PROTEINS: Structure, Function, and Genetics 15:360–373 (1993)

# Theoretical Probes of Conformational Fluctuations in S-Peptide and RNase A/3′–UMP Enzyme Product Complex

**John E. Straub [1] and D. Thirumalai[2]**

*[1]Department of Chemistry, Boston University, Boston, Massachusetts 02215, and [2]Department of Chemistry and Biochemistry, and Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742*

***ABSTRACT*** **The dynamic properties of the RNase A/3′–UMP enzyme/product complex and the S-peptide of RNase A have been investigated by molecular dynamics simulations using suitable generalization of ideas introduced to probe the energy landscape in structural glasses. We introduce two measures, namely, the kinetic energy fluctuation metric and the force metric, both of which are used to calculate the time needed for sampling the conformation space of the molecules. The calculation of the fluctuation metric requires a single trajectory whereas the force metric is computed using two *independent* trajectories. The vacuum MD simulations show that for both systems the time required for kinetic energy equipartitioning is surprisingly long even at high temperatures. We show that the force metric is a powerful means of probing the nature and relative importance of conformational substates which determine the dynamics at low temperatures. In particular the time dependence of the nonbonded force metric is used to demonstrate that at low temperatures the system is predominantly localized in a single cluster of conformational substates. The force metric is used to show that relaxation of long range (in sequence space) interactions must be mediated by a sequence of local dihedral angle transitions. We also argue that the time needed for compact structure formation is intimately related to the time needed for the relaxation of the dihedral angle degrees of freedom. The time for nonbonded interactions, which drive protein molecules to fold under appropriate conditions, to relax becomes extremely long as the temperature is lowered suggesting that the formation of maximally compact structure in proteins must be a very slow process.** © 1993 Wiley-Liss, Inc.

**Key words: conformational substates, protein conformations, molecular dynamics, nonbanded relaxations, ergodic measures, s-peptide, RNase A**



## INTRODUCTION

The dynamics of internal motions in proteins and nucleic acids have been extensively investigated theoretically[1,2] and experimentally[3] by a variety of techniques. Much of the theoretical insight into these complex systems has been obtained using several different simulation methods. Results of molecular dynamics (MD) simulations have shown that the fluctuations in internal motions of proteins can span several decades in time. Furthermore the possible relationship between these fluctuations and the ultimate biological function of proteins is beginning to be unraveled.[4] For example, the dynamics at different temperatures of the rebinding of ligands to myoglobin and hemoglobin following flash photolysis has revealed a rich array of motions on timescales from subpicoseconds to microseconds. Perhaps the most significant aspect that has emerged from these studies is that the energy (or free energy) landscape is very rough possessing many minima in which the protein has somewhat different structure although the precise differences between these structures is unknown. It has been argued that the spectra of these structures [referred to as conformational substates (CS)] are arranged in tiers,[3] and that individual substates are separated by barriers ranging from hundredths to many kcal/mol. These general characteristics of the free energy landscape are believed to result in nonexponential kinetics, and non-Arrhenius temperature dependence of the rebinding rate constant.[5,6] The rough free energy landscape that exhibits these CS has also been probed by MD simulations.[7,8] Although all of the conclusions regarding the general energy landscape for proteins have been reached by a series of studies on the heme proteins myoglobin and hemoglobin it is likely that the generic features will be observed in all proteins.[9]

Received June 8, 1992; revision accepted July 30, 1992.

Address reprint requests to Dr. D. Thirumalai, Department of Chemistry and Biochemistry, Institute for Physical Science and Technology, University of Maryland at College Park, College Park, MD 20742-2431.

In this paper we present general methods that can be used to obtain some insight into the nature of minima explored by proteins by using trajectories from MD simulations. All of the simulations in this paper have been done in vacuum. However the methods introduced here can be conveniently used even if the effects of the solvent molecules are explicitly taken into account. The basic methodology used in this paper to analyze the dynamics of relaxation of the different degrees of freedom in proteins was first introduced to understand the nature of potential minima in structural glasses.[10,11] The adaptation of ideas developed in the context of glasses to study protein dynamics is natural because it is suspected that for these seemingly disparate systems the low temperature dynamics is essentially controlled by a rough energy surface, i.e., the presence of many minima separated by barriers of varying heights. For structural glasses this scenario was postulated a long time ago.[12] The stretched exponential behavior seen in the rebinding kinetics of ligands to myoglobin and the associated non-Arrhenius temperature dependence of the rate constant are often used as evidence of a rough energy landscape in proteins.[13] In addition both neutron scattering experiments[14] and computer simulations[15] of proteins show that below a certain (glass transition) temperature the dynamic structure factor develops a nondecaying component at long times. The features described above are the hallmark of glasses[16] and it is in this sense that one believes when dynamic properties of proteins are examined over a wide temperature range these systems appear to be glass-like.

There are two important issues in MD simulations which must be addressed before the results can be usefully interpreted. The first is the need to know the potential energy function that adequately describes interactions in protein molecules. The second issue, which is the one dealt with here, is given a potential energy function and the temperature what is the minimum time required for the system to reach equilibrium? By equilibrium we mean that the fluctuations in the various degrees of freedom have been sampled sufficiently so that the computed averages of physical observables can be expected to agree with experimental measurements (provided the potential energy function is adequate). In essence our methods provide a theoretical means for obtaining the minimum time, $t_{min}$, required for equipartitioning to be reached in systems with a large number of degrees of freedom as is the case in proteins. This is an extremely important problem because MD simulations are commonly used to interpret experiments, and based on such comparisons the adequacy of the potential energy functions is often assessed.[15] If the value of $t_{min}$ is not large enough such comparisons are meaningless, and therefore it is of interest to understand in a quantitative way the time scale needed for sampling the available conformational space. In part the present paper is devoted to addressing this issue.

The major objectives of our study are the following: (1) To demonstrate using the generalized ergodic measures (GEM) that one can assess the approximate time scale for sampling the configuration space of a protein molecule. Currently MD simulations can only probe motions in proteins that occur on the time scale of nanoseconds and consequently only those degrees of freedom that relax on this time scale can be monitored. (2) We show using the GEM that there is a direct correlation between the dynamics of nonbonded interactions and dihedral angle transitions. We argue that the ideas presented here are quite general and may be used to analyze MD (as well as Monte Carlo or hybrid dynamics) trajectories of other heterogeneous systems. (3) We demonstrate that the GEM provides a natural way of probing the nature of the energy landscape in proteins. There have been previous attempts to probe the topography of the potential energy surface in Mb.[7] However, these approaches have used static measures whereas the GEM are dynamic. Thus our methods can offer insight into the recent claim based on experiments[4] that the energy landscape appears to be time and temperature dependent. By combining various aspects of the GEM we show how the dynamics in different CS can be characterized.

## THEORY AND METHODOLOGY

In the molecular dynamics method, in which trajectories are generated by solving the classical equations of motion, it is commonly assumed that the time averages of physical observables are equivalent to the corresponding averages over all spatial conformations of the system. This is the celebrated ergodic hypothesis[17] and it is almost always assumed that classical many body systems (such as the protein models considered here) are ergodic. Strictly speaking the equivalence between time and phase space averages stated in the ergodic hypothesis is expected to be satisfied only when the averaging time, $\tau_{ave}$, is essentially infinite [or more precisely scales as $\tau_{ave} \sim \exp(N)$ where $N$ is the number of atoms in the system]. However, this requirement is far too strict to be of practical utility. In practice one simply requires that a "representative" number of conformations be sampled and the time required for this is always far less than that needed for the ergodic hypothesis to be satisfied. The "representative number" of conformations will depend on the problem. For example, it is suspected that in the process of folding of proteins a random sampling of all conformations does not take place. Rather only a restricted number of conformations, determined by energetic considerations as well as excluded volume effects, are sampled. In general in biological applications certain severe constraints, which are often

difficult to determine, drastically reduce the number of relevant conformations.

In MD simulations, in which $\tau_{ave}$ is finite, for meaningful computations we require that $\tau_{ave}$ be long enough so that typical conformations of the protein molecules corresponding to a fixed total energy (microcanonical ensemble) are adequately sampled. In practice this means that the trajectory should uniformly sample all the accessible regions of the conformation space of the protein. If $\tau_{ave}$ is long enough for this to be satisfied then we assume that effective ergodicity is achieved.

The second point, which is especially pertinent to the analysis of fluctuations in proteins, concerns the range of motions that can be adequately sampled on the simulation time scale $\tau_{ave}$. Currently the time scale obtainable in MD simulations using present day computers is at best on the order of nanoseconds. Thus if there are motions that have relaxation times significantly greater than a few nanoseconds the physical quantities dependent on such motions will appear frozen when observed on the time scale $\tau_{ave}$. Consequently these quantities are essentially nonergodic on the time scales accessible to computers, and thus at this time simulations are not capable of probing the dynamics of such motions. This is why the study of protein folding (which apparently involves large scale slow collective motions) on realistic systems using computer experiments has so far proved intractable. However, there are numerous situations involving biological molecules where a variety of dynamic processes occur on the time scales accessible to computer simulations.[2] Although the methods we present here are applicable without regard to limitations set by current computing methods the present applications to the S-peptide and RNase A/3′–UMP enzyme/product complex (referred to as enzyme/product from now on) are restricted to degrees of freedom that can be probed on the order of hundreds of picoseconds.

With the above important comments in mind we want to address the following questions: Given a potential energy function for a particular protein and the temperature $T$ how can one estimate the minimum value of the averaging time for which effective ergodicity is achieved? In the time $t_{ave}$, what is the interplay between the dynamics of relaxation of the dihedral degrees of freedom and the nonbonded interactions? Finally is it possible to recognize when the system is not effectively ergodic and probe into the nature of the conformational substates which characterize the low temperature protein dynamics? To address these and related issues we propose a generalization of the ergodic measures which were introduced to study related problems in structural glasses.[10,11] Generalization of the ergodic measures (referred to as GEM) is necessary because proteins are heterogeneous systems. We will argue that to a large degree the questions raised above can be answered using the dynamic scaling law obeyed by the GEM.

## Fluctuation Metric

The fluctuation metric, which is an example of the GEM, is constructed from quantities that are readily calculable in a typical MD simulation. The fluctuation metrics are calculated from time averaged values of the relevant space variables. Suppose we have a variable $F_j(t)$ for the $j$th atom of a system of $N$ atoms. (Later in our discussion we choose $F_j(t)$ to be the kinetic energy for which $F_j(t) = m_j v^2_j/2$ where $v_j(t)$ is the velocity of the $j$th particle at time $t$ and $m_j$ is the mass.) We write the time average of $F_j(t)$ as

$$f_j(t) = \frac{1}{t}\int_0^t ds F_j(s). \tag{1}$$

Further writing the average of $f_j(t)$ over all $N$ atoms of the system as

$$\bar{f}(t) = \frac{1}{N}\sum_{j=1}^{N} f_j(t) \tag{2}$$

we define the mean-square difference of the individual $f_j(t)$s from the average $f(t)$ as the function[18]

$$\Omega_f(t) = \frac{1}{N}\sum_{j=1}^{N} [f_j(t) - \bar{f}(t)]^2. \tag{3}$$

The function $\Omega_f(t)$ is referred to as the fluctuation metric. The reason $\Omega_f(t)$ is called the fluctuation metric is the following: if $N$ is sufficiently large, i.e., $N >> 1$ then the results obtained by measurements on a single sample will differ negligibly (at most of the order $1/\sqrt{N}$) from those computed using averages over an ensemble of systems. This is the law of large numbers. Physical properties that satisfy this criterion are self-averaging. Thus one can replace Eq. (2) by $<f>$ where $< >$ denotes an ensemble average, and if the system behaves ergodically on the time scale $\tau_{ave}$ then $f_j(\tau_{ave}) \rightarrow <f>$. From the above discussion we understand that for systems which are effectively ergodic $\Omega_f(t)$ measures the mean-square fluctuations of $f$ from its ensemble average.

It follows from the law of large numbers that $\Omega_f(0)$ gives the equilibrium fluctuation in the variable $f$

$$\Omega_f(0) = <f^2> - <f>^2. \tag{4}$$

The above equation is valid only when $N$ is sufficiently large, but in practice even for simulations of small proteins (containing on the order of 1000 atoms) in vacuum Eq. (4) is obeyed.

The reason $\Omega_f(t)$ is useful in analyzing certain aspects of the internal motions in proteins is that one can show that $\Omega_f(t)$ obeys a simple dynamic scaling law for $t$ greater than a certain characteristic time. If a particular observable, $F$, is self-averaging then

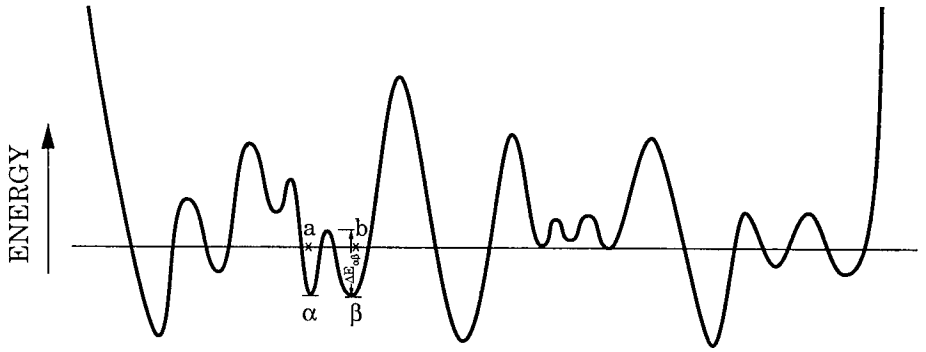


Fig. 1. Schematic of the energy surface of a protein. The solid horizontal line corresponds to a fixed value of the total energy. Isoenergetic conformers belonging to the minima labeled α and β, respectively, are separated by the potential barrier $\Delta E_{\alpha\beta}$. In general there are numerous minima and there is a distribution of barrier heights.

the fluctuation metric decays to zero at long times as[11]

$$\Omega_f(0)/\Omega_f(t) \approx D_f t \tag{5}$$

where $D_f$ is the rate at which the fluctuations involving the physical variable $F$ are sampled, and will be referred to as the generalized diffusion constant. Alternatively if $\Omega_f(t)$ does not obey the scaling relation given in Eq. (5) we can assume that the physical property $f$ is not self-averaging at least on the time scale of the simulations. We will show later that the lack of decay of $\Omega_f(t)$ at long times is a useful probe of the rough energy surface in proteins.

The reason for referring to $D_f$ as a generalized diffusion constant is that the relationship given in Eq. (5) is indicative of a diffusive process. Physically this implies that the approach to equilibrium for the phase space variable $F$ proceeds by a random walk process, and Eq. (5) is then the usual result for the long time behavior of such a Markov process. The slope of $1/\Omega_f(t)$ gives the rate of exploration of available conformations of the protein, and the approximate time needed for adequate sampling of these conformations is proportional to $D_f^{-1}$.

A necessary, but not sufficient, condition for the system dynamics to be ergodic is that $\Omega_f(t)$ obey Eq. (5). For example, imagine the dynamics of proteins in a rough free energy surface which is schematically shown in Figure 1. The two minima labeled α and β in Figure 1 are separated by a barrier. Assume that a particular $\tau_{ave}$ is not long enough for the system (so that the protein, which is initially in say α, is unable to overcome the barrier separating the two minima). Within either minima, given enough time, any trajectory will explore all of the allowed space. For a set of trajectories started in α, $\Omega_f(t)$ will decay to zero obeying Eq. (5), and the property $F(t)$ will be self-averaging. However, unless we have started one of our trajectories in region β (or $\tau_{ave}$ is long enough to overcome the barrier $\Delta E_{\alpha\beta}$) we cannot know that the partition exists and the system is effectively ergodic. Therefore, the decay of $\Omega_f(t)$ according to Eq. (5) is only a necessary condition for ergodicity. For a probe of the situation described above one has to compute the energy or force metrics as discussed in the next subsection. However, $\Omega_f(t)$ is easily calculable and is a powerful measure of the convergence of a given property that is dominated by dynamics in a single minimum. This would be the case at low temperatures or at temperatures such that the barriers separating the minima are large enough that they cannot be overcome even on experimental time scales. The latter situation pertains to glassy systems, and under these circumstances the conformations belonging to the two minima are truly disjoint. It should be noted that the computation of alternative measures of stochasticity such as Lyapunov exponents[19] in a protein are much more involved, and do not appear to be as relevant to the convergence of thermodynamic properties as the ergodic measure.

### Force Metric

In order to introduce the force metric let us assume that there are two distinct minima corresponding to two different conformational substrates of the molecule such as the ones labeled α and β in Figure 1. If the simulation time $\tau_{ave}$ is less than the typical time needed to cross the barrier separating the two minima then configurations belonging to a given minimum will in the process of evolution remain in the same basin for times on the order of $\tau_{ave}$. It should be noted that configurations belonging to the other minima are energetically accessible but if $\tau_{ave}$ is less than $\exp(\Delta E_{\alpha\beta}/k_B T)$ then the bottleneck prevents the system from sampling these configurations. If on the other hand the system is ergodic on the time scale $\tau_{ave}$ the bottleneck can be overcome and the conformations belonging to both α and β will be sampled. (Note that the arguments presented above explicitly assume that the number of degrees of freedom is very large and it is possible to ex-

change energy between the various degrees of freedom.) If there is a barrier separating the conformations belonging to two distinct minima, as is the case with the minima labeled α and β in Figure 1, then the simulation time must exceed $\tau_0 \exp(\Delta E_{\alpha\beta}/k_B T)$ for effective ergodicity to be reached where $\tau_0$ is a microscopic time scale like the typical vibrational time in one of the minima.

In order to probe whether there are two distinct minima separated by barriers it is necessary to compare properties of trajectories that are independent. Since the conformations belonging to the two minima have the same total energy it is desirable to construct an energy metric as was done in the context of structural glasses.[10] For proteins modeled by the standard potential energy function[2] it proves more convenient to consider the force metric which is computed using the time averaged values of the force on the $i$th atom. In introducing the force metric we first focus on the force experienced by the $i$th atom arising from the nonbonded potential only. Two *independent* initial states of the system corresponding to two distinct conformational states are chosen and are labeled a and b. For example, in Figure 1 the conformation labeled a belongs to the minimum labeled α while b belongs to the minimum β. The method of generating the two independent initial conformations of the system is discussed below. Let

$$\mathbf{f}^{a}_{iN}(t) = 1/t \int_0^t \mathbf{F}^{a}_{iN}(s)\, ds \tag{6}$$

be the time averaged value of the force due to the nonbonded potential acting on the $i$th atom. Similarly we compute $\mathbf{f}^{b}_{iN}(t)$ starting from the other independent starting configuration b. A measure that can distinguish between the forces (and hence the dynamics) in the two distinct minima (if they are present) is the nonbonded force metric

$$d_{FN}(t) = 1/N \sum_{i=1}^{N} \|\mathbf{f}^{a}_{iN}(t) - \mathbf{f}^{b}_{iN}(t)\|^2. \tag{7}$$

Suppose the two initial conformations a and b belong to the same minimum. Then $d_{FN}(t)$ (or a suitable measure constructed by following the dynamics of two independent trajectories) should vanish as t approaches $\tau_{ave}$ in a time averaged sense, i.e., one expects $f^{a}_{iN}(\tau_{ave}) = f^{b}_{iN}(\tau_{ave})$. If the initial conformations a and b belong to distinct minima then $d_{FN}(t) \to 0$ only when each trajectory crosses the barriers separating the two minima or when $\tau_{ave} > \tau_0 \exp(\Delta E_{\alpha\beta}/k_B T)$. On the other hand if $d_{FN}(\tau_{ave}) \neq 0$ then one can conclude that the initial conformations belong to two distinct conformational states separated by a barrier $\Delta E_{\alpha\beta}$ where on average $\tau_{ave} < \tau_0 \exp(\Delta E_{\alpha\beta}/k_B T)$. The long time behavior of the metrics of the form given by Eq. (7), which explicitly compares the time evolution of two independent trajectories of the system, can serve to distinguish between two different minima, and therefore can be used as a probe of the energy landscape in proteins. We have shown previously that for any Hamiltonian system, independent of the nature of the potential energy function, the metrics of the form defined by Eq. (7) also obey the scaling law given by[11]

$$d_{FN}(0)/d_{FN}(t) \approx D_{FN} t \tag{8}$$

where as before $D_{FN}$ is the rate at which the conformations responsible for the nonbonded degrees of freedom are sampled. The above equation is applicable only for ergodic systems, and if Eq. (8) is not obeyed then one can conclude that the system is not ergodic on the time $\tau_{ave}$. We have also computed the metric for the total force, $d_{FT}(t)$, which is defined as

$$d_{FT}(t) = \frac{1}{N} \sum_{i=1}^{N} \|f^{a}_{iT}(t) - f^{b}_{iT}(t)\|^2 \tag{9}$$

where $f^{a}_{iT}(t)$ is the time-average value of the total force on the $i$th atom. Interestingly, the total force metric proves to be of little value in examining the time scale for sampling conformation space. Decomposition of the force metric into its various components (like the nonbonded force) is much more useful. In the next section we also consider the total force metric and the dihedral angle force metric.

## Simulation Details

The MD simulations were performed at several temperatures, ranging from 40 to 400K for the bovine pancreatic RNase A/3′–UMP enzyme/product complex[20] and the corresponding S-peptide of RNase A. The X-ray structure of RNase A shows that it contains three short helices. When RNase A[21] is cleaved at the peptide bond between residues 19 and 20 one obtains the so called S-peptide[22] (residues 1–19) and the enzymatically active RNase S (residues 20–124). The S-peptide forms α-helical structure between residues 3 and 13[23] as can be seen in Figure 2. All the simulations were done in vacuum, and hence are not as realistic as the simulations that explicitly include the effect of solvent. Nevertheless for the purposes of demonstrating the analysis methods suggested here, and to show the interrelation between the various internal motions in proteins, the vacuum simulations should suffice. The starting structures for our simulation of both the full RNase A/3′–UMP enzyme/product system and the S-peptide was the 1.9 Å resolution X-ray crystal structure.[24] The initial configuration of the S-peptide was obtained by retaining the coordinates of the atoms of the first 19 amino acid residues of the full enzyme crystal structure. The initial configuration of the S-peptide was equilibrated at each temperature for approximately 50 psec (150 psec for the

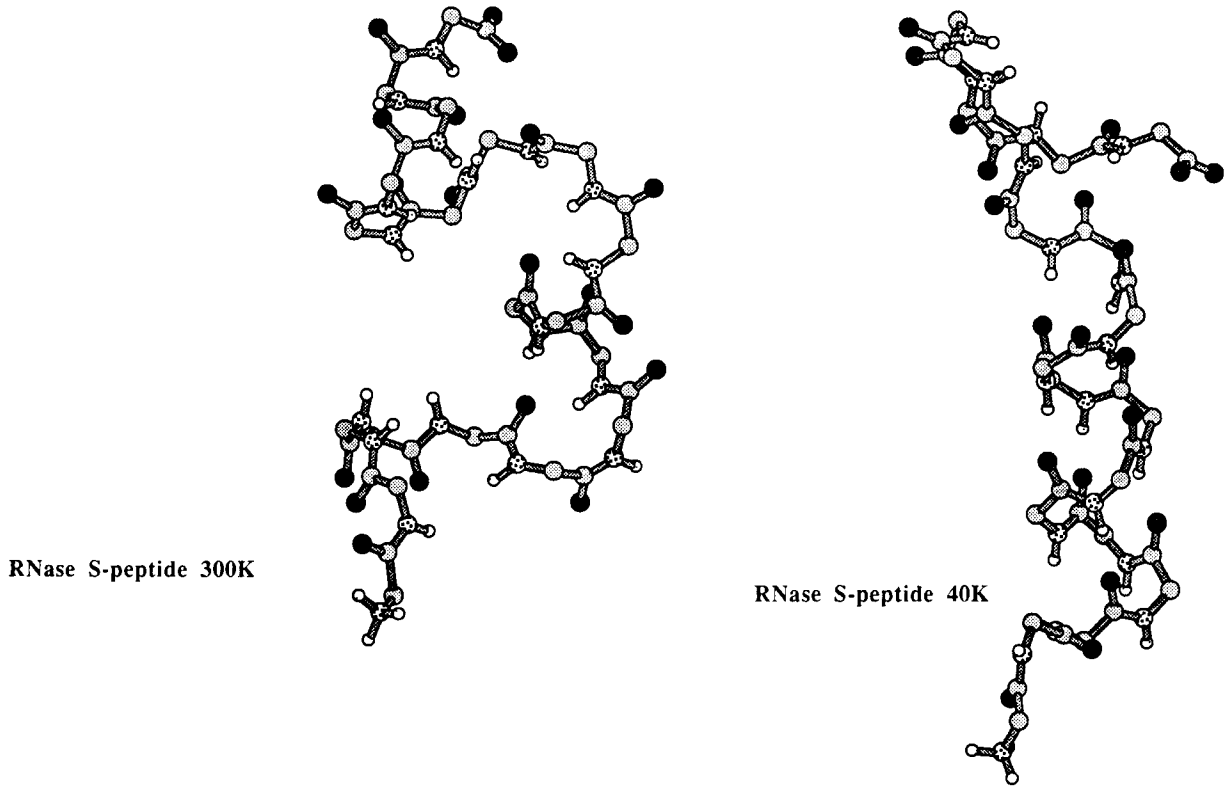


Fig. 2. Ball and stick representation of snapshots of the S-peptide backbone atoms at 40K and 300K.

40K trajectory) while the enzyme/product complex was equilibrated for 50 psec at each temperature. The equilibration was performed using uniform velocity rescaling every 0.5 psec. After the systems were equilibrated, two independent trajectories were generated for both systems at each temperature. In our model all atoms were free to move, i.e., no bond length constraint was enforced. The classical equations of motion were integrated using the Verlet algorithm with a time step of 0.5 fsec. The nonbonded interaction energy was cut off at $r_c = 12$ Å using a shifting function of the form

$$s(r) = \begin{cases} 1-2(r/r_c)^2 + (r/r_c)^4 & r \le r_c \\ 0 & r > r_c. \end{cases} \quad (10)$$

In these simulations an empirical distance-dependent dielectric constant was used, and a modification of version 20 of the CHARMM program[25] was employed with the version 19 polar-hydrogen potential function.

The computation of the force metric [cf. Eq. (7)] requires the generation of two independent trajectories. The generation of independent trajectories is not an issue at higher temperatures where the correlation times between various configurations are expected to be short. However, at low temperatures, especially when the dynamics is dominated by bottlenecks separating the different conformational substates (see Fig. 1), these correlation times can be extremely long. It is therefore important to ensure as far as possible the independence of the trajectories. For the S-peptide, the first trajectory was initiated from the endpoint of the equilibration run, and continued at constant energy for 75 psec. The second "independent trajectory" was generated by using a starting configuration of the endpoint of an equilibration run at a higher temperature. This higher temperature structure was then equilibrated at the desired temperature for 50 psec before generating a 75 psec trajectory at constant energy. For the trajectories at 40K and 80K the equilibrated 120K structure was used as a starting point; for the 120K, 160K, and 240K trajectories the equilibrated 300K structure was used; at 300K the equilibrated 400K structure was used; at 400K an equilibrated 500K structure was used.

For the enzyme/product complex system the first trajectory was initiated from the endpoint of the equilibration run and continued at constant energy for 30 psec. The second "independent" trajectory was generated for the 40, 120, and 240K runs starting from the equilibrated 300K structure followed by 25 psec equilibration at 40K or 12.5 psec equilibration at 120 and 240K. At 300K the equilibrated 300K structure was heated to 400K and run for 12.5 psec with velocity rescaling, followed by 12.5 psec equil-

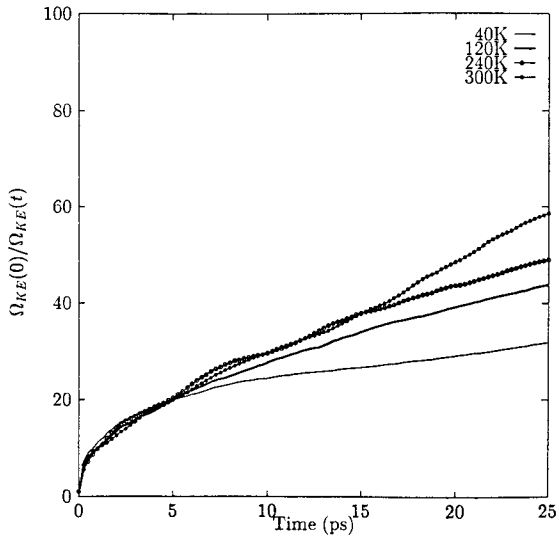


Fig. 3. The reciprocal of the normalized kinetic energy metric, $\Omega_{KE}(0)/\Omega_{KE}(t)$, for the S-peptide as a function of $t$ in psec. The thin solid line is for $T$ = 40K, the thick solid line for $T$ = 120K, the open circles for $T$ = 240K, and the closed circles for $T$ = 300K.

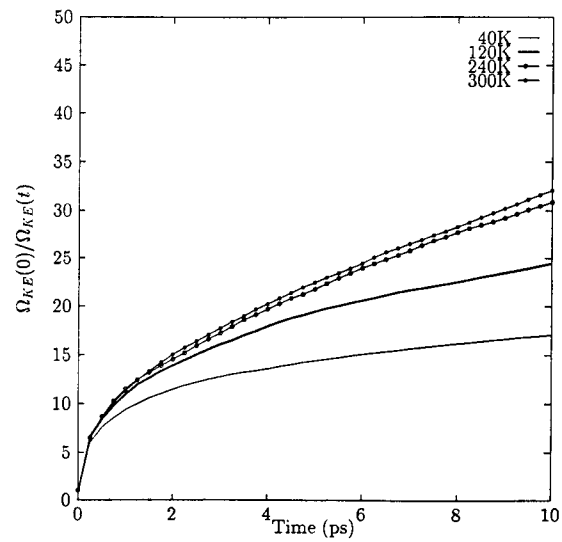


Fig. 4. Plots of $\Omega_{KE}(0)/\Omega_{KE}(t)$ at several temperatures for the RNase A/3′–UMP enzyme/product (referred to as enzyme/product) as a function of t. For an explanation of the symbols see Figure 3 caption.

ibration at 300K. These quasiindependent equilibrated configurations were used as the starting points for the second 30 psec trajectories.

## RESULTS AND DISCUSSION

In this section we present results for the kinetic energy fluctuation metric, the total force metric, and the decomposition of the force metric into the contribution arising from the dihedral angle potential and the nonbonded potential. In addition the temperature dependence of the generalized ergodic diffusion coefficients is analyzed.

### Kinetic Energy Fluctuation Metric:

The reciprocal of $\Omega_{KE}(t)$, normalized to unity at $t$ = 0, $\Omega_{KE}(0)/\Omega_{KE}(t)$, for the S-peptide as a function of $t$ is given in Figure 3. The results for $\Omega_{KE}(0)/\Omega_{KE}(t)$ for the enzyme/product complex system are given in Figure 4. There are several comments that are worth making based on the results displayed in Figures 3 and 4: (1) We see that both systems are effectively ergodic as far as the sampling of the kinetic energy is concerned. This means that within some finite time, which is set by $D_{KE}$ (see below), equipartitioning of the kinetic energy is expected to occur. (2) Analysis of the short time behavior of $\Omega_{KE}(t)$ yields some insight into the short-time librational motion undergone by the individual atoms or group of atoms. Since both of these systems adopt well-defined compact structures at low temperatures it is clear that many atoms are tightly confined in a "cage" of neighboring atoms. Thus at short enough times we expect a small amplitude rattling type motion in which an individual atom is perceived to move in a cage created by the neighboring atoms.[2] For short times the expected behavior for $\Omega_{KE}(t)$ is given by

$$\Omega_{KE}(t)/\Omega_{KE}(0) \approx 1 - (\text{Const})\, t^2 \qquad (11)$$

where the constant is independent of temperature and is related to the density of normal modes of the protein. The duration for which the above equation is valid is roughly the lifetime of the cage and can be estimated from Figures 3 and 4. Inspection of these figures shows that for up to $t_c \approx 3$ psec for the S-peptide, and up to $t_c \approx 1.5$ psec for the enzyme/product complex system, we find the behavior predicted by Eq. (5) in $\Omega_{KE}(t)$. For $t \leq t_c$ one is basically observing a rattling type motion in a cage, and for times greater than $t_c$ the atoms begin to collide inelastically with the neighbors resulting in exchange of energy which in turn leads to equipartitioning of the kinetic energy. The amplitude of the rattling type motion can be estimated using $t_c$ and this turns out to be in the range 0.5–1.2 Å. We note that the above picture of the short time dynamics has been well established using different methods of analysis.

The temperature dependence of $D_{KE}$ is shown in Figure 5 for both the S-peptide and the enzyme/product complex system. Before we analyze the results shown in Figure 5 we first argue that the fluctuation metric for the kinetic energy is a unique measure for heterogeneous systems. As discussed before $\Omega_{KE}(t)$ measures the mean-square fluctuations of the kinetic energy from the average value $K = 1/N \sum \frac{1}{2} m_i v^2_i(t)$, which according to the law of large numbers is $\langle K \rangle = 3k_B T/2$. If the system is effec-

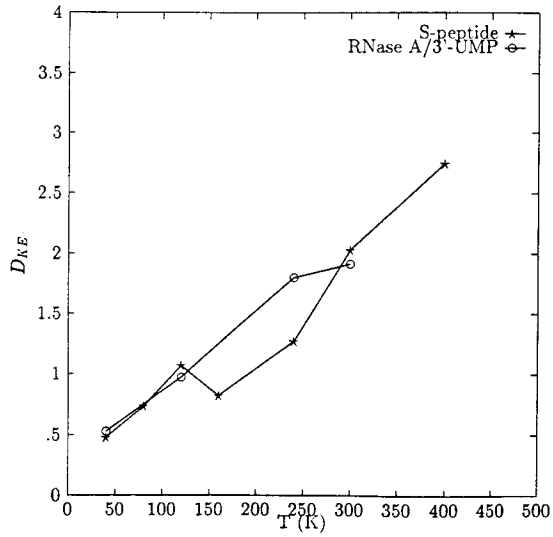


Fig. 5. The kinetic energy ergodicity diffusion constant $D_{KE}$ (in units of inverse psec) as a function of temperature. The circles correspond to the enzyme/product complex and the stars are for the S-peptide. The solid lines connect the various symbols as a guide to the eye.

tively ergodic on the averaging time scale $\tau_{ave}$ then we expect

$$\frac{1}{t_{ave}} \int_0^{t_{ave}} k_i(s)ds \rightarrow <K> \qquad (12)$$

*independent* of $i$. This is the equipartition theorem. In general for all other properties of heterogeneous systems the time average property of a physical variable will depend on the residue, and one will have a distribution of equilibrium values of the physical variable. It is for this reason that $\Omega_{KE}(t)$ plays a special role, and it also implies that we expect the equipartitioning of the kinetic energy to be the fastest of any property in proteins. The approximate time in which equipartitioning happens is proportional to $D_{KE}^{-1}$. Figure 4 shows that $D_{KE}$ is larger at all temperatures for the enzyme/product complex system than for the S-peptide. This implies that the time needed for equipartitioning of the kinetic energy is smaller for the enzyme/product complex system than for the S-peptide. The process of energy exchange requires collisions with neighboring atoms. The number of such collisions is expected to be frequent in the more compact enzyme/product complex system than in the S-peptide thus resulting in the smaller values of $D_{KE}^{-1}$. In general given roughly the same number of amino acids comprising a protein we can conjecture that $D_{KE}^{-1}$ would be smaller for the more compact structure. Structures with a larger proportion of surface residues should have smaller values of $D_{KE}^{-1}$. This picture is also supported by calculation of $\Omega_{KE}(t)$ on BPTI and myoglobin.[26] The above conclusion may not be valid (especially for the less compact structures) when the solvent molecules are included. In such cases collisions with solvent molecules can effectively increase $D_{KE}$ resulting in a decrease in time needed for equipartitioning.

It is of interest to obtain a quantitative estimate of $t_{ave}$ from the values of $D_{KE}$ shown in Figure 5. If we assume that equipartitioning is achieved if $\Omega_{KE}(0)/\Omega_{KE}(t) \approx 100$ then $t_{ave} \approx 100/D_{KE}$. This criterion gives $t_{ave} \approx 50$ psec for the enzyme/product complex at $T = 300$K. The arbitrariness involved in writing $\tau_{ave} = 100/D_{KE}$ can be removed by comparing the values of $t_{ave}$ at two different temperatures. The value of $t_{ave}$ increases by a factor of four in the temperature range 40–300K. From Figures 2 and 3 it is obvious $t_{ave}$ has to be much greater than $t_c$, and thus very optimistically we would estimate that a minimum acceptable $t_{ave}$ must be in the range of 10–30 psec. These values of $t_{ave}$, even for the equipartitioning of the kinetic energy, which is perhaps the simplest dynamic process that involves no obvious activated or cooperative mechanism for relaxation, are surprisingly long. It is likely that these times could be reduced, while not drastically[27] in a simulation that includes solvent molecules explicitly. For vacuum simulations it might be prudent to use noisy MD (low friction Langevin dynamics) to decrease $t_{ave}$ as was recently done in the simulation of the folding of a model heteropolymer[28] and as is commonly employed in calculating equilibrium properties of proteins.[29] However, preliminary results using the force metric indicate that the gain is small and that other simulation strategies such as hybrid MC[30] or mean-field algorithms[31] must be employed to significantly increase the rate of sampling conformation space.

## Force Metric

The inverse of the total force metric for the S-peptide, normalized to its initial value, as a function of time at several temperatures is shown in Figure 6. This figure shows that for all times

$$d_{FT}(0)/d_{FT}(t) = \delta t^2 \qquad (13)$$

where $\delta$ is a temperature independent constant. This is a very striking result. In particular from the $t^2$ dependence of $d_{FT}(t)$, which is very different from the scaling behavior expected for the GEM for effectively ergodic systems [cf. Eq. (8)], we may be tempted to conclude that as far as the force metric is concerned the system appears to sample only configurations belonging to a single valley. However the results given in Figure 6 follow from the notion of connectivity i.e., the atoms in a protein are connected to their neighbors by covalent bonds. The connectivity in our model is described by a harmonic potential with appropriate force constants. Because

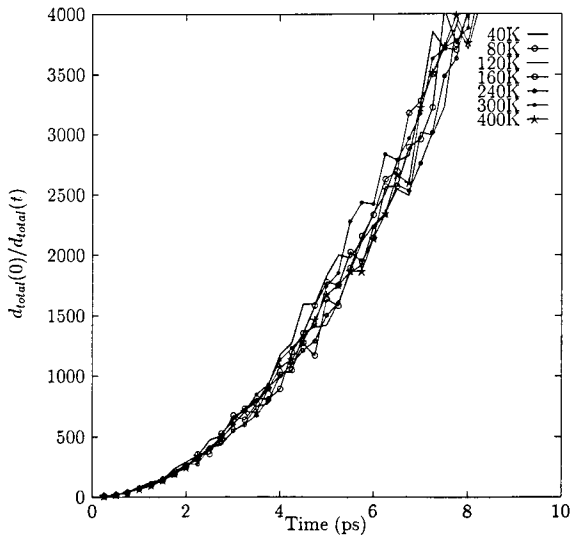


Fig. 6. The reciprocal of the total force metric, $d_{FT}(0)/d_{FT}(t)$, as a function of time for the S-peptide at seven temperatures. The plot demonstrates that the inverse of the normalized force metric is identical at all temperatures.

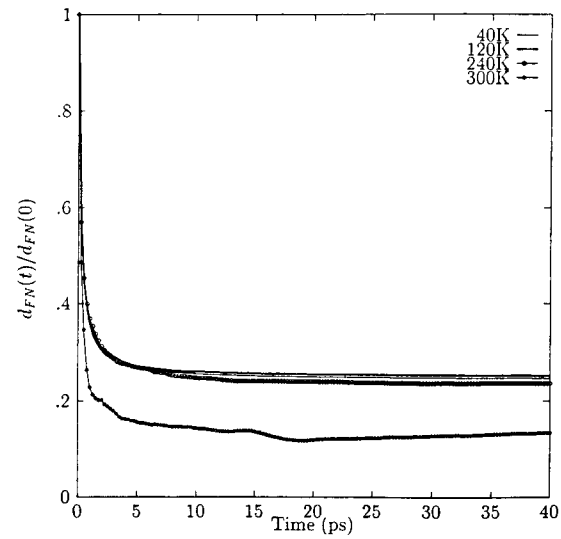


Fig. 7. Plots of the normalized nonbonded force metric, $d_{FN}(t)/d_{FN}(0)$, as a function of time for the S-peptide. The solid line corresponds to $T$ = 40K, the thin line is for $T$ = 120K, the open circles are for $T$ = 240K, and the closed circles are for $T$ = 300K.

of this constraint the displacement of individual atoms may be small enough that a normal mode picture is appropriate as far as the total force metric is concerned. If the normal mode picture is accurate then it can be easily shown that Eq. (13) results. The constant $\delta$ is related to the second moment of the normal mode spectrum. A far more general derivation of Eq. (13) will be provided elsewhere. From this result it follows that the total force should be dominated by the harmonic potential describing the connectivity in protein molecules. Clearly this result is fairly general and should hold for other proteins as well as polymeric systems.

In order to examine the dynamic interplay between the relaxation of dihedral angle motion and the nonbonded (Coulomb and van der Waals) interactions, which are responsible for enabling the systems to become compact at low temperatures, we have examined the force metrics corresponding to these degrees of freedom. In particular we have computed

$$d_{FD}(t) = \frac{1}{N}\sum_{i=1}^{N} \|\mathbf{f}_{iD}^{a}(t) - \mathbf{f}_{iD}^{b}(t)\|^2 \qquad (14)$$

where $N$ is the total number of atoms in the molecule, and $f_{iD}(t)$ is the time-average value of the force from the dihedral angle potential on the $i$th atom. The superscripts a and b correspond to two independent initial conformations of the system.

In addition we have also computed the fluctuation metric for the nonbonded force

$$\Omega_{FN}(t) = \frac{1}{N}\sum_{i=1}^{N} \|f_{iN}(t) - \bar{f}_N(t)\|^2 \qquad (15)$$

where $\bar{f}_N(t) = (1/N)\sum_{i=1}^{N} f_{iN}(t)$. Notice that $\bar{f}_N(t) \equiv 0$ for all $t$ even for the force arising from nonbonded interactions alone. This is because the force on particle $i$ due to particle $j$ is exactly opposite to that on particle $j$ due to particle $i$. The fluctuation metric for the force measures the mean-square fluctuations in the force, and therefore unlike the kinetic energy metric the long-time limit of $\Omega_{FN}(t)$ will not be zero but will have a well-defined value independent of the starting configuration for systems for which $t_{ave}$ is adequate for sampling the conformational space. On the other hand we expect $d_{FN}(t)$ [or $d_{FD}(t)$] would decay to zero at long times provided the conformations a and b mix on the time $t_{ave}$. More precisely, in order for $d_{FN}(t)$ [or $d_{FD}(t)$] to vanish at long times requires not only that the magnitude of the time average forces on the $i$th particle be equal for the initial conformation but also that their directions have to be parallel. The force metric thus enables one to obtain insight into the average correlations in the relative local orientations of the two conformations.

The quantities $d_{FN}(t)/d_{FN}(0)$ and $\Omega_{FN}(0)/\Omega_{FN}(t)$ for the S-peptide as a function of $t$ at various temperatures are displayed in Figures 7 and 8 respectively. The corresponding results for the enzyme/product complex system are given in Figures 9 and 10. The results for the S-peptide were calculated to 75 psec and are shown to 40 psec. There appears to be little convergence between 40 and 75 psec. Figures 7 and 9 show that for both systems $d_{FN}(t)$ decays extremely slowly. In both systems we observe a rather fast initial decay in about 5–10 psec followed by a very slow decay. The rapid decrease at short

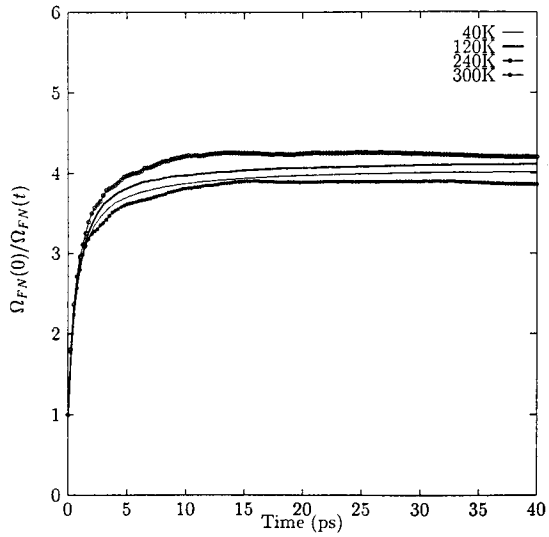


Fig. 8. Reciprocal of the fluctuation force metric for the nonbonded interactions as a function of time for the S-peptide at four temperatures. For an explanation of the various curves see Figure 7 caption.

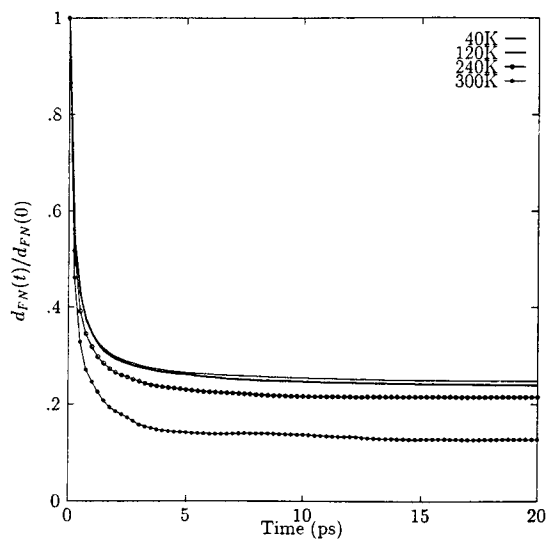


Fig. 9. Plots of $d_{FN}(t)/d_{FN}(0)$ as a function of time for the enzyme/product complex. The various curves correspond to different temperatures, and the symbols are the same as in Figure 7.

times implies that there are conformations that can be sampled easily without encountering any significant barriers. The nonvanishing of $d_{FN}(t)$ at long times (out to 75 psec) implies that significant cooperative motion may be involved in the process of sampling conformations belonging to the two distinct conformational states. It is interesting that even at the highest temperature $d_{FN}(t)$ decays extremely slowly at long times indicating the presence of significant barriers which require times of 75 psec

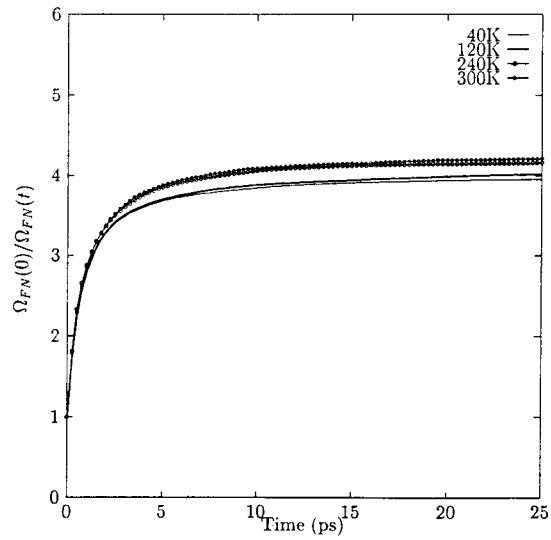


Fig. 10. Plots of $\Omega_{FN}(0)/\Omega_{FN}(t)$ as a function of time for the enzyme/product complex. The explanation of the various curves is given in Figure 7.

or longer to traverse. As discussed below the average barrier between the conformational substates can be estimated from the values of $D_{FN}$ [cf. Eq. (8)]. For the enzyme/product complex we find that $\Delta E_{\alpha\beta}$ is in the range of 2.3–3.5 kcal/mol.

To further characterize the motion in the two distinct conformational substates it is convenient to write $d_{FN}(t)$ as

$$d_{FN}(t) = \Omega^a(t) + \Omega^b(t) - X^{ab}(t) \qquad (16a)$$

where

$$\Omega^j = 1/N \sum_{i=1}^{N} \| \mathbf{f}^j_{iN}(t) - \bar{f}^j_N(t) \|^2 \qquad j = a,b \qquad (16b)$$

and

$$X^{ab}(t) = 2/N \sum_{i=1}^{N} \mathbf{f}^a_{iN}(t) \cdot \mathbf{f}^b_{iN}(t). \qquad (16c)$$

The qualitative features of the long time properties of $d_{FN}(t)$ can be predicted based on the decomposition given in Eq. (16). The cross term $X^{ab}(t)$ can be written as $2/N \sum_{i=1}^{N} |f^a_{iN}(t)| |f^b_{iN}(t)| \cos \theta_i(t)$ where $|f^a_{iN}(t)|$ is the magnitude of $\mathbf{f}^a_{iN}(t)$ and $\theta_i(t)$ is the angle between the vectors $\mathbf{f}^a_{iN}(t)$ and $\mathbf{f}^b_{iN}(t)$. The term $X^{ab}(t)$ provides information on correlations in the local time-averaged forces between conformations a and b. At high temperatures we expect that such correlations will persist only for short times. This should lead to the conclusion that the long time value of $X^{ab}(t)$ should be zero. However, at low temperatures we expect longer correlation times implying that $X^{ab}(t)$ either goes to a constant at long times or decays slowly. Furthermore if the conformers a and b mix on the time scale $\tau_{ave}$ implying adequate sam-

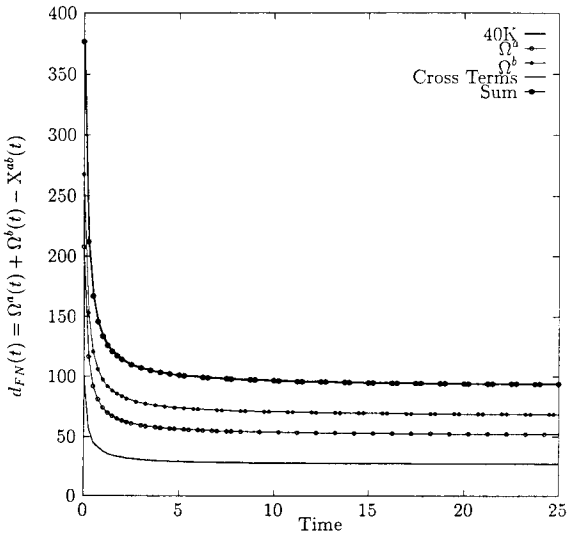


Fig. 11. Plots of $d_{FN}(t)$, $\Omega^a(t)$, $\Omega^b(t)$, and $X^{ab}(t)$ as a function of time for the S-peptide at $T$ = 40K. The thin solid line represents $X^{ab}(t)$, the open circles correspond to $\Omega^a(t)$, the closed circles are $\Omega^b(t)$, and the thick solid line is $d_{FN}(t)$.

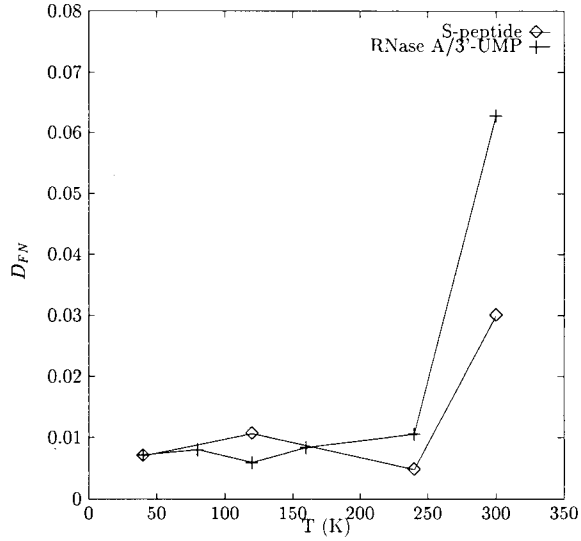


Fig. 12. Same as Figure 11 except the temperature corresponds to $T$ = 300K. The symbols are explained in Figure 11.

pling has occurred then $\Omega^a(\tau_{ave}) \simeq \Omega^b(\tau_{ave})$ independent of a and b. From the above arguments it follows that the conformations a and b belong to the same cluster of CS if either of the following conditions are satisfied: (1) $d_{FN}(\tau_{ave}) = 0$ which is a very stringent criterion and the time for this to happen can scale very roughly as exp($N$); (2) $X^{ab}(\tau_{ave}) = 0$ *and* $\Omega^a(\tau_{ave}) \simeq \Omega^b(\tau_{ave})$. If either of the above conditions is not satisfied then we can conclude that the conformations a and b belong to *distinct* conformational substates, and longer times are needed for the mixing of a and b.

The various terms in Eq. (16a), namely $\Omega^a(t)$, $\Omega^b(t)$, and $X^{ab}(t)$, for the S-peptide at $T$ = 40K and $T$ = 300K are plotted in Figures 11 and 12, respectively. It is obvious from Figure 11 that at 40K $\Omega^a(t$ = 25 psec) $\neq \Omega^b(t$ = 25 psec) and that the cross term $X^{ab}(\tau_{ave})$ is significantly different from zero. There is little change over the full 75 psec duration of our simulation. This clearly demonstrates that a and b belong to two different conformational substates separated by barrier(s) which cannot be overcome on a 75 psec time scale. On the other hand we notice in Figure 12 that at 300K $\Omega^a$(25 psec) = $\Omega^b$(25 psec) and $X^{ab}(t) \approx 0$ for all $t$. It therefore follows that the conformations a and b belong to the same conformational substate (or at least to dynamically equivalent substates). The vanishing of $X^{ab}(t)$ for all times implies that the conformations a and b are truly independent, and the fact that $\Omega^a(\tau_{ave}) = \Omega^b(\tau_{ave})$ suggests that the two conformers mix in a time on the order of 25 psec indicating that a and b belong to dynamically equivalent substates.

Because the systems explore conformations belonging to two distinct CS at low temperatures it is of interest to ascertain the structures that are being sampled during the time evolution of the two independent trajectories. The easiest way to examine this is to obtain average structures of the systems for the two independent trajectories. In the case of the S-peptide we find that the helical structure is maintained during the time evolution of the trajectories. For the enzyme/product complex we also find that the substrate UMP is always bound to the enzyme at all times. This suggests that the relevant conformations that are sampled starting from two independent initial conformations maintain the overall three-dimensional structure of the proteins. Thus the average structure of the protein in the distinct conformational substates probably differs only in the relative orientations of the backbone and the side chains of the protein.

The process of relaxation in the protein molecule involves primarily the dynamics of the dihedral angles, and the relaxation associated with nonbonded degrees of freedom. We have shown that metrics for the nonbonded interactions exhibit an initial fast relaxation followed by a slower decay which presumably involves large scale collective motion. It is of interest to see if the relaxation of nonbonded interactions is coupled to dihedral angle transitions. In order to investigate this aspect we have computed the dihedral angle force metric [cf. Eq. (14)]. The results for the S-peptide are shown in Figure 13. At all temperatures the time scale for the relaxation of $d_{FD}(t)$, $\tau_D$, appear to be very similar to that for $d_{FN}(t)$, namely $\tau_N$. To demonstrate more succinctly this interrelationship between the nonbonded dy-

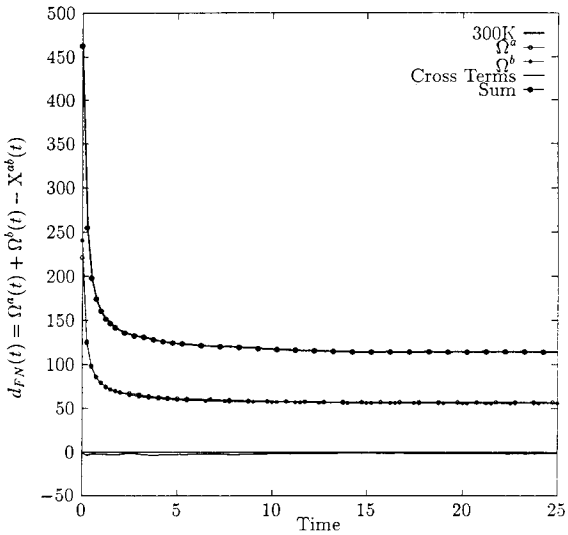


Fig. 13. Reciprocal of the dihedral angle force metric $d_{FD}(0)/d_{FD}(t)$ as a function of time for the S-peptide. For an explanation of the symbols representing the different curves see Figure 7.

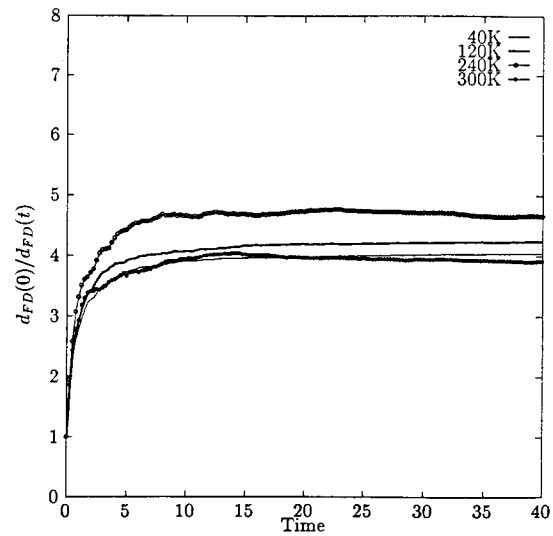


Fig. 14. Plot of $\Delta_{FD}(t)$ [cf. Eq. (17)] showing the difference between the nonbonded force metric and the dihedral angle force metric for the S-peptide. The symbols representing the various curves are explained in Figure 7.

namics and the relaxation of the dihedral angle degrees of freedom we plot the quantity

$$\Delta_{ND}(t) = \frac{d_{FN}(0)}{d_{FN}(t)} - \frac{d_{FD}(0)}{d_{FD}(t)} \tag{17}$$

as a function of $t$ for the S-peptide at four different temperatures in Figure 14. By comparing Figures 8, 12, and 13 we can conclude that within fluctuations $d_{FN}(0)/d_{FN}(t) \approx d_{FD}(0)/d_{FD}(t)$ for all $t$. This suggests a mechanism for the relationship between the relaxation of dihedral angle transitions and the nonbonded degrees of freedom. If $\tau_D$ [given by the slope of $d_{FD}(0)/d_{FD}(t)$] is much larger than $\tau_N$ then the relaxation in the protein molecule would first involve dynamics involving the nonbonded forces followed by the relaxation associated with dihedral angle degrees of freedom. If $\tau_N/\tau_D >> 1$ then the situation would be reversed. However we find $\tau_N \sim \tau_D$ at all temperatures and thus we conclude that the relaxation of the nonbonded interaction is mediated by dihedral angle transitions. This finding has an important consequence for the dynamics of the folding of proteins. It is believed that the conformation of proteins in the native state is maximally compact. The formation of compact structure is possible because certain residues that are well separated in sequence space (i.e., nonbonded) prefer to be close to each other in configuration space. Thus it is reasonable to assume that the process of achieving a state of higher compactness (i.e., having a spatial arrangement with higher packing fraction) involves the relative motion of nonbonded residues. Based on our results we can assert that the time needed for reaching the maximally compact structure should be roughly proportional to $\tau_N$.

The temperature dependence of $D_{FN}$, which is obtained from the slope of $d_{FN}(0)/d_{FN}(t)$, for the two systems is shown in Figure 15. The results in Figure 15 indicate that for the S-peptide the time needed for compact helical structure formation ($\tau_c \alpha D_{FN}^{-1}$), which happens for $T \leq 120$K, is extremely long. In fact $\tau_c$ is much longer than our simulation times. By comparing Figures 5 and 15 we find that the ratio $D_{KE}/D_{FN} \sim 100$ which suggests that the time needed for relaxation of the nonbonded degrees of freedom is a much slower process than the equipartitioning of the kinetic energy. This is in accord with our intuition that the relaxation of nonbonded degrees of freedom involves large length scale cooperative motion, and hence can have significant barriers. On the other hand equipartitioning of kinetic energy can take place by independent random collisions of residues with their neighbors and thus should be a more facile process. The values of $D_{FN}$ can be used to estimate the average activation energy separating the conformational substates $\alpha$ and $\beta$. Since $D_{FN}^{-1}$ is the time required for sampling the conformations belonging to the two distinct CS we can write $D_{FN}^{-1} \sim \tau_0 \exp(\Delta E_{\alpha\beta}/k_B T)$. If $\tau_0$ is taken to be about 1 psec the values of $D_{FN}$ in Figure 15 lead to an estimate for $\Delta E_{\alpha\beta}$ in the range of 2.2–3.5 kcal/mol. These values fall in the range of activation energies calculated more directly for the alanine tetrapeptide.[32]

## SUMMARY AND CONCLUSIONS

In this paper we have introduced a very general method for analyzing the dynamics of heterogeneous systems such as proteins and nucleic acids. The spe-

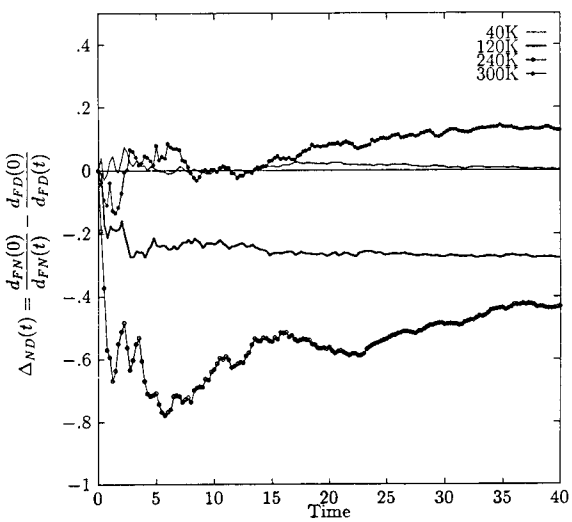


Fig. 15. The nonbonded force metric diffusion constant, $D_{FN}$ (psec$^{-1}$) [cf. Eq. (8)] as function of temperature. The open circles correspond to the enzyme/product complex and the crosses are the S-peptide.

cific application of these concepts to the study of various motions in the two systems—the S-peptide and the RNase A/3′–UMP enzyme/product complex—have yielded results which we believe to be general properties of proteins.

We have shown that the time needed for equipartitioning of kinetic energy is quite long even at biological temperatures. This is somewhat surprising because the kinetic energy is the simplest property that will have a unique temperature-dependent long time average value independent of the residue. The long times involved in the kinetic energy equipartitioning are probably due to the presence of low frequency normal modes, where $\nu \sim 3$ cm$^{-1}$ and the period $T = 1/\nu \sim 10$ psec, which take a significantly longer time to average over. We have also found that at a given temperature the more compact a structure is the smaller is the value of $D_{KE}^{-1}$. The time needed for kinetic energy equipartitioning is therefore a very rough indication of the degree of compactness in proteins. It also follows from our analysis that these results will be true of homopolymer systems as well. It is likely that $D_{KE}^{-1}$ can be decreased by using low friction Langevin dynamics or a hybrid algorithm that uses a combination of the standard MD algorithm and Monte Carlo method.[30]

The use of the force metric has revealed completely new ways of probing multiple conformational substates in proteins. We have shown that by comparing the dynamics of two conformations that are initially independent we can infer whether these conformations belong to the same CS or belong to two distinct CS separated by a barrier. More importantly the approximate time scale $D_{FN}^{-1}$ (and the barrier $\alpha \ln D_{FN}^{-1}$) for two conformations belonging to the two distinct CS to mix can be obtained from the scaling law obeyed by the force metric. The time $D_{FN}^{-1}$ can be interpreted as the approximate average residence time in a conformational substate. These aspects have been demonstrated by examining the time dependence of the force metric for both the nonbonded potential and the dihedral angle potential. On the other hand the time dependence of the total force metric for all temperatures is superimposable on to a curve that decays as $t^{-2}$, and hence this measure while interesting is not a useful indicator of the degree of conformation sampling.

There is an interesting correlation between the dynamics of the dihedral angle motion and the process of nonbonded relaxation. In particular the dynamics of both processes happen on approximately the same time scale implying that the nonbonded relaxation is parametrically dependent on dihedral angle transitions. The formation of folded structures in proteins is possible because of the presence of favorable nonbonded interactions. It is reasonable to suggest that the time scale for achieving a structure of maximal compactness is intimately related to the number of dihedral angle transitions required and the activation energy for each transition. The large scale cooperative motion associated with the relaxation of nonbonded interactions is in fact driven by a series of "local" dihedral angle transitions. The nonbonded force metric can be used as a novel analytical means of studying the kinetics of folding of a protein into compact native structures.

In the simulation of proteins one has to contend with the possibility of a complex energy landscape with many minima separated by a wide range of barrier heights. Our results suggest that in order to optimize the conformation space sampling it is better to perform averages over several relatively short, independent MD trajectories rather than one very long one. The rationale behind this is that by generating several independent trajectories (by a combination of heating and cooling methods) one can sample more efficiently isoenergetic conformations belonging to independent conformational substates.[33] Thus in proteins in which one has to contend with a rugged energy landscape meaningful simulations should use a combination of the molecular dynamics method in conjunction with the simulated annealing technique. It is also likely that a combination of energy minimization method and molecular dynamics technique[34] can enhance the rate of sampling of conformation space. The other important lesson is that the GEM can be used to design MD simulations. Because the scaling behavior for the GEM is obtained in relatively short times for systems which are ergodic the associated generalized diffusion constant and the value $\tau_{ave}$ can be easily found. This sort of analysis is particularly useful in the calculation of the free energy difference

between reactant and product by free energy perturbation methods in which several MD simulations corresponding to different Hamiltonians are needed.

## ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation (Grant NSF CHE-86-09722) and the Camille and Henry Dreyfuss Foundation.

## REFERENCES

1. McCammon, J.A., Harvey, S.C. "Dynamics of Proteins and Nucleic Acids." London: Cambridge University Press, 1987.
2. Brooks, C.L., Karplus, M., Pettit, M. "Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics." New York: John Wiley, 1988.
3. Frauenfelder, H., Parak, F., Young, R.D. Conformational substates in proteins. Annu. Rev. Biophys. Chem. 17:451–479, 1988.
4. Steinbach, P.J., Ansari, A., Berendzen, J., Braustein, D., Chu, K., Cowen, B.R., Ehrenstein, D., Frauenfelder, H., Johnson, J.B., Lamb, D.C., Luck, S., Mourant, J.R., Nienhaus, G.V., Ormos, P., Phillip, R., Xie, A., Young, R.D. Ligand binding to heme proteins: Connection between dynamics and function. Biochemistry 30:3988–4001, 1991.
5. Ansari, A., Berendzen, D., Braustein, D., Cowen, B.R., Frauenfelder, H., Hong, M.K., Iben, I.E.T., Johnson, J.G., Ormos, P., Sanke, T.B., Scholl, R., Schulte, A., Steinbach, P.J., Vittitow, J., Young, R.D. Rebinding and relaxation in the myoglobin pocket. Biophys. Chem. 26:337–355, 1987.
6. Berendzen, J., Braustein, D. Temperature derivative spectroscopy: A tool for protein dynamics. Proc. Natl. Acad. Sci. U.S.A. 87:1–5, 1990.
7. Elber, R., Karplus, M. Multiple conformational states of proteins: A molecular dynamics analysis of myoglobin. Science 235:318–321, 1987.
8. Noguti, T., Go, N. Structural basis of hierarchical multiple substates of a protein. I: Introduction. Proteins 5:97–103, 1989.
9. Frauenfelder, H., Sligar, S.G., Wolynes, P.G. The energy landscapes and motions of proteins. Science 259:1598–1603, 1991.
10. Thirumalai, D., Mountain, R.D., Kirkpatrick, T.R. Ergodic behavior in supercooled liquids and in glasses. Phys. Rev. A. 39:3563–3574, 1989.
11. Thirumalai, D., Mountain, R.D. Ergodic Convergence properties of supercooled liquids and glasses. Phys. Rev. A. 42:4574–4587, 1990.
12. Goldstein, M. Viscous liquids and the glass transition: A potential energy barrier picture. J. Chem. Phys. 51:3728–3739, 1969.
13. Austin, R.H., Beeson, K.W., Eisenstein, L., Frauenfelder, H., Grunslaus, I.C. Dynamics of ligand binding to myoglobin. Biochemistry 19:5147–5157, 1975.
14. Doster, W., Cusak, S., Petry, W. Dynamical transition of myoglobin revealed by inelastic neutron scattering. Nature (London) 337:754–756, 1989.
15. Loncharich, R.L., Brooks, B.R. Temperature dependence of dynamics of hydrated myoglobin. J. Mol. Biol. 215:439–455, 1990.
16. Angell, C.A. Perspective on the glass transition. J. Phys. Chem. Solids 49:863–871, 1988.
17. Ma, S.K. "Statistical Mechanics." Singapore: World Scientific, 1989.
18. Mountain, R.D., Thirumalai, D. Measures of effective ergodic convergence in liquids. J. Phys. Chem. 93:6975–6979, 1989.
19. Lichtenberg, A.J., Lieberman, M.A. "Regular and Stochastic Motion." New York: Springer-Verlag, 1983.
20. Richards, F.M., Wyckoff, H.W. Bovine pancreatic ribonuclease. Enzymes 4:647–806, 1971.
21. Cantor, C.R., Schimmel, P.R. "Biophysical Chemistry." New York: Freeman, 1980.
22. Nelson, J.W., Kallenbach, N.R. Stabilization of the ribonuclease S-peptide α-helix by trifluroethanol. Proteins 1:211–217, 1986.
23. Kim, P.S., Baldwin, R.L. A helix stop signal in the isolated S-peptide of ribonuclease A. Nature (London) 307:329–334, 1984.
24. Gilbert, W.A., Fink, A.L., Petsko, G.A. Unpublished.
25. Brooks, B.R., Braccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., Karplus, M. CHARMM: A program for macromolecular energy, minimization and dynamics calculations. J. Comp. Chem. 4:187–217, 1983.
26. Straub, J.E., Kuczera, K., Thirumalai, D., Karplus, M. Unpublished.
27. Swaminathan, S., Ichiye, T., van Grunsteran, W.F., Karplus, M. Time dependence of atomic fluctuations in proteins: Analysis of local motions in bovine pancreatic trypsin inhibitor. Biochemistry 21:5230–5241, 1982.
28. Honeycutt, J.D., Thirumalai, D. Metastability of folded states of globular proteins. Proc. Natl. Acad. Sci. U.S.A. 87:3526–3529, 1990.
29. Straub, J.E., Lim, C., Karplus, M. Simulation analysis of the binding interactions in the RNase A/3'-UMP enzyme-product complex as a function of pH. J. Am. Chem. Soc., submitted.
30. Duane, S., Kennedy, A.D., Pendelton, B.J., Roweth, D. Hybrid Monte Carlo. Phys. Lett. B. 195:216–222, 1987.
31. Elber, R., Karplus, M. Enhanced sampling in molecular dynamics: Use of time-dependent Hartree approximations for simulation of carbon monoxide diffusion through myoglobin, J. Am. Chem. Soc. 112:9161–9175, 1990.
32. Czerminski, R., Elber, R. Reaction path study of conformational transitions in flexible systems: Application to peptides. J. Chem. Phys. 92:5580–5601, 1990.
33. Frauenfelder, H., Petsko, G.A., Tsernoglou, D. Temperature dependent X-ray diffraction as a probe of protein structural dynamics. Nature (London) 280:558–563, 1979.
34. Levitt, M. Protein folding by restrained energy minimization and molecular dynamics. J. Mol. Biol. 170:723–764, 1983.