15

Group Report: How Does Complexity Lead to an Apparently Simple Function?

K. MOFFAT, Rapporteur

J.-P. CHANGEUX, D.M. CROTHERS, H. GRUBMÜLLER, G.U. NIENHAUS, M.U. PALMA, F.G. PARAK, K. SCHULTEN, and A. WARSHEL

Keywords: macromolecular dynamics, thermodynamic averaging, energy landscapes, biological selection, quantum treatments, simplification, molecular environments, macromolecular assemblies, common rules

INTRODUCTION: FROM COMPLEXITY TO SIMPLICITY

A bewildering complexity is revealed by even a first, casual glance at the information available on macromolecular structure and function, on integration of these structures into higher-order molecular and cellular assemblies, and on the more elaborate functions and modes of regulation displayed by such assemblies. It appears as though each entity has to be considered alone, its uniqueness more important than its similarities to its neighbors. If this were true, then the study of biology at the molecular and cellular level by the techniques of physics and chemistry would be daunting indeed. In this chapter we consider to what extent, and how, this complexity may be reduced to a simpler view of both structure and function. Can "conceptual averaging" capture the essential properties? How may apparently simpler functions, all of which are associated at some level with flexibility and the motion of atoms, arise from appropriate statistical averaging over vastly more complicated, underlying motions? With such a reductionist viewpoint, there is always the danger that injudicious simplification may lose the key elements of the problem; one ends up with naïveté rather than simplicity.

We choose not to venture a formal definition of the word "complexity." There are too many pitfalls in potential conflict with existing definitions in physics and

Simplicity and Complexity in Proteins and Nucleic Acids
Edited by H. Frauenfelder, J. Deisenhofer, and P.G. Wolynes

mathematics. In this chapter, we use the word primarily in the sense of our inability to understand intuitively those systems that are properly described by large numbers of variables, for example in coordinate and momentum space. Such highly multidimensional systems are likely to be complex and yet, reduction of dimensionality by averaging may lead to a description of a macromolecule and its functions that exhibits simplicity. Although our approach is largely that of the physical chemist, we illustrate our points with both simpler and more complex biological examples.

CLASSICAL AND QUANTUM DYNAMICS, THERMODYNAMICS, LANDSCAPES, AND KINETICS

Simplification through Thermodynamic Averaging

The motion of biopolymers is governed by potential surfaces with many thermally accessible minima. Averaging over the resulting heterogeneities may produce a simpler effective behavior. Computational execution of such averaging includes integration over phase space to yield the partition function from which the thermodynamic parameters are obtained, or reduction of a complex reaction profile to a low-dimensional reaction coordinate. A protein's structural change may be captured by a two-state model with a single rate constant connecting the states, even though there may be billions of substates in each "state" and a vast number of trajectories that connect the two states.

An example is the R to T transition in hemoglobin. In a simple model, one may consider two reaction coordinates: a metal out-of-plane coordinate and the distance between a ligand and the heme iron. One can then calculate the corresponding free-energy barrier for moving from the R to T minima in this two-dimensional diagram (see, e.g., Warshel and Weiss 1981). Simplification, however, inevitably leads to a loss of information, and it is possible to carry the process too far and lose the key features. Indeed, some systems may be irreducible, in that any averaging comes at the price of loss of fundamental features.

Many processes in proteins acquire simplicity since they can be dealt with in terms of first-order kinetics, in contrast to cases in which the detailed dynamics of the system must be considered in order to explain the observed behavior. This question can be reformulated in terms of the permitted order of averaging over the substates of a system. If, for example, the substates of states A and B equilibrate rapidly compared to the rate of the A to B transition, then the system should be adequately described by a single rate constant, and a single decay time in the relaxation spectrum. If, however, equilibration over substates is only as fast as the A to B transition or is slower, the detailed dynamics is important and will be reflected in a multiexponential or even continuous relaxation spectrum. In the first of these cases one can, in principle, get a reasonable estimate of the rate constant by calculating the probability of the transition state, whereas the second case requires an explicitly time-dependent treatment. However, at the limit of slow equilibration we may average over rate constants of reactions that start from different substates.

Biological Selection Acting on Simple Function

The "apparently simple function" that characterizes proteins and nucleic acids is a consequence of natural selection of polymers from a highly multidimensional sequence space in a self-replicating system. Since each polymer sequence is also highly multidimensional in coordinate and momentum space, the selection of a function that confers reproductive advantage reflects an extraordinary degree of simplification. The basis for selection of simple function can be viewed in simple terms. Two obvious properties are the stability of a protein or nucleic acid, and the rate at which it can carry out a process. It is obvious that a protein must be selected for stability of the native structure in the environment where the organism lives. On the other hand, maximizing stability is clearly not an objective, since the unusual stability of proteins from extremophiles demonstrates that much higher stabilities are possible than seen in mesophiles. Furthermore, more stable mutants frequently lead to less efficient enzymes (Shoichet et al. 1995).

Selection for kinetic properties is obvious in the case of enzymes, since organisms may benefit from a substantial acceleration of a chemical reaction rate. On the other hand, maximizing all reaction rates is not a viable general objective, since the metabolism of an organism must retain a steady state balance and evolution actually operates at the level of metabolic pathways. Selection of a specific relaxation time for successful function is illustrated by neurotransmitter proteins, which display fast and slow conformational changes as a basis for developing a "memory effect" (see below).

An illustration of the balance between kinetic and stability selection can be found in the interaction of two RNA loops which is the first step in a sense–antisense interaction that leads to control of copy number of the *ColE1* plasmid in *Escherichia coli* (Polisky 1998). The sequences of the complementary 7-nucleotide loops are not optimized for maximum stability, but rather for a high reaction rate constant to form the loop–loop interaction (Eguchi and Tomizawa 1991).

Selection for explicitly dynamic properties can also be identified. It is obvious that the polymers on which life is based must be able to search configuration space rapidly enough to solve the folding problem. As a more specific example, protein dynamics must be sufficiently rapid to open a channel to the active site on a time scale appropriate to the needed turnover rate of the enzyme.

Energy Landscapes

In the case of a protein, a description in terms of energy landscapes has proved useful. Contact between a microscopic model and measurable macroscopic quantities is provided by weighted thermodynamic averaging among a huge variety of configurations. This concept greatly helps our understanding of how a protein works at physiological temperatures, if we average thermodynamically accessible configurations on a time scale short compared to the time scale of protein function. This averaging over "fast" variables is expected to smooth out the energy landscape as described through the

potential energy function. The averaging gradually transforms the energy landscape into an effective free energy.

The "energy landscape" picture of proteins has been introduced to describe a potential energy surface in an essential subset of degrees of freedom that is structured over wide ranges of energy (barriers) and conformational length scales (see Eaton et al., this volume). This potential energy landscape turns into a free-energy surface for temperatures high enough or times of conformational interconversions long enough to ensure thermodynamic averaging within quasi-equilibrium intermediate states. In this limit, the landscape picture is equivalent to a kinetic scheme with multiple states; an example is provided by the bacteriorhodopsin photocycle (Oesterhelt et al. 1992; Schulten et al. 1995).

However, experiments have shown that complete thermodynamic averaging is more the exception than the rule for actual protein processes so that simple schemes give only a zero-order description. For example, in reaction centers for photosynthetic bacteria, sequential electron transfer over several steps occurs with transfer rate coefficients spanning the time range between picoseconds and seconds. Even at room temperature, only a small fraction of averaging over protein configurations occurs on the picosecond time scale so that the fast events occur in a (largely) heterogeneous ensemble of proteins (Parson et al. 1998; McMahon et al. 1998).

A second example of nontrivial kinetic behavior is ligand binding to heme proteins (see chapters by Nienhaus and Eaton, both this volume). After breaking of the covalent bond between the heme iron and the oxygen, the ligand either rebinds in a geminate process or exits. The nonexponentiality of the geminate process, even at physiological temperature, shows that fluctuational averaging is not complete at the time the ligand decides to either rebind or exit. Consequently, a simple three-well description of the process of ligand dissociation (and association) cannot adequately describe this process either.

These are just two illustrations of a problem that is general for biomolecules and arises from the fact that proteins are complex structures with relaxations stretched out over the entire time range, starting with picoseconds and extending beyond kiloseconds.

The energy landscape picture is clearly necessary to adequately describe experimental results. A quantitative and predictive use of this concept in theory seems to be very difficult, but the question of how to reduce the complexity of an energy landscape to a few relevant coordinates able to describe protein function is posed as a challenge to the theorists among us. Following pioneering work by Go and Scheraga (1969), initial steps have been taken (see, e.g., Karplus and Kushick 1981; Garcia 1992; Amadei et al. 1993; Orešic and Shalloway 1994; Grubmüller 1995; Balsera et al. 1996).

In modeling chemical processes in enzymes, one usually starts the evaluation of the activation free energy by performing a free-energy perturbation/umbrella sampling calculation and forcing the system to move from the reactant to the intermediate (or product) state. Such calculations are done starting from one protein structure and sometimes averaging over several initial coordinates and utilizing drastically

simplified solvent models, by simulated annealing. The brute force approach will involve generating more and more coordinates. In principle, one may look for landscape concepts to determine how much averaging should be done in the direction orthogonal to the reaction coordinate, if suitable algorithms can be provided (Eaton et al., this volume). By sampling a (possibly small) fraction of (relative) energy minima and properly employing energy landscape concepts, one may obtain an estimate of the global energy minimum without actually having to visit all minima.

Since the pioneering work of Frauenfelder and coworkers, it is known that substates can play an important role in protein dynamics as a function of temperature (Austin et al. 1975). It is often stated that the substates occur in a hierarchical fashion with a series of separate tiers characterized by markedly different heights of barriers that separate the substates within each tier (Ansari et al. 1985; Frauenfelder et al. 1991). This view seems to arise at least in part from the limited measurements that can be made. In myoglobin, for example, each tier is characterized by a single experimental marker. Simulations suggest, however, that there is a continuum of substates, at least in the range of fluctuations that can be studied. The concept of substates is clearly important and one in which additional research is needed.

One may question, however, the extent to which the landscape view offers more than the potential energies employed in simulations, when the latter are combined with straightforward thermodynamic analyses. A method for selection of essential coordinates has not been provided by the advocates of the landscape picture and there appear to be few, if any, examples where the landscape picture has yielded insight into protein mechanisms that could not have been gained otherwise.

Determining Essential Degrees of Freedom

One way to depart from the overwhelming complexity of 3N-6 degrees of freedom of a protein (N being the number of atoms of the protein molecule) to simpler — yet sufficiently accurate — descriptions is dimension reduction, which is considered a particularly promising route by which to base the landscape concept on microscopic models. Success here rests on the ability to identify the essential degrees of freedom and to eliminate the irrelevant ones. Normal mode analysis or the more general tool of principal component analysis (applied to Boltzmann structure ensembles) can provide such essential degrees of freedom (Amadei et al. 1993; Hayward et al. 1994), but only after sufficient sampling and not in a predictive manner (Balsera et al. 1996). In particular, principal component analysis has been used to predict slow collective conformational motions (Grubmüller 1995) that cannot otherwise be predicted by, for example, conventional brute force, nanosecond molecular dynamics simulations.

Once essential degrees of freedom are defined, by integration over the orthogonal subspaces, a free-energy landscape for the essential degrees can be derived from the potential (enthalpic) energy landscape for all atomic coordinates (for a good review, see Elber 1996). It is assumed that such a dimension-reduced description of the dynamics of the essential degrees of freedom generally will have to be corrected for

memory effects arising from the neglect of degrees of freedom considered irrelevant. Here the crucial question is how many "essential" coordinates suffice for a proper description of protein dynamics.

Nucleic Acids

Nucleic acids in their unfolded or single-stranded state are just as complex in coordinate and momentum space as are proteins. However, they differ from proteins in that simple base-pairing and base-stacking interactions dominate the structural organization. The "folding problem" for double helical DNA, whose logic was solved by Watson and Crick, proceeds by linear helix growth from a (randomly chosen) double helical nucleus. Folding of RNA molecules generally nucleates from hairpin helices, followed by double helix formation between more distant sequences, and formation of tertiary structural interactions. The secondary structure of complex RNAs, such as those in the ribosome, has been established by phylogenetic comparisons, in which it is observed that a sequence change at one site is accompanied by a complementary change at the site to which it is base paired. Thus, because of the dominance of strongly polar interactions in nucleic acids, the logic and pathway of their folding is much simpler than for proteins for which nonpolar interactions with their short range and nondirectional character are of great importance (Zarrinkar and Williamson 1994; Pyle and Green 1995; Doudna and Doherty 1997; Batey and Doudna 1998).

Is Dynamics Essential?

In many cases, and especially when the barrier for a reaction is sufficiently high, one finds that the reaction rate is determined by the probability of reaching the transition state (the activation free energy, $\Delta G^{\neq}$) rather than by pure dynamical effects. To address this point, we first need to define "dynamical effects." Several definitions are acceptable and we can, for example, state that "if we can determine a rate constant by using configurational sampling approaches (e.g., Monte Carlo simulation) then dynamical effects do not play a major role." This will be the case if the transmission factor in transition state theory is near unity and the prefactor does not vary too much in protein-solvent systems (Warshel 1991). The latter assumption is, however, the subject of controversial discussions. In other words, all processes involve motion but in many cases we can determine the probability of a process from the relevant activation energy.

In describing chemical reactions one has to account for the fact that a molecule fluctuates for a long time in the reactant state until the rare event occurs in which a fluctuation carries the molecule to the transition state and to the product state. The chance for such fluctuation is proportional to $\exp\{-\Delta G^{\neq}/RT\}$. The rate constant is then given by this expression multiplied by a preexponential factor that reflects the time that a trajectory needs to pass the transition state region. This time can be determined by running downhill trajectories and monitoring how long it takes to fall from the transition state

to the reactant or product state minima. One can discern from such simulations how many times a system goes back and forth at the transition state, thereby providing the transmission factor. Once the barrier is sufficiently high, the time of the downhill trajectories is similar for barriers of different heights and, thus, we have a similar prefactor (usually not much different than unity) for many processes with moderate or large activation barriers (Warshel 1991).

A space and time-average structure in the native state, presumably at the global minimum of free energy, is revealed by such techniques as X-ray crystallography and magnetic resonance. The structure itself does not reveal directly the energy or the free energy; these must be derived from the structure by the application of suitable potential functions and estimates of the entropic contributions. If this can be accurately performed, then a comparison of two structures will yield a measure of the free-energy difference between them and hence (if appropriate) the equilibrium constant that governs their interconversion. This procedure is the more accurate, the more closely related the structures are; the free-energy perturbation method has proved effective in such cases (see, e.g., Radmer and Kollman 1997).

In relating structure to kinetics, the additional difficulty arises that the structural comparison is between structures at a free-energy minimum (a valley in the free energy vs. reaction coordinate cartoon) and a maximum (a peak or transition state). Although the minima may be directly accessible to experiment via either a conventional, static structure determination in the case of reactants or products, or via a time-resolved structure determination in the case of transient intermediates, the maxima will always be experimentally inaccessible (see Dill et al., this volume). Transition state geometries can only be inferred by suitable extrapolation from accessible structures, based on further chemical and enzymological knowledge. Such inferences, however, may be hard to draw. The prediction of rate constants from structure, or the understanding of observed rate constants in terms of structure, is therefore a hazardous process.

A first example may be provided by the reactivity of the α-and β-chains of deoxyhemoglobin in the T quaternary structure towards oxygen. The complete structure of deoxyhemoglobin reveals that the prospective binding pocket is open in the α-chain, but obstructed by the side chain of a valine residue in the β-chain. One might, therefore, predict that the α-chain would react with ligands more rapidly than the β-chain (Perutz 1970). However, kinetic experiments yielded the opposite result (Olson and Gibson 1972). This discrepancy may be accounted for by noting that the free energies of activation associated with transfer of ligand from bulk solvent to the ligand pocket of the α- and β-chains, adjacent to the iron, are required. Thus, the free energies of activation involved in the percolation of ligand through the α- and β-chains towards the ligand pocket, and the free energy required to displaced the valine are needed — and these are by no means obvious from a simple inspection of the deoxyhemoglobin structure.

Superoxide dismutase (SOD) provides a second example for a clear, yet surprising, structure–kinetics relationship. The very high binding rate relies on the electrostatic

field surrounding the protein dimer. This field is shaped through the dielectric inhomogeneity of the protein–solvent interface as well as through salt effects; its theoretical description requires solution of the Poisson–Boltzmann equation. Brownian dynamics of the ligand–enzyme system in the force field along both translational and rotational degrees of freedom can explain the observations of Getzoff et al. (1992) on the action of human SOD and its mutants, binding rates of which differ by factors of 2–4. However, in the case of a mutant SOD from *Propionibacterium shermanii*, a reduction of the activity by 90% was observed, but only 17% could be accounted for by electrostatic effects (Gabbianelli et al. 1997). Nevertheless, no structural differences between the wild type and the mutant could be seen (M. Schmidt, F. Parak, and A. Desideri, unpublished results).

More challenging is an attempt to determine the rate constants for an enzyme mechanism by knowing the structures of the free enzyme and an enzyme-substrate complex. Here, one can take the initial structure, define potential surfaces for reactants and intermediate states in an assumed chemical mechanism, and perform free-energy perturbation studies of the relevant activation barriers. Such studies may be quite reliable when implemented with proper theoretical methods (Warshel 1991; Åqvist and Warshel 1993). Such calculations would be even more effective if the structures of the enzyme-product complex and of short-lived intermediates were also available (Dill et al., this volume).

In cytochrome c, the structure of the reduced and the oxidized form, cyt_{red} and cyt_{ox} respectively, has been determined by X-ray structure analysis and by NMR. Within the experimental error it is difficult to see any structural difference. Is it nevertheless possible to calculate from these structures the barrier height between the oxidized and the reduced form? This issue has been addressed by taking the two relevant structures and performing molecular dynamic simulations with the reduced system around the reduced structure and the oxidized system around the oxidized structure (Muegge et al. 1997). One can determine from these calculations the Marcus theory parabolas and, hence, the reaction barrier at the crossing point of the two parabolas.

Need for Quantum Treatment

Simulation tools currently employed for the dynamics of biomolecular systems are based mainly on classical mechanics. The use of those tools has fueled a remarkable progress in computational biology and led to many successes in the interpretation of experimental data. However, for systems having three to four atoms, say, quantum mechanical descriptions are required in order to make quantitative comparisons with experiments that yield detailed trajectories. Relatively accurate quantum mechanical calculations may also be required for many-atom biopolymer systems. We mention a few quantum effects which underscore this requirement.

So-called zero point quantum effects arise from the difference in classical and quantum mechanics in the low energy behavior of vibrational modes. In quantum mechanics there exists a finite, nonvanishing minimal energy content in such modes

which leads to the specific heat approaching zero at temperatures less than $h\omega/k_B$, i.e., for many modes at physiological temperatures. This implies that classical mechanics overestimates energy transfer rates.

Tunneling is a second characteristic quantum process without a classical equivalent. Tunneling reactions have been widely reported, often reflected in strong isotope effects. The geminate rebinding of CO to myoglobin at helium temperature is a classical example for the tunneling of entire molecules (Alben et al. 1983). Recent studies by Klinman and coworkers on tunneling of hydrogen in the glucose oxidase reaction have attracted much attention (Bahnson et al. 1997). Strong isotope effects have also been observed for the kinetics of the proton pump cycle of bacteriorhodopsin.

Interpretation of spectroscopic observations, in particular in regard to force fields, needs to be based on a quantum mechanical description of the underlying vibrational motion, even in cases where the dynamical motions of the atoms involved are not subject to large quantum effects. Intensities of infrared and, in particular, resonance Raman spectra require accurate quantum mechanical descriptions. Vast amounts of vibrational spectra are available for certain protein systems, such as bacteriorhodopsin.

Nonradiative transitions between different electronic states occur in the course of many protein reactions, e.g., in electron transfer or in primary photobiological processes. The transition is an inherently nonclassical effect which typically takes place at intersections or near intersections of potential energy surfaces corresponding to different electronic states. Important photobiological processes where nonadiabatic transitions are known or believed to play a major role are the photoisomerization of retinal in rhodopsin and bacteriorhodopsin (Ebrey 1993; Ben-Nun et al. 1998), of the 4-hydroxycinnamic acid chromophore in photoactive yellow protein, PYP (Perman et al. 1998) and CO binding to myoglobin and hemoglobin (Buhks and Jortner 1985; Frauenfelder and Wolynes 1985).

Following the Born–Oppenheimer approximation, nuclear and electronic degrees of freedom are separated in most present-day simulations of biopolymers: establishment of a force field based on electron properties is succeeded by a description of nuclear motion. However, the increased power of computers already permits electronic and nuclear degrees of freedom to be treated simultaneously for systems of less than about hundred atoms, e.g., using semi-empirical electron Hamiltonians (Warshel and Levitt 1976; Field et al. 1990; Warshel 1991), density functional methods (Stanton et al. 1993; Wesolowski et al. 1996), or the Car–Parrinello method (Molteni and Parrinello 1998).

The described shortcomings will be overcome through a new generation of quantum simulation methods. The feasibility of such simulations has already been documented in a few cases, e.g., for a full quantum mechanical treatment of protein vibrations (Roitberg et al. 1996), for a quantum simulation of all nuclear motions in bacteriorhodopsin during the primary photoprocess (Ben-Nun et al. 1998), for a full description of electronic-nuclear degrees of freedom for an oligonucleotide crystal (Parrinello 1997), and for quantum mechanical tunneling in enzymes by path integral approaches (Hwang and Warshel 1996).

Quantum Mechanics as an Example of Simplification

The quintessential reaction process *reactant* → *product* in chemistry involves electrons switching between two states. Examples are electron transfer reactions in proteins, when an electron moves from an orbital on the donor moiety D to an orbital on the acceptor moiety A, and bond formation or bond breaking in an enzyme, when electrons shift from a nonbonding state to a bonding state or vice versa. When the reaction proceeds in a protein environment, systematic interactions with the protein environment influence the reaction. For example, electrostatic forces stabilize the transition state, but there arise also unsystematic contributions due to the thermal motion of the protein. The latter contribution is not necessarily unwanted. Rather like a diffusion controlled reaction driven by Brownian motion due to collisions with a condensed environment, a two-state quantum system in a protein can be driven by thermal vibrations, as demonstrated in Schulten (1995) and earlier work cited therein.

In fact, monitoring the time-dependent energy gap between two quantum states, $\Delta(t) = E_1(t) - E_2(t)$, provides a useful way of modeling chemical reactions in solution and in proteins (Warshel 1984; Treutlein et al. 1988). Fluctuations in this energy gap are usually determined by the fluctuation of the field from the protein and solvent, arising from nearby dipoles as well as charged residues. These fluctuations occur along structural coordinates $q_j, j = 1, 2, \ldots 3N$ (N = number of protein atoms). Fluctuations across essentially the entire protein are coupled. In the case in which only $\Delta(t)$ is affected and the coupling to the protein is linear, i.e., $\Delta(t) = \Sigma_j c_j q_j$, the influence of the protein on the process can be captured succinctly through stochastic quantum mechanics as outlined in Warshel and Hwang (1986) and Leggett et al. (1987), and applied to proteins by Xu and Schulten (1994); see also Warshel et al. (1989). The theory requires knowledge of the weighted spectral density $J(\omega) \sim \Sigma j\ (c_j^2/\omega_j)\ \delta(\omega - \omega_j)$, where ω_j are the vibrational frequencies of the protein. However, one can express $J(\omega)$ through the correlation functions $<\Delta(t)\ \Delta(0)>$ and RMSD values of $\Delta(t)$ of the protein such that the theory does not rest on the assumption of ad hoc parameters (Xu and Schulten 1994). This electron transfer theory has been successfully applied to experimental data by McMahon et al. (1998). The stated description reduces to the Marcus theory in the limit of high temperature.

The dynamical effects of the protein fluctuations may be converted to activation barriers at the semiclassical high temperature limit. That is, integration of $J(\omega)$ gives the so-called Marcus reorganization energy (Warshel and Hwang 1986) and this can be used in the evaluation of the Marcus activation barrier. If the integrated $J(\omega)$ gives the same rate constant as the more rigorous treatment with the detailed $J(\omega)$, we may conclude that dynamical effects are not critical. This implies that quantum effects do not play a role for all vibrations involved in electron transfer ($h\omega << kT$).

The development of a potential stochastic quantum mechanics description for more general quantum systems in proteins is an active field of research and molecular biophysics could benefit greatly from advances (Schulten 1995). An example for a quantum system that requires a more general description is electronic excitations in

chlorophyll aggregates in the photosynthesis apparatus (as discussed in Hu et al. 1997), where thermal fluctuations introduce incoherencies and localization of quantum states.

WATER MOLECULES TO ORGANELLES: FROM RELATIVE SIMPLICITY BACK TO COMPLEXITY

Molecules and Their Environment

The structure, dynamics, and function of biological macromolecules depend crucially on the properties of their environment. Macromolecules therefore cannot be treated separately from their environment; they form an inseparable system. The environment may differ for different regions of a single protein. For example, an integral membrane protein such as the acetylcholine receptor has three principal domains: a transmembrane domain in which the environment is that of the hydrophobic interior and the more hydrophilic head groups of the lipid bilayer, an extracellular domain which binds the neurotransmitter, and a cytoplasmic domain. The environments of the latter two domains are hydrophilic, aqueous solution. The environment may vary in time; for example, a protein may transiently associate with a membrane surface, or with other macromolecules as part of a multimeric complex, such as a transcription initiation complex or a spliceosome. The environment may be chosen as part of the experimental setup. For example, rarely does one investigate structure–function relations in natural environments. Normally, the protein is isolated from its native environment, purified, and (for example) caused to aggregate in the form of a crystal for structure determination. The crystal structure then provides the basis for a mechanistic explanation of function, although the function itself may have been examined in quite different solvents and environments. For a few systems like hemoglobin and myoglobin that have been studied in more detail, reaction rates in the crystal are similar to those in solution (Rivetti et al. 1993; Nienhaus et al. 1998) which is encouraging, but this fact should not encourage generalization of this observation. Indeed, ligand binding to hemoglobin crystals does not display the key functional feature of cooperativity (Rivetti et al. 1993).

Proteins are often embedded in viscous solvents like glycerol in experiments that study such protein functions as ligand binding, electron transfer, and proton transfer over a wide range of temperature. In metmyoglobin, even the electronic state of the heme iron is modified by surroundings (Bizzari et al. 1995). In many experiments, it has become obvious that protein motions are strongly affected by the dynamics of the surrounding glassy matrix, and the term "slaved glass transition" has been coined to described this situation (Ansari et al. 1987). However, the details are less clear. Which motions are arrested, and which are not affected? In small globular proteins like bacteriorhodopsin, myoglobin and blue copper proteins, most motions at the active site in the interior are frozen in upon vitrification of the solvent. However, for larger

proteins, lesser constraints of the dynamics in the interior by the arrested surface atoms are expected.

One may distinguish two temperature regimes separated by the glass transition temperature, T_g, of the solvent. This separation is by no means sharp. Below T_g, the dynamics can be described essentially by a normal mode analysis. Jumps between conformational substates are rare; a great part of the functional properties is absent; and the molecule is nearly dead. Above T_g, all functions are present although the rates depend on temperature and slow down as T approaches T_g of the surrounding medium. In this physiological temperature region, practically all modes of motion present below T_g are retained but an additional channel of motions is activated. These motions may be described by a diffusion of parts of the molecule with respect to each other, limited in space by the forces which try to preserve the average protein structure. For example, the efficiency of the electron transfer Q_A to Q_B in the reaction center of *Rhodospirillum rubrum* (Parak et al. 1980; Novella and Schulten 1991) or photosystem II of spinach (Garbers et al. 1998) correlates with these motions. In this picture, the glass transition of the environment suppresses functionally important motions. However, it is not necessarily enforcing a glass transition in the protein. A "slaving" of protein dynamics by the environment of the surrounding medium occurs only below T_g (see Parak and Palma, this volume).

Proteins can be embedded in rigid polymer or sugar matrices (see, e.g., Cordone et al. 1998) that restrict or suppress fluctuations between conformational substates even at room temperature, or incorporated into other non-native environments like surfaces or interfaces. Altered, controllable environments are of great interest not only for certain types of experiments, but also for technological applications to the preservation of biomolecules in the food and pharmaceutical industries, and to the development of biocatalysts and biosensors. A better understanding of the interactions between proteins and their environments could greatly assist in their technological application.

Solvent-induced Forces

The known flexibility of water molecules in making hydrogen-bonded structures is the reason for the particular complexity and ruggedness of the energy landscape of liquid water. Solutes such as macromolecules alter this energy landscape and related free energy of water. Consequently, a thermodynamic force arises that attempts to bring solutes toward the nearest spatial arrangement which minimizes the solute-solvent contributions to free energy. The gradient of the latter expresses the pertinent solvent-induced force (SIF).

An example of the interaction between the solvent and a protein is provided by hemoglobin. In its transition from the R (oxy) state to the T(deoxy) state, hemoglobin exposes some 1,000 Å^2 of additional hydrophobic surface to the solvent (Chothia et al. 1976). Experiments estimate the free-energy cost of the additional hydration of this surface at about 60 kJ mol^{-1} (Parak and Palma, this volume). It involves the statistical transfer of about 75 water molecules from the bulk to the hydrophobic surface (Bulone

et al. 1993). This solvent-based contribution is essential to the structural transition and to the associated functional properties of hemoglobin. The simplifying view is reached by averaging over the relatively large hydrophobic surface. At single residue resolution, however, local effects of the solvent are best elicited by SIFs due to their intrinsically differential nature. This allows the notable simplification of expressing the local effects of interactions between the two highly complex systems of protein and water with a limited set of force vectors. Complexity reappears, nevertheless, under the aspect of a strong non-pair additivity of SIFs and related strong many-body character of hydration free energy. Strong evidence for such complexity has been obtained recently (Brugé et al. 1997; Martorana et al. 1996, 1997, 1998; Bulone et al. 1997; San Biagio et al. 1998). As a consequence, SIFs are characterized by a marked dependence upon the specific spatial arrangements (or "context") of solutes. This causes highly selective interactions and unexpected modifications of the free-energy landscape of solutes (Parak and Palma, this volume).

This view can be integrated over a scale even larger than the 1000 $Å^2$ mentioned above, e.g., over entire proteins. In this case, rewarding simplification is obtained in terms of a mean field approximation (Bulone et al. 1993). Simultaneous consideration over two scales allows eliciting interwined simplicity/complexity effects (Parak and Palma, this volume).

Macromolecular Assemblies

Many cellular functions are achieved by assemblies of biopolymers, for example, multi-enzyme complexes, ribonucleoproteins, individual proteins or groups of proteins embedded in membranes, viruses, and complexes of proteins with DNA. Through advances in crystallography and electron microscopy, structures of these large systems are being solved today with increasing frequency. Due to their important functions, the extension of structural biology towards cell biology through study of such assemblies will assume center stage during the next decade.

This will pose new conceptual and technical challenges to researchers. For example, computational methodologies need to be developed further, based on the general increase of processor power and the advent of massively parallel computing (Heller et al. 1990) as well as highly efficient molecular dynamics algorithms (Grubmüller et al. 1991; Eichinger et al. 1995, 1997; Luty and Van Gunsteren 1996; Schlick et al. 1997; Grubmüller and Tavan 1998; Schlick et al. 1999; Kale et al. 1999). Molecular graphics must be adapted to the study of very large assemblies, with the main goal being support in navigating through large structures, up and down the spatial resolution scale. New conceptual (cartoon) and detailed (physical) representations may combine geometrical and physical data. For example, mechanical models of protein domains as the building blocks of aggregates are valuable for examining virus capsids and light-harvesting systems (Bailey et al. 1998).

Simulation tools must be capable of dealing with the numerical task posed by accurate many-atom simulations (10^6 atoms), and of manipulating individual components,

e.g., pulling components apart while adhesion forces are monitored. For the latter purpose new simulation tools, so-called steered (targeted) molecular dynamics methods, are presently being developed and have been successfully applied (Grubmüller et al. 1996; Izrailev et al. 1997; Lu et al. 1998; Marrink et al. 1998; Marszalek et al. 1999; see also Dill et al., this volume). The main challenge at present is the short time scale of computational manipulations which implies a large dissipative component to all forces measured (Grubmüller et al. 1996; Izrailev et al. 1997). This component must be discounted in order to furnish the effective force fields relevant for cellular processes (Gullingsrud et al. 1999; Heymann and Grubmüller 1999).

An example that demonstrates the necessity for modeling and simulation on a larger scale is provided by the problem of domain mobility in the cytochrome bc_1 complex. The cytochrome bc_1 complex from mitochondria consists of 11 protein subunits and four cofactors. As part of the respiratory chain, it oxidizes ubiquinol (QH_2) to quinone (Q) and pumps four protons from the matrix to the intermembrane space. It has two major quinone binding sites, Q and Q_0. At the quinone oxidation site Q_0, one electron of QH_2 is transferred to the iron sulfur protein (ISP) in a strongly exothermal reaction ($\Delta E \sim 300$ mV); the second electron of QH_2 is transferred via the two hemes of cytochrome b to the quinone reduction site Q in a much less exothermal reaction ($\Delta E \leq 100$ mV) (Brandt and Trumpower 1994). X-ray structure analyses (Xia et al. 1997; Kim et al. 1998; Zhang et al. 1998; Iwata et al. 1998) of bovine bc_1 in the presence and absence of different Q_0 inhibitors showed that the extramembrane domain of the ISP can switch between a fixed state, where it is bound to the surface of cytochrome b, and a loose state, where it does not populate a structurally well-defined state. In the loose state, it appears to be free to undergo a hinged motion that could achieve two goals. It could assist electron transfer from the ISP to cytochrome c_1 by allowing transient contacts between these proteins; or, transition from the fixed to the loose state could prevent the transfer of both electrons to the iron-sulfur protein which would halve the enzymatic efficiency. Different conformations of the ISP were also reported in crystal structures of chicken bc_1 complex in the presence and absence of a Q_0 inhibitor and of the complex from rabbit (Zhang et al. 1998)

The nature and kinetics of the ISP motion can be simulated by large-scale computational modeling (Izrailev et al. 1999) and may eventually explain the electron transfer kinetics between QH_2 and cyt c_1.

Levels of Organization

In the biological sciences, the simplifying concept of level of organization has been found rather useful. In the brain, for instance, one may distinguish the molecular from the cellular level, the "small networks" level from the level of neuronal assemblies, and the level of "assemblies of assemblies." Theoretical and experimental work will attempt to relate a particular level of organization to a relevant function, knowing *a priori* that these hierarchical levels are "nested" and interdependent. The Hodgkin and Huxley model for transmission of the nerve impulse was proposed at the cellular level

without knowing the actual structure of ion channels, and the Hebb rule for learning was proposed at the neuron assembly level without knowing what neurotransmitter receptors are as molecular entities.

One may try to apply such views to protein molecules themselves, where hierarchical levels of organization may also be recognized. For example, individual subunits of the acetylcholine receptor are split into spatially distinct, functional domains; the extracellular N-terminal domain contributes to ligand binding, the transmembrane domain contributes to the ion channel, and the cytoplasmic domain to regulation. These domains display both folding and functional autonomy. Indeed, one may build a functional receptor by linking together the binding domain of the $\alpha 7$ neural nicotinic receptor and the transmembrane or channel domain of the $5HT_3$ (serotonin) receptor, despite considerable sequence differences (Eiselé et al. 1993). In the case of the glutamate receptor, the evolutionary origin of these two distinct domains can be traced back to bacteria (Paas 1998). The glutamate-binding domain strikingly resembles the bacterial periplasmic lysine/arginine/ornithine-binding protein or the leucine/isoleucine /valine-binding protein. On the other hand, the channel domain shares such structural features with the bacterial potassium channel as its tetrameric organization and presence of a P loop. One may say that "tinkering" with two bacterial proteins suffices to create a brain receptor!

At the next highest structural level, that of quaternary structure, "emergent" properties may be recognized which are directly dependent upon specific features of that structure. To build a binding site, the single subunit level suffices. On the other hand, to achieve cooperativity between binding sites seems to require the assembly of subunits (protomers) into limited, closed, symmetrical assemblies referred to as oligomers. This is the case for hemoglobin, for many regulatory enzymes such as aspartate transcarbamoylase, phosphorylase b and L-lactate dehydrogenase, for receptors of neurotransmitters, and so on. Without a doubt, the cooperative binding of specific ligands to these flexible oligomers is dependent upon their quaternary organization. As proposed in the Monod–Wyman–Changeux (MWC) model, their equilibrium binding and kinetic data can be accounted for by a "concerted" or cooperative conformational transition of the quaternary assembly (Eaton, this volume). In other words, the "simple" functional feature of responding to the threshold concentration of a regulatory ligand is "determined" by the more complex assembly of subunits into an oligomeric structure – a common feature of many regulatory, allosteric proteins. From an evolutionary point of view, only a few surface mutations may suffice to confer on an independent subunit the ability to associate with itself, to acquire quaternary structure, and thus to form a cooperative oligomer.

In the case of the acetylcholine receptor, the situation is even more complex. The receptor prepared from fish electric organ (Changeux 1990) is a 300 kDa heteropentamer made up of 4 homologous subunits in the clockwise order $\alpha\gamma\alpha\delta\beta$ (Hucho et al. 1996). It carries two acetylcholine binding sites at the boundaries between subunits and possesses an ion channel located along the fivefold rotation axis of pseudosymmetry of the molecule. The assembly of a fivefold pseudosymmetrical

oligomer from four chemically different subunits remains a challenge for the physical chemists as well as for the molecular biologists. Early work suggests that: (1) the assembly of the subunits takes place through initial formation of an $\alpha\gamma$ and an $\alpha\delta$ dimer, their association together, and inclusion of the β subunit to form the functional pentameric oligomer; (2) folding and assembly of the subunits occurs in the endoplasmic reticulum and is a very slow process, requiring hours; and (3) many molecules do not reach the correct conformation and are destroyed (only about 20% of adequately assembled molecules "survive" in the cytoplasmic membrane). The folding and assembly processes of such complex molecules include a long series of molecular and cellular processes which result in drastic final selection of the functional molecule. It would be a mistake to view the assembly process as a strictly computational problem.

The distance between acetylcholine binding sites and the ion channel is in the range of 20–30 Å (Herz et al. 1989). Interaction typically takes place between topographically distinct sites; it is, by definition, allosteric (Monod et al. 1963). The signal transduction mechanism thus emerges at the level of the quaternary organization of the protein.

From a kinetic point of view, the receptor oligomer undergoes a cascade of conformational transitions upon binding the neurotransmitter: a fast opening of the ion channel in the millisecond time scale (activation transition), and a slower "desensitization" transition in the 100 ms to second (or even minute) time scale which leads to a high affinity "refractory state." A simple representation of the transition includes four states denoted R, A, I, and D (Figure 15.1) which, in agreement with the MWC scheme, may preexist in the absence of ligand binding (Changeux 1990; Edelstein et al. 1996).

Many, if not all, ligand-gated ion channels exhibit slow desensitization transitions. Their functional significance is evident. They regulate the "efficiency" of the receptor response to the neurotransmitter. They may serve as elementary "memory" devices for the storage of short-term memories in neural networks (see Heidmann and Changeux 1982; Dehaene and Changeux 1997). Moreover, the transmembrane organization of the receptor is such that they may "integrate" multiple extracellular as well as intracellular signals through such transitions, thus serving as time coincidence detectors (see Edelstein and Changeux 1998). A more complicated structure thus allows for a more diverse range of function and much more sophisticated regulation.

Selection may not exert its action solely on fast reactions. Slow transitions may have important physiological functions in the brain, in photoreceptors and in other systems, and thus be directly selectable by evolution. Indeed, as a rule of thumb the larger a molecular entity or complex, the slower are certain of its structural transitions. Quaternary structural transitions are usually slower than (and preceded by) tertiary structural transitions.

Further structural and functional complexities arise at the level of the whole cell. Proteins are not necessarily dispersed within the cytoplasm in a random, nonorganized manner. Membrane proteins may, for instance, cluster in small-scale assemblies or

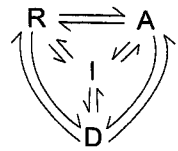

Figure 15.1 Conformational transitions in the acetylcholine receptor (Changeux 1990; Edelstein et al. 1996). For acetylcholine, the R to A activation equilibrium constant has a value of around 10^{-5} M; the I to D desensitization equilibrium constant has a value around 10^{-8} M.

even in well-defined edifices limited, on the outside, to the basal lamina or to the connective tissue environment; or, on the inside, to the cytoskeleton. A particularly remarkable example is the very restricted topological distribution of neurotransmitter receptors at the synapse (Duclert and Changeux 1995). At the neuromuscular junction, the surface density of acetylcholine receptor molecules under the nerve ending where neurotransmitter release takes place is near to close packing (several thousand receptors per μm^2); the lateral diffusion of the receptors is negligible; and the receptor protein is metabolically stable with a half-life of about 11 days. The stability of the postsynaptic aggregate of receptor molecules relies upon multiple interactions with other extracellular structures (agrin, collagen, "cleft substances"...) and with intracellular, cytoplasmic, structures (43K-rapsyn, cytoskeletal proteins...). The receptor molecules traverse the lipid bilayer and in the aqueous environment on both sides of the membrane, they may interact with other protein or polysaccharide species through specialized binding sites.

We emphasize that this rather static picture of the higher-order assemblies of macromolecules does not fully suffice to understand their properties and stability. In particular, we must take into account the developmental aspects of synapse formation, in particular the contribution of local regulation of gene transcription in this process. Indeed, the myofiber is a multinucleate structure but there exists a differentially high transcriptional activity of acetylcholine receptor genes at the level of the subjunction of nuclei. In general, understanding protein structure and superstructure requires an understanding of its past evolutionary and developmental histories.

ARE THERE GENERAL LAWS?

In the physical sciences, one is looking for the simplifying concepts embodied in "general laws." The Maxwell equations describe, for instance, the field of electrodynamics in a unique and definitive manner. The question is, can more or less general laws also be formulated for macromolecules and their assemblies? This can be doubted on both philosophical and practical grounds (see, e.g., Parak and Palma, this volume). The concept of "general laws" that rule the universe is rather misleading; the implicit reference to a Platonist world of "Ideas" and "Forms" is no longer acceptable. In general, we refer to more limited and modest "models" which capture regularities in the world on the basis of, for example, a mathematical formalism. No mathematical equation, as

"beautiful" as it may be, can pretend to give an exhaustive description of physical reality. We should, therefore, be cautious and humble in our attempt to evaluate regularities and diversities in the world, and in particular, in the world of macromolecules. The eminent evolutionary biologist Ernst Mayr posited that pitfalls might be avoided in the biological sciences by rejecting the so-called "physicist view of the world." We must recognize that living organisms are not homogeneous systems but display an immense structural diversity, unlike systems studied conventionally in physics, and such diversity cannot be understood without rejecting the standard static view of the world and incorporating biological evolution with its very long time scales. (It must be stated that the above paragraph does not reflect the opinion of all authors).

At a more practical level, the question regarding general laws can also be asked in regard to the physical properties of macromolecules. One may ask, for example, if descriptions valid over a wide length scale can be found, e.g., if a conventional full atom molecular dynamics description can be converted into simplified descriptions which include averages over, for example, the 10 to 50 Å scale. This question had been posed in the theory of liquids with the expectation that a full atom molecular dynamics description could be replaced by sequence of pictures which, after coarse graining at a 1000 Å length scale, would assume the well-known forms of hydrodynamics. Despite years of efforts, the theory of liquids did not reach this goal; it turned out to be either extremely difficult or impossible to develop simplified descriptions of liquids valid over the entire length scale of 10 to 1000 Å. Unfortunately, this length scale is that of proteins. One might conclude that similarly it would be hopeless to derive coarse-grained descriptions of proteins. However, certain computationally useful approaches on a variety of length scales are being pursued, such as multiple time step algorithms (Tuckerman et al. 1990; Grubmüller et al. 1991; Humphreys et al. 1995; Zhou and Berne 1995; Littell et al. 1997; Schlick et al. 1997), multipole methods (Appel 1985; Barnes and Hut 1986; Greengard and Rokhlin 1989; Board et al. 1992; Eichinger et al. 1997), lattice sum and grid methods (Luty et al. 1995; Luty and Van Gunsteren 1996), "essential dynamics" (Amadei et al. 1993), "conformational flooding" (Grubmüller 1995), and elastic theory as applied to larger stretches of DNA, both supercoiled and in loops (Schlick 1995; Olson 1996; Balaeff et al. 1999).

With these cautionary tales, we should try to carefully limit our ambition to the recognition of common functional properties (catalysis, cooperativity, signal transduction...) and to the identification of the common structural and dynamical features to which the common functions may be carefully related. This is a rather reasonable program. It is preferable to search for "common rules" rather than to strive for "general laws."

Some Common Rules

We supply a few common rules and invite the reader to add to the list.

1. The function of enzymes, transport proteins, receptors, motor proteins, and many other systems is associated with some structural flexibility. Therefore,

the free energy of structural stabilization is very small and corresponds to 2 to 3 hydrogen bonds only. Exceptions include so-called mechanical proteins, like titin, that are protected against mechanical strain.

2. Proteins possess features such that they fold, function, and evolve. Only defined parts of the molecular structure contribute directly to a specific function (see, however, Martorana et al. 1996, 1998). Other parts may become significant for function upon changes of the environment.
3. Many proteins make clever and highly sophisticated use of the entropic component of free energy. Without, of course, violating thermodynamic laws, they nevertheless appear to come closer to Maxwell's Demon than most other many-body systems. An example is the light-driven proton pump, bacteriorhodopsin, which after light absorption stores only a small fraction of the photon's energy. It apparently primes its pump through photochemical selection of a state with high entropy content. The universality of ATP in driving cell processes likewise seems to be based on the high degree of topochemical specificity (directedness) of its endothermic hydrolysis reaction.
4. Catalysis arises from transition state stabilization rather than from ground state destabilization. It appears that increasing the catalytic rate constant k_{cat} involves electrostatic complementarity between the charge distribution of the transition state and the surrounding protein active site.
5. The reduction of complexity to simplicity in the self-organization of polymers is made easier if polar, directional interactions dominate, as is the case for nucleic acids as opposed to proteins.
6. Assembly of a protein into a finite oligomeric structure confers symmetry properties and cooperative behavior.
7. In the case of metalloproteins, the electronic configuration of the metal has a large number of low energy levels (2–600cm^{-1}) in contrast to inorganic and organic metal components (first excited level $\approx$ 10,000 cm^{-1}). This is necessary for reversibility of binding small molecules or accepting additional electrons under conditions of constant temperature or pressure.

CONCLUSIONS

The examples in this chapter present complexity at a variety of experimental and theoretical levels. The prospects of uncovering "general laws" are apparently still distant, as evidenced by the paucity and relative specificity of even the "common rules", above. Perhaps the biologist Stephen Jay Gould said it best, in a recent essay (Gould 1999): "The fallacies and foibles of human thinking lead us into systematic and predictable trouble when we try to grasp the complexities of external reality. Among these foibles, our persistent attempts to build abstractly beautiful, logically impeccable, and comprehensively simplified systems always lead us astray .... Nature, to cite a modern

cliché, always bats last. She will not succumb to the simplicities of our hopes or mental foibles, but she remains eminently comprehensible."

Observations may be relatively simple, but elucidating the connections between the undoubted complexity of macromolecular structure and dynamics and these observations remains a major research challenge.

ACKNOWLEDGMENTS

We thank Hans Deisenhofer and Martin Karplus for their thoughtful comments and contributions.

REFERENCES

Alben, J.O., D. Beece, S.F. Bowne, L. Eisenstein, H. Frauenfelder, D. Good, M.S. Marden, P.P. Moh, L. Reinisch, A.H. Reynolds, and K.T. Yue. 1983. Isotope effect in molecular tunneling. *Phys. Rev. Lett.* **44**:1157–1160.

Amadei, A., A.B.M. Linssen, and H.J.C. Berendsen. 1993. Essential dynamics of proteins. *Proteins* **17**:412–425.

Ansari, A., J. Berendzen, S.F. Bowne, H. Frauenfelder, I.E.T. Iben, T.B. Sauke, E. Shyamsunder, and R.D. Young. 1985. Protein states and protein quakes. *Proc. Natl. Acad. Sci. USA* **82**:5000–5004.

Ansari, A., J. Berendzen, D. Braunstein, B.R. Cowen, H. Frauenfelder, M.K. Hong, I.E.T. Iben, J.B. Johnson, P. Ormos, T.B. Sauke, R Scholl, A. Schulte, P.J. Steinbach, J. Vittitow, and R.D. Young. 1987. Rebinding and relaxation in the myoglobin pocket. *Biophys. Chem.* **26**:337–355.

Appel, A.W. 1985. An efficient program for many-body simulation. *SIAM J. Sci. Stat. Comput.* **6**:85–103.

Åqvist, J., and A. Warshel. 1993. Simulation of enzyme reactions using valence bond force fields and other hybrid quantum/classical approaches. *Chem. Rev.* **93**:2523–2544.

Austin, R.H., K.W. Beeson, L. Eisenstein, H. Frauenfelder, and I.C. Gunsalus. 1975. Dynamics of ligand binding to myoglobin. *Biochemistry* **14**:5355–5373.

Bahnson, B.J., T.D. Colby, J.K. Chin, B.M. Goldstein, and J.P. Klinman. 1997. A link between protein structure and enzyme catalyzed hydrogen tunneling. *Proc. Natl. Acad. Sci. USA* **94**:12,797–12,802.

Bailey, M., U. Schulten, and J.E. Johnson. 1998. The use of solid physical models for the study of macromolecular assembly. *Curr. Opin. Struct. Biol.* **8**:202–208.

Balaeff, A., L. Mahedevan, and K. Schulten. 1999. Elastic rod model of a DNA loop in the lac operon. *Phys. Rev. Lett.* **83**:4900–4903.

Balsera, M.A., W. Wriggers, Y. Oono, and K. Schulten. 1996. Principal component analysis and long time protein dynamics. *J. Phys. Chem.* **100**:2567–2572.

Barnes, J., and P. Hut. 1986. A hierarchical O(N log N) force-calculation algorithm. *Nature* **324**:446–449.

Batey, R.T., and J.A. Doudna. 1998. The parallel universe of RNA folding. *Nature Struct. Biol.* **5**:337–340.

Ben-Nun, M., F. Molnar, H. Lu, J.C. Phillips, T.J. Martinez, and K. Schulten. 1998. Quantum dynamics of retinal's femtosecond photoisomerization in bacteriorhodopsin. *Faraday Soc. Disc.* **110**:447–462.

Bizzari, A.R., O.A. Iakovleva, and F. Parak. 1995. Spin lattice relaxation in Mössbauer spectra of metmyoglobin: Investigation of crystals, water, and water-glycerol solutions. *Chem. Phys.* **191**:185–194.

Board, J.A., J.W. Casey, J.F. Lethrum, A. Windemuth, and K. Schulten. 1992. Accelerated molecular dynamics simulation with the fast multipole algorithm. *Chem. Phys. Lett.* **198**:89–94.

Brandt, U., and B.L. Trumpower. 1994. The protonmotive Q cycle in mitochondria and bacteria. *CRC Crit. Rev. Biochem. Mol. Biol.* **29**:165–197.

Brugé, F., S.L. Fornili, G.G. Malenkov, and M.B. Palma-Vittorelli. 1997. Solvent-induced forces on a molecular scale: Non-additivity, modulation, and causal relation to hydration. *Chem. Phys. Lett.* **254**:283–291.

Buhks, E., and J. Jortner. 1985. The dynamics of carbon monoxide binding to hemoglobin and to cytochrome. *J. Chem. Phys.* **83**:4456–4462.

Bulone, D., V. Martorana, P.L. San Biagio, M.B. Palma-Vittorelli. 1997. Effects of electric charges on hydrophobic forces. *Phys. Rev. E* **56**:R4939–R4942.

Bulone, D., P.L. San Biagio, M.B. Palma-Vittorelli, and M.U. Palma. 1993. On the role of water in hemoglobin function and stability. *Science* **259**:1335–1336.

Changeux, J.-P. 1990. Functional architecture and dynamics of the nicotinic acetylcholine receptor: An allosteric ligand-gated ion channel. Fidia Research Foundation Neuroscience Award Lectures. **4**:21–168.

Changeux, J.-P., and S.J. Edelstein. 1998. Allosteric receptors after 30 years. *Neuron* **21**:959–980.

Chothia, C., S. Wodak, and J. Janin. 1976. Role of subunit interfaces in the allosteric mechanism of hemoglobin. *Proc. Natl. Acad. Sci. USA* **73**:3793–3797.

Cordone, L., P. Galajda, E. Vitrano, A. Gassmann, A. Ostermann, and F. Parak. 1998. A reduction of protein specific motions in CO-ligated myoglobin embedded in a trealose glass. *Eur. Biophys. J.* **27**:173–176.

Dehaene, S., and J.-P. Changeux. 1997. A hierarchical neuronal network for planning behavior. *Proc. Natl. Acad. Sci. USA* **94**:13,293–13,298.

Doudna, J.A., and E.A. Doherty. 1997. Emerging themes in RNA folding. *Folding and Design* **2**:R65–R70.

Duclert, A., and J.-P. Changeux. 1995. Acetylcholine receptor gene expression at the developing neuromuscular junction. *Phys. Rev.* **75**:339–368.

Ebrey, T. 1993. Light energy tranduction in bacteriorhodopsin. In: Thermodynamics of Membranes, Receptors and Channels, ed. M. Jacobson. New York: CRC Press.

Edelstein, S., O. Schaad, E. Henry, D. Bertrand, and J.-P. Changeux. 1996. A kinetic mechanism of nicotinic acetylcholine receptors based on multiple allosteric transitions. *Biol. Cybern.* **75**:361–379.

Eguchi, Y., and J. Tomizawa. 1991. Complexes formed by complementary RNA stem loops: Their formations, structures, and interactions with Col E1 Ram protein. *J. Mol. Biol.* **220**:831–842.

Eichinger, M., H. Grubmüller, and H. Heller. 1995. User manual for EGO_VIII, Release 2.0, electronic access: http://www.mpibpc.gwdg.de/abteilungen/o71/ego.html.

Eichinger, M., H. Grubmüller, H. Heller, and P. Tavan. 1997. FAMUSAMM: A new algorithm for rapid evaluation of electrostatic interactions in molecular dynamics simulations. *J. Comp. Chem.* **18**:1729–1749.

Eiselé, J.L., S. Bertrand, J.L. Galzi, A. Devillers-Thiéry, J.-P. Changeux, and D. Bertrand. 1993. Chimaeric nicotinic-serotonergic receptor combines distinct ligand binding and channel specificities. *Nature* **366**:479–483.

Elber, R. 1996. Reaction path studies of biological molecules. In: Recent Developments in Theoretical Studies of Proteins, ed. R. Elber. Tai Seng, Singapore: World Scientific.

Field, M.J., P.A. Bash, and M. Karplus. 1990. A combined quantum-mechanical and molecular-mechanical potential for molecular-dynamics simulations. *J. Comp. Chem.* **11**:700–733.

Frauenfelder, H., S.G. Sligar, and P.G. Wolynes. 1991. The energy landscapes and motions of proteins. *Science* **254**:1598–1603.

Frauenfelder, H., and P.G. Wolynes. 1985. Rate theories and puzzles of hemeprotein kinetics. *Science* **229**:337–345.

Gabbianelli, R., A. Battistoni, E. Polticelli, B. Meier, M. Schmidt, G. Rotolio, and A. Desideri. 1997. Effect of Lys 175 mutation on structure-function properties of Propionibacterium shermanii superoxide dismutase. *Protein Eng.* **10**:1067–1070.

Garbers, A., F. Reifarth, J. Kurreck, G. Renger, and F. Parak. 1998. Correlation between protein flexibility and electron transfer from QA to QB in PSII membrane fragments from spinach. *Biochemistry* **37**:11,399–11,404.

Garcia, A.E. 1992. Large-amplitude nonlinear motions in proteins. *Phys. Rev. Lett.* **68**:2696–2699.

Getzoff, E.D., D.E. Cabelli, C.L. Fisher, H.E. Parge, M.S. Viezzoli, L. Banci, and R.A. Hallewell. 1992. Faster superoxide dismutase mutants designed by enhancing electrostatic guidance. *Nature* **358**:347–351.

Go, N., and H.A. Scheraga. 1969. Analysis of the contribution of internal vibrations to the statistical weights of equilibrium conformations of macromolecules. *J. Chem. Phys.* **51**:4751–4767.

Gould, S.J. 1999. Branching through a wormhole: Lamarck's ladder collapses. *Nat. History* **108**:24ff.

Greengard, L., and V. Rokhlin. 1989. On the evaluation of electrostatic interactions in molecular modeling. *Chem. Scr.* **29A**:139–144.

Grubmüller, H. 1995. Predicting slow structural transitions in macromolecular systems: Conformational flooding. *Phys. Rev. E* **52**:2893–2906.

Grubmüller, H., H. Heller, A. Windemuth, and K. Schulten. 1991. Generalized Verlet algorithm for efficient molecular dynamics simulations with long-range interactions. *Mol. Simul.* **6**:121–142.

Grubmüller, H., B. Heymann, and P. Tavan. 1996. Ligand binding: Molecular mechanics calculation of the streptavidin-biotin rupture force. *Science* **271**:997–999.

Grubmüller, H., and P. Tavan. 1998. Multiple time step algorithms for molecular dynamics simulations of proteins: How good are they? *J. Comp. Chem.* **19**:1534–1552.

Gullingsrud, J., R. Braun, and K. Schulten. 1999. Reconstructing potentials of mean force through time series analysis of steered molecular dynamics simulations. *J. Comp Phys.* **151**:190–211.

Hayward, S., A. Kitao, and N. Go. 1994. Harmonic and anharmonic aspects in the dynamics of BPTI: A normal mode analysis and principal component analysis. *Physica Scripta* **3**:936–943.

Heidmann, T., and J.-P. Changeux. 1982. Un modèle moléculaire de régulation d'efficacité d'un synapse chimique au niveau postsynaptique. *C.R. Acad. Sci. Paris* **295**:665–670.

Heller, H., H. Grubmüller, and K. Schulten. 1990. Molecular dynamics simulation on a parallel computer. *Mol. Simul.* **5**:133–16.

Herz, J.M., D.A. Johnson, and P. Taylor. 1989. Distance between the agonist and noncompetitive inhibitor sites on the nicotinic acetylcholine receptor. *J. Biol. Chem.* **264**:12,439–12,448.

Heymann, B., and H. Grubmüller. 1999. ANO2-Hapten unbinding forces studied by molecular dynamics AFM simulations. *Chem. Phys. Lett.* **303**:1–9.

Hu, X., T. Ritz, A. Damajanovic, and K. Schulten. 1997. Pigment organization and transfer of electronic excitation in the purple bacteria. *J. Phys. Chem. B* **101**:3854–3871.

Hucho, F., V.I. Tsetlin, and J. Machold. 1996. The emerging three-dimensional structure of a receptor, the nicotinic acetylcholine receptor. *Eur. J. Biochem.* **239**:539–557.

Humphreys, D.D., R.A. Friesner, and B.J. Berne. 1995. Simulated annealing of a protein in a continuum solvent by multiple-time-step molecular dynamics. *J. Phys. Chem.* **99**:10,674–10,685.

Hwang, J.-K., and A. Warshel. 1996. How important are quantum mechanical nuclear motions in enzyme catalysis? *J. Am. Chem. Soc.* **118**:11,745–11,751.

Iwata, S., J.W. Lee, K. Okada, J.K. Lee, M. Iwata, B. Rasmussen, T.A. Link, S. Ramaswamy, and B.K. Jap. 1998. Complete structure of the 11-subunit bovine mitochondrial cytochrome bc1 complex. *Science* **281**:64–71.

Izrailev, S., A.R. Crofts, E.A. Berry, and K. Schulten. 1999. Steered molecular dynamics simulation of the Rieske subunit motion in the cytochrome bc1 complex. *Biophys. J.* **77**: 1753–1768.

Izrailev, S., S. Stepaniants, M. Balsera, Y. Oono, and K. Schulten. 1997. Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophys. J.* **72**:1568–1581.

Kale, L., R. Skeel, M. Bhandarkear, R. Brunner, A. Gursoy, N. Krawelz, J. Phillips, A. Shinozaki, K. Varadarajan, and K. Schulten. 1999. NAMD2: Greater capability for parallel molecular dynamics. *J. Comp. Phys.* **151**:283–312.

Karplus, M., and J.N. Kushick. 1981. Method for estimating the configurational entropy of macromolecules. *Macromolecules* **14**:325–332.

Kim, H., D. Xia, C.-A. Yu, J.-Z. Xia, A.M. Kachurin, L. Zhang, L. Yu, and J. Deisenhofer. 1998. Inhibitor binding changes domain mobility in the iron-sulfur protein of the mitochondrial bc1 complex from bovine heart. *Proc. Natl. Acad. Sci. USA* **95**:8026–8033.

Leggett, A.J., S. Chakravarty, A.T. Dorsey, M.P.A. Fisher, A. Garg, and W. Zwerger. 1987. Dynamics of the dissipative two-state system. *Rev. Mod. Phys.* **59**:2–86.

Littell, T., R. Skeel, and M. Zhang. 1997. Error analysis of symplectic multiple time stepping. *SIAM J. Numerical Analysis* **34**:1792–1807.

Lu, H., B. Isralewitz, A. Krammer, V. Vogel, and K. Schulten. 1998. Unfolding of titin immunoglobin domains by steered molecular dynamics simulation. *Biophys. J.* **75**:662–671.

Luty, B.A., I.G. Tironi, and W.F. van Gunsteren. 1995. Lattice-sum methods for calculating electrostatic interactions in molecular simulations. *J. Chem. Phys.* **103**:3014–3021.

Luty, B.A., and W.F. van Gunsteren. 1996. Calculating electrostatic interactions using the particle–particle particle–mesh method with non-periodic long-range interactions. *J. Phys. Chem.* **100**:2581–2587.

Marrink, S.-J., O. Berger, P. Tieleman, and F. Jähnig. 1998. Adhesion forces of lipids in a phospholipid membrane: Studies by molecular dynamics simulations. *Biophys. J.* **74**:931–943.

Marszalek, P.E., H. Lu, H. Li, M. Carrion-Vazquez, A.F. Oberhauser, K. Schulten, and J.M. Fernandez. 1999. Mechanical unfolding intermediates in titin modules. *Nature* **402**: 100–103.

Martorana, V., D. Bulone, P.L. San Biagio, M.B. Palma-Vittorelli, and M.U. Palma. 1997. Collective properties of hydration: Long range and specificity of hydrophobic interactions. *Biophys. J.* **73**:31–37.

Martorana, V., G. Corongiu, and M.U. Palma. 1996. Correlated solvent-induced forces on a protein at single residue resolution: Relation to conformation, stability, dynamics, and function. *Chem. Phys. Lett.* **254**:292–301.

Martorana, V., G. Corongiu, and M.U. Palma. 1998. Interaction of explicit solvent with hydrophobic/philic/charged residues of a protein: Residue character vs. conformational context. *Proteins* **32**:129–135.

McMahon, B.J., J.D. Müller, C.A. Wraight, and G.U. Nienhaus. 1998. Electron transfer and protein dynamics in the photosynthetic reaction center. *Biophys. J.* **74**:2567–2587.

Molteni, C., and M. Parrinello. 1998. Glucose in aqueous solution by first principles molecular dynamics. *J. Am. Chem. Soc.* **120**:2168–2171.

Monod, J., J.-P. Changeux, and F. Jacob. 1963. Allosteric proteins and cellular control systems. *J. Mol. Biol.* **6**:306–320.

Muegge, I., P.X. Qi, A.J. Wand, Z.T. Chu, and A. Warshel. 1997. The reorganization energy of cytochrome c revisited. *J. Phys. Chem. B* **101**:825–836.

Nienhaus, G.U., K. Chu, and K. Jesse. 1998. Structural heterogeneity and ligand binding in carbonmonoxy myoglobin crystals at cryogenic temperatures. *Biochemistry* **37**:6819–6823.

Novella, M., and K. Schulten. 1991. Molecular dynamics simulation of electron transfer in proteins: Theory and application to QA → QB transfer in the photosynthetic reaction center. *J. Phys. Chem.* **95**:2059–2067.

Oesterhelt, D., J. Titov, and E. Bamberg. 1992. A unifying concept for ion translocation in retinal proteins. *J. Bioenerg. Biomembr.* **24**:181–191.

Olson, W.K. 1996. Simulating DNA at low resolution. *Curr. Opin Struct. Biol.* **6**:242–256.

Olson, J.S., and Q.H. Gibson. 1972. The reaction of n-butyl isocyanide with human hemoglobin. *J. Biol. Chem.* **247**:1713–1726.

Orešic, M., and D. Shalloway. 1994. Hierarchical characterization of energy landscapes using Gaussian packet states. *J. Chem. Phys.* **101**:9844–9857.

Paas, Y. 1998. The macro- and microarchitectures of the ligand-binding domain of glutamate receptors. *Trends Neurosci.* **21**:117–125.

Parak, F., E.N. Frolov, A.A. Kononenko, R.L. Mössbauer, V.I. Goldanskii, and A.B. Rubin. 1980. Evidence for a correlation between the photoinduced electron transfer and dynamic properties of the chromatophore membranes from Rhodospirillum rubrum. *FEBS Lett.* **117**:368–372.

Parrinello, M. 1997. From silicon to RNA: The coming of age of ab initio molecular dynamics. *Solid State Commun.* **102**:107–120.

Parson, W.W., Z.T. Chu, and A. Warshel. 1998. Reorganization energy of the initial electron-transfer step in photosynthetic bacterial reaction centers. *Biophys. J.* **74**:182–191.

Perman, B., V. Srajer, Z. Ren, T.-Y. Teng, C. Pradervand, T. Ursby, D. Bourgeois, F. Schotte, M. Wulff, R. Kort, K. Hellingwerf, and K. Moffat. 1998. Energy transduction on the nanosecond timescale: Early structural events in a xanthopsin photocycle. *Science* **279**:1946–1950.

Perutz, M.F. 1970. Stereochemistry of cooperative effects in hemoglobin. *Nature* **228**:726–739.

Polisky, B. 1998. ColE1 replication control circuiting: Sense from antisense. *Cell* **55**:929–932.

Pyle, A.M., and J.B. Green. 1995. RNA folding. *Curr. Opin. Struct. Biol.* **5**:303–310.

Radmer, R.J., and P.A. Kollmann. 1997. Free energy calculation methods: A theoretical and empirical comparison of numerical errors and a new method for qualitative estimates of free energy changes. *J. Comp. Chem.* **18**:902–919.

Rivetti, C., A. Mozzarelli, G.L. Rossi, E.R. Henry, and W.A. Eaton. 1993. Oxygen binding in single crystals of hemoglobin. *Biochemistry* **32**:2888–2906.

Roitberg, A., R.B. Gerber, R. Elber, and M.A. Ratner. 1996. Anharmonic wave functions of proteins: Quantum self-consistent field calculations of BPTI. *Science* **268**:1319–1322.

San Biagio, P.L., D. Bulone, V. Martorana, M.B. Palma-Vittorelli, and M.U. Palma. 1998. Physics and biophysics of solvent induced forces: Hydrophobic interactions and context-dependent hydration. *Eur. Biophys. J.* **27**:183–196.

Schlick, T. 1995. Modeling superhelical DNA: Recent analytical and dynamic approaches. *Curr. Opin. Struct. Biol.* **5**:245–262.

Schlick, T., E. Bartha, and M. Mandziuk. 1997. Biomolecular dynamics at long timesteps: Bridging the timescale gap between simulation and experimentation. *Ann. Rev. Biophys. Biomol. Struct.* **26**:181.

Schlick, T., R. Skeel, A. Brünger, L. Kale, J. Board, J. Hermans, and K. Schulten. 1999. Algorithmic challenges in computational molecular biophysics. *J. Comp. Phys.* **151**:9–48.

Schulten, K. 1995. Curve crossing in a protein: Coupling of the elementary quantum process to motions of the protein. In: Proc. Ecole de Physique des Houches, ed. D. Bicout and M.J. Field, Les Editions de Physique, pp. 85–118. New York: Springer.

Schulten, K., W. Humphrey, I. Logunov, M. Sheves, and D. Xu. 1995. Molecular dynamics studies of bacteriorhodopsin's photocycles. *Isr. J. Chem.* **35**:447–464.

Shoichet, B.K., W.A. Baase, R. Kuroki, and B.W. Matthews. 1995. A relationship between protein stability and protein function. *Proc. Natl. Acad. Sci. USA* **92**:452–456.

Stanton, R.V., D.S. Hartsough, and K.J. Merz, Jr. 1993. Calculation of solvation free energies using a density functional/molecular dynamics coupled potential. *J. Phys. Chem.* **97**:11,868–11,870.

Treutlein, H., K. Schulten, C. Neidermeier, J. Deisenhofer, H. Michel, and D. Devault. 1988. Electrostatic control of electron transfer in the photosynthetic reaction center of *Rhodopseudomonas viridis*. In: The Photosynthetic Bacterial Reaction Center: Structure and Dynamics, ed. J. Bredon and A. Vermeglio, NATO ASI Series A 149, pp. 139–150. New York: Plenum.

Tuckerman, M.E., G.J. Martyna, and B.J. Berne. 1990. Molecular dynamics algorithm for condensed systems with multiple time scales. *J. Chem. Phys.* **93**:1287–1291.

Warshel, A. 1984. Dynamics of enzymatic reactions. *Proc. Natl. Acad. Sci. USA* **81**:444–448.

Warshel, A. 1991. Computer Modeling of Chemical Reactions in Enzymes and Solutions. New York: John Wiley.

Warshel, A., Z.T. Chu, and W.W. Parson. 1989. Dispersed polaron simulations of electron transfer in photosynthetic reaction centers. *Science* **246**:112–116.

Warshel, A., and J.-K. Hwang. 1986. Simulation of the dynamics of electron transfer reactions in polar solvents: Semiclassical trajectories and dispersed polaron approaches. *J. Chem. Phys.* **84**:4938–4957.

Warshel, A., and M. Levitt. 1976. Theoretical studies of enzymatic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **103**:227–249.

Warshel, A., and R.M. Weiss. 1981. Energetics of heme–protein interactions in hemoglobin. *J. Am. Chem. Soc.* **103**:446–451.

Wesolowski, T., R.P. Muller, and A. Warshel. 1996. Ab Initio frozen density functional calculations of proton transfer reactions in solution. *J. Phys. Chem.* **100**:15,444–15,449.

Xia, D., C.A. Yu, J.Z. Xian, A.M. Kachurin, L. Zhang, L. Yu, and J. Deisenhofer. 1997. Crystal structure of the cytochrome bc1 complex from bovine heart mitochondria. *Science* **277**:60–66.

Xu, D., and K. Schulten. 1994. Coupling of protein motion to electron transfer in a photosynthetic reaction center: Investigating the low-temperature behavior in the framework of the spin-boson model. *Chem. Phys.* **182**:91–117.

Zarrinkar, P.P., and J.R. Williamson. 1994. Kinetic intermediates in RNA folding. *Science* **265**:918–924.

Zhang, Z.L., L.S. Huang, V.M. Shulmeister, Y.I. Chi, K.K. Kim, L.W. Hung, A.R. Crofts, E.A. Berry, and S.H. Kim. 1998. Electron transfer by domain movement in cytochrome bc1. *Nature* **392**:677–684.

Zhou, R., and B.J. Berne. 1995. A new molecular dynamics method combining the reference system propagator algorithm with a fast multipole method for simulating proteins and other complex systems. *J. Phys. Chem.* **103**:9444–9459.