Rep. Prog. Phys., Vol. 47, pp 1–46, 1984. Printed in Great Britain

Protein dynamics

J A McCammon

Department of Chemistry, University of Houston, Houston, Texas 77004, USA

Abstract

The biological activity of protein molecules depends on their structural fluctuations. Recent theoretical studies have helped to clarify the nature and function of these fluctuations. Because proteins are large densely-packed structures, their atomic motions can be compared to those that occur in other dense materials. Small motions at short times are similar to what is observed in liquids. Larger motions in proteins are opposed by the forces that stabilise their native structures, resulting in solid-like features. Of special importance is the strong coupling observed between local and collective displacements; this coupling governs the character of many ligand-binding processes and structural transformations that are essential to biological function.

This review was received in June 1983.

0034-4885/84/010001+46 © 1984 The Institute of Physics

Contents

1. Introduction

1.1. Protein function

Globular proteins are essential components of all living organisms. These molecules are responsible for a remarkably wide range of biological functions, as may be seen by listing a few of the major groups within this vast family. Enzyme molecules are proteins that catalyse biochemical reactions. They act to build the structural elements of organisms and to provide the energy necessary for life processes. Familiar examples include the digestive enzymes that degrade foodstuffs to simple, assimilable compounds; the biosynthetic enzymes that build complex molecules from simpler compounds; and muscle proteins that produce mechanical work from chemical reactions. Transport proteins such as haemoglobin facilitate the movement of molecular oxygen and other essential compounds to their sites of utilisation. Antibody molecules are proteins that bind to and neutralise foreign materials that may be harmful to an organism. Other globular proteins play essential roles in genetic expression, nerve conduction and all other biological processes.

An important characteristic of proteins is their specificity of function. A particular enzyme will bind specific substrate molecules and catalyse a specific chemical transformation of the substrate. A particular antibody molecule will bind specific antigens. For many proteins, this specificity of action is so narrowly defined that a small change in a ligand molecule that binds strongly to the protein (e.g. replacement of a hydrogen atom by a methyl group) leads to a dramatic reduction in binding. A number of proteins also display regulatory character. For example, their primary activity may be increased or decreased by the binding of specific auxiliary 'effector' ligands. Together with the spatial ordering of proteins imposed by the anatomy of an organism, the specificity and regulation of protein function are largely responsible for the required coherence of biochemical processes.

1.2. Protein structure and dynamics

Given the functional richness of proteins, one would expect to observe a corresponding complexity in the detailed structure of these molecules. This expectation has been confirmed by x-ray diffraction studies, which have provided the crystal structures of more than 100 proteins during the past 25 years (Bernstein *et al* 1977, Richardson 1981). Proteins are very large molecules; their molecular weights are often in the tens of thousands. The basic component of these molecules is the polypeptide chain, an unbranched polymer consisting of a sequence of amino-acid residues. There are 20 commonly occurring amino acids, and a typical chain will contain a few hundred of these elementary structural units. Protein molecules consist of one or a small number of such polypeptide chains, complemented in some cases by one or more 'prosthetic' groups (e.g. metal ions or special organic molecules). For a given globular protein, the polypeptide chain of each molecule is folded compactly into a characteristic three-dimensional structure. Although the resulting structures are complicated, it is commonly observed that the packing density of the protein components is nearly

maximised, subject to the requirement that those amino-acid residues which have a favourable free energy of interaction with water tend to remain near the protein surface (Richards 1977). In many cases, it has been possible to carry out x-ray diffraction studies of globular proteins with bound ligands (e.g. substrate analogues). These studies show that the folding of a protein is such that key amino acids with chemically active groups are strategically located in a well-defined 'active site', where the groups can interact in a coordinated fashion with the ligand. Such studies have been invaluable in the development of structural interpretations of protein function.

During the past six or seven years, increasing attention has been focused on the dynamic aspects of protein structure and function. It has long been inferred from a variety of experimental studies that substantial structural fluctuations occur in proteins, and that these fluctuations are essential to protein function (Edsall 1968, Careri *et al* 1975, Weber 1975). Until recently, the exact nature of the structural fluctuations has proved elusive. The recent surge of interest in protein dynamics has largely been stimulated by theoretical studies that have provided a detailed picture of the atomic motion in proteins. These theoretical studies are the primary focus of the present review. The theoretical work on proteins involves a combination of methods from theoretical chemical physics and protein structure theory. The methods from chemical physics include techniques that have been used successfully in the past to study atomic motion in dense materials such as liquids and solids. These methods are appropriate in view of the high density and large size of globular proteins. Along with the theoretical developments, new experimental techniques that provide detailed insights to protein dynamics have become available. Indeed, the present robustness of this field is largely a result of the interplay of modern theoretical and experimental work. Theory has successfully predicted a number of fundamental properties such as the average magnitude of atomic thermal displacements, the variation of these magnitudes throughout a protein and the time scales of certain group displacements. Recent experiments have presented new challenges (e.g. concerning solvent effects on protein dynamics and function (Beece *et al* 1980)) that are stimulating further theoretical work. The results achieved during the past few years and the history of corresponding efforts for systems such as simple liquids both suggest that the theoretical work on proteins will become increasingly sophisticated and useful in the coming years.

1.3. Scope of this review

The number of publications on dynamic aspects of protein structure and function is growing at an extraordinary rate. The present review is not intended to provide an all-inclusive catalogue of this activity. It is rather intended to provide a reasonably self-contained introduction to the theoretical foundations of the subject, and to highlight a representative selection of important new theoretical results within an integrated framework. The reader may wish to consult other recent reviews for additional material (Careri *et al* 1979, McCammon and Karplus 1980a, 1983, Karplus and McCammon 1981a, b, 1983, Levitt 1982, Welch *et al* 1982, van Gunsteren and Berendsen 1982). A limited number of experimental results are also described to illustrate the types of data available and the degree of overlap with the theoretical results. Again, excellent reviews that focus on various aspects of the experimental work have recently been published (Campbell *et al* 1978, Peticolas 1978, Woodward and Hilton 1979, Gurd and Rothgeb 1979, Williams 1980, Karplus and McCammon 1981a, Jardetzky 1981, Cooper 1981, Debrunner and Frauenfelder 1982, Hilinski and Rentzepis 1983, Huber

and Bennett 1983). A useful, brief introduction that will be of particular interest to physicists has been presented by Frauenfelder and Marden (1981).

2. Protein structure

2.1. Polypeptide chain

The structure of a short section of polypeptide chain is illustrated in figure 1. The polypeptide chain is intrinsically flexible because many of the covalent bonds that occur in its backbone and sidechains are rotationally permissive. The residues of which the polypeptide is composed are chosen from the 20 commonly occurring amino acids. A given protein is characterised by a definite sequence of residues; this is termed the primary structure of the chain. The amino-acid residues are distinguished by the structures and chemical properties of their sidechains. The sidechains can be divided into two broad classes. Sidechains that are relatively soluble in water are termed hydrophilic, while those that are less soluble in water are termed hydrophobic. The hydrophilic sidechains include electrically charged groups (acidic groups or basic groups, which typically bear a full negative or positive charge, respectively) and neutral groups with a substantial electric-dipole moment. The hydrophobic sidechains are neutral and relatively non-polar.

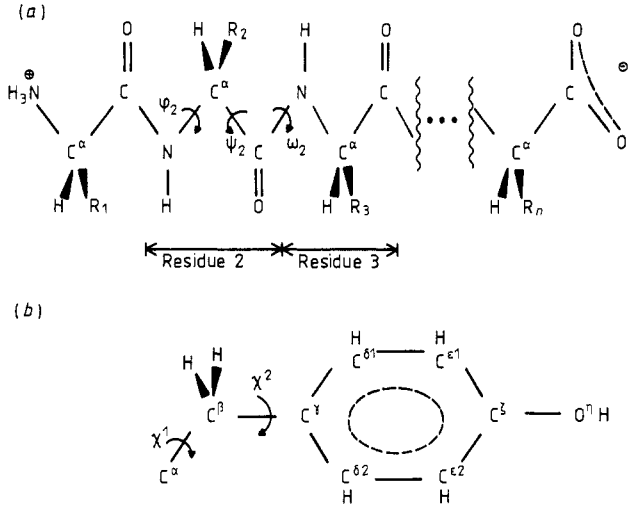

Figure 1. Structure of the polypeptide chain. The backbone or main chain is shown in (*a*). The covalent bonds and bond angles are rather rigid, but sizeable rotations can occur *around* certain bonds. The dihedral angles φ_i and ψ_i measure the torsion about the rotationally permissive bonds in the backbone of residue *i*. The dihedral angles ω_i exhibit little variation because the C—N bond has partial double-bond character; each peptide group (CONH) and its adjoining C^α atoms therefore tend to remain in a common plane. The labels R_i represent the sidechains, one of which is shown in detail in (*b*). This tyrosine sidechain has two rotationally permissive bonds; the corresponding dihedral angles are χ^1 and χ^2. The ring remains relatively flat due to partial double-bond character in its C—C bonds.

2.2. Chain folding and interactions

As has been mentioned, the polypeptide chain of a given type of protein folds into a characteristic 'native' three-dimensional structure in water. For simple proteins, this

folding is a spontaneous process (Kim and Baldwin 1982). In this subsection, we mention the principal interactions that stabilise the native structure. Of special importance are two types of interaction between the polypeptide and the surrounding water. The first of these is the so-called 'hydrophobic interaction', which causes hydrophobic sidechains to tend to cluster together in the protein interior. This entropy-driven effect is due to the restrictions on the allowed orientations of water molecules adjacent to a hydrophobic group (Stillinger 1980). The tendency of hydrophobic groups to cluster together is quite strong (Cantor and Schimmel 1980). Removal of a single methyl group from water to a non-polar region in the protein interior results in a net free energy of stabilisation of the order of 10 kJ mol^{-1}. This effect is important because a large fraction of the residues in a typical protein are hydrophobic. The second type of solvent interaction, that between charged groups and water, has an opposite character (Mao *et al* 1982b). Due to the large dipole moments of its molecules, liquid water has a high dielectric constant. Charged groups will therefore be attracted into the water and away from the protein interior, which has a relatively low dielectric constant. A typical protein has a relatively small fraction of charged residues, but the magnitude of this interaction makes the effect important. Transfer of a small hydrated ion from water to the protein interior would have a free energy cost of the order of 30 kJ mol^{-1}.

In addition to the polypeptide–solvent interactions described above, a number of important interactions between non-bonded atoms of the polypeptide help to determine the native structure of a protein. One of these is the Coulombic interaction between partial electrical charges on the polypeptide atoms. The most important interaction of this type is hydrogen bonding (Schulz and Schirmer 1979). When a hydrogen atom is bonded to oxygen, nitrogen or certain other electronegative atoms, the bonding electron density is largely shifted onto the heavier atom. The hydrogen is left with a significant partial positive charge and can approach other atoms relatively closely because of the drawn-in character of its electron cloud. Such a hydrogen atom can therefore have a strong electrostatic attraction for other oxygen or nitrogen atoms; such interactions are termed hydrogen bonds. For a hydrogen bridging two oxygen atoms, the maximum interaction amounts to about 20 kJ mol^{-1} when the oxygen atoms are separated by 0.28 nm; the interaction becomes neglible for oxygen distances greater than 0.4 nm. One effect of hydrogen bonding is to restrict the possible arrangements of hydrogen bonding groups in the protein interior. These groups must interact with one another in such a way that the loss of hydrogen bonds to solvent water molecules (which can act as hydrogen bond donors or acceptors) is compensated by the formation of internal hydrogen bonds in the folding of the protein. For the protein backbone, such arrangements often result in the formation of what are termed elements of secondary structure (Schulz and Schirmer 1979, Cantor and Schimmel 1980). A familiar example is the alpha helix, in which the peptide oxygen of residue i forms a hydrogen bond with the peptide nitrogen of residue $i+4$. Another example is the beta sheet structure, in which extended strands of the polypeptide chain lie next to one another and are cross-linked by hydrogen bonds between their peptide groups.

Finally, there are the important, non-specific van der Waals interactions between non-bonded atoms in the protein. This type of interaction is initially attractive for approaching atoms, and then becomes steeply repulsive as the electron clouds of the atoms begin to overlap substantially. Considering two methyl groups, a maximum stabilisation of the order of 1 kJ mol^{-1} is obtained when the two carbon atoms are separated by 0.42 nm; reducing this distance by 0.05 nm results in a destabilisation of the order of 10 kJ mol^{-1}. To compensate for the loss of attractive van der Waals

interactions with solvent molecules in the folding process, the atoms in the protein interior tend to pack together quite closely.

The three-dimensional arrangement of the groups in the native protein is termed the tertiary structure of the molecule. This structure represents the optimum balance among the various interactions described above. Typical features include the clustering of hydrophobic sidechains in several regions of the protein interior. Charged sidechains tend to remain exposed to the solvent at the protein surface; the few charged groups observed in protein interiors are usually paired with oppositely charged groups. Sections of the polypeptide backbone that are buried in the protein interior typically form secondary structure to compensate for the loss of their hydrogen bonds with solvent water molecules. The packing density of the atoms is quite high, particularly in the hydrophobic clusters where non-specific interactions predominate. The high packing density is clear in space-filling representations of protein structure (figure 2).

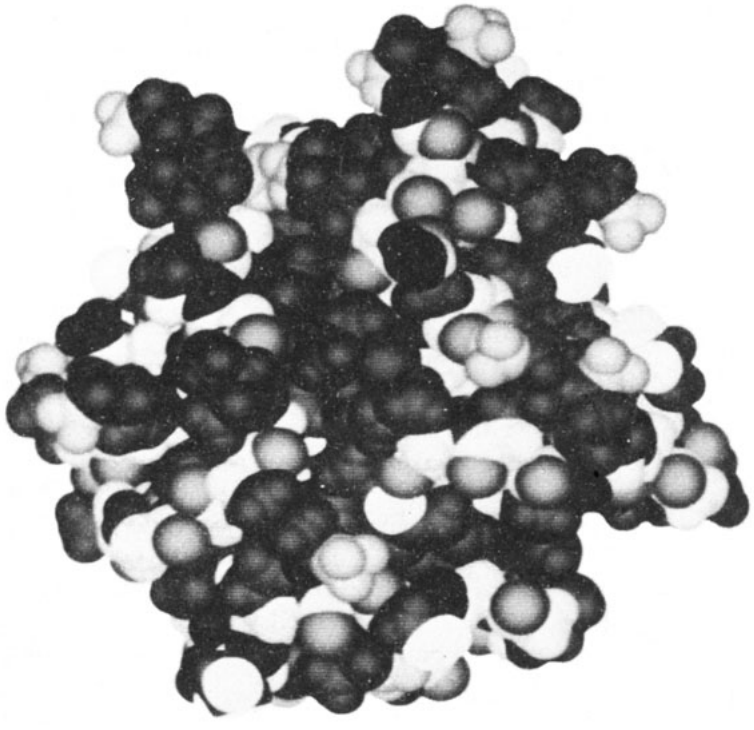

Figure 2. Space-filling representation of the native structure of a relatively small protein, cytochrome c. The sizes of the atomic spheres correspond to the distances of closest approach allowed by the van der Waals forces between non-bonded atoms. (Photograph courtesy of Richard J Feldmann, NIH.)

2.3. Model potential energy functions

More quantitative descriptions of the interactions that stabilise the native structure of a globular protein are provided by expressions that give the potential energy of the molecule as a function of its atomic positions. Ideally, such a function could be obtained by quantum-mechanical calculations using the Born–Oppenheimer approximation. That is, the ground-state energy of the molecule would be computed at every possible set of atomic positions. Such calculations are, however, not yet feasible for molecules

as large as proteins. A practical alternative is to construct model potential energy functions of the 'molecular mechanics' type (Schulz and Schirmer 1979, Cantor and Schimmel 1980). In this approach, one specifies a form for the potential function that is a sum of types of terms often used in the description of simpler systems. A typical model potential function has the form

$$V = \tfrac{1}{2}\sum_{\text{bonds}} K_b(b-b_0)^2 + \tfrac{1}{2}\sum_{\substack{\text{bond}\\\text{angles}}} K_\theta(\theta-\theta_0)^2 + \tfrac{1}{2}\sum_{\substack{\text{dihedral}\\\text{angles}}} K_\varphi[1+\cos(n\varphi-\delta)]$$

$$+ \sum_{\substack{\text{non-bonded}\\\text{pairs}}} \left(\frac{A}{r^{12}} - \frac{C}{r^6} + \frac{q_1 q_2}{Dr}\right). \tag{2.1}$$

The first sum includes a term for every covalent bond in the protein. Each bond is treated as a simple Hookesian spring with a characteristic force constant and equilibrium bond length; this is a very good approximation at normal biological temperatures where the bond length fluctuations will be quite small. The second term accounts for the deformation energy of angles between the covalent bonds to a given atom, and the third term represents the intrinsic deformation energy for twisting about an axis through covalently bonded atoms. Together, these three sums account for variations in the covalent bonding energy of the protein. The remaining sum corresponds to the van der Waals and Coulombic interactions described in the previous subsection; these interactions are particularly important determinants of protein structure due to the high packing density in proteins. The parameters that appear in a model potential function (e.g. the force constants and equilibrium bond lengths) are obtained from experimental and quantum-mechanical studies of small molecules that are chemically similar to segments of the protein.

Model potential energy functions such as that described above have two advantages over those that would be obtained from a full quantum-mechanical calculation. First, the energy can be computed very rapidly for a given configuration of atoms. This makes possible a determination of the relative stability of different possible protein conformations. Second, the simple form of the function allows one to write down analytic expressions for the spatial derivatives of the energy. Thus, the forces acting on the atoms of the protein can also be computed very rapidly. These advantages have led to the use of similar model potential functions in the analysis of vibrational spectra of small molecules (Wilson *et al* 1955) and the computer simulation of atomic and molecular liquids (Hansen and McDonald 1976).

As it stands, an equation such as (2.1) is strictly appropriate only for a protein in a vacuum. Such an equation is nevertheless useful in the analysis of small structural changes in the interior of a globular protein. More generally, it is necessary to incorporate the effects of the solvent surroundings. This can be done in one of two ways. More approximate, but less demanding of computation time, is the adjustment of the energy function parameters to reflect the average effects of the solvent surroundings. In terms of statistical mechanics, this corresponds to replacing the potential energy function by a potential of mean force that is averaged with respect to the possible configurations of the water molecules that surround the protein (Lifson and Oppenheim 1960). The potential of mean force is a type of free energy and is dependent on temperature (McQuarrie 1976). Among the parameter adjustments that have been used are the following. For the Coulombic interaction between two atoms, the dielectric

constant D has been set equal to the magnitude of the separation of the atoms expressed in ångström units (1 Å = 0.1 nm) (McCammon *et al* 1979). Physically, this corresponds to the tendency of the field lines connecting interior atoms to spread into the high dielectric solvent region as the atoms are drawn apart. In some calculations, this modification has been supplemented by a reduction of the charges of atoms near the protein surface because the field lines between such atoms always extend into the solvent region (Northrup *et al* 1981). Another modification that has occasionally been used is the elimination of the attractive part of the van der Waals energy for non-bonded interactions that involve a hydrophilic group; the remaining net attraction between hydrophobic groups reflects the expected tendency of such groups to aggregate in water (Mao *et al* 1982b). Some effort has been devoted to the development of 'solvation shell' models that should allow a more quantitative description of this type of effect (Paterson *et al* 1981). In principle, the most satisfactory representation of the solvent surroundings of the protein would be obtained by explicit incorporation of solvent molecules in the basic model (Wood 1979). The water molecules would then be treated on the same level as groups within the protein and represented by additional terms in the model potential function. This approach has been fully realised in a small number of very recent studies; the encouraging results obtained and improvements in computer technology are likely to make this a popular approach in future work (van Gunsteren *et al* 1983). Although considerably more computer time is required, this approach allows a detailed analysis of the structural and dynamical effects of solvent upon the protein surface.

3. Overview of dynamics

3.1. Length and time scales

As is illustrated in figure 1, a polypeptide chain consists of a large number of groups linked by covalent bonds that are intrinsically permissive of rotation. The groups linked by such bonds are themselves comparatively rigid and constitute the fundamental dynamical elements in a protein molecule. Examples of such groups are the CONH peptide groups that link successive residues and the ring in the tyrosine sidechain (figure 1). The typical thermal motions observed within a protein are dominated by the torsional oscillations of these groups about the single bonds that link them together.

Table 1. Typical features of some internal motions of proteins.

Motion	Spatial extent (nm)	Amplitude (nm)	lg of characteristic time (s)
Relative vibration of bonded atoms	0.2–0.5	0.001–0.01	−14 to −13
Elastic vibration of globular region	1–2	0.005–0.05	−12 to −11
Rotation of sidechains at surface	0.5–1	0.5–1	−11 to −10
Torsional libration of buried groups	0.5–1	0.05	−11 to −9
Relative motion of different globular regions ('hinge-bending')	1–2	0.1–0.5	−11 to −7
Rotation of medium-sized sidechains in interior	0.5	0.5	−4 to 0
Allosteric transitions	0.5–4	0.1–0.5	−5 to 0
Local denaturation	0.5–1	0.5–1	−5 to +1

Within such groups, only small atomic displacements occur due to the large energy cost of deforming bond lengths, bond angles and dihedral angles about multiple bonds. Superposition of the rigid group oscillations yields a remarkably rich dynamical spectrum that ranges from the rapid local motions of the individual groups to slow collective distortions of large regions within the molecule. Some of the motions that occur are indicated in table 1; these motions will be considered in more detail in what follows.

3.2. Comparison with motions in other dense materials

Because of the high packing density in protein molecules, their atomic motion displays certain similarities to that seen in other dense materials. Over short periods of time (less than 0.5 ps), the small-amplitude motions display similarities to the motions of molecules in a simple liquid (Hansen and McDonald 1976, Chandler *et al* 1983). Each group is temporarily trapped, rattling in a cage that consists of other groups in the protein and (at the surface of the protein) molecules of the surrounding solvent. The cage atoms are, of course, not stationary. They frequently collide with the encaged group, rapidly randomising its motion. Such collisions are the microscopic basis of the frictional effects that limit the rate of net displacement of the group.

For many processes with longer characteristic times, solid-like components of motion appear in the dynamics of the protein. Such components are expected, because the atoms of the protein have definite average positions corresponding to the native molecular structure. Because of the interactions that maintain this structure, the protein matrix displays only limited compliance in the larger deformations that occur at long times. The solid-like aspects of protein behaviour appear on both the local and global scales. Local group motions of large amplitude are typically opposed by substantial restoring forces associated with the distortion of the cage surrounding the group. Consider, for example, the rotational isomerisation of a tyrosine ring (figure 1) corresponding to a 180° change in χ^2. If the ring is located in the interior of a protein, a number of protein atoms are located in the volume that would be swept out by the ring during this rotation. The necessary displacement of these cage atoms associated with ring rotation produces substantial strain energy in the protein matrix; thus, there is a large energy barrier to the rotation of the ring (Gelin and Karplus 1975, McCammon *et al* 1983). A similar situation occurs in the movement of small molecules from one site to another within a protein. Such movements are of biological importance in the binding of oxygen to myoglobin or haemoglobin, a process that has been studied extensively in elegant experiments by Frauenfelder and co-workers (Debrunner and Frauenfelder 1982) and calculations by Case and Karplus (1979). The rate of such movements is again limited by energy barriers of steric origin. The underlying mechanism is quite similar to that which obtains for vacancy diffusion in crystals, i.e. hopping from one site to another with a strained intermediate state (Bennett 1975). In certain large-scale motions, the distortion of the protein is distributed over many residues and the relative displacements of neighbouring atoms are small. The protein can then be pictured as behaving somewhat like a continuous, elastic material. For a compact, globular protein, one simple example would be a 'breathing' motion analogous to the fundamental mode of radial oscillation of a sphere (de Gennes and Papoular 1969, Suezaki and Gō 1975). Another example is the 'hinge-bending' motion that occurs in proteins having two globular regions that are linked by a region having a smaller cross section (McCammon *et al* 1976). In these

types of motion, the protein may display a simple Hooke's law character. Such motions typically display significant damping due to the solvent surrounding the protein.

As has been mentioned, both frictional forces and mechanical restoring forces play roles in the dynamics of proteins. The relative importance of these forces depends on the particular physical process considered. For local oscillations about a stable conformation, both underdamped and overdamped motions occur. Examples include the underdamped vibrations of covalent bonds (a result of the large restoring forces together with the small size of the structural units and correspondingly small frictional effects involved) (McCammon 1976, McCammon *et al* 1977), the nearly critically damped oscillations of tyrosine rings relative to their cages (McCammon *et al* 1979) and the overdamped collective distortions of hydrophobic clusters within a protein. Global distortions of the hinge-bending type are typically overdamped because of the large surface displacements and solvent damping involved (McCammon *et al* 1976, McCammon and Wolynes 1977). Other global motions, such as the breathing modes of individual globular regions, may be underdamped (de Gennes and Papoular 1969, Suezaki and Gō 1975). A similar spectrum of behaviour obtains for structural transitions from one stable conformation to another. The rotations of tyrosine rings over the large barriers imposed by the protein matrix are of predominantly inertial character, although frictional effects reduce the rate of such transitions (Northrup *et al* 1982a). Frictional effects are often dominant in transitions that involve larger groups or smaller energy barriers. Simple examples include the rotation of sidechains at the protein surface (where the substantial restoring forces associated with a protein matrix are absent) (Levy *et al* 1979) or the unwinding of regions of the polypeptide chain from the protein surface (local 'denaturation') (McCammon *et al* 1980). Another important example is the initial diffusional encounter between a protein molecule and a ligand molecule with which the protein must interact in biological processes. Such encounters typically involve relatively weak mechanical interactions between the protein and the ligand, but are subject to substantial solvent frictional effects (Wolynes and McCammon 1977).

The phenomena described above can be characterised in more quantitative fashion by the use of phenomenological models of the motion. Such models have a long history in the analysis of liquids, polymer solutions and other dense materials. One such model is summarised by the Langevin equation of motion (Chandrasekhar 1943, McCammon *et al* 1979). For a single coordinate (e.g. one component of the centre-of-mass vector of a group in the protein), the equation takes the following form:

$$m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = F(x) - f\frac{\mathrm{d}x}{\mathrm{d}t} + R(t)\,. \tag{3.1}$$

Here, m and x are the mass and position of the group, respectively, and t is the time. The term $F(x)$ represents the effective mechanical restoring force acting on the group. The terms $-f\,\mathrm{d}x/\mathrm{d}t$ and $R(t)$ represent the direct effects of rapidly varying forces acting on the group (e.g. those due to collisions); the first term is the average frictional force due to the relative motion of the group through its surroundings (f is the friction coefficient) and $R(t)$ represents the remaining randomly fluctuating force. This model, and other phenomenological models (e.g. the diffusion equation and the Kramers and transition-state theories for the crossing of energy barriers), will be considered in more detail later. All of these phenomenological models incorporate effective restoring and frictional forces without attempting a detailed description of the microscopic origin of these forces. These models are of great value in the analysis of protein dynamics.

By using such models, one can focus on particular motions, assign definite magnitudes to the various forces involved and develop simplified physical pictures that help to deepen our understanding of the dynamics. Analysis of protein dynamics by use of a phenomenological model is often a helpful intermediate step before analysis by use of more detailed microscopic theories. Phenomenological models are also of value as a basis for simplified simulation studies of protein dynamics. For example, if one has available (e.g. from the Stokes and Einstein relationships) estimates of the diffusion constants of key particles in an overdamped system, then one can generate typical motions of these key particles without having to follow every detail of less interesting particles in their surroundings. Such 'stochastic' dynamics simulation methods will be described more fully in a later section.

4. Fast motions

4.1. Methods

Two theoretical methods have been used to study the details of motions in proteins with characteristic times less than about 10 ps. The first method is molecular dynamics simulation, in which the classical equations of motion for the atoms in the system are solved by numerical techniques for time intervals of 10–100 ps (McCammon *et al* 1977, 1979). The second method is the normal-mode approach, in which the motion is described as a superposition of harmonic vibrations whose frequencies are determined by the multidimensional parabolic shape of the potential surface near an energy minimum (Levy *et al* 1982b, Gō *et al* 1983). The normal-mode approach has the advantage that, once the modes of vibration are determined, many time-average and dynamic properties can be computed easily by analytic techniques. Also, the separation of actual motions into their principal normal-mode components may be helpful in the analysis of protein dynamics. The normal-mode approach has two major disadvantages relative to the dynamical simulation approach, however. The matrix eigenvalue calculations that are required for determination of the normal-mode displacements and frequencies are not possible for systems having more than a few hundred coordinates. This difficulty can be circumvented by simplifying the system or using special techniques to extract selected normal modes. A more fundamental difficulty is that anharmonic effects are known to be important in protein dynamics at room temperature. That is, the system moves on regions of the potential surface that are poorly approximated by parabolic extrapolations from the region of a local minimum. We will therefore focus primarily on molecular dynamics results in what follows.

A variety of specific algorithms have been used in dynamical simulations of protein molecules. These include the algorithms due to Gear, Verlet and Beeman that have been used in simulation studies of simpler systems (McCammon and Karplus 1980a, b, Levitt 1982). In some calculations, modified algorithms have been used to hold bond lengths fixed at their ideal values during the numerical integration of the atomic equations of motion (van Gunsteren and Karplus 1982a). This has an insignificant effect on the resulting dynamics but affords some savings in computer time. The corresponding freezing of bond angles produces significant distortions of the resulting atomic motions.

In a typical simulation, one begins with an x-ray structure for the protein of interest. The atomic positions are usually adjusted by use of an energy minimisation algorithm

to relax any large stresses in the initial structure. If solvent is included, the solvent atoms may be located initially near regular lattice points or at positions derived from a simulation of the pure solvent. Solvent molecules that overlap with atoms from the protein are excluded in this initial structure. Prior to the actual simulation, the system is equilibrated to ensure that the atoms have representative velocities and are subject to representative forces at the beginning of the simulation. Typically, the equilibration process involves atomic velocity reassignment from Maxwellian distributions for increasing temperatures alternated with short periods of dynamical propagation (Northrup *et al* 1981). When the system is equilibrated at a desired temperature, the dynamical propagation is continued for a period of 10–100 ps to produce the trajectory for analysis.

4.2. Single group motions

Many common features are apparent in the dynamical simulations that have been studied to date. The atomic displacements are larger near the surface of the protein than they are in the interior. For cytochrome c at 300 K, the average root-mean-square (RMS) position fluctuation of the atoms is of the order of 0.05 nm in the protein interior (Northrup *et al* 1981). This average remains nearly constant with increasing distance from the protein centre until the region of the rough molecular surface is reached. In this exterior region of the protein, the average RMS position fluctuation of the atoms grows rapidly to values of the order of 0.1 nm. The increased magnitude of the exterior fluctuations is due to the diminished steric constraints associated with the irregular topography of the surface. Similar results have been obtained in molecular dynamics studies of other proteins; for example, the largest motions in the small protein BPTI occur in loops of the polypeptide chain at the surface of the protein (McCammon *et al* 1977, van Gunsteren and Karplus 1982b, Swaminathan *et al* 1982). Another important generalisation is that the RMS position fluctuations of the atoms are larger in sidechains than in the backbone of the protein. Moreover, within a typical sidechain, the fluctuations become larger as one moves along the sidechain away from the backbone.

The motion of a typical atom in the protein is quite anisotropic (Northrup *et al* 1981, Morgan *et al* 1983). That is, the displacements are significantly larger in some directions than in others. The anisotropy can be described quantitatively in terms of the mean square displacement matrix of each atom. The mean square displacement matrix for a particular atom is (Willis and Pryor 1975)

$$\begin{pmatrix} \langle \Delta x \Delta x \rangle & \langle \Delta x \Delta y \rangle & \langle \Delta x \Delta z \rangle \\ \langle \Delta y \Delta x \rangle & \langle \Delta y \Delta y \rangle & \langle \Delta y \Delta z \rangle \\ \langle \Delta z \Delta x \rangle & \langle \Delta z \Delta y \rangle & \langle \Delta z \Delta z \rangle \end{pmatrix}$$

where Δx, Δy and Δz are the displacements of the atom from its mean position along the x, y and z axes, respectively, and the angular brackets indicate time averages. The sum of the diagonal elements is equal to the mean square position fluctuation of the atom. For each atom, a rotated set of Cartesian axes can be defined in which the mean square displacement matrix is diagonal; these axes are the principal axes of the matrix (Willis and Pryor 1975). The three elements of the diagonalised matrix are the matrix eigenvalues. Each eigenvalue is equal to the mean square fluctuation of the atom along one of the principal axes. For cytochrome c, the ratio of the mean square

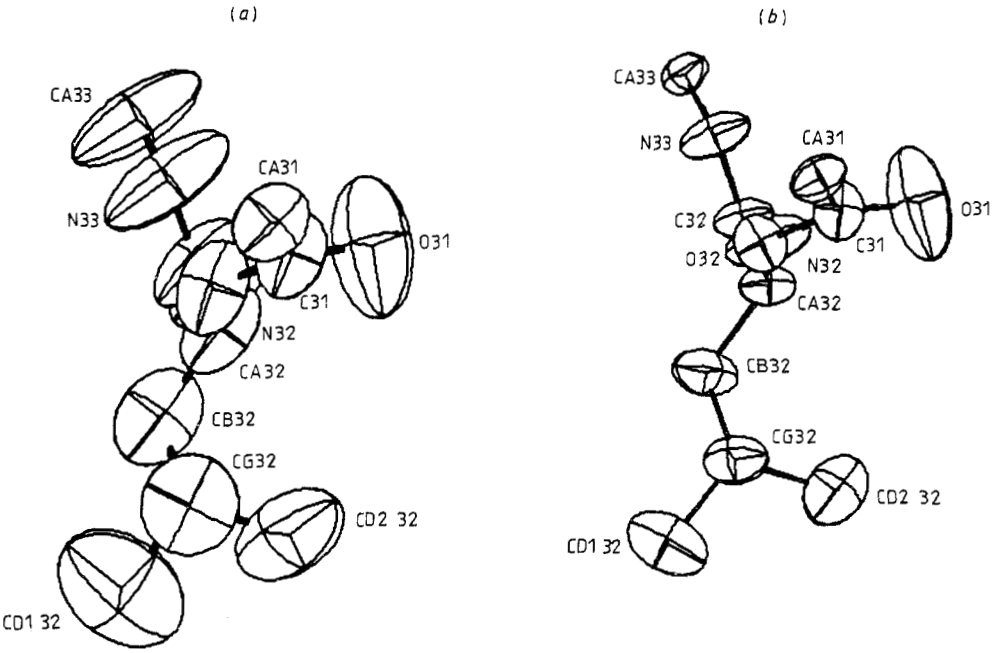

Figure 3. Thermal ellipsoids representing the position fluctuations of the atoms in the leucine-32 residue of cytochrome c, based on a molecular dynamics simulation (Morgan *et al* 1983). The lengths of the principal semiaxes correspond to (*a*) the RMS fluctuations observed during periods of 32 ps and (*b*) twice the RMS fluctuations observed during a period of 0.2 ps. The importance of collective motion during the longer time period is evident in the similar orientations of the longest axes of a number of atoms.

displacements along the axes of largest and smallest displacement is about 0.25 for a typical atom; the range of such ratios observed is 0.03–0.68 (Northrup *et al* 1981).

For certain classes of atoms, the direction of preferred displacement can be correlated with the local protein structure (Morgan *et al* 1983). An example is shown in figure 3(*a*). The atoms CD1 and CD2 at the end of the sidechain have major axes that are roughly consistent with rotational oscillations about the single bond between the CB and CG atoms. In general, however, the directions of preferred displacement are determined by collective motions that are unrelated to local bonding; this effect is discussed below.

As mentioned before, the atomic motion in proteins is substantially anharmonic at room temperature. This can be seen directly by calculating moments of the atomic position distribution functions. For harmonic motions, the atomic position distribution functions will be Gaussian. In this case, the first few moments of the position distribution functions along a given direction satisfy the following equalities:

$$\langle(\Delta x)^3\rangle = 0$$

$$\langle(\Delta x)^4\rangle = 3\langle(\Delta x)^2\rangle^2. \tag{4.1}$$

Calculations based on molecular dynamics simulations of cytochrome c show that one or both of these equalities are substantially violated for about half of the atoms in the protein (Mao *et al* 1982a). Morever, the anharmonic effects are related to the anisotropy; for a typical atom, the anharmonic effects are largest for motion in the direction of largest displacement. The atoms with largest anharmonicity fall into two

classes. In the first class are those atoms whose effective potential wells are distorted toward a square well shape. This is especially common near the surface of the protein or at the ends of sidechains, where structural constraints are expected to be small. The second class includes those atoms whose effective potentials have multiple minima separated by barriers greater than or equal to about 2 kJ mol^{-1}.

The importance of anharmonic effects in protein dynamics has also been shown in a recent study of the shape of the potential surface of BPTI near an energy minimum corresponding to the native state (Gō *et al* 1983). In this study, a normal-mode calculation was first carried out in the dihedral angle space of the protein. The atomic positions were then displaced in accordance with each normal mode; the actual potential energy of the distorted protein was calculated and compared with that obtained by a quadratic fit to the surface near its minimum. For displacements of typical thermal magnitude, significant differences in the two energies were found for the softest directions. Of the 36 modes with frequencies below $\nu = 10^{12}\ s^{-1}$, only one is close to being harmonic. These low-frequency anharmonic modes were found to dominate the atomic displacement magnitudes. Superimposed on these low-frequency motions are localised motions with higher frequency and smaller amplitude; these localised motions are more harmonic in character.

Also of interest is a recent study of the decaglycine alpha helix (Levy *et al* 1982b). Here, the atomic position fluctuations from a molecular dynamics simulation (reflecting motion on the true potential surface) were compared with those from a normal-mode calculation. Substantial anharmonicity was apparent even in this comparatively rigid element of secondary structure. At room temperature, the mean square displacements observed in the dynamical simulation are approximately twice as large as those in the corresponding harmonic model. Again, detailed analysis showed that the low-frequency motions are primarily responsible for the atomic displacements and display the largest anharmonicities.

The preceding discussion has only considered time-average properties of the atomic displacements. In considering the time dependence of the atomic motions, it is appropriate first to examine the dependence of the apparent amplitude of atomic motion upon the observation period (Morgan *et al* 1983). The RMS displacements of the protein atoms approach their limiting values in a time of the order of 10–20 ps. About one-half of this limiting value is achieved in a time of 1 ps. For the short-time (less than 1 ps) position distribution functions, the anisotropy of atomic displacement is simply related to the local bonding. That is, the atomic displacements are dominated by local rotations about single bonds. This result is quite apparent in figure 3(*b*). These correlations are largely washed out by collective motions in the protein at longer times, as is apparent in figure 3(*a*).

More detailed information on the dynamics of short-time motions can be obtained by calculation and analysis of appropriate time correlation functions (McCammon 1976, McCammon *et al* 1977, Swaminathan *et al* 1982). The time correlation functions for displacements of atom positions or for rotations about single bonds typically indicate the presence of different processes occurring on two or more different time scales. These time correlation functions generally decay significantly in the first 0.2 ps. During the subsequent 1–2 ps, a variety of behaviours are observed, ranging from slow, nearly monotonic decay to damped oscillation ($\nu \leqslant 10^{12}\ s^{-1}$). Inclusion of the solvent surroundings in the simulation results in some slowing of the motion (correlation times increase by a factor of two or so), but relatively little change in amplitude (Swaminathan *et al* 1982).

4.3. Collective contributions to local motions

Several of the results described in the previous subsection indicate that collective motions are of importance in protein dynamics. For example, collective motions are suggested by the collinearity of the preferred axes of displacement of a number of atoms in figure 3(*a*). Another indication is the existence of components with long decay times in the atom position fluctuation time correlation functions. That such collective components are present is consistent with the large amplitude of the atomic position fluctuations in the closely packed structure of the protein; large-amplitude displacements must involve the collective motion of an atom and its neighbours. A number of recent studies have focused on these collective motions.

In one study, the atomic displacements in a dynamical simulation of cytochrome c were decomposed into two parts (Morgan *et al* 1983). The first part is the displacement of an atom relative to the centroid of the amino-acid residue to which it belongs, and the second part is the displacement of the residue centroid itself. The former component largely reflects localised motions while the latter component is a probe of the collective motion. This study showed that the local components dominate the magnitude and directionality of the net atomic displacements during time intervals less than 1 ps. Over longer time intervals, the collective component tends to dominate the atomic displacements. These results are illustrated in figure 3. The relative contributions of the local and collective motions vary throughout the protein; the collective contributions are most predominant for larger, rigid groups and for the polypeptide backbone. The spatial extent of the collective motions in cytochrome c has been probed by calculating the displacement correlations of the residue centroids (Ransom-Wright and

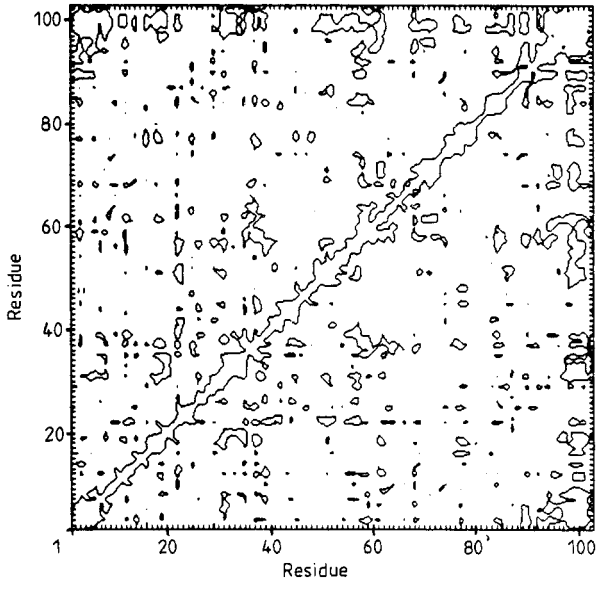

Figure 4. Map showing the extensive residue displacement correlations in a 32 ps dynamical simulation of cytochrome c (Ransom-Wright and McCammon 1984). The correlation of residues i and j is computed as the normalised quantity $C_{ij} = \langle R_i R_j \rangle / \langle R_i^2 \rangle^{1/2} \langle R_j^2 \rangle^{1/2}$, where R_i is the instantaneous displacement of the centroid of residue i from the mean position of the centroid. The values of C_{ij} range from 0.6–1.0. A single contour is drawn at the value 0.87; larger correlations occur along the diagonal and within the small closed regions off the diagonal.

McCammon 1984). The results, displayed in figure 4, show that many residues have large cross-correlations. Some of these correlations are easily understood. For example, the correlated motion of residues 92–103 seems to be associated with the C-terminal alpha helix, and the correlation of these residues with residues 1–10 reflects the contact between the C-terminal and N-terminal helices in the native structure of the protein. Other correlations that involve spatially distant residues may be best rationalised in terms of large-scale normal modes of oscillation (see below).

In another study, based on a molecular dynamics simulation of BPTI, it was shown that distinct regions of the protein could be identified within which groups of atoms

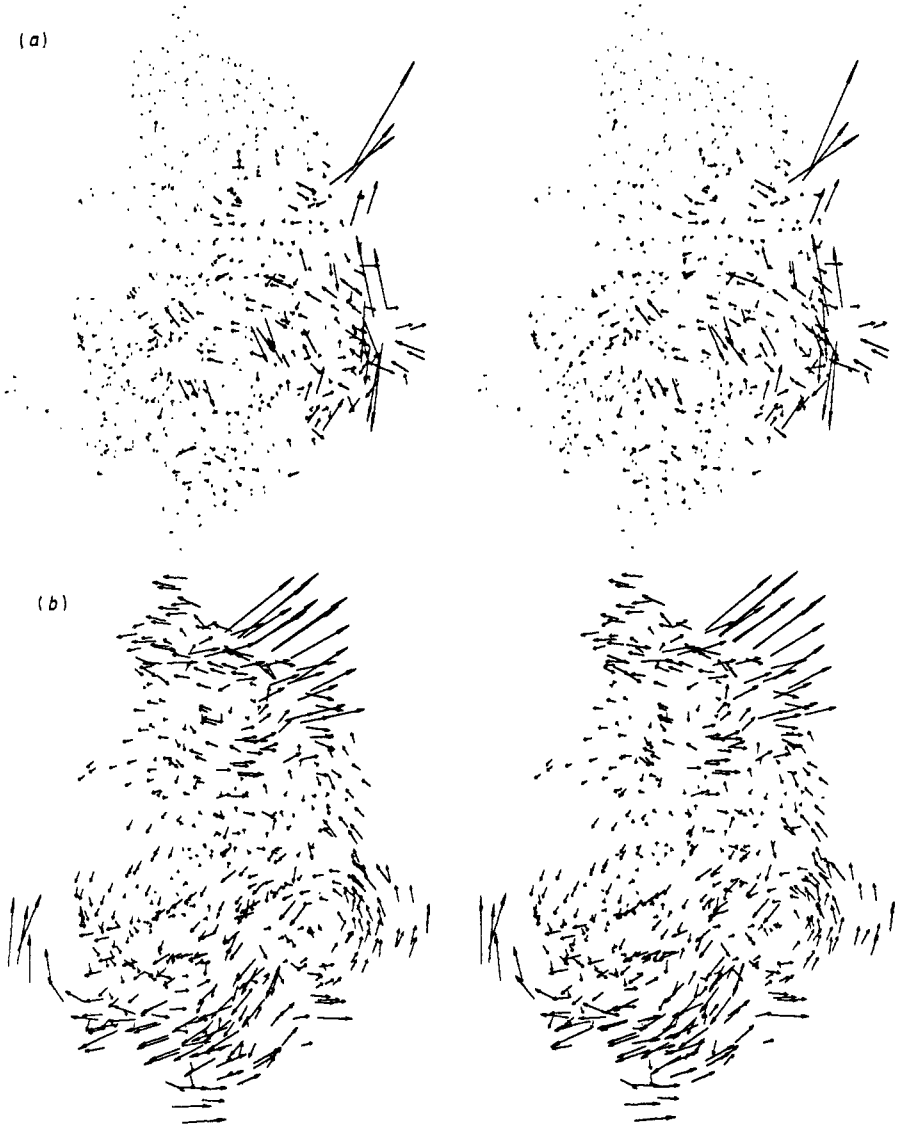

Figure 5. Stereoviews of the directions of atomic displacement in two low-frequency vibrational modes of the pancreatic trypsin inhibitor, as determined by a normal-mode calculation (Gō *et al* 1983). Results are given for modes at frequencies (*a*) $\nu = 3.56 \times 10^{12}\ s^{-1}$, (*b*) $\nu = 2.07 \times 10^{11}\ s^{-1}$.

displayed very similar displacement time correlation functions (Swaminathan *et al* 1982). This result suggests that collective motion dominates the displacements within each such region. These regions include atoms from as many as seven or more residues. The importance of collective motions in BPTI has also been demonstrated by normal-mode calculations on the protein (Gō *et al* 1983). In this study, it was shown that most of the low-frequency modes involve concerted displacement of many atoms. For modes with frequencies in the range $\nu = 3.6 \times 10^{12}\,s^{-1}$ to $6 \times 10^{12}\,s^{-1}$, the motions typically involve several neighbouring residues in the three-dimensional structure. For modes with frequencies below $3.6 \times 10^{12}\,s^{-1}$, the motions are more global in character. The patterns of collective displacement in two of these low-frequency modes are illustrated in figure 5. The low-frequency motions make the dominant contribution to the net atomic displacements. This dominant role of slow, collective motions is consistent with the results obtained from molecular dynamics studies of proteins (Morgan *et al* 1983, Swaminathan *et al* 1982) and isolated alpha helices (Levy *et al* 1982b). Although the normal-mode approach provides a useful characterisation of some important collective motions, it is limited by the harmonic approximation. Molecular dynamics simulations show that collective motions with anharmonic character do occur in proteins. An important example is the concerted shifting of nearby groups of atoms from one local minimum to another (Mao *et al* 1982a).

4.4. Langevin equations

As mentioned in § 3, a useful first step in the quantitative analysis of motions in proteins is the application of phenomenological models. An analysis of this kind has been carried out for the torsional librations of buried tyrosine residues observed in a molecular dynamics simulation of BPTI (McCammon *et al* 1979). The librations examined are those about the axis passing through the gamma and zeta carbon atoms of each ring (cf figure 1). The motion is analogous to rotation of a benzene molecule about an axis passing through two atoms at opposite vertices of the ring. Within the protein, the amplitude of libration is limited by steric hindrance between atoms in the

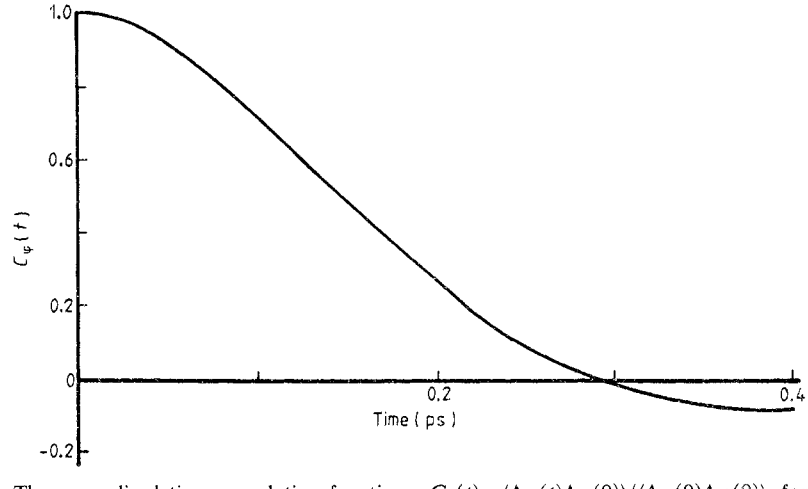

Figure 6. The normalised time correlation functions, $C_\varphi(t) = \langle \Delta\varphi(t)\Delta\varphi(0)\rangle / \langle \Delta\varphi(0)\Delta\varphi(0)\rangle$, for torsional fluctuations $\Delta\varphi$ of the tyrosine-21 ring in a molecular dynamics simulation of the pancreatic trypsin inhibitor (McCammon *et al* 1979).

ring and those in the cage surrounding the ring. A typical amplitude observed at 300 K is 12°. The time correlation functions for the torsional fluctuations decay to small values in a short time (approximately 0.2 ps); subsequent decay occurs on a time scale of several picoseconds. This result, illustrated in figure 6, indicates that the torsional motion is predominantly local in character, with the ring rattling in its cage. On the longer time scale, collective distortions of the protein produce small changes in the overall shape of the cage with a concomitant alteration of the ring orientation. From the distribution of observed torsional librations $p(\varphi)$, one can calculate the potential of mean torque $W(\varphi)$ for the ring librations:

$$W(\varphi) = -k_B T \ln p(\varphi). \tag{4.2}$$

The potential $W(\varphi)$ is roughly quadratic, which suggests that the ring libration can be analysed in terms of the Langevin equation for a harmonic oscillator. This equation is

$$I\ddot{\varphi} = -k\varphi - f\dot{\varphi} + N(t) \tag{4.3}$$

where $\varphi(t)$ is the torsional displacement, I is the moment of inertia of the ring about the torsional axis, f is a friction constant, k is a harmonic restoring force constant obtained from the potential of mean torque, and $N(t)$ represents the random torques acting on the ring due to fluctuations in its environment. The Langevin equation is appropriate if $N(t)$ varies more rapidly than $\varphi(t)$; in this limit, $N(t)$ may be regarded as a Gaussian random process and it is not necessary to specify the mechanism by which the torque fluctuations arise. This condition is approximately satisfied for the tyrosine rings; the duration of a typical collision between the ring atoms and the

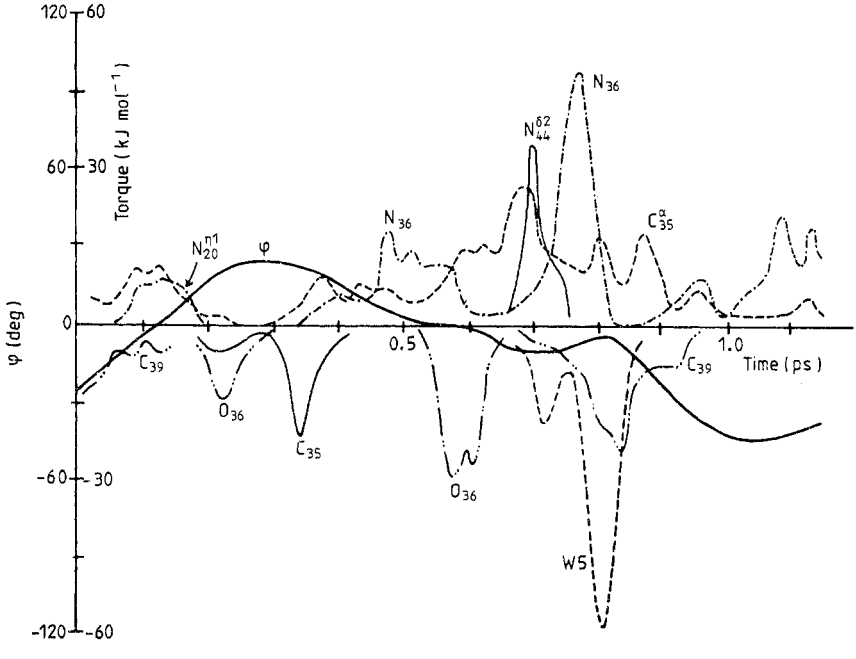

Figure 7. Torques exerted on the tyrosine-35 ring due to van der Waals interactions with the surrounding atoms in a molecular dynamics simulation of the pancreatic trypsin inhibitor (McCammon and Karplus 1980b).

surrounding cage atoms is 0.1ps, which is shorter than the ring libration correlation time (cf figure 7). The moment of inertia is readily calculated from the known atomic masses and geometry of the ring, and the force constant k is known from the shape of the potential of mean torque. The friction constant f can be calculated from the decay time τ of the torsional fluctuation correlation function by the relation $\tau = f/k$. The numerical values of these three coefficients indicate that the rotational motion of the ring within its cage is nearly critically damped; this is consistent with the shape of the time correlation function in figure 6. The friction constant may be related to an angular diffusion constant by use of the Einstein formula $D = k_B T/f$. The calculated value, $D = 2 \times 10^{11}\ s^{-1}$, is of the same order of magnitude as the experimental diffusion constants for the corresponding rotational motions of small aromatic molecules in organic solvents (e.g. $D = 7.9 \times 10^{10}\ s^{-1}$ for benzene in isopentane). Physically, this result makes sense because the protein rings examined are located in hydrophobic clusters within the protein. That the rotational diffusion constant in the protein is somewhat larger than that in a typical organic solvent is apparently due to the covalent connectivity of the surrounding cage atoms within the protein; there are fewer matrix atoms in positions that allow for effective perturbation of the ring angular velocity than would be found in typical solvent surroundings. Similar effects have been observed in spectroscopic studies of small molecules in large-molecule solvents (Moog *et al* 1982).

The above analysis is restricted to motion on a time scale of a few tenths of a picosecond. It is only for these short intervals that the calculated time correlation functions are accurate according to strict criteria (Zwanzig and Ailawadi 1969). An attempt has been made, however, to extend the above analysis to longer times by examination of atomic displacement time correlation functions over periods of up to 10 ps (Swaminathan *et al* 1982). A number of interesting qualitative conclusions emerged from this work. Dynamical simulations of BPTI in the presence and absence of a simple model solvent environment were found to yield similar RMS atomic displacements. Time correlation functions for the atomic displacements decay somewhat more slowly in the presence of solvent, particularly for atoms near the surface of the protein. In the vacuum calculation, the range of correlation times was 0.2–5 ps whereas the solvent calculation results ranged from 0.4–10 ps. For the protein in solvent, the correlation times for atoms in the exterior tend to be substantially larger than those for atoms in the interior of the protein. Effective friction constants for the atoms were estimated using a simple harmonic oscillator Langevin equation approach. From these friction constants, it was estimated that the collective motions that dominate the atomic displacements over long times typically involve structural units containing of the order of 100 heavy atoms.

4.5. Detailed microscopic models

Applications of the Langevin equation allow one to identify some of the important contributions to the dynamics of particular motions in protein molecules. In principle, a more detailed decomposition of the dynamics is possible by microscopic models of the type employed in kinetic theory. Such a microscopic analysis has been attempted for the sub-picosecond ring librations that were described in the previous subsection (McCammon *et al* 1979). Because the ring environment bears some resemblance to an organic solvent, the microscopic model employed is of the Enskog type that has been used successfully in the study of simple liquids (Hynes 1977). In this model, the reorientation of a molecule in a liquid is assumed to occur as a result of successive

binary collisions between the repulsive van der Waals cores of neighbouring molecules. The successive collisions are assumed to be uncorrelated and of instantaneous duration and to randomise the angular velocity of the reorienting molecule; this molecule moves freely between collisions. For the tyrosine ring in the protein, the model was modified to take into account the fact that, between collisions, the ring moves in a harmonic manner due to the torsional restoring force. In the absence of collisions, the ring reorientation is described by a simple harmonic oscillator equation of motion

$$I\ddot{\varphi} = -k\varphi \tag{4.4}$$

where we have assumed that $\langle\varphi\rangle = 0$. The unperturbed oscillator exhibits undamped librations at a frequency $\omega_0 = (k/I)^{1/2}$. The time correlation function for the torsional angle is simply

$$C_\varphi^0(t) = \cos \omega_0 t. \tag{4.5}$$

If the oscillator is subject to instantaneous, uncorrelated collisions, and if each collision randomises the angular momentum of the oscillator, then the displacement time correlation function of the perturbed oscillator is

$$C_\varphi(t) = \exp(-\nu t/2)[\cos at + (\nu/2a)\sin at] \tag{4.6}$$

where ν is the reciprocal of the mean time between collisions and $a^2 = \omega_0^2 - (\nu^2/4)$. With reasonable parameters, this model successfully reproduces the simulation time correlation function for periods of several tenths of a picosecond. The simple microscopic model is, however, inadequate in several respects. As can be seen from figure 7, most of the collisional torques are not sufficiently strong to randomise the angular velocity of the ring. Also, several impulses in rapid succession are sometimes observed for particular ring-atom, matrix-atom pairs (most commonly when the matrix atom is part of the local backbone). It would be desirable to incorporate these features into a more satisfactory microscopic model.

4.6. Nature of fast motions

The forces that act to stabilise the native folding patterns of globular proteins are sufficiently weak that the tertiary structures of these molecules are only marginally stable under optimal conditions. Consequently, sizable structural fluctuations occur in proteins, especially near the protein surface where steric constraints are somewhat relaxed. The small amplitude displacements of the atoms ($\leqslant 0.02$ nm) observed during short periods of time ($\leqslant 0.2$ ps) are similar to those observed in liquids. The motion of a typical group is chaotic, reflecting the nearly impulsive collisions between the group and the atoms that encage it. The largest displacements are in the directions corresponding to local dihedral angle fluctuations. Larger displacements of the atoms ($\leqslant 0.1$ nm) observed during longer periods of time ($\leqslant 10$ ps) are more similar to those observed in solid materials at temperatures approaching their melting points. The motion of a typical group is subject to significant restoring forces but is damped by collisional effects. The magnitudes and directions of displacement are dominated by collective motions of clusters of amino-acid residues. The solid-like behaviour is manifest most strongly in the protein interior, where the steric constraints are strongest. This qualitative picture of a protein having hybrid solid/liquid character, with the solid character most pronounced in the interior, is consistent with a simple theoretical model of protein structure developed some time ago by Lifshitz (1969).

The large contribution of collective motions to the net atomic displacements suggests that such motions are of importance in protein function. This supposition is strengthened by studies of activated processes in which collective motions have been found to play an essential role, as will be discussed in subsequent sections. Although more work is needed to adequately characterise such motions, some important results are available. Many collective motions in a protein relax in times of the order of 10 ps. This implies that dynamic coupling between solvent and the protein interior, or between different regions within a protein, will be cut off at frequencies above about $10^{11}\ s^{-1}$. Because the chemical bonding interactions that stabilise the structure of individual groups within a protein are much stronger than the non-bonded interactions that stabilise the tertiary structure, the collective motions consist of anharmonic displacements of groups whose internal motions are harmonic. In some cases, however, collective motions may have 'quasiharmonic' character. That is, the potential of mean force for collective coordinates (averaged over the faster localised motions) may be approximately quadratic even though the motion extends to non-quadratic regions of the underlying potential energy surface. Thus, it may be possible to analyse such motions in terms of normal modes on a (temperature-dependent) effective potential surface (Levy *et al* 1983).

5. Slow, localised motions

5.1. Methods

Displacements of local groups of atoms that have long characteristic times often represent activated processes. In such processes, the system must surmount a rate-limiting energy barrier that separates certain initial and final configurations. Such processes are of great importance in biology. Examples range from local conformational changes associated with the binding of ligands (Debrunner and Frauenfelder 1982) to the rearrangement of covalent bonds that occurs in enzymatic reactions (Schulz and Schirmer 1979, Cantor and Schimmel 1980).

To determine the dynamical details of activated processes, it is desirable to carry out simulations of the atomic motions involved. However, the conventional molecular dynamics technique is not sufficient for such simulations. The difficulty arises from the infrequent occurrence of the transitions. Although the barrier crossing process itself is often intrinsically rapid (an individual transition may require less than one picosecond), the time required for random fluctuations within the system to produce the local atomic momenta required for the subsequent barrier crossing may be of the order of milliseconds or longer. Accordingly, the process will not be observed in a simulation of a few hundred picoseconds duration.

Recently, a modified molecular dynamics method has been developed for simulation studies of such processes (McCammon and Karplus 1979, 1980b, Northrup *et al* 1982a). In this method, the calculation is divided into two parts. In the first part, one calculates the probability of finding the system in a region near the top of the energy barrier. In the second part, trajectory calculations are initiated from the region of the barrier top, thereby avoiding the long activation step. The approach is most easily understood in terms of the following expression for the rate constant (Northrup *et al* 1982a):

$$k = \tfrac{1}{2}\kappa\langle|\dot{\xi}|\rangle\left[\rho(\xi^{\ddagger})\left(\int_i \rho(\xi)\,\mathrm{d}\xi\right)^{-1}\right]. \tag{5.1}$$

Here, κ is a 'transmission coefficient' (see below), ξ is a 'reaction coordinate' that measures the progress of a transition, $\dot{\xi}$ is the time derivative of ξ, $\xi^{\ddagger}$ designates the value of ξ at the peak of the barrier and $\rho(\xi)\,\mathrm{d}\xi$ is the probability of finding the system in the interval ξ to $\xi+\mathrm{d}\xi$. Except for κ, all of the factors that appear in this expression are time-average or equilibrium quantities. The transmission coefficient κ depends on the detailed dynamics of the trajectories passing through the barrier region. The first step of the modified technique consists of defining a suitable reaction coordinate for the process and calculation of $\rho(\xi)$; this provides a value for the factor in square brackets in equation (5.1). The reaction coordinate is a function of the system coordinates and has the following ideal properties: (*a*) the reaction coordinate varies monotonically as the system moves from the initial-state region over the energy barrier and into the final-state region, and (*b*) all configurations with $\xi<(>)$ the barrier top value $\xi^{\ddagger}$ experience a mean force averaged with respect to coordinates other than ξ in the direction of the initial-(final-)state region. From property (*b*), it is desirable that ξ include terms corresponding to all interactions that make sizeable and systematic contributions to the barrier. Given the reaction coordinate, the calculation of $\rho(\xi)$ can be carried out by an umbrella sampling procedure (Northrup *et al* 1982a). In this procedure, one carries out a sequence of simulations in which the system is constrained to remain within small intervals of ξ, but with unrestricted sampling of all other coordinates. By piecing together data from overlapping 'windows' of ξ, $\rho(\xi)$ can be constructed over a range extending from the reactant region through the barrier region. In the second part of the calculation, a large number of independent trajectories are calculated by the molecular dynamics technique, starting from a representative set of configurations with $\xi=\xi^{\ddagger}$. The trajectory data are used to calculate $\langle|\dot{\xi}|\rangle$ and to evaluate κ according to the following equation (Chandler 1978):

$$\kappa(t)=D\langle\dot{\xi}(0)\delta[\xi(0)-\xi^{\ddagger}]H_p[\xi(t)]\rangle. \tag{5.2}$$

Here, D is a normalisation constant chosen such that $\kappa(0_+)=1$, δ is the Dirac delta function and H_p is a step function which is equal to 1 or 0 for ξ greater than or less than $\xi^{\ddagger}$, respectively. Thus, $\kappa(t)$ is a time correlation function that measures the net flux into the final-state region for trajectories that start at the barrier top. This reactive flux time correlation function decays rapidly (within a picosecond) to a plateau value that is used in equation (5.1).

5.2. Rotational isomerisation

The method outlined in the prevous subsection has been applied in a detailed study of rotational isomerisation of tyrosine sidechains in the interior of the protein BPTI (Northrup *et al* 1982a, McCammon *et al* 1983). In this process, the tyrosine ring rotates such that $\Delta\chi^2=180°$ (cf figure 1). The energy barrier to this rotation arises from non-bonded repulsions between atoms in the ring and in the surrounding protein matrix. An early analysis, using energy minimisation methods, showed that the very large barriers encountered in a hypothetical rigid protein are reduced to 40 to 100 kJ mol^{-1} when deformations of the matrix are allowed (Gelin and Karplus 1975). Although this is a comparatively simple activated process, the full dynamical analysis has provided a number of important general results (Northrup *et al* 1982a, McCammon *et al* 1983). A significant initial finding was that the dihedral angle χ^2 is not by itself a good reaction coordinate. Examination of structural models showed that, during rotation of the tyrosine-35 ring, certain ring atoms come into close contact with atoms of the adjacent polypeptide backbone, particularly with N_{36}. Thus, a new reaction

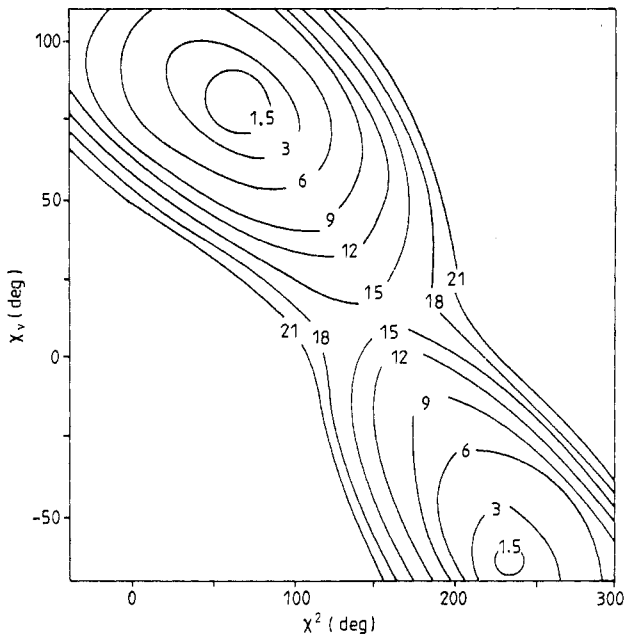

Figure 8. Adiabatic potential surface for rotational isomerisation of the tyrosine-35 ring in the pancreatic trypsin inhibitor (McCammon *et al* 1983). The contours are in units of kcal mol^{-1} (1 kcal = 4.184 kJ).

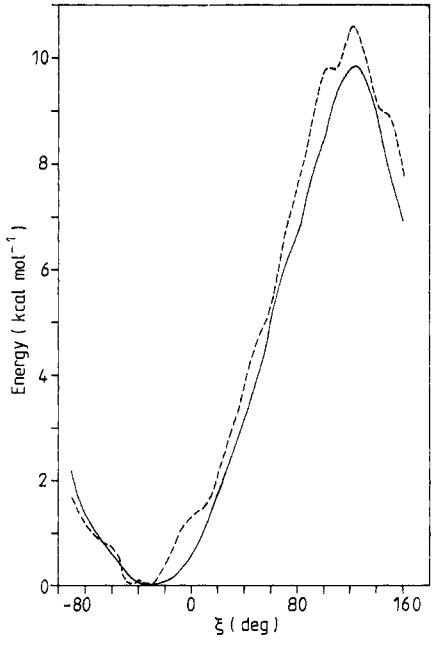

Figure 9. The potential of mean force $W(\xi)$ (——) and mean potential energy $\langle V(\xi) \rangle$ (– – –) as functions of the tyrosine-35 ring rotational isomerisation reaction coordinate ξ (Northrup *et al* 1982a). The energies are given in units of kcal mol^{-1} (1 kcal = 4.184 kJ).

coordinate was constructed as $\xi = \chi^2 - \chi_v$; χ_v is an angle that measures how far N_{36} is from the plane of the ring. Energy minimisation of the system was carried out for fixed values of χ^2 and χ_v. The resulting 'adiabatic map' provides a qualitative indication of the effective potential surface for the motion of these two degrees of freedom. From the shape of this potential surface (figure 8) the reaction coordinate is seen to provide a good description of the most likely path of motion from the initial state to the final state of the system. The umbrella sampling calculations carried out using this reaction coordinate yield the results shown in figure 9. The potential of mean force $W(\xi)$ is related to the probability $\rho(\xi)$ by the equation

$$W(\xi) = -k_B T \ln \rho(\xi). \tag{5.3}$$

Here, W describes the work function or Helmholtz free energy variation of the system as a function of the reaction coordinate. The second curve in the picture describes the variation of the average potential energy of the system as a function of the reaction coordinate. From the relation $\Delta W = \Delta E - T\Delta S$, and the fact that $\Delta E = \Delta\langle V\rangle$ when T is constant, it is seen that the intrinsic entropy variation during the rotation is small. The barrier in the potential of mean force is somewhat smaller than what one would expect based on NMR lineshape analysis of the ring rotation; the latter yields an experimental barrier that is about 30% larger than the theoretical result (Wagner *et al* 1976). This quantitative discrepancy appears to be associated with the omission of solvent water from the initial calculations (Northrup *et al* 1982a). In the absence of hydrogen bonding and other interactions with the solvent, the surface of the protein (which includes part of the ring environment) changes somewhat from the expected solution structure; the slight distortions that result in the protein matrix are sufficient to account for the discrepancy.

The calculation of trajectories starting from the barrier region showed that a substantial fraction of these did not cross the barrier in the smooth, uninterrupted manner assumed in the 'ideal transition-state' model (see below). Frictional effects, particularly collisions between the ring and the local backbone, were found to interrupt

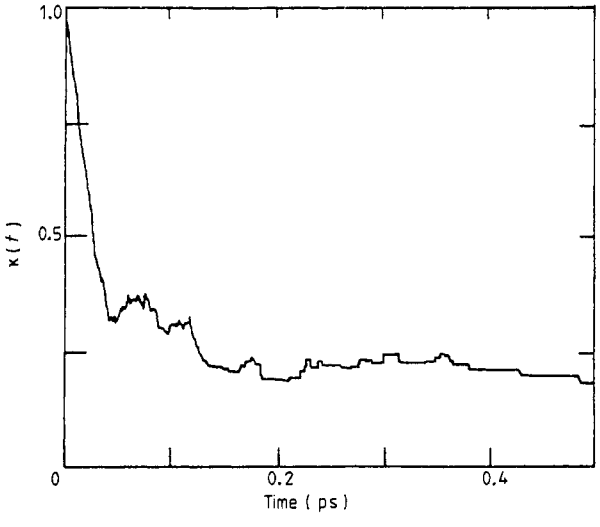

Figure 10. The normalised reactive flux time correlation function $\kappa(t)$ for the tyrosine-35 ring rotational isomerisation (Northrup *et al* 1982a).

a significant number of the crossing attempts. The transmission coefficient κ for the process was found to be about 0.2 (figure 10). Detailed analysis of the individual trajectories has helped considerably to clarify the mechanism for this isomerisation process (McCammon *et al* 1983). The ring rotation was found to be preceded by important structural fluctuations in the matrix around the ring. Particularly significant is a spontaneous displacement of a section of backbone that lies above one face of the ring. This displacement systematically precedes the ring rotation and appears to make two important contributions. The first effect of the backbone fluctuation is to displace atoms slightly from the path of the ring, resulting in a substantial reduction of the barrier to be crossed. Thus, the ring rotation appears to be gated by a collective motion in the surrounding matrix. During the rotation, bond angle deformations occur within the sidechain and adjacent backbone; the effect of these is to increase the distance of closest approach between the ring and N_{36}, resulting in a further reduction of the effective energy barrier. The second effect of the backbone fluctuation is to create a transient packing defect that helps to initiate the transition. The defect takes the form of a small volume (approximately 10 $Å^3$) into which the ring can rotate in response to collisions with the remaining matrix atoms. The resulting collisional bias allows the ring to accumulate the rotational kinetic energy necessary to surmount the residual barrier and is reminiscent of the mechanism of displacement of atoms in simple liquids (Rahman 1966). The coupling of the collective fluctuation in the protein matrix and the local isomerisation reaction is apparent in a sequence of structural 'snapshots', figure 11.

5.3. Nature of local activated processes

Studies of fast motions in proteins have shown that the atomic displacements resemble those observed in the liquid and solid states, as has been discussed in § 4. It is therefore useful to consider the general characteristics of local activated processes in liquids and solids as a reference point for the corresponding processes in proteins. Activated processes in condensed phases can be analysed in terms of equations (5.1) and (5.2). Processes for which the transmission coefficient $\kappa = 1$ correspond to an 'ideal transition-state theory' limit. This theory assumes that the frictional effects (e.g. collision frequencies) within a system are large enough to ensure that an equilibrium population of activated initial states is maintained in an ensemble of reacting systems and that systems that have crossed the energy barrier are quenched in the final-state region before they can rebound back across the barrier. It is also assumed that the frictional effects are small enough that systems crossing the barrier region in the direction of the final state are not interrupted and deflected back toward the initial state. These assumptions can not be perfectly realised in any real system, so that κ is always less than 1. Much work in chemical physics has been devoted to clarifying the nature of departures from the ideal transition-state theory (Kramers 1940, Chandler 1978, Skinner and Wolynes 1978, Northrup and Hynes 1980, Grote and Hynes 1980, Garrity and Skinner 1983).

In a liquid, the collision frequencies are sufficiently large that the rates of some processes may be reduced substantially by interruption of system trajectories in the barrier region. Such effects are most pronounced for the motion of large groups (which have large collision cross sections or friction coefficients) over broad energy barriers. For a highly damped, one-dimensional process in which the effective potential surface is parabolic in the initial, barrier and final-state regions, the rate constant can be

Figure 11. Stereoviews of atoms in the vicinity of tyrosine-35 at a sequence of times during the rotational isomerisation (McCammon *et al* 1983). From top to bottom in each panel, the times in picoseconds are (*a*) 0.50, 0.75, 1.00, (*b*) 1.25, 1.38, 1.50, (*c*) 1.62, 1.75, 2.00. The 'gate', which includes the backbone of residue 37, opens about 0.5 ps before the tyrosine-35 ring begins to rotate.

expressed as (Kramers 1940, Chandrasekhar 1943)

$$k = \kappa(\omega_i/2\pi)\exp(-\Delta W/k_B T) \qquad (5.4(a))$$

$$\kappa = \omega_b m/f. \qquad (5.4(b))$$

Here, ω_i and ω_b are the angular frequencies of oscillation associated with the initial well and inverted barrier, respectively, ΔW is the height of the potential of mean force barrier on the reaction coordinate, k_B is Boltzmann's constant, T is absolute temperature, m is the mass of the particle moving along the reaction coordinate, f is the friction coefficient of the particle and κ is a transmission coefficient which would be equal to one in the simple transition-state theory. The high damping limit is obtained, for example, for certain conformational changes of flexible molecules in solution; the rate constant is reduced from the transition-state value because such molecules cross the energy barriers in an erratic, Brownian manner (Levy *et al* 1979). The simple transition-state picture may also break down for processes that involve relative motion of charged groups in polar solvents (van der Zwan and Hynes 1982, Warshel 1982, Calef and Wolynes 1983). Here, the relative motion of the charges will be strongly influenced by fluctuations in the polarisation of the solvent. If one defines a primitive reaction coordinate in terms of the charge separation, the polarisation fluctuations will again lead to an erratic, Brownian motion along this coordinate. In extreme cases, the rate of the process will be controlled by the collective motions of the solvent dipoles.

In solids, one again observes certain significant departures from simple transition-state theory due to the nature of the underlying atomic motion. Consider, for example, vacancy diffusion in a simple monatomic crystal (Bennett 1975, Becker and Hoheisel 1982). Typically, an atom must squeeze between several nearest neighbours to move into the vacancy site. The necessary displacement of the neighbours occurs first and engenders strain in the lattice that provides the dominant contribution to the activation energy. The mobile atom then passes between its neighbours, experiencing much smaller non-bonded repulsions than would occur in the undistorted lattice. Detailed studies show that a significant number of barrier recrossings occur; the mobile atom retains enough kinetic energy to return to its initial site before the permissive configuration of the neighbouring atoms collapses.

Activated processes in proteins can be expected to display many of the properties observed in liquids and solids as well as additional properties that result from the unique structural features of these molecules. Processes that involve small displacements of non-polar groups are likely to be subject to relatively small frictional forces. The rotational isomerisation of a methyl group, for example, can be expected to have a nearly ideal transition-state character although some recrossing of the barrier may occur. Processes that involve large displacements of non-polar groups will generally be more complicated in several respects. First, the group will be subject to direct frictional effects that will act to reduce the rate constant. Second, the motion is likely to involve systematic coupling to collective distortions in the surrounding protein matrix. The covalent connectivity of the matrix atoms can be expected to result in more extensive collective distortions than would be seen in, for example, a corresponding small-molecule crystal. This property is apparent in the tyrosine ring isomerisation study described in § 5.2; in this process, the displacement of a sizeable region of polypeptide backbone is coupled to the local rotation. The extensive character of the collective component leads to a significant time lag between the backbone and ring motions; the collective displacement 'gates' the ring rotation. Such effects complicate

the interpretation of the activation energies and other experimental parameters because the collective motion is likely to make an important contribution. To analyse such processes, it may be essential to consider their intrinsic multidimensional character (Northrup and McCammon 1984) (figure 12). Because of the relatively weak interactions between discrete structural domains in many proteins, seemingly localised activated processes may, in some cases, be coupled to global motions of proteins, e.g. the transitions of groups located at the interface of two domains may be coupled to the relative motion of those domains. An extreme example is the coupling of ligand molecule displacements in hinge-bending proteins (cf § 6). In cases where the collective motion involves the protein surface, the rate of the local activated process is expected generally to display a dependence on solvent viscosity and other environmental factors that affect displacements of the surface (McCammon *et al* 1976, Beece *et al* 1980).

Activated processes that involve displacements of charged groups can be expected to display all of the above features plus a dependence on polarisation fluctuations within the protein and the solvent. As with reactions that involve charge reorganisation in polar solvents, such reactions may be limited by the frequency of occurrence of permissive environmental polarisations; this represents another example of collective 'gating' of local reactions.

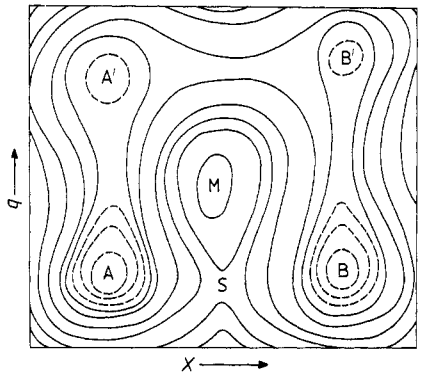

Figure 12. Schematic representation of the potential surface corresponding to a simple 'gated' reaction (Northrup and McCammon 1984). The spatial coordinates used to describe the reaction are a 'primitive' coordinate X and a 'gate' coordinate q. The local minima at A and B correspond to stable states within the 'gate-closed' conformation; A′ and B′ are corresponding states within the 'gate-open' conformation. M is a local maximum and S is a saddle point. Transitions from A to B will typically proceed through the intermediate states A′ and B′ rather than proceeding directly through the higher energy region around S. In many cases, motion along the primitive coordinate corresponds to local motion (e.g. ring rotation) while that along the gate coordinate corresponds to collective motion (e.g. displacement of a section of the protein adjacent to the ring). The overall rate of the A to B transition will therefore often reflect global influences (e.g. solvent viscosity, alterations of the protein rigidity that result from binding of effector molecules, etc).

6. Slow, extensive motions

6.1. Methods

Large-scale motions in proteins are generally slow and must be studied by techniques other than conventional molecular dynamics simulations. The fastest such motions are small-amplitude vibrations involving all or much of the molecule (de Gennes and

Papoular 1969, Suezaki and Gō 1975, McCammon *et al* 1976). Due to the large effective masses and solvent damping involved, these motions have characteristic times that often exceed 10 ps. Larger amplitude motions (e.g. the local denaturation of chain segments near the protein surface) may involve the diffusional exploration of a large number of low-energy conformations; such motions often have characteristic times in the microsecond to second range (Creighton 1978).

The type of method used to study slow, extensive motions depends on the complexity of the motion of interest. For low-frequency vibrations with amplitudes $\lesssim 0.1$ nm, normal-mode calculations can be used to identify some of the important features (Levy *et al* 1982b, Gō *et al* 1983). Quantitative difficulties arise, however, due to the importance of anharmonic effects (cf § 4.2). These difficulties can be circumvented by use of a quasiharmonic approach in which one considers slow oscillations near the minima of a potential surface that has been averaged with respect to the fast motions in the system. There are several different ways to do this. One can take the familiar approach of continuum mechanics, in which the system is characterised by suitable elastic constants (Young's modulus, Poisson's ratio) and damping parameters. The elastic constants describe the strain in the material in response to a slowly varying applied stress and therefore implicitly include the averaging with respect to fast motions. This approach has been used to characterise some of the low-frequency oscillations of spherical proteins, for example (de Gennes and Papoular 1969, Suezaki and Gō 1975). A second approach is to monitor the large-scale displacements that occur in a full molecular dynamics simulation, and to use this information to construct a potential of mean force for appropriate collective coordinates in the protein (Levy *et al* 1983). The resulting temperature-dependent potential typically has a quadratic minimum in the region of the native structure and can be used as a basis for 'effective' normal-mode calculations.

In cases where it is possible to focus on a specific mode or displacement path, but where the motions may be of large amplitude, there are again two useful methods. The first is the adiabatic mapping technique in which one displaces the system along the path of interest while minimising the energy of other degrees of freedom (McCammon *et al* 1976, Mao *et al* 1982b). The resulting potential may provide an approximation to the potential of mean force for displacements along the path. This approach has been used to study both large- and small-amplitude 'hinge-bending' motions; these motions involve the relative displacement of globular regions that are linked by flexible sections of a protein. Given the effective potential surface, the dynamics of the motion in the presence of solvent surroundings can be modelled by use of the Langevin equation (McCammon *et al* 1976, McCammon and Wolynes 1977). A second possible approach, which has not been applied to date, would be to calculate the potential of mean force more rigorously by using the type of approach described in connection with the sidechain isomerisation motions (cf § 5).

For large-amplitude motions where one is not able to specify the path in advance, it may be useful to apply Brownian dynamics simulation methods (Ermak and McCammon 1978). Brownian dynamics is the diffusional analogue of molecular dynamics and is an appropriate technique for the highly damped situations that are ordinarily encountered. Here, one initially simplifies the potential function by averaging over the rapid motions in the system and over displacements of solvent molecules and other groups whose detailed motions are not of interest. The degrees of freedom that have been averaged over are not explicitly represented in the final simulation model; rather, they are included in a 'thermal bath'. The simulation is then carried out by a

numerical integration of the Langevin equation; the bath is represented not only by the modified potential but also by suitable frictional and random forces. For extensive, activated processes where the motion involves barrier crossing, special methods analogous to those described in § 5.1 for inertial systems are useful (Northrup and McCammon 1980).

6.2. Hinge-bending motions

Many enzymes and ligand-binding proteins have two or more globular domains connected by relatively flexible strands of polypeptide (Ptitsyn 1978, Richardson 1981). The binding sites for substrate or ligand are typically located in a cleft between two such domains, so that the relative motion of the domains is likely to play an important role in binding and subsequent activity. Similar 'hinge-bending' motions may be important in the function of antibodies and certain muscle proteins.

The first protein for which a dynamical analysis of hinge-bending motions was attempted is lysozyme, a small enzyme with two globular domains separated by an active site cleft (McCammon *et al* 1976). An approximate potential for the hinge-bending was calculated by the following procedure. A bending axis was first identified by examination of the crystal structure of the protein. One lobe of the protein was then rigidly rotated through various angles about this axis. After each rotation, the outer parts of the two lobes were held fixed and the remainder of the protein was relaxed by use of an energy minimisation procedure. This relaxation process relieved local strains (e.g. close contacts between non-bonded atoms, bond-angle deformations). Because the characteristic time scale for these local motions (less than 0.1 ps) is much shorter than the time scale for the hinge-bending motion itself (greater than 10 ps, see below), the resulting adiabatic potential is a reasonable approximation to the effective potential for the hinge-bending. The effective potential is approximately quadratic in shape. An analysis of the calculated stress–strain ratio in terms of the equations for the deformation of a thick, linear beam shows that the Young's moduli are of the order of 10^{-11} dyn cm^{-2}, in the range expected for proteins, and that bending dominates shear by about an order of magnitude in determining the stiffness of the protein. The calculated force constant is such that fluctuations of the cleft width of the order of 0.1 nm are expected at room temperature.

Because the hinge-bending motion of lysozyme involves a substantial movement of the protein surface, it is essential to consider the solvent in a description of the hinge-bending dynamics (McCammon *et al* 1976, McCammon and Wolynes 1977). This was done within the framework of the harmonic oscillator Langevin equation by calculating the effective moment of inertia for the protein bending and the friction constant associated with solvent damping. The damping was treated by modelling the two globular domains as spheres and calculating the viscous frictional drag opposing the relative motion of these spheres by use of a modified Stokes law. The relative motion of the two globular domains in lysozyme was found to be overdamped, with a characteristic relaxation time of 20 ps. Thus, the hinge-bending motion is expected to be Brownian in character. A typical fluctuation will open or close the cleft by about 0.1 nm and will persist for approximately 20 ps.

Recent studies suggest that much larger hinge-bending motions occur in other proteins. For the L-arabinose-binding protein, calculation and experiment both indicate that the binding cleft, which is open in the unliganded state, is induced to close upon ligand binding; the two lobes of the protein swing through a relative angle of

approximately 30° upon closing (Mao *et al* 1982b). The open and closed structures of the protein are illustrated in figure 13.

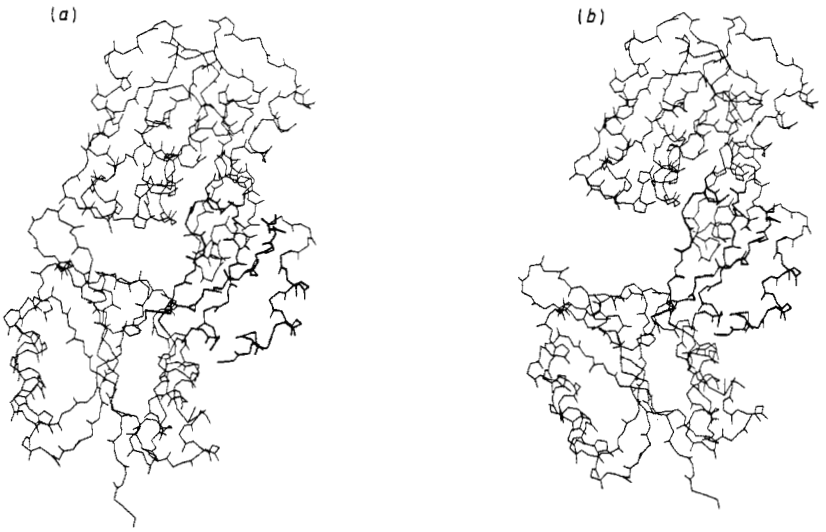

Figure 13. Structures of the L-arabinose-binding protein corresponding to hinge-bending angles of 0° (*a*) and 28° (*b*) (Mao *et al* 1982b). Sidechain atoms are omitted, and the hinge region is indicated by darker lines.

6.3. Single-strand motions

Certain large-scale motions in proteins involve structural changes that are less well defined than, for example, the hinge-bending motions described in the previous subsection. One example is the transient unfolding of a region of polypeptide chain at the protein surface. Such local denaturations may allow complete rotations of the backbone dihedral angles in the displaced polypeptide strand. Approximate simulations of such motions are possible using the Brownian dynamics procedure (Ermak and McCammon 1978).

The Brownian dynamics analysis of large-scale motions can be illustrated by recent studies of the growth of alpha helices in aqueous solution (McCammon *et al* 1980, Pear *et al* 1981). These calculations made use of a simplified structural model for the polypeptide chain. Each residue in the chain was represented by a single interaction site and the sites were linked by virtual bonds (Flory 1969). A set of energy parameters for the simplified model was developed by recognising that local motions such as sidechain rotations are faster than large-scale changes in the chain conformation (Levitt 1976, McCammon *et al* 1980). This separation of time scales was exploited to construct an approximate potential of mean force for the virtual bond rotations. The potential of mean force includes the average effects of the local motions and solvent interactions.

The large-scale motions of the polypeptide chain are subject to large frictional forces due to the high viscosity of the solvent. Inertial effects are negligible, so that the Brownian dynamics simulation approach is appropriate (McCammon *et al* 1980). To study the dynamics of helix coil transitions in polyvaline, a series of simulations was produced, each starting at the all-helix configuration. The diffusional motion of residues out of and back into the helical configuration was monitored. A pronounced end effect was observed; the rate constants for helix coil transition of the terminal

residue were about two orders of magnitude larger than those for residues closer to the middle of the chain. A more detailed analysis of the simulations revealed that adjacent residues often move from the helix to the coil or *vice versa* in nearly concerted fashion (Pear *et al* 1981). Such transitions are not consistent with the conventional idea that successive transitions occur independently (Schwarz and Engel 1972). It has

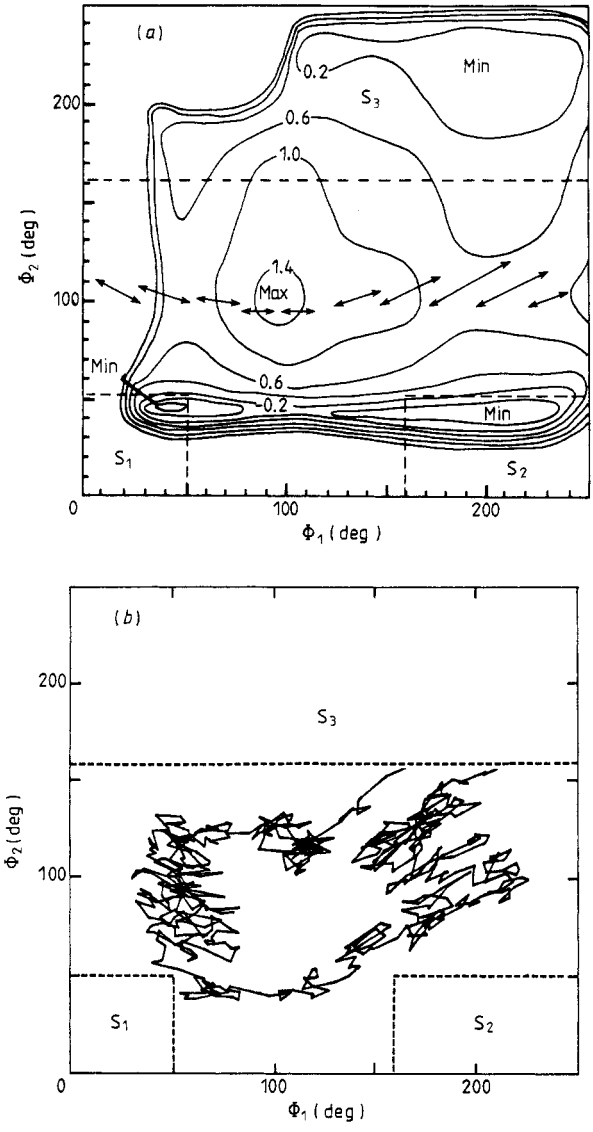

Figure 14. Configuration space defined by the virtual dihedral angles for the two residues at the end of an α-helical region of polyvaline (Pear *et al* 1981). (*a*) The potential of mean force with contour intervals of 0.4 kcal mol^{-1} (1 kcal = 4.184 kJ). The minima at S_1, S_2 and S_3 correspond, respectively, to the most stable configurations with all residues in the helix, the end-most residue out of the helix, and the two residues at one end out of the helix. Arrows indicate the preferred directions of diffusion (displacements subject to the smallest frictional forces in the solvent). (*b*) Diffusional trajectories for two helix-coil transitions. The trajectories begin just to the right of region S_1. Neither trajectory follows the commonly assumed sequence $S_1 \rightarrow S_2 \rightarrow S_3$.

been shown that the frequent occurrence of the correlated transitions results from the relatively small frictional forces associated with these motions (Pear *et al* 1980, 1981). The trajectories of two such transitions are shown in figure 14.

6.4. Nature of large-scale motions

Because large-scale motions typically involve substantial displacements of the protein surface, the coupling between the protein and its solvent surroundings is an important determinant of the nature of such motions. This coupling will, in general, influence both the probabilities of occurrence of different configurations of the protein and the rates of transition among these configurations.

The simplest kinetic effect of the solvent is the slowing of protein motions due to viscous damping. For the enzyme lysozyme, the lifetime of an open conformation of the active site cleft is increased from about 8 ps for the underdamped motion expected *in vacuo* to about 20 ps for the overdamped motion in water at room temperature (McCammon *et al* 1976). The hinge-bending motions in antibody molecules involve smaller restoring forces and larger frictional forces; the characteristic times for these motions are of the order of 10 ns (McCammon and Karplus 1977). In the case of motions that have an intrinsically multidimensional nature, the solvent damping effects may also change the preferred pathways of displacement from what would be expected based on the potential surface alone (Northrup and McCammon 1983). An example is the solvent-induced correlation in dihedral angle rotations shown in figure 14. It has recently been shown that the preferred path P of displacement for a highly damped system is that for which (Berkowitz *et al* 1983)

$$\int_{\mathrm{P}} f_{\mathrm{tt}} \exp(\beta W)\, \mathrm{d}s = \text{minimum} \tag{6.1}$$

where f_{tt} is the tangential component of the friction tensor along the path, $\beta^{-1} = k_{\mathrm{B}}T$ is the Boltzmann constant multiplied by temperature and W is the potential of mean force. This variational formula explicitly displays the dependence of the pathway of conformational change upon both frictional and potential effects. As will be discussed in § 8, such effects are expected to influence the mechanistic details of ligand-binding and other biological functions of proteins.

In the analyses of damping effects described above, the solvent has been treated as a viscous continuum. This simple model must be replaced by more sophisticated descriptions of the solvent in certain cases. For example, the large restoring forces involved in the lysozyme hinge-bending result in sufficiently rapid motions that the simple viscous response model for the solvent begins to break down (McCammon and Wolynes 1977). The solvent can still be modelled as a continuum, but one which has an inertial as well as viscous character. Other extensions of the continuum model are possible in principle, e.g. incorporating solvent dielectric features as has been done in studies of electrolyte solutions (Wolynes 1980). In some cases, it will be essential to introduce a detailed molecular model for the solvent. This will be required for protein motions that result in significant alterations of the surface exposure. Such motions will involve changes in the solvation of the surface groups with concomitant energetic and kinetic effects. An approximate consideration of such effects upon the hinge-bending motions of the L-arabinose-binding protein suggests that solvation changes of charged sidechains in the binding site cleft may regulate the overall conformation of this protein (Mao *et al* 1982b).

7. Experimental results

7.1. X-ray diffraction

X-ray diffraction studies provide information on the spatial distribution of electronic density in protein crystals. The method has been extensively applied to determine the average positions of the atoms in protein molecules from the corresponding peaks in the electron density. It has long been recognised that the shape of such peaks provides information on the distributions of the displacements of the individual atoms in the crystal. Recently, diffraction data from a number of proteins have been analysed in terms of the thermal motions of their constituent atoms (Frauenfelder *et al* 1979, Artymiuk *et al* 1979, Hartmann *et al* 1982, Glover *et al* 1983).

In the simplest case of isotropic, harmonic displacement, the width of an electron density peak is characterised by a single Debye–Waller factor (or 'temperature' factor), B. This factor is related to the apparent mean square atomic displacement by (Willis and Pryor 1975)

$$B = (8\pi^2/3)\langle(\Delta \boldsymbol{r})^2\rangle. \tag{7.1}$$

The apparent mean square atomic displacement includes a contribution $\langle(\Delta \boldsymbol{r})^2\rangle_{ld}$ from lattice disorder in the crystal and other effects in addition to the contribution $\langle(\Delta \boldsymbol{r})^2\rangle_{cv}$ for thermal motion (conformational and vibrational fluctuations) within a single molecule. The two terms may be of comparable magnitude at room temperature, but the small apparent variation of the disorder term allows one to estimate $\langle(\Delta \boldsymbol{r})^2\rangle_{ld}$ and make comparison with dynamical simulation results. In the one case where a detailed comparison has been made, the experimental and theoretical results were found to be similar (Northrup *et al* 1980, 1981), cf figure 15.

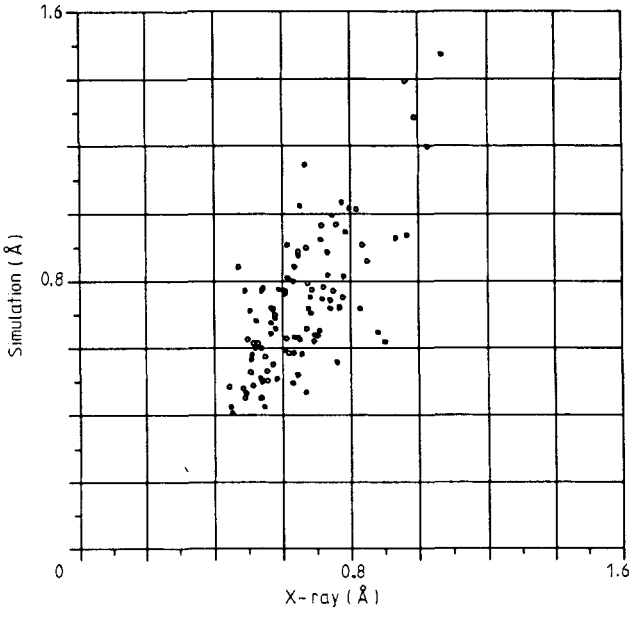

Figure 15. Correlation diagram for the calculated and experimental values of the average RMS atomic position fluctuation for each residue in cytochrome c (Northrup *et al* 1981). The correlation coefficient is $R = 0.73$.

Some effort has been made to go beyond the simple assumptions of isotropic and harmonic motion in the analysis of experimental results. These include x-ray structure refinements with partial or full analysis of the atomic displacement anisotropy (Artymiuk *et al* 1979, Konnert and Hendrickson 1980, Glover *et al* 1983). The results obtained to date are consistent with the theoretical finding that the atomic motion is, in general, quite anisotropic. Experiment and theory are also in agreement in concluding that collective motions within proteins tend to wash out the correlations that might be expected between local structure and the preferred direction of atomic displacement. Studies of the temperature dependence of the electron density peak widths provide valuable information on the character of the effective potentials constraining the atoms to their mean positions (Frauenfelder *et al* 1979, Hartmann *et al* 1982). Such studies show that the atomic displacements are generally anharmonic in character. Frauenfelder *et al* (1979) have characterised the effective atomic potentials by simple power law expressions. Among other findings, they show that the potentials of mean force for atomic displacement near the protein surface often have a square well character, in agreement with theoretical findings.

7.2. Nuclear magnetic resonance

Nuclear magnetic resonance is particularly valuable as a probe of protein dynamics (Campbell *et al* 1978, Gurd and Rothgeb 1979, Jardetzky 1981, Nagayama 1981, Andrew *et al* 1982). The method can be used on the normal solution state of proteins and, as a result of recent developments in spectral assignment techniques, can provide structural and dynamical information on specific parts of protein molecules. Moreover, the method is useful for study of both fast and slow local motions. Fast motions modulate the magnetic environment of nuclei at frequencies that stimulate nuclear magnetic relaxation. Thus, these fast motions are reflected in NMR relaxation times and in the nuclear Overhauser enhancement (NOE) factors that characterise the cross-relaxation of specific pairs of nuclei. Studies of ^{13}C nuclei have been particularly useful, because the magnetic-field fluctuations are often dominated by the relative motion of protons that are bonded to the carbon with a well-characterised geometry. NMR studies also provide useful information on the dynamical effects of solvent water at the protein surface (Bryant *et al* 1982, Andrew *et al* 1983).

In general, the detailed analysis of motion based on NMR data is complicated by the need to specify a physical model for the underlying dynamics. In this connection, the recent development of a theoretical formalism that allows model-free analysis is particularly useful (Lipari and Szabo 1982a, b). Application of this and more detailed approaches to ^{13}C NMR relaxation data from the small protein BPTI has shown that the motion of a number of groups is generally compatible with what has been found in dynamical simulations (Levy *et al* 1981, 1982a, Lipari *et al* 1982).

Local motions on long time scales can be probed by use of techniques that are sensitive to large displacements or 'exchange' of nuclei. The 180° rotations of tyrosine rings provide a good example (Campbell *et al* 1978). At low temperatures, the rings remain for a sufficiently long time in given orientations so that the delta and epsilon protons can be individually distinguished because of their different magnetic environments. The rings rotate more rapidly as the temperature is increased, producing spectral changes that lead eventually to a collapse of some of the lines when rapid rotation produces an equivalent average magnetic environment for the corresponding protons. Analysis of these spectral changes allows determination of the rate constants, activation energies and activation entropies for the ring rotations.

7.3. Mössbauer spectroscopy

There have been a number of attempts in recent years to use Mössbauer spectroscopy as a probe of protein dynamics (Knapp *et al* 1982, Bauminger *et al* 1983). This technique provides information on the mean square displacements of certain atomic nuclei during intervals corresponding to the lifetime of the nuclear excited state produced by gamma-ray absorption. Most investigations have focused on the displacements of iron atoms in haeme proteins during periods of 10^{-7}–10^{-9} s. As with x-ray diffraction and NMR studies, some difficulty arises in any attempt to specify a unique microscopic description of the atomic motion, although temperature-dependent studies help to reduce the number of possibilities. The conclusions emerging from these studies emphasise the dominant contribution of collective motions within the proteins to the iron displacements. For a number of proteins, the apparent displacements decrease dramatically below certain critical temperatures (200 K for metmyoglobin and deoxymyoglobin); this has been attributed to a freezing of the water at the protein surface and concomitant suppression of the important collective motions. Because the Mössbauer technique is not sensitive to low-frequency motions, it has been useful in the approximate separation of the static and dynamic contributions to the atomic displacements observed in x-ray diffraction studies (Frauenfelder *et al* 1979, Hartmann *et al* 1982).

7.4. Hydrogen exchange

Many of the hydrogens bonded to oxygen or nitrogen atoms in a protein will be replaced by deuterium if the molecule is suspended in heavy water. Because such exchange requires transient breaks in any hydrogen bonds at the exchange site and the presence of solvent species at this site, the hydrogen exchange phenomenon has long been recognised as an indicator of conformational fluctuations in proteins (Englander *et al* 1972, Woodward and Hilton 1979). The detailed character of the conformational fluctuations has been the subject of some debate, in part due to the difficulty of identifying which hydrogens are exchanged in a given period of time. Recent applications of proton NMR methods to this problem, together with advances in NMR assignment techniques, offer some hope of making hydrogen exchange a more detailed tool for probing protein dynamics (Hilton and Woodward 1978, Wüthrich and Wagner 1979, Wagner and Wüthrich 1979, 1982). The NMR techniques have more recently been complemented by neutron diffraction techniques, which indicate the locations of exchanged hydrogens in protein crystals soaked in heavy water (Kossiakoff 1982, Wlodawar and Sjölin 1982). Overall, the results of the recent studies are consistent in showing that the slowest exchange is associated with backbone hydrogens that are buried in the protein interior and centrally located in substantial elements of secondary structure. This is consistent with the observation by Levitt (1981) that there is some correspondence between the fluctuations in hydrogen bond lengths in dynamical simulations and the hydrogen exchange rates. Kossiakoff (1982) suggests that the slow exchange of buried hydrogens proceeds by a 'regional melting' mechanism that involves the local breaking of hydrogen bonds and the formation of a solvent-filled cleft to the protein surface. This model is intermediate between the traditional 'penetration' and 'local unfolding' models, which, respectively, suggest less and more disruption of the protein matrix. Although the dynamic interpretation of the hydrogen exchange results is still somewhat imprecise, an important finding in the recent work is that surprisingly rapid exchange is observed from some well-buried

atoms with small temperature factors in the x-ray diffraction data. Thus, despite the results of Levitt (1981) mentioned above, it is not always possible to extrapolate from the relatively rapid, low-energy motions measured by temperature factors (or RMS displacements in conventional dynamical simulations) to the slow, high-energy motions probed by hydrogen exchange.

7.5. Other experiments

It was shown a number of years ago that oxygen molecules can diffuse into the interior of proteins freely enough to quench the fluorescence of buried tryptophan sidechains on a nanosecond timescale (Lakowicz and Weber 1973). The apparent diffusion constant for oxygen in the interior of a protein is nearly as large as that for diffusion in bulk water. Many subsequent studies (summarised by Calhoun *et al* (1983b)) have examined the quenching of tryptophan fluorescence and phosphorescence by other molecules and ions. Calhoun *et al* (1983a, b) have provided results which indicate that, while oxygen can move through native proteins fairly freely, ions and polar molecules larger than oxygen are strongly excluded from the interior of native proteins and depend upon substantial disturbances of the protein structure for tryptophan contact and quenching.

Three new techniques for probing large-scale motions in proteins have recently been described. In crystals of proteins that have globular lobes connected by hinges, the pattern of crystallisation may be such as to produce mechanical anisotropy. This has been demonstrated for triclinic lysozyme crystals, which are relatively easily deformed by compressive forces that act in the direction corresponding to hinge-bending of the molecules in the crystal (Morozova and Morozov 1982). The apparent flexibility of the individual molecules is roughly consistent with a theoretical estimate (McCammon *et al* 1976). Other evidence of hinge-bending motions has recently been obtained in inelastic neutron scattering studies of the enzyme hexokinase (Jacrot *et al* 1982). Finally, evidence that the loops of the activation domain in trypsinogen are dynamically disordered has recently been obtained by analysis of the perturbed angular correlation of mercury atom labels (Butz *et al* 1982). The apparent characteristic time for the loop motion is about 11 ns, which is comparable to the time required for diffusional reorganisation of other polypeptide chains in simulation studies (McCammon *et al* 1980).

7.6. Comment on experiments

The studies mentioned in this section (and in § 8) represent a small sample from the total experimental effort on protein dynamics. More comprehensive reviews of this effort have been provided recently (see references in § 1.3). The purpose of this section has been to illustrate the type of information presently available from experiments and the degree of overlap with the theoretical work. It is clear that the x-ray diffraction and NMR studies provide the most information concerning the structural aspects of motions in proteins. Even in these studies, however, theoretical modelling has been helpful in the development of detailed structural interpretations. Theory is likely to play a larger role in the future with respect to the interpretation of other types of experiments. Among the experimental approaches mentioned here, inelastic neutron scattering has perhaps the greatest unrealised potential. By appropriate isotopic labelling studies, this approach should be able to provide the same type of detailed

dynamical information for proteins that it has for other dense materials (Leadbetter and Lechner 1979).

8. Protein dynamics and the kinetics of protein function

8.1. Ligand binding

The rate of initial encounter between a ligand and a protein is diffusion-controlled. For some proteins, this encounter rate limits the overall rate of protein function. This is true, for example, of enzymes that have been 'perfected' as catalysts by evolutionary selection (Knowles and Albery 1977, Brouwer and Kirsch 1982). The general factors that determine the diffusional encounter rate are understood (Calef and Deutch 1983), but research continues on the detailed effects of protein–ligand interaction potentials (Neumann 1981, Matthew and Richards 1982, Simondsen *et al* 1982, Chou and Zhou 1982, Matthew *et al* 1983), the saturation of binding sites at high ligand concentration (Berg and Ehrenberg 1983), and other factors. Shoup and Szabo (1982) have analysed the partially diffusion-controlled case in which there is a free energy barrier that prevents stable binding for some fraction of the protein–ligand encounters; they have provided a useful connection between the customary treatment in terms of a partially absorbing boundary condition and a more detailed treatment based on the average interaction potential between protein and ligand.

The internal motions of a protein may influence some details of its diffusional encounters with ligands (as a result of hydrodynamic coupling through the solvent, charge fluctuations, etc), but more important effects are expected in the subsequent binding steps. The structural fluctuations of the protein will generally include variations in the available volume of its binding sites, as well as in the disposition of charged, hydrogen bonding and other groups that are involved in ligand binding. The probability that a properly oriented ligand will actually be bound upon collision with the initial binding region at the protein surface will therefore display a time dependence. The character of the relevant binding site motions will depend on the particular protein and ligand involved. For the initial entry of molecular oxygen into proteins such as myoglobin or haemoglobin, relatively localised distortions of the protein matrix may be sufficient to open the necessary pathway (Frauenfelder *et al* 1979, Case and Karplus 1979). A number of enzymes have mobile surface loops that must adopt certain conformations to allow ligand entry and then other conformations to secure the ligand in place. Examples include the 'flap' regions at the active sites of penicillopepsin (James *et al* 1982, James and Sielecki 1983) and triosephosphate isomerase (Banner *et al* 1975) and the activator-binding loops in trypsin (Bode 1979, Huber and Bennett 1983). In yet other cases, hinge-bending motions involving the relative displacement of large globular regions are required to open binding sites (McCammon *et al* 1976, Anderson *et al* 1979, Newcomer *et al* 1981, Huber and Bennett 1983).

In all of the examples mentioned above, the protein fluctuates among a variety of conformations, only some of which allow ligand binding. Thus, the protein acts as a kind of gate that regulates access to the binding site. Such gating is often a necessary consequence of the functional design of a protein. The protein must be able to open the binding site easily to allow ligand entry and release, and yet be able temporarily to trap the ligand and (in the enzyme case) bring the necessary catalytic groups into the correct positions around the ligand.

The effects of gate fluctuations have recently been studied by analysis of the partially diffusion-controlled reaction of particles whose intrinsic reactivity fluctuates with time (McCammon and Northrup 1981, Northrup *et al* 1982b, Szabo *et al* 1982). In the simplest treatment, the reaction is described by the diffusion equation with a gated sink term:

$$\frac{\partial \rho}{\partial t} = r^{-2} \frac{\partial}{\partial r}\left(r^2 D \frac{\partial \rho}{\partial r}\right) - k_s h(t) \rho \delta(r-R) \tag{8.1}$$

and a reflecting wall boundary condition at $r = R$. Here, $\rho(r, t)$ is the density of ligands at a distance r from the protein at time t, D is the relative diffusion constant of protein and ligand, k_s is the specific rate constant in the gate-open (reactive) state, R is the ligand–protein contact distance and $h(t)$ is a characteristic gating function that fluctuates between values of 0 (gate closed) and 1 (gate open). The average binding rate that would be observed in a conventional experiment is characterised by the bimolecular rate constant:

$$k = \langle 4\pi R^2 k_s h(t) \rho(R, t) \rangle \tag{8.2}$$

where the brackets indicate a time average. The effects of gating on k are found to depend on such factors as the typical lifetimes of the gate-open and gate-closed states and the net rate of motion of ligands relative to the protein. In the limit of slow gate dynamics, k is just the rate constant for the gate-fixed-open case, multiplied by the fraction of the time that the gate is open. In other cases, substantial deviations from this 'intuitive' result occur.

8.2. Structural transformations

Structural transformations are involved in the function of many proteins. Examples include the atomic rearrangements that occur in protein and substrate during catalytic steps in enzymes (Lipscomb 1982), the shifting of enzymes or other proteins from less active to more active conformations upon the binding of substrates (Koshland 1963, Lipscomb 1982) or effector molecules (Monod *et al* 1963, Perutz 1970, Huber and Bennett 1983), and the large-scale structure changes in myosin that are involved in muscle contraction. Such transformations typically consist of one or more activated processes, in which the protein must cross free energy barriers that separate different possible stable states. The experimental rates of such processes have generally been analysed in terms of the 'thermodynamic' formulation of transition-state theory, in which the rate constant is expressed as (Fersht 1977)

$$k = (k_B T/h) \exp(\Delta S^{\ddagger}/R) \exp(-\Delta H^{\ddagger}/RT). \tag{8.3}$$

Here, k_B is Boltzmann's constant, T is absolute temperature, h is Planck's constant, R is the gas constant, $\Delta S^{\ddagger}$ is the entropy of activation and $\Delta H^{\ddagger}$ is the enthalpy of activation. The latter two quantities are determined by measuring the temperature dependence of the rate constant. Other similar quantities are determined by varying other conditions, e.g. measurement of the pressure dependence of the rate constant leads to an 'activation volume' defined by

$$\Delta V^{\ddagger} = -k_B T \left(\frac{\partial \ln k}{\partial \ln p}\right)_T. \tag{8.4}$$

Although this approach allows a useful reduction of experimental data in terms of a small number of parameters, recent studies have shown that such parameters must be interpreted with caution in any attempt to make contact with the microscopic dynamics of the system. Thus, $\Delta S^{\ddagger}$ and $\Delta H^{\ddagger}$ will not be simply related to the actual entropy and enthalpy costs of crossing a barrier if the actual enthalpy barrier exhibits a significant temperature dependence. Such a temperature dependence can arise, e.g. from thermal expansion and contraction effects that modulate steric hindrance in dense materials (Northrup *et al* 1982). Similarly, an activation volume does not in general correspond to a physical volume that must open up to allow an activated process to proceed (Karplus and McCammon 1981b). Because of the intimate coupling between local and collective motions in proteins (cf § 5.3), caution is also required in attempts to interpret parameters such as $\Delta H^{\ddagger}$ (or even the actual enthalpy barriers) in local terms (Northrup and McCammon 1984). For example, collective polarisation fluctuations appear to play a role in certain of the catalytic steps of lysozyme (Warshel and Levitt 1976).

A second type of difficulty in the simple thermodynamic transition-state theory approach arises from the neglect of non-equilibrium effects in equation (8.3). Frictional effects generally decrease the rate of a reaction from what would be expected based on the potential energy surface alone (cf § 5.3) and, in some cases, may even determine the dominant mechanism or pathway of reaction (cf § 6). That such effects are important in biology is clear, for example, from recent experiments on myoglobin reported by Frauenfelder's group (Beece *et al* 1980). In these experiments, the rate of rebinding of ligands (e.g. O_2 or CO) to the haeme iron is monitored spectroscopically following an initial laser flash that breaks the iron–ligand bond. Analysis of the temperature dependence of rebinding indicates that the ligand can hop among several stable locations within the protein before exiting to the solvent or rebinding to the iron. The stable locations are separated by energy barriers that are mostly of steric origin and that fluctuate in magnitude as a consequence of the normal internal motion of the protein atoms. The experiments show that the rate of hopping over each barrier depends on the solvent viscosity, even for barriers corresponding to transitions well within the protein interior. These results are most simply understood in terms of a coupling of the local ligand displacements to collective fluctuations that involve motion of the protein–solvent interface. In the language used elsewhere in this review (cf §§ 5 and 8.1), the ligand hopping can be viewed as a gated process and the dynamics of the gate is influenced by the viscosity of the solvent.

9. Summary and future directions

As the preceding sections will have made clear, there is ample evidence for a wide variety of motions in protein molecules. These molecules are constantly shifting among different conformations in the general neighbourhood of the native structure. Some of these conformations represent different stable structures or substates into which the protein can be frozen at sufficiently low temperatures. Such substates may be analogous to the defective lattice configurations or 'hidden structures' that are obtained upon quenching liquids (Stillinger and Weber 1982). From this point of view, a protein differs from a liquid primarily as a result of the constraints imposed by the covalent connectivity of the polypeptide chain. These constraints broaden the spectrum of structural relaxation times for the protein by increasing the energy barriers for larger

conformational transitions and introduce strong couplings between local and collective displacements.

It is also clear that the motions in protein molecules are important in terms of biological activity. The biological functions of protein mobility range from the transient occurrence of structures that allow ligand binding to the preferential population of active substates induced by the binding of effector molecules. In this connection, it is of interest that the regions of proteins that are involved in ligand binding and catalysis often exhibit thermal displacements of particularly large magnitude (Artymiuk *et al* 1979, Frauenfelder *et al* 1979, Huber and Bennett 1983).

Many of the theoretical techniques needed to study protein dynamics are now established and some key results have been obtained. More dramatic developments should appear in the forseeable future, as will be discussed below. These developments will require the extension of current methods in several directions. First, it will be essential to have a detailed treatment of the solvent surroundings of the protein (Rupley *et al* 1983). A useful step in this direction has been taken in a recent molecular dynamics simulation of protein and water (van Gunsteren *et al* 1983). Second, it will be important to extend the current energy functions with quantum-mechanical calculations so that covalent transformations in enzymes, electron transfer reactions and similar problems can be studied (Warshel 1981). Third, it will be necessary to develop appropriate simplified thermal-bath models so that detailed calculations need be performed only for the vicinity of an active site or other region of special interest (Berkowitz and McCammon 1982). Fourth, it will be important to develop techniques to determine optimum reaction coordinates to facilitate the analysis of activated processes (Northrup *et al* 1982a). These and other technical developments are the subjects of active investigation.

There are numerous areas for application of the existing and developing theoretical methods. These include fundamental aspects of protein dynamics such as the detailed characterisation of the space–time correlations involved in collective motions in proteins, the effect of solvent on these motions and the mechanisms of accumulation and disposal of kinetic energy during activated processes. Even more needs to be done in the general area of biological applications. The details of protein–ligand interactions, including the associated solvation changes, must be investigated. For particularly flexible binding sites, the effects of interactions between ligand and protein upon the gating transitions of the protein await study. The dynamical details of the protein structural transitions involved in ligand binding, activation and enzyme function have not yet been appreciably explored by theoretical means.

In time, one can expect theoretical protein dynamics to be of practical value in such areas as drug design and protein engineering. Recent attempts at drug design based on the examination of x-ray structures of receptor molecules have largely been frustrated by the delicate balance of the interactions involved (Abraham *et al* 1983), although incorporation of basic physicochemical data (Smith *et al* 1982) and molecular mechanics models that allow for some distortion of the binding sites (Deiters *et al* 1982, Wipff *et al* 1983) have proven helpful. The introduction of techniques for calculating free energy changes and rate constants, together with further developments in computing technology, are likely to make this an increasingly active area in the future. The advent of experimental methods for making arbitrary alterations in the amino-acid composition of proteins (Winter *et al* 1982, Ulmer 1983), together with the theoretical advances mentioned above, raise the intriguing possibility of designing proteins with predetermined functionality. This would, of course, have major implications in medicine and industry.

In his Dublin lectures of 1943, Erwin Schrödinger (1967) eloquently laid out much of the conceptual framework of molecular biophysics. This framework included the description of biological macromolecules as aperiodic crystals, and emphasised the functional importance of such dynamical processes as diffusion and activated structural transformations. Although Schrödinger's remarks were focused on genetic activity, it is now clear that protein molecules are well-accomodated within this same framework. Physicists and chemical physicists have made essential contributions to the explication of protein structure and dynamics. Their continuing contributions will increasingly illuminate the detailed nature of protein function.

Acknowledgments

The author has benefited greatly from discussions or collaborative efforts with many of the scientists whose names appear in the reference list. He is also grateful to colleagues who have sent preprints of their papers. The author's research has been supported in part by grants from the National Science Foundation, the National Institutes of Health and the Robert A Welch Foundation. He has received additional support from a Sloan Fellowship and from Dreyfus Teacher-Scholar and NIH Research Career Development Awards.

References

Abraham D J, Perutz M F and Phillips S E 1983 *Proc. Nat. Acad. Sci., USA* **80** 324
Anderson C M, Zucker F H and Steitz T A 1979 *Science* **204** 375
Andrew E R, Bone D N, Bryant D J, Cashell E M, Gaspar R and Meng Q A 1982 *Pure Appl. Chem.* **54** 585
Andrew E R, Bryant D J and Rizvi T Z 1983 *Chem. Phys. Lett.* **95** 463
Artymiuk P J, Blake C C F, Grace D E P, Oatley S J, Phillips D C and Sternberg M J E 1979 *Nature* **280** 563
Banner D W, Bloomer A C, Petsko G A, Phillips D C, Pogson C I, Wilson I A, Corran P H, Furth A J, Milman J D, Offord R E, Priddle J D and Waley S G 1975 *Nature* **255** 609
Bauminger E R, Cohen S G, Nowik I, Ofer S and Yariv J 1983 *Proc. Nat. Acad. Sci., USA* **80** 736
Becker K D and Hoheisel C 1982 *J. Chem. Phys.* **77** 5108
Beece D, Eisenstein L, Frauenfelder H, Good D, Marden M D, Reinisch L, Reynolds A H, Sorensen L B and Yue K T 1980 *Biochem.* **19** 5147
Bennett C H 1975 *Diffusion in Solids* ed J J Burton and A S Nowik (New York: Academic) p 73
Berg O G and Ehrenberg M 1983 *Biophys. Chem.* **17** 13
Berkowitz M and McCammon J A 1982 *Chem. Phys. Lett.* **90** 215
Berkowitz M, Morgan J, McCammon J A and Northrup S H 1983 *J. Chem. Phys.* at press
Bernstein F C, Koetzle T F, Williams G J B, Meyer E F, Brice M D, Rodgers J R, Kennard O, Shimanouchi T and Tasumi M 1977 *J. Molec. Biol.* **112** 535
Bode W 1979 *J. Molec. Biol.* **127** 357
Brouwer A C and Kirsch J F 1982 *Biochem.* **21** 1302
Bryant R G, Brown R D and Koenig S H 1982 *Biophys. Chem.* **16** 133
Butz T, Lerf A and Huber R 1982 *Phys. Rev. Lett.* **48** 890
Calef D F and Deutch J M 1983 *Ann. Rev. Phys. Chem.* **34** in press
Calef D F and Wolynes P G 1983 *J. Chem. Phys.* **78** 470
Calhoun D B, Vanderkooi J M and Englander S W 1983a *Biochem.* **22** 1533
Calhoun D B, Vanderkooi J M, Woodrow G V III and Englander S W 1983b *Biochem.* **22** 1526
Campbell I D, Dobson C W and Williams R J P 1978 *Adv. Chem. Phys.* **39** 55
Cantor C R and Schimmel P R 1980 *Biophysical Chemistry* (San Francisco: Freeman)
Careri G, Fasella P and Gratton E 1975 *CRC Crit. Rev. Biochem.* **3** 141
—— 1979 *Ann. Rev. Biophys. Bioengng* **8** 69
Case D A and Karplus M 1979 *J. Molec. Biol.* **132** 343

Chandler D 1978 *J. Chem. Phys.* **68** 2959
Chandler D, Weeks J D and Andersen H C 1983 *Science* **220** 787
Chandrasekhar S 1943 *Rev. Mod. Phys.* **15** 1
Chou K C and Zhou G P 1982 *J. Am. Chem. Soc.* **104** 1409
Cooper A 1981 *Sci. Prog., Oxf.* **66** 473
Creighton T E 1978 *Prog. Biophys. Molec. Biol.* **33** 231
Debrunner P G and Frauenfelder H 1982 *Ann. Rev. Phys. Chem.* **33** 283
de Gennes P G and Papoular M 1969 *Polarization, Matériere et Rayonnement* (Paris: Presses Universitaire de France)
Deiters J A, Gallucci J C and Holmes R R 1982 *J. Am. Chem. Soc.* **104** 5457
Edsall J T 1968 *Structural Chemistry and Molecular Biology* ed A Rich and N Davidson (San Francisco: Freeman) p 88
Englander S W, Downer N W and Teitelbaum H 1972 *Ann. Rev. Biochem.* **41** 903
Ermak D L and McCammon J A 1978 *J. Chem. Phys.* **69** 1352
Fersht A 1977 *Enzyme Structure and Mechanism* (San Francisco: Freeman)
Flory P J 1969 *Statistical Mechanics of Chain Molecules* (New York: Wiley)
Frauenfelder H and Marden M C 1981 *Physics Vade Mecum* ed H L Anderson (New York: AIP) p 108
Frauenfelder H, Petsko G A and Tsernoglou D 1979 *Nature* **280** 558
Garrity D K and Skinner J L 1983 *Chem. Phys. Lett.* **95** 46
Gelin B R and Karplus M 1975 *Proc. Nat. Acad. Sci., USA* **72** 2002
Glover I, Haneef I, Pitts J, Wood S, Moss D, Tickle I and Blundell T 1983 *Biopolymers* **22** 293
Gō N, Noguti T and Nishikawa T 1983 *Proc. Nat. Acad. Sci., USA* **80** 3696
Grote R F and Hynes J T 1980 *J. Chem. Phys.* **73** 2715
Gurd F R N and Rothgeb T M 1979 *Adv. Protein Chem.* **33** 73
Hansen J P and McDonald I R 1976 *Theory of Simple Liquids* (New York: Academic)
Hartmann H, Parak F, Steigemann W, Petsko G A, Ponzi D R and Frauenfelder H 1982 *Proc. Nat. Acad. Sci., USA* **79** 4967
Hilinski E F and Rentzepis P M 1983 *Nature* **302** 481
Hilton B D and Woodward C K 1978 *Biochem.* **17** 3325
Huber R and Bennett W S Jr 1983 *Biopolymers* **22** 261
Hynes J T 1977 *Ann. Rev. Phys. Chem.* **28** 301
Jacrot B, Cusack S, Dianoux A J and Engelman D M 1982 *Nature* **300** 84
James M N G and Sielecki A 1983 *J. Molec. Biol.* **163** 299
James M N G, Sielecki A, Salituro F, Rich D H and Hofmann T 1982 *Proc. Nat. Acad. Sci., USA* **79** 6137
Jardetzky O 1981 *Acc. Chem. Res.* **14** 291
Karplus M and McCammon J A 1981a *CRC Crit. Rev. Biochem.* **9** 293
—— 1981b *FEBS Lett.* **131** 34
—— 1983 *Ann. Rev. Biochem.* **52** 263
Kim P S and Baldwin R L 1982 *Ann. Rev. Biochem.* **51** 459
Knapp E W, Fischer S F and Parak F 1982 *J. Phys. Chem.* **86** 5042
Knowles J R and Albery W J 1977 *Acc. Chem. Res.* **10** 105
Konnert J H and Hendrickson W A 1980 *Acta Crystallogr.* A **36** 344
Koshland D E 1963 *Cold Spring Harbor Symp. on Quantitative Biology* vol 28 (New York: Academic) p 473
Kossiakoff A A 1982 *Nature* **296** 713
Kramers H A 1940 *Physica* **7** 284
Lakowicz J R and Weber G 1973 *Biochemistry* **12** 4171
Leadbetter A J and Lechner R E 1979 *The Plastically Crystalline State* ed J N Sherwood (New York: Wiley) p 285
Levitt M 1976 *J. Molec. Biol.* **104** 59
—— 1981 *Nature* **294** 379
—— 1982 *Ann. Rev. Biophys. Bioengng* **11** 251
Levy R M, Dobson C M and Karplus M 1982a *Biophys. J.* **39** 107
Levy R M, Karplus M and McCammon J A 1979 *Chem. Phys. Lett.* **64** 4
—— 1981 *J. Am. Chem. Soc.* **103** 994
Levy R M, Perahia D and Karplus M 1982b *Proc. Nat. Acad. Sci., USA* **79** 1346
Levy R M, Srinivasan A R, Olson W K and McCammon J A 1983 *Biopolymers* submitted
Lifshitz I M 1969 *Sov. Phys.-JETP* **28** 1280
Lifson S and Oppenheim I 1960 *J. Chem. Phys.* **33** 109
Lipari G and Szabo A 1982a *J. Am. Chem. Soc.* **104** 4546

Lipari G and Szabo A 1982b *J. Am. Chem. Soc.* **104** 4559
Lipari G, Szabo A and Levy R M 1982 *Nature* **300** 197
Lipscomb W N 1982 *Acc. Chem. Res.* **15** 232
McCammon J A 1976 *Models for Protein Dynamics* ed H Berendsen (Orsay: CECAM, Université de Paris IX) p 137
McCammon J A, Gelin B R and Karplus M 1977 *Nature* **267** 585
McCammon J A, Gelin B R, Karplus M and Wolynes P G 1976 *Nature* **262** 325
McCammon J A and Karplus M 1977 *Nature* **268** 765
—— 1979 *Proc. Nat. Acad. Sci., USA* **76** 3585
—— 1980a *Ann. Rev. Phys. Chem.* **31** 29
—— 1980b *Biopolymers* **19** 1375
—— 1983 *Acc. Chem. Res.* **16** 187
McCammon J A, Lee C Y and Northrup S H 1983 *J. Am. Chem. Soc.* **105** 2232
McCammon J A and Northrup S H 1981 *Nature* **293** 316
McCammon J A, Northrup S H, Karplus M and Levy R M 1980 *Biopolymers* **19** 2033
McCammon J A and Wolynes P G 1977 *J. Chem. Phys.* **66** 1452
McCammon J A, Wolynes P G and Karplus M 1979 *Biochem.* **18** 927
McQuarrie D A 1976 *Statistical Mechanics* (New York: Harper and Row)
Mao B, Pear M R, McCammon J A and Northrup S H 1982a *Biopolymers* **21** 1979
Mao B, Pear M R, McCammon J A and Quiocho F A 1982b *J. Biol. Chem.* **257** 1131
Matthew J B and Richards F M 1982 *Biochem.* **21** 4989
Matthew J B, Weber P C, Salemme F R and Richards F M 1983 *Nature* **301** 169
Monod J, Changeux J P and Jacob F 1963 *J. Molec. Biol.* **6** 306
Moog R S, Ediger M D, Boxer S G and Fayer M D 1982 *J. Phys. Chem.* **86** 4694
Morgan J D, McCammon J A and Northrup S H 1983 *Biopolymers* **22** 1579
Morozova T Y and Morozov V N 1982 *J. Molec. Biol.* **157** 173
Nagayama K 1981 *Adv. Biophys.* **14** 139
Neumann E 1981 *Structural and Functional Aspects of Enzyme Catalysis* ed H Eggerer and R Huber (Berlin: Springer-Verlag) p 47
Newcomer M E, Lewis B A and Quiocho F A 1981 *J. Biol. Chem.* **256** 13218
Northrup S H and Hynes J T 1980 *J. Chem. Phys.* **73** 2700
Northrup S H and McCammon J A 1980 *J. Chem. Phys.* **72** 4569
—— 1983 *J. Chem. Phys.* **78** 987
—— 1984 *J. Am. Chem. Soc.* at press
Northrup S H, Pear M R, Lee C Y, McCammon J A and Karplus M 1982a *Proc. Nat. Acad. Sci., USA* **79** 4035
Northrup S H, Pear M R, McCammon J A, Karplus M and Takano T 1980 *Nature* **287** 659
Northrup S H, Pear M R, Morgan J D, McCammon J A and Karplus M 1981 *J. Molec. Biol.* **153** 1087
Northrup S H, Zarrin F and McCammon J A 1982b *J. Phys. Chem.* **86** 2314
Paterson Y, Némethy G and Scheraga H A 1981 *Ann. NY Acad. Sci.* **367** 132
Pear M R, Northrup S H and McCammon J A 1980 *J. Chem. Phys.* **73** 4703
Pear M R, Northrup S H, McCammon J A, Karplus M and Levy R M 1981 *Biopolymers* **20** 629
Perutz M F 1970 *Nature* **228** 734
Peticolas W L 1978 *Methods Enzymol.* **61** 425
Ptitsyn O B 1978 *FEBS Lett.* **93** 1
Rahman A 1966 *J. Chem. Phys.* **45** 2585
Ransom-Wright L J and McCammon J A 1984 *Biopolymers* to be submitted
Richards F M 1977 *Ann. Rev. Biophys. Bioengng* **6** 151
Richardson J S 1981 *Adv. Protein Chem.* **34** 167
Rupley J A, Gratton E and Careri G 1983 *Trends Biochem. Sci.* **8** 18
Schrödinger E 1967 *What is Life*? (Cambridge: Cambridge University Press)
Schulz G E and Schirmer R H 1979 *Principles of Protein Structure* (Berlin: Springer-Verlag)
Schwarz G and Engel J 1972 *Angew. Chem. Int. Edn* **11** 568
Shoup D and Szabo A 1982 *Biophys. J.* **40** 33
Simondsen R P, Weber P C, Salemme F R and Tollin G 1982 *Biochem.* **21** 6366
Skinner J R and Wolynes P G 1978 *J. Chem. Phys.* **69** 2143
Smith R N, Hansch C, Kim K H, Omiya B, Fukumura G, Selassie C D, Jow P Y C, Blaney J M and Langridge R 1982 *Arch. Biochem. Biophys.* **215** 319
Stillinger F H 1980 *Science* **209** 451

Stillinger F H and Weber T A 1982 *Phys. Rev.* A **25** 978
Suezaki Y and Gō N 1975 *Int. J. Peptide Protein Res.* **7** 333
Swaminathan S, Ichiye T, van Gunsteren W and Karplus M 1982 *Biochem.* **21** 5230
Szabo A, Shoup D, Northrup S H and McCammon J A 1982 *J. Chem. Phys.* **77** 4484
Ulmer K M 1983 *Science* **219** 666
van der Zwan G and Hynes J T 1982 *J. Chem. Phys.* **76** 2993
van Gunsteren W F and Berendsen H J C 1982 *Biochem. Soc. Trans.* **10** 301
van Gunsteren W F, Berendsen H J C, Hermans J, Hol W G J and Postma J P M 1983 *Proc. Nat. Acad. Sci., USA* **80** 4315
van Gunsteren W F and Karplus M 1982a *Macromolec.* **15** 1528
—— 1982b *Biochem.* **21** 2259
Wagner G, DeMarco A and Wüthrich K 1976 *Biophys. Struct. Mech.* **2** 139
Wagner G and Wüthrich K 1979 *J. Molec. Biol.* **134** 75
—— 1982 *J. Molec. Biol.* **160** 343
Warshel A 1981 *Acc. Chem. Res.* **14** 284
—— 1982 *J. Phys. Chem.* **86** 2218
Warshel A and Levitt M 1976 *J. Molec. Biol.* **103** 227
Weber G 1975 *Adv. Protein Chem.* **29** 1
Welch G R, Somogyi B and Damjanovich S 1982 *Prog. Biophys. Molec. Biol.* **39** 109
Williams R J P 1980 *Chem. Soc. Rev.* **9** 325
Willis B T M and Pryor A W 1975 *Thermal Vibrations in Crystallography* (Cambridge: Cambridge University Press)
Wilson E B, Decius J C and Cross P C 1955 *Molecular Vibrations* (New York: McGraw-Hill)
Winter G, Fersht A R, Wilkinson A J, Zoller M and Smith M 1982 *Nature* **299** 756
Wipff G, Dearing A, Weiner P K, Blaney J M and Kollman P A 1983 *J. Am. Chem. Soc.* **105** 997
Wlodawer A and Sjölin L 1982 *Proc. Nat. Acad. Sci., USA* **79** 1418
Wolynes P G 1980 *Ann. Rev. Phys. Chem.* **31** 345
Wolynes P G and McCammon J A 1977 *Macromolec.* **10** 86
Wood D W 1979 *Water: A Comprehensive Treatise* ed F Franks (New York: Plenum) p 279
Woodward C K and Hilton B D 1979 *Ann. Rev. Biophys. Bioengng* **8** 99
Wüthrich K and Wagner G 1979 *J. Molec. Biol.* **130** 1
Zwanzig R and Ailawadi N K 1969 *Phys. Rev.* **182** 280