Vinokurov, L. M. (1982) *FEBS Lett. 139*, 130–135.
Parmeggiani, A., & Sander, G. (1981) *Mol. Cell. Biochem. 35*, 129–158.
Riordan, J. F. (1973) *Biochemistry 12*, 3915–3923.
Rippa, M., Spanio, L., & Pontremoli, S. (1967) *Arch. Biochem. Biophys. 118*, 48–57.
Risuleo, G., Gualerzi, C., & Pon, C. (1976) *Eur. J. Biochem. 67*, 603–613.
Rohrbach, M. S., & Bodley, J. W. (1976) *Biochemistry 15*, 4565–4569.
Rohrbach, M. S., & Bodley, J. W. (1977) *Biochemistry 16*, 1360–1363.
Rohrbach, M. S., Bodley, J. W., & Mann, K. C. (1975) *J. Biol. Chem. 250*, 6831–6836.
Skar, D. C., Rohrbach, M. S., & Bodley, J. W. (1975) *Biochemistry 14*, 3922–3926.
Studier, F. W. (1973) *J. Mol. Biol. 79*, 237–248.
Tsou, C. L. (1962) *Sci. Sin. (Engl. Ed.) 11*, 1535–1558.

Time Dependence of Atomic Fluctuations in Proteins: Analysis of Local and Collective Motions in Bovine Pancreatic Trypsin Inhibitor†

S. Swaminathan, T. Ichiye, W. van Gunsteren,‡ and M. Karplus*

ABSTRACT: An analysis is made of the time dependence of the atomic motions obtained from two molecular dynamics simulations of the bovine pancreatic trypsin inhibitor; one simulation is for the protein in vacuum and the other for the protein in a van der Waals solvent with the atom size and density corresponding to those of water. Time series, correlation functions, and the time development of the mean square fluctuations are examined. A wide range of relaxation times is found for the displacement correlation functions (0.2–10 ps); the values of the relaxation times correlate with aspects of the structure and with the magnitude of the mean square displacements, as expected from the Langevin equation for an oscillator. It is shown that the atomic fluctuations which contribute to the temperature factor (thermal ellipsoid) can be separated into local oscillations superposed on motions with a more collective character. The former have a subpicosecond time scale; the latter, which can involve only a few neighboring atoms, a residue, or groups of many atoms in a given region of the protein, have time scales ranging from 1 to 10 ps or longer ($\bar{\nu} \simeq 3$–30 cm^{-1}). By following the time development of the atomic fluctuations from 0.2 to 25 ps, it is shown that the high-frequency oscillations, which contribute about 40% of the average root mean square fluctuations of main chain atoms, tend to be uniform over the structure. It is the longer time scale, more collective motions which introduce the variations in the fluctuation magnitudes that characterize different parts of the protein structure. These distributed modes are likely to be most sensitive to the external medium and to other environmental perturbations. Implications of these results for the function of proteins and analysis of anisotropic temperature factors are discussed.

For the interpretation of the structural and functional properties of proteins, it is essential to have a knowledge not only of the average positions of the atoms but also of the magnitudes and time scales of the fluctuations about the average positions (Karplus & McCammon, 1981). Recent X-ray diffraction studies of temperature factors for proteins have provided estimates of the magnitude of the fluctuations, expressed in terms of an isotropic and harmonic model (Frauenfelder et al., 1979; Artymiuk et al., 1979). The magnitudes of the mean square displacements calculated from molecular dynamics simulations have been shown to be in general agreement with the X-ray results (Northrup et al., 1980; van Gunsteren & Karplus, 1981, 1982a,b). Further, the simulation studies have demonstrated that the distribution functions for the atomic fluctuations tend to be highly anisotropic (Karplus & McCammon, 1979; Northrup et al., 1981, van Gunsteren & Karplus, 1981, 1982a,b) and somewhat anharmonic (Karplus & McCammon, 1979; Mao et al., 1982; van Gunsteren & Karplus, 1981, 1982a,b). X-ray estimates of the anisotropic character of the atomic motions have recently become available (Artymiuk et al., 1979; Konnert & Hendrickson, 1980; S. E. Phillips, unpublished experiments). By examining the temperature dependence of mean square displacements, Frauenfelder et al. (1979) have attempted to determine the anharmonic contributions to the local motions.

In the present paper, we focus on the time dependence of the atomic fluctuations that make the dominant contribution to the observed mean square displacements. These have been shown by simulation methods to occur on the picosecond time scale (McCammon & Karplus, 1980; Karplus & McCammon 1981). No data on the time scale of the motions are obtained from X-ray diffraction results, though other techniques (e.g., nuclear magnetic resonance, fluorescence depolarization, and Mössbauer spectroscopy) have shown that high-frequency fluctuations occur in proteins (Gurd & Rothgeb, 1979; Karplus & McCammon, 1981). We report here a detailed examination of the time dependence of the atomic fluctuations in the bovine pancreatic trypsin inhibitor (PTI), a small protein that has been used in a variety of motional studies (Karplus & McCammon, 1981). We analyze two simulations of this protein (van Gunsteren & Karplus, 1981, 1982a,b); the first simulation is for PTI in vacuum and the second for PTI in a van der Waals solvent composed of atoms with the size and particle density of water. The primary effect of the solvent

† From the Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138. *Received March 8, 1982.* Supported in part by the National Institutes of Health.
‡ Present address: Department of Physical Chemistry, University of Groningen, The Netherlands.

0006-2960/82/0421-5230$01.25/0

is that the average structure of the protein remains closer to the X-ray structure than it does in vacuum; the magnitudes of the fluctuations in the protein interior were found to be altered only slightly by the presence of solvent.

The time dependence of the motions in PTI is examined by making use of the time series for the atomic fluctuations, the correlation functions calculated from them, and the time development of the mean square displacements. It is shown that the fluctuations that contribute to the temperature factors are composed of a local subpicosecond librational component and a longer time, collective part that has contributions extending to 10 ps or longer. This is in accord with the results of a preliminary analysis of PTI (Karplus & McCammon, 1979) and of a recent simulation of an isolated α helix (Levy et al., 1982).

The present study of the time dependence of the positional fluctuations of the atoms complements the analysis by molecular dynamics simulations of experimentally measurable parameters, such as chemical shifts of aromatic residues (Snyder et al., 1975; Gelin & Karplus, 1975; Wagner et al., 1976; Campbell et al., 1976) and NMR relaxation times (Levy et al., 1981a,b).

Materials and Methods

The methodology of the molecular dynamics simulations of proteins has been presented previously (McCammon et al., 1979; Karplus & McCammon, 1981). Details of the specific PTI trajectories analyzed here are given by van Gunsteren & Karplus (1981, 1982a,b); the solvent run used periodic boundary conditions and surrounded the protein by 2647 van der Waals particles with radii and well depths corresponding to those in ST2 water (Stillinger & Rahman, 1972). The stimulations were done in the extended atom model with fixed bond lengths by means of the SHAKE algorithm; this eliminated the high-frequency, bond-length vibrational component which is of little interest for the fluctuations being examined. The analysis of the vacuum and solvent simulations covered a period of 25 ps; during this period, the average temperature, as measured by the kinetic energy of the protein, was equal to 305 K in the vacuum run and 306 K in the solvent run.

To analyze the time evolution of the atomic fluctuations, we examined time series, correlation functions, and time development of the mean square displacements. From the trajectories, 625 coordinate sets at 0.04-ps intervals were used in most cases; for the short time development of the mean square displacements, an augmented set at 0.008 ps was used. Time series for the Cartesian displacements [$\Delta x(t)$, $\Delta y(t)$, and $\Delta z(t)$] relative to the average position were calculated; the local principal axis system for the mean square displacements, corresponding to the thermal ellipsoid (Willis & Pryor, 1975; Dunitz, 1979), was employed. From these time series, correlation functions for the fluctuations were determined (Zwanzig, 1965; Kushick & Berne, 1977). For a variable A, the time correlation function, $C_A(t) = \langle A(s)\, A(s + t)\rangle_s$, is obtained by multiplying $A(s)$, the value of A at time s, by $A(s + t)$, the value taken by A at a time t later, evaluating such products for a set of representative initial times s and averaging the result. If the averaging is done over a sufficiently long dynamical simulation of an equilibrium system, $C_A(t)$ is independent of the initial times s used in the calculation; it is then customary to write $C_A(t) = \langle A(0)\, A(t)\rangle$. For the present study, the variable A represents the fluctuation of a quantity from its mean value, so that $\langle|A(0)|^2\rangle$ is the mean square fluctuation and $C_A(t)$ describes the average decay behavior of the fluctuation. If the fluctuations were to involve undamped simple harmonic motion, the correlation function would be correspondingly harmonic. In the interior of the protein, as in fluid media in general, it is expected that the motion of an atom will be perturbed by interactions with both bonded and nonbonded neighbors. As a result, the correlation between the displacement at a given time and a later time is disturbed, and the correlation function decays, often with intermediate oscillations that are indicative of the presence of important frequency components that contribute to the motions. Correlation functions have been used previously for interpreting molecular dynamics results for proteins, e.g., in the analysis of the tyrosine ring fluctuation of PTI (McCammon et al., 1979).

From the present simulation results, which provide values of A at N discrete times, t_n ($n = 1, 2, ..., N$), the normalized correlation $C_A(t)$, is calculated by means of the expression

$$C_A(t_n) = \frac{\sum_{m=1}^{N-n} A(t_m)A(t_m + t_n)}{\frac{N-m}{N}\sum_{m=1}^{N}[A(t_m)]^2} \qquad (1)$$

where $A(t_m)$ is the value of the variable at t_m. If A is a vector (e.g., $\Delta\vec{r}$), the product $A(t_m)A(t_m + t_n)$ is replaced by $\vec{A}(t_m)\cdot\vec{A}(t_m + t_n)$ in eq 1. We have examined the correlation functions for the displacement vector, $\Delta\vec{r} = (\Delta x, \Delta y, \Delta z)$ and its components. For estimation of the relaxation time of a given atomic fluctuation correlation function, a single exponential can be fitted to the initial decay of $C_A(t_n)$. Since the time dependence of the correlation function involves significant oscillations in many cases, the estimate from the initial decay can be improved by use of the time-integral definition (Zwanzig, 1965):

$$\tau = \int_0^\infty C_A(t)\, dt \qquad (2)$$

It has been pointed out in earlier analyses of the internal motions of a protein (McCammon et al., 1979; Karplus & McCammon, 1981) that the observed fluidlike behavior suggests that an equation of the Langevin type (Chandrasekhar, 1943) can be introduced to aid in interpreting the results. For the present case, in which the atoms are connected by relatively strong forces, a generalized Langevin equation (Adelman, 1980; Tully, 1981) would be required for a complete description. However, the simple harmonic oscillator Langevin equation can be used to obtain some insight into the nature of the motions. In one dimension, we have for the displacement Δx

$$m\frac{d^2\Delta x}{dt^2} + f\frac{d\Delta x}{dt} + k\Delta x = A_r(t) \qquad (3)$$

where m is the effective mass, f is the effective frictional coefficient, k is the effective force constant, and $A_r(t)$ represents the random force on the particle due to rapid fluctuations in the surrounding environment.

The correlation function corresponding to the Langevin equation for the harmonic oscillator (eq 3) has the form (Chandrasekhar, 1943)

$$C_x(t) = e^{-\beta t/2}\left[\cos \omega t + \frac{\beta}{2\omega}\sin \omega t\right] \qquad (4)$$

where $\beta = f/m$, $\omega = (\omega_0^2 - \beta^2/4)^{1/2}$, and $\omega_0^2 = k/m$. The relaxation time, τ, associated with the correlation function $C_x(t)$ is found from eq 2 and 4 to be

$$\tau = \int_0^\infty C_x(t)\, dt = \beta/\omega_0^2 = f/k = f\frac{\langle\Delta x^2\rangle}{k_B T} \qquad (5)$$

Table I: Relaxation Time Results

classes of atoms[a]	vacuum run[b]		solvent run[b]	
	$\langle\Delta r^2\rangle^{1/2}$ (Å)	τ ("ps")[c]	$\langle\Delta r^2\rangle^{1/2}$ (Å)	τ ("ps")[c]
all atoms (458)	0.72 (0.83)	1.54 (1.14)	0.78 (0.86)	1.92 (1.40)
backbone atoms (230)	0.57 (0.42)	1.65 (1.26)	0.57 (0.48)	1.79 (1.05)
side-chain atoms (228)	0.84 (0.79)	1.42 (0.98)	0.94 (0.72)	2.04 (1.69)
C^{α} (58)	0.54 (0.38)	1.68 (1.20)	0.54 (0.44)	1.94 (1.11)
C^{β} (52)	0.62 (0.39)	1.37 (0.81)	0.69 (0.65)	1.93 (1.57)
C^{γ} (42)	0.74 (0.58)	1.39 (0.86)	0.89 (0.81)	1.98 (1.89)
C^{δ} (39)	0.86 (0.84)	1.50 (1.18)	1.18 (1.00)	2.13 (1.62)
internal water O (4)	1.07 (0.66)	1.60 (0.89)	0.53 (0.28)	0.96 (0.59)
backbone O (57)	0.66 (0.48)	1.54 (1.28)	0.66 (0.56)	1.58 (1.03)

[a] Values in parentheses correspond to the number of atoms in the class. [b] Values in parentheses correspond to the rms deviations in the quantity. [c] The symbol "ps" represents 0.98 ps, a natural unit of time relative to the other units used in the molecular dynamics simulation (van Gunsteren & Karplus, 1982a,b).

where the last equality follows from the classical equipartition result for a harmonic oscillator. From eq 4 and the expression for ω, it is evident that there are two important regimes for the decay of the correlation function and the motion to which it corresponds. If ω is real (i.e., $\omega_0^2 > \beta^2/4$), the system is "underdamped" and eq 4 can be used directly. It shows that the correlation function is oscillatory with frequency ω and decays to zero with damped oscillations. For such an underdamped system, in which the effective force constant is large and the frictional drag is relatively weak, the particle would relax to its equilibrium position by damped oscillations in the absence of random forces. If ω is imaginary ($\omega_0^2 < \beta^2/4$), it is convenient to rewrite eq 4 in the form

$$C_x(t) = e^{-\beta t/2}\left[\cosh \omega' t + \frac{\beta}{2\omega'} \sinh \omega' t\right] \quad (6)$$

where $\omega' = (\beta^2/4 - \omega_0^2)^{1/2}$. Equation 6 shows that in this case the correlation decays monotonically to zero. For such an overdamped system, in which the effective force constant is small relative to the frictional drag of the medium, the particle would relax to its equilibrium position without oscillating in the absence of random forces. Actually, the particle experiences random forces due to collisions with the surrounding protein matrix in both the overdamped and underdamped cases, so that Brownian fluctuations are expected to be superposed on the ideal decay behavior.

In the following section, we apply the Langevin equation to interpret the forms of some of the correlation functions obtained for the atom fluctuations from the dynamic simulation. Both the overdamped and underdamped cases occur in the simulation results for PTI, and the Langevin equation provides relations between the properties of the correlation functions and physical characteristics of the system. However, because of the complicated nature of the dynamics, the correlation functions for the underdamped motions deviate considerably from the simple Langevin expression (eq 4) in many cases. Of particular interest are correlation functions whose low-frequency oscillations do not show the exponential damping expected from the Langevin equation. For such motions, a model involving a superposition of undamped normal mode oscillations can provide a useful limiting description. The form of the correlation function in one dimension is given by (Wang & Uhlenbeck, 1945; Levy & Karplus, 1979)

$$C_x(t) = k_B T \sum_i \frac{\alpha_i^2}{\omega_i^2} \cos \omega_i t \quad (7)$$

where ω_i is the frequency of the ith normal mode, Q_i, expressed in terms of mass-weight Cartesian coordinates, and α_i is the component of Q_i in the direction of the displacement coordinates, Δx:

$$\Delta x(t) = \sum_i \alpha_i Q_i(t) \quad (8)$$

For this case, the mean square fluctuation is

$$\langle\Delta x^2\rangle = k_B T \sum_i \frac{\alpha_i^2}{\omega_i^2} \quad (9)$$

As we see in the following, a useful model for many of the atomic fluctuations is given by the superposition of two or more modes, one local and high frequency in character, the others more global and of lower frequencies. The frequencies and effective masses associated with two types of modes are sufficiently different that the interaction between them is small and a Born–Oppenheimer-type uncoupling provides a good approximation (Pauling & Wilson, 1935). For the simplest case of two particles (an atom attached to an effective "particle" composed of a group of atoms undergoing collective motions), the high-frequency motion of the atom is calculated by assuming that it is moving relative to the effective particle, which is at rest, and the motion of the latter is treated independently, and its effect is introduced as a time-dependent change in the average position of the atom. An analogous approach has been used in the vibrational analysis of smaller molecules in which the frequencies separate into two groups (Wilson et al., 1955).

Either introduction of damping effects into the coupled mode description or coupling between oscillators in the Langevin model would lead to the generalized Langevin equation approach (Wang & Uhlenbeck, 1945; Peticolas, 1979; Adelman, 1980; Tully, 1981; Berkowitz & McCammon, 1981). However, the limited length of the run and the actual complexity of the protein motions (anharmonic and anisotropic, superposition of many modes) are such, in accord with the observed form of the time series and correlation functions, that the more complete models do not at this stage provide greater insight into the dynamics than the limiting cases considered here.

Results and Discussion

In this section, we present the results for the time development of the fluctuations and relate them to the structure of the protein. For many aspects of the time dependence, we focus on the solvent run since the results obtained from it are likely to be most realistic. To demonstrate the effect of the solvent, we compare the vacuum and solvent results where appropriate.

Relaxation Times. To obtain a general idea of the time dependence of the positional fluctuations, we survey the es-

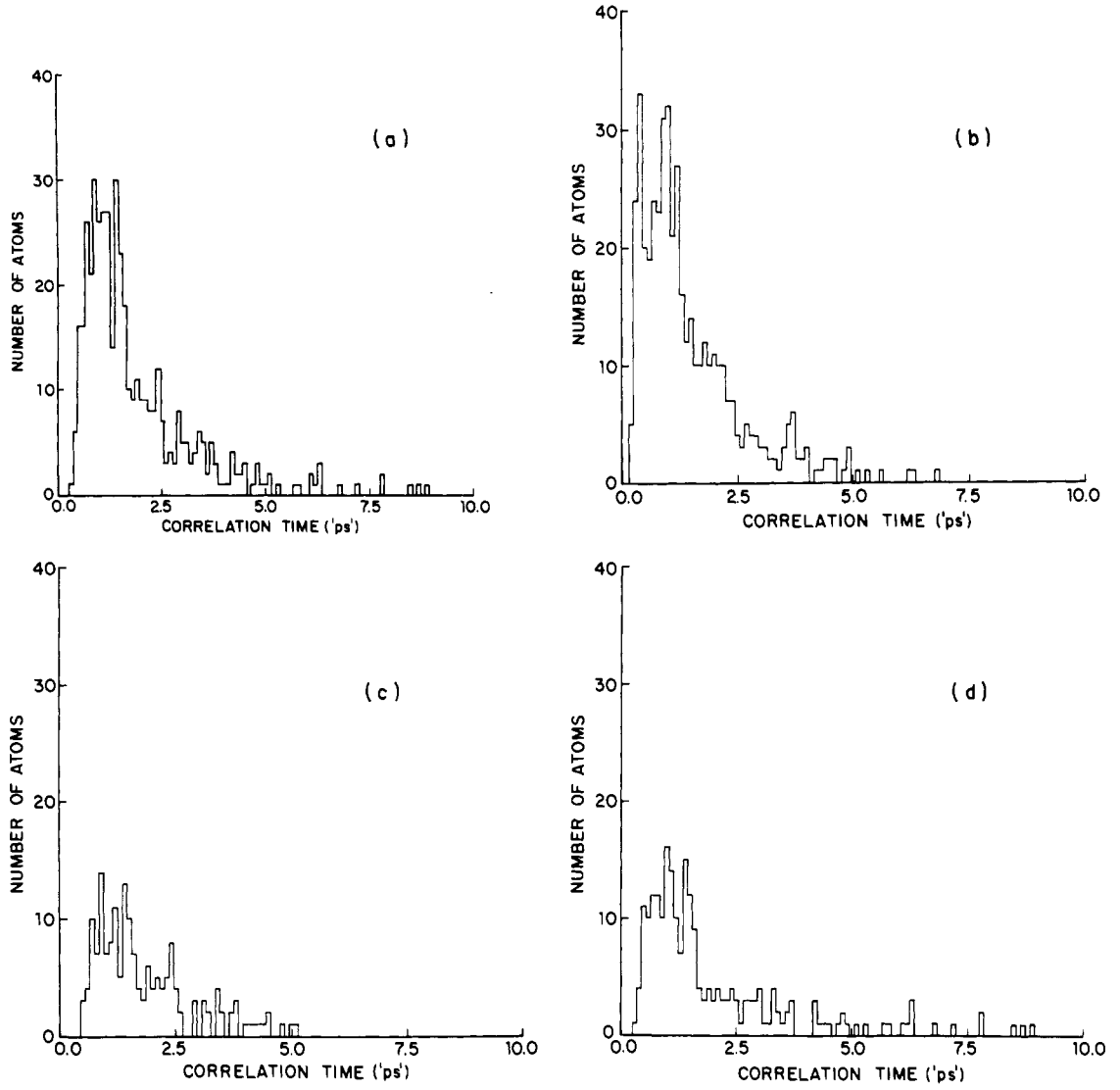

FIGURE 1: Histogram of relaxation time (τ) distribution showing the number of atoms with a given value of τ as a function of τ: (a) all atoms, solvent run; (b) all atoms, vacuum run; (c) backbone atoms, solvent run; (d) side-chain atoms, solvent run.

timated relaxation times associated with the correlation functions for the vectorial displacements, $\Delta\vec{r}$, of all the PTI heavy atoms in the solvent and vacuum runs. Table I shows the values for the root mean square (rms) fluctuation magnitudes, $\langle\Delta r^2\rangle^{1/2}$, and relaxation times, τ, averaged over various classes of atoms in PTI; the rms deviations from the mean are also given. It can be seen that the overall time scale of the fluctuations is on the order of 1 ps, in accord with previous results (Karplus & McCammon, 1981; van Gunsteren & Karplus, 1981, 1982a,b). However, the range of values is quite large; for the solvent run, the shortest relaxation time is 0.4 ps (Asn-24, terminal NH_2), and the longest is 8.9 ps (Asp-3, carboxyl oxygen). Figure 1 gives histograms of the number of atoms as a function of the relaxation time; histograms for all the atoms in the solvent and vacuum runs are given, as well as separate plots for the backbone (N, C^α, C) and the side-chain atoms from the solvent run. Most of the atoms have τ values in the range between 0.4 and 2.5 ps, although the distribution is skewed toward longer times. For the vacuum run, the range of relaxation times is shifted toward shorter values; i.e., the smallest τ value is 0.19 ps (Thr-32, hydroxyl oxygen), the maximum τ value is 6.5 ps (Leu-6, carbonyl C), and there are 22 atoms with $\tau < 0.35$ ps. There seems to be no association of short relaxation times with backbone vs. side-chain atoms; as to the very long relaxation times, they are primarily associated with side-chain atoms.

A number of characteristics of the relaxation time distributions are evident from the mean values for different atom types given in Table I. In the solvent run, the τ values for the side-chain atoms are longer than those for the backbone; this is not true for the vacuum run. There is some tendency of the relaxation time to increase as one goes out along a side chain. The averages for the backbone atoms N, C^α, and C are very similar, while the backbone carbonyl oxygen value is somewhat shorter. Three of the internal water molecules, which are localized in one part of the structure and have hydrogen bonds to each other, have relaxation times less than 1 ps, while the fourth internal water molecule, which is isolated, has a τ value of 1.9 ps, in the range of that of the surrounding residues.

Table II shows the relaxation time and rms fluctuation averages obtained by dividing the atoms into groups according to secondary and tertiary structural characteristics. Values are given for groups of atoms occupying concentric shells of 3-Å thickness about the center of mass. The relaxation times of the solvent run have a slow increase in the interior of the protein with distance from the center (up to 9 Å) and then rise more rapidly as the solvent-exposed exterior is reached. A scatter plot (Figure 2) of τ vs. the distance R from the centroid shows the relationship most clearly; there are short relaxation times ($\leq$2.0 ps) at all distances from the center, but

Table II: Fluctuation Behavior as a Function of Structure

	vacuum[b]		solvent[b]	
	$\langle\Delta r^2\rangle^{1/2}$ (Å)	τ ("ps")[c]	$\langle\Delta r^2\rangle^{1/2}$ (Å)	τ ("ps")[c]
shells from centroid (Å)				
0–3	0.71 (0.69)	1.81 (0.61)	0.62 (0.61)	1.15 (0.32)
3–6	0.64 (0.59)	1.71 (0.65)	0.53 (0.48)	1.16 (0.65)
6–9	0.61 (0.52)	1.62 (1.04)	0.54 (0.45)	1.38 (0.82)
9–12	0.76 (0.93)	1.44 (1.14)	0.73 (0.71)	1.94 (1.17)
12 to ∞	0.82 (0.88)	1.48 (1.34)	1.06 (1.02)	2.61 (1.84)
secondary structure[a]				
β-sheet backbone (63)	0.49 (0.32)	1.52 (0.93)	0.52 (0.42)	1.69 (0.85)
β-sheet side chain (88)	0.83 (0.80)	1.63 (0.99)	0.85 (0.94)	1.33 (0.88)
α-helix backbone (27)	0.46 (0.27)	1.09 (0.62)	0.43 (0.25)	1.90 (1.01)
α-helix side chain (30)	0.91 (1.25)	1.01 (0.75)	1.28 (1.25)	3.65 (2.04)

[a] See footnote *a* to Table I. [b] See footnote *b* to Table I. [c] See footnote *c* to Table I.

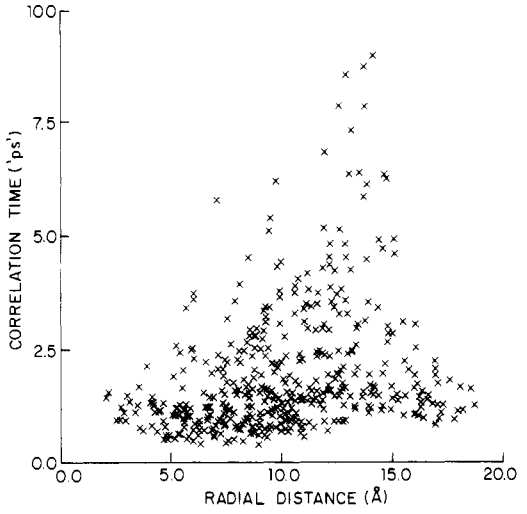

FIGURE 2: Scatter plot for relaxation time (τ) vs. the distance R from the centroid in the solvent run; each cross represents an atom.

the number of large τ values increases sharply with distance. These solvent results contrast with those for the vacuum run in which the interior atoms show longer relaxation times than the exterior residues. The cause for the differences in the solvent and vacuum results, which involves the collective character of the motions, will be discussed under Analysis of Results.

Considering the PTI secondary structural elements (i.e., the extensive β sheet corresponding to residues 16–24 and 26–37 and the α helix corresponding to residues 47–55), we find in Table II that the relaxation times for these backbone atoms do not have any special characteristics relative to the remainder of the molecule, although as pointed out previously (McCammon et al., 1977) the rms fluctuations are smaller for secondary structural elements than for the protein as a whole. As to the side chains, the β-sheet relaxation times are somewhat shorter and those for the α helix somewhat longer than the average; the latter correlates with the larger value obtained for the magnitude of the rms fluctuations (see Table II).

Relaxation Times and Fluctuation Amplitudes. Tables I and II list the magnitudes of the rms fluctuations in addition to the relaxation times. There is a strong correlation between the two variables; i.e., the larger the fluctuation, the longer the relaxation time. A more detailed picture of the τ vs. $\langle\Delta r^2\rangle$ correlation is given in a scatter plot for the solvent run in Figure 3; we include the all-atom results and the residue averages for all atoms, the backbone atoms, and the side-chain atoms. The relation found for the averages in Table II is clearly demonstrated in the figure for the individual atom and residue values. However, the scatter that is present indicates that there are additional effects superposed on the τ vs. $\langle\Delta r^2\rangle$ correlation. These will be analyzed further when we consider models for the atom fluctuations.

It is important to note that the τ vs. $\langle\Delta r^2\rangle$ correlation is less evident in the vacuum run, although it is present for the side chains. More generally for the two runs, it is found that the magnitudes of the rms fluctuations are highly correlated but the relaxation times are less so. Figure 4 illustrates this result by scatter diagrams for all of the backbone atoms; Figure 4a shows the correlation of the τ values and Figure 4b that of the $\langle\Delta r^2\rangle$ values obtained in the two runs. Only in the tightly coupled β-sheet region does the solvent have little effect on the dynamics; i.e., both τ and $\langle\Delta r^2\rangle$ are well correlated between the solvent and vacuum runs.

Form of the Correlation Function, Time Series, and Mean Square Fluctuation Time Development. Although the relaxation times and their variation with structural properties give an overview of the time dependence of the atom motions, it is of interest to examine the form of the correlation functions and the time series from which they were derived. In addition, the time development of the mean square fluctuations themselves can provide information concerning the nature of the motions. We briefly present the results and then analyze them under Analysis of Results.

(*A*) *Correlation Functions.* The correlation functions for $\Delta\vec{r}$ obtained in the vacuum and solvent runs for all the atoms of PTI have functional forms that fall into a small number of classes. We show examples of these classes in Figure 5. The first class (I), which corresponds to half of the heavy atoms in PTI, has a simple, approximately exponential decay to zero, often with small, high-frequency oscillations superposed on the overall functional dependence. Typical of class I correlation functions are those for Ala-16 C^β, Gly-37 N, and Arg-20 C^γ (see Figure 5a); the exponential fit to the correlation functions is given as a dashed line in the figure. This type of correlation function is found for both side-chain and backbone atoms, with most examples being in the interior of the protein. A second class (II) of correlation functions shows an initial rapid decay followed by large, relatively well-defined oscillations, as exemplified by Pro-13 C^β, Arg-39 N, Ile-19 C^γ, and Tyr-21 $C^\epsilon{}_2$ (Figure 5b) from the solvent run and Thr-11 N, Cys-14 C^α, and Lys-15 C^α (Figure 5c) from the vacuum run. There are about 150 atoms in this class; as is evident from Figure 5b,c, both the amplitude and time course of the correlation functions vary over a wide range. Finally in the third class (III), there are correlation functions for nearly 100 atoms which show an approximately linear decrease over the entire simulation; typical of these are Gly-28 C^α and Asp-50 C^β shown in Figure 5d. In certain cases, there are significant

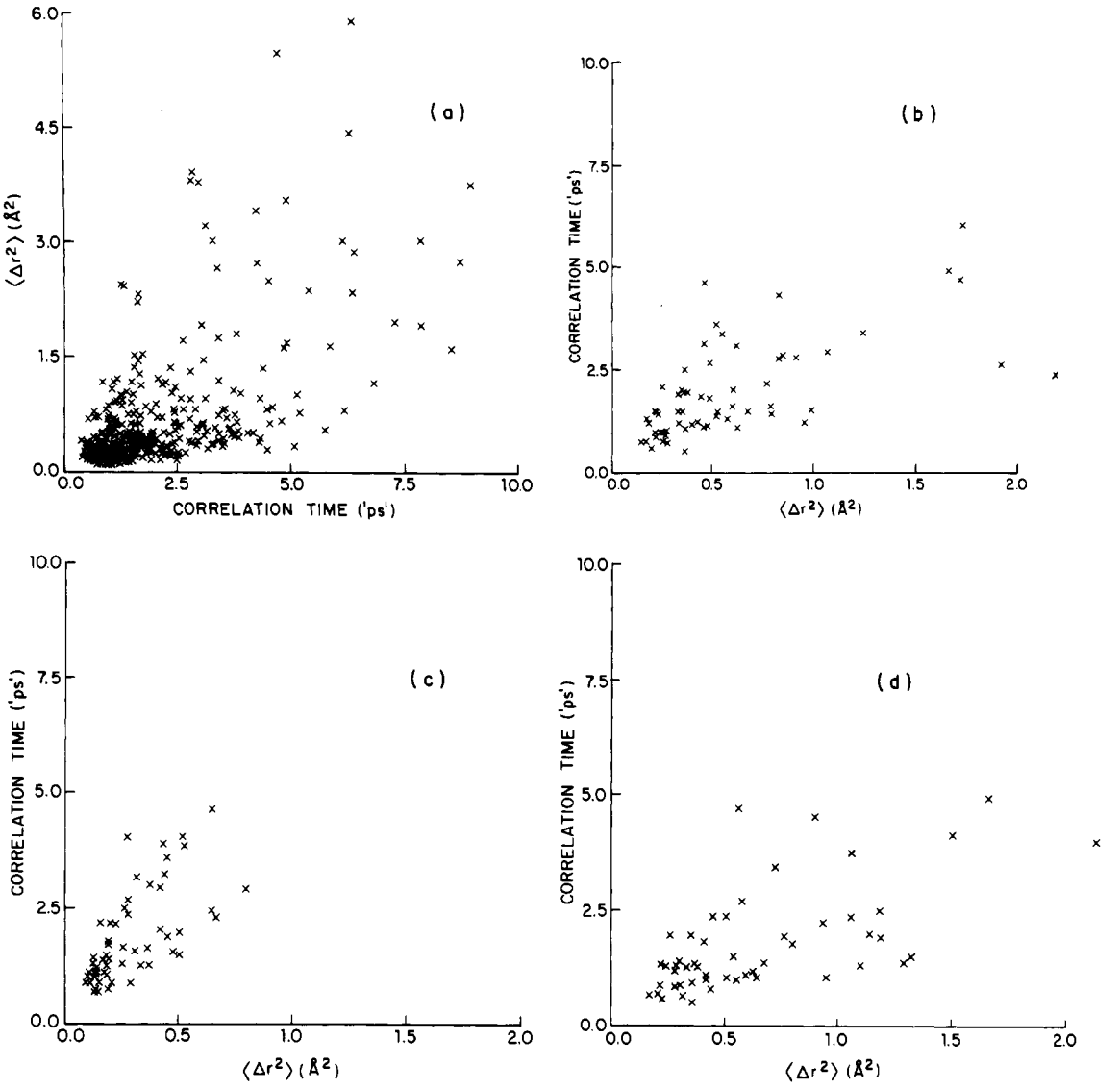

FIGURE 3: Scatter plots for relaxation time (τ) vs. the mean square fluctuation $\langle \Delta r^2 \rangle$ for the solvent run: (a) all atoms; (b) residue averages for all atoms; (c) residue averages for backbone atoms; (d) residue averages for side-chain atoms.

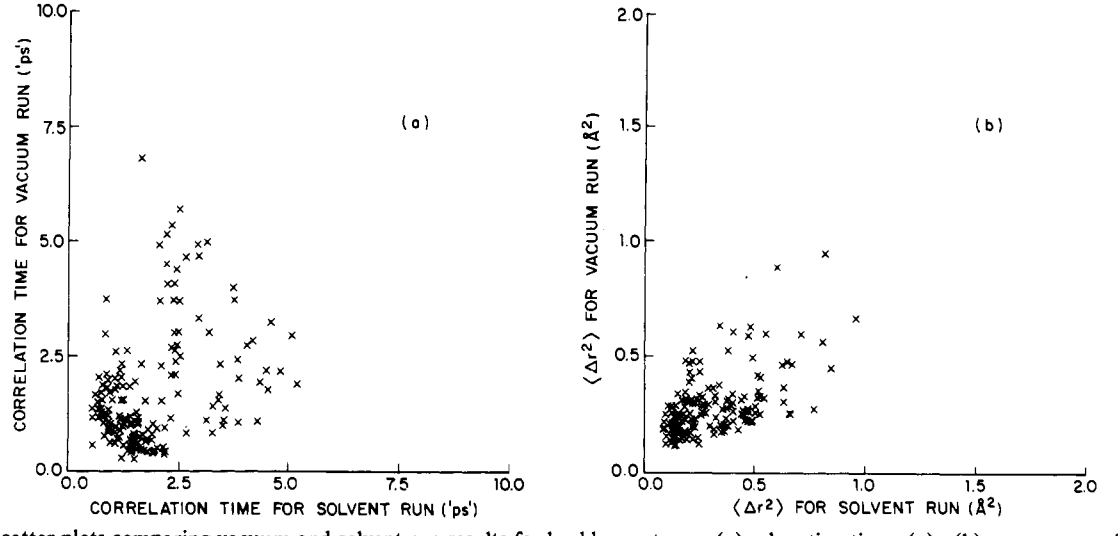

FIGURE 4: Scatter plots comparing vacuum and solvent run results for backbone atoms: (a) relaxation times (τ); (b) mean square fluctuations ($\langle \Delta r^2 \rangle$).

oscillations superposed on the dominant, essentially linear behavior.

(*B*) *Time Series.* To obtain additional insight into the nature of the motions, we illustrate the time series for certain atoms. For Tyr-21 C^{δ_1} (class I, solvent), which has an exponentially decaying correlation function, the time series for fluctuations in the directions of the principal axes of the thermal ellipsoid are shown in Figure 6a. As can be seen,

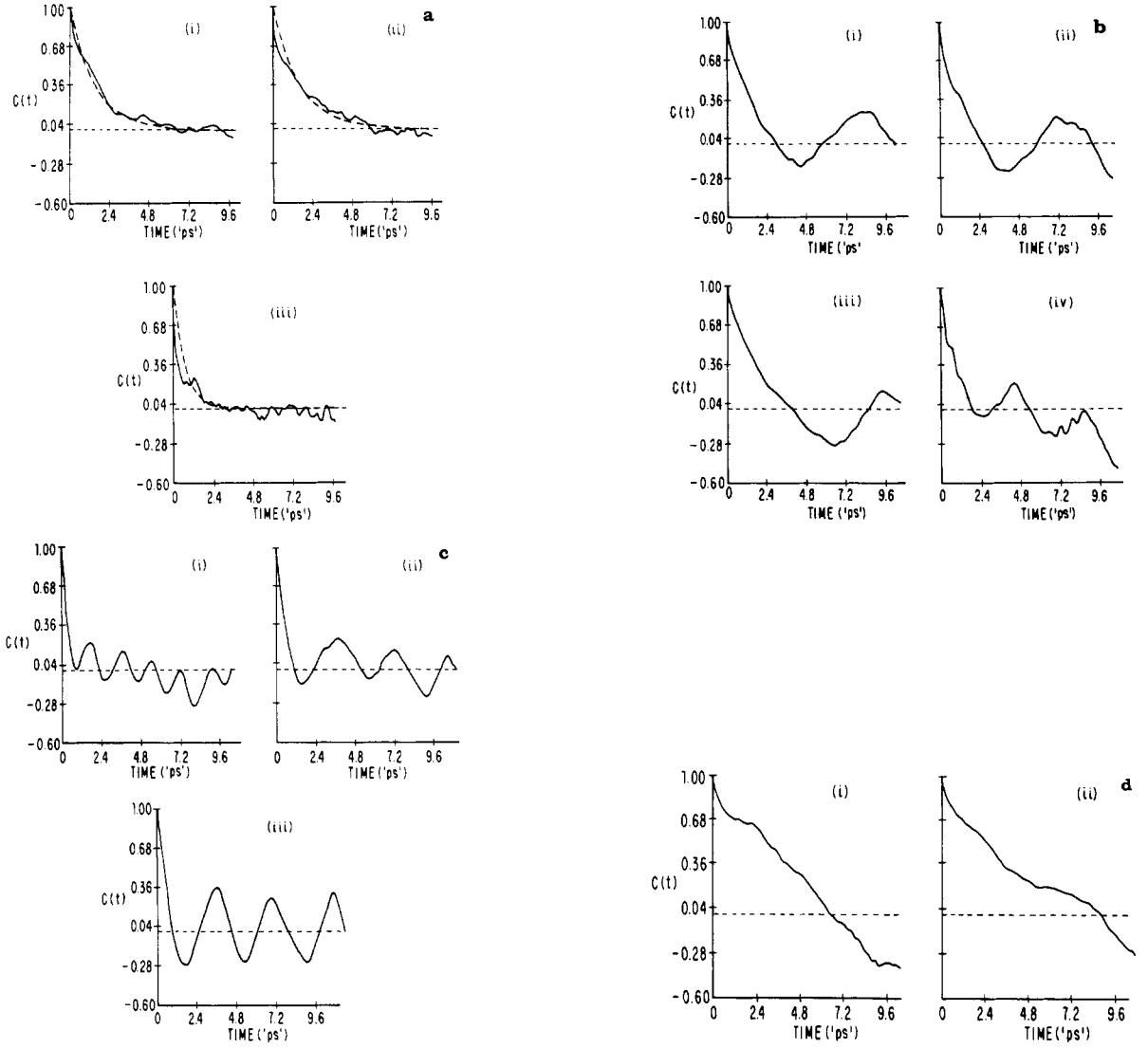

FIGURE 5: Displacement correlation functions. (a) Class I (overdamped decay; solvent run): (i) Ala-16 C^{β}; (ii) Gly-37 N; (iii) Arg-20 C^{γ}. (b) Class II (underdamped oscillations; solvent run): (i) Pro-13 C^{β}; (ii) Arg-39 N; (iii) Ile-19 C^{γ}; (iv) Tyr-21 C^{ϵ_2}. (c) Class II (underdamped; vacuum run): (i) Thr-11 N; (ii) Cys-14 C^{α}; (iii) Lys-15 C^{α}. (d) Class III (solvent run): (i) Gly-28 C^{α}; (ii) Asp-50 C^{β}.

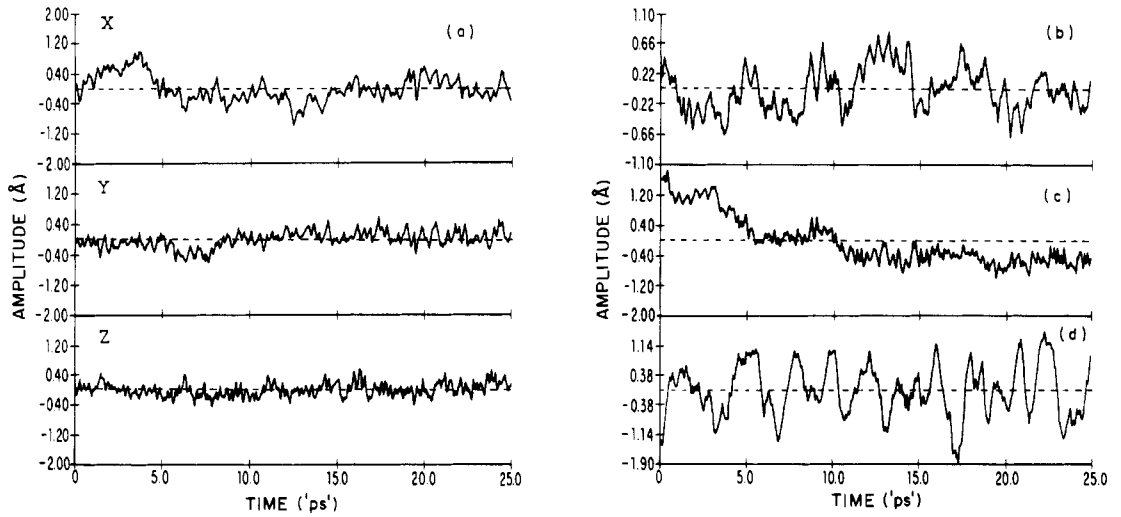

FIGURE 6: Time series for Cartesian components in the principal axis system of the thermal ellipsoid: (a) Tyr-21 C^{δ_1} (solvent run), three components; (b) Tyr-21 C^{δ_2} (solvent run), largest component only; (c) Asp-50 C^{β} (solvent run), largest component only; (d) Lys-15 C^{δ} (vacuum run), largest component only.

all three components show high-frequency, small-amplitude oscillations ($\leq$0.3 ps); such high-frequency oscillations are present for all atoms. In addition, there is some indication of lower frequency, large amplitude contributions to the motion, though generally only for the major axis component (X) are they significant. Figure 6b shows the time series for

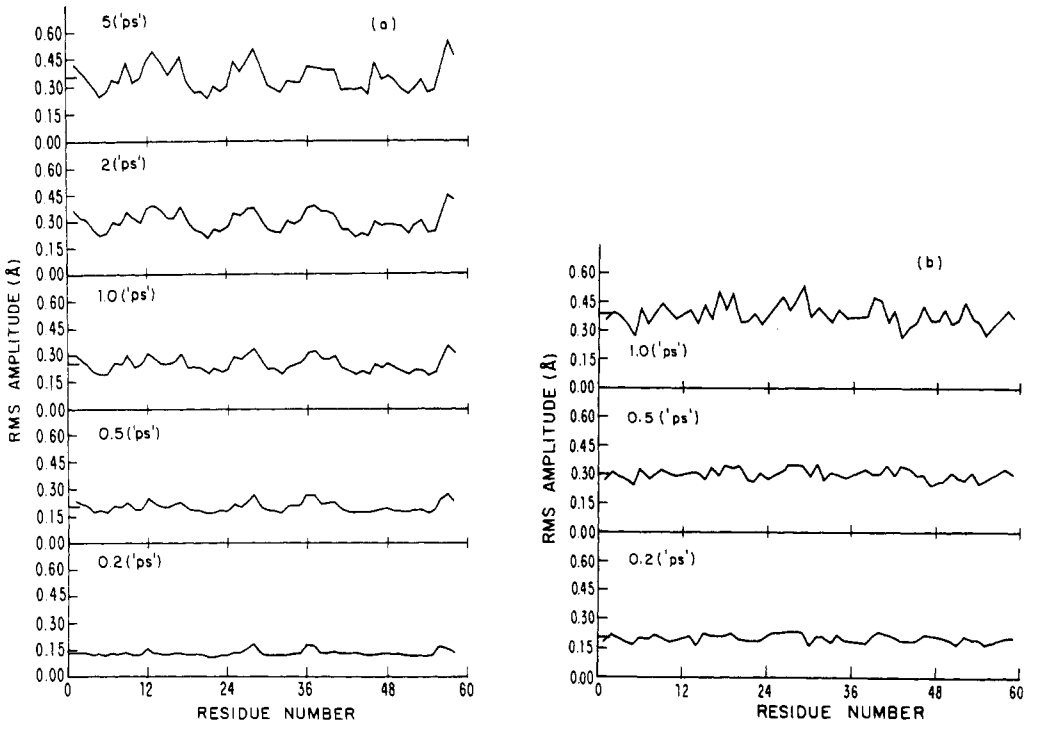

FIGURE 7: Root mean square displacement subaverages as a function of residue number: (a) C^{α} for solvent run (0.2, 0.5, 1, 2, and 5 "ps"); (b) side-chain averages for solvent run (0.2, 0.5, and 1 "ps").

the largest principal axis component of Tyr-21 C^{δ_2} (class II, solvent). It is clear that, in addition to the high-frequency oscillations, much lower frequencies are very important in producing the total flutuation amplitude; this result is consistent with the form of the correlation function shown for Tyr-21 C^{ϵ_2}, which is very similar to that of Tyr-21 C^{δ_2}, in Figure 5b (iv). Figure 6c gives the time series for the major component of Asp-50 C^{β} (class III, solvent), which has the correlation function shown in Figure 5d (ii). It illustrates the presence of high-frequency oscillations, superposed on the monotonic decrease in the mean value of the coordinate. Finally, we show the principal component of Lys-15 C^{δ} (class II, vacuum) which has large, very well-defined oscillations corresponding to those present in the correlation function; the very similar correlation function for Lys-15 C^{α} (vacuum) is given in Figure 5c (iii).

(C) Time Development of Mean Square Displacement. From the time series and correlation functions, it is evident that the atomic motions represent a superposition of high-frequency oscillations and lower frequency fluctuations. To separate the contribution of these two types of motion to the mean square displacements, we can determine subaveraged mean square displacement values; that is, the entire trajectory is divided into a series of intervals of a given length, the mean square displacement relative to the mean for each interval are calculated, and the results are averaged for the entire trajectory (van Gunsteren & Karplus, 1982a). Figure 7 shows the results obtained for the C^{α} atoms and for averages over the side-chain atoms of each residue from the solvent run; the vacuum results are very similar.

For the C^{α} atoms, there is a significant contribution to the rms value from the subpicosecond motion. At 0.2 ps, the C^{α} rms averaged over all residues is already 0.13 Å, about 40% of the value at 5.0 ps. Comparing the relative values of the C^{α} fluctuations, we see that at 0.2 ps, the results are rather uniform. Since the local high-frequency oscillations are making the main contribution to this subaverage, it appears that the effective potential does not vary significantly throughout the protein. This is in accord with expectations if the dominant factor for the high-frequency oscillations is the local librational potential associated with torsional motion of the backbone atoms. It should be noted, however, that even on the 0.2-ps time scale, there is some suggestion of inhomogeneity in that in the neighborhood of atoms 14, 28, and 36, as well as of the N- and C-terminal ends, slightly larger fluctuations are found; these regions all have greater than average fluctuations when the low-frequency modes are included (as can be seem from a comparison of the 0.2- and 5.0-ps values). From the time series results (Figure 6), the 0.5-ps averages include all of the high-frequency contributions, and the lower frequency motions are becoming more important. This is clearly indicated in Figure 7 by the greater variation in the fluctuations along the polypeptide chain. As longer time subaverages are examined, we see that the magnitudes of the fluctuation continue to increase in certain regions, in accord with the long relaxation times and low-frequency, collective character of the modes involved; examples are the loop region (residues 24–28) and the region at the top of the molecule (around residues 14 and 38). On the order of 10 ps is required for convergence. By contrast, for other parts of the protein (e.g., the β-sheet region, 16–24), the C^{α} fluctuations have already reached their asymptotic value by 2 ps.

For the side-chain averages (Figure 7b), the general behavior is similar to the backbone C^{α} results. However, even at 0.2 ps, the distribution of fluctuations as a function of residue number is more nonuniform than it is for the C^{α} atoms. This corresponds to the fact that the effective local potentials for the side-chain atoms have a greater variability than those for the main chain atoms.

Analysis of Results. The correlation functions, time series, and time development of the mean square displacements for the atoms of PTI suggest that the fluctuations generally involve the superposition of two types of motions. One is a high-frequency oscillation of relatively small magnitude, and the other has a considerably lower frequency and larger amplitude.

From the characteristics of the individual atom fluctuations, and from the relation between the displacements of different atoms, we can draw qualitative conclusions concerning the nature of the two types of motional contributions. The high-frequency oscillations are local, in the sense that they correspond to the librational motion of the individual atoms in an effective potential. This potential is a summation of dihedral angle terms in the potential function for the backbone or side chain of which the atom is a part and of nonbonded interactions with the atoms of the surrounding protein matrix. By contrast, the lower frequency components have a nonlocal, more collective origin, in that they involve the correlated motion of groups of atoms, ranging from a small number next to each other along the backbone and/or parts of a single side chain to a much larger number of atoms in certain regions of the protein.

(*A*) *Local Oscillations.* The high-frequency contribution to the fluctuations, which is present for all atoms, can be seen most easily in the time series themselves (Figure 6). It is irregular in character, due to perturbations of the librational motion by collisions with neighboring atoms. This results in the fast initial decay observed in most of the correlation functions (Figure 5). To analyze these local motions, we use a one-dimensional harmonic oscillator model, assuming that $\langle \Delta r^2 \rangle = 3\langle \Delta x^2 \rangle$. Applying the equipartition result (eq 5), we can estimate the effective force constant, k, from the results of Figure 7. For the C^α atoms at $T_{av} = 0.5$ ps, we have $\langle \Delta x^2 \rangle_{0.5} \simeq 0.013$ Å^2, which yields $k \simeq 3 \times 10^4$ erg cm^{-2}. With an average extended atom mass of 14 amu, the definition of ω_0 yields a frequency $\nu \simeq 6 \times 10^{12}$ s^{-1} and a period of 0.2 ps, in good accord with the values evident in the time series (Figure 6). This suggests that the Einstein-independent oscillator model is approximately valid for the high-frequency oscillations of the atoms. Further, the value of the mean frequency, which corresponds to $\tilde{\nu} = 200$ cm^{-1}, is of the order of those associated with localized, dihedral angle modes (Levy & Karplus, 1979).

(*B*) *Nonlocal Modes of Motion.* The general behavior of the correlation functions appears not to be associated with the high-frequency oscillations of individual atoms (except for the initial rapid decay) but rather with the slower motions that must involve correlated displacements of more than one atom. With the Langevin equation, we can interpret the relation between the relaxation time τ and the mean square magnitude of the fluctuations (Figure 3). Although the simple Langevin equation is apparently applicable only to the overdamped (class I) cases, we can study general trends with this equation, especially since about half of the atoms are class I. The analogue of eq 5 for $\langle \Delta r^2 \rangle$ is given by

$$\tau \simeq \frac{f}{3k_B T}\langle \Delta r^2 \rangle \qquad (10)$$

For a constant frictional coefficient, this would imply that the relaxation time is directly proportional to the mean square amplitude of the fluctuations. Since there is a large range of τ values for each $\langle \Delta r^2 \rangle$ as shown in Figure 3, it appears that the frictional coefficient is not constant throughout the protein, due to the variation in the effective particle size and in the local viscosity. Figure 8 shows a plot from the solvent run of the ratio $\tau/\langle \Delta r^2 \rangle$ for the C^α atoms as a function of the distance from the center of mass. This plot suggests that there is a decrease in the average friction coefficient with increasing distance from the center (i.e., from the interior to the exterior of the protein).

The friction coefficient obtained from the Langevin equation can also be used to estimate the effective particle size. Since

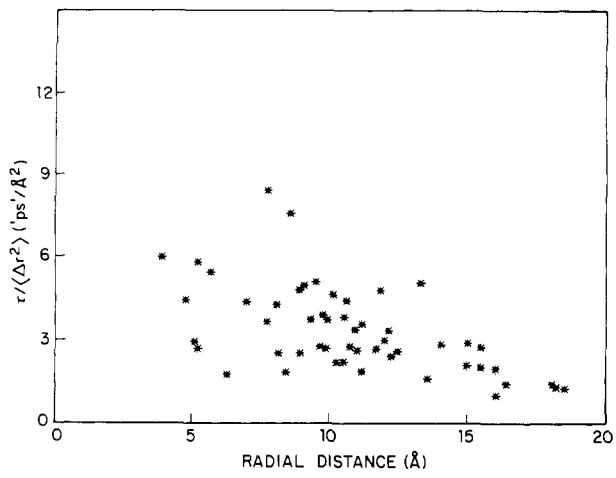

FIGURE 8: Ratio $\tau/\langle \Delta r^2 \rangle$ for α carbons as a function of the distance from the center.

the average value of $\tau/\langle \Delta r^2 \rangle$ for the solvent run is about 3 ps/Å^2 (see Figure 8), we find a frictional coefficient $f \simeq 4 \times 10^{-9}$ g s^{-1} from eq 10. This value can be related to the viscosity η and effective radius a of the moving body (assumed spherical) by Stokes law with slip boundary conditions, $f = 4\pi\eta a$. If we introduce the protein interior viscosity ($\eta \simeq 0.3$ cP) estimated in an earlier molecular dynamics simulation (McCammon et al., 1979), we find an effective radius $a \simeq 1 \times 10^{-7}$ cm, about 6 times the average van der Waals radius of the extended atoms used in the simulation. This order of magnitude result suggests that there is a collective character to the larger scale motions of the atoms in PTI, even for the overdamped correlation functions. The wide range of relaxation times obtained in the calculations indicates that the groups of atoms that move as a unit vary considerably in size; this point is analyzed further below.

The interpretation of the variation of τ with the distance from the center for the vacuum and solvent runs (Table II) is complicated by the fact that τ is dependent on the local viscosity and effective particle size, as well as $\langle \Delta r^2 \rangle$. The relaxation times in the interior of the vacuum structure are longer than those of the solvent structure, probably mostly due to viscosity effects since the interior of the vacuum structure is denser (van Gunsteren & Karplus, 1982a). In the solvent run, since the exterior atoms have a more collective character than the interior atoms, as indicated by larger values of $\langle \Delta r^2 \rangle$, the relaxation times are longer for exterior atoms. This accords with the finding that the exposure of atoms for the average dynamical structure, as measured by the Lee and Richards criterion (Richards, 1977), shows no correlation with relaxation time in either the solvent or the vacuum run, even though both exposure and relaxation time correlate with distance from the center. The longer correlation times for exterior atoms in the solvent run occur for all the atoms of the effective particle, even if some are not exposed. In vacuum, the viscosity experienced by the exterior atoms is less than that experienced in the presence of solvent, which reduces the relaxation times of exterior atoms relative to the solvent run. Furthermore, the size of the effective particles in the exterior apparently is smaller in vacuum than in solvent, possibly because the solvent acts to correlate the atomic motions. Some combination of reduced viscosity and effective particle size apparently leads to a decrease in τ with increasing distance from the center in vacuum, just the reverse of what is found in the solvent run.

The oscillatory correlation functions found for many of the atoms correspond to underdamped motion, which provide additional information on the nature of the protein fluctuations.

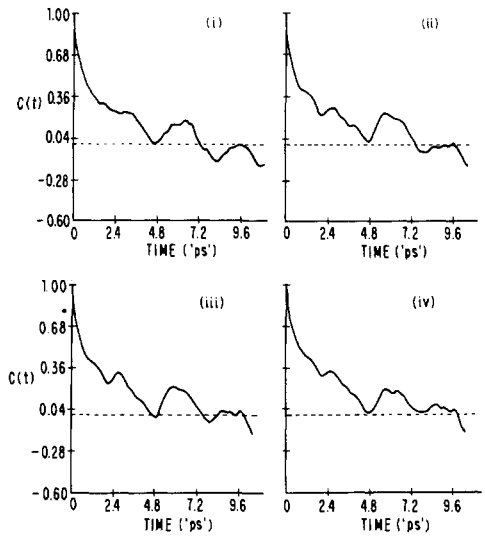

FIGURE 9: Displacement correlation functions showing collective motion (solvent run): (i) Tyr-10 C; (ii) Thr-11 N; (iii) Thr-11 C^{α}; (iv) Thr-11 C.

Because of their characteristic forms (class II; Figure 5b,c), it is possible to find groups of atoms that appear to be moving together simply by examining the correlation functions. This is true on a relatively local level, as well as on a considerably larger scale in certain parts of the protein.

There are many cases where two or three residues or parts of residues (e.g., backbone or side-chain atoms), either contiguous in sequence or in contact as a result of the tertiary structure of PTI, have essentially identical correlation functions. A clear case of such correlated motions for main chain atoms is illustrated in Figure 9 for the solvent run. The atoms involved are Tyr-10 C and O and Thr-11 N, C^{α}, C, and O as well as the side-chain atoms Thr-11 C^{β} and C^{γ}; Thr-11 O, although with a similar correlation function, is somewhat perturbed by its hydrogen-bonding interaction. For the neighbors of such a group, the correlation function of each atom is likely to have a somewhat different form, until the transition has been made to another highly correlated group.

For many side-chain atoms, the underdamped correlation functions also demonstrate concerted motions. Figure 5c shows the correlation function for Lys-15 C^{α} in the vacuum run; the entire Lys-15 side chain (except N^{ζ}) has the same correlation function, as do the Lys-15 backbone atoms and the backbone atoms of Cys-14. Of interest also are cases like Tyr-21, where one side of the ring (C^{δ_2} and C^{ϵ_2}; see Figure 5b), which is in the protein interior, has an oscillatory correlation function, while the other side (C^{δ_1} and C^{ϵ_1}), which is more exterior, has an overdamped form. Pairs of atoms that are hydrogen bonded tend to have similar oscillatory correlation functions that often differ from those in the rest of the side chain.

Some sets of correlation functions in class II indicate that much larger regions with many atoms are involved in the concerted motion. One of these groups showing collective effects is in the "top" part of the PTI molecule surrounding the 14–38 disulfide bond in the region that binds to trypsin (see Figure 10). Representative correlation functions are illustrated for Pro-13 C^{β} and Arg-39 N in Figure 5b. There are about 35 atoms in residues Gly-12, Pro-13, Cys-14, Lys-15, Cys-38, Arg-39, and Ala-40 that have the same form of correlation function, indicative of a very low-frequency oscillation in this part of the molecule. The vacuum run also suggests there is a collective mode in the same region, but the frequency is about twice that found in the solvent run due to the fact that a smaller number of atoms are coupled together.

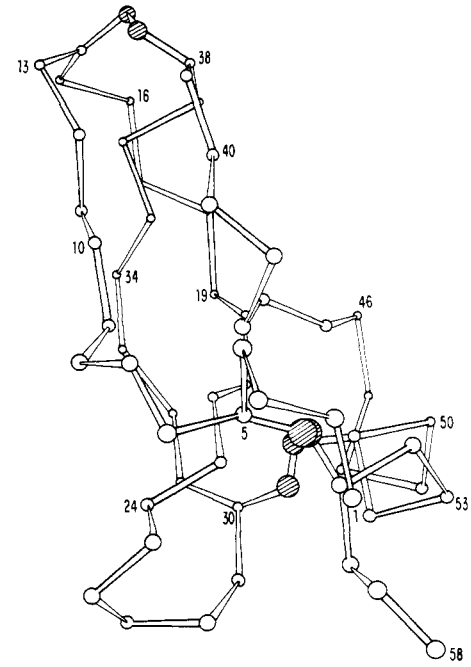

FIGURE 10: X-ray structure of PTI; only α carbons and disulfide bridges are shown; coordinates are from Deisenhofer & Steigemann (1975).

Finally, we mention the monotonically decreasing correlation function (class III, Figure 5d), shown by atoms in the "bottom" of the molecule (Figure 10) for both the vacuum and the solvent runs. Most atoms in the N-terminal region (1–4), the loop region (24–29), and the C-terminal region (53–58) are involved. Examination of the components of the radius of gyration perpendicular to the long axis of PTI suggests that the molecule is contracting in the bottom region during the simulation. Whether this contraction is an equilibration phenomenon or whether it involves a low-frequency oscillation with a period on the order of the length of the run is not established.

Although the correlation functions for the atoms participating in the collective modes that we have been examining are oscillatory in character, they do not satisfy the simple Langevin equation for underdamped motion (eq 4). As shown in Figures 5 and 9, there is an initial very rapid decay of the correlation function followed by a low-frequency oscillation, which can be very regular, particularly for the vacuum run (Figure 5c), or more irregular, particularly for the solvent run (Figures 5b and 9). However, in none of these cases do the low-frequency oscillations in the correlation function show significant damping, as is expected from the Langevin equation. To analyze these correlation functions, we consider a superposition of modes (eq 7–9) and make use of the simplified model consisting of two oscillators, one representing the local high-frequency oscillation of the atom and the other corresponding to the lower frequency collective motion. We assume that the collective motion can be represented by that of a single massive "particle" to which is attached the lighter particle (atom) that undergoes the high-frequency oscillations. For the latter, we know the characteristics (mass, force constant, and frequency) from the Einstein oscillator analysis given above. Using the Born–Oppenheimer approximation (see Materials and Methods), we treat the heavy particle separately and consider the two examples cited above: Lys-15 C^{α} in the vacuum run (Figure 5c) and Cys-14 C^{α} in the solvent run. From the time series for the dominant component, we find the maximum displacements (vibrational amplitudes) are ~ 1.2

Å for Lys-15 C^{α} and ~0.9 Å for Cys-14 C^{α}. Introducing these values into the one-dimensional harmonic oscillator equipartition formula, we find effective force constants $k \simeq 5.8 \times 10^2$ erg cm^{-2} for Lys-15 and $k \simeq 1.0 \times 10^3$ erg cm^{-2} for Cys-14. From a comparison of individual correlation functions, we estimate that there are about 12 atoms involved in the Lys-15 low-frequency motion and about 35 atoms in that of Cys-14. Introducing these values into the harmonic oscillator expression with an average mass per atom of 14 amu, we obtain frequencies of $\nu \cong 2.3 \times 10^{11}$ s^{-1} for the Lys-15 mode and $\nu \cong 1.8 \times 10^{11}$ s^{-1} for the Cys-14 mode. These results are to be compared with the frequencies obtained from the correlation functions or directly from the time series; they are $\nu \cong 3.1 \times 10^{11}$ and 1.5×10^{11} s^{-1}, respectively. Thus, the frequencies obtained by the two approaches are in satisfactory agreement, considering the approximations involved.

The frequencies for the collective mode contributions to the correlation functions vary from $\sim 1 \times 10^{11}$ to $\sim 1 \times 10^{12}$ s^{-1}; in wavenumbers, this corresponds to values between 3 and 30 cm^{-1}. The class III correlation functions have not been included in the analysis; if they corresponded to an oscillatory mode, the frequency would be about 1.5 cm^{-1}.

Although the model treating portions of the protein as a single massive entity provides a simple description of the low-frequency contribution to the motion, the true situation is certainly more complicated. Analysis of the correlation coefficients for the displacements of different atoms suggests that there are some low-frequency relative motions within the portions of the protein that we have approximated by a rigid body, particularly for the cases where many atoms are involved.

Conclusions

The analysis of the time dependence of the atomic fluctuations calculated by molecular dynamics simulations of the bovine pancreatic trypsin inhibitor in vacuum and in a van der Waals solvent leads to a number of important results. The relaxation times for the displacement correlation functions obtained from the solvent run range from 0.4 to 10 ps; those of the vacuum run tend to be somewhat shorter, with a range of 0.2–5 ps. Comparing the mean square displacements and the relaxation times, we find that the former are less sensitive to the solvent than the latter. Since an aqueous solvent has a higher viscosity than the van der Waals model used here, a greater increase in the relaxation times might be expected in water.

Analysis of the correlation functions for the atomic fluctuations suggests a simple model involving the superposition of two types of motions. One component corresponds to the high-frequency oscillations ($\tilde{\nu} \simeq 200$ cm^{-1}) of the individual atoms in the potential involving its dihedral angles and the surrounding protein atoms. The other component has a more collective character that can be approximated by the motion of a rigid effective particle to which the atom is attached. These effective particles may consist of only a few atoms (e.g., as in an individual side chain or a small part of the backbone) or of a larger fraction of the protein (e.g., about 35 atoms making up the portion of PTI in the neighborhood of the 14–38 disulfide bond). The frequencies of these collective motions are found to be in the 3–30-cm^{-1} range.

Analysis of the time development of the rms fluctuations has shown that the high-frequency, local oscillations with periods up to 0.2 ps contribute about 40% of the average magnitude and are essentially uniform throughout the protein. It is the lower frequency collective modes with periods between 1 and 10 ps that introduce the large amplitude fluctuations which differentiate one part of the protein from another. Because the latter are associated with groups of atoms, their anisotropic character may have no relation to the covalent bonding of individual atoms. This has already been noted in the orientation of thermal ellipsoids obtained in simulation studies (Northrup et al., 1981) and suggests that care is needed in the development of constrained models for the refinement of anisotropic temperature factors (Konnert & Hendrickson, 1980).

The results of the present analysis, which suggest an interpretation of protein atom fluctuations as composed of two types of components, are of general interest. Although the local high-frequency oscillations are rather uniform throughout the protein, the collective, lower frequency modes have a more inhomogeneous character. The latter may well be of particular importance in the biological function of the protein; they may be involved in both the binding (opening and closing) and catalytic steps of enzyme reactions. Further, the extended nature of these motions makes them more sensitive to the environment (e.g., the observed differences in the present simulations between the vacuum and solvent runs). It is possible that they are involved in transmitting the effect of the surrounding solvent into the interior of the protein, as has been observed in the oxygen binding by myoglobin (Beece et al., 1980). Also, because they involve sizable portions of the surface of the protein, they are expected to be quenched at low temperature by freezing of the solvent. Their contribution to the mean square fluctuations may explain the transition in the temperature dependence of the fluctuations in myoglobin near 200 K (Parak et al., 1981).

This superposition of the motional contributions to the atomic fluctuations in PTI is analogous to that observed in the study of the dynamics of an isolated α helix as a function of temperature (Levy et al., 1982). By comparing the results of a complete molecular dynamics simulation with the harmonic approximation to it, the mean square fluctuations were separated into a high-frequency, harmonic part (<1 ps) and a much lower frequency anharmonic contribution. The latter is unimportant below 50 K but approximately doubles the magnitude of the mean square atomic fluctuations at room temperature.

A caution about the present analysis needs to be started. Since the lowest frequency collective motions ($\tau \sim 5$–10 ps) have a time scale that is of the same order as the dynamics simulation itself, they may not have been adequately sampled during the run; longer simulations will be necessary to obtain more complete information. However, the essential conclusion that collective motions play an important role in determining the atomic fluctuations in proteins appears unequivocal.

References

Adelman, S. A. (1980) *Adv. Chem. Phys. 44*, 143–253.

Artymiuk, P. J., Blake, C. C. F., Grace, D. E. P., Oatley, S. J., Phillips, D. C., & Sternberg, M. J. E. (1979) *Nature (London) 280*, 563–568.

Beece, D., Eisenstein, L., Frauenfelder, H., Good, D., Marden, M. C., Reinisch, L., Reynolds, A. H., Sorensen, L. B., & Yue, K. T. (1980) *Biochemistry 19*, 5147–5157.

Berkowitz, M., & McCammon, J. A. (1981) *J. Chem. Phys. 75*, 957–961.

Campbell, I. D., Dobson, C. M., Moore, G. R., Perkins, S. J., & Williams, R. J. P. (1976) *FEBS Lett. 70*, 96–100.

Chandrasekhar, S. (1943) *Rev. Mod. Phys. 15*, 1–89.

Deisenhofer, J. O., & Steigemann, W. R. (1975) *Acta Crystallogr., Sect. B B31*, 238–250.

Dunitz, J. D. (1979) *X-ray Analysis and the Structure of Organic Molecules*, Cornell University Press, Ithaca, NY.

Frauenfelder, H., Petsko, G. A., & Tsernoglou, D. (1979) *Nature (London) 280*, 558–563.
Gelin, B. R., & Karplus, M. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 2002–2006.
Gurd, F. R. N., & Rothgeb, T. M. (1979) *Adv. Protein Chem. 33*, 73–165.
Karplus, M., & McCammon, J. A. (1979) *Nature (London) 277*, 578.
Karplus, M., & McCammon, J. A. (1981) *CRC Crit. Rev. Biochem. 9*, 293–349.
Konnert, J. H., & Hendrickson, W. A. (1980) *Acta Crystallogr., Sect. A A36*, 344–350.
Kushick, J., & Berne, B. J. (1977) *Mod. Theor. Chem. 6*, 41–63.
Levy, R. M., & Karplus, M. (1979) *Biopolymers 18*, 2465–2495.
Levy, R. M., Karplus, M., & McCammon, J. A. (1981a) *J. Am. Chem. Soc. 103*, 994–996.
Levy, R. M., Karplus, M., & Wolynes, P. G. (1981b) *J. Am. Chem. Soc. 103*, 5998–6011.
Levy, R. M., Perahia, D., & Karplus, M. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 1346–1350.
Mao, B., Pear, M. R., McCammon, J. A., & Northrup, S. H. (1982) *Biopolymers* (in press).
McCammon, J. A., & Karplus, M. (1980) *Annu. Rev. Phys. Chem. 31*, 29–45.
McCammon, J. A., Gelin, B. R., & Karplus, M. (1977) *Nature (London) 267*, 585–590.
McCammon, J. A., Wolynes, P. G., & Karplus, M. (1979) *Biochemistry 18*, 927–942.
Northrup, S. H., Pear, M. R., McCammon, J. A., Karplus, M., & Takano, T. (1980) *Nature (London) 287*, 659–660.
Northrup, S. H., Pear, M. R., Morgan, J. D., McCammon, J. A., & Karplus, M. (1981) *J. Mol. Biol. 153*, 1087–1109.
Parak, F., Frolov, E. N., Mössbauer, R. L., & Goldanskii, V. I. (1981) *J. Mol. Biol. 145*, 825–833.
Pauling, L., & Wilson, E. B., Jr. (1935) *Introduction to Quantum Mechanics*, McGraw-Hill, New York.
Peticolas, W. L. (1979) *Methods Enzymol. 61*, 425–458.
Richards, F. M. (1977) *Annu. Rev. Biophys. Bioeng. 6*, 151.
Snyder, G. H., Rowan, R., Karplus, S., & Sykes, B. D. (1975) *Biochemistry 14*, 3765–3777.
Stillinger, F. H., & Rahman, A. (1972) *J. Chem. Phys. 57*, 1281.
Tully, J. C. (1981) *Acc. Chem. Res. 14*, 188–194.
van Gunsteren, W. F., & Karplus, M. (1981) *Nature (London) 293*, 677–678.
van Gunsteren, W. F., & Karplus, M. (1982a) *Biochemistry 21*, 2259–2274.
van Gunsteren, W. F., & Karplus, M. (1982b) *Macromolecules* (in press).
Wagner, G., DeMarco, A., & Wüthrich, K. (1976) *Biophys. Struct. Mech. 2*, 139–158.
Wang, M. C., & Uhlenbeck, G. E. (1945) *Rev. Mod. Phys. 17*, 323–342.
Willis, B. T. M., & Pryor, A. W. (1975) *Thermal Vibrations in Crystallography*, Cambridge University Press, London.
Wilson, E. B., Jr., Decius, J. C., & Cross, P. C. (1955) *Molecular Vibrations*, McGraw-Hill, New York.
Zwanzig, R. (1965) *Annu. Rev. Phys. Chem. 16*, 67–102.

Analysis of Electrostatic Interactions and Their Relationship to Conformation and Stability of Bovine Pancreatic Trypsin Inhibitor†

Keith L. March, David G. Maskalick, Richard D. England, Stephen H. Friend,‡ and Frank R. N. Gurd*

ABSTRACT: The modified Tanford–Kirkwood electrostatic theory has been employed to evaluate pK values for all charge sites in the bovine pancreatic trypsin inhibitor (BPTI). ^{13}C NMR titration data were obtained for all titrating groups except arginine residues in BPTI at nearly constant ionic strength in 0.1 M NaCl, at 41 °C. The chemical shifts of 46 resonances were found to be sensitive to pH. The pK values of these titrating resonances compared well with those computed by the modified Tanford–Kirkwood electrostatic theory. A conformational change involving the NH_2- and COOH-terminal and nearby residues is shown to be partly electrostatically driven by the formation of a salt bridge between the α-amino and α-carboxyl groups at mid-pH values. The computed total electrostatic free energy of the molecule is found to be stabilizing at neutral pH despite the substantial net positive charge borne by the protein under such conditions.

Electrostatic interactions in biological macromolecules are widely recognized as crucial to structure–function relationships at every level. The Tanford–Kirkwood electrostatic theory as modified by Shire predicts in a detailed manner the pK values of individual charge sites within a protein (Shire et al., 1974a,b, 1975; Botelho et al., 1978; Matthew et al., 1979a,b) as well as the overall contribution of each site to the total electrostatic free energy of the molecule (Friend & Gurd, 1979a,b; Friend et al., 1980, 1981; Flanagan et al., 1981). Because of the exceptional stability of the bovine pancreatic trypsin inhibitor (BPTI) to extremes of pH and temperature (Vincent et al., 1971; Masson & Wüthrich, 1973; Harina et al., 1980; Wagner & Wüthrich, 1978) as well as its small size (Huber et al., 1970; Deisenhofer & Steigemann, 1975), it presents an opportunity to test the prediction of pK values for acidic and basic residues. The titrations of various amino acid residues in BPTI have been reported under many different conditions of ionic strength and temperature (Maurer et al., 1974; Snyder et al., 1975; Wagner & Wüthrich, 1975; Brown et al., 1976, 1978; Richarz

† From the Department of Chemistry and the Medical Sciences Program, Indiana University, Bloomington, Indiana 47405. *Received February 16, 1982.* This is the 129th paper in a series dealing with coordination complexes and catalytic properties of proteins and related substances. For the preceding paper, see Brown & Gurd (1981). This work was supported by U.S. Public Health Service Research Grants HL-05556 and HL-14680. K.L.M. and R.D.E. gratefully acknowledge support from Northwestern National Life Insurance Co. and American United Life, respectively, through the Insurance Medical Scientist Scholarship Fund.

‡ Present address: Department of Pediatrics, Children's Hospital of Philadelphia, Philadelphia, PA.

0006-2960/82/0421-5241$01.25/0