

articles

Dynamics of folded proteins

J. Andrew McCammon, Bruce R. Gelin & Martin Karplus

Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138

The dynamics of a folded globular protein (bovine pancreatic trypsin inhibitor) have been studied by solving the equations of motion for the atoms with an empirical potential energy function. The results provide the magnitude, correlations and decay of fluctuations about the average structure. These suggest that the protein interior is fluid-like in that the local atom motions have a diffusional character.

RESULTS of X-ray crystallography provide a picture of a globular protein in its native conformation as a well defined, densely-packed structure. Other experimental data¹⁻¹⁰ and theoretical considerations¹¹⁻¹³ indicate that there is considerable local motion inside a protein at ordinary temperatures. Moreover, the structural data themselves show that significant residue or subunit displacements have an important role in the activity of proteins (for example, enzyme catalysis¹⁴, haemoglobin cooperativity¹⁵, immunoglobulin action¹⁶). To obtain a more complete understanding of proteins, it is essential to have a detailed knowledge of their dynamics. In spite of the considerable effort directed toward protein folding¹⁷, very little has been done to investigate the motions of a protein in the neighbourhood of its equilibrium configuration. For certain cases in which the displacements along a suitably chosen coordinate can be isolated (for example, aromatic side chain rotations in the pancreatic trypsin inhibitor¹⁸, the opening and closing of the active site cleft in lysozyme¹⁹), it has been demonstrated that empirical energy functions can be used to analyse the motion involved. Here we undertake a more general examination of the internal dynamics of a folded globular protein. The approach used is of the molecular dynamics type¹⁹. In such a study the classical equations of motion for all the atoms of an assembly are solved simultaneously for a suitable time period and detailed information is extracted by analysing the resulting atomic trajectories. The full interatomic potential can be used to obtain the forces on the atoms, so that the method is applicable even when the system is highly anharmonic. The molecular dynamics approach has been very successful in revealing structural and dynamical characteristics of fluids¹⁹. As we demonstrate in what follows, molecular dynamics can have a corresponding role for the internal motions of proteins.

Bovine pancreatic trypsin inhibitor (PTI) (see Fig. 1a) was selected for study because of its small size (58 amino acid residues), high stability and accurately determined X-ray structure^{20,21}. Four water molecules, which are strongly bound in the interior, were included in the dynamical simulation. The potential energy was represented by an empirical function¹⁸ composed of a sum of terms associated with bond lengths, bond angles, dihedral angles, hydrogen bonds, and non-bonded (van der Waals and electrostatic) interactions. Hydrogen atoms are not explicitly considered, but are combined with the heavy atoms to which they are bonded by a suitable adjustment of

number of interactions which must be calculated and also permits larger steps in the trajectory calculation since the high frequency hydrogen vibrations have been eliminated. Integration of the equations of motion was performed by means of the Gear algorithm²² with time steps of 9.78×10^{-16} s. X-ray coordinates²¹ were used for the initial positions and the initial velocities were set equal to zero. After 100 equilibration steps, the stresses in the initial structure had partly relaxed and the system had an internal kinetic energy corresponding to a temperature of 140 K. At this point, all velocities were multiplied by a factor of 1.5 and 250 more equilibration steps were taken. The added kinetic energy ($250.6 \text{ kcal mol}^{-1}$) partitioned itself between kinetic and potential terms during this interval, and an average kinetic temperature of 285 K was reached. The actual simulation consisted of 9,000 additional steps, corresponding to 8.8 ps. Some equilibration of the bond lengths, bond angles and electrostatic interactions continues during the first 3 ps of the simulation, producing a small rise in temperature. Over the whole simulation, the average temperature is 295 K and the total energy is well conserved, changing by only $0.7 \text{ kcal mol}^{-1}$.

In what follows we first describe results concerned with the relation between the time-averaged and X-ray structures and with the magnitudes of structural fluctuations. We then examine the details of the dynamics, the correlation and damping of fluctuations, and some examples of the larger changes that occur. Of primary importance are the conclusions reached concerning the fluid-like nature of the internal motions, which will clearly have to be considered in developing dynamical models for biological processes.

Time-averaged structure and fluctuations

The time-averaged structure obtained in the dynamics run is near the X-ray structure but not identical with it; the root mean square (r.m.s.) deviation of the α carbons is 1.2 Å and that for all the atoms is 1.7 Å. The largest deviations come from the two ends of the molecule, the external loop (residues 25-28) and exposed sidechains. The two parts of the β sheet (residues 18-24, 29-35) and the α helix (residues 47-56) have significantly smaller deviations.

Use of the X-ray structure as the starting configuration corresponds to $0.4 \text{ kcal mol}^{-1}$ of kinetic energy per atom after equilibration (140 K); for the main dynamics run (295 K), the mean kinetic energy per atom is about $0.9 \text{ kcal mol}^{-1}$. Although there is no tendency to unfold, the dynamic development with this kinetic energy permits the molecule to make small rearrangements throughout; a picture of the structure after 3.2 ps is shown in Fig. 1b. Apparently the PTI molecule samples a series of neighbouring conformations, the average of which is not identical with the X-ray structure. It is not clear how much of the difference is a real effect (in the sense that the structure in the crystal is constrained) and how much is due to inaccuracies in the potential function and the absence of solvent. In lysozyme, comparison of the X-ray structures for

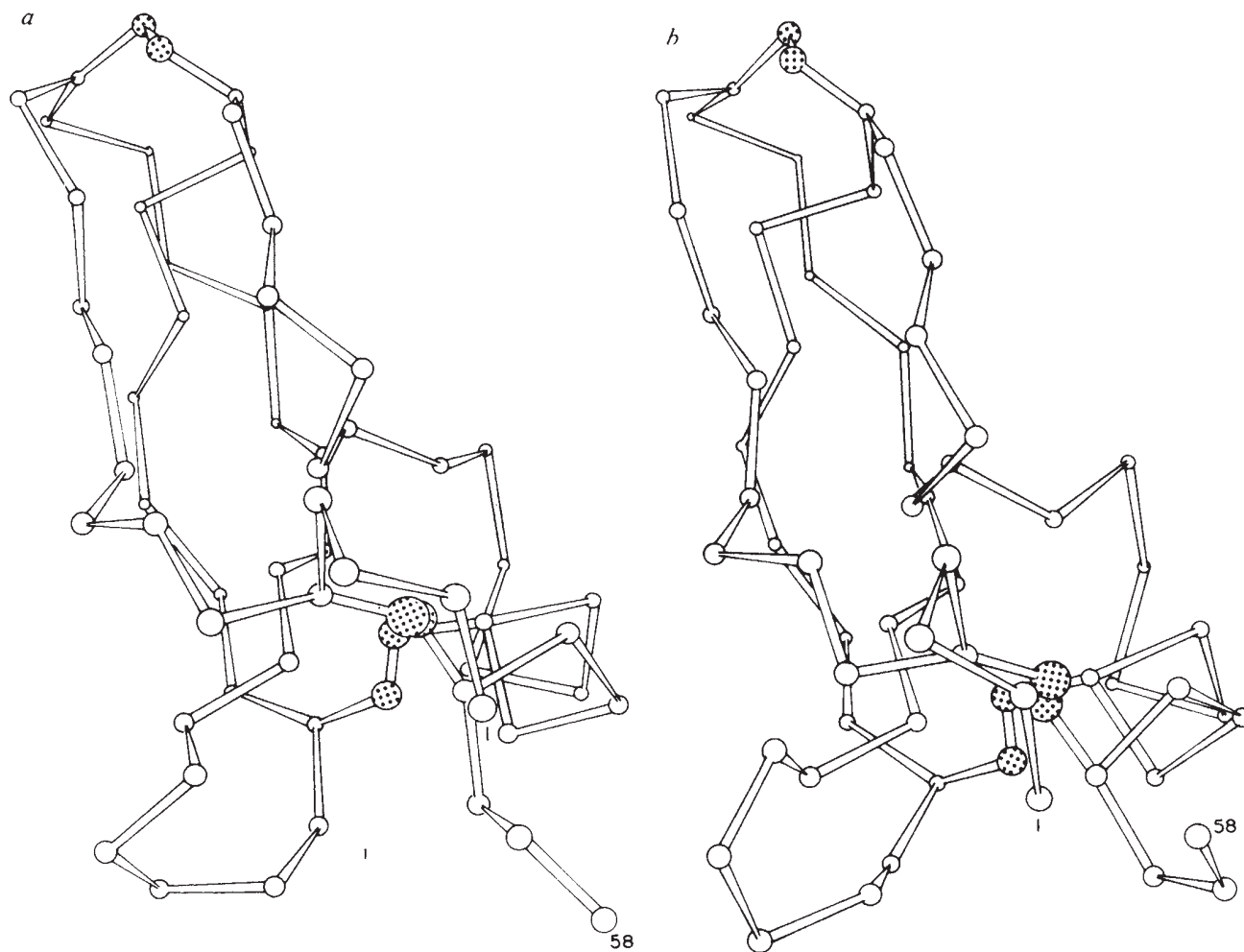


Fig. 1 The peptide backbone (α carbons) and disulphide bonds of PTI. *a*, X-ray structure²¹. *b*, Time evolved structure after 3.2 ps of dynamical simulation.

deviation of backbone atoms equal to 0.5 Å, with the largest changes at the surface of the protein. Also, a variety of other studies²⁵⁻²⁷ support the picture of a protein accommodating its structure to different environments. In any case, the X-ray and dynamic average structures are sufficiently similar that the present results should give a meaningful approximation to the short-time dynamics of PTI.

Figure 2 shows the r.m.s. displacements of the α carbons, relative to the dynamical average structure. The three residues at the carboxyl end of the chain have particularly large fluctuations; these are due to the presence of excess kinetic energy resulting from local strain in the initial structure. Some of this kinetic energy was transferred to the nearby amino end of the chain, which also exhibits relatively large r.m.s. displacements. The important fluctuations in residues 25-28 of the loop connecting the two strands of β sheet seem to reflect an intrinsic softness in this part of the molecule. The smaller fluctuations of the β sheet and α helical regions are also evident. The side-chain atom displacements are somewhat larger than those of the α carbons, but generally show a similar pattern of variation. The average r.m.s. fluctuation of all atoms is 0.9 Å. Particularly large r.m.s. displacements (1.8-3.1 Å) occur in the side-chains of Lys 26, Arg 39, Arg 42, Lys 46 and Met 52; all of these are at the surface of the protein. Overall, the large fluctuation regions correspond to those deviating most from the X-ray structure. In principle, the temperature factors from the X-ray analysis can be compared with the calculated

obtained from the experimental temperature factor is 0.4 Å. Also, some of the more mobile regions correspond to groups of atoms which were not located or had unusually high temperature factors (J. Deisenhofer, personal communication).

The range of variation of the internal coordinates in the static structure can be compared with the dynamical averages and fluctuations. Since the X-ray results for bond lengths and angles correspond to standard values for the different amino acids²¹, we use for comparison the energy-refined geometry (ERG)¹⁸, which reflects the effect of the environment on the coordinates. Selected examples for backbone coordinates are listed in Table 1. It can be seen that the ERG and dynamical averages and r.m.s. variations are in close correspondence. It should also be noted that the average r.m.s. dynamical fluctuation of any given coordinate is two to five times larger than the variation in the coordinate throughout the molecule. The results for the side chains are similar. For the dihedral angle ω , neither the ERG refinement nor the dynamic results show any significant distortions or unusual fluctuations for residues 14-17; this differs from the X-ray results²¹.

The dynamic averages for individual ϕ and ψ dihedral angles differ from the corresponding ERG results by less than 15° for 70 of the 115 backbone angles and by somewhat larger values for the rest (up to 30°). The behaviour of the sidechain dihedral angles is similar to ϕ and ψ .

The r.m.s. dynamic fluctuations of the bond lengths, the bond angles, and the ω angles are essentially constant along the backbone, except for a slight (20%) larger fluctuation in

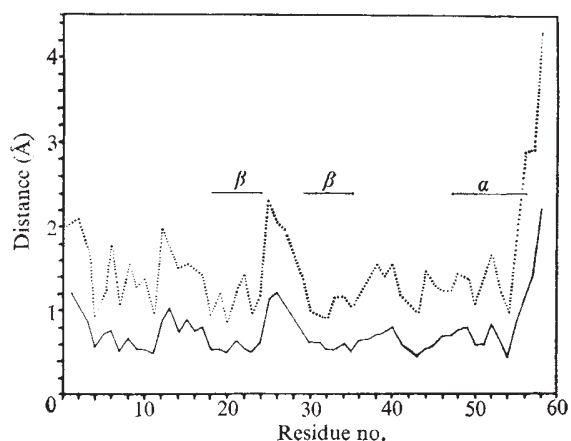


Fig. 2 R.m.s. displacements (—) and maximum displacements (.....) of the α carbons, relative to the dynamical average structure.

residues 56–58. For the ϕ and ψ dihedral angles, the r.m.s. dynamic fluctuations range from 10° to 40° and show more variation along the chain; in both the β sheet and α helix regions these fluctuations are about 18° on the average, while in the loop region the fluctuations average about 25° .

Bond lengths calculated from the dynamically averaged set of Cartesian coordinates were uniformly 0.02 – 0.03 Å shorter than the average bond lengths obtained directly. A similar discrepancy arises in calculating interatomic distances from X-ray diffraction data because of thermal atomic motion²⁸.

Time-dependence of motions

The fact that the individual PTI internal coordinates exhibit dynamic fluctuations of appreciable magnitude raises questions concerning the details of the motions, including their time evolution and spatial correlation.

To illustrate the time-dependence of individual coordinate fluctuations, we present results obtained for the Phe 22. This residue is relatively well buried in the protein and so should exemplify the effect of interactions between different parts of the molecule; other residues examined were found to have qualitatively similar dynamic behaviour. Figure 3a shows the time-dependence of several Phe 22 internal coordinates; the considerable variation in the frequencies, magnitudes of fluctuations, and irregularities of the behaviour of the different coordinates is evident.

To quantitate the average decay characteristics of fluctuations, so as to best analyse the coupling between the coordinate and the rest of the molecule, it is helpful to introduce time

correlation functions²⁹. The time correlation function $C_A(t) = \langle A(\tau+t)A(\tau) \rangle$, for a dynamical variable A is obtained by multiplying the value of $A(\tau)$ by $A(\tau+t)$, the value taken by A after the system has evolved for an additional time t , calculating such products for a representative set of initial times τ and averaging. If A is the fluctuation of a variable from its mean value, $\langle [A(\tau)]^2 \rangle$ is the mean-square fluctuation of the variable for an equilibrated system, while the time correlation function $C_A(t)$ describes the average decay behaviour of the fluctuation. Time correlation functions for the fluctuations of the internal coordinates of Phe 22 shown in Fig. 3a are given in Fig. 3b. Fluctuations of the backbone bond length $N_{22}-C_{22}^\alpha$ exhibit a partial loss of correlation in the first 0.1 ps, followed by a very slow decay of the residual small amplitude oscillation. Other backbone and sidechain bond lengths behave similarly. Although the specific oscillation frequencies depend on the force constants and the masses of the bonded atoms, all have long-lived oscillations; among others, the disulphide bond $S_{14}-S_{38}$ was examined and showed no atypical features. The other stiff internal coordinates (bond angles, dihedral angle ω) also exhibit partial loss of correlation in the first 0.1 ps but have persistent though less regular small-amplitude oscillation at longer times.

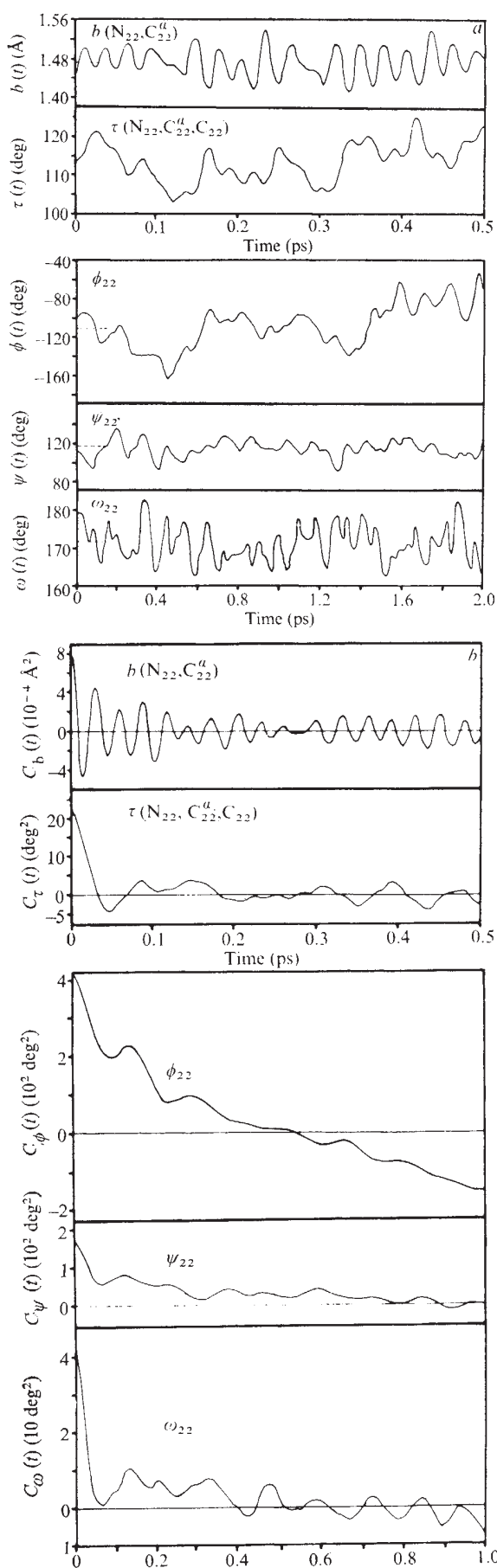
The relaxation behaviour can be analysed further by obtaining the Fourier transform of the correlation function and finding the dominant frequency components. For the $N_{22}-C_{22}^\alpha$ bond, the main peaks are found at 800 cm^{-1} , $1,010$ cm^{-1} , $1,110$ cm^{-1} , $1,170$ cm^{-1} , and $1,430$ cm^{-1} . These frequencies correspond to the calculated normal modes that involve the largest displacements of the $N-C^\alpha$ bond length of the isolated fragment $\text{CH}_3-\text{CO}-\text{Phe}-\text{NH}-\text{CH}_3$ (787 ; 952 ; $1,099$; $1,161$; and $1,425$ cm^{-1}). In addition, there are relatively small contributions from many low frequency motions. Thus, the $N_{22}-C_{22}^\alpha$ bond fluctuations are those expected for a strongly coupled system of local oscillators that is weakly coupled to the rest of the protein molecule. For the bond angles and the dihedral angle ω , there are important frequency contributions in the range 300 – $1,200$ cm^{-1} , which correspond to the fragment vibration frequencies associated with these degrees of freedom. Since frequencies below 300 cm^{-1} are also important, it seems that the bond angle and the dihedral angle ω fluctuations are significantly coupled to the protein matrix. Backbone dihedral angles (ϕ , ψ) are seen in Fig. 3b to exhibit considerable damping, though ϕ_{22} preserves long term correlation in a low frequency oscillation of the order of 2 ps. For the sidechain dihedral angle, χ_{22}^2 , associated with the rotation of the phenyl ring, damping produces a nearly monotonic decay with a relaxation time of 0.2 ps. It is clear from the decay behaviour of the correlation functions that the fluctuations of the ϕ , ψ and χ dihedral angles are dominated by interactions with the protein matrix. It is reasonable that torsional motions which involve substantial displacements of large groups inside proteins

Table 1 Backbone internal coordinates*

Internal coordinate	Average†		r.m.s. variation†		Average dynamical r.m.s. fluctuation‡
	ERG	Dynamics	ERG	Dynamics	
Bond lengths					
N-C $^\alpha$	1.476	1.478	0.005	0.004	0.028
C $^\alpha$ -C	1.531	1.533	0.004	0.004	0.027
C-O	1.237	1.237	0.002	0.003	0.015
C-N	1.321	1.323	0.003	0.003	0.018
Bond angles					
C-N-C $^\alpha$	124.6	125.8	1.6	1.7	4.8
N-C $^\alpha$ -C	113.8	114.9	2.9	2.4	4.8
C $^\alpha$ -C-N	115.4	114.6	1.8	1.7	4.6
C $^\alpha$ -C-O	119.7	118.7	1.5	1.5	4.6
Dihedral-angle ω	177.9	177.8	5.8	4.7	8.5

* Units are Å for bond lengths, degrees for bond and dihedral angles.

† Average and r.m.s. variation over the 58 amino acids of a given type of coordinate.



(for example, aromatic sidechains) will have a collective, diffusion-like behaviour, while those subject to smaller steric interactions will retain more of a local character.

The relaxation of individual atom fluctuations seems to be dominated by rotations of the dihedral angles. Atom displacement time correlation functions exhibit qualitative features and characteristic times similar to those shown for the dihedral angles. In many cases, the correlation functions show nearly monotonic decay with relaxation times of the order of a ps.

Correlated fluctuations

The large fluctuations found in the individual dihedral angles might seem to suggest a degree of structural mobility that is greater than that corresponding to the r.m.s. fluctuations of the individual atoms. In fact, the fluctuations of neighbouring dihedral angles are generally correlated so as to minimise disturbances of the backbone and sidechain atoms. For the typical backbone dihedral angle ϕ_{33} , which is at the centre of a strand of β sheet, the observed equal-time correlations of fluctuations of neighbouring ϕ angles and of neighbouring ψ angles are shown in Fig. 4. The large negative correlation $\langle \Delta\phi_{33}(\tau)\Delta\psi_{32}(\tau) \rangle$ is striking and has as a consequence the conservation of the general direction of the backbone and of the approximate position of the sidechains; the larger the individual dihedral fluctuations, the stronger the anti-correlation $\langle \Delta\phi_i(\tau)\Delta\psi_{i-1}(\tau) \rangle$ is found to be. The correlation pattern seen in Fig. 4 is similar to that obtained for an isolated harmonic α helix³⁰. A corresponding pattern of positive and negative dihedral angle correlations is likely to occur generally in strongly conserved regions of secondary structure. The anti-correlation $\langle \Delta\phi_i(\tau)\Delta\psi_{i-1}(\tau) \rangle$ is, in fact, present in other parts of the protein, though somewhat weaker outside the β sheet and α helical regions.

Structure-conserving correlations between dihedral angles also exist in the sidechains. For the eight aromatic sidechains with dihedral angles χ^1 and χ^2 , the normalised correlations $\langle (\Delta\chi^1\Delta\chi^2) \rangle / \langle (\Delta\chi^1)^2 \rangle^{1/2} \langle (\Delta\chi^2)^2 \rangle^{1/2}$ range from -0.05 to -0.63 ; again, dihedral angles with larger r.m.s. fluctuations tend to be more strongly anti-correlated. This correlation is the one required to conserve approximately the orientation of the aromatic rings with respect to the polypeptide backbone.

Concerted motions

In addition to the structure-preserving correlations, a number of concerted atom motions were observed. One of these involves the relatively flexible loop region (residues 25–29) which seems to oscillate as a whole with a period of approximately 6 ps, as well as undergoing changes in shape (see Fig. 1a, b). Some damping of this motion would be expected in solution due to the high solvent accessibility of the loop.

Time correlation functions involving different internal coordinates (for example, $\langle A(\tau+t)B(\tau) \rangle$) can be used to probe the dynamical character of non-local, collective motions in the protein. One such case concerns time correlations between the fluctuations of the dihedral angle ϕ_{33} in the β sheet and the fluctuations of other backbone and sidechain dihedral angles in the molecule. The results show that both strands of the β sheet and some additional residues participate in a low-frequency 'rippling motion' ($\nu = 15 \text{ cm}^{-1}$); within each strand of the β sheet, backbone dihedral angles tend to oscillate with the phase relationships suggested by Fig. 4. Examination of the correlation functions for the atoms involved also shows damped oscillations of the same frequency.

A feature of the dynamics is that certain groups in the molecule undergo transitions from one minimum to another. Such behaviour was found for a number of sidechains;

Fig. 3 *a*, Time development of selected Phe 22 internal coordinates; the initial time is 4.9 ps after the beginning of the dynamical simulation; average values are indicated by broken line segments. *b*, Time correlation functions for fluctuations of

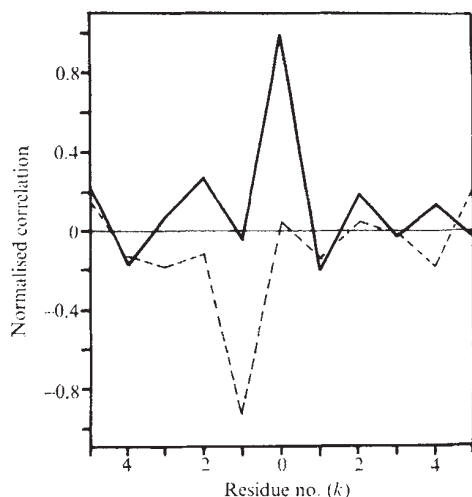
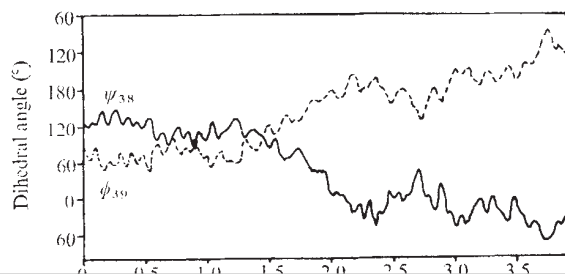


Fig. 4 Normalised equal-time correlations of fluctuations in ϕ_{33} with fluctuations in ϕ_{33+k} (—) and with fluctuations in ψ_{33+k} (---); the quantities plotted are, for example, $\langle \Delta\phi_{33}\Delta\phi_{33+k} \rangle / \langle (\Delta\phi_{33})^2 \rangle^{1/2} \langle (\Delta\phi_{33+k})^2 \rangle^{1/2}$.

for example, Met 52, where χ_{52}^2 moved successively from $-60^\circ \rightarrow +60^\circ \rightarrow -60^\circ \rightarrow 180^\circ \rightarrow 60^\circ \rightarrow 180^\circ$ and Arg 39, where χ_{39}^2 moved successively from $-60^\circ \rightarrow 180^\circ \rightarrow -60^\circ \rightarrow 180^\circ \rightarrow 60^\circ$. The times involved in these rotations vary between 0.2 and 2 ps. Since these sidechains are on the surface of the molecule, they have low barriers due to the protein; their detailed dynamics would be expected to be altered by interactions with the solvent.

In addition to the sidechain reorientations, there occurred a few concerted transitions in the backbone involving ψ_{i-1}, ϕ_i dihedral angle pairs. In these transitions, the anti-correlated angles ψ_{i-1} and ϕ_i rotate in the opposite sense by about 180° so that the intervening amide group flips over, but the backbone and sidechain directions are only slightly altered. The pairs involved are ψ_{38}, ϕ_{39} (near the 14–38 disulphide bond); ψ_{44}, ϕ_{45} ; and ψ_{56}, ϕ_{57} . All of these amide groups are located near the surface of the protein, so that the probability and dynamics of their transitions could be affected by solvent. Nevertheless, it is worthwhile to examine the behaviour since it can serve as a model of structural alterations in the presence of the protein matrix. Figure 5 shows the results for the ψ_{38}, ϕ_{39} pair. The transition is seen to take place in about 1 ps. During this period, the variation in the angles is roughly linear in time and the kinetic energy of the amide group does not show any significant departure from the thermal value. The amide flip is seen to be followed by a conformational adjustment period of 3–4 ps, during which ψ_{38}, ϕ_{39} drift over a relatively small angular interval into their new equilibrium ranges. It is apparent that there is a smooth change from one minimum to the other, while the protein relaxes so that the transition can occur without a large activation energy. The details of the driving torques and other matrix effects remain to be analysed, as do a variety of other larger scale motions that occur.

Fig. 5 Amide group transition; the time development of ψ_{38} and ϕ_{39} during the transition is shown.



Discussion

A molecular dynamics study of the pancreatic trypsin inhibitor has revealed a rich variety of motional phenomena that occur on the atomic level at ordinary temperatures. PTI has been shown to exhibit significant structural fluctuations, and detailed information on the magnitude, correlations and decay of these fluctuations has been obtained.

One important result of the dynamical simulation is that many aspects of the internal motion of PTI suggest that the folded protein is fluid-like at ordinary temperatures³¹. By this we mean that the fluctuations in atom positions, although confined to the neighbourhood of the average structure, have a diffusional character. In other words, the dynamics of atomic displacements are dominated by collisions with neighbouring atoms, at least on the picosecond time scale.

Also of significance is the heat capacity of the protein, which may be calculated from the magnitude of fluctuations in the kinetic energy³². Over the final 2 ps of the simulation, these fluctuations had the average value $\langle (\Delta KE)^2 \rangle = 87 \text{ kcal}^2 \text{ mol}^{-2}$; the corresponding heat capacity per atom is $C = 2.3 k$, where k is Boltzmann's constant. This result, which is midway between those expected for a harmonic solid ($C = 3k$) and for a hard-sphere fluid ($C = 1.5k$) indicates that the potential governing the atomic motions is significantly anharmonic. Similar values have been measured in real fluids; for example, at 84 K in liquid argon, $C = 2.32k$ (ref. 33). The calculated heat capacity of $0.33 \text{ cal g}^{-1} \text{ K}^{-1}$ is consistent with experimental values for proteins¹³.

An idealised model for proteins is that of a dense hard-sphere fluid composed of particles that are connected by flexible links. This implies that many of the dynamical properties (though not necessarily the correct average structure) can be obtained from any potential function which includes the forces that depend strongly on distance (covalent and hydrogen bonds, non-bonded repulsions) and provides sufficient attractive interactions to preserve the compact structure of the native system. Such a representation of a protein has much in common with the hard-sphere model of dense fluids, which has had considerable success recently in reproducing their equilibrium and dynamic properties³⁴. The close correspondence of the protein dynamics with that of the high temperature (300 K) solid phase of normal alkanes is also of interest^{35,36}. The importance of motions in which hard-sphere interactions dominate suggests that in conformational changes the non-bonded contacts will determine the transmission of effects from one part of the molecule to another. One such case of biological interest is the recent determination of the mechanism of tertiary structural change in a haemoglobin subunit, where non-bonded contacts were found to have the dominant role³⁷.

Although the time scale of the simulation is rather short, it provides essential details concerning the dynamics that should be of considerable utility in analysing longer time scale phenomena. A basic result is the r.m.s. atom fluctuation of 0.9 \AA . In a stochastic model that assumes independent atom motions, this provides a means for estimating the probability of larger fluctuations. Of corresponding significance is the nearly monotonic decay of the atom fluctuations with a time constant of $\sim 1 \text{ ps}$. It is also of interest that the kinetic energy of residues 49–58 remains 10–20% above the equipartition value for more than 9 ps; this indicates that vibrational energy can localise for relatively long time periods in proteins; this may be compared with the vibrational relaxation times in polyatomic fluids, which are often of the order of many picoseconds, even when efficient relaxation pathways exist³⁸.

It is clear from the dynamic simulation that the protein with room temperature kinetic energy samples highly anharmonic regions of the potential surface. Further, from an examination of the details of the coordinates during the run, as well as the dynamic average structure, it follows that the protein is able to wander over a range of conformations that are in the neighbour-

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts



Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research



With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips



Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.