., and Vali, H., 1990, *Nature* **343:**161–163.
37. Matsunaga, T., Sakaguchi, T., and Tadokoro, F., 1991, *Appl. Microbiol. Biotechnol*. **35:**651–655.