RESEARCH ARTICLES

Chaos in Protein Dynamics

Michael Braxenthaler,1 Ron Unger,2 Ditza Auerbach,3 James A. Given,1 and John Moult1*
1Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, Rockville, Maryland
2Department of Life Sciences, Bar-Ilan University, Ramat-Gan, Israel
3Department of Mathematics and Computer Science, Bar-Ilan University, Ramat-Gan, Israel

ABSTRACT MD simulations, currently the most detailed description of the dynamic evolution of proteins, are based on the repeated solution of a set of differential equations implementing Newton’s second law. Many such systems are known to exhibit chaotic behavior, i.e., very small changes in initial conditions are amplified exponentially and lead to vastly different, inherently unpredictable behavior. We have investigated the response of a protein fragment in an explicit solvent environment to very small perturbations of the atomic positions (10^-3–10^-9 Å). Independent of the starting conformation (native-like, compact, extended), perturbed dynamics trajectories deviated rapidly, leading to conformations that differ by approximately 1 Å RMSD within 1–2 ps. Furthermore, introducing the perturbation more than 1–2 ps before a significant conformational transition leads to a loss of the transition in the perturbed trajectories. We present evidence that the observed chaotic behavior reflects physical properties of the system rather than numerical instabilities of the calculation and discuss the implications for models of protein folding and the use of MD as a tool to analyze protein folding pathways. Proteins 29:417–425, 1997. © 1997 Wiley-Liss, Inc.

Key words: nonlinear dynamics; chaotic motion in complex systems; protein folding pathways; molecular dynamics; Lyapunov exponent

INTRODUCTION

It is widely accepted that protein folding must follow a pathway through conformational space from any accessible unfolded conformation to the native conformation. Only in this way can the functional conformation be found on a biologically useful time scale.1

Although this general paradigm is accepted, the actual mechanisms involved are highly controversial. A range of models have been suggested to describe the nature of the folding pathway: jigsaw,2 framework,3–5 initial random collapse,6 nucleation7 followed by propagation,7 and/or diffusion or collision.8 A fundamental question about the folding process is to what extent the events along the folding pathway are invariant, given similar initial conditions.

Access to details of the nature of transient folding events is hampered by experimental difficulties in characterizing partially folded intermediates and the computational complexity of even small protein systems. The very early folding events often seem to occur locally in particular regions of the polypeptide chain.9,10 Study of such small regions is more tractable than the study of complete proteins.

MD simulations provide the most detailed description of the dynamic evolution of protein systems currently available. Many such systems of nonlinear differential equations are known to exhibit chaotic behavior.11 These are deterministic systems but have long-term behavior that is practically impossible to predict. One of the hallmarks of such systems is that two almost identical starting configurations evolve very differently in time. A system is defined as chaotic if small perturbations in its initial configuration are amplified exponentially with time.

We have used the trajectory of the 1,000-ps MD simulation of a 13-residue barnase fragment in the folded state12 and a 600-ps trajectory of the same fragment started from a representative unfolded conformation (unpublished results) to examine sensitivity to small perturbations.

In the technical part of the work we have three aims. The first aim is to show specifically that MD...
simulations of protein folding exhibit chaotic behavior. The second aim is to calculate the divergence rate of the system at different stages of the folding process, and the third aim is to study the effect of this chaotic behavior on the occurrence of certain folding events during a particular simulation. Taken together, conclusions from these studies influence our perspective on protein folding.

The structure of the article is as follows. In Materials and Methods we discuss how signatures of chaos can be detected in simulations of complex heterogeneous systems, such as protein molecules, and describe the MD simulations used to do this. In Results, we show how the system responds to small perturbations and what happens if such perturbations are introduced. Finally, in Discussion we present the implications of these observations for models of protein folding pathways.

**METHODS**

**Measurement of Chaos in MD Simulations**

Intuitively, a deterministic system is considered chaotic if a very small perturbation in its initial condition has a large effect on the subsequent behavior of the system. In particular, a small perturbation is amplified exponentially in time. More formally, for a one-dimensional discrete time system (a system characterized by a single variable x): Assume that at time \( t = 0 \), a small change \( \Delta x_0 \) is made in the value of the variable, and two trajectories are run, one starting from \( x = x_0 \) and the other from \( x = x_0 + \Delta x_0 \). For a chaotic system the difference in x increases with each time step, and successive iterations result in an exponential deviation:

\[
\Delta x_t \sim \Delta x_0 \exp(\lambda t) \tag{1}
\]

where \( \Delta x_t \) is the difference in the value of \( x \) between the original and the perturbed trajectory after time step \( t \). The coefficient \( \lambda \) is known as the Lyapunov exponent of the system. For a multidimensional system, the exponential deviation is characterized by many Lyapunov exponents (as many as the number of degrees of freedom in the system). For example, a two-dimensional system may have different sensitivity to perturbations in the x and y directions and thus needs two exponents. It must be noted, however, that the Lyapunov exponents are properties of the whole system and do not generally map onto specific dimensions. A direct way to show that a system behaves chaotically is to show that it has at least one positive global Lyapunov exponent.\(^{13} \)

We seek to demonstrate qualitatively that an MD description of protein folding is chaotic. There are established procedures for calculating the value of the largest global Lyapunov exponent in dynamic systems by introducing small perturbations at many points along a trajectory in a coupled manner (the rescaling method\(^{13,14} \)) so as to track the direction of maximum deviation through phase space. In principle, these procedures may be applicable to large systems, such as proteins, but this has not yet been investigated.

More easily calculated are the finite time Lyapunov exponents \( \kappa(X) \).\(^{15} \) (We use here the term finite time Lyapunov exponents; the term local Lyapunov exponents also is frequently used in the literature in the same sense.) These measure the deviation rate between perturbed trajectories over short times (relative to the time needed to explore the whole phase space), starting at specific points:

\[
\kappa(X) = \log \frac{d(\vec{X}' - \vec{X})}{dt} \tag{2}
\]

where \( \vec{X}' \) is the vector of coordinates and momenta of the particles in the perturbed trajectory, and \( \vec{X} \) is the corresponding quantity in the unperturbed trajectory.

For a wide range of systems, given enough finite time Lyapunov exponents derived from appropriate intervals of the trajectory, global properties of these systems, such as the global maximal Lyapunov exponent, can be approximated.\(^{16} \)

We estimate a set of maximal finite time Lyapunov exponents by introducing very small perturbations at many, uncoupled points along the MD trajectories and measuring the resulting deviation for finite time intervals (Fig. 1).

The rate of separation of two trajectories in phase space involves deviations in both positions and momenta. For simpler particle systems, such as a Lennard-Jones gas, it has been shown that the same...
average rate of separation is found when one takes into account only the position variables. Thus, in this study we approximate distance in the phase space by the RMSD of atomic coordinates of the peptide and do not use the solvent or any of the momenta. We have estimated the maximal finite time Lyapunov exponents under different conformational conditions. We return to the relationship between the finite time exponents and the largest global Lyapunov exponent for this system in the discussion.

MD Simulations

We previously reported the analysis of a 1-ns MD simulation of the helix-loops fragment of barnase, a folding initiation site, in its native conformation isolated from the rest of the protein. The fragment was chosen based on experimental and computational data supporting its involvement in folding initiation events. Numerical experiments to examine its sensitivity to small perturbations were made on this trajectory. In addition, a 600-ps trajectory of the same protein fragment starting at a representative denatured conformation (unpublished results) was used to compare the sensitivity with perturbations of native-like vs. denatured states.

The native-like system consists of the protein fragment in the experimental conformation and 1,598 water molecules in a box of $37 \times 37 \times 37 \, \text{Å}^3$; the denatured system contains the fragment and 2,042 water molecules in a box of $40 \times 40 \times 40 \, \text{Å}^3$. Both systems were equilibrated as described previously, and trajectories were created for 1,000 ps for the native-like simulation and 600 ps for the denatured simulation, at 300 K with a time step of 1 fs under periodic boundary conditions, using the program DISCOVER (Biosym Technologies, Inc., San Diego, CA) with the CVFF force field. Restart files were saved every 10 ps.

The initial conformation for the simulation of the denatured state was generated by randomly assigning torsion angles according to residue type, by using a library of observed values in native proteins. This starting conformation was largely extended (30 Å between termini) and did not exhibit long-range side-chain to side-chain contacts. In the course of the simulation the fragment gradually deviated from the starting conformation, reaching a C$_\alpha$ RMSD of 5.5 Å over 600 ps (as opposed to 2.7 Å after 1,000 ps in the native-like simulation). The conformations encountered during the simulation range between 5.5 and 8.5 Å C$_\alpha$ RMS difference from the native conformation. Over the whole course of the simulation the fragment maintained a noncompact state, although an open loop between the N- and C-terminal regions was formed.

Perturbations and Measurements

The restart files of the original long MD runs were used as branching points to introduce perturbations in atomic coordinates and rerun parts of the trajectories under otherwise identical conditions. In all cases perturbations were introduced as translations of atomic coordinates by a defined distance in a random direction.

The RMSD on all atoms of the peptide between time-corresponding sets of Cartesian coordinates in the unperturbed and perturbed trajectories was used as a measure of deviation of the two trajectories from each other. No transformation to obtain the best superposition of the sets of coordinates was used. This RMSD, defined over a partial Cartesian subspace, serves as a distance function $d(t)$ between two trajectories.

In addition to the Cartesian RMSD measure, deviation was also measured in torsion space as follows:

$$d(t) = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\tau_{pi}(t) - \tau_{ui}(t))^2}$$

where $u$ denotes conformations in the unperturbed trajectory, $p$ denotes conformations in the perturbed trajectory at time $t$ after perturbation, and the summation is over all the $N$ torsion angles $\tau$. Torsion angles were unlimited in the sense that they were allowed to extend from $-\infty$ to $\infty$ as opposed to closing the angle space at $-180^\circ/180^\circ$.

The slope over the linear region of a log-plot of $d(t)$ vs. $t$, where $d(t)$ is the Cartesian or torsion space distance between perturbed and unperturbed trajectories at time $t$ after perturbation, provides an estimate of the maximal finite time Lyapunov exponent at a point on the trajectory.

To study the effects of chaotic behavior on the robustness of conformational transitions, we also ran longer trajectories starting from perturbed conformations and compared the occurrence of folding events in these and the original trajectories. We use as an example significant changes in the value of a particular backbone torsion angle that was considered important to the folding pathway of the barnase fragment.

RESULTS

Chaotic Behavior of MD Simulations

The initial experiment was designed to establish which set of atoms should be perturbed. The full spectrum, from perturbing every atom including solvent in the system, via perturbing all the protein atoms, to perturbing only a single atom was examined. The various sets of atoms were perturbed by 0.001 Å per atom, followed by a short (0.5 ps) run of MD. The perturbation was done every 10 ps between 230 and 300 ps of the original MD run, yielding eight
pairs of short trajectories. After a short (approximately 0.2 ps) lag phase (the reason for which will be discussed below), the deviations grow exponentially with a very similar slope, approximately $5 \text{ ps}^{-1}$, for the very different perturbations (Fig. 2).

Considering the similarity in deviation characteristics, we settled for perturbation of all protein atoms in all subsequent experiments.

**Effect of Magnitude of Perturbation**

The next point we examined is the effect of the size of the perturbation on the behavior of the system. Figure 3 shows the effect of changing the perturbation size from $10^{-3}$ Å to $10^{-9}$ Å. The conditions of this experiment were as in the first experiment, and the perturbation was made to all of the protein atoms. Again, we see exponential growth with similar slopes for all the perturbations. Note that the plateau is reached later for smaller perturbations, although at a constant RMSD value, as expected.

It can be shown that the lag time before the start of exponential growth is related to the method of introducing the perturbations: The $10^{-9}$ Å perturbation shows the same initial lag as the $10^{-6}$ Å perturbation, yet the rate of deviation after the $10^{-9}$ Å perturbation trajectory is exponential in the $10^{-6}$ Å RMS range, i.e., if the point where the $10^{-9}$ Å perturbation reaches a $10^{-6}$ Å deviation is considered as a starting point for a subsequent simulation, no lag period is observed. One explanation of the initial lag is that after perturbation the initial forces introduced into the system tend to return it to its previous state. A second explanation is that the lag reflects the time required for the initial arbitrary perturbation direction to change to align with the direction of maximum deviation rate. We have not attempted to establish the relative importance of these two mechanisms.

**Measuring Deviation in Torsion Space**

The deviations were also calculated in terms of torsion RMSD rather than Cartesian RMSD. Figure 4 shows the behavior of the system after perturbations of various magnitudes. The same qualitative picture as for Cartesian RMSD (Fig. 3) emerges: an initial lag phase followed by exponential deviation at a rate of approximately $5 \text{ ps}^{-1}$. Note that the deviation flattens out at approximately $20^\circ$ RMSD, which is in the range of the radii of the allowed regions in a Ramachandran plot. We call this effect "cage rattling" and elaborate on it in Discussion.
Behavior at Different Stages of Folding

To investigate the rate of deviation at different stages of folding of the barnase fragment, the following perturbations were made: from the trajectory started at the experimental conformation, 12 points between 230 and 300 ps, which we consider the native-like state; 8 points between 430 and 500 ps, where a significant conformational change had occurred; and 8 points from the end of the simulation (920–990 ps), where the native structure has been partially reconstructed. Another 12 points were taken from the 600-ps simulation (every 40 ps between 40 and 480 ps) of the denatured conformation. As before, a deviation of 0.001 Å on all protein atoms was introduced. It is evident from Figure 5 that the rate of deviation is virtually the same at all stages of folding, approximately 5 ps⁻¹, and the duration of the exponential phase is also very similar in all cases.

Effects of Chaotic Behavior on Folding Events

Here we deal with the question of sensitivity of folding events to small perturbations. Can major folding events be avoided if the simulation is slightly perturbed just before they would otherwise occur?

To investigate this question the following experiments were performed. In the original trajectory there are two major transitions, both affecting the same residue: a change from α to β in the torsion angles of tyrosine 13 at approximately 472 ps and a back transition at approximately 921 ps. Having restart files every 10 ps, we can use the points of 470, 910, and 920 ps to introduce 1.5, 0.8, and 10.8 ps "upstream" perturbations. This scheme is shown in Figure 6.

For each restart, 10 different sets of 0.001 Å random perturbations of all protein atoms were made. The 0.8- and 1.5-ps upstream perturbations were followed by 25-ps MD simulations, and the 10.8-ps upstream perturbations were followed for 50 ps. The analysis focused on the robustness of the Y13 ψ transition, i.e., how many of the perturbed simulations show the original transitions within the analyzed period of time. If the perturbation was made very close (0.8 ps) to the original transition, all of the trajectories contained the transition (Fig. 7A). When the perturbations were made 1.5 ps before the original transition, the transitions still occurred, but their timing was spread over a period of 25 ps (Fig. 7B), i.e., there was a significant delay in the occurrence of the event. If the perturbations were made 10.8 ps before the transition, the change in behavior was dramatic. Only one of the 10 trajectories contained the transition at the end of the 50-ps observation period (Fig. 7C). No other alternative major folding events were observed during the perturbed simulations.

Numerical Stability

Observed chaotic behavior in a computer simulation can result from physical properties of the system or from numerical instabilities in the calculation. In
this section we explore the numerical stability of our observations to variation of a series of numerical parameters: CPU, compiler, floating point precision, cutoff radius of nonbonded interactions, force field, and atomic representation (all-atom/united-atom).

The original simulations were performed by using the DISCOVER CVFF force field \(^{21}\) with an all-atom representation and a cutoff radius for nonbonded interactions of 8.5 Å in a water box under periodic boundary conditions. Simulations were run on a Cray Y-MP supercomputer with an executable compiled with double precision for floating point variables (128-bit representation).

In a first experiment, a snapshot conformation from the original trajectory of the native-like simulation at 230 ps was used to set up a simulation with the GROMOS87 software and force field\(^ {22}\) on a SGI Indigo2 workstation using an executable compiled with single precision for floating point variables (32-bit representation), under otherwise identical conditions (solvated system, cutoff radius 8.5 Å, periodic boundary conditions). Instead of an all-atom representation the GROMOS force field uses a united-atom representation for nonpolar CH\(_2\) groups.

The deviation characteristics in the GROMOS system are indistinguishable from those of the DISCOVER/CVFF system (Fig. 8) establishing the independence of the observations from software and force field and from the computer system/CPU.

In a second experiment the influence of the nonbonded cutoff radius was investigated. For this purpose, a simplified system, omitting the solvent, was used. The GROMOS system on a SGI Indigo2 workstation was again used but now omitting solvent. The rates of deviation were compared for 8.5 and 100 Å (effectively infinite) and were found to be the same.

**DISCUSSION**

**Chaotic Behavior of MD Simulations**

The results clearly show that the behavior of a MD simulation of folding is chaotic, in the sense that

---

Fig. 7. Sensitivity of two conformational transitions to the upstream distance of a perturbation. Multiple random perturbations were introduced at different distances in time upstream from the original transitions in the ψ torsion angle of Y13. The arrows mark the time of perturbation, and the horizontal bars show the upstream distance from the transition (A: 0.8 ps; B: 1.5 ps; C: 10.8 ps; full lines: original trajectory; dotted lines: perturbed trajectories). For the perturbations close to the original transition point the event still occurs (A), although at different times (B). For the perturbation far from the original transition point the event is not observed (C).

Fig. 8. Demonstration of independence of sensitivity to perturbations from numerical factors. Deviation behavior of the solvated system under identical molecular dynamics parameters using the CVFF force field on a Cray Y-MP or C90 supercomputer with 128-bit floating point precision (dashed line) or the GROMOS force field on a SGI Indigo2 workstation with 32-bit floating point precision (full line). Due to heavy computing time requirements, perturbed trajectories on the workstation were run only for 2 ps.

The original simulations were performed by using the DISCOVER CVFF force field\(^ {21}\) with an all-atom representation and a cutoff radius for nonbonded interactions of 8.5 Å in a water box under periodic boundary conditions. Simulations were run on a Cray Y-MP supercomputer with an executable compiled with double precision for floating point variables (128-bit representation).

In a first experiment, a snapshot conformation from the original trajectory of the native-like simulation at 230 ps was used to set up a simulation with the GROMOS87 software and force field\(^ {22}\) on a SGI Indigo2 workstation using an executable compiled with single precision for floating point variables (32-bit representation), under otherwise identical conditions (solvated system, cutoff radius 8.5 Å, periodic boundary conditions). Instead of an all-atom representation the GROMOS force field uses a united-atom representation for nonpolar CH\(_2\) groups.

The deviation characteristics in the GROMOS system are indistinguishable from those of the DISCOVER/CVFF system (Fig. 8) establishing the independence of the observations from software and force field and from the computer system/CPU.

In a second experiment the influence of the nonbonded cutoff radius was investigated. For this purpose, a simplified system, omitting the solvent, was used. The GROMOS system on a SGI Indigo2 workstation was again used but now omitting solvent. The rates of deviation were compared for 8.5 and 100 Å (effectively infinite) and were found to be the same.
very small changes in initial conditions lead to exponential divergence of trajectories and to a different sequence of folding events. This finding will not come as a surprise to practitioners of MD. It has been a common frustrating experience that when a computer or a compiler has been slightly changed, trajectories of MD cannot be reproduced. The chaotic behavior of our system offers an explanation for these experiences.

There are two possible sources for the chaotic behavior. Either the MD description and/or the computational methods of solving them introduce artifacts into folding trajectories, or that indeed the actual folding behavior is chaotic. If the former were true, we would be forced to the strong conclusion that MD does not reflect the real nature of protein folding.

The tests we have made for independence of the conclusions from the numerical simulation process support though do not conclusively prove the latter conclusion: The deviation characteristics of a solvated peptide system using different force fields with different atom representations on different architectures with very different floating point precision are virtually indistinguishable.

In either case our observations suggest that individual MD trajectories of folding are too sensitive to small perturbations to have significant predictive quality. That is, the occurrence of a particular event at a certain stage in a trajectory does not mean that it would be found at that point, or indeed at all, in trajectories run under very slightly different conditions. Single trajectories may provide insight into the type of events possible in a system, but averaging over a number of independent runs is essential to begin to obtain data on event likelihood. It is not clear from the present results how extensive the variability of event behavior is and therefore how extensive a set of trajectories would be required.

Surprisingly, under a variety of conformational conditions virtually the same rate of local exponential divergence is found, approximately 5 ps⁻¹. Thus, it appears that the maximal finite time Lyapunov exponent of the system is a constant. In such a situation it is reasonable to assume⁴⁰ that the maximum global Lyapunov exponent of the system is, in fact, similar to the maximal finite time Lyapunov exponent. Although these results are for a protein fragment, the consistency under different conformational states suggests that this behavior would also be found for a complete protein.

**Origins and Extent of Chaos in Polypeptide Chain Motion**

The single most likely underlying cause for chaotic behavior is the interactions between the van der Waals walls of the atoms.²³ In a polypeptide chain, the atomic collisions occur primarily because of motion around the torsion angles (the ψ and φ angles of the backbone and the χ angles of the side chains). The range of motion around these angles is coupled, and there are multiple restrictions on each torsion angle, from both van der Waals collisions and attractive interactions, such as hydrogen bonds. Further, the set of relevant interactions and the extent to which they restrict motion change continuously. So, it is not surprising (some might say obvious) that chaotic behavior is found.

In the MD simulations of both the folded and unfolded states of the barnase peptide, we observe the same rate of exponential divergence of the conformations up to approximately 1 Å RMSD on all atoms, corresponding to approximately 20° RMSD in the torsion angles, independent of the environment sampled. This behavior raises the following questions.

First, why is the local rate of divergence so independent of the sampled environments? A possible explanation is as follows: We are measuring the microscopic behavior of individual atoms. At this level, the environment is similar, whether it consists mainly of other protein atoms as in the folded state, or mainly of water molecules as in the unfolded state.

Second, what determines the limit of exponential divergence? The finite size of the phase space always enforces a maximum difference between states. For our peptide, a random all-atom RMSD between conformations is approximately 5–6 Å. The exponential deviation does not last for nearly as long as this, so we are not seeing total phase space limitations. Apparently, on the time scale of these simulations, the system is trapped in a small volume of phase space. Two pieces of data support the view that this volume is essentially that represented by the local minima in the torsions angles. Analysis of both trajectories shows that events in which a torsion angle changes from one local minimum to another are rather rare (in the order of 10 per ns). Also, the loss of exponential behavior occurs after approximately 20° of divergence between trajectories, a scale commensurate with the torsion minima widths.

Third, what controls the rate of divergence, and how does this affect the likelihood that the course of folding events will be changed? The magnitude of the maximum Lyapunov exponent of a system is related to the time scale on which the system dynamics become unpredictable.¹⁵ In other words, it reflects the length of time after perturbation when the original and the perturbed trajectories are still quite similar. The deviation rate that we have measured is approximately 5 ps⁻¹. Indeed, this time period of the order of a picosecond coincides with the measured effect of chaos on folding events. As long as the perturbation preceded the events by approximately a picosecond (the 0.8- and 1.5-ps upstream perturbations) most of the folding events still occur, i.e., the behavior is predictable. But when the time of perturbation is an order of magnitude earlier (the 10.8-ps

---

**CHAOS IN PROTEIN DYNAMICS**

---

**423**
Implications for Folding Models

We now explore the implications of the chaotic behavior for protein folding models. Consider the following Gedanken experiment. Two identical polypeptide chains start to fold at the same time from the same initial conformation, except for one atom differing in position by 10^{-9} \text{Å}. If the sequence is that of a normal functional soluble protein, both chains will end up folded in the same native conformation. However, our results suggest that along the way the folding pathways will rapidly diverge and become significantly distinct.

Because the timing of individual folding events cannot be predicted, it is very hard to assume that the relative order of folding events along the folding pathway can be maintained. It would require a very specific design of an energy landscape to ensure that each folding event must be followed by a specific other folding event. Such a strong requirement does not seem realistic.

How can we reconcile the chaotic nature of the folding trajectory and the fact that all trajectories end at essentially the same folded functional conformation? This combination of properties can be found in dissipative systems and is known as transient chaos. Systems that exhibit transient chaos typically have long trajectories showing chaotic characteristics, but all converge to a set of points somewhere in the phase space. These points constitute an attractor for the system. In this sense, the folded state (which can be defined as the set of points in phase space that represent conformations with similar main-chain dihedral angles to those of the native conformation) is an attractor for the protein dynamics.

A useful analogy is a ball rolling down from the top of a rugged crater to its bottom. Because the bottom of the crater is an attractor of the system, it is clear that the ball will hit the bottom, but it can follow very different routes downhill. The trajectory of a single ball is chaotic, but we can still analyze the statistical properties of various downhill paths. Protein folding is similar, in the sense that the precise pathway of folding of each molecule is unpredictable, but the collective behavior can be analyzed probabilistically. In such a model each possible folding event has a probability of occurring determined by the current state of the system. For example, if all the n-torsion angles of the chain are in the alpha conformation, then torsion \( i \) has a probability \( P(s_i = \alpha \rightarrow \beta | s_{1<i} = \alpha) = \alpha \) of switching from \( \alpha \) to \( \beta \). Thus, if one starts with a nonnative conformation, it is possible to use these transition probabilities to construct the probability of an entire pathway leading from the nonnative conformation to the final native conformation. In this sense, we suggest that, although protein motion can be described by deterministic equations of motion, it is more informative to use this type of probabilistic Markovian model. A similar viewpoint has recently been suggested.\(^{25}\)

These results imply that the most useful role of MD in studies of protein folding is not that of a simulation tool for the whole folding process. Rather, it can be used to estimate the probability of transitions between states by running many short trajectories and in this way provides a quantitative tool for analyzing folding pathways.

ACKNOWLEDGMENTS

We thank William G. Hoover for helpful discussions.

REFERENCES

9. Wright, P., Dyson, H., Lerner, R. Conformation of peptide fragments of proteins in aqueous solution: Implications for...