THE ANATOMICAL RECORD (NEW ANAT.) 261:198–216, 2000

FEATURE ARTICLE

The New Architectonics: An Invitation To Structural Biology

CLARENCE E. SCHUTT* AND UNO LINDBERG

The philosophy of art might offer an epistemological basis for talking about the complexity of biological molecules in a meaningful way. The analysis of artistic compositions requires the resolution of intrinsic tensions between disparate sensory categories—color, line and form—not unlike those encountered in looking at the surfaces of protein molecules, where charge, polarity, hydrophobicity, and shape compete for our attentions. Complex living systems exhibit behaviors such as contraction waves moving along muscle fibers, or shivers passing through the growth cones of migrating neurons, that are easy to describe with common words, but difficult to explain in terms of the language of chemistry. The problem follows from a lack of everyday experience with processes that move towards equilibrium by switching between crystalline order and chain-like disorder, a commonplace occurrence in the submicroscopic world of proteins. Since most of what is understood about protein function comes from studies of isolated macromolecules in solution, a serious gap exists between what we know and what we would like to know about organized biological systems. Closing this gap can be achieved by recognizing that protein molecules reside in gradients of Gibbs free energy, where local forces and movements can be large compared with Brownian motion. *Architectonics*, a term borrowed from the philosophical literature, symbolizes the eventual union of the structure of theories—how our minds construct the world—with the theory of structures—or how stability is maintained in the chaotic world of microsystems. *Anat Rec (New Anat) 261:198–216, 2000.* © 2000 Wiley-Liss, Inc.

KEY WORDS: architectonics; cytoskeleton; functional microanatomy; molecular motor; structural biology; proteins; hydrophobic effect; Gibb's free energy; thermodynamics; proteomics

REDUCTIONISM IN THE POST-GENOMIC ERA

We are coming to grips with our molecular destiny. Just one hundred years ago, when it appeared impossible that atoms could ever be seen, a plausible case could be made that they didn't exist at all, and even if they did, it wasn't clear how mere aggregations could give rise to distinct qualities (Duhem, 1914). The application of x-ray analysis to crystals by Max von Laue, William Bragg, Sr. and Jr., changed all of that in short order. Their legacy, delivered by subsequent generations of x-ray crystallographers, is that structures of complex biological macromolecules, whose presence was barely suspected a century ago, can not only be seen, but obey the ordinary laws of physical chemistry. Genes themselves, a mathematical abstraction needed to explain regularities in plant heredity, were revealed as elaborate three-dimensional structures containing nucleic acids wrapped around proteins. The human genome—nothing less than a highly encoded map of our material being, its history and development—will soon be freely available to everyone to probe, exploit, or to study as new form of knowledge about ourselves. While it is not at all obvious how we will read this map, it is a certainty that a sound understanding of the structure of proteins will be needed, because all of life's intermeshed processes, including the natural decoding of the message, act through them.

We are now accustomed to hearing about "the post-genomics era" and words like "proteomics" and "emergomics" are meant to alert us to where the real intellectual or technological (not to say investment) opportunities lie. The "proteome" is the set of all proteins actually present in a given cell, as revealed by electrophoresis gels capable of resolving thousands of distinct spots, each reporting the presence of a single protein. Similarly, patterns of gene expression, both spa-

Dr. Lindberg is Professor of Zoological Biology at the Wenner-Gren Institute of Stockholm University. He is First Vice-President and Foreign Secretary of The Royal Swedish Academy of Sciences, and Past Chair of the Stockholm branch of the Swedish Folk University. Dr. Schutt is Professor of Chemistry at Princeton University. He is a Member of the Program in Neuroscience, an Associated Faculty Member in the Molecular Biology Department, and former Director of the Molecular Biophysics Program. He is Chairman of the Board of the National Alliance for Autism Research and Annual Co-Chair and Founder of the Princeton-Eden Lecture Series on Autism.
***Correspondence to: Clarence E. Schutt, The Henry H. Hoyt Laboratory, Washington Road, Princeton University, Princeton, NJ, 08544, USA. E-mail: schutt@chemvax.princeton.edu**

© 2000 Wiley-Liss, Inc.

tially and temporally, can be found by expanding the repertoire of messenger RNA present in individual cells and identifying them with antibodies or anti-sense oligomers of DNA. The amount of information will be staggering and its implications for understanding human diseases and affecting cures can be hardly exaggerated.

All of this appears to represent the triumph of *reductionism*; all components and sub-components of living systems can be described, in principle, by atomic structures. Atoms combine to form molecules, some of them very large indeed, such as proteins, which require for their synthesis miniature machine-like ribosomes. Proteins are synthesized as long chains, annealing into compact structures in tiny oven-like chambers called chaperonins. Recently, the structures of some of the chaperonins (Sigler et al., 1998), and even the ribosome itself (Wimberly et al., 2000; Ban et al., 2000), have yielded their secrets to crystallographers, now armed with powerful synchrotron x-ray sources, and computers with awesome capabilities for data reduction and visualization. These structures show us the arrangement of the working parts and set the stage for understanding how these movements are coordinated to carry out the functions of these complex particles.

WHAT IS ARCHITECTONICS?

Architectonics is a venerable term in the philosophical literature. It pertains to both "the theory of structure", as in the architectural literature, and "the structure of theory", used by philosophers in a variety of ways to denote the processes by which ideas become fixed in our consciousness (Kant, 1787). Suzanne K. Langer (1967) in her magisterial three-volume work on the philosophy of art uses the word "architectonics" to expound on the discursive, non-sequential, quality of visual art forms: "Their complexity, consequently, is not limited, as the complexity of discourse is limited, by what the mind can retain from the beginning of an apperceptive act to the end of it. Of course, such a restriction on discourse sets bounds to the complexity of speakable ideas. An idea that contains too many minute yet closely related parts, too many relations within relations, cannot be 'projected' onto discursive forms; it is too subtle for speech." Langer (1967) is claiming that the boundary conditions set by the linear structure of sentences limits what we can communicate about the arts through language alone. The great works of art must speak for themselves directly through the multiple portals of our senses to some consciousness beyond words.

In many ways, the surfaces of proteins share this discursive aspect with the visual arts (see Figures 1 and 2). It is difficult to find words to describe the complexity of form one sees. One is clearly in the presence of a composition, but even the canonical forms of sculpture, including Alexander Calder's mobiles, are inadequate to the task of describing all of the events that take place during protein folding or during the encounter of two proteins. Even if the surfaces are color-coded for charge, hydrophobicity, and polarity, the presence of ridges, clefts, knobs, and holes adds so much texture and landscape to the chemical complexity that one is at a loss for where to begin.

***Architectonics* pertains to both "the theory of structure," as in the architectural literature, and "the structure of theory," used by philosophers to denote the processes by which ideas become fixed in our consciousness.**

In recent years, crystallographers have determined structures for a large number of functional protein complexes. What these structures reveal is a high degree of complementarity, both chemical and topological, between the surfaces of interacting proteins (see Figure 2) . These surfaces are somewhat easier to describe because they are shaped to conform to one another. Even though each is complicated in itself, the pleasing "clasped hands" nature of their mutual interactions, allows us to grasp the meaning of the functional association. As one becomes familiar with different classes of interactions and motifs by studying many crystal structures, a kind of visual gallery develops, which can form the basis for a language having its own syntax and semantic rules. To participate in the discourses of structural biology, one must learn how to express what one means to say about biological function in this language.

It is this aspect of structural biology that invites comparison with formal criticism in the arts or in literature. When an art critic assesses a new work, he or she assumes that the reader knows something of the history of art. To see the new demands knowledge of the old. But there is more to it. References to earlier works carry with them contexts of discovery, breakthrough and failure, drawing from the reader's consciousness a broad spectrum of memories and feelings. It is the intention of the critic to set the stage for these evoked responses. The mind then becomes receptive to the critic's message. In structural biology, certain structures are "classic" in the same way that the paintings of Cezanne, Van Gogh, Picasso, Rothko, or Stella define unique compositional styles, immediately recognizable, evocative. We will refer to these classic structures as "structural paradigms."

But, architectonics is not the same as architecture; "tonics" refers to tensions inherent in a structure. Returning to Langer's analysis of artistic composition: "The most fundamental elements seem to be tensions; and upon closer inspection, tensions show some peculiarly interesting traits. By their very occurrence they immediately engender a structure. They act on each other in a great variety of ways—they can be handled so as to intersect without losing their identity, or contrariwise, so that they fuse and compose entirely new elements. They

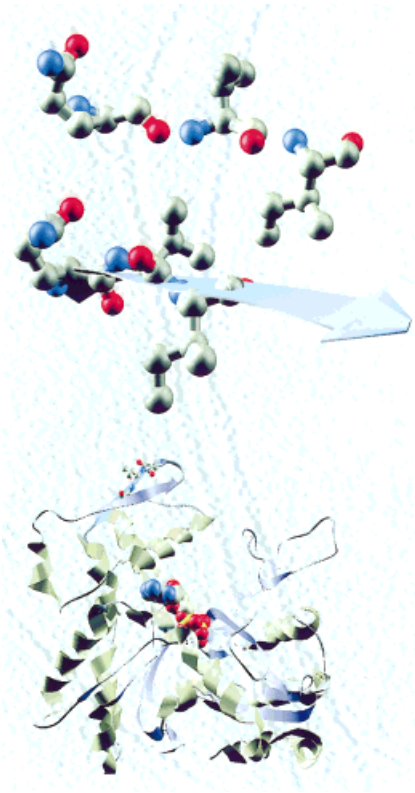

Figure 1. Hierarchical structure of proteins. Proteins are covalently-bonded polymers of the twenty naturally-occurring amino acids. Polypeptide chains fold into three-dimensional structures, unique for each sequence of amino acids. The actin molecule is shown as an example. Higher order structures consist of subunits of proteins held together by extensive non-covalent interactions. The helical actin filament is shown in light relief in the background. Figure originally appeared in Kreatsoulas et al. (1999).

can be intensified or muted, resolved either by being spent or by being counterbalanced, modified by a touch, and all the while they make for structure. This appears to be true in all the great orders of art; in every one of them, a general range of tensions is set up by the first element —line, gesture, or tone—which the artist establishes" (Langer, 1967).

Proteins can be described as hierarchical structures using similar language (Figure 1). The lowest level of the hierarchy is the polypeptide chain itself. Under the proper conditions, the chain can fold back upon itself forming elements of secondary structure, such as helices and sheets, which subsequently pack against each other to form tertiary, or fully-folded, structures. Proteins interact with each other to form higher-order structures like filaments, tubules, and viral shells.

These "first elements" of biological structure are not rocks or bricks, but dynamic units with internal tensions between the competing tendencies for order and disorder that lie at the heart of life's rhythmic processes. The structures of DNA and RNA are the primary elements of storage and retrieval of genetic information. Their twists and turns, and sequence variations, define a world of sufficient template complexity to encode life's design. In the cytoplasm, actin is the first element of a system of meshed and cross-linked filaments that enables cells to move, to pull, and to change shape. These molecules are "bobbing" in a torrent of Gibbs free energy, a quantity of immense importance for understanding the emergence and maintenance of biological complexity (see Box I). The flow of free energy starts in the nuclear transformations in the sun's core, which produce a vast stream of ultraviolet photons that are harnessed by the photosynthetic machinery of plants. These little factories produce glucose molecules, rich in chemical bond energy, that are converted by mitochondria in plant and animal cells into ATP, adenosine trisphosphate, a universal fuel whose breakdown into ADP and inorganic phosphate by muscles and other kinds of transport machinery makes life possible.

STRUCTURAL PARADIGMS

It is sometimes claimed by those molecular biologists who can perform the neat teleological trick of decanting "function" from "structure" that structural biology is nothing more than a tool that provides scaffolds upon which to hang notions of information flow. This view presupposes "structural paradigms," structures whose interpretations have led to new principles and ways of looking at biological processes. The discovery of new structural paradigms is akin to discoveries in the world of art because the microscopic world never looks the same once they have appeared. The lessons learned from these classic structures, passively absorbed, are unconsciously at play in the formulation of molecular mechanisms for all biological processes.

As an exercise in thinking without structural paradigms, try to define *enzyme specificity* without recourse to specific molecules. A good start might be to invoke the notion of "lock and key". Fair enough, this idea served enzymologists rather well during the decades before the structures of the first enzymes were obtained. But, exactly how is the lock fashioned from bits of polypeptide chain? How unique is the lock, how secure from tampering (e.g., by oxidants)? Once the key is inserted (a process signposted by a decreasing Gibbs free energy, described below), what makes it turn? And how does it get out? Are molecules of a certain softness or pliability required to meet a particular functional need? Are these requirements in line with the properties of bonded carbon, nitrogen, oxygen, and hydrogen? The question boils down to this: given complete knowledge of the laws of physical chemistry, can you provide a plausible basis for the diversity of biological reactions carried out by macromolecules (gene regulation, force generation, signal transduction), *without referring to results obtained by x-ray crystallography?*

For example, compare the structures of trypsin, chymotrypsin, and elastase (Stryer, 1988). Notice how naturally the description of specificity follows, how easily the common *catalytic triad* is distinguished from the *specificity pocket*—no fuzzy thinking, no dubious assumptions. While its function could conceivably be deduced from its structure (it might be easy to spot an icosahedral virus, or a globin e.g.), the reverse operation cannot yet be performed without relying on family resemblances within an established paradigm. This is the aspect of biology that makes it so different from physics, where First Principles generally can be specialized to the case at hand.

The helix-loop-helix DNA-binding motif is another classic piece of protein architecture. In looking at a new DNA-binding protein, one looks first

to see if this element is present, and considers other precedents in trying to classify the new in terms of the old. But in doing so, what should echo in the mind are the struggles of the early band of protein-DNA crystallographers. Carl Pabo, Tom Steitz, Brian Matthews, Steve Harrison and others were trying to find a structural code in the patterns of charge and polarity on the surfaces of these proteins that might be used to predict how the edges of specific base pairs could be recognized in the major groove. Implicit in these attempts was the belief that the phosphodiester sugar backbone of DNA was an invariant, sequence-independent, structural feature.

The famous Watson-Crick B-form of DNA was originally deduced from x-ray fiber patterns. Fiber diffraction, though clearly of value in defining helical parameters such as helical pitch and repeat distances, utilizes specimens in which the fibrous molecules are randomly oriented about the fiber axis. Many details are averaged out when fiber patterns are used to deduce structural models. The importance of these details did not become clear until Richard Dickerson and Horace Drew solved the crystal structures of several short DNA stretches by x-ray crystallography. They found sequence-dependent variation in the structural parameters for DNA, suggesting that more subtle geometric aspects of the structure might be involved in base pair recognition than just the matching of rigid surfaces (Dickerson and Drew, 1981).

As crystal structures for more and more protein-DNA complexes were determined, it was clear that base pair recognition was based, not just on the geometric possibilities of the helix-loop-helix motif, but on other motifs as well, and especially on the potential for DNA to undergo changes in structure (Anderson et al., 1987; Davies et al., 2000). When the energy of deformation enters into the balance of factors determining whether two surfaces can bind, the search for a simple structural recognition code (e.g., arginine side-chains recognize GC base pairs on a rigid DNA molecule) had to be abandoned (Pabo and Neludova, 2000). Indeed, the reverse side of the coin, how one set of proteins (the histones) can package up all possible DNA sequences into chromatin, has been revealed in an elegant crystal structure analysis by Tim Richmond and colleagues (Luger et al., 1997).

Chris Calladine, a structural engineer working a stone's throw away from the Cavendish Laboratory in Cambridge where it all started, developed a set of simple rules that made it possible to describe deformations of DNA in terms of base twists and tilts and other operations on the molecule (for recent review see Lu and Olson, 1999). While these rules have a mechanical flavor easy to comprehend, a bridge has been crossed where, even though the machine-like character of intermolecular interactions involved in "recognition" is self-evident, everyday words to communicate the sequential steps in these processes are becoming increasingly more difficult to find (Kono and Sarai, 1999).

To participate in the discourses of structural biology, one must learn how to express what one means to say about biological function in this language.

MICHAEL POLYANI AND LIFE'S IRREDUCIBLE STRUCTURE

We will argue here that the Gibbs free energy concept, when applied to protein molecules, is the First Principle of structural biology. However, to apply thermodynamic concepts, such as free energy, to complex hierarchical systems, it is necessary to draw boundaries (physical or conceptual) between levels of the hierarchy. Over thirty years ago, just as the revolution in structural biology was getting underway, Michael Polyani (Polyani, 1969) wrote about the machinery of life as having to satisfy "boundary conditions" in a multi-leveled hierarchy. The analogy to boundary conditions, a concept taken from mathematical physics, is apt. When a taut wire is strung between two points and plucked, the possible waveforms on the wire are restricted to those which are integral fractions of the distance between the endpoints; the imposition of a boundary condition, in this case the pinned ends of the wire, results in the selection of a particular subset of possible harmonic motions.

Similarly, the solutions to Schrodinger's wave equation, the distinctively shaped atomic orbitals (s,p,d,f) familiar from freshman chemistry, arise from the restriction that negatively charged electrons stay in the potential well of the positively charged nucleus. When electrons are forced to accommodate the restrictions of multi-centered positive charges, as when six atoms of carbon come together to form benzene rings, new electronic forms emerge, imposing a planar hexagonal structure on the nuclear centers themselves. The benzene molecule is frequently used in the science of "complexity" to illustrate "emergent properties", or the idea that the whole is greater than the sum of its parts, in this case the aromaticity of the de-localized electrons is the emergent property.

Polyani saw "DNA structure", the double helical pattern of hydrogen-bonded bases projecting inwards from a common phosphodiester backbone, as organizing atoms (P,O,N,C,H) from a lower, information-free, level of the hierarchy into a readable blueprint. Any other constellation of the same atoms would be meaningless when presented to the transcription apparatus, a complex ensemble of protein molecules. Polyani stressed the *irreducibility* of the information coded in the sequence of bases, the fact that knowing everything about the quantum mechanical laws governing atomic structure tells us nothing about the content of the coded message or its origin. He wrote that: "morphogenesis, the process by which the structure of living beings develops, can then be likened to the shaping of a machine which will act as a boundary for the laws of inanimate nature". For Polyani, all working parts of a living

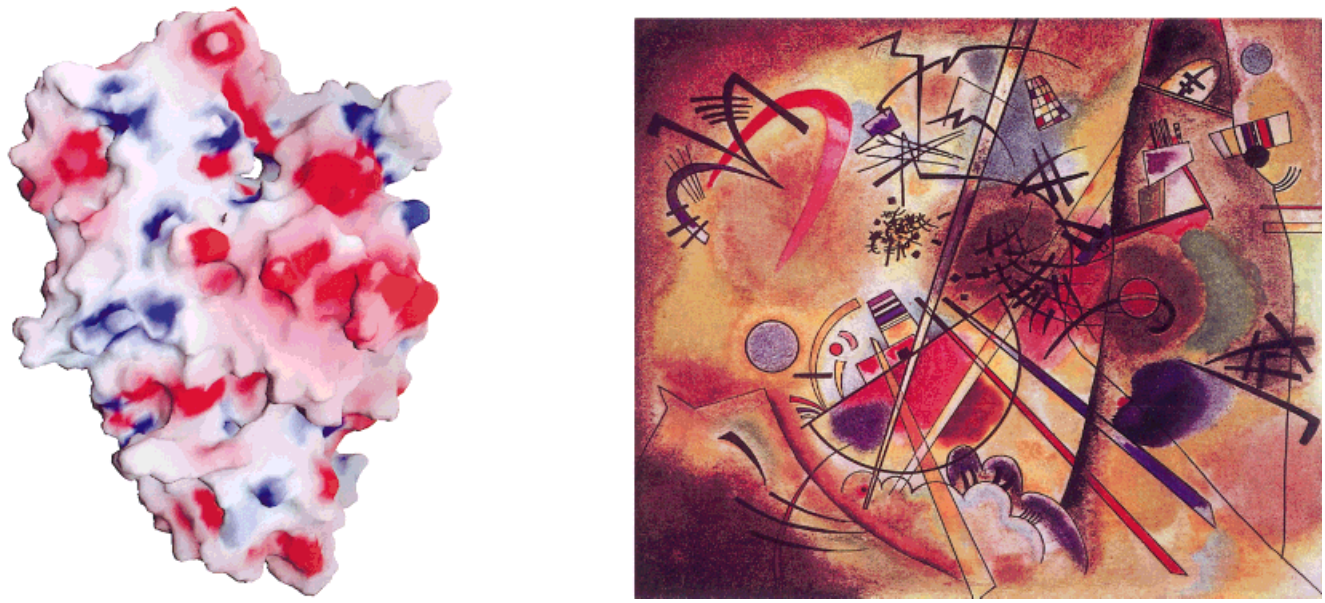

Figure 2. Surface complexity. Proteins reflect a degree of complexity similar to the paintings of Kandinsky. The profilin:actin complex is shown on the left as a solid object color-coded for electric charge (blue is positive, from lysine and arginine amino acid sidechains, and red for the acidic carboxylate groups of aspartic acid, glutamic acid, and the carboxy termini of actin and profilin). The corrugated surface, pattern of electrostatic charges, hydrophobicity, and hydrogen-bonds of proteins comprise a unique interaction site for other proteins that have complementary surfaces. Kandinsky explored how the diverse elements of painting (color, line, and form) worked harmoniously to convey artistic meaning. Shown on the right is "Small Dream in Red, 1925), which appeared in the first edition of Kandinsky's *Point and Line to Plane* (from *Kandinsky 1866-1944 The Journey to Abstraction* by Ulrike Becks-Malorny, Köln: Benedikt Taschen Verlag GmbH (1994), p. 152. ISBN 3-8228-9045-6 Permission to publish obtained from the Kunst museum, Bern, Germany.).

system are under "dual control", whereby higher principles work on principles at a lower, more fundamental level such as the laws of physics and chemistry. Using written language as an example of a hierarchic system (letters-to-words-to-sentences-to-meaning), he argued that boundary conditions establish a semantic relation between lower and higher levels.

The concept of "higher principles" may sound old-fashioned to the modern mind, since we can now "see" (owing to the power of electron microscopy and crystallography) genes packaged into chromatin, the basic unit of structure of chromosomes, and how they are transcribed and translated into proteins. Also, Polyani was speaking about holistic concepts such as "epigenetic landscapes" and "morphological fields", which have been recast by modern developmental biologists in strictly molecular terms (Edelman, 1988). For example, a system of "semaphorin" molecules establishes signposts for migrating neurons as the developing brain connects itself. Yet, Polyani's notion of boundary conditions in functional hierarchies is a fertile one, and represents the Second Principle of structural biology. Protein molecules can be regarded as key elements of "Polyani machines" shaped to apply boundary conditions in the functional hierarchy of life. For example, cell-cell interactions are mediated by protein-protein interactions at the cell surface, which transfer information across cell membranes to initiate intracellular signaling cascades.

The discovery of new structural paradigms is akin to discoveries in the world of art because the microscopic world never looks the same once they have appeared.

PROTEINS AS "BOUNDARY CONDITIONS"

The eukaryotic cell is the basic organizing unit of higher life forms. Thousands of proteins on the surfaces of cells continuously sample their environments, searching for molecular cues on which directions might be favorable for finding nutrients, or the correct turn to take during development. These results are passed through the plasma membrane to specialized protein structures, called "signal transduction complexes," that shape the message that is transmitted to the nucleus. The concept of limiting surfaces formed from protein molecules allows us to see signal transduction as an "assembly process", in which the signal on the outside, perhaps a growth factor, induces the formation of structures on the inside of

BOX I: The Laws of Thermodynamics

All processes in nature are subject to the first and second laws of thermodynamics. The first law, the equivalence of work and energy and the conservation of their sum, is easy to apply once one learns how to quantify heat and to calculate the work done on or by an isolated system (Klotz, 1967). The second law, that the entropy of the universe must increase in any spontaneous process, is inherently impossible to apply because the whole universe is not accessible to measurement.

Willard Gibbs, in a remarkable scientific treatise, showed how the concept of free energy allows a neat solution to this problem. The first and second laws can be combined into one formula expressed in terms of system variables only (i.e. "changes in the rest of the universe" are not explicitly used) (Gibbs, 1906; Denbigh, 1971):

$$\Delta G_{sys} = \Delta H_{sys} - T\Delta S_{sys} + \textbf{work done on "system"}$$

(Gibbs' formula)

where, $\Delta G_{sys} = G_{final} - G_{initial}$ is the difference in free energy between initial and final states, $\Delta H_{sys} = H_{final} - H_{initial}$ is the difference in internal energy (enthalpy, at constant pressure), T is the temperature, $\Delta S_{sys} = S_{final} - S_{initial}$ is the difference in entropy of the 'system' alone, and the work done on the 'system' can be by a variety of external agents.

$\Delta G_{sys} \leq 0$ expresses the notion that a process described by the initial and final states of the "system" will not occur spontaneously unless the free energy decreases. When $\Delta = 0$ the 'system' is in equilibrium. Transferring material between two states of equal Gibbs free energy, though they may differ in structure, as with two phases in equilibrium (liquid $\leftrightarrow$ vapor), will not involve any net differences in enthalpy or entropy from which useful work can be derived.

The "work" terms most commonly used in biophysics (Katchalsky and Curran, 1964; Klotz, 1967; Morowitz, 1970) are

Mechanical

$$\begin{cases} P\Delta V (\textit{volume change in 'system' against a pressure } P \\ \sigma\Delta A (\textit{change in surface area at a surface tensions} \\ \tau\Delta l (\textit{change in length at a tension } \tau \end{cases}$$

Electrical

$\Psi\Delta q$
(movement of charge against an electric potential)

Chemical gradients

$\sum_i \mu_i dn_i$
(movement of molecule against a concentration gradient: $\mu_i = \mu_i^0 + RT \ln C_i$; chemical potential of the ith species)

The "work" terms include all externally wrought changes in the structure of a system, measured in terms of volume, surface area and length. The fixed attributes of the system are the pressure, P, surface tension, σ, and tension τ. It takes work, forces resisted over distances, to bring about these changes. The re-ordering of matter and charge involved in these processes, can be driven by chemical gradients and reactions. The free energy of the system will be changed by the work done on it. This will affect the heat capacity of the system. In the final analysis, this appears as changes in bond energy in the molecules comprising the system or in the intrinsic disorder. An internal combustion engine is a perfect example of system in which a complex chemical liquid is converted to simple gases and heat in a device producing mechanical work. The total free energy available from the bond energy in the molecules, expressed as chemical potentials, balances the net change in heat released and the work produced.

Ludwig Boltzmann proposed a definition of entropy that enables changes in *molecular structure* to be used in calculating free energy changes

$$S = R \ln W \ (\textit{Boltzmann's formula})$$

where R is the gas constant (the product of Boltzmann's constant times Avogadro's number) and W is the number of ways in which the elements of a system can be arranged (Hill, 1960. The field of statistical thermodynamics is concerned with a careful evaluation of W under all sorts of conditions). The entropy of a perfect crystal ($W_{crystal} = 1$) of argon atoms, for example, is zero at a temperature of absolute zero (we will ignore entropy associated with the zero point energies of quantum mechanics). By introducing a small number of vacancies or imperfections (by decreasing the crystal density) the disorder of the system increases by a *calculable* amount (see Denbigh, 1971 for a rigorous discussion of the perfect crystal). Vacancies increase the heat capacity of the crystal lattice because they allow atoms to more easily hop about. A cornerstone of information theory is a statement that "information" can be quantified using Bolzmann's formula. The basic idea is that a system can be completely described in the form of answers to a series of *yes* or *no* questions expressed as a string of ones and zeroes. Thus, entropy provides a measure of our ignorance of atomic positions, quantifiable as an increase in entropy.

Stretching a rubber band results in its radiating heat in the amount $-T\Delta S$ because stretching reduces the number of possible ways that the cross-linked polypropylene polymer chains can arrange themselves is smaller ($W_{chain} = 1$, in the limit). The stretched system no longer needs as much random energy to remain in equilibrium with its surroundings (Hill, 1960). Its Gibbs free energy is higher because of the work done on it in stretching. Slowly releasing the rubber band allows it to lower its free energy as it performs recoverable work ($-\tau\Delta l$) on some external object. In the course of shortening, the polymer chains take up heat from their surroundings, increasing the total number of allowed configurations ($W_{chain} \gg 1$). The unstretched rubber band in thermal equilibrium with its surroundings is most stable when the polymer chains are least structured!

Rubber elasticity illustrates the idea that contracting polymers absorb heat at constant temperature by increasing the number of degrees of rotational freedom about their bonds. The heat capacity of crystals, on the other hand, has two aspects: volume changes, which allow more degrees of freedom through vacancy creation, and increases in internal energy of the molecules in the lattice. Proteins continuously interconvert enthalpy and entropy in performing work in cells, where the Gibbs free energy is kept well above the minimum by external sources. Without structures such as proteins and lipids that can perform the feat of changing degrees of internal order, life would be impossible at ordinary temperatures. In fact, it appears to be generally true that already-folded proteins lose their ability to function at a temperature below 220 Kelvin (Rasmussen et al., 1992).

the plasma membrane. These protein complexes initiate and catalyze further change within the cell, such as increased protein synthesis, directed cytoplasmic movements, or cell division.

What is remarkable is that the signaling machinery comes together as the process proceeds. The chemically and topologically complex surfaces of interacting proteins select the parts that form signaling complexes at the cell surface. If a different factor is encountered at the surface, some other combination of parts forms, tailored to pass on the requisite signal to the nucleus. Only one combination of parts is selected from the multitude of possibilities, enabling a subtle blending and ordering of cellular responses to an incredible range of external threats and opportunities. At each step in the assembly of these response complexes, protein surfaces must match up in a precise way; unused parts stand in readiness to react to future or concurrent contingencies.

The word "machinery " conjures up many images. After Newton formulated his Laws of Motion, it became fashionable to speak of the universe as a clockwork in which all that has happened, or whatever will happen, is contained in solutions to differential equations. This Laplacean world challenged the concept of free will and made determinism a rude guest in moral dialogues. As a result, Life was held to be separate from, and even in opposition to, the workings of the mechanized world. Living beings were seen as being built on a separate plan. Even as the nature of heat and its role in work-producing chemical reactions became clarified in the 19th Century, it was hard to avoid vitalistic notions of what made cells tick. Curiously, thermodynamics, the physics of heat, held up acceptance of Darwinism, because the Earth was thought to be cooling at too high a rate to allow enough time for life to have developed. Lord Kelvin's famous resistance to evolutionary ideas was gracefully dropped when he learned that radioactive processes in the Earth's core could maintain a constant surface temperature for eons (Burchfield, 1975).

All machines consume energy. Biological cells are open thermodynamic systems, meaning that they can exchange energy and matter across their boundaries. Much like the perennial waterfall beloved by philosophers of process, like Whitehead and Bergson, cells can appear unchanging even as their contents remain continually in flux. It was Willard Gibbs' great distinction to have discovered the laws of chemical equilibrium and the subtle interplay of order and thermal energy in chemical systems. The protein machines operating in biological cells are best analyzed in terms of their consumption of Gibbs free energy, generally in the form of adenosine triphosphophosphate (ATP). It is remarkable that hydrolysis of GTP by "G-proteins" is integral to the activation of internal response networks in response to external signals processed by many transmembrane signaling complexes. A cost must be paid to send a signal. This is a superb example of a "Polyani machine" acting as a gate-keeper or constituting a boundary condition, in this case across the physical surface separating outside from inside.

A Third Principle of structural biology emerges from this analysis, namely that Gibbs free energy must be spent in sorting cellular proteins to produce differences across boundaries. The life of a cell can be broken down into the "economic" process of separating components, say viral shell proteins and RNA, or the proteins involved in passing signals across a membrane, until they are later brought together by transport mechanisms to "self-assemble" into functional complexes. The economist Georgescu-Roegen, in a book of astonishing intellectual virtuosity (Georgescu-Roegen, 1971), tried with some success to unify economics with the physical and biological sciences. Central to his thinking was the concept of "sorting" and "recombining" material and energy resources as the main driving force in economic life. He quite clearly understood the importance of Gibbs free energy stored over hundreds of millions of years in petroleum and coal, and how humankind has used it to separate metals from ores, the core physical process driving economic development.

A PRIMER ON THE HISTORY OF PROTEIN STRUCTURE

When asked what a protein molecule looks like, most of us respond with a vigorous, gesticulating personification of the polypeptide chain coursing through space, describing coils and sheets, and generally folding back upon itself into a sweaty heap. The rigid appearance of helices and sheets gives a misleading picture of the forces stabilizing proteins and the extent of the motions these structures exhibit. For an isolated system at equilibrium, the Gibbs free energy is a minimum. This thermodynamic gemstone provides the currency for an economical description of biophysical processes. In particular, it enables a precise definition of "structural stability" and some clue as to why Nature needs twenty different kinds of amino acid side-chains.

The preoccupation with protein "folds" is the legacy of the victory of the polymer theory over the colloid theory in explaining the mystery of denaturation, or the ease with which proteins can be induced to give up their selective powers. Owing to the pioneering work of Emil Fischer, Jacques Loeb, Linus Pauling, Fred Sanger, Walter Kauzmann, Max Perutz, and others, the principles of protein structure are now mostly understood from the point-of-view of the chemist. Proteins are polymers with any one of 20 naturally occurring amino acids spaced regularly along a backbone spine. The great diversity of life on Earth is possible because even the class of modestly sized polypeptides having only100 amino acid side-chains contains 20 to the 100th power different protein sequences. One of the mysteries of life is why so few protein sequences are found in the biosphere compared with the astronomically large number of possibilities. Is it simply that only a few were selected for during evolution, or are the rules governing how polypeptide chains fold up into three-dimensional structures so restrictive that only a small number of sequences qualify?

How did the modern idea of a protein as a folded polypeptide chain having a specific sequence of amino acid side-chains develop? John Edsall (Edsall, 1962) has provided the perfect jumping off point in an essay that traces the vicissitudes of the idea that proteins are very large molecules. He points out that, even though proteins were crystallized in the last century and were determined to have large molecular weights, there was still a great reluctance to consider them as distinct molecules. Indeed, scientific philosophy continued to entertain serious doubts as to the ultimate reality of atoms (Pais, 1982), as held by the schools of Ernst Mach and Pierre Duhem (Duhem, 1914), much less of molecules.

Nineteenth century biological chemists considered the central mystery to be the process by which enzymes could lose their highly specific properties ("denaturation") by very gentle treatments. In Edsall's analysis (Edsall, 1962), the belief that proteins were colloids, self-aggregating masses of heterogeneous peptides, was the prevalent one because it led to plausible explanations for many of the properties of proteins. For example, the variegated surfaces could provide sites for catalysis and safe harbors for the rigid organic structures that carried out the "real" chemistry. Denaturation was simply explained as the colloid losing its grip on the active agent and reaggregating into a non-productive substance. Organic chemists were very reluctant to accept the idea that long chain-like polymers could provide the specificity exhibited by enzymes. Even the great Emil Fischer, who discovered the peptide bond and established the chirality of the natural amino acids, felt that short polymers no longer than twenty amino acids or so could *by combination* account for the diversity of biological molecules (Fruton, 1972).

Jacques Loeb's direct attack on the colloid school (Loeb, 1922), resembling in tone the fable of the "Emperor's New Clothes", established that proteins are polyelectrolytes. The wild variation in the colligative properties of protein solutions (e.g., osmotic pressure), exhibited such a strong sensitivity to metal ions and their salts that apparently explanations could only be provided in terms of colloidal behavior. Loeb showed that these effects were simply consequences of the inability of large charged proteins to cross semipermeable membranes, resulting in a redistribution of ions across membranes separating the proteins from bulk solvents. Furthermore, he showed that all of the phenomena, discussed in exquisite terminology long since abandoned, could be quantitatively accounted for when the pH was adequately controlled since it is the controlling variable for surface charge on the protein molecules. Ironically, it was a colloid chemist, The Svedberg who produced the clearest evidence that proteins were molecules with distinct molecular weights. The first photographic plates from a centrifuge run on a solution of hemoglobin, surprisingly, showed sharp boundaries for hemoglobin rather than the broad 'smear' that would have the distribution of smaller peptides making up the colloid.

To master structural biology requires familiarity with the history of its development—much like within the world of art—because the questions posed by the early pioneers still linger and color the language we use to describe proteins.

Yet, according to Edsall (Edsall, 1962), the majority of chemists doubted whether polymers could achieve the precision displayed by proteolytic enzymes in attacking their substrates. The question became, could a protein alone, without a prosthetic group, catalyze a reaction? The answer was provided by John Northrup whose systematic crystallization of enzymes demonstrated unambiguously that proteins are true molecules (Northrop et al., 1938) with distinct masses and properties. This proof rests on appreciating the Gibbs Phase Rule (Denbigh, 1971). Essentially, Northrup showed that a saturated protein solution was a true phase in equilibrium with the solid crystalline phase. As more protein is added, it cannot dissolve and must contribute mass to the growing crystals. If it does dissolve, then it cannot be pure, and another chemical species must be present. In this way, it was established that protein molecules alone, rather than bound impurities or groups, could be the sole source of the catalytic power found in protein solutions.

The first rule of protein structure was formulated by Linus Pauling, who predicted the existence of regular elements of structure (alpha helices, beta sheets, and gamma turns) from a consideration of the implications of the *planarity of the peptide bond*. Discovered by Fischer at the turn of the century, the peptide bond was shown in the 1930's to have a planar structure based on Bernal's determination of the structure of diketopiperazine. Planarity was explained in terms of quantum resonance between the amide and carbonyl groups contributed by each amino acid along the chain. Maximal overlap of electronic orbitals, related to bond strength, occurs when the NH and CO atoms lie in a plane. Pauling reasoned that planar peptide bonds stiffened the polypeptide backbone just enough to enable stable structures to form. Otherwise, if there were three flexible bonds per amino acid, rather than the two present when the peptide bond is fixed, hydrogen bonds between adjacent chains would be too weak to stabilize a distinctly folded polymer structure in the face of the universal tendency of matter to expand into the space of all possibilities (the 2nd Law of Thermodynamics), in this case of the polypeptide chain to take on all rotational angles about its main-chain bonds.

The key to Pauling's analysis was that each unit of the repeating backbone structure contains both a single hydrogen bond donor and an acceptor. Since the most stable structures are those for which the greatest number of bonds are formed, a repeating pattern stabilized by bonds formed between different units along the chain naturally provides an acceptor for each donor. Thus, two side-by-side polypeptide chains running in opposite directions can bond each other along their lengths as part of a beta-sheet. It is a remarkable fact that these regular structures are possible independently of the exact sequence and chemical nature of the amino acid side-chains, which project out and away from the hydrogen-bonded backbones. This indifference to specific sequence is what allows three-dimensional protein structures having quite different amino acid sequences to be classified according to backbone chain topology.

GIBBS FREE ENERGY: NATURE'S ACCOUNTANT OF 'PROCESS'

From the point-of-view of a physical chemist, the most remarkable aspect of biology is that energy conversions, principally the production of work (mechanical, osmotic, or electrical) from chemical bond breakage, take place under *isothermal* conditions. As argued below, there must be a constant interchange of chemical bond energy and entropy within molecules that can undergo order-disorder transitions. For example, the energy released when a bond forms between a solvated metal ion and a protein can bring about changes in the flexibility of some part of the polypeptide chain at the expense of changes in the ordered structure of the solvation sphere. Thus, changes in structure alter the heat capacity of a system, forcing it to radiate or absorb heat energy from the surroundings to maintain a fixed temperature.

Of course, there is a net exudation in the form of broken down molecules (waste) and randomized energy (heat) but, within the organism, steady-state processes take place (memory formation, muscle movements, nerve conduction) without local changes in temperature (Morowitz, 1970). There exists a theoretical apparatus of deceptive simplicity for analyzing and describing processes taking place under isothermal conditions. These tools are called irreversible thermodynamics (Katchalsky and Curran, 1965; Prigogine and Stengers, 1984) and the centerpiece is the Gibbs free energy. To venture into structural neurobiology or physiology without understanding the idea that life is an adaptation of free energy flow through macromolecules is to invite confusion and disappointment (See Box I).

Thermodynamics, the science of what drives the process of change, originated in a commonplace observation that was difficult to explain in terms of Newton's laws: *heat spontaneously flows from a hotter to a cooler object and not the other way* (Denbigh, 1971). The first crude steam engines were constructed from simple expanding cylinders. When cooled by small children throwing buckets of water at it, a collapsing cylinder could be used to lift water (via a lever) from flooded coal mines, providing access to desperately needed stores of chemical energy. The children, using water from the mines to cool the cylinders, to get the coal, to heat the cylinder, were the first of the "Maxwell Demons." Although a few deep thinkers, notably the economist Jevons, worried about the supplies of coal running out, most persons were happy just to keep warm (Georgescu-Roegen, 1971).

Obviously, force production and other physiological processes do not depend upon heat flows through temperature gradients as in simple "heat engines." Although muscle fibers do lengthen and contract, they do so without large changes in volume. Furthermore, muscle fibers utilize chemical energy from the hydrolysis of ATP, the universal biological fuel, to perform mechanical work at constant temperature. Rather than with volume or temperature changes, it is cyclic changes in the structure of muscle proteins that enable them to transduce chemical free energy into work.

Can a purely physical description of proteins capture our intuitive need to predict or quantify function? Thermodynamics is often said to be irrelevant for analyzing molecular mechanisms in detail since it involves relationships between bulk properties. But this sidesteps the point that Gibbs free energy provides a "plausibility index" that is useful in rooting out non-sensical or purely mechanical models for phenomena that manifestly require a non-equilibrium treatment. Maybe, some day, free energy changes involved in storing and retrieving information from genes and in sorting cellular constituents, will be added to a free energy "cost of living index" and the study of life's processes will have an economic flavor.

An extreme post-modern structuralist might now say, without fear of ridicule, that proteins do not 'fold', a pedestrian view, but rather 'condense' to the maximally dense set of side-chain clusters compatible with the sequence. As the condensation ensues, diffusion is not in the three-dimensional space of the solvent, but rather in the one-dimensional space along the polypeptide, which continually writhes upon itself while hydrophobic editing of side-chains takes place (see Box II). Protein factors, such as the chaperonins (Pelham, 1986) might work simply by holding onto the N- and C-terminal ends of the polypeptide while supplying enough free energy to shake the self-reptating chain enough to keep it out of local minima. This might explain why a major domain in chaperonins resemble actin molecules in having a strategically-placed hydrolyzable ATP molecule that potentially can mediate changes in protein conformation.

Thus, a protein molecule is a unique object for chemical study—a colloid on a string—a fixed length polymer having a specific sequence of surface-active chemical groups that can aggregate so as to find a minimum Gibbs free energy in an aqueous milieu. The unique colloidal collapse of the polymer places catalytic residues and recognition sites at distinct regions on the surface. The early chemists who struggled with the baffling questions posed by biological enzymes caught a glimpse of the truth with their notion

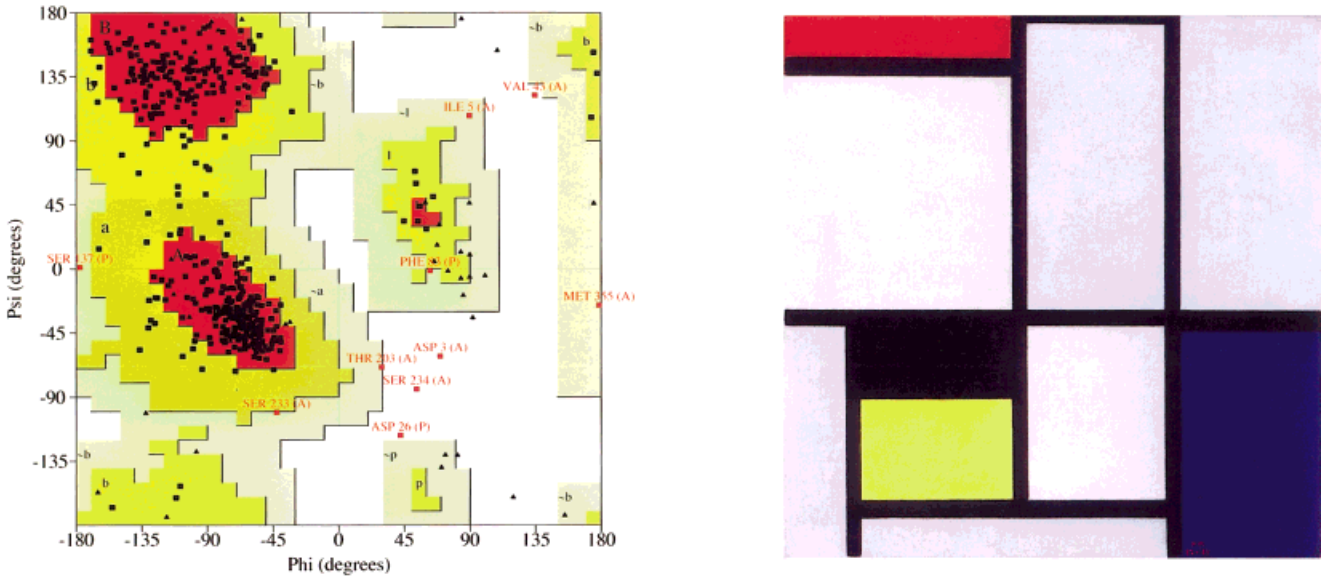

Figure 3. Tension resolution. The polypeptide backbones of proteins are semi-flexible. Restricted rotations, termed phi and psi, about the two covalent bonds flanking the planar peptide bond are possible. The course of the polypeptide chain can be completely specified by the set of phi/psi angles for each protein. When these are plotted as points on a two-dimensional "Ramachandran Diagram" (shown on the left, with the values plotted for profilin and actin), it is seen that they cluster predominantly in three regions (highlighted in red). These regions characterize alpha helices (lower left), beta sheets (upper left), and reverse turns (upper right), which are folds that allow the chain to turn without clashes between side chains and backbone atoms. The allowed regions in the Ramachandran diagram thus represent a resolution of the tension between the freedom of the backbone chains to take on all rotational possibilities and the restrictions that the volumes of the side chains place on allowed folds. Mondrian believed that the intersections of lines to form planes, and the coloration of these enclosed areas, allowed only a limited number of aesthetic possibilities. He created paintings which he though best resolved these internal tensions in composition. Shown is "Tableau No. 111; Composition No. 14; Composition with Red, Black, Yellow, Blue, and Gray (1921)" from *Piet Mondrian, 1872-1944* by Yve-Alain Bois, Joop Joosten, Angelica Zander Rudenstine and Hans Janssen. Boston: Bullfinch Press, 1994 ISBN 0-8212-2164-7. Permission to publish obtained from the Phillips Collection, Washington, D.C.

of achieving specificity through combination, but they considered polymerization to be too uncontrolled, yielding too heterogeneous a mixture, to provide dedicated single macromolecules. The existence of molecular mechanisms to control the ordering of amino acids in heteropolymers was apparently too speculative for any nineteenth century thinker to imagine (Edsall, 1962; Fruton, 1972). The "crystoidal" character of proteins, arising from definite sequences of amino acids, is an endowment only understandable when viewed as the product of a genetically encoded information system.

But an artistic tension, precisely in the sense meant by Langer, is revealed in the composition (Figure 3). We cannot look at a protein structure piecewise, either as bits of structure, or as resolved components of force. Water molecules, squeezed out in the process, invisible to the eye, are as much a part of what presents itself to the mind as the writhing polypeptide chain or the aperiodically packed side-chains in the structural core of the protein. And how is the composition held as a piece? Its Gibbs free energy is lower than some other arrangement of parts.

THE IMMUNE SYSTEM: "SELF-NON-SELF" AS A PROBLEM IN ARCHITECTONICS

Polyani's analysis of biological complexity leads us to examine the question of boundaries between different levels in a hierarchy. In adopting particular folds, protein molecules create unique surfaces out of which boundaries in the living hierarchy are established. One of the most fascinating questions is how an individual animal can distinguish its own proteins from those of foreign invaders. Where is the boundary between "self" and "non-self"? How can it be described in terms of the laws of chemistry? The function of the immune system is to detect and take countermeasures against threats to the health of the organism posed by viruses and bacteria. How do cells "decide" on the basis of molecular interactions to clear the body of an antigen encountered in the bloodstream?

The surfaces of most cells contain molecules of the cellular immune system, such as the major histocompatibility complexes (MHCs), which grip between "Jaws-like" alpha helices small lengths of polypeptide chain (Stern and Wiley, 1994; Stern et al., 1994; Strominger and Wiley, 1995). These peptides are sampled from the collection arising from the ceaseless "grinding up" of a large fraction (perhaps up to 30%) of all newly synthesized proteins into short peptides by maw-like structures called proteasomes. MHC molecules, once loaded

with peptide, are transported to the surface of the cell where they are presented to receptors on circulating T-cells. T-killer cells are commissioned to pass a life-or-death sentence on the peptide-presenting cells, depending upon the "meaning" carried by the presented peptide. The "boundary" between an individual animal and its pathogenic invaders is thus not one contiguous surface, but rather the set of all cellular surfaces presenting MHC complexes.

A key aspect of this "clonal selection theory of immunology" is that each T-cell recognizes only one MHC-peptide surface. T-cells bearing receptors that could potentially bind MHC-peptide complexes derived from normally occurring cellular proteins are selected against during a prenatal "learning" phase. Before birth, any T-cell that binds to an MHC-peptide complex, presumably presenting only "self" peptides, is killed. After birth, a developmental "switch" is thrown that causes the activation of any surviving T-cell line that successfully binds an MHC-peptide complex. Since there are no "self" recognizing T-cells in circulation after the switch is thrown, owing to clonal selection before birth, only "foreign" peptides, originating from infecting viral or bacterial proteins, are recognized in the context of an MHC surface. Notice the essential role played by molecular "sorting" in establishing the boundary between "self" and "non-self". Selection and transport of sampled peptides are central mechanisms governing the flow of information from the cell interior to its outer surface, a stunning example of Georgescu-Roegen's architectonics of economic process. The 19th Century chemists even got part of it right in recognizing that combinations of peptide fragments can generate the diversity needed to explain Life's contingent responses.

THE PROTEIN CLASSIFICATION PROBLEM

One of the great challenges to protein science, especially now that the human genome has become available, is to apply these general thermodynamic and structural principles to the problem of predicting three-dimensional structures of proteins. One aim of physical chemistry is to predict all states of organization of matter from a knowledge of the forces between molecules (represented as derivatives of potential energies, discovered by extrapolation from simpler systems). The achievement of these goals requires an analytical procedure for calculating the Gibbs free energy for various states of the polypeptide chain and the solvent molecules surrounding it. Stated this way the problem is unlikely to be solved in terms of the atomic interactions of all molecules involved (but see Kono et al., 2000). Why then is there a sense of excitement in the field of structure prediction, and a feeling that "proteomics" will be able to shoulder its load after the genome appears? The reasons are several.

The first is that a useful rule-of-thumb has gained wide currency and computer programs have been written to apply it (Richmond, 1984; Eisenberg and McLachlan, 1986). The idea is that, since the structure of water is extremely sensitive to the properties of surfaces with which it is in contact (water on wax), *changes in solvent accessible surface area* of aliphatic side-chains during protein-folding should be proportional to the decrease in Gibbs free energy accompanying the formation of the hydrophobic core (Chothia, 1984; See Box II). This simple relationship provides a means for estimating Gibbs free energy changes, without having to calculate the actual changes in the average positions and velocities of the solvent molecules in order to arrive at enthalpy and entropy values separately.

A second path-breaking development is the introduction of "tertiary filters," a set of computer programs that can be used to determine whether a given sequence of amino acids qualifies the folded protein as a member of some family of structurally-homologous proteins (Bowie et al., 1991). These programs go beyond mere comparisons of location of helices and sheets in the linear sequence of amino acids and incorporate the specific *context* of each side-chain. This represents an exciting step forward in applying thermodynamic principles (see Box II) to validating proposed structures. It brings closer the day when a gene sequence alone can be used to produce a structural model, provided that at least one representative of the structural class is known to atomic resolution. Atomic detail, particularly at domain/domain or subunit/subunit interfaces, will certainly require more structural data, but essential catalytic residues or recognition sites might be readily located in many instances. The main point is that substantial progress will be made in our ability to predict three-dimensional structures *by classification* well before the "protein folding problem" is solved. Whewell may have anticipated our point-of-view when he wrote "Classification is the architectonic science, to which Crystallography and the Doctrine of External Characters are subordinate (Whewell, 1857)."

Thus, the reductionist approach is running the risk of being bypassed by taxonomy, a source of biological truth even before Linnaeus. Furthermore, selected-site mutagenesis of proteins, combined with measurements of changes in their heat capacity as they fold is allowing a precise dissection of the free energy changes involved in the folding-condensation process (Matsumura et al., 1988; Fersht et al., 1992; Fersht, 1993) and during protein/protein association. Growth hormone binding to its receptor provides an especially apt example (Clackson et al., 1998; Bass et al., 1991). The idea is that the free energy of proteins in solution can be shifted by changes in pH or the addition of denaturants, such as urea and guanidine hydrochloride. The resulting changes in protein structure can be monitored spectroscopically as equilibrium is approached. When temperature is changed instead of chemical potentials, changes in heat capacity can be resolved into enthalpy and entropy changes.

Finally, our knowledge of the protein folding process is being put to the test by direct design of novel proteins, having amino acid sequences that have not been refined by natural selection (Hecht et al., 1990). A major conclusion from these studies is that it is

BOX II: "The hydrophobic effect:" looking at proteins with thermodynamic eyes

The concept of Gibbs free energy can be invoked to explain how the tradeoff between disorder, represented in the entropy of the backbone chain and the surrounding solvent molecules, and order, apparently conferred by hydrogen bonds, salt-bridges, and disulfide links is achieved. Since polar water molecules can interact with the hydrogen bond donors and acceptors along the backbone of the unfolded chain, there is no net gain in stability in forming helices and sheets via hydrogen bonds as the chain folds up. There must be something else that explains how the polypeptide chain eludes the dictates of the 2nd Law of Thermodynamics. That something else is to be found in the chemical differences amongst the 20 amino acid side-chains, which can be charged, neutral, wax-like, polar, or be able to form covalent disulfide bonds.

A folded protein molecule is a very improbable state for a polymer, representing only one set of torsional angles ($W_{chain} = 1$) out of a vast multitude. Substituting Boltzmann's formula into Gibbs equation at equilibrium ($\Delta G = 0$), allows a neat argument (Schulz and Schirmer, 1979) showing that hydrogen-bonding is only marginally effective at stabilizing rather long α-helices against the pull of entropy in a vacuum at a temperature of 300 kelvin ($\Delta HH-_{bonds} = T\Delta S_{chain}$). Unless the peptide bond is assumed planar, thereby freezing out one of the three torsional degrees of freedom along the chain per amino acid, the randomness of the disordered chain overwhelms any tendency to form stable helices or sheets; hydrogen bonds simply are not strong enough to overcome the entropy of the freely-jointed chain ($W_{chain} \gg 1$).

According to the quantum theory of radiation, the 'vacuum' of outer space contains radiation with associated degrees of freedom (Dyson, 1954) and the disordered energy ("heat") at any temperature is partitioned between the vacuum radiation and the material structures with which it is in equilibrium. At lower temperatures, helices could form in a vacuum because thermal radiation would predominate over chain entropy in the equitable distribution of heat. Of course, the ordering of protein chains does not take place in a vacuum and the folding process does not involve radiating disordered chain energy into the frozen reaches of space. Some other vehicle must transport heat ("disordered energy") away from the chain as it takes on a more ordered form (Figure 6).

Indeed, it was Walter Kauzmann's celebrated paper on "the hydrophobic effect" that ushered in the modern synthesis of the concepts of solution thermodynamics and polypeptide chain flexibility (Kauzmann, 1959). He showed that the most stable state of the polypeptide chain in solution arises when the Gibbs free energy of the complete system, *including the contributions arising from the entropy and changes in structure of the solvent water molecules*, is a minimum. The essential breakthrough was to recognize that the proper "system" to study was protein plus solvent. The chain adopts a "fold" that, given the constraints of the peptide bond, frees up the greatest number of water molecules to roam freely in solution rather than having to adopt more restricted 'iceberg-like' environments in the vicinity of aliphatic side-chains. Therefore, $WH_2O > O$ even though W_{chain} 1, with the result that $\Delta SH_2O +$ $_{chain} > O$ and the folding process is spontaneous.

The apparent stability of a soluble protein cannot be explained unless the entropic contribution of water molecules is included in Nature's free energy ledger. Structurally this manifests itself as a partial segregation of hydrophobic, or wax-like hydrocarbon side-chains, into the interior of the protein molecule where they cannot, by virtue of their surface activity, induce local order in solvent water molecules. Thus, the origins of protein stability lie, not necessarily in the strengths of internal chemical interactions, as in the structures of everyday life such as bridges, towers, and walls, but equally in the freedom of water molecules to take on a larger number of arrangements in the surrounding medium. This is the basis for the second rule of protein structure, the "*hydrophobic effect*", or entropy stabilization.

The dawn of the "crystallographically correct" synthesis of structure and thermodynamics is heralded by the work of Frederic Richards at Yale, who asks somewhat elliptically why Nature has selected a particular set of hydrophobic amino acids. After all, the most austere application of Kauzmann's principle only requires that a certain relative number of fatty side-chains be buried within the compact "folded" structure. Richards takes the analysis one step further by showing that for most proteins the number of packing defects or voids is rather small in the hydrophobic core (Lee and Richards, 1971). This implies that there might be some relatively restricted set of side-chain packing *combinations* that are to be found in natural protein structures (Ponder and Richards, 1987).

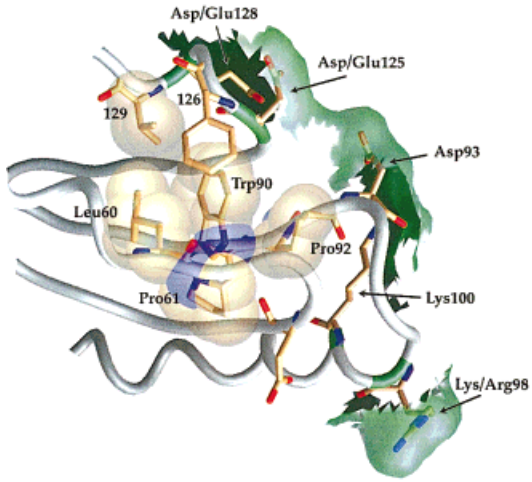

Figure 6. Close packing of hydrophobic sidechains in the interior of cofilin. Cofilin (also known as actin depolymerization factor—ADF) controls the dynamic turnover of actin subunits on one of the ends of the filament. Shown here are a set of conserved hydrophobic sidechains that form the core of ADF/cofilin. Thes sidechains are shaded light brown. Part of the ADF/cofilin polypeptide chain is shown as a silver wire. The requirement for a closely-packed core dictates the path of the polypeptide chain which must also obey the "law of the chain" to remain in allowed regions of the Ramachandran diagram (Fig. 3). The particular fold of ADF/cofilin results in the placement of key sidechains on the surface involved in actin binding (shown in green). Atoms are color-coded: carbon-brown, nitrogen-blue, and oxygen-red (Bowman et al., 2000).

difficult to design sequences that achieve the final degree of close-packing in the hydrophobic core characteristic of natural proteins (DeGrado et al., 1991). The protein engineers have yet to master the intricate tension between packing distinctively-shaped hydrophobic sidechains in the interior of the protein and bending of the polypeptide backbone within allowed limits.

THE FREE ENERGY BOUNDARY BETWEEN SUN AND EARTH

The apparently spontaneous appearance of order in the living world was for many centuries adduced as evidence for vitalistic forces, and later as support for the argument that living systems run counter to the 2nd law of thermodynamics. This is no longer an issue. To see why, it is useful to consider the planet Earth as an "open system (Schrödinger, 1946; Penrose, 1989). The earth radiates energy in the form of radiant heat in an amount approximately equal to what it absorbs from the sun. It is in an energetic steady state. However, there is a difference in entropy between the absorbed light, peaking in the ultraviolet region of the electromagnetic spectrum, and the heat radiated in the form of infrared photons. For a fixed amount of energy, each ultraviolet photon carts a bigger share of the total than the more numerous infrared photons streaming from the earth into the frigid void (Dyson, 1954). The number of ways (*W*: see Box I) of distributing a fixed amount of energy into packets is smaller for ultraviolet photons than for infrared ones. Hence the entropy of the incoming light is substantially less than the heat radiated by the earth. To a first approximation, the increase in the entropy of light itself drives all the living processes on earth without a net warming of the planet. This argument is valid even though other processes such as natural radioactivity in the Earth's crust and geothermal heating contribute to maintaining the earth's temperature.

The "boundary" between the sun and the biosphere is the set of all photosynthetic proteins in plants toiling away in the flux of radiation from the sun. The photoreaction center is the "Polyani machine" that extracts Gibbs free energy from ultraviolet light. The crystal structure of this protein machine is known and can be used to explain how the absorption of photons is converted into a flow of electrons used in the synthesis of ATP (Deisenhofer et al., 1985).

THE ORIGIN OF CELLULAR TENSION

Viruses typify the value of the factory metaphor for understanding cellular economics; parts are gathered in staging areas and then brought together for assembly processes driven by ordinary crystallization. The work is done in the sorting (see Box III). However, something radically new happens when protein subunits can carry free energy with them into polymerizing structures. Tubulin-GTP monomers form microtubules. Actin-ATP monomers form microfilaments. Both of these filament systems can generate forces if these subunits can change shape in converting NTP to NDP, where N can be adenosine (A) or guanine (G). These forces result from a conversion of Gibbs free energy from nucleotide hydrolysis into length changes along some direction (say along x) since f = -dG/dx. The structural basis for length changes in actin is the principle of subdomain rotations about hinge and shear points in the actin molecule (Page et al, 1998) following the structural paradigm of viral capsid proteins (see Box III). These forces can be rather large compared with Brownian forces, 100 piconewtons in the case of actin (Schutt and Lindberg, 1998). It is curious that microfilaments and microtubules are considered by most experts in the field of "motor proteins" to be merely "tracks" along which transporting motors such as myosin and kinesin move. Since microtubules display highly unusual dynamic motions in GTP-containing solutions (Mitchison and Kirschner, 1984; Desai and Mitchison, 1997) it is natural to wonder what role they could play in more "engine-like" mechanisms, where kinesin and tubulin work cooperatively to produce forces.

The paradigm for chemomechanical process in biology is the "sliding filament model of muscle contraction," in which cross-bridges projecting out of the myosin thick filaments bind to actin thin filaments and pull them towards the center of the sarcomeres, the basic units of contraction in muscle fibers (See Figure 4). Actin is generally thought to be an inert rod-like element in this process. The myosin cross-bridges bind ATP as they detach from actin and hydrolyze it in the unattached state. Upon rebinding actin, the myosin head 'rotates' through several binding sites on actin of successively lower energy while stretching a molecular "spring" that then pulls on the actin filament (Huxley and Simmons, 1971). In this manner, converting bond energy into elastic energy, it is believed that the free energy of ATP hydrolysis is transduced into work. Myosin is often called a "motor molecule," because the macroscopic forces generated by muscle fibers could be explained as the summed effect of hundreds of myosin heads independently pulling on each actin filament. That situation has changed very recently. New measurements on the extensibility of actin filaments, and reconsideration of the thermodynamics of muscle, (Baker et al, 1999) have cast doubts on the validity of the conventional cross-bridge theory of contraction. Attention is being increasingly focused on models that take into account the overall spatial and temporal organization in muscle lattices, and the possibility of cooperativity amongst the myosin motors.

We have analyzed the energetics of muscle contraction in terms of a new, quite different, model (Schutt and Lindberg, 1992; 1993; 1998; Kreatsoulas et al., 1999) in which actin is not a passive rope-like element but rather an active force generator. Actin not only possesses a three-dimensional structure similar to other ATPases, such as hexokinase, but actually hydrolyses ATP during the assembly of actin filaments. In our model, actin itself hydrolyzes ATP successively subunit by subunit along actin filaments, each triggering phosphate release into solution by its neighbor, as

BOX III: An example of a structural paradigm: self-assembly of simple viruses

Explaining the existence of symmetric, highly ordered structures in biology (simple viruses, microtubules, muscle fibers) typifies the difficulty of applying the laws of thermodynamics to the living process. After all, it would seem that crystalline order, more characteristic of the inorganic chemistry of inert materials, must represent an unlikely occurrence in the interior of cells, where Brownian motion, the incessant random movement of all bits of matter at ordinary temperatures, is the dominant phenomenon. There are two structural paradigms, tobacco mosaic virus (TMV) and tomato bushy stunt virus (TBSV), two unlikely-sounding characters for the center stage, that exemplify self-assembling structures. TMV (Bloomer et al., 1978; Namba and Stubbs, 1986) and TBSV (Harrison et al., 1978) have been solved at atomic resolution by x-ray diffraction, and there are many lessons to be learned from these simple isometric viruses (Harrison, 1991; Perutz, 1992).

Bernal deduced from the x-ray diffraction patterns that viruses had multiple copies of single protein subunits arranged according to some rule of regularity, not unlike those found in crystals. The analogy with crystal growth eventually led to the concept of self-assembly; i.e., the ability of a large structure to spontaneously assemble from its components, without the need for templates upon which to grow, or enzymes to direct bond formation. The subunits aggregate by non-covalent interactions (hydrophobic, for example) by which the system minimizes its Gibbs free energy (Janin and Chothia, 1990). Thus, in the case of simple viruses, the design for the whole particle is specified by the direction and strength of 'bonds' between subunits; *the design of the whole is contained within the parts*, an idea advocated by Buckminster Fuller for the blueprint-free assembly of geodesic domes (Caspar and Klug, 1962).

A system composed of identical subunits can most easily minimize its Gibbs free energy by forming symmetric structures. This can be understood by considering the simple case of linear polymerization where each subunit associates end-to-end with two others. If the directions of the 'bonds' between these subunits are changed, the linear structure can spiral outwards or inwards depending on the bond angles. For a particular set of angles, the subunits will fold into a helix generating another class of stabilizing contacts between turns; i.e., each subunit now has four or six neighbors in the helix. Any distortion in the bond angles that destroys the equivalence of the subunits, that is the helical symmetry, necessarily leads to a less stable particle since the solvent accessible area increases. Therefore, greater particle stability is a direct consequence of having a subunit capable of forming symmetric structures. This argument can be extended to the case of closed isometric (or quasi-spherical) shells through the application of icosahedral point group symmetry.

This is an example of Polyani's notion that events at lower levels in a hierarchical system, in this case the self-association of proteins, entail functional "meaning" at higher levels that cannot be deduced from full knowledge of the lower level. There are a number of functional advantages in having a structure constructed from identical subunits (Caspar and Klug, 1962): (1) genetic efficiency - a particle of the size and composition of TMV could never have its protein shell coded for by the TMV-RNA without some element of repetition in the capsid, (2) error editing - an improperly folded subunit, say one with a misincorporated amino acid, not tailored to fit into the niche presented by the growing particle , would be rejected, and (3) self-assembly - no further protein scaffolds or enzymatic steps are required.

The simple picture of single subunits fitting into the growing helical ramp has been supplanted by a scheme involving a disk-like assembly intermediate (Butler and Klug, 1978). TMV protein subunits first assemble into disks which have the capacity to nucleate assembly by binding to a specific sequence on TMV-RNA, providing selectivity, while at the same time overcoming the entropic barrier in getting the first few subunits to form a helical nucleus (Namba and Stubbs, 1986). Lowering the pH of a solution of TMV subunits in the absence of RNA results in the assembly of empty helical shells. Caspar (1963) observed that a pair of carboxyl groups exhibited abnormal pKs (near 7.0) in the intact virus, but behaved normally for the disks (near 4.5). Since the carboxyl groups are charged at neutral pH in the disk, he postulated that a 'negative electrostatic switch' prevents "lockwashering" into helices. However, when RNA is added to the protein solution, the binding energy of RNA to the disk overcomes the electrostatic repulsion, forcing the carboxyl groups into environments in the helical virus that resemble the abnormally-titrating carboxylate groups in maleic acid (Namba and Stubbs, 1986). It is clear that the TMV assembly process in solution can be fully understood in terms of the system flowing down gradients of Gibbs free energy. The free energy involved in melting stems and loops of the free RNA as it becomes packaged in the virus includes the role of water, bound and free, in the process. The subunits in the disk assembly are less tightly-packed compared with the helix and have more flexible side-chains exposed to solvent. They are entropy stabilized. When RNA becomes encapsulated, the formation of stronger, better-packed, inter-subunit bonds in the (more symmetrical) helical state lowers the free energy of the particle. *In vivo*, the orchestrated process of separating proteins and nucleic acids into different cellular compartments (via vesicular transport mechanisms) raises the free energy locally, setting the stage for a return to equilibrium that appears as "self-assembly". The driving force for the process, really just a flow down a chemical gradient, is set up with the original work expended in segregating the viral components.

TMV is the paradigm for helically-symmetric particle assembly. Tomato bushy stunt virus (TBSV) is the paradigm for spherical viruses, which are based on icosahedral point group symmetry. Formally, each subunit can be in an identical environment in an icosahedrally-symmetric shell only if it contains exactly 60 subunits, three arranged about each of the 20 three-fold axes. The striking aspect of TBSV is that it contains 180 protein subunits, all with the same amino acid sequences, and therefore *the subunits in the TBSV shell cannot all have identical three-dimensional structures* (Caspar and Klug, 1962) and must fall into three different 'quasi-equivalent' classes. These deviations from strict symmetry have important biological implications. The three classes of subunit structure found in the intact TBSV capsid derive from two features of the protein (Harrison, 1991). First, the N-terminal 66 residues are highly charged positively and are involved in binding RNA, while the next 35 residues form "arms" *that have different degrees of packing order in the shell* depending on the local environment. In two of the three "quasi-equivalent" subunits, the arms are disordered. However, every third subunit has an ordered arm that, with its 60 symmetry-related partners, knits the whole particle together. Secondly, there is a hinge between the two major domains of the TBSV subunits that allows them to take up different positions in the viral capsid without relative slippage in the subunit-subunit bonds. This design feature, maintained contacts between variably-linked domains, confers flexibility to the viral shell under conditions mimicking those in the host cell.

The expansion of the shell is thought to be important during the assembly and unpackaging of the virus (Harrison, 1991). The design of TBSV, like TMV, incorporates controlled polypeptide flexibility in two different ways, variable terminal arms and internal hinges, that enable proteins to encapsulate RNA as they move down gradients of Gibbs free energy. The 2nd law of thermodynamics is not violated when all factors are taken into account. These include the work required to keep the components separated before assembly, the heat lost during the initial ATP hydrolysis involved in establishing gradients, and the change of entropy of the system (virus+solvent) as the particle anneals. More complicated viruses, such as poliovirus, use proteolytic cleavage to "trigger" transitions in viral capsids during the assembly process, and other elaborate strategies, but are understandable "mechanisms" in the context of free energy flow.

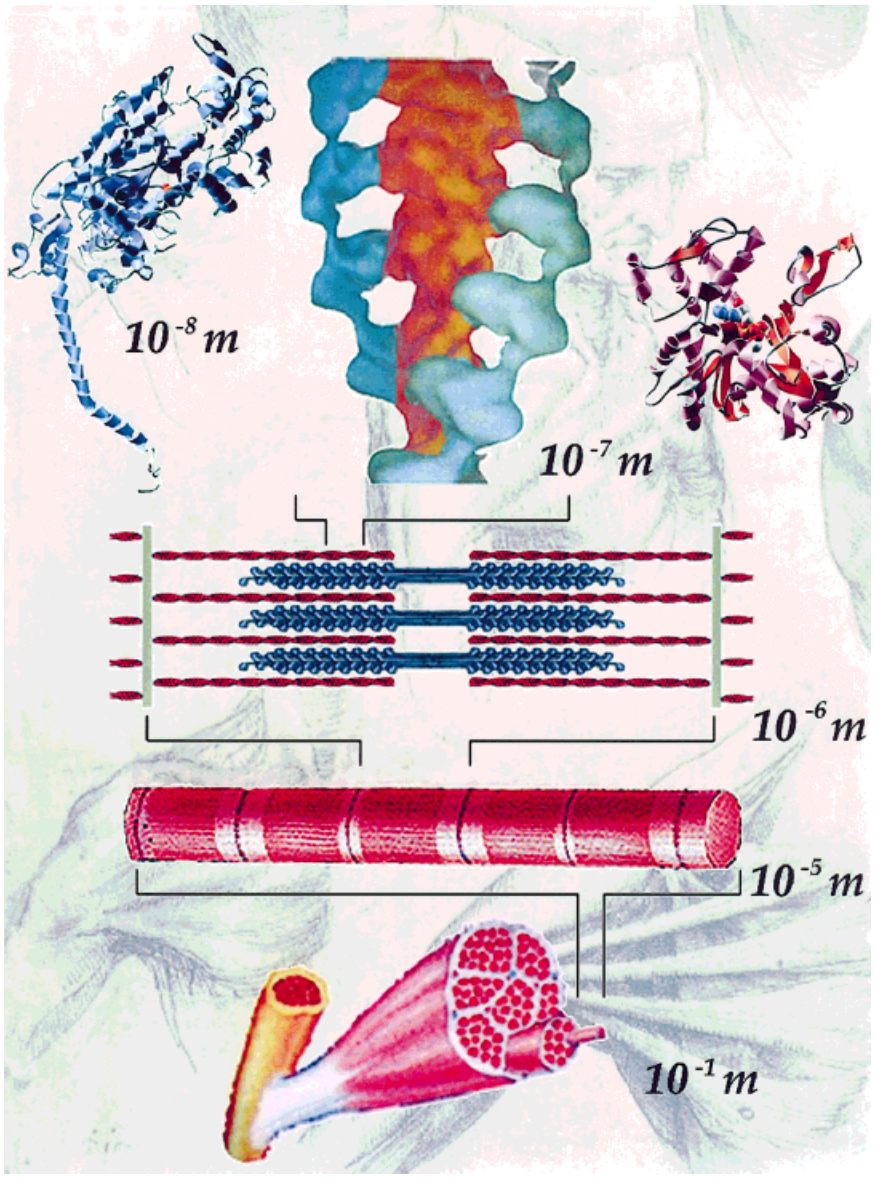

Figure 4. Muscle as a hierarchical system. Muscle cells contain hierarchical arrays of fibers. The basic functional unit of muscle is the sarcomere, which consists of interdigitating arrays of actin and myosin filaments. When a fiber contracts, each of the thousands of sarcomeres arranged end-to-end contract in unison. In this way, the ten micron per second sarcomeric contractions are amplified into meter per second shortenings of the whole fiber. At the molecular level, the actin filaments move along the myosin filaments, powered by the hydrolysis of ATP. The background image of Leonardo Da Vinci reminds us of mankind's centuries-long search for the secret of animal locomotion. Figure originally appeared in Kreatsoulas et al. (1999).

it moves relative to myosin. The role of the well-known actin-activated myosin ATPase is to catalyze the uptake of ATP on actin and to provide traction points for the crawling actin filaments. This suggests a new "gatekeeping" role for ATPases in controlling the kinetics of biophysical processes, "Polyani machines" operating at boundaries in the time domain. The implications of this new model for studying actin-based movements are many, but one idea is that *actin subunits generate forces that can be used to drive enzymes through their catalytic cycles*. In the standard theories of enzyme kinetics, it is assumed that binding to the substrate helps to lift the free energy of the enzyme into an activated state and that random Brownian motion 'kicks the protein' with sufficient force to bring about the required transitions. Perhaps the free energy of activation (a form of work) can be supplied by attached actin filaments undergoing force-producing length changes. In effect, the actin-rich cell cortex is a "tension bath" analogous to a thermodynamic "heat bath," but capable of directed vectorial action that enables the enzyme to not only respond more quickly, but to be synchronized to other enzymes elsewhere in the cell and drawing free energy from the same source.

The actin-rich cell cortex is a "tension bath" analogous to a thermodynamic "heat bath," but capable of directed vectoral action that enables the enzyme to not only respond more quickly, but to be synchronized to other enzymes elsewhere in the cell and drawing free energy from the same source.

Eukaryotic cells are notable for the broad range of motile activities in which they can engage over the course of their lifetimes. Cell division is certainly the most balletic but on the local scale membrane ruffling and microspike movements probably qualify as the most violent. Neurons and immune cells are especially rich in motile activity (Smith, 1988; Davenport et al., 1993; Matus, 1999). Neurotransmitters, synthesized in the cell soma of motor neurons, are transported to

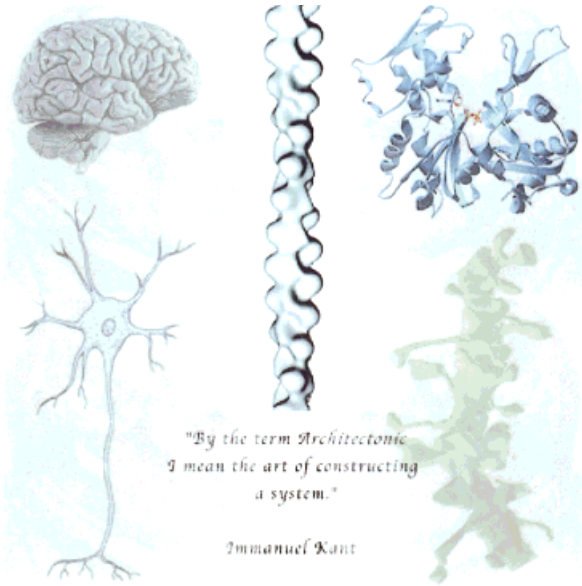

Figure 5. Mind as a problem in architectonics. Neurons (shown on the bottom left) connect up with each other along their dendrites (shown enlarged on bottom right). The small, micron-sized, variously-shaped projections from the dendrites are called spines. Spines are the sites of contact (synapses) with the axons of other neurons. Synaptic plasticty or changes in synaptic strength are believed to form the basis for learning and memory. At the molecular level, these changes in spine shape are mediated by the polymerization of actin, a universal "architectonic element" (the backbone representation of an actin subunit is shown on the upper right). Actin forms helical filaments (shown at low resolution in the center. Immanuel Kant revolutionized philosophy with his views on how the structure of our minds filters the stream of sensory information in categorizing external reality (his multiple images form a tableau for offering a structural biological view of the mind's hierarchy).

neuromuscular synapses along the microtubules in axons. The vast majority of all excitatory synapses in cortical neurons are located at the tips of actin-rich structures called "dendritic spines" which are densely distributed along the dendritic arbors of pyramidal cells (Harris and Kater, 1994) (See Figure 5, lower right). Spines, which are 1-2 microns in length, change shape on a second by second basis as they adjust to change the strength of their synaptic junctions with neighboring neurons (Fischer et al., 1998). This process lies at the heart of learning and memory as first outlined by E. O. Hebb in 1951. The precise mechanisms underlying "synaptic strength" have not been established, although polymerization and cross-linking of actin filaments are certainly factors, and many neurotransmitter receptors are physically linked to actin via specialized proteins, such as gephyrin, in the case of GABA receptors (Lindberg and Schutt, 1999).

Actin-binding proteins, such as members of the formin and dystrophin families, link membrane-bound enzymes and channels with the actin microfilament system (Lindberg and Schutt, 1999). Actin's role may extend beyond merely providing basement support. Microfilament tension transmitted directly to ion channels in the plasma membrane could affect the magnitude of the electrochemical resting voltages that govern the propagation of action potentials in the nervous system. Actin binding proteins are important in mediating vesicle transport of neurotransmitters in pre-synaptic termini. Is it really too fanciful to wonder how mechanical, membrane electrical potential, and chemical neurotransmission, might mutually shift the overall free energy balance sheet in distributed networks of neurons? Hebbian memory must after all be Gibbsian!

CONCLUSION

Structural Biology is a system of thought. It is not a simple system. It is an architectonic science because it is concerned with how living systems are built up. To master it requires familiarity with the history of its development—much like within the world of art—because the questions posed by the early pioneers still linger and color the language we use to describe proteins. Statistical thermodynamics is essential in order to understand how structure can be sustained in a micro-world where Brownian fluctuations are the dominant processes. Ultimately, if we are to understand how thoughts and beliefs can be fixed in our minds, we must look to proteins as the material substrates for mental processes and devise theories that recognize their special character as hierarchical structures under tension. Change is possible because we live in a flux of Gibbs free energy, generated in the sun, captured by plants, and parceled out as fuel to drive the sorting and motile activities of cells.

Architectonics began with Kant (1787) and his preoccupation with the problem of how our minds devise theories about the external world from sense-given information; the structure of our minds limits what we can know of the world (Figure 5). In Kant's words: "By the term *Architectonic* I mean the art of constructing a system. Without systematic unity, our knowledge cannot become science; it will be an aggregate, and not a system. Thus Architectonic is the doctrine of the scientific in cognition, and therefore necessarily forms part of our methodology."

Susanne Langer's great mentor, Ernst Cassirer, carried Kant's program forward by seeking the universal metaphorical structures behind language, arts, and mathematics. Although he could never anticipate what kinds of physical structures in the brain could filter sensory data into categories, Cassirer (1957) may have captured the essence of modern thinking about how neurons construct the mind. He wrote in concluding his magnum opus: "We have so far tried to show how the individual symbolic

forms—language, myth, theoretical knowledge—are aspects in the structure of the intelligent organization of reality. Each of them presented us with an independent, architectonic principle, an ideal "structure", or, better, — since we are here never dealing with describing purely static relationships, but rather with exposing dynamic processes—a characteristic way of "structuring" itself. . .". Cassirer's language suits our present purposes because we are concerned with "structuring" itself, but as it occurs at the surfaces of communicating cells. Applied to neurons, the dynamic changes at the tips of dendritic spines are brought about by changes in the architecture of the underlying actin filaments.

There is a strain of modern philosophy that continues the tradition of Cassirer. It is the "hermeneutical school" whose central premise is that our minds build worlds out of historically learned knowledge and fresh sensory experiences. Rather than the fixed categories of cognition imagined by Kant, these philosophers (Gadamer, 1976; Goodman, 1978) posit an infinite plenitude of frames of reference depending on the conversations in which one engages. To understand Structural Biology requires a close study of the theoretical problems that beset its founders, a kind of dialogue with the past reflected in terms like "denaturation", "folding", "self-assembly", "lock-and-key", and "sorting."

Perhaps the highest level of the functional hierarchy is the brain itself and the place where theory-building takes place. The modern cellular theory holds that the combinatorial complexity inherent in networks of billions of neurons connected by thousands of synapses is enough to sort the received sensory streams from the outside world into learned and remembered categories. Whether one believes that the synapses are changeable nodes in a network that encodes concepts (Churchland, 1984), or link up neurons into groups, as in Edelman's theory of neuronal selection (Edelman, 1987), the boundary between our outer and inner worlds is the set of all synaptic surfaces (see Figure 5). The boundary is ever-changing as we encounter new experiences, sift through our memories, form new concepts. In the architectonic view, our minds are written on this surface by microscopic dendritic spines, small actin-rich "Polyani machines" reaching out to connect with axons bearing gifts of precious neurotransmitters, signaling changes in the external world. Actin is thus a prototypical architectonic element: structuring and shaping the synaptic surfaces between communicating neurons, building new worlds from our present experiences and past memories.

ACKNOWLEDGMENTS

This work was supported by grants from The Nancy Lurie Marks Family Foundation (to CES), the National Institutes of Health (GM44038 to CES), the Swedish Cancer Foundation (to UL), and the Swedish Natural Science Research Council. Figure Credits: Figures 1 and 2 were prepared by Constantine Kreatsoulas and originally appeared in Kreatsoulas et al. (1999). Figures 3,4,6 were prepared by Gregory D. Bowman. Figure 5 was produced by Alex Plakantonakis.

LITERATURE CITED

Anderson JE, Ptashne M, Harrison SC. 1987. Structure of the repressor-operator complex of bacteriophage 434. Nature 326:846–852.

Ban N, Nissen P, Hansen J, Moore PB, Steitz TA. 2000. The complete structure of the large ribosomal subunit at 2.4 Angstroms resolution. Science 289(5841):905–920.

Bass SH, Mulkerrin MG, Wells JA. 1991. A systematic mutational analysis of hormone-binding determinants in the human growth hormone receptor. Proc Natl Acad Sci U S A. May 15; 88:4498-4502.

Bloomer AC, Champness JN, Bricogne G, Staden R, and Klug A. 1978. Protein disk of tobacco mosaic virus at 2.8 Å resolution showing interactions within and between subunits. Nature 276:362-368.

Bowie JU, Luthy R, Eisenberg D. 1991. A method to identify proteins that fold into a known 3-dimensional structure. Science 253:164-170.

Bowman GD, Nodelman IM, Hong Y, Chua N-H, Lindberg U, Schutt CE. 2000. A comparative structural analysis of the ADF/cofilin family. Proteins 41:374–384.

Burchfield JD. 1975. Lord Kelvin and the Age of the Earth. Univ. of Chicago Press. Chicago, IL.

Butler PJG, Klug A. 1978. The assembly of a virus. Sci Am 239:62-69.

Caspar DLD. 1963. Assembly and stability of the tobacco mosaic virus particle. Adv in Protein Chem 18:37-121.

Caspar DLD, Klug A. 1962. Physical principles in the construction of regular viruses. Cold Spring Harbor Symp. Quant Biol 27:1-27.

Cassirer E. 1957. The Philosophy of Symbolic Forms. Yale University Press. New Haven. Volume 4. 50p.

Chothia C. 1984. Principles that determine the structure of proteins. Annu Rev Biochem 53:537-572.

Churchland PM. 1984. Matter and Consciousness. MIT Press. Cambridge, MA.

Clackson T, Ultsch MH, Wells JA, de Vos AM. 1998. Structural and functional analysis of the 1:1 growth hormone: receptor complex reveals the molecular basis for receptor affinity. J Mol Biol. 277:1111-28.

Davenport RW, Dou P, Rehder V, Kater SB. 1993. A sensory role for neuronal growth cone filopodia. Nature 361:721-724.

Davies DR, Goryshin IY, Reznikoff, Rayment I. 2000. Three-dimensional structure of the Tn5 synaptic complex transposition intermediate. Science 289:77–85.

DeGrado WF, Raleigh DP, Handel T. 1991. *De novo* protein design: what are we learning? Curr Opin Struct Biol 1:984-993.

Denbigh K. 1971. The principles of chemical equilibrium. Cambridge University Press, Cambridge.

Deisenhofer J, Epp O, Miki K, Huber R, Michel H. 1984. The structure of the photoreaction center from rhodopseudomonas viridis. Nature 318:618-624.

Desai A Mitchison TJ. 1997. Microtubule polymerization dynamics. Annu Rev Cell Dev Biol 13:83-117. Review.

Dickerson RE, Drew HR. 1981. Kinematic model for B-DNA. Proc Natl Acad Sci USA 78:7318–7322.

Duhem P. 1914. The aim and structure of physical theory. Atheneum (Second edition reprinted by Princeton University Press, 1954).

Dyson FJ. 1954. What is heat? Sci Am 58-63.

Edelman GM. 1987. Neural Darwinism. Basic Books. New York.

Edelman GM. 1988. Topobiology: an introduction to molecular embryology. Basic Books. New York.

Edsall JT. 1962. Proteins as macromolecules: an essay on the development of the macromolecular concept and some of its vicissitudes. Arch Biochem, Suppl. 1:12-20.

Eisenberg D, McLachlan AD. 1986. Solvation energy in protein folding and binding. Nature 319:199-203.

Fersht AR. 1993. Protein folding and stability: the pathway of folding of barnase. FEBS Lett 325:1, 2, 5-16.

Fersht AR, Matouschek A, Serrano L. 1992. The folding of an enzyme . 1. Theory of protein engineering analysis of stability and pathway of protein folding. J Mol Biol 224:771-782.

Fischer M, Kaech S, Knutti D, Matus A. 1998. Rapid actin-based plasticity in dendritic spines. Neuron 20:847-854.

Fruton JS. 1972. Molecules and life: historical essays on the interplay of chemistry and biology. Wiley-Interscience, New York.

Gadamer HG. 1976. Philosophical Hermeneutics. Univ. of California Press. Berkeley, CA.

Georgescu-Roegen N. 1971. The entropy law and the economic process. Harvard University Press, Cambridge, MA.

Gibbs JW. 1906. The Scientific Papers of J. Willard Gibbs (Longmans, Green C.; reprinted as Dover paperback, New York, 1961).

Goodman N. 1978. Ways of Worldmaking. Hacket Publishing Company. Indianapolis, IN.

Harris KM, Kater SB. 1994. Dendritic spines: cellular specializations imparting both stability and flexibility to synaptic function. Annu Rev Neurosci 17:341-371.

Harrison SC. 1991. What do viruses look like? Harvey Lectures 87:127-152.

Harrison SC, Olson AJ, Schutt CE, Winkler FK, Bricogne G. 1978. Tomato bushy stunt virus at 2.9 Å resolution. Nature 276:368-373.

Hecht MH, Richardson JS, Richardson DC, Ogden RC. 1990. *De novo* design, expression, and characterization of Felix: a four-helix bundle protein of native-like sequence. Science 249:884-891.

Hill TL. 1960. An introduction to statistical thermodynamics. Addison-Wesley, Reading, Mass.

Huxley AF, Simmons R. 1971. Proposed mechanism of force generation in striated muscle. Nature 233:533-538.

Janin J, Chothia C. 1990. The structure of protein-protein recognition sites. J Biol Chem 265:16027-16030.

Kant I. 1787. Critique of Pure Reason. Everyman's Library. J.M. Dent & Sons Ltd London (reprinted in 1964, 471p).

Katchalsky A, Curran PA. 1965. Nonequilibrium thermodynamics in biophysics. Harvard University Press, Cambridge, Mass.

Kauzmann W. 1959. Some factors in the interpretation of protein denaturation. Protein Chem 14:1-63.

Kirschner M, Mitchison T. 1986. Beyond self-assembly: from microtubules to morphogenesis. Cell 45:329-342

Klotz IM. 1967. Energy changes in biochemical reactions. Academic Press Inc., New York.

Kono H, Sarai A. 1999. Structure-based prediction of DNA target sites by regulatory proteins. Proteins 35:114–131.

Kono H, Saito M, Sarai A. 2000. Stability analysis of the cavity-filling mutations of the Myb DNA-binding domain utilizing free energy calculations. Proteins 38:197–209.

Kreatsoulas C, Lindberg U, Schutt CE. 1999. "Nouveau Regard sur le Muscle". *Les Cahiers de Science & Vie*. 53:78-87.

Langer SK. 1967. Mind: an essay on feeling. Volume 1. p. 93 and 155.

Lee B, Richards FM. 1971. The interpretation of protein structures: estimation of static accessibility. J Mol Biol 55:379-400.

Lindberg U, Schutt CE. 1999. "Actin-binding Proteins". Encyclopedia of Molecular Biology. (T. Creighton, Ed). John Wiley & Sons, Inc. p.32-48 .

Loeb J. 1922. Proteins and the Theory of Colloidal Behavior. McGraw-Hill, New York.

Lu X-J, Olson W. 1999. Resolving the discrepancies among nucleic acid conformational analyses. J Mol Biol 285:1563–1575.

Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature 389:251-260

Matsumura M, Becktel WJ, Matthews BW. 1988. Hydrophobic stabilization in T4 lysozyme determined directly by multiple substitutions of ile 3. Nature 334: 406-410.

Matus A. 1999. Postsynaptic actin and neuronal plasticity. Curr Opin Neurobiol 9:561-565.

Mitchison T, Kirschner M. 1984 Dynamic instability of microtubule growth. Nature 312:237-42.

Morowitz HJ. 1970. Entropy for biologists: an introduction to thermodynamics. Academic Press, New York.

Namba K, Stubbs G. 1986. Structure of tobacco mosaic virus at 3.6 Å resolution: implications for assembly. Science 231: 1401-1406.

Northrup JH, Kunitz M, Herriott RM. 1938. Crystalline Enzymes. Columbia University Press, New York.

Pabo CO, Nekludova L. 2000. Geometric analysis and comparison of protein-DNA interfaces: why is there no simple code for recognition? J Mol Biol 301:597–624.

Page R, Lindberg U, Schutt CE. 1998. Domain motions in actin. J Mol Biol 280: 463-474.

Pais A. 1982. "Subtle is the Lord": the science and the life of Albert Einstein. Oxford University Press, Oxford.

Pauling L, Corey RB, Branson HR. 1951. The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain. Proc Natl Acad Sci USA 37:205-211.

Pelham HRB. 1986. Speculation on the functions of the major heat shock proteins and glucose-regulated proteins. Cell 46:959-961.

Perutz MF. 1992. Protein structure: new approaches to disease and therapy. W. H. Freeman & Co., New York.

Polyani M. 1968. Life's Irreducible Structure. Chapt. 14 Knowing and Being (1969): essays by Michael Polyani. Grene, M. ed. Routledge & Kegan Paul Ltd. London (1969).

Ponder JW, Richards FM. 1987. Tertiary templates for proteins: use of packing criteria in the enumeration of allowed sequences for different structural classes. J Mol Biol 193:775-791.

Prigogine I, Stengers I. 1984. Order out of chaos: man's new dialogue with nature. Bantam, New York.

Rasmussen BF, Stock AM, Ringe D, Petsko G. 1992. Crystalline ribonuclease A loses function below the dynamical transition at 220 k. Nature 357:423-424.

Richmond TJ. 1984. Solvent accessible surface area and excluded volume in proteins: analytical equations for overlapping spheres and implications for the hydrophobic effect. J Mol Biol 178:63-89.

Schrödinger E. 1946. What is life? Reprinted by Cambridge University Press, Cambridge 1967.

Schulz GE, Schirmer RH. 1979. Principles of Protein Structure. Springer-Verlag (New York Berlin Heidelberg Tokyo).

Schutt CE, Lindberg U. 1991. Scratching the back of the cortex. Curr Biol 4:257-258.

Schutt CE, Lindberg U. 1992. Actin as the generator of tension during muscle contraction. Proc Natl Acad Sci USA 89:319-323.

Schutt CE, Lindberg U. 1993. A new perspective on muscle contraction. FEBS Lett 325:59-62.

Schutt CE, Lindberg U. 1998. Muscle contraction as a Markov process: energetics of the process. Acta Physiol Scand 163: 307-323.

Sigler PB, Xu Z, Rye HS, Burston SG, Fenton WA, Horwich AL. 1998. Structure and function in GroEL-mediated protein folding. Annu Rev Biochem 67:581–608.

Smith SJ. 1988. Neuronal cytomechanics: the actin-based motility of growth cones. Science 242:708-715.

Stern LJ, Wiley DC. 1994. Antigenic peptide binding by class I and class II histocompatibility proteins. Structure 2:245-251.

Stern LJ, Brown JH, Jardetzky TS, Gorga JC, Urban RG, Strominger JL, Wiley DC. 1994. Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. Nature 368:215-221.

Strominger JL, Wiley DC. 1995. The class I and class II proteins of the human major histocompatibility complex. The 1995 Albert Lasker Medical Research Award. JAMA 274:1074-6.

Stryer L. 1988. Biochemistry, 3rd ed. W. H. Freeman, New York.

Whewell W. 1847. The Philosophy of the Inductive Sciences. 2nd. Ed. Johnson reprint Corp. New York.

Wimberly BT, Broderson DE, Clemons Jr, WM, Morgan-Warren RJ, Carter AP, Vonrheim C, Hartsch T, Ramakrishnan V. 2000. Structure of the 30S ribosomal subunit. Nature 407:327–339.