Molecular dynamics simulations in biology

Martin Karplus & Gregory A. Petsko

Molecular dynamics—the science of simulating the motions of a system of particles—applied to biological macromolecules gives the fluctuations in the relative positions of the atoms in a protein or in DNA as a function of time. Knowledge of these motions provides insights into biological phenomena such as the role of flexibility in ligand binding and the rapid solvation of the electron transfer state in photosynthesis. Molecular dynamics is also being used to determine protein structures from NMR, to refine protein X-ray crystal structures faster from poorer starting models, and to calculate the free energy changes resulting from mutations in proteins.

SCIENTIFIC theory has as its ultimate goal the understanding and prediction of natural phenomena. In physics and chemistry, theory and experiment have worked well together to interpret the results of measurements and to provide predictions that aid in their design. Biology has been slower to embrace theoretical approaches, in part because of a certain scepticism as to the validity of the assumptions and methods that are applicable to large and complex systems, such as the macromolecules of biological interest, not to mention organelles or cells. The most successful theories make use of a combination of analytical and computational approaches. They have given rise to 'theoretical biophysics', a rapidly growing field that is having a significant impact on many areas of biological research.

Supercomputer availability has contributed to the development of this field, but the main reason for its growth is the introduction and application of molecular dynamics simulation methods[1,2]. Biological macromolecules are inherently dynamic systems, and it was the demonstration that calculations based on simple physical models could give new, and testable, insights into their internal motions that convinced many biochemists that biophysical theory had become relevant. Dynamical simulations have now been extended to include not only molecular motion, but also the refinement of macromolecular structures based on X-ray and NMR data and the sampling of configuration space required for the evaluation of free energy changes induced by mutations and other perturbations.

Macromolecular motions cover a large range in magnitude (from hundredths of an ångström to tens of ångströms), and an enormous range of time (from sub-picoseconds to seconds and even longer). Some of these motions are known to have a functional role. Oxygen cannot reach the binding site in myoglobin or haemoglobin unless the protein structure fluctuates to open a transient pathway. Hexokinase transfers a phosphate from ATP to glucose much faster than it transfers one to water because the binding of glucose induces a large conformational change in the protein that brings two domains together and creates a protected catalytic site. The domain closure is slow, but it would not be possible without the numerous rapid atomic motions that allow groups of atoms to slide past each other despite their repulsive interactions. There are many other examples in biology in which dynamics is involved, either as the essential element of the phenomenon or as a secondary element (that is, the timescale of the motion is not important). Molecular dynamics simulations are important for understanding these motions, particularly because detailed experimental studies are often very difficult.

This review focuses on the methodology of molecular dynamics and its applications to biological problems. We begin with an overview of the methodology, including some remarks on its experimental validation, and continue with an account of some important applications. We focus on proteins in this review, although similar but less extensive studies have been made for nucleic acids, oligosaccharides and lipids. We close with a discussion of the present limitations of molecular dynamics simulations and some comments on likely future developments.

What is molecular dynamics?

Molecular dynamics is the science of simulating the motions of a system of particles. It has been applied to systems as small as an atom and a diatomic molecule undergoing a chemical reaction, and as large as a galaxy[3]. In all cases, the essential elements for a molecular dynamics simulation are a knowledge of the interaction potential for the particles, from which the forces can be calculated, and of the equations of motion governing the dynamics of the particles. The interaction potential may vary from the simple gravitational interaction between stars to the complex many-body forces between atoms and molecules. Classical newtonian equations of motion are adequate for many systems, including the biomolecules of primary concern here, but for some problems (such as reactions involving tunnelling) quantum corrections are important, and for others (such as galaxy evolution) relativistic effects may have to be included.

Two attributes of molecular dynamics simulations have played an essential part in their explosive development and wide range of applications. Simulations provide individual particle motions as a function of time so they can be probed far more easily than experiments to answer detailed questions about the properties of a system. Further, although the potential used in a simulation is approximate, it is completely under the user's control, so that by removing or altering specific contributions their role in determining a given property can be examined. 'Computer alchemy'—changing the potential from that representing one system to another during a simulation—is a powerful tool for calculating free energy differences.

Potential energy function

Molecular dynamics simulations begin with a knowledge of the energy of the system as a function of the atomic coordinates. The potential energy surface determines the relative stabilities of the different possible stable or metastable structures. The forces acting on the atoms of the system, which are related to the first derivatives of the potential with respect to the atom positions, can be used to calculate the dynamic behaviour of the system by solving Newton's equations of motion for the atoms as a function of time.

Although quantum mechanical calculations can yield potential surfaces for small molecules, only empirical energy functions can provide this information for proteins and their environment. As at ordinary temperatures most motions leave the bond lengths and angles of polypeptide chains near their equilibrium values, the accuracy of the energy function representation of bonding is comparable to that of vibrational analyses of small molecules. In globular proteins, contacts between non-bonded atoms also make a large contribution to the potential energy of the folded or native structure. The success of hardsphere non-bonded radii in conformational studies suggests that relatively simple functions may adequately describe them.

The energy functions used for proteins are generally composed of bonding terms representing bond lengths, bond angles, and torsional angles, and non-bonding terms consisting of van der Waals interactions and electrostatic contributions. One widely used expression[4] is:

$$E(\mathbf{R}) = \tfrac{1}{2}\sum_{\text{bonds}} K_b(b-b_0)^2 + \tfrac{1}{2}\sum_{\substack{\text{bond}\\\text{angles}}} K_\Theta(\Theta-\Theta_0)^2 + \tfrac{1}{2}\sum_{\text{torsional}} K_\phi[1+\cos(n\phi-\delta)] + \sum_{\text{nb pairs}}\left(\frac{A}{r^{12}}-\frac{B}{r^6}+\frac{q_1q_2}{Dr}\right) \quad (1)$$

The energy, E, is a function of the Cartesian coordinate set, R, specifying the positions of all the atoms, from which are calculated the internal coordinates for bond lengths (b), bond angles (Θ), dihedral angles (ϕ) and interparticle distances (r).

The first term in equation (1) represents instantaneous displacements from the ideal bond length, b_0, by a Hooke's law (harmonic) potential. Such a harmonic potential is the first approximation to the energy of a bond as a function of its length. The bond force constant K_b determines the flexibility of the bond and can be evaluated from infrared stretching frequencies or quantum mechanical calculations. Ideal bond lengths can be inferred from high resolution, low temperature crystal structures or microwave spectroscopy data. The energy associated with alteration of bond angles given by the second term in equation (1) is also represented by a harmonic potential. For rotations about bonds, torsion angle potential functions given by the third term in equation (1) are used. This potential is assumed to be periodic and modelled by a cosine or sum over cosine functions. The final term in equation (1) represents the contribution of non-bonded interactions and has three parts: a repulsive term preventing atoms from interpenetrating at very short distances; an attractive term accounting for the London dispersion forces between atoms; and an electrostatic term that is attractive or repulsive depending on whether the charges q_1 and q_2 are of opposite or the same sign. The first two non-bonded terms combine to give the familiar Lennard-Jones 6–12 potential, which has a minimum at an interatomic separation equal to the sum of the van der Waals radii of the atoms; parameters A and B depend on the atoms involved and have been determined by a variety of methods, including non-bonding distances in crystals and gas-phase scattering measurements.

Electrostatic interactions between pairs of atoms are represented by a Coulomb potential with D the effective dielectric function for the medium and r the distance between the two charges. Use of atomic partial charges avoids the need for a separate term to represent the hydrogen bond interaction; that is, when the positive hydrogen attached to an electronegative atom comes within van der Waals distance of a negative acceptor atom, the Coulomb attraction adds to the Lennard-Jones potential and results in a hydrogen bond.

The usefulness of empirical energy functions depends on the extent to which the parameters determined for equation (1) by the study of model systems, such as amino acids, can be employed for macromolecules, such as proteins. Evidence from a number of comparisons suggest that this transferability condition is satisfied in many applications.

Simulation methods

Given a potential energy function there are several approaches to the dynamics. The most exact and detailed information is provided by molecular dynamics simulations in which Newton's equations of motion are solved for the atoms of the system and any surrounding solvent. For a simple homogeneous system, such as a box of water molecules with periodic boundary conditions, average structural and dynamic properties can be determined in simulations of only a few picoseconds[5]. Inhomogeneous systems such as proteins require considerably longer simulations. Modern computers allow simulations of up to nanoseconds, long enough to completely characterize the librations of small groups in a protein and to determine the dominant contributions to atomic fluctuations[1].

To begin a dynamic simulation an initial set of atomic coordinates and velocities is required. The coordinates can be obtained from X-ray crystallographic or NMR structure data, or by model-building (based on the structure of a homologous protein, for example). Given a set of coordinates, a preliminary calculation serves to equilibrate the system. The structure is first refined using an iterative minimization algorithm to relieve local stresses due to overlaps of non-bonded atoms, bond length distortions, and so on. Next, atoms are assigned velocities ($\mathbf{v}$) taken at random from a maxwellian distribution for a low temperature, and a simulation is performed for a few picoseconds. This is done by finding the acceleration $\mathbf{a}_i$ of atom i from Newton's law $\mathbf{F}_i = m_i\mathbf{a}_i$. ($\mathbf{F}_i$, the force on the atom, is computed from the derivatives of equation (1) with respect to the position; m_i is the atomic mass), and introducing it into the equation for the

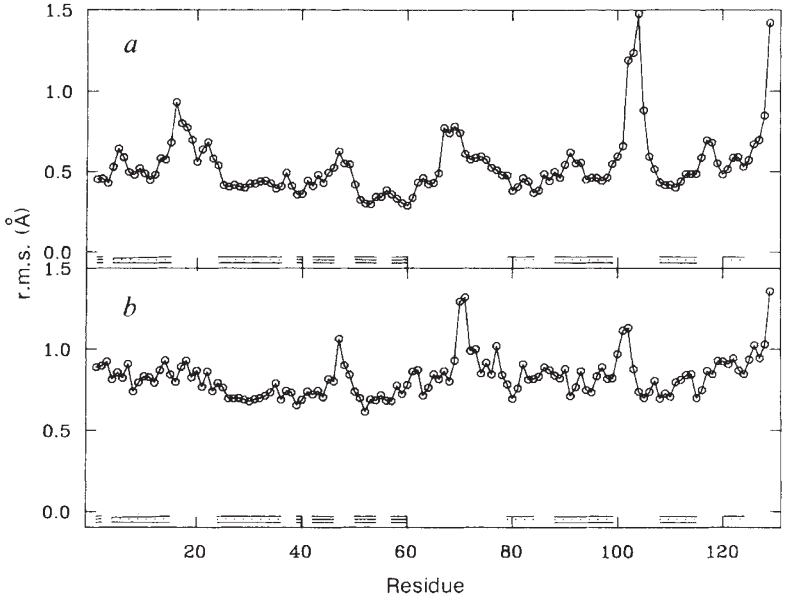

FIG. 1 Calculated and experimental r.m.s. fluctuations of lysozyme. Backbone averages are shown as a function of residue number. They were obtained *a*, from a molecular dynamics simulation, and *b*, from X-ray temperature factors without correcting for disorder contributions (from ref. 28).

position $\mathbf{r}_i$ at time $t+\Delta t$, given $\mathbf{r}_i$ at time t:

$$\mathbf{r}_i(t+\Delta t)=\mathbf{r}_i(t)+\mathbf{v}_i\Delta t+\tfrac{1}{2}\mathbf{a}_i(\Delta t)^2+\cdots \quad (2)$$

The equilibration is continued by alternating new velocity assignments, chosen from maxwellian distributions for temperatures that are successively increased to some chosen value, with intervals of dynamical relaxation. The temperature T of the system is measured by the mean kinetic energy,

$$\tfrac{1}{2}\sum_{i=1}^{N} m_i\langle v_i^2\rangle=\tfrac{3}{2}Nk_BT \quad (3)$$

where N is the number of atoms in the system, $\langle v_i^2\rangle$ is the average velocity squared of the ith atom and k_B is the Boltzmann constant. The equilibration period is considered finished when the temperature is stable for longer than about 10 ps, the atomic momenta obey a maxwellian distribution and different regions of the protein have the same average temperature.

Continued integration of the equations of motion after equilibration generates the coordinates and velocities of the atoms as a function of time. With modern computers, equation (2) can be solved simultaneously for all atoms in a solvated macromolecule by use of a number of standard programs[4]. The quantity Δt must be very small so that the potential energy does not change too much during each time step; a satisfactory value for Δt is one femtosecond (10^{-15} s), which corresponds to about 30 steps for one vibration of a carbon–carbon bond. Even on a supercomputer such as the Cray X-MP a single time-step in the simulation of myoglobin (which has roughly 1,500 atoms) takes 0.2 s of CPU time, so that the 100,000 time-steps of a 100-ps simulation would take about 6 h of supercomputer time. The time increases roughly linearly with the number of atoms in the system.

For large systems the required computer time can be almost prohibitive. Nowhere is this problem more manifest than in the treatment of the solvent. Proteins and nucleic acids have evolved to function in an aqueous environment; large numbers of water molecules are tightly bound to their surfaces, forming a first hydration shell that is an integral part of the structure and should be included in the simulation. That can add hundreds of atoms to the computation. Bulk solvent is an even more serious problem—if not included the simulation is '*in vacuo*', which may not affect interior atoms but is likely to have a serious effect on the conformations and motions of surface residues, especially those in exposed loops. Solvent must be included in simulations concerned with the quantitative analysis of thermodynamic properties, particularly of those portions of the protein at the solvent interface. For myoglobin, a satisfactory solvated simulation system includes about 1,000 water molecules (3,000 atoms) with a corresponding increase in computer time.

One strategy commonly used for simplifying the treatment of such large systems is to restrict the computation to a subset of atoms close to the site of interest (such as the active site of an enzyme). The atoms in a surrounding 'buffer' region are treated as stochastic particles, and all other atoms in the system are eliminated; their effect is approximated by boundary forces, frictional drag and stochastic terms acting on atoms in the buffer-region[1]. This stochastic boundary molecular dynamics approach can reduce the computing time by a factor of 10 to 100, relative to a fully solvated simulation.

Experimental verification

For homogeneous systems, such as a box of water molecules, detailed verification of the structural and dynamic behaviour predicted by simulations has been possible by comparison with experiment[5], but this is more difficult for the inhomogeneous systems of biology. Nevertheless, the approximate agreement between the average structure obtained in the first protein simulation[6], that of the bovine pancreatic trypsin inhibitor (BPTI), and the structure determined by X-ray crystallography suggested that such simulations were stable and that meaningful results could be obtained.

The fast motions sampled by the simulations leave their trace in X-ray diffraction data as they cause atoms to occupy more space on average than in a rigid molecule, smearing out the electron density, the quantity determined in crystallographic studies. If a gaussian model is used to fit the spread of density around the average position of each atom, the width of the gaussian can be expressed in terms of a 'Debye–Waller factor', or 'isotropic temperature factor', B, of the form:

$$\langle\Delta r_i^2\rangle=3B_i/8\pi^2 \quad (4)$$

where $\langle\Delta r_i^2\rangle$ is the mean-square fluctuation of atom i. Comparisons of $\langle\Delta r_i^2\rangle$ values from molecular dynamics simulations and measured B factors show good qualitative agreement (Fig. 1). Larger deviations occur in regions of the structure where molecules make contact with neighbours in the crystal lattice: contact residues have systematically low experimental B factors. Surface side chains protruding from the protein also show a less good fit, but the discrepancy is reduced if the simulation is carried out in the presence of solvent and a crystalline environment[7].

Internal motions of proteins can be studied also by NMR, which unlike X-ray diffraction is sensitive to the timescale as well as the magnitude of the motions[8]. NMR parameters such as nuclear spin–spin coupling and chemical shift are affected by molecular motion (Fig. 2). If interconversion of different

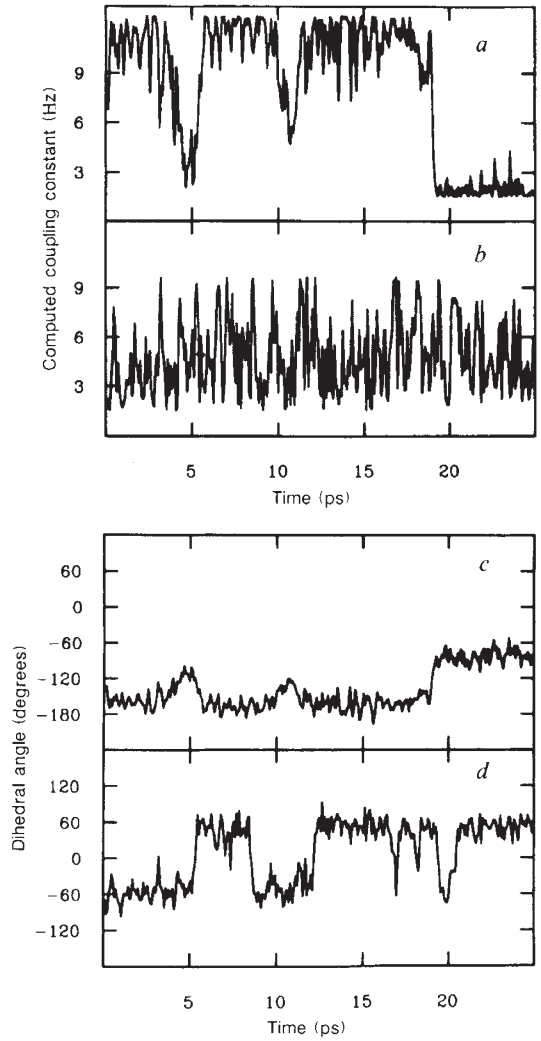

FIG. 2 Fluctuations of three-bond proton coupling constants and associated dihedral angles during a molecular dynamics simulations of BPTI. *a*, Glu 7 $^3J_{\beta\gamma}$; *b*, Arg 42 $^3J_{\beta\gamma}$; *c*, χ_2 of Glu 7; *d*, χ_2 of Asp 42 (from ref. 29).

local conformations is rapid on the NMR timescale, average values are obtained. For motions on the NMR timescale, transition rates can be determined directly by NMR. For example, aromatic residues are sometimes observed to 'flip' by NMR (on the timescale of milliseconds) and in long or activated dynamics simulations. That such ring flips occur, even in the tightly packed interior of proteins, came as a surprise to X-ray crystallographers and provided an early example of the importance of packing defects and correlated motions[8,9].

The ring flips constitute a rare event that is measurable because the result of flips (averaging of certain chemical shifts) has a significant effect on the NMR spectrum. One can measure even less frequent events if they are essential to the phenomenon under consideration. A case in point is hydrogen exchange, which may occur on a timescale of minutes or hours or even longer. Some hydrogens in the interior of proteins apparently can exchange only if a relatively large fluctuation takes place. Even though this happens only very rarely for a protein under physiological conditions, the rate can be measured when it controls a measurable phenomenon such as hydrogen exchange. Such complex events are still outside the range of what is directly accessible to molecular dynamics simulations.

Macromolecular motions

The picture that has emerged for the internal motions of proteins (and nucleic acids) is that the atoms are fluctuating on the picosecond timescale so that the instantaneous structures differ significantly from the time-average. This view has now become part of the dogma of biology and is often introduced in a descriptive fashion to explain a wide range of phenomena, even when there is no evidence that motions play a part.

During simulations the protein samples many conformations close to but not identical with the X-ray crystal structure. Atoms in the interior of proteins move less than atoms on the surface. Most atoms move anisotropically, and the directions of motion are determined more by non-bonded interactions with neighbouring groups than by the nature of the covalent bonds. Root-mean-square (r.m.s.) atomic displacements of 0.5–1 Å are not uncommon in the protein interior, and even larger ones occur for surface residues. The motions are somewhat anharmonic at 300 K, usually due to side chains 'hopping' between two or more alternative minima. A residue executes many small fluctuations about the minimum energy position before hopping to another minimum. As expected for motion involving a barrier the transition is very rapid, so that the side chain spends most of its time in the minima, and it is these positions that are observed in the X-ray structure, as for the atoms of atomatic residues involved in ring flips described above. Below 160 K, simulations indicate that the dynamics are essentially harmonic. These features are confirmed by the temperature dependence of the B factors determined by X-ray crystallography and inelastic neutron scattering[10–12].

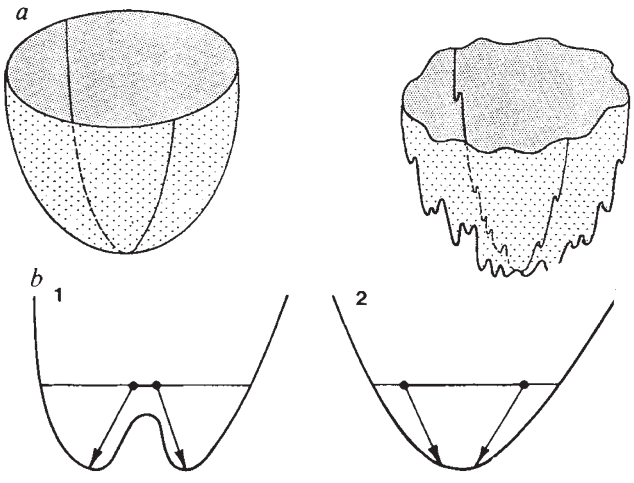

FIG. 3 Model potentials for protein motions. *a*, Two-dimensional representation of harmonic potential (left) and multimimum (substate) potential (right). *b*, representation of the r.m.s. difference criterion for different minima: 1, r.m.s. after the minimization is larger than the initial r.m.s., implying that the two conformations correspond to different minima; 2, r.m.s. after the minimization is smaller than the initial r.m.s., implying that the two conformations correspond to the same minimum (from ref. 13).

1. Myoglobin and haemoglobin. Myoglobin and haemoglobin bind oxygen or carbon monoxide reversibly, yet the haem group is so deeply buried within the protein that there is no 'path' from outside the molecule to the ligand binding site. In their original papers, Perutz and Kendrew recognized that these proteins must 'breathe' to open channels connecting the surface and haem pocket. In the X-ray structure the energy barriers to ligand migration are high, with a half-time of approximately 10^{50} years required for entrance or exit of a ligand in the rigid protein. When the side chains surrounding the haem pocket are allowed to fluctuate, however, the barriers are substantially reduced and low energy pathways appear. One of these requires partial rotation of the distal histidine, His 64, and takes the ligand past one of the haem propionic acid side chains and out of the molecule in the vicinity of Arg 45, a pathway for which there is some experimental support from structural studies.

The mechanism can be examined in more detail by a molecular dynamics simulation of a ligand migrating into or out of the fluctuating protein. This was done by breaking the ligand–iron bond and giving the ligand a high initial velocity. In the absence of a bond to the iron, the non-bonding van der Waals repulsions prevent rebinding and the ligand starts to explore the distal pocket. Eventually some of the surrounding atoms move out of the way enough to allow the ligand to escape from the pocket and migrate through the protein until it finally reaches the outside.

To delineate the full range of possible pathways would seem to require many ligand plus protein simulations. A simplification is to run one simulation with hundreds of ligand molecules in the haem pocket, each with a different initial velocity[17]. Interaction terms between the ligand molecules are eliminated so that they do not 'feel' each other's presence, but each ligand molecule has the full interaction with the surrounding protein. As it is the dynamics of the protein that takes most time in the simulation, calculation for one myoglobin and many ligand molecules is very efficient. Such a 100-ps simulation of myoglobin with 60 carbon monoxide molecules revealed five principal routes for entrance and exit of a ligand. In all cases the ligand spent a considerable time rattling around inside one or another of the many atom-sized internal cavities known to exist in the protein before jumping over a barrier to a new position, usually another cavity or, finally, the protein exterior (Fig. 5).

Two general models have been considered for the internal motions of proteins, and these are illustrated in Fig. 3*a*. In one model the fluctuations are assumed to occur within a single multidimensional well that is harmonic or quasiharmonic as a limiting case. The other model assumes there are multiple minima or substates, with the internal motions corresponding to a superposition of oscillations within wells and transitions between them. Experimental data have been interpreted according to both models, and it has proved difficult to distinguish between them. The two models can be distinguished by comparing structures sampled in a simulation, and determining whether they correspond to the same or different minima[13]. This is done by minimizing the coordinate sets obtained from the simulation (see Fig. 3*b*). In a 300 ps simulation of myoglobin at 300 K, different minima were accessed every few tenths of a picosecond, consistent with the model in which there are many structurally similar substates in the neighbourhood of the absolute minimum. The large number of minima sampled by the room temperature simulation suggests that the effective barriers separating them are low. As the energy differences between many of the wells are small, individual molecules may be trapped in metastable states at low temperature. This could lead to a glass-like substance.

The backbone structural differences between the minima correspond to loop displacements coupled to changes in the relative orientation of helices which individually behave as nearly rigid bodies. The helix displacements are made possible by the loop rearrangements and by side-chain reorientations at helix–helix contacts. Specific loop motions are initiated by side-chain transitions in the helix contacts, main chain dihedral angle transitions of the loops themselves, or a combination of the two.

The dynamical results for the helix motions can be compared with structural data. The maximum dynamical displacements in the simulation are larger than those observed in different X-ray structures of the same protein and are of the same order as the differences found in the known X-ray structures of various globins. A comparison of the globin structures suggested that the observed range of helix packings is achieved primarily by changes in side chain volumes resulting from amino acid substitutions[14,15]. In the dynamics, it is the correlated motions of side chains that are in contact, plus the rearrangements of loops, that make possible the observed helix fluctuations. As more than one set of side-chain orientations is consistent with a given set of helix positions, the known globin crystal structures represent only a very small subset of the possible local minima. Overall, the range of conformations sampled by a single myoglobin trajectory at 300 K is similar to that found in the evolutionary variation among structures of the globin series (Fig. 4). This 'molecular plasticity' is likely to have played a part in the evolution of protein sequences.

Functionally relevant motions

It would not be surprising if internal motions had been subjected to selective pressure during evolution. Just as structure is selected on the basis of function, there could be selection for certain internal motions, a consequence of the structure, if they had specific functional roles. Two examples of functionally relevant motions, one in globins and the other in photosynthetic reaction centres, are described in Boxes 1 and 2.

The myoglobin simulations (Box 1) illustrate the general result that individual dynamic events in biology that are relatively slow on a macroscopic scale (having, for example, a microsecond rate constant) actually occur very rapidly on a microscopic scale (in, for example, a picosecond). The much slower macroscopic rate is due to the rarity of the microscopic event resulting from the presence of energy barriers. Events involving a few particles are limited by activation barriers if the rate constant is on the order of 10^{10} s^{-1} or less.

The general permeability of proteins to ligands suggested by the myoglobin simulation is supported by several experiments, particularly aromatic side-chain fluorescence quenching and amide hydrogen exchange data. Observations that solvent molecules can quench the fluorescence of even deeply buried groups and that amide protons completely inaccessible to solvent in the static protein structure exchange with solvent hydrogens at a relatively high rate can be explained either by assuming the protein unfolds and refolds or that solvent can penetrate rapidly into the protein interior. Molecular dynamics simulations suggest that the latter mechanism is correct, at least at temperatures much below that at which denaturation occurs. At higher temperatures transient denaturation can be important.

The primary event of photosynthesis is another biologically important phenomenon where the motions of protein atoms seem to be essential[18]. Photosynthetic bacteria trap solar energy in pigments bound to light-harvesting antenna proteins. The resulting excitation is transferred to a 'special pair' of bacterial chlorophylls. The excited electron is ejected from the special pair and rapidly transferred to a bacterial pheophytin molecule 17 Å away, from whence it moves on to a quinone. Although many biological processes are slow, the primary event in photosynthesis is observed to occur on the picosecond timescale[19]. This is because light provides the energy necessary to trigger the activated event involved in the electron transfer. Once the charge separation has occurred, the charged species (for example, the positive special pair and negative pheophytin) are stabilized by the femtosecond relaxation of protein polar groups (Box 2).

Structure determination and refinement

No one knows how to predict the native structure of a protein from its amino-acid sequence, unless that sequence has significant homology with another protein of known structure. It is impossible, for example, to take the extended polypeptide chain, run a molecular dynamics simulation, and have it fold automatically into the correct structure. Even if there were enough supercomputer time to run a simulation for the time required to fold up in solution (for example, a 10^{-3} s simulation uses about 10^8 hours of computer time) it is not clear whether the correct structure would be obtained, as the potential functions are not perfectly accurate and the free-energy of denaturation is very small (of the order of 0.01 kcal per atom). If, however, the potential function in equation (1) is supplemented by extra information, such as a subset of the interatomic distance constraints characteristic of the folded molecule, a simulation may be able to find the correct structure in 10 picoseconds. Box 3 describes how this approach can be used with distance constraints from NMR data to determine the structure of a protein[20,21]. Simulations have been used also to show that the internal protein motions, which alter the instaneous interproton

2. Photosynthetic reaction centre. The first electron-transfer step in photosynthetic bacteria has extremely high efficiency, owing to two factors. The very fast rate of forward transfer (about 1 ps) prevents de-excitation from competing with electron transfer. Of equal importance is the slow rate of back transfer (1 ns to 1 μs), which is necessary because the second step of the process, transfer from the pheophytin to the quinone, is relatively slow (200 ps). Molecular dynamics has been used to show that relaxation of the reaction centre proteins stabilizes the separated electron-hole pair and thereby can contribute to the slowing down of the back reaction rate. Such a simulation is made possible by the determination of the high-resolution X-ray structure of the reaction centre from *Rhodopseudomonas viridis*[27]. The pigments, prosthetic groups of transmembrane proteins, are organized in a roughly linear fashion across the membrane with their redox potentials ordered so as to funnel the electrons in the appropriate direction.

Although full molecular dynamics simulations of the reaction centre are very costly due to the very large size of the complex, the stochastic dynamics method described in the text can be used if the assumption is made that transfer of the electron to and from a chromophore produces only local structural changes. Such a simulation has been carried out to compare the charge-separated state with the initial state[18]. No large changes in structure took place during the simulation although there are numerous small adjustments in the atom positions surrounding the charged chromophores. Analysis of the chromophore energies showed that this structural relaxation, complete in femtoseconds, altered the electrostatic interactions between the protein and the chromophores so as to stabilize the special pair and pheophytin in their charged states. With such a stabilization, which corresponds to 'solvation' of the charge pair, analogous to solvation in aqueous solution, the energy matching between the charge transfer and neutral chromophore states required for rapid electron transfer is destroyed and the back transfer rate is reduced sufficiently so that it cannot compete with efficient forward transfer to the quinone. A corresponding low temperature simulation shows that the forces due to the charges are large enough for the adjustment of atomic positions and the resultant stabilization to occur at 10 K, in accord with the experimental result that the electron-transfer efficiency is nearly temperature independent.

The computed change in electrostatic energy is the result of small motions of many residues, rather than the large displacement of one or a few charged side chains. This implies that it will be difficult to confirm the simulation results experimentally either by structural studies or site-specific mutagenesis. The photosynthetic reaction centre thus demonstrates the importance of simulations for understanding functional dynamics.

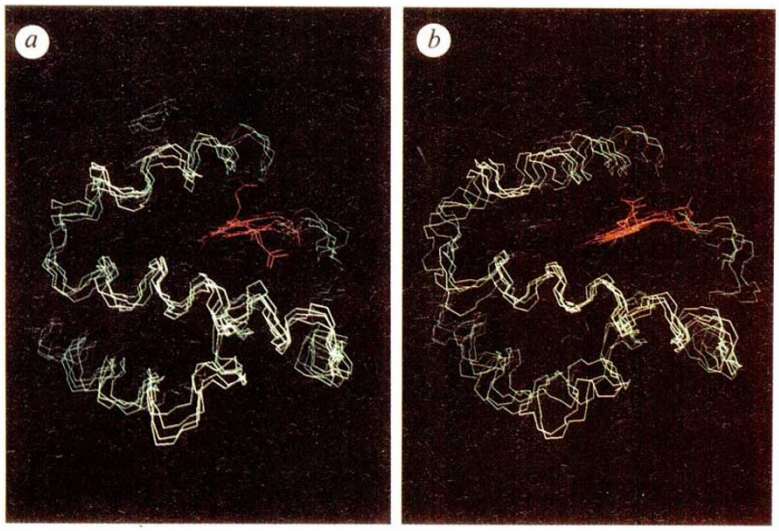

FIG. 4 *a*, Set of structures, shown as backbone traces from the myoglobin minimizations separated by about 10 ps (R. Elber and M. K., unpublished; see also ref. 13). *b*, set of structures, shown as backbone traces, from X-ray data for different globins (see refs 14, 15).

distances, usually introduce only a small error in the interpretion of the experimental data in terms of a rigid structure[8].

In principle any information can be included as the effective potential, $E_{\text{Tot}}(\mathbf{R})$, in a molecular dynamics simulation by way of a constraint term, $E_{\text{const}}(\mathbf{R})$, such that

$$E_{\text{Tot}}(\mathbf{R}) = E(\mathbf{R}) + E_{\text{const}}(\mathbf{R}) \tag{5}$$

where $E(\mathbf{R})$ is the potential energy function in equation (1). One possibility might be theoretical constraints given by rules developed from the study of known protein structures. The simulations would then generate a refined structure consistent with that information. Protein crystallography, whether by X-ray or neutron diffraction, is an area where structure refinement is a major problem and where molecular dynamics with equation (5) has been shown to be able to assist in the refinement process. Box 4 outlines this important application[22,23].

Free-energy simulations

Molecular dynamics methods can also provide a tool for calculating free-energy differences between similar structures. Examples include the determination of the effect of a site-directed mutation on protein stability or the rate of an enzymatic reaction. A related problem concerns the effect of chemical change to a drug molecule on its affinity for the target binding site, where the relevant parameter is the free-energy difference between the bound drug and the drug in solution. Both quantities are changed when the drug is altered so that observed tighter binding could be due to a decreased solvation free energy of the modified drug rather than an increase in the free-energy of binding to the protein target. The two types of contributions to binding and protein folding have been elucidated by free-energy simulations[24]. Of particular importance has been the capability of free-energy simulations to determine the role of solvation and of specific protein residues to the measured free-energy changes, which generally are small and mask the individual components.

It is useful to express the free-energy change in terms of a thermodynamic cycle. For protein stability, for example, we have:

$$\begin{array}{ccc} N_{wt} & \xrightarrow{\Delta G_{wt}(N\to D)} & D_{wt} \\ \Delta G_N(wt\to mut)\downarrow & & \downarrow \Delta G_D(wt\to mut) \\ N_{mut} & \xrightarrow[\Delta G_{mut}(N\to D)]{} & D_{mut} \end{array}$$

where N_{wt}, D_{wt} and N_{mut}, D_{mut} denote the wild-type and mutant proteins in the native and denatured states, respectively, and the various ΔG values in the diagram correspond to the free-energy associated with the changes of state indicated by the arrows. This type of thermodynamic cycle, an example of Hess's law, is analogous to the Born cycle used to determine the solvation energy of an ion. The measured quantity is:

$$\Delta\Delta G = \Delta G_{\text{mut}}(N\to D) - \Delta G_{\text{wt}}(N\to D) \tag{6}$$

but this corresponds to reactions that are difficult to simulate because the change from the native (folded) state to the denatured state is a complex large-scale process. To avoid this

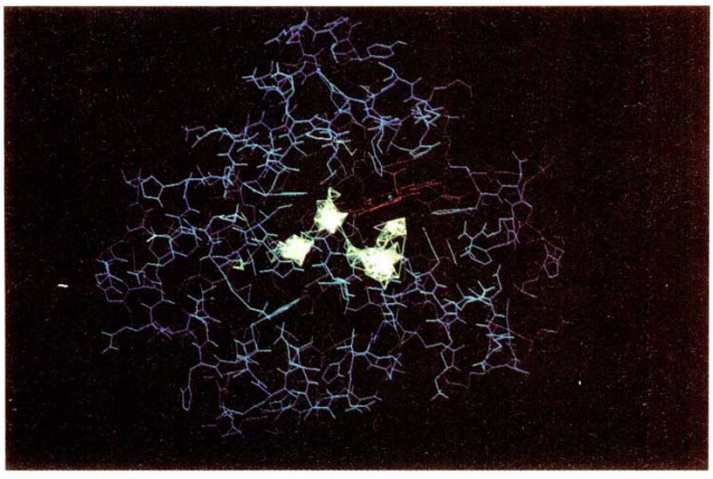

FIG. 5 Trajectory of a CO molecule escaping from myoglobin (R. Elber and M.K., unpublished).

problem, one can resort to the flexibility of the molecular dynamics methodology and do the calculation by following an unphysical path that makes use of 'computer alchemy'. Instead of changing base metal into gold (easily done in a simulation), one amino acid is changed into another. Thus, one calculates:

$$\Delta\Delta G = \Delta G_D(\text{wt} \rightarrow \text{mut}) - \Delta G_N(\text{wt} \rightarrow \text{mut}) \quad (7)$$

which, by the thermodynamic cycle, equals $\Delta\Delta G$ in equation (6).

As the desired quantities are thermodynamic functions, their evaluation by the computer 'experiment' that converts the wild type into the mutant form must be done reversibly. This is accomplished in practice by making the transformation in small increments. The free-energy change, ΔG, from state A to state B is given by:

$$\Delta G = G_B - G_A = \int_0^1 \langle \Delta V \rangle_\lambda \, d\lambda \approx \sum_i \langle \Delta V \rangle_{\lambda_i} \Delta_{\lambda_i} \quad (8)$$

where

$$\Delta V = V_B - V_A \quad (9)$$

and

$$V_\lambda = (1-\lambda) V_A + \lambda V_B \quad (10)$$

Here λ is a fractional coordinate describing the transition from the initial state ($\lambda = 0$) to the final state ($\lambda = 1$), and V_A and V_B are the empirical potential energy functions (see equation (1)) for the wild-type and mutant form, respectively.

The linear potential scaling (equation (10)) and the thermodynamic integration formula (equation (8)) make possible a decomposition of the free energy into contributions from the solvent and individual protein residues[25]. The resulting free-energy differences include the entropic as well as the enthalpic terms. They are obtained by determining the ensemble average of the potential energy differences between A and B, $\langle \Delta V \rangle_\lambda$, for the phase space corresponding to the hybrid system V_λ, and then integrating over the values of λ. By the ergodic theorem the time average behaviour obtained from a molecular dynamics simulation is equivalent to the spatial average over the configurational space in the long-time limit.

Free-energy simulations have been performed for drug–macromolecule interactions and for the effect of site-specific mutations on binding and catalysis, generally with good agreement between computationally and experimentally obtained values of $\Delta\Delta G$. The real power of the approach, however, lies in the possibility of detailed analysis of the results. Comparison of the crystal structure of a mutant protein with that of the wild type can often provide experimental clues as to the origin of the free energy changes. But the structural changes, *per se*, do not provide a direct measure of the thermodynamic quantities of interest[25]. This is particularly true because solvation effects

3. Structure determination. Two-dimensional NMR techniques are able to determine approximate short-range (less than 5 Å) interatomic distances by the nuclear Overhauser effect (NOE). To obtain an NOE the spectroscopist irradiates the protein at a frequency corresponding to the resonance of a particular proton. Resonances arising from protons within 5 Å of the target proton will have their magnitude affected by a through-space (spin dipole–spin dipole) transfer of magnetization. An NOE between two protons is therefore diagnostic of the short distance between them, even if they are far apart in the primary sequence. For a typical small protein of 50 amino acids it may be possible to obtain 300 or more approximate NOE distances to use as constraints in a folding simulation. Many of these will be between residues close together in sequence, which will help to establish the secondary structure of the protein; the others are important for determining the tertiary structure.

The NOE distance information is incorporated into the dynamics simulation as in equation (5) with $E_{const}(\mathbf{R})$ a sum of terms for the NOE distances, each of which is like the bond energy terms in $E(\mathbf{R})$. In the simplest case, a harmonic interaction is used, such that:

$$E_{const}(\mathbf{R}) = E_{NOE}(\mathbf{R}) = \sum_i A_i (R_i - R_{NOE})^2$$

Here R_i is the distance between the ith pair of protons in the dynamics structure, R_{NOE} is their distance estimated from the NOE experiment, and A_i is a 'force constant' chosen to give the forces due to the NOE terms a weight that is similar in magnitude to that of the rest of the empirical potential energy function. Thus, $E_{NOE}(\mathbf{R})$ provides a continuous energy penalty that can be differentiated to obtain a force that tends to zero as the structure satisfies the NOE distance constraints.

This method has been applied to determine solution structures of many proteins. In an early model study of the small (46 residue) protein crambin, the molecular dynamics simulations converged to the known crambin structure from several different initial extended structures (Fig. 6).

4. Structure refinement. The initial model used for solving a protein crystal structure is determined by fitting to an electron density map that usually contains numerous ambiguous regions and some incorrect density. It is possible to improve this model by an iterative procedure that compares the observed X-ray diffraction data with those calculated by Fourier transformation of the assumed structural model. The refinement process is performed by minimzing a residual function, the sum of the squares of the differences between the observed ($|\mathbf{F}_{obs}(h, k, l)|$) and calculated ($|\mathbf{F}_{calc}(h, k, l)|$) structure factor amplitudes. Refinement progress is monitored by the value of the R-factor,

$$R = \sum_{h,k,l} ||\mathbf{F}_{obs}(h, k, l)| - |\mathbf{F}_{calc}(h, k, l)|| \Big/ \sum_{h,k,l} |\mathbf{F}_{obs}(h, k, l)|$$

where h, k, l are the reciprocal lattice points of the crystal and $||$ indicates magnitudes. Protein crystallographers have used non-linear least squares methods to minimize R. As this has a small radius of convergence, repeated cycles of time-consuming manual repositioning of the atoms that are too far from their correct positions and least-squares minimization are required.

The crystallographic residual function is strikingly similar to the NOE constraint term, suggesting that a dynamics calculation carried out with equation (5) and

$$E_{const} = E_{Xray}(\mathbf{R})$$

where

$$E_{Xray}(\mathbf{R}) = S \sum_{h,k,l} (|\mathbf{F}_{obs}(h, k, l)| - |\mathbf{F}_{calc}(h, k, l)|)^2$$

would produce a structure with good stereochemistry that also satisfied the observed X-ray data; the factor S is chosen to balance the two terms in equation (5). Use of molecular dynamics in a simulated annealing mode[22] extends the radius of convergence of this refinement procedure and reduces the amount of manual intervention required. In simulated annealing, the protein is heated to a high temperature (up to 10,000 K) to allow it to traverse a wide range of coordinate space and overcome the local minimum problem before cooling it down to 'anneal' the structure to the lowest energy state. It is easy to apply this methodology in a molecular dynamics simulation as velocity scaling allows the simulation to be carried out at any effective temperature. After minimization of the initial model to remove any bad interatomic contacts, a dynamics simulation with the X-ray constraint term is carried out for several picoseconds at high temperature. The X-ray constraints pull the structure toward agreement with the data, while the energy function prevents it from assuming impossible geometries. The protein samples many different conformations in a few picoseconds. After the dynamic search, the protein is cooled during a further short simulation, and finally the structure is energy-minimized to ensure good stereochemistry. As in the case of the NMR analysis (Box 3), simulated annealing was first tested on crambin and has now been sucessfully applied to the X-ray refinement of numerous proteins[23].

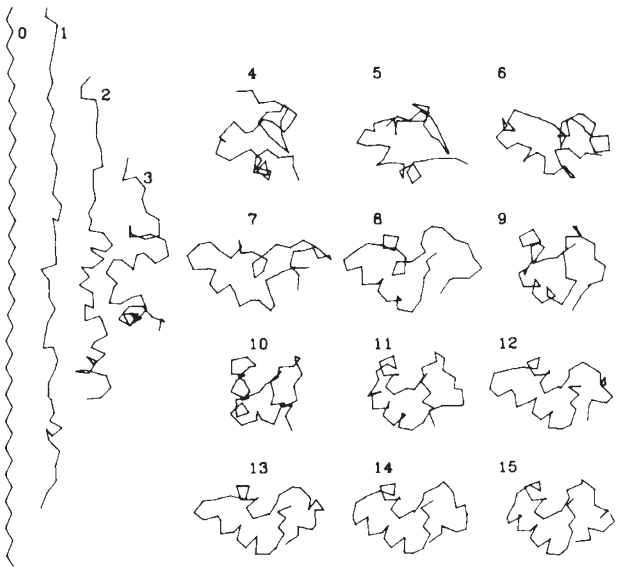

FIG. 6 Folding of crambin backbone by restrained molecular dynamics. Starting with the extended initial structure (0) snapshots are shown at 1-ps intervals; only the Cα atoms are shown (from ref. 20).

are difficult to determine and entropic, as well as energetic, effects can be important. Simulations are necessary to determine which of the structural changes are important and to evaluate the entropic contributions. Insights provided by simulations are illustrated by the effect on cooperativity of the mutation of residue Asp 99 to Ala in the β subunit of haemoglobin (Box 5).

Limitations and prospects

Two limitations in existing simulations are the approximations in the potential energy functions and the lengths of the simulations. The first introduces systematic errors and the second, statistical errors. As the questions that are being asked have changed from those concerned with the qualitative characterization of the dynamics to more quantitative determinations of thermodynamics and kinetics, improved accuracy based on refined potential functions is needed. Much effort is being expended by many groups to improve the form of the potential functions and to obtain more accurate parameters. There is also a shift from using gas-phase experiments and *ab initio* calculations for isolated molecules toward parametrization with emphasis on solution and crystal data. More complex bonding terms are being used and polarization and higher-than-monopole terms are being introduced for the electrostatic interactions. Improved evaluations of the long-range electrostatic contributions are being made by the use of Ewald sums in crystals and multistep algorithms with multipole expansions for inhomogeneous systems. The combination of such refinements with the solution of the Poisson–Boltzman equation[26] may permit the inclusion of the dominant effect of solvation (including ionic strength) and reduce the need for full simulations with explicit solvent. But here, as in other attempts to use continuum models, the frequency dependence (of the dielectric constant, for example) must be considered in dynamic versus thermodynamic applications. For reactions, quantum calculations can be used to obtain the forces on the parts of the system undergoing bond breaking or bond making. Moreover, where necessary, quantum dynamics can be introduced, particularly for thermodynamic (high-frequency vibrations) and reaction studies (tunnelling) where classical approximations to the dynamics break down.

For many problems, simulations of 100 picoseconds are sufficient. Longer simulations would, however, be useful in the study of the relaxation phenomena involved in NMR and fluorescence depolarization, because the experimental time scale is on the order of nanoseconds. Extended simulations would be very useful for characterizing the motions involved so as to make the interpretation of experiments more meaningful. Of equal interest are structural changes (such as the relaxation of myoglobin after photodissociation of carbon monoxide), which could be probed directly by such simulations.

There are some biologically relevant problems at the molecular level to which simulations have so far made only very limited contributions. These include the study of large-scale conformational changes (such as that from the deoxygenated to oxygenated structure in haemoglobin), details of enzymatic kinetics, very rare events such as hydrogen exchange and the 'ultimate' problem of protein folding. In all these it is likely to be necessary, as well as expedient, to replace brute force dynamics by methodologies specialized for the problem under consideration.

For enzymatic reactions, for example, which tend to be localized in character, techniques analogous to those used for simple gas-phase reactions should be adequate in most cases. They are based on activated dynamics[1], whose essential element is the determination of a transition state for the reaction and the performance of trajectories starting at the transition state to obtain satisfactory statistics for evaluating the transmission coefficient. Hydrogen exchange is of interest, not because of its intrinsic functional importance but because its wide range of timescales suggests it may be a useful probe of many aspects of protein dynamics. For conformational changes a number of simplified studies (such as hinge bending, which occurs in lysozyme, and the allosteric transition in haemoglobin) that treat the protein as composed of quasi-rigid units have had some success. For haemoglobin, an analysis which treats the helices

5. Free-energy simulations. Asp 99 β is situated at the α–β interface of the haemoglobin tetramer, a region involved in the transmission of information between subunits when haemoglobin makes the transition from the deoxygenated to oxygenated state. The mutation Asp 99 β → Ala (Hb Radcliffe) decreases the cooperativity and increases the oxygen affinity of haemoglobin. On the basis of the crystal structures of deoxygenated and oxygenated forms of haemoglobin, Perutz suggested that the essential role of Asp 99 β is to stabilize the deoxygenated form, by making hydrogen bonds to two specific residues (Tyr 42 and Asn 97) of the α subunit. The experimentally determined magnitude of the free-energy difference, $\Delta\Delta G$ (3.5 kcal per mol of interface), between the deoxygenated and oxygenated states of the mutant haemoglobin relative to the wild-type protein is consistent with the interpretation that the mutant has lost two hydrogen bonds.

Free-energy simulations[25] reveal that the effect of changing Asp 99 β to Ala is more complex than analysis of crystal structures had suggested. The mutation destabilizes both the deoxygenated and oxygenated forms of the haemoglobin tetramer, owing in large part to the loss of solvation free energy, but the destabilization of the deoxy tetramer (66 kcal per mol of interface) is greater than the destabilization of the oxy tetramer (60.5 kcal per mol of interface). The observed $\Delta\Delta G$ is a small difference (5.5 kcal per mol of interface) between the large individual ΔG values. In addition to the solvation effect, interactions with protein residues are also significant. Tyr 42 and Asn 97 are found to be important, but a number of other residues that have not been considered previously make larger contributions.

Although the structural change between oxygenated and deoxygenated haemoglobin may make the Asp 99 β mutation unusual in its complexity, the simulation results make clear that small measured $\Delta\Delta G$ values can arise from large 'hidden' contributions. In particular, partly cancelling changes in solvent–protein interactions and electrostatic repulsive interactions that may not be evident from a crystal structure need to be considered. Further, the results suggest that the observed free-energy changes may be small even when they contain many large contributions because the plasticity in the protein–protein and protein–solvent interactions lead to approximate cancellation of the effects for mutants that are stable.

as approximately rigid bodies attached to flexible hinges would seem to be the next step.

Finally, for protein folding it may be that the best approach to the dynamics is to study unfolding, so that at least the initial state is well defined. To achieve the required timescale it may be useful to introduce simplified models, some of which have already been used in brownian dynamics simulations and energy minimization[1]. As it is likely that the rate of this slow process (as of any other) is limited by barriers to metastable intermediates, it may be useful to follow the transitions in a stepwise fashion.

Many applications of molecular dynamics require only straightforward use of the standard methodology to solve problems of biological interest and aid in the interpretation and analysis of experimental data (such as NMR and X-ray diffraction data). There are a variety of problems that can clearly be solved but that require special methods to obtain meaningful results. Finally, there are problems, some of fundamental interest, for which molecular dynamics can be useful but for which new ideas will be required (probably beyond those suggested here) before a full solution is obtained.

It is an exciting time in the rapidly developing field of molecular dynamics simulations of biomolecules. This review will have served its purpose if it stimulates even a few biophysicists and biochemists, not to mention molecular biologists, to use molecular dynamics simulations as an aid in solving their problems. □

Martin Karplus is at the Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138, USA, and Gregory Petsko is at the Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, Massachusetts 02254, USA.

1. Brooks, C. L., III, Karplus, M. & Pettitt, B. M. *Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics* (Wiley, New York, 1988).
2. McCammon, J. A. & Harvey, S. *Dynamics of Proteins and Nucleic Acids* (Cambridge University Press, Cambridge, 1987).
3. Barnes, J. E. *Nature* **338,** 123–126 (1989).
4. Brooks, B. R. *et al. J. comput. Chem.* **4,** 187–217 (1983).
5. Stillinger, F. H. & Rahman, A. *J. chem. Phys.* **60,** 1545–1557 (1974).
6. McCammon, J. A., Gelin, B. R. & Karplus, M. *Nature* **267,** 585–590 (1977).
7. Aqvist, J., van Gunsteren, W. F., Leijonmarck, M. & Tapia, O. *J. molec. Biol.* **183,** 461 (1985).
8. Dobson, C. M. & Karplus, M. *Meth. Enzym.* **131,** 362–389 (1986).
9. Gelin, B. R. & Karplus, M. *Proc. natn. Acad. Sci. U.S.A.* **74,** 801–805 (1977).
10. Ringe, D. *et al. Trans. Am. cryst. Ass.* **20,** 109–122 (1984).
11. Doster, W., Cusack, S & Petry, W. *Nature* **337,** 754–755 (1989).
12. Smith, J., Kuczera, K. & Karplus, M. *Proc. natn. Acad. Sci. U.S.A.* **87,** 1601–1605 (1990).
13. Elber, R. & Karplus M. *Science* **235,** 318–321 (1987).
14. Chothia, C. & Lesk, A. M. *Trends biochem. Sci.* **10,** 116 (1985).
15. Lesk, A. M. & Chothia, C. *J. molec. Biol.* **136,** 225 (1980).
16. Case, D. A. & Karplus, M. *J. molec. Biol.* **132,** 343–368 (1979).
17. Elber, R. & Karplus, M. *J. Am. chem. Soc.* (in the press).
18. Teutlein, H. *et al.* in *The Photosynthetic Bacterial Reaction Center* (ed. Vermeglio, A.) 101–118 (Plenum, London, 1988).
19. Fleming, G. R., Martin, J.-L. & Breton, J. *Nature* **333,** 190–192 (1988).
20. Brünger, A. T., Clore, G. M., Gronenborn, A. M. & Karplus, M. *Proc. natn. Acad. Sci. U.S.A.* **83,** 3801–3805 (1986).
21. Wüthrich, K. *Accts chem. Res.* **22,** 36–44 (1989).
22. Brünger, A. T., Kuriyan, J. & Karplus, M. *Science* **235,** 458–460 (1987).
23. Brünger, A. T. & Karplus, M. *Accts chem. Res.* (in the press).
24. Beveridge, D. L. & Di Capua, F. M. *A. Rev. Biophys. biophys. Chem.* **181,** 431–492 (1989).
25. Gao, J., Kuczera, K., Tidor, B. & Karplus M. *Science* **244,** 1069–1072 (1989).
26. Harvey, S. *Proteins* **5,** 78–92 (1989).
27. Deisenhofer, J. & Michel, H. *Science* **245,** 1463–1473 (1989).
28. Post, C. B. *et al. J. molec. Biol.* **190,** 455–479 (1986).
29. Hoch, J. C., Dobson, C. M. & Karplus, M. *Biochemistry* **24,** 3831–3841 (1985).

ACKNOWLEDGEMENTS. We are most grateful to our students, postdoctoral fellows and senior colleagues, many of whom are listed in the references, for their contributions to the work described in this review.

The three-dimensional seismic velocity structure of the East Pacific Rise near latitude 9° 30′ N

Douglas R. Toomey*§, G. M. Purdy†, Sean C. Solomon* & William S. D. Wilcock‡

* Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
† Department of Geology and Geophysics, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA
‡ MIT/WHOI Joint Program in Oceanography, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Three-dimensional images of crustal seismic structure beneath the East Pacific Rise show pronounced axial heterogeneity over distances of a few kilometres. A linear high-velocity anomaly, approximately 1–2 km in width and restricted to the uppermost 1 km of the crust, is centred on the rise axis. An axial low-velocity anomaly at depths of 1–3 km varies in amplitude along the axis, consistent with a zone of higher crustal temperatures midway between two discontinuities in the morphology of the rise axis. This apparent thermal segmentation along axis is consistent with injection of mantle-derived melt midway along a locally linear, 12-km-long segment of the rise.

OVER the past two decades our view of sea-floor spreading has been transformed from being two-dimensional and time invariant to three-dimensional and time-dependent[1–8]. Most current hypotheses for the well documented along-axis variability in the morphology and geology of mid-ocean ridges invoke some form of mantle and crustal-level magmatic segmentation[1,8–12]. The specific relation between sites of vertical injection of magma and tectonic segmentation of the rise axis, however, is not known. Nor is it understood whether the extent of lateral crustal-level magma migration controls axial discontinuities in sea-floor morphology such as overlapping spreading centres or deviations from axial linearity[8,12,13]. Such issues may be addressed by mapping in detail the characteristics of both axial thermal structure and sea-floor morphology, enabling the correlation of morphological features with the structure of underlying magmatic systems. Because magmatism and thermal structure are

§ Present address: Department of Geological Sciences, University of Oregon, Eugene, Oregon 97403, USA.