Study of Pair Contact Formation among Hydrophobic Residues in a Model HP-36 Protein: Relationship between Contact Order Parameter and Rate of Folding and Collapse

Goundla Srinivas[†] and Biman Bagchi*

Solid State and Structural Chemistry Unit, Indian Institute of Science, Bangalore, India 560 012

Received: October 30, 2002; In Final Form: May 12, 2003

Recent analysis of the structures of a large number of proteins in their native state has demonstrated a close relation between relative contact order parameter (COP) (which gives the average contact distance among hydrophobic residues) and the rate of protein folding (Grantcharova et al. *Curr. Opin. Struct. Biol.* **2001**, *11*, 70). We have explored the existence of such a relationship by carrying out Brownian dynamics simulations of a model protein. The model consists of 36 amino acid residues and mimics a thermostable single domain headpiece of chicken villin (HP-36) protein. Long range interactions across the protein are obtained by using a simplified hydropathy scale. The heteropolymer exhibits qualitative features of folding and correlates with the COP. We have defined and calculated the distance dependent pair correlation function among the hydrophobic residues. Nativelike states are characterized by distinct pair contacts which are mostly absent in other collapsed non-native states. Contact pair formation shows a strong dependence on contour distance separation between the pairs. Study of the dynamics of specific contact pair formation during folding shows different characteristics depending on whether the final state is (near) native or far from the native configuration. Approach to the final collapsed state is almost always faster when the final state is nativelike. Fluctuations in pair contacts are found to be smaller in nativelike conformations compared to those of higher energy configurations. This is in agreement with a recent theoretical prediction.

1. Introduction

The protein folding problem is regarded as a problem of great importance in natural sciences. The folding of an extended, unfolded protein to its unique three-dimensional folded native state is a highly complex problem which has attracted a great deal of interest in recent years.[1−14] Despite numerous theoretical and experimental studies, a comprehensive understanding of many aspects of the protein folding is still lacking. However, recent protein folding motivated studies have addressed many stimulating problems of dynamics in polymers and strongly correlated systems.[5−7]

The protein folding problem has been studied and analyzed from many different angles, and different perspectives exist. It is now generally agreed that the folding proceeds not via a hierarchical and sequentially prearranged unique path but by exploring pathways in a funnel-like free energy surface. While the former picture predicts a very slow folding (the Levinthal paradox[1]), the latter approach can explain the fast rate of folding found in small single domain proteins. The initial approaches of Dill and co-workers[8,9] and of Bryngelson and Wolynes[10] were based on hydrophobic collapse of the heteropolymer and reordering of the residues to the native state. The resultant two order parameter free energy surface is successful in providing a qualitative picture of folding where the native state is the lowest free energy state which is separated from the extended unfolded state by a significant free energy barrier.[11] Subsequent studies have further refined this approach, and several quantitative studies have now been performed.[12,13] It has also been suggested that the native state is distinct from non-native states in having much less amplitude of vibrational motion around the average position than that in the non-native state.[6] This prediction is yet to be verified. There are several other aspects which need to be investigated. A recent study[14] investigated the relation between the rate of folding and the contact order parameter. This study established a highly interesting relation between the contact order parameter (which effectively gives the average distance between hydrophobic contacts) and the rate of protein folding. Clearly, understanding of this relation requires further theoretical and experimental study, particularly from the point of view of the dynamics of contact formation during folding. Recently Mukherjee and Bagchi[15] studied the correlation between rate of folding, energy landscape, and topology in the model HP-36 protein folding along these lines by introducing a specific force field. Their model could successfully capture multistage folding dynamics and the correlation of folding rate with topology.

The relative contact order parameter (COP) of a protein is defined by the following relation,[16]

$$\mathrm{COP} = \frac{1}{LN}\sum^{L}\Delta S_{i,j} \tag{1}$$

where $\Delta S_{i,j}$ is the contour distance between the residues (i, j) which form the contact in the native state and L is the total number of residues in the protein. It was found that, for a large number of proteins, the rate of protein folding is linearly related to the COP—the rate decreases as the COP increases. The faster rate of contact formation of the nearby residues has been

* To whom correspondence should be addressed. E-mail: bbagchi@sscu.iisc.ernet.in.

[†] Present address: Center for Molecular Modeling, Department of Chemistry, University of Pennsylvania, Philadelphia, PA 19104.

10.1021/jp022333j CCC: $25.00
Published on Web 10/02/2003

TABLE 1

nature of the interaction	$\epsilon_{i,j}$
hydrophobic−hydrophobic	2.0ϵ
weakly hydrophilic−weakly hydrophilic	0.3ϵ
strongly hydrophilic−strongly hydrophilic	0.3ϵ
hydrophobic−weakly hydrophilic	1.0ϵ
hydrophobic−strongly hydrophilic	0.8ϵ
strongly hydrophilic−weakly hydrophilic	0.3ϵ

discussed earlier.[8,9,17] However, the formation of long distance contacts is a complicated topological process, especially after the formation of other contacts. Clearly, the stiffness of the chain will also play an important role.

In a recent work, we have studied folding[18] and unfolding[19] of a thermostable chicken villin head piece, a 36-residue single domain protein (HP-36), through a minimalistic model by introducing a simplified hydropathy scale[20] to account for various interactions. It was found that this minimalist model reproduces several features of folding reliably. In particular, we could construct the folding funnel by calculating the energy, the number of states, and the number of hydrophobic contacts. In addition, it was found that the dynamics of approach to the final folded state was markedly oscillatory, both in total energy and in the radius of gyration. These oscillations agree with the rugged funnel landscape. Motivated by the success of this model, in this work we have studied both the static and the dynamic aspects of contact pair formation between the hydrophobic residues, both in equilibrium and during the folding.

As before,[18] we model the HP-36 protein as a necklace of different kinds of beads. Each bead in the sequence represents the corresponding amino acid in the protein sequence. There are 36 beads in the chain, since the number of residues in the original protein sequence (MLSDEDFKAV FGMTRSAFAN LPLWKQQNLK KEKGLF) is 36. All the beads are assumed here to be of the same mass and size.

As pointed out by Kauzmann[21] and also by Tanford[22] many years ago, one of the major driving forces of protein folding in aqueous media is the hydrophobic/hydrophilic nature of amino acids. This can be best represented by the hydropathy scale.[20,22] This scale arranges the standard free energies of transfer from aqueous solutions to pure liquid hydrocarbons and provides a measure of hydrophobicity, that is, the liking of a particular amino acid for water. Depending on the hydropathy values, we have categorized all the amino acids present in the HP-36 sequence into three classes: (i) hydrophobic, (ii) weakly hydrophilic, and (iii) strongly hydrophilic. The classification criterion remains the same as that in our earlier study.[19] If the hydropathy value is positive, the amino acid is hydrophobic. On the other hand, among the hydrophilic amino acids (the hydropathy value is negative) if the hydropathy value is smaller than -2.5, it is strongly hydrophilic; otherwise, it is weakly hydrophilic.

It should be stressed at this point that the transfer of the hydropathy scale to the intermolecular potential is to be understood only as a technique to incorporate the effects of interactions of amino acids with solvent (here water), and thus, the interactions between the amino acids themselves are to be regarded as "solvent averaged potentials". This can also be considered as a potential of mean force, well-known in liquid state theory.[23] Also, such a transfer of hydrophobicity to interatomic potential was perhaps first done by Dill and coworkers[24] in their lattice simulations. As can be seen from Table 1 (where we list the interaction strength parameters for all the six interactions), the interaction between two strongly hydrophilic groups is least favored because water will shield them while that between two hydrophobic groups is strongly attractive. Thus, these potentials are all water averaged potentials.

The nativelike states are characterized by distinct contact pairs which are mostly absent in other collapsed non-native states. It is found that contact pair formation indeed shows a strong dependence on the contour distance separation between the pairs. The study of dynamics of specific contact pair formation during folding shows different characteristics depending on whether the final state is (near) native or far from the nativelike configuration. The approach to the final state is almost always faster when the final state is nativelike. The fluctuations in pair contacts are found to be rather small in nativelike conformations compared to higher energy configurations. This is in agreement with a recent theoretical prediction.[5,6]

The rest of the paper is organized as follows. In the next section the simulation details are described. In section 3 the results on pair distributions (3.1), pair correlations (3.2), and also the dynamics of contacts both during folding and in the native state (3.3) are presented. This paper is closed with a brief discussion in section 4.

2. Simulation Details

The details of the simulation are similar to those of our earlier study.[18] Nevertheless, we briefly discuss the main aspects of the present simulations. The HP-36 protein is modeled as a necklace of three different kinds of beads; the six interaction energies are given in Table 1. The beads in HP-36 interact via a site−site Lennard-Jones potential. Neighboring beads are connected via harmonic springs. Each bead represents an amino acid in the actual protein sequence. The total potential energy of the chain can be written as

$$U = U_{\mathrm{b}} + U_{\mathrm{LJ}} + U_{\mathrm{s}} \tag{2}$$

where U_{b} represents the bonding potential,

$$U_{\mathrm{b}} = \sum_{i=2}^{N} \kappa(|\mathbf{r}_i - \mathbf{r}_{i-1}|)^2 \tag{3}$$

$\kappa = 9$ is chosen in this study. The interaction between nonbonded beads is represented by the Lennard-Jones like potential,

$$U_{\mathrm{LJ}}(r) = \epsilon_{i,j}\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{4}$$

where $\epsilon_{i,j}$ represents the interaction strength and σ is the LJ collision diameter. N is the number of beads in the chain, $\mathbf{r}_i$ is the position of bead i, and $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$. The stiffness is introduced through the bending potential U_{s},

$$U_{\mathrm{s}} = \mathcal{J}(\cos\theta - 1)^2 \tag{5}$$

where we set the chain stiffness $\mathcal{J} = 1$, that is a soft bending potential with minimum at $\theta = 0$ (i.e. straight chain configuration). The unit of time, τ, is b^2/D_0, where the single bead diffusion coefficient is denoted by D_0. The length is scaled by b, the bead diameter, as usual. We define $\epsilon^* = \epsilon/k_{\mathrm{B}}T$, where $k_{\mathrm{B}}T$ is the thermal energy.

The time evaluation of the model protein is done according to the following equation of motion,[25]

$$\mathbf{r}_j(t + \Delta t) = \mathbf{r}_j(t) + \frac{D_0}{k_{\mathrm{B}}T}F_j(t)\Delta t + \Delta X^{\mathrm{G}}(t) \tag{6}$$

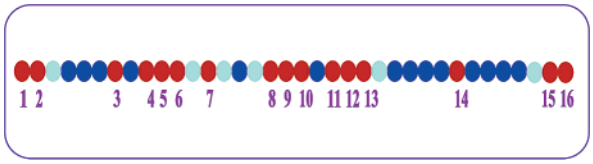

Figure 1. Schematic representation of modeling of the HP-36 protein by using the hydropathy values. The above shown color code of the hydropathy values of the simplified sequence is used in the present study. Only the hydrophobic residues (red) in the sequence are shown by the labels. Note that the pair contact formation among the *hydrophobic beads* is studied. The color code for the beads is as follows: red = hydrophobic, dark blue = strongly hydrophilic, and light blue = weakly hydrophilic.

where $\mathbf{r}_j(t)$ is the position of the j-th bead at time t and the systematic force on j is denoted by $F_j(t)$. The random Brownian displacement, $\Delta X^G(t)$, is taken from a Gaussian distribution with zero mean and $2\Delta t$ variance.

The course of the simulation is as follows. For each trajectory, an initial configuration is selected from the Monte Carlo generated equilibrium configurations at $\epsilon^* = 0.1$ (high temperature). The temperature of the initial configuration is then instantaneously reduced by 0.1ϵ, after 2.5×10^5 BD steps. Five such quenches, each with a gap of 2.5×10^5 steps, have been incorporated to facilitate the folding. Thus, the final $\epsilon^* = 0.6$. Further simulations for 2.5 million BD steps are carried out (after the quenching) to obtain the final configuration. Such a procedure is repeated for the model proteins with 600 different configurations. More details on the simulation scheme can be found in a similar study on homopolymers.[25,26]

In each simulation run, after choosing an initial configuration, the folding is followed till a stable final state is reached. It is important to note that in each simulation only one *protein* is simulated to obtain the corresponding final energy.

To obtain a list of final energies, such simulations are repeated for $\mathcal{N}$ number of independent *single proteins* sampled from an equilibrium distribution. In other words, we have carried out $\mathcal{N}$ different simulation runs with independent protein configurations (with the same sequence) to obtain $\mathcal{N}$ number of final energies. The results on dynamics are obtained by studying the folding of the model protein during the stepwise quenches. Although the transition from the initial extended to the final compact structure involves five discrete quenches, it is important to note that the total run time remains equal to the sum of the times taken for the five individual quenches.

3. Results

In Figure 1 a pictorial representation of the simplified sequence of HP-36 is shown. This code is used to construct the six interaction potentials required in the simulations. The color representation of the beads is as follows: hydrophobic, red; strongly hydrophilic, dark blue; weakly hydrophilic, light blue. Since we aimed to study the pair contact formation among the hydrophobic beads, only the hydrophobic beads (red) are labeled in the sequence. In the rest of the paper, the labels of beads refer only to the hydrophobic beads as shown in this figure. As mentioned earlier, the relevant interaction parameters are shown in Table 1. To study the properties of contact formation/breaking, the scalar distance between the hydrophobic residues is monitored as a function of time.

3.1. Characterization of the Nativelike State. Folding/collapse of 600 extended configurations have been carried out. Among these, only six (that is, 1%) of the configurations are found to have attained the nativelike conformation. The nativelike configurations are characterized by (a) the energy criterion and (b) the number of specific hydrophobic contact pair formations. The minimum energy criterion employed is to consider those configurations which have energy below a certain low energy cutoff. The cutoff itself is determined by the lowest energy state obtained in our simulations. Thus, the five other states selected as nativelike state are the ones close in energy to the lowest energy configuration. The second criterion is based on the number of hydrophobic contacts—the lowest energy native state is also found to have the largest number (20) of hydrophobic contacts. The nativelike states are chosen to have at least 15 contacts.

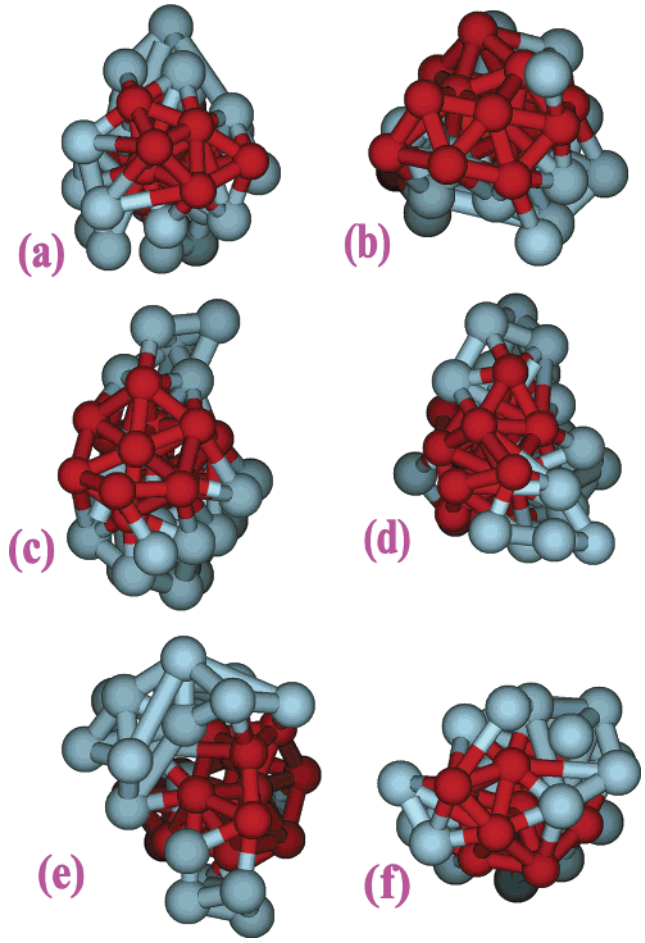

Figure 2. Snapshots of six nativelike final conformations of model HP-36 as observed in Brownian dynamics simulations. Note the formation of a specific contact pair in all the final configurations. The configurations from a to f are arranged in order of ascending energy.

In Figure 2, six nativelike configurations observed in the present simulations are shown. Note the hydrophobic core (red spheres) and the hydrophilic surface (blue spheres). Minimal energy configurations are characterized by the larger number of hydrophobic contacts (HC) compared to the case of the maxima, which are characterized by the smaller number of HC.

3.2. Equilibrium Distribution of Contacts. The assigned interaction energies among various amino acids are given in Table 1. After the temperature quenches, a fully extended initial configuration gradually folds into a compact rigid structure by forming the native contacts. We define the probability distribution that a specific (i, j) pair contact is separated by a distance R as

$$G_{i,j}(R) = \langle \delta(R - R_{ij}) \rangle \tag{7}$$

where R_{ij} is the distance between the hydrophobic residues i

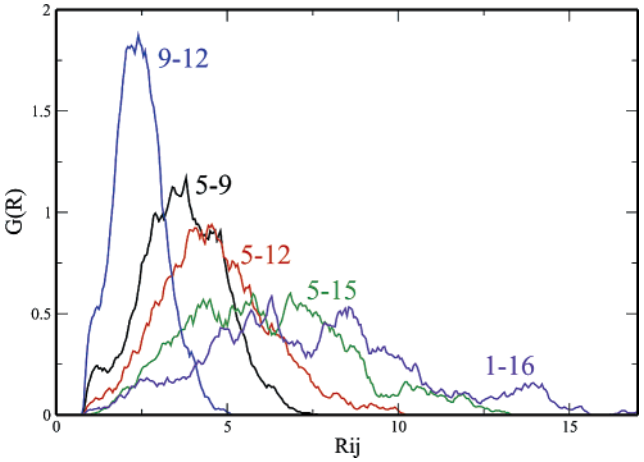

Figure 3. Pair distribution functions among various hydrophobic residues. Representation of all the curves is shown in the figure. On going from shorter separations to longer separations, the peak height decreases while the width increases.

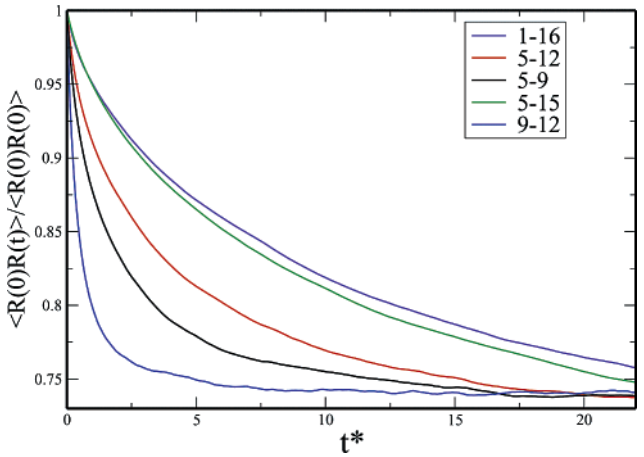

Figure 4. Pair correlation function $C_{i,j}(t)$ defined by eq 9 for various pairs of hydrophobic residues plotted against the reduced time. The representation of the curves is indicated in the figure. As can be seen, the pair with the smallest sequence separation (9−12) in the HP-36 folds faster (blue) while the pair with the longest separation (1−16) is the slowest (violet) to reach the native contact. Results for the rest of the pairs follow a similar trend.

and j in the final collapsed state. In the above equation averaging is defined by

$$\langle x \rangle = \frac{1}{\mathcal{N}} \sum_{n=1}^{\mathcal{N}} x^n \tag{8}$$

where $\mathcal{N}$ is total number of configurations, which includes the subsets of either nativelike configurations or configurations that are far from the native state or all the configurations.

In Figure 3 equilibrium pair distribution functions ($G_{ij}(R)$) for various pairs of hydrophobic residues are plotted. The notations of the pair distribution functions are shown in the figure. As can be seen from the figure, the pair distribution functions provide considerable insight into the pair dynamics. As shown in the figure, the largest peak corresponds to the contact pair 9−12 while the pair 1−16 shows the smallest peak. On going from the smaller sequence separation to the larger sequence separation, the height of the peak decreases while the width increases. From the observed peak heights, it is clear that most of the conformations have the 9−12 contact formed. On the other hand, the contact 1−16 has a smaller peak, spread over a wide range. This reflects the importance of the sequence separation in the contact formation during the folding. It is interesting to note that from the HP-36 protein sequence (shown in Figure 1), among the hydrophobic residues studied, 9 and 12 are separated by just 3 monomers, while the hydrophobic residues 1 and 16 are separated by 34 monomers. In other words, the pair 9−12 involves local interaction while the interaction between 1 and 16 is nonlocal. The rest of the peaks shown in the figure (5−9, 5−12, and 5−15) have peak heights proportional to the respective sequence separation in the HP-36 sequence.

3.3. Dynamics of Contact Pair Formation. The unfolded protein obviously folds by forming pair contacts between the specific beads. The dynamics of contact pair formation provide a better window to study the protein folding. However, these dynamics are much more involved. For the purpose of the present study, it is sufficient to monitor the distance between the specific pairs as a function of time. The study of the pair correlations reveals the information on the contact formation between the specific hydrophobic residues. To examine this point, we have studied the dynamics of each of these pairs. To this end, we define the contact pair time correlation function (CPCF) as

$$C_{i,j}(t) = \frac{\langle R_{i,j}(0) R_{i,j}(t) \rangle}{\langle R_{i,j}(0) R_{i,j}(0) \rangle} \tag{9}$$

The dynamics followed by studying the folding from time $t = 0$ to $t = \infty$, the time required for all the quenches. In Figure 4 the pair correlation $C_{i,j}(t)$ is shown as a function of reduced time for all the hydrophobic pair contacts that are shown in Figure 3. The representation of curves is shown in the figure. From this figure it is clear that the pair 9−12 (blue) is the fastest to attain its final position in the native conformation while the pair 1−16 is the slowest. The rest of the pairs (5−9, 5−12, and 5−15) show a similar trend to what has been observed in Figure 3. Close observation of pair 9−12 in the HP-36 sequence (shown in Figure 1) shows that these two hydrophobic beads are surrounded by hydrophobic beads on either side. Moreover, the hydrophobic blocks containing 9 and 12 are separated by just one hydrophilic bead. The fast contact formation of the 9−12 pair reveals the importance of the sequential arrangement of amino acid residues in the protein in determining folding rate. Note that for the smaller separations folding is faster compared to that for the longer ones. This reveals that the initial part of the folding is dominated by the local interactions, while the nonlocal interactions come into play in the later stages. This suggests that the protein folding involves two stages. First, the formation of the pair contacts between the residues due to the local interactions. Second, the pair contacts mediated by the nonlocal interactions. However, folding continues for a relatively longer time even after the formation of the (local) pair contacts. Thus, it is clear that the protein folding involves (1) the formation of the local contacts, which is a fast and perhaps the first step, and (2) the formation of nonlocal contacts, which is a relatively slower step. The rate-determining step in the protein folding is the formation of the nonlocal contacts. This is in accordance with the studies of Baker and co-workers[14] and Makarov et al.,[27] as discussed later.

In Figure 5 we have plotted the average $C_{i,j}(t)$ as a function of reduced time in two very different situations. The full line shows the result obtained by averaging over five nativelike configurations selected in the vicinity of the native state. In the same figure, the result obtained averaging over five configura-

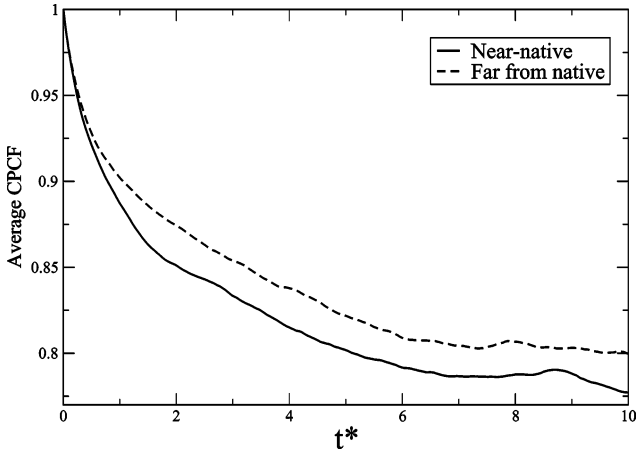

Figure 5. Comparison of the pair correlation function calculated by averaging over five configurations near the native state (full line) and that for five configurations far from the native state (dashed line). This figure shows that the native configuration results from the initial fast contact formation followed by slower rearrangement. On the other hand, energetically unfavorable configurations originate due to slow monotonic contacts, as in the case of collapse of simple homopolymers.

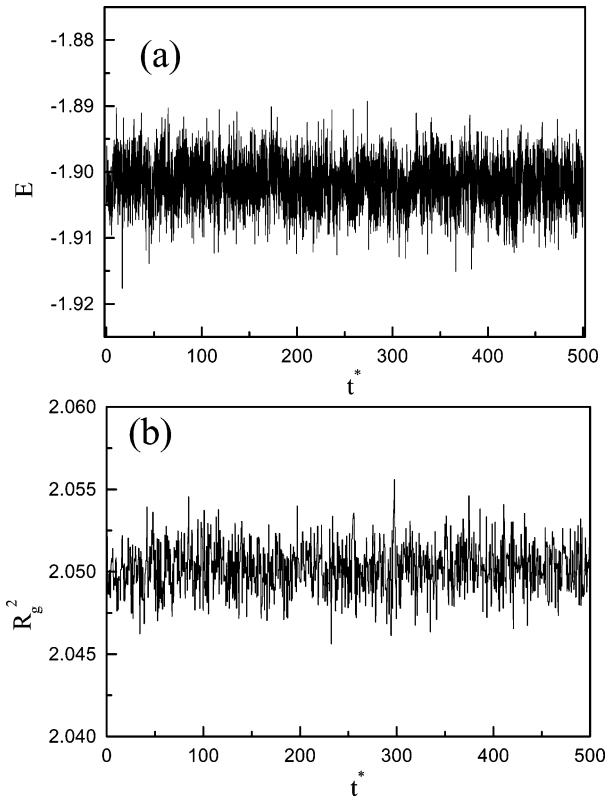

Figure 6. Fluctuations in (a) energy and (b) radius of gyration of the native state shown as a function of reduced time. The purpose of this figure is to show the stability of the native configuration obtained in this study.

tions that are far from the native state is shown by the dashed line. While the average $C_{i,j}(t)$ is found to show exponential-like decay, $C_{i,j}(t)$ in these two cases exhibits very different behavior. This comparison reveals that the formation of native-like configurations results from the fast formation of the initial contacts (as we have already described, these are the local contacts) followed by the slow (nonlocal) contacts, which results in a long time tail in $C_{i,j}(t)$. On the other hand, the folding of the protein into energetically unfavored configurations (far from the native state) results from the formation of the arbitrary pair contacts, resulting in a monotonic decay of $C_{i,j}(t)$, which resembles the collapse of a simple homopolymer.[19] Thus, the formation of the "correct" pair contacts plays a vital role in the protein folding, which is in accordance with the study of Zwanzig.[7]

3.4. Stability of the Native State. To study the stability of the native state in more detail, we have selected a native configuration from the resulting final configurations by following the minimum energy criterion. It is found that in the native state most of the pairs show a relatively faster approach toward the final state compared to the average trend. To study the stability of the native-configuration, we have carried out further simulations of the native state for the 5×10^6 steps. Figure 6a shows the energy as a function of the reduced time while that of the radius of gyration (R_g^2) is shown in Figure 6b. This figure reveals the stability of the nativelike conformation obtained in the presented study.

4. Conclusions

In this work, we have studied pair distributions and pair correlations between the hydrophobic residues both in equilibrium and during the folding of a model protein. The study of the pair correlations during the folding revealed that the folding of a protein involves two stages: faster local contact formation and the slower nonlocal contacts. The protein folding rate is mainly controlled by the latter contacts. This is further confirmed by comparing the $C_{i,j}(t)$ for the states near and far from the native state. This conclusion is in accordance with the findings of a recent study by Makarov et al.,[27] which states that the formation of all the native contacts is the rate-limiting step in the protein folding.

It is observed that the residues with a larger separation in the sequence require longer time to come closer and form a contact (i.e. a nonlocal contact). This implies that the fewer the nonlocal contacts, the greater the protein folding rate and vice versa. In other words, a larger distance between a contacting pair implies a greater physical distance in the sequence, which takes a longer time for the contact to form. This is in accordance with a recent study of Zhou and Zhou,[28] who introduced a total contact order parameter (based on similar concept) and found a significant correlation.

The minimalist model studied here could capture many of the essential features of protein folding. Future studies will explore the sensitivity to the potential employed and also generalizing the set of potentials to accommodate more realistic potentials. A recent study incorporated the sizes of amino acids that constitutes the HP-36 protein, which could capture the multistage folding dynamics and the correlation of folding rate with topology.[15]

The present study suggests that it might be possible to obtain qualitative information on folding mechanism and folding rates and also about the stability by modeling more complex proteins in a similar way.

Note Added in Proof. Subsequent to this work, a study of the dynamics of contact order formation has been carried out by Mukherjee and Bagchi by using a more detailed model of HP-36. The details are available elsewhere.[29]

Acknowledgment. The financial support from DST, India, is gratefully acknowledged. G.S. thanks CSIR for a research fellowship.

References and Notes

(1) Levinthal, C. In *Mossbauer Spectroscopy in Biological Systems*. Proceedings of a Meeting held at Allerton House, Monticello, IL; Debrunner,

P., Tsibris, J. C. M., Munck, E., Eds.; University of Illinois Press: Urbana, IL, 1969; p 22.
(2) Alm, E.; Baker, D. *Proc. Natl. Acad. Sci.* **1999**, *96*, 11305.
(3) Zhou, Y.; Hall, C. K.; Karplus, M. *Phys. Rev. Lett.* **1996**, *77*, 2822.
(4) Dinner, R. A.; Karplus, M. *Nat. Struct. Biol.* **1998**, *5*, 236.
(5) Portman, J. J.; Takada, S.; Wolynes, P. G. *Phys. Rev. Lett.* **1998**, *81*, 5237.
(6) Portman, J. J.; Takada, S.; Wolynes, P. G. *J. Chem. Phys.* **2001**, *114*, 5069; **2001**, *114*, 5082. Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 6170.
(7) Zwanzig, R. *Proc. Natl. Acad. Sci.* **1995**, *92*, 9801; **1997**, *94*, 148.
(8) Dill, K. *Biochemistry* **1990**, *29*, 7133. Chan, H. S.; Dill, K. A. *J. Chem. Phys.* **1994**, *100*, 9238. Dill, K. A.; Fiebig, K. M.; Chan, H. S. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1942.
(9) Honeycutt, J. D.; Thirumalai, D. *Biopolymers* **1992**, *32*, 695. Honeycutt, J. D. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 3526.
(10) Bryngelson J. D.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 7524. Bryngelson J. D.; Wolynes, P. G. *J. Phys. Chem.* **1989**, *93*, 6902.
(11) Leopold, P. E.; Montal, M.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 8721.
(12) Onuchic, J. N.; Wolynes, P. G.; Luthey-Schulten, Z.; Socci, N. D. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 3626.
(13) Socci, N. D.; Onuchic, J. N.; Wolynes, P. G. *J. Chem. Phys.* **1996**, *104*, 5860.
(14) Grantcharova, V.; Alm, E. J.; Baker, D.; Horwich, A. L. *Curr. Opin. Struct. Biol.* **2001**, *11*, 70.
(15) Mukherjee, A.; Bagchi, B. *J. Chem. Phys.* **2003**, *118*, 4733.
(16) Plaxco, K. W.; Simons, K. T.; Baker, D. *J. Mol. Biol.* **1998**, *277*, 985.
(17) Dill, K. A.; Chan, H. S. *Nat. Struct. Biol.* **1997**, *4*, 10.
(18) Srinivas, G.; Bagchi, B. *J. Chem. Phys.* **2002**, *116*, 8579.
(19) Srinivas G.; Bagchi, B. *Curr. Sci.* **2002**, *82*, 179.
(20) Stryer, L. *Biochemistry*, 4th ed.; W. H. Freeman and Company: New York, 1995.
(21) Kauzmann, W. *Adv. Protein Chem.* **1959**, *14*, 1.
(22) Tanford, C. In *The Hydrophobic Effect*, 2nd ed.; Wiley: New York, 1980.
(23) Hansen, J. P.; McDonald, I. R. *Theory Of Simple Liquids*; Academic Press: New York, 1986 and references therein.
(24) Dill, K. A. *Protein Sci.* **1999**, *8*, 1166.
(25) Srinivas, G.; Yethiraj, A.; Bagchi, B. *J. Phys. Chem. B* **2001**, *105*, 2475; *J. Chem. Phys.* **2001**, *114*, 9170.
(26) Srinivas, G.; Bagchi, B. cond-mat 0105138.
(27) Makarov, D. E.; Keller, C. A.; Plaxco, K. W.; Metiu, H. *Proc. Nat. Acad. Sci. U.S.A.* **2002**, *99*, 3535.
(28) Zhou, H.; Zhou, Y. *Bio. Phys. J.* **2002**, *82*, 458.
(29) Mukherjee, A.; Bagchi, B. *J. Chem. Phys.* Submitted for publication.