Folding Mechanism of β-Hairpins Studied by Replica Exchange Molecular Simulations

Jian Zhang,[1] Meng Qin,[1] and Wei Wang[1,2]*
[1]*National Laboratory of Solid State Microstructure and Department of Physics, Nanjing University, China*
[2]*Interdisciplinary Center of Theoretical Studies, Chinese Academy of Sciences, Beijing, China*

ABSTRACT **The folding process of trpzip2 β-hairpin is studied by the replica exchange molecular dynamics (REMD) and normal MD simulations, aiming to understand the folding mechanism of this unique small, stable, and fast folder, as well as to reveal the general principles in the folding of β-hairpins. According to our simulations, the TS ensemble is mainly characterized by a largely formed turn and the interaction between the inner pair of hydrophobic core residues. The folding is a zipping up of hydrogen bonds. However, the nascent turn has to be stabilized by the partially formed hydrophobic core to cross the TS. Thus our folding picture is in essence a blend of hydrogen bond-centric and hydrophobic core-centric mechanism. Our simulations provide a direct evidence for a very recent experiment (Du et al., Proc Natl Acad Sci USA 2004;101: 15915–15920), which suggests that the turn formation is the rate-limiting step for β-hairpin folding and the unfolding is mainly determined by the hydrophobic interactions. Besides, the relationship between hydrogen bond stabilities and their relative importance in folding are investigated. It is found that the hydrogen bonds with higher stabilities need not play more important roles in the folding process, and vice versa. Proteins 2006;62:672–685.**

Key words: β-hairpin; folding mechanism; trpzip2; replica exchange method

INTRODUCTION

Understanding the folding of key secondary structural elements such as β-hairpins can provide glimpses into the folding process of large proteins. This is because that (1) many features of folding kinetics in proteins can be found in β-hairpins,[1,2] and (2) the time scales of folding of these peptides are within microseconds, which are accessible to both fastest time-resolved experiments and to advanced simulation techniques.[3,4] Among various β-hairpins, the C-terminal β-hairpin of protein G B1 domain has been extensively studied due to its apparent two-state behavior and high stability in water. However, although many works have been done,[1,2,5–16] the folding mechanism of this β-hairpin remains elusive. Whether it folds by a hydrogen bond-centric zipper mechanism[1,2,5,6] or a hydrophobic core-centric mechanism[7–11] is still under debate. For example, Eaton and coworkers[1,2] suggested a zipper mechanism in which folding is initiated from the turn and propagates toward the tail by forming hydrogen bonds, the hydrophobic cluster forms relatively later although it also plays important role. In contrast, Karplus and coworkers[7] suggested that the hydrophobic assembly forms early, this nascent structure brings together the main chain in this region, from which point the sheet propagates outward. Pande et al.[8,9] also argued that the establishment of interactions between hydrophobic core residues is the central event and most probably the rate-limiting step. Lately, some works have shown that both the hydrogen bonds and hydrophobic interactions play important roles in the folding process.[12–16] Very recently, by comparing the folding and unfolding rates of trpzip series and protein G β-hairpin, it is suggested that the rate-limiting event in the β-hairpin folding process corresponds to the formation of the turn, whereas the hydrophobic cluster increases the stability of a β-hairpin by decreasing its unfolding rate.[17] Therefore, it is clear that there are many inconsistent pictures concerning the folding mechanism of β-hairpins.

Very recently, another kind of stable β-hairpins, tryptophan zippers, had been designed[21] and found to exhibit reversible and highly cooperative thermal unfolding transitions in water. These tryptophan zippers (trpzips) are very special because their stabilities are significantly higher than those reported for other small β-structures, including the C-terminal β-hairpin of protein G. On a per-residue basis, the trpzips have stabilities comparable to much larger protein domains ($\Delta G_{unf,\max}$ = 60 − 120 cal/mol × residue).[21] Their remarkable stabilities are attributed to the cross-strand pairs of indole rings. Because of their small size, unusual stability, and very favorable spectroscopic properties, trpzips should prove useful and quite simple systems to study β-structure and folding. Moreover, the trpzips are the smallest peptides to adopt unique tertiary fold without requiring metal binding, unusual amino acids, or disulfide crosslinks. There-

Grant sponsor: National Natural Science Foundation of China; Grant numbers: 10504012, 90403120, 10021001, 10474041; Grant sponsor: the Nonlinear Project (973) of the NSM.

*Correspondence to: Wei Wang, National laboratory of Solid State Microstructure and Department of Physics, Nanjing University, 210093, China; E-mail: wangwei@nju.edu.cn

Received 11 March 2005; Revised 14 July 2005; Accepted 16 August 2005

Published online 16 December 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20813

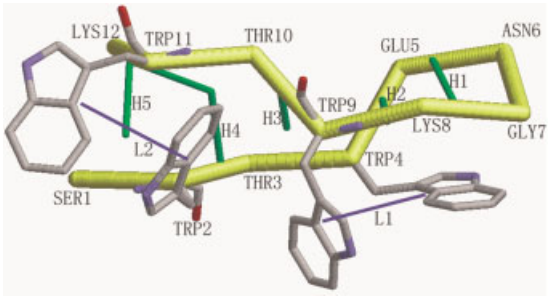

Fig. 1. The native structure of trpzip2 β-hairpin. The hairpin is strongly stabilized by two pairs of cross-strand Trp residues, i.e., Trp4 and Trp9, and Trp2 and Trp11. The five native backbone hydrogen bonds are also plotted. [Color figure can be viewed on the online issue, which is available at www.interscience.wiley.com.]

fore, it is of great interest to study the folding of these β-structures.

In this article, we choose trpzip2 (PDB code: 1LE1, SWTWENGKWTWK) as a model system, aiming to understand the folding mechanism of this fast folder, as well as to shed light on the general principles in the folding of β-hairpins. The structure of trpzip2 is shown in Figure 1. In our simulations, all-atom representation with explicit water model is used. To overcome the multiple-minima problem, or the quasi-ergodicity problem, the recently developed replica exchange MD method (REMD)[18] is used to sample the conformation space of trpzip2 β-hairpin. By using the REMD technique, the multiple-minima problem can be overcame since the MD runs at low temperatures have a probability to be exchanged to high-temperature regions. Moreover, the REMD simulations converge much faster than normal Monte Carlo runs since global conformation exchanges are introduced in the algorithm. By sampling a wide conformation space of trpzip2 β-hairpin, the folded, unfolded, and transition regions are located on the free-energy surfaces. Then the states in the transition region are investigated. These thermodynamics analyses are combined with the kinetics results obtained from the REMD trajectories and normal MD simulations, to provide an integral picture of the TS. It is found that the TS ensemble is mainly characterized by the formation of turn and establishment of hydrophobic interaction between the inner pair of core residues. This is a direct evidence for a very recent experiment,[17] which suggests that the turn formation is the rate-limiting step for β-hairpin folding. However, it should be noted that the stability of the nascent turn is strongly related to the partly formed hydrophobic core, which stabilizes the nascent turn by decreasing its "unfolding rate." Besides, the relationship between hydrogen bond stabilities and their relative importance in folding process is investigated. It is found that the hydrogen bonds with higher stabilities need not play more important roles in the folding process, and vice versa.

METHOD

The REMD method enables us to efficiently sample the conformation space of small proteins and is especially useful to study the temperature dependence of protein properties. It can be summarized as the following two-step algorithm.

(1) Each replica at fixed temperature is simulated simultaneously and independently for a certain number of MD steps.
(2) The neighbor replica pairs are exchanged with the acceptance probability,

$$P(x,x') = \min(1,\exp(-\Delta)),$$

where $\Delta = (\beta - \beta')[V(x') - V(x)]$, β and β′ are the two reciprocal temperatures, x and x' are the related conformations at β and β′, and $V(x)$ and $V(x')$ are potential energies at these two configurations, respectively. In our work, a parallel replica-exchange script is used to drive the MD simulations using the AMBER 7.0 package. The explicit water model is used under periodic boundary conditions. The 12-residue peptide is contained in a cubic box containing 2433 water molecules and 7515 atoms. A Cl^- ion is added to neutralize the charge of the system. The force field parm96 is used in our simulations since the parm94 and parm99 will exaggerate helix formation, it is more appropriate to use parm96 for the trpzip2 β-hairpin system.

REMD simulation is carried out at constant (N,T,V), with a water density of ~1.0 g/cm. The temperature coupling is set to 0.1 psec. Electrostatic interactions are calculated by using PME method with a cutoff of 8.0 Å, and the same cutoff is also used in the calculation of nonbonded list. We simulate this system with 62 replicas, with temperature ranging from 250–640 K. Starting from an energy-minimized native structure, a 2-nsec simulation is performed first at 300 K at constant pressure. Then the final configuration of this system is subjected to 2-nsec simulations at constant T in the temperature range indicated above. The final configurations for the 62 replicas are used as the starting structures of the REMD simulations. Replica exchanges are attempted every 1 psec.

A WHAM method[19] is used to combine data of all replicas below temperature T = 500 K to calculate the properties presented in this article. The reason we drop the data obtained at higher temperatures than T = 500 K is that the force field is normally parameterized at RT and some works have shown that the resulting potential function cannot reproduce some properties well at high temperatures.[13,20] This choice of temperature cutoff is somewhat arbitrary but has little influence on the results.

We calculated the potential of mean force of trpzip2 hairpin by projecting it onto seven reaction coordinates, including number of native backbone hydrogen bonds nHB, radius of gyration R_g, radius of the gyration of the hydrophobic core R_g^{core} (i.e., the four Trp residues), fraction of native contacts Q, RMSD with respect to the native states, the distance L1 between two CD2 atoms in the Trp pair Trp4–Trp9, and the similar distance L2 in Trp2–Trp11 pair (see Fig. 1). When calculating the radius of gyration, we only consider the Cα atoms. The two residues are assumed to form a native contact if the distance between any pair of heavy atoms of two residues is within

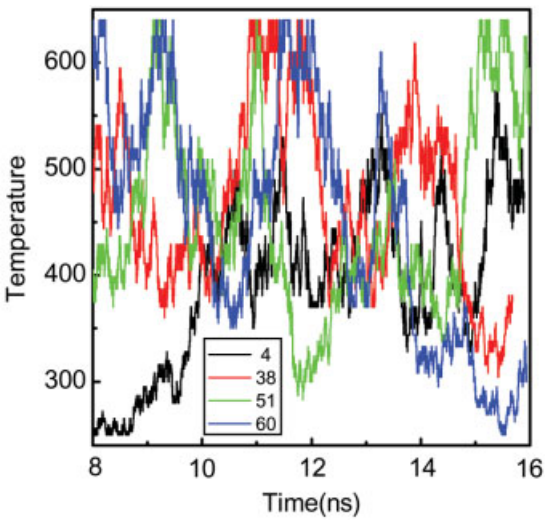

Fig. 2. The temperature evolution of four replicas, showing the random walk of each replica on the temperature space. [Color figure can be viewed on the online issue, which is available at www.interscience.wiley.com.]

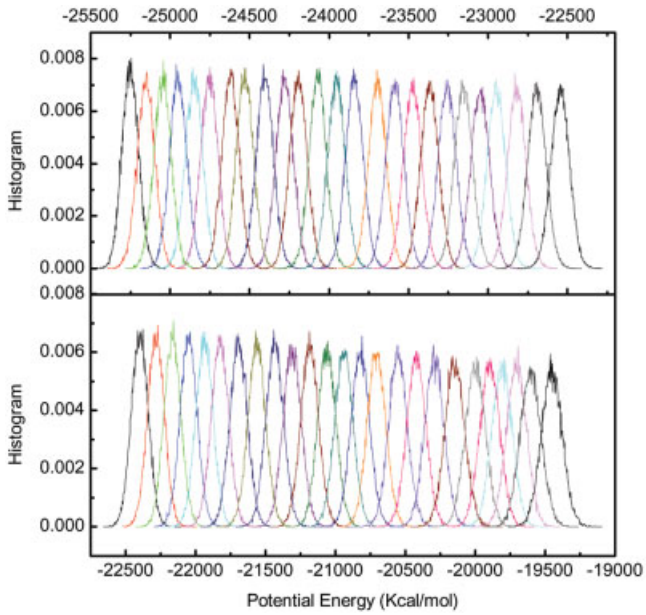

Fig. 3. The energy histogram of the first 50 replicas. Different curves correspond to different temperatures, ranging from 250 K–500 K. [Color figure can be viewed on the online issue, which is available at www.interscience.wiley.com.]

5 Å in the native structure. The total number of native contacts is 15 for trpzip2 hairpin. A hydrogen bond is counted if the distance between two heavy atoms (N and O in this case) is <3.5 Å and the angle N-H· · ·O is larger than 120°. There are five native backbone hydrogen bonds in the trpzip2 hairpin, numbered from turn to tail. For example, the H-bond 1 is between O (Glu5) and N (Lys8), the H-bond 5 is between O (Ser1) and N (Lys12). To characterize the structure of the basins of attractions identified from the PMF surface, we cluster the corresponding ensembles by using the algorithm proposed by Snow et al.[4] The conformations are clustered by their structural similarity, measured by RMSD. The related RMSD cutoff is chosen 2.0 Å.

RESULTS AND DISCUSSION

Replica Exchange Diagnostics

Figure 2 displays the temperature evolution of four replicas, demonstrating that the replicas sample a wide range of temperatures. The other replicas show similar behaviors. The energy histograms of replicas at each target temperature are shown in Figure 3. They significantly overlap with each other, ensuring a relatively high exchange rate, the average exchange rates between replicas range from 12%–25%. The Boltzmann distribution of the energy of each replica indicates that the equilibrium is achieved.

The convergence of REMD is also tested by monitoring the running average of the fraction of native contact, ⟨Q⟩. As shown in Figure 4, the curves represent ⟨Q⟩ averaged from 4 nsec to the time indicated by the legend. According to this diagram, it is safe to assume that the simulation has converged at about 8 nsec. In our REMD simulation, each replica is carried out a 16-nsec long run, the last 8 nsec of each replica is recorded for data collection. The total simulation time is about 1μsec.

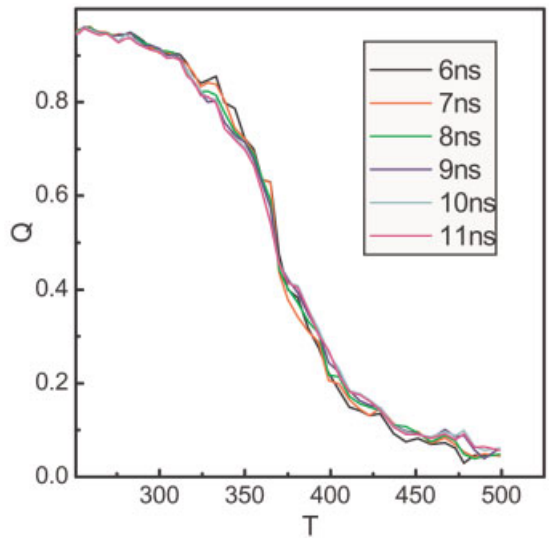

Fig. 4. The running average of the fraction of native contact, ⟨Q⟩, as a function of temperature. From top to bottom, the first curve corresponds to the ⟨Q⟩ averaged on the simulation data collected between 4–6 nsec, the second curve corresponds to that averaged on between 4–7 nsec, the third 4–8 nsec, and so on. The data before 4 nsec is thrown away. This figure shows that it is safe to assume that the calculation has converged at 8 nsec. [Color figure can be viewed on the online issue, which is available at www.interscience.wiley.com.]

Comparing with Experiments

Since the REMD simulation samples a wide range of phase space, we would like to calculate the thermodynamical parameters and compare them with experiments. Figure 5 shows the calculated unfolding free energy $\Delta G = G(U) - G(F)$ as a function of temperature, where the $G(U)$

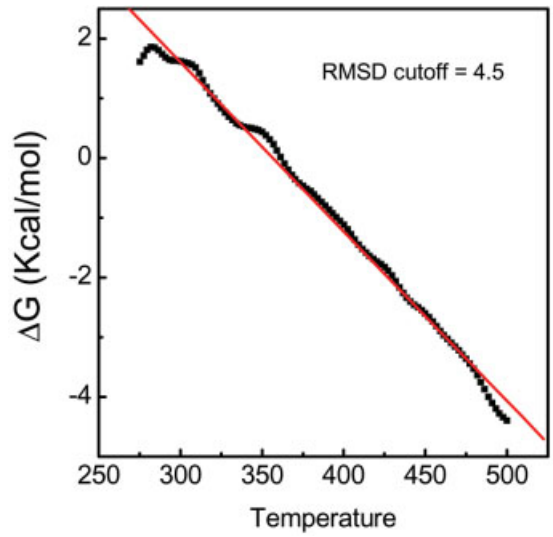

Fig. 5. The unfolding free energy $\Delta G = G(U) - G(F)$ (indicated by the square) at temperatures ranging from 273 to 500 K, where the G(U) and G(F) are the free energy of the denatured and native ensembles, respectively. The straight line is a linear fit of the ΔG. [Color figure can be viewed on the online issue, which is available at www.interscience.wiley.com.]

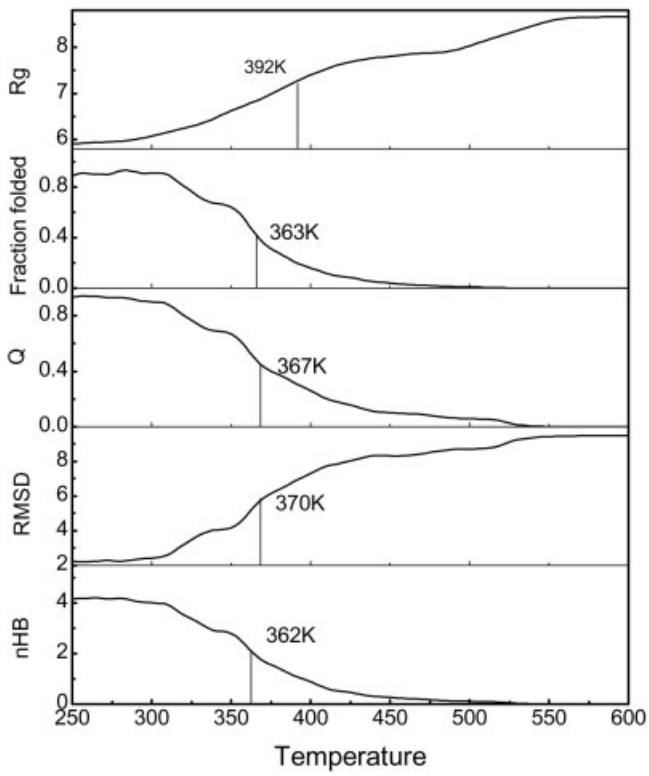

Fig. 6. The calculated radius of gyration R_g, fraction of folded Q, RMSD and number of hydrogen bonds as a function of temperature. The melting temperature T_ms are marked by the vertical lines. They correspond to the midpoints of the values of observables at very low and high temperatures.

and $G(F)$ are the free energy of the unfolded and native ensembles, respectively. To calculated the free energy, a state is assumed to be in native ensemble if its RMSD $<$ 4.5 Å, and is assumed to be in unfolded ensemble otherwise. The choice of RMSD cutoff is based on the PMF plot against RMSD discussed later in this article. This partition of unfolded and native states is rather rough; however, we find that the calculated thermodynamics is only weakly dependent on the RMSD cutoff. For example, the enthalpy ΔH_m varies $<$0.4 Kcal $\cdot$ mol^{-1} if the RMSD cutoff is decreased to 3.5 Å. This is because the population near RMSD cutoff is very small. Figure 5 shows a good linear dependence of ΔG on temperature, thus the unfolding enthalpy and entropy at the melting temperature can be obtained by a linear fit. We get ΔH_m = 10.123 Kcal $\cdot$ mol^{-1}, ΔS_m = *0.028* Kcal $\cdot$ mol^{-1}, and T_m = 361 K. Both ΔH_m and ΔS_m are smaller than that obtained by CD spectroscopy measurements, where ΔH_m = 16.77 Kcal $\cdot$ mol^{-1} and ΔS_m = 0.049Kcal $\cdot$ mol^{-1}, respectively.[21] The T_m obtained from our simulation is slightly larger than the experimental value, 361 K versus 345 K. Although there are discrepancies, we feel that the simulated thermodynamics is roughly in agreement with the experiments.

The discrepancies are attributed to the heterogeneity in the folding of trzip2 hairpin found in both experiments and simulations. By monitoring 12 different experimental and computational observables, it was found that the value of melting temperature T_m is dependent on the observables.[25] Then it is reasonable to image that other thermodynamics such as unfolding enthalpy or entropy will be also dependent on the chosen observables. This heterogeneity will inevitably cause difficulties when comparing computational results with experiments, and may contribute to most of the discrepancies discussed above.

We also find this heterogeneity in our simulations with explicit water model, as shown by Figure 6. It can be seen that when monitored by the fraction of folded (those with RMSD $<$ 4.5 Å), Q, RMSD, and number of formed hydrogen bonds, the calculated T_m agree roughly with each other and with the value calculated from ΔG, spanning a narrow temperature range from 362 K–370 K. They form the first cluster of melting transition. However, when monitored by the radius of gyration, the T_m is found to be about 392 K, deviating far from the first cluster of melting transition. This heterogeneity is similar to that observed in the previous work,[25] where by carrying out REM simulations with implicit water model, it was found there are three clusters of melting transition, ranging from 15°C–160°C. Remarkably, in that work, the T_m obtained by monitoring the radius of gyration is much higher, about 433 K. This is even higher than the corresponding value we obtained. We tentatively attribute this discrepancy to the different water model used in two simulations.

The simulations with explicit, implicit water model and experiments all demonstrate the heterogeneities in the folding of trpzip2 hairpin, that is, different observables give different thermodynamics. This indicates a rough free-energy landscape of this hairpin or a low overall free-energy barrier for folding, and suggests it may be a candidate for downhill folder under some optimal condi-

tions. However, the discussion of this topic is beyond the scope of this article.

The H1 and H2 States

In the following sections, we investigate the structure of the basins of attractions identified on the PMF surface. Figure 7(a–d) shows the free-energy contour plots at 300 K as a function of nHB, R_g^{core}, Q, L1, and L2. According to Figure 7(a), there are six states or basins of attractions: the native state (N), the unfolded state (U), a compact state H (including H1 and H2), a partially folded state P, and a possibly misfolded state M. Here, we define these states according to their positions on Figure 7(a). That is, the H1 state is composed of these conformations with their number of native backbone hydrogen nHB $\leq$ 1 and radius of gyration of the hydrophobic core $6.0 < R_g^{core} < 7.5$; the H2 state, nHB $\leq$ 1 and $5.4 < R_g^{core} < 6.0$; the P state, nHB = 3 and $5.2 < R_g^{core} < 5.5$; the M state, nHB $<$ 1 and $R_g^{core} < 5.0$; the N state, nHB = 5 and $5.3 < R_g^{core} < 5.5$. The conformations in each basin of attraction are clustered by using the algorithm proposed by Snow et al.[4] Then the distribution of nHB, R_g, R_g^{core}, Q, RMSD, L1, and L2 are calculated for each cluster. These are used to determine the position of states such as H and P on other free-energy contour plots than Figure 7(a). The positions of these states thus obtained are plotted on Figure 7(b–d).

According to Figure 7(a,b), during trpzip2 folding, the peptide first collapses to a H1 state. (Note here the ordering of folding events is based on the progression along reaction coordinates, not a progression in time.) The radius of gyration of the hydrophobic core R_g^{core} of this state ranges from 6–7.5 Å and no stable native backbone hydrogen bonds are formed. Besides the H1 state, an H2 state can be observed. Its R_g^{core} is ~5.6 Å, close to the value of the native state (~5.36 Å). The corresponding fraction of native contacts Q changes from $Q < 0.1$ for H1 state to $Q \sim 0.2$ for H2 state. The transition from H1 to H2 state is very smooth, with the barrier $<1k_BT$ [Fig.7(b)]. The free energy of the H2 state is slightly higher than that of H1 state, thus H2 has a smaller population. The H2 state might be viewed as a short-lived intermediate state lying on the pathway from the H1 state to transition state. The observation that the peptide quickly collapses to a compact H state (including H1 and H2) at the very early folding stage demonstrates that the essential driving force initiating the folding process is the hydrophobic interactions. The hydrogen bond is not the driving force at the very early folding stage since no stable native backbone hydrogen bonds are observed.

The trpzip2 hairpin is very special because of its high stability compared to most other β-hairpins. Its per residue thermodynamic parameters are comparable to those of large protein domains,[21] for example, the unfolding activation enthalpy is about 50 or 74 KJ/mol.[4] The extremely high stability is attributed to the packing of two cross-strand pairs of indole rings (the first or inner pair is between Trp4 and Trp9 and the second or outer pair is between Trp2 and Trp11),[21] which construct the hydrophobic core of trpzip2. Therefore, it is crucial to investigate the packing process of two pairs of indole rings to understand the folding and stability of this β-hairpin. This process can be described by the reaction coordinates L1 and L2 indicated in Figure 1. The free energy plots as a function of nHB and L1, L2 are shown in Figure 7(c,d), respectively. To characterize the H1 and H2 state more clearly, the distributions of L1 and L2 in H1 and H2 states are calculated and shown in the insets of the Figure 7(c,d), respectively.

Clearly, in H1 state, the distributions of L1 and L2 are rather broad, indicating that there are many kinds of packing patterns of core residues in the H1 state. For the H2 state, however, there is a suddenly shrink of the L1 distribution, indicating the packing of the inner pair of indole rings. It is highly possible that this is related to the searching process for the TS (one important character of the TS is the packing of the inner pair of indole rings, more about this later). However, although the distribution of L1 in H2 state concentrates around the native value, the packing pattern of the inner pair of indole rings is not correct (i.e., native-like) yet.

In order to make clear the structures of the H1 and H2 states further, we classify the conformations belonging to the H1 and H2 states by using the cluster algorithm. It is found that most conformations in the H1 ensemble can be clustered into four clusters, their relative populations are 60%, 20%, 8%, and 3%, respectively. The structures of these clusters are shown in Figure 8. It can be seen that the packing of the hydrogen core is apparent, but the packing pattern varies between clusters. The structure of the H1 state indicates that it is a very early intermediate state along the folding pathway; the packed hydrophobic clusters have to be readjusted in the later folding process. The H2 ensemble is relatively homogeneous and most (>86%) conformations are within one cluster. The incorrect packing pattern of the inner pair of indole rings is also obvious in H2 ensemble. It is interesting to see that there is a short helix formed from residue 2–7. Since the H2 state is not heavily populated, this semi-helical state is short-lived and may not be observed in experiments.

The P State

There is a P state beside the native state N with its R_g^{core} almost as same as that of the native state and with three out of five hydrogen bonds formed, as shown in Figure 7. This P state is reminiscent to the P state in protein G β-hairpin,[13] which also has its R_g^{core} almost equal to that of the native state and two to three hydrogen bonds of five total. Besides, the P states observed in two peptides are both separated from the native states by a low barrier. By clustering all the conformations within the P basin of attraction, it is found that they can be clustered into two main clusters, with the first cluster taking account of about 64% populations and the second about 30%. The structures of two clusters are shown in Figure 9(a,b), respectively. Generally speaking, the structures of both clusters are native-like except the low formation probabilities of the first two hydrogen bonds, which are related to the incorrect ψ value of residue Glu5 and ϕ of residue

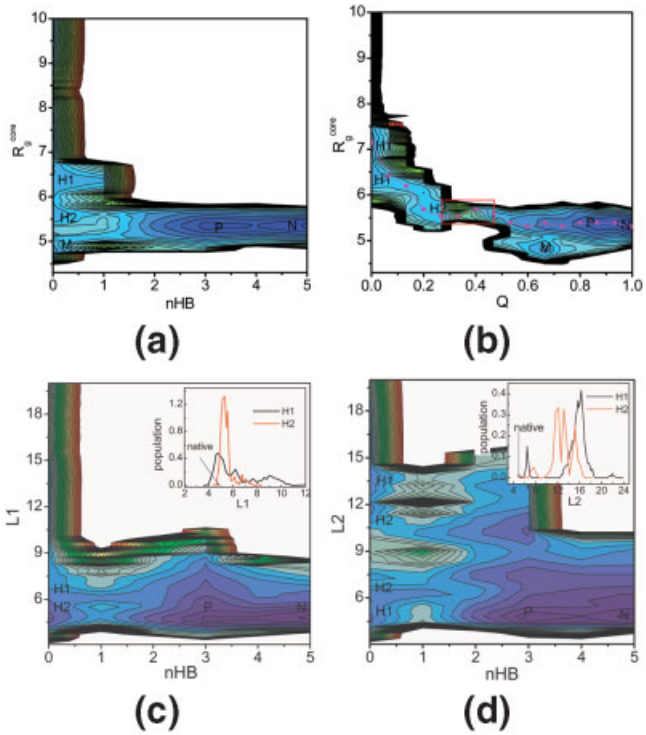

Fig. 7. The contour plots of the free energy at 300 K projected onto the reaction coordinates nHB, R_g^{core}, Q, L1, and L2. The basins of attractions H1, H2, P, M, and N are marked in the figures. The energy separation between neighbor contour lines is $1k_BT$. The insets in (**c**,**d**) show the distributions of L1 and L2 in H1 and H2 states, respectively. The values of L1 and L2 of the native state are marked by blue arrows.

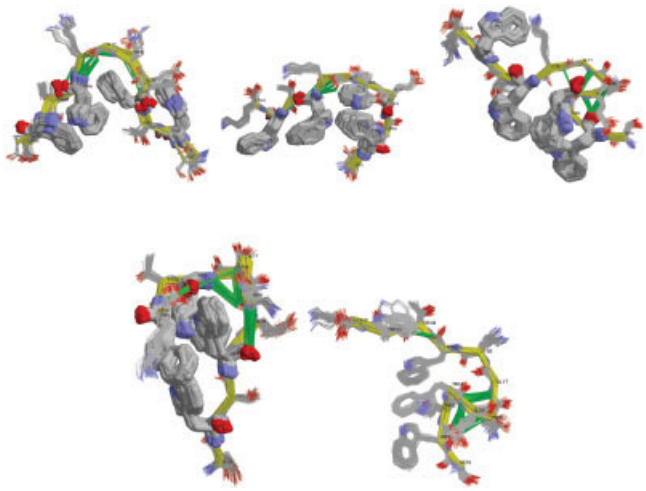

Fig. 8. The representative structures of the H1 state (the first four structures) and the H2 state (the last structure). The four Trp residues are plotted by sticks, the hydrogen bonds are colored green, and the backbone is colored yellow. All the structures in this article are plotted by the RasMol software.

Asn6. These two structures are slightly different in the packing pattern of the first pair of indole rings. The packing pattern of the first cluster is native-like, namely, with the CG-CD2-CE3 edge of the indole ring of Trp4 pointing toward the face of indole ring of Trp9. The

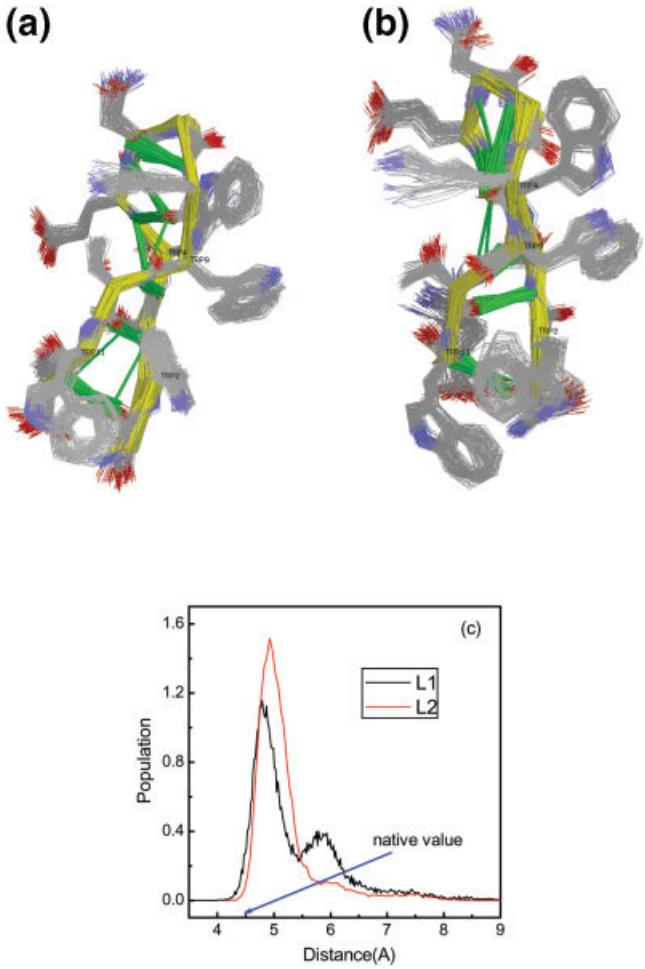

Figure 9.

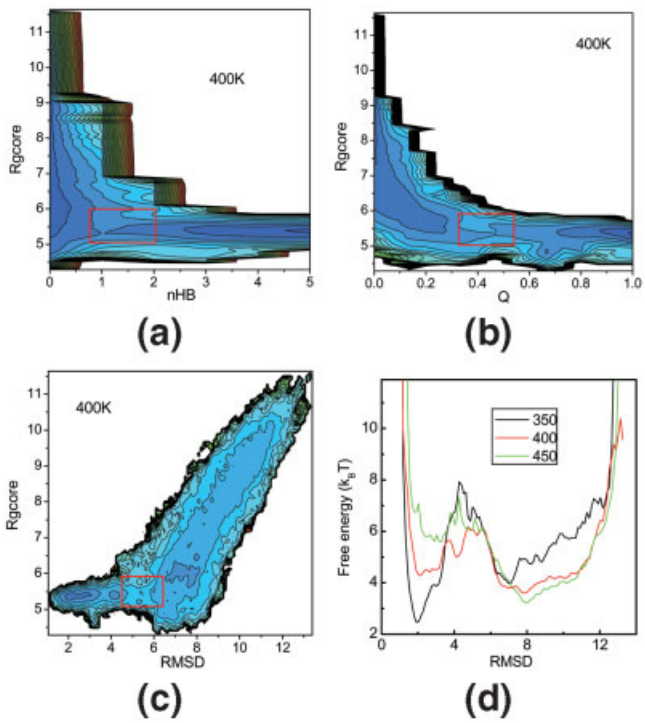

Figure 10.

packing pattern in the second cluster, however, is with the CG-CD1-NE1 edge of Trp4 toward the ring face of the Trp9. These two different packing patterns are also supported by the distribution of L1 and L2 in P state shown in Figure 9(c). It can be seen that the distribution of L2 is unimodal, demonstrating that only one major packing pattern exists for the second pair of indole rings. For the L1 distribution, however, there are two peaks. They are related to the two packing patterns indicated above. The energy barrier separating P state from the native state is within the range of thermal fluctuation ($<1k_BT$, see Fig. 7) and will disappear as temperature increases to 350 K. Therefore, the P state must be interconverting rapidly with the native state and can be assumed belong to the native basin of attraction.

The Transition State and Folding Mechanism

Kinetic analysis

Based on the analysis of the H1 and H2 states, it is concluded that at the very early stage, the folding is initiated by the hydrophobic collapse. The following events will take place within the compact ensemble (H1 and H2). Now it comes to the question what the next folding event is, the formation of turn or the establishment of the hydrophobic core, or something else? To answer such a question, the structure of transition state (TS) ensemble is investigated in this section.

It is well known that the REMD samples a wide conformation space at the cost of eliminating the information on dynamics. The trajectory at a fixed temperature is discontinuous due to exchange of conformations between neighboring temperatures. However, if we follow the trajectory into a new temperature region when the exchange occurs, we get a continuous "trajectory." Such a trajectory is different from the normal one since the related temperature is not a constant but varies when the exchange occurs. However, from such kinds of trajectories, we can extract the information on TS if the transition between states occurs at not too high temperatures (for example, <500 K). Previously, this kind of trajectory has been used by García et al. to study the transition of Met-Enkephalin.[26]

There are many techniques to identify the TS from MD trajectories. In this paper, we employ the method invented by Dokholyan and Shakhnovich and colleagues to identify the folding nucleus.[27] In detail, by monitoring the REMD trajectories mentioned above, we search for these processes that originate from the folded region and enter the transition region, but eventually return back to the folded region without descending to the unfolded region. This is called an "FF process" by Dokholyan et al. For these FF processes, the states in the transition region contain ample information on the folding nucleus, because it is the nucleus that guarantees the returning back to the folded region of the trajectories. These states in the transition region are the transition states. This is because physically, the folding nucleus is defined as the minimal number of contacts, which once formed will drive the molecules into folded region; without these critical contacts, the molecules will descend to the unfolded region, thus the complete nucleus must be formed at the top of the free-energy landscape and is the subset of the contacts of the TS. By collecting the states in the transition region during FF process, we can characterize the TS. It should be noted that in the original method used by Dokholyan et al., the non-nucleus contacts in the TS have to be figured out by taking account of the UU processes simultaneously. Here this procedure is not needed since we are only interested in the TS.

To implement the method described above, we have to first define the folded, unfolded, and transition region. Figure 10(a,c) gives the free-energy plots projected on several reaction coordinates at 400 K. According to the projected free-energy surfaces, we define the folded region at $Q > 0.9$ and RMSD < 3.5, the transition region between $0.32 < Q < 0.55$, and the unfolded region at $4.5 <$ RMSD < 6.5, RMSD > 6.5. The RMSD ranges of the folded and transition region are insensitive to the temperature ranging from 350 K–450 K, as shown by Figure 10(d). The Q parameter is more insensitive to the temperature. In our REMD simulations, many transitions occur in the temperature range between 350 K–450 K.

Figure 11 shows an example of FF process found in the REMD trajectories. Note that the trajectory is continuous and the temperature is not a constant. It originates from the folded region, then enters the transition region and finally returns back to the folded region again. The related temperatures during transition range from 400 K–450 K. Now it is interesting to check the time evolution of various parameters. As for the hydrogen bonds, when the trajectory runs from the folded to transition region, the terminal hydrogen bond 5 breaks first, then the hydrogen bond 4. At almost the same time, the outer pair of Trp residues drifts from each other (indicated by the reaction coordinate L2). After a short time of staying at the transition region, the molecule returns back to folded region eventually. This is because the turn region is kept intact and the hydrophobic core is still partly formed, which is sufficient to drive the molecule back to the folded region.

Since in the FF process the farther the molecule is from the folded region, the closer it is to the top of the barrier, we pick out the 30 structures with largest RMSD from the transition region, then classify these structures by the clustering algorithm. The largest cluster with 21 structures is shown in Figure 12. Basically, the turn as well as the first three hydrogen bonds keeps intact, and the fist pair of Trp residues packs together whereas the second pair departs from each other. This character of the TS

Fig. 9. The structures of two main clusters of the P states, shown in (**a**) and (**b**), respectively. The four Trp residues are plotted by sticks, the hydrogen bonds are colored green, and the backbone is colored yellow. The two clusters correspond to two different packing patterns of the inner pair of indole rings, respectively. These two different patterns are illustrated by the bimodal character of the distribution of L1, as shown in (**c**).

Fig. 10. (**a**–**c**) The free energy contour plots calculated at 400 K, the transition regions are marked by the rectangles. (**d**) The free energy as a function of RMSD, showing that the RMSD ranges of the folded and transition region are insensitive to the temperature ranging from 350 K–450 K.

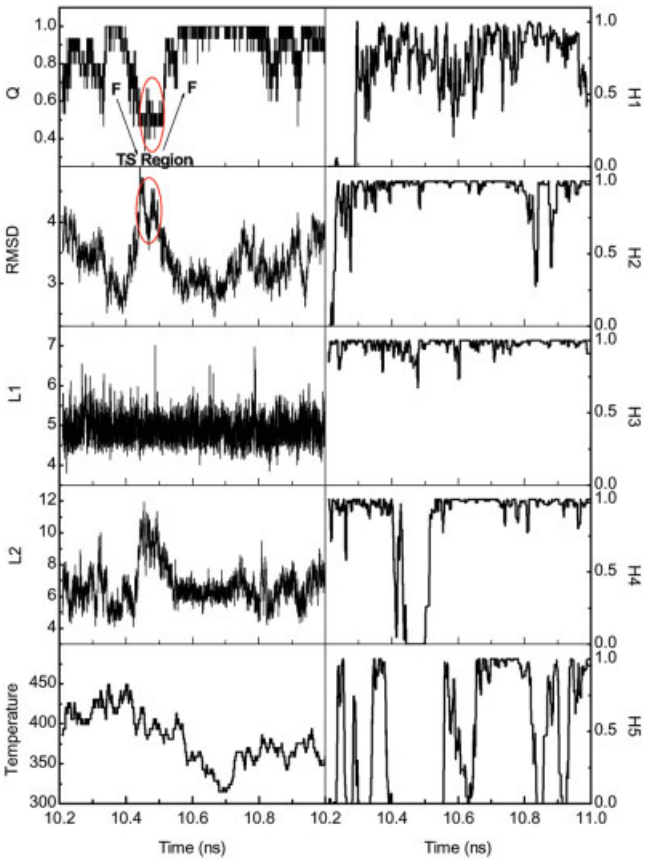

Fig. 11. A REMD trajectory involving FF process. Different figures correspond to different simulation observables. H1–H5 are the formation probability of each hydrogen bond, respectively. The left bottom figure shows how temperature varies in this REMD trajectory. [Color figure can be viewed on the online issue, which is available at www.interscience.wiley.com.]

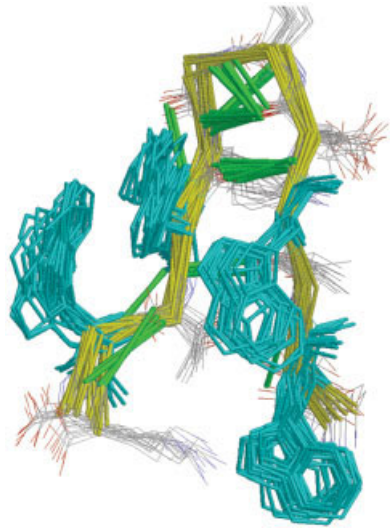

Fig. 12. The largest cluster of states taken from the transition region in the FF process, showing the typical character of the TS. [Color figure can be viewed on the online issue, which is available at www.interscience.wiley.com.]

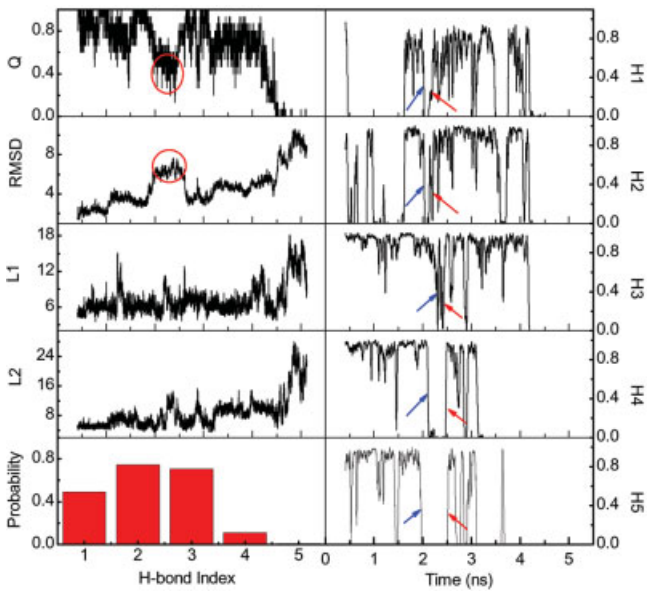

Fig. 13. A trajectory of normal MD simulations, carried out at 500 K. Different figures correspond to different MD observables. The left bottom figure shows the average probability of each hydrogen bond for the states in the transition region. [Color figure can be viewed on the online issue, which is available at www.interscience.wiley.com.]

illustrates the importance of turn as well as the interaction between first pair of hydrophobic core residues in the folding process.

We repeat the above algorithm to search for all possible TSs in the 62 trajectories in our REMD simulation, totally we find about 1200 structures. Most of them resemble the structures shown in Figure 12, except the flip of indole rings of the first pair of Trp residues and the great fluctuation of the second pair of Trp residues. This confirms the characteristic of TS shown in Figure 12. Besides, we do find exceptions. Some structures have their inner hydrogen bonds and interaction between inner pair of hydrophobic core residues broken, whereas the outer hydrogen bonds and hydrophobic interaction are kept intact. These exceptions indicate a mechanism that unfolds hairpins from the inner part and folds them from the terminal.

To confirm the previous observations obtained from the REMD trajectories, we also carry out 20 normal MD simulations at 500 K, each last for several nanoseconds. When doing the simulations, the exactly same model of protein and water is used. The temperature is chosen at 500 K to accelerate the sampling process. Then we repeat the above algorithm to search for the TS. We also find several FF processes in the normal MD trajectories; one of them is shown in Figure 13. This FF process is similar to that found in REMD trajectories. The molecule enters the transition region at about 2 nsec and stays there for about 500 psec, and then goes back to the folded region. It is interesting to notice that during the unfolding process (from the folded region to the transition region); the central hydrogen bond 3 is the last one to break, as may relate to its closeness to the hydrophobic core. This obser-

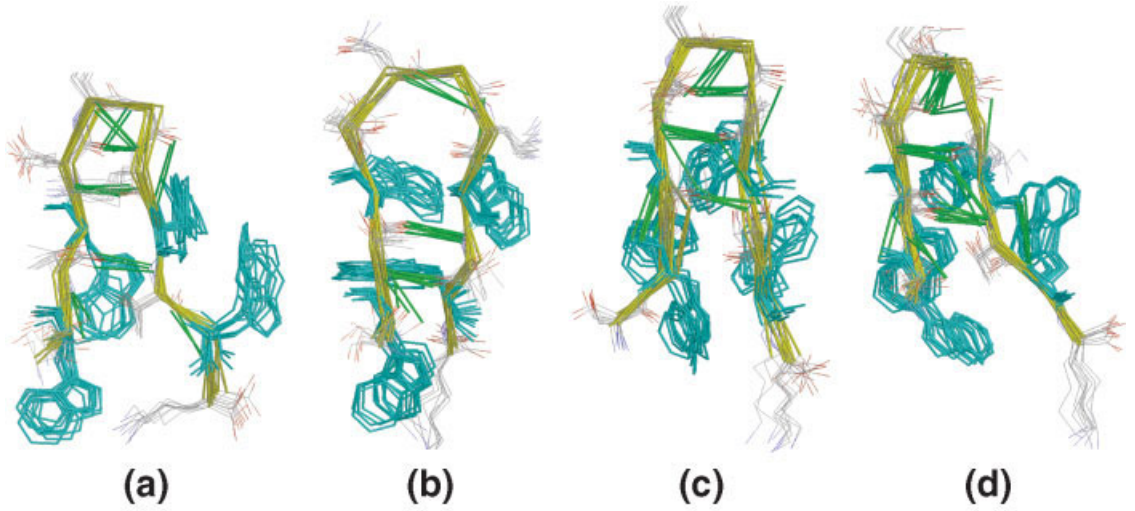

Fig. 14. The structures of the largest four clusters of states, showing the major characters of the TS. [Color figure can be viewed on the online issue, which is available at www.interscience.wiley.com.]

vation indicates that the unfolding rate is mainly determined by the hydrophobic cluster. Remarkably, during the refolding process (from the transition region back to folded region), the hydrogen bonds 1 and 2 form first, then hydrogen bond 3, followed by the outer two hydrogen bonds, 4 and 5. This clearly demonstrates a zipper folding mechanism. The hydrogen bond probabilities averaged on the states in the transition region are given in Figure 13. Statistically, the inner three hydrogen bonds have high formation probabilities compared to the outer two, demonstrating the turn formation in the TS.

The states in the transition region are also subjected to the cluster procedure, the structures of the largest four clusters are given in Figure 14. It can be seen that most clusters resemble the structures shown in Figure 12 except the second one. It confirms the character of TS discussed above again. The states in the second cluster, unlike the others, do not have a formed turn. The two inner hydrogen bonds are absent, whereas the hydrogen bond 3 and 4 are kept intact. These structures are related to the unfolding process indicated above (from folded region to transition region), during which the hydrogen bond closer to the hydrophobic core is the last one to break. Consistent with the unfolding order of hydrogen bonds, the structures in the second cluster suggest the unfolding process is controlled by the hydrophobic interactions.

Very recently, by comparing the folding and unfolding rates of trpzip series and protein G β-hairpin, it is suggested that the rate-limiting event in the β-hairpin folding process corresponds to the formation of the turn, whereas the unfolding rate is mainly determined by the hydrophobic cluster.[17] Here our simulations provide a direct evidence for this declaration. Besides, the observation in one single trajectory that the hairpin unfolds by a hydrophobic core-centric mechanism and folds by a zipper mechanism may question the studying of protein folding mechanism by unfolding simulations.

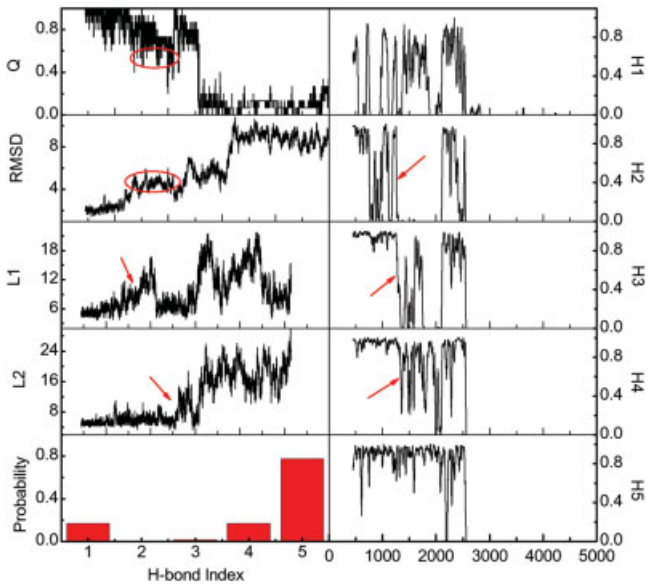

Fig. 15. A trajectory of normal MD simulations, carried out at 500 K. Different figures correspond to different MD observables. The left bottom figure shows the average probability of each hydrogen bond for the states in the transition region. [Color figure can be viewed on the online issue, which is available at www.interscience.wiley.com.]

Interestingly, in the normal MD trajectories, we also find another kind of FF process, as shown in Figure 15. This trajectory shows an unfolding process that begins from the inner part of the hairpin and propagates toward the terminal. In contrast to the trajectory described in Figure 13, the outmost hydrogen bond 5 have a high formation probability in the TS. This trajectory is related to the mechanism proposed previously[16] which unfolds hairpin from the turn and refolds it from the terminal. In our opinion, this mechanism may be a minor one for hairpin folding, because physically the hairpins folding by this mechanism do not follow the minimal-entropy-cost

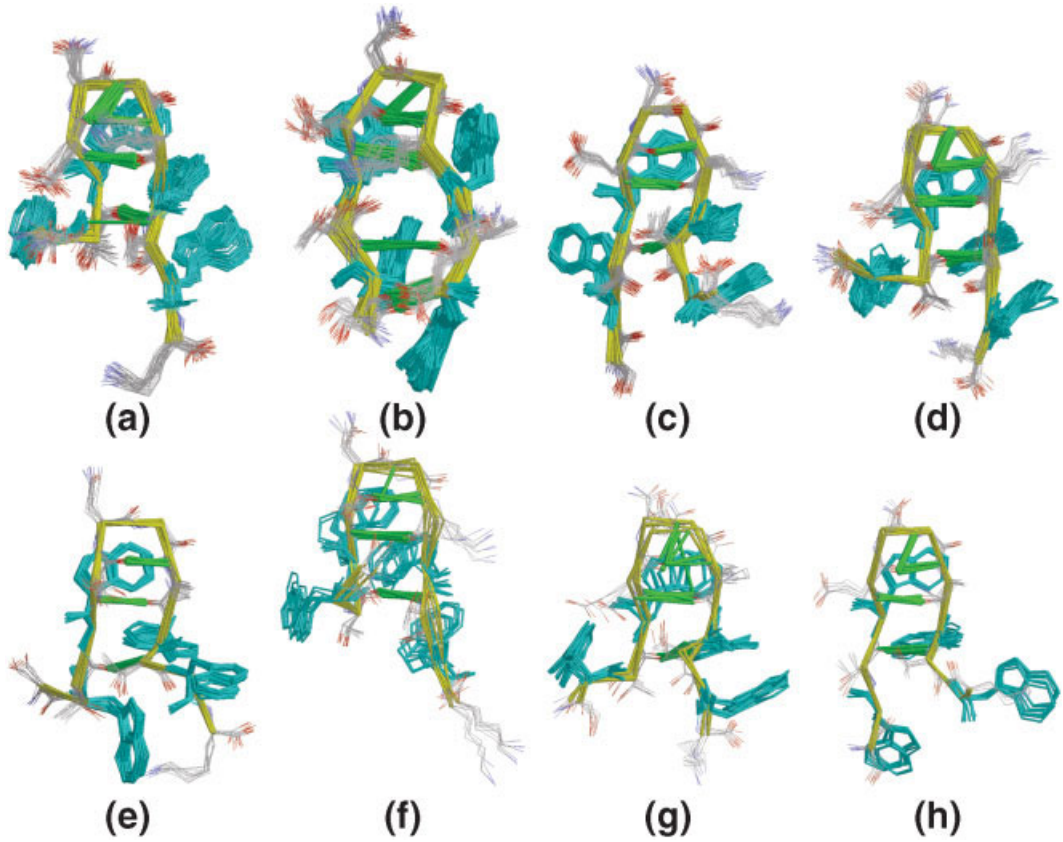

Fig. 16. The structures of the largest eight clusters of TRS at 400 K. These clusters contain 163, 135, 37, 30, 13, 12, 9, and 8 conformations, respectively. The four Trp residues are plotted by sticks, the hydrogen bonds are colored green, and the backbone is colored yellow. [Color figure can be viewed on the online issue, which is available at www.interscience.wiley.com.]

pathway, the formation of contacts between hairpin terminals first will lead to a large entropy cost.

Although the kinetic analysis of the trajectories provides many insights into the structures of TS as well as the folding mechanisms, we feel it may suffer from the small number of trajectories being studied. Since the REMD simulation samples a wide conformation space of the hairpin and reaches thermodynamic equilibrium, we try to extract the information on TS by thermodynamic analysis in the following section.

Thermodynamic analysis

Our main idea to extract information on TS by thermodynamic analysis was to characterize the clusters of conformations at the top of barrier in the projected free-energy surface. It should be emphasized that thus obtained states are only approximations but not direct proofs of the TS, since it is hard to prove that the related reaction coordinates are good ones. Although this problem can be overcome to some extent by using many reaction coordinates simultaneously and using relatively strict rules to locate the transition region, the obtained states are still an approximation of the TS, since the TS refers to a kinetic mechanism, not a thermodynamic one. In this section, we prefer to use the term "transition region" (TR) or "states in transition region" (TRS) and avoid using "TS" to emphasize this approximation. We hope the thermodynamic study of the TRS, combined with the kinetic results on TS in the previous section, can provide a more integral description of the hairpin folding. Keeping the limitations of the method in mind, we try to locate the transition region in the free-energy surfaces. Figure 10 has given the projected free-energy surfaces at 400 K. The reason we investigate the TRS at 400 K is that the overall free-energy barrier at high temperature is relatively lower than at low temperature, thus the transition region will be more populated.

By investigating the projected free-energy surface shown in Figure 10, we locate the transition region by a relatively strict rule, that is, $5.0 < R_g^{core} < 6.0$, $1 \leq nHB \leq 2$, $0.32 < Q < 0.55$, and $4.5 < RMSD < 6.5$. The states in the TR are subjected to the cluster algorithm and classified into 11 clusters. The large number of clusters indicates that the TRS are rather heterogeneous. As shown in Figure 16, these clusters are commonly characterized by the following features: (1) Most importantly, it can be seen that for most clusters, the three hydrogen bonds as well as the turn have been largely formed; (2) The inner pair of indole rings (belong to Trp 4 and Trp9) almost pack together, with the packing patterns similar to that of the native state; (3) The distance between the outer pair of indole rings is in great fluctuation.

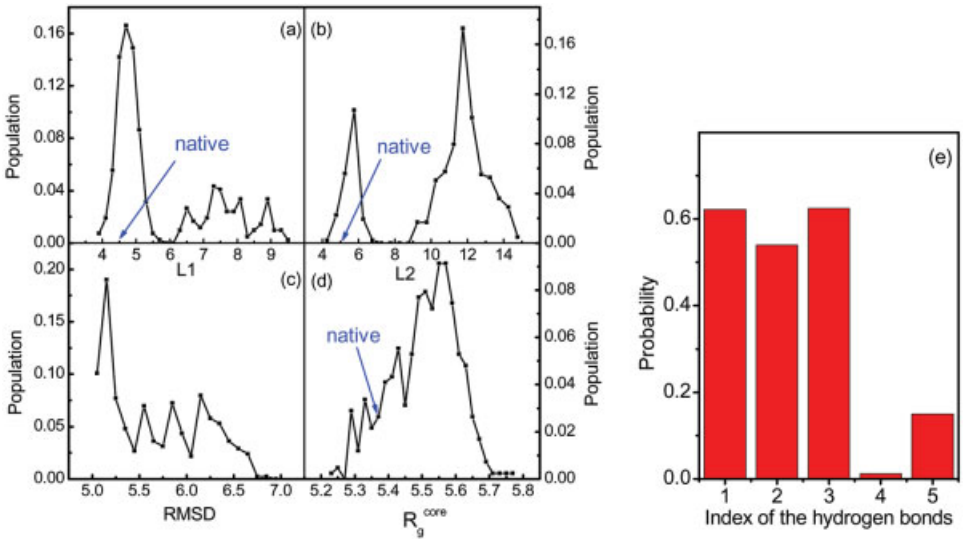

Fig. 17. **a–d:** The distributions of L1, L2, RMSD, and R_g^{core} of the states in transition region. **e:** The average probability of each hydrogen bond of the TRS. [Color figure can be viewed on the online issue, which is available at www.interscience.wiley.com.]

To quantitatively characterize the structures of TRS, Figure 17(a,d) shows the distributions of L1, L2, RMSD, and R_g^{core} of the TRS, respectively. The major peak in Figure 17(a) clearly illustrates that most states have their inner pair of indole rings packing together, whereas more than half of them have the outer pair dissociated [the second peak in Fig. 17(b)]. Besides, as shown by Figure 17(d), the peptide mainly populates the states with $R_g^{core} \sim 5.55$ Å. All of these demonstrate that the hydrophobic core is only partly formed in the TRS, that is, the hydrophobic interaction between Trp4 and Trp9 is almost established whereas that between Trp2 and Trp11 is not. The formation probabilities of the hydrogen bonds are also calculated [Fig.17(e)]: As expected, the inner three native hydrogen bonds have remarkably high formation probabilities compared to the outer two.

The structures of the TRS show almost the same character with that of TS extracted from the FF trajectories, the similarity indicates that the TRS thus obtained is a good approximation of the TS. Put together the kinetic information on TS and the thermodynamic results of the TRS, we can delineate a picture for the TS: The turn and the interaction between the inner pair of hydrophobic core residues are largely formed, whereas the second pair of hydrophobic core residues are dissociated.

The folding mechanism

According to the above kinetic and thermodynamic analyses, the folding picture of the trpzip2 β-hairpin can be described as follows: at the early folding stage, the hydrophobic interactions play dominant roles and initiate the early collapse of the β-hairpin, resulting in a compact H state. Although the H state is rather compact, the packing pattern of hydrophobic residues is not correct since the hydrophobic collapse is nonspecific. The next folding event is the formation of turn and three inner hydrogen bonds, as revealed by the direct trajectory analysis and statistics of the TRS. However, it should be noted that the stability of the nascent turn is strongly related to the establishment of the partly formed hydrophobic core (Trp 4 and Trp9). This is supported by the observation that the formation of turn is always accompanied by the native-like packing of the inner pair of core residues, as shown in Figures 11–14, 16, and 17. Physically, without these hydrophobic interactions and the resulted energy gain, the zipping of the remain strands will continue to be an uphill process in free-energy surface, thus the formed hydrogen bonds are not stable and will disappear soon. The formation of turn and establishment of hydrophobic interactions are cooperative process and take place at almost the same time: It is not appropriate to say which event occurs earlier. Thus our folding model is in essence a blend of the hydrogen bond-centric and hydrophobic core-centric mechanism.

Although these two events occur almost simultaneously, the roles they play in the folding process are different. In our opinion, the roles of the hydrophobic interactions are twofold. Firstly, they provide a nest of compact ensemble for the native hydrogen bonds to register and hatch the β-hairpins. Secondly, they stabilize the nascent turn. On the other side, the sequence of the turn determines the folding rate in that it determines the size of conformation space that the peptide has to search within to find the TS, a peptide with rigid turn can effectively reduce the conformation space of the unfolded states and find the bottleneck of TS faster. This takes account of the evidence that the mutant of D46 slows the folding rate of protein G β-hairpin 20-fold[22] and a recent experiment that shows that the β-hairpins with turn sequence DDATKT and EPNK fold faster than the others.[17]

Our folding picture are almost as same as that proposed by Munoz and Eaton when studying the folding of the protein G β-hairpin.[1] Based on the experiment and a simple statistical mechanical model, they suggested that

both the hydrogen bonds and hydrophobic core residues are important in the zipping up mechanism. They also stated the TS species represents an ensemble of structures which contains the 45–52 stretch of native peptide bonds and all possible non-native ψ, ϕ pairs for the remaining peptide bonds, and the TS is crossed right at the time of making the first hydrophobic interaction. After that, any additional interaction, whether hydrogen bond or hydrophobic, stabilizes the hairpin by slowing down the unfolding rate. Our simulations essentially confirm all of these points. Taking into consideration the simple model they employ, the picture they provide is striking insightful.

Our folding picture is in good agreement with the previous study on β-hairpin using the explicit water model where a blend of hydrogen bond-centric and hydrophobic core-centric mechanisms is proposed.[13] The picture is also consistent with previous experiments and simulations that suggest that the polar residue in turn region and the turn sequence is crucial for β-hairpin folding,[5,12,14,23,24] and to some extent agree with the simulations that generally support a hydrophobic core-centric mechanism.[7,9] Moreover, our simulations provide a direct evidence for a recent experiment which suggests that the rate-limiting event in the β-hairpin folding process corresponds to the formation of the turn, whereas the hydrophobic cluster increases the stability of a β-hairpin by decreasing its unfolding rate.[17]

It should be pointed out that the structure of TS we obtained is different from that of Snow et al.[4] By using the distribution computing technique, they studied the folding rates of trpzips (TZ1, TZ2, and TZ3). Their simulations suggested that the structures of TS are very close to the native state. They are characterized by almost the same backbone twist as the native state and the larger fluctuation of all four Trp residues comparing to that in the native state. In contrast, according to our simulations, the TS is characterized by the native-like packing of the inner pair of core residues and the large fluctuation of the second pair. The difference may be attributed to a different water model, force field, or sampling method used, or a different method used to identify TS. Additional experiments and simulations are needed to make clear which picture is close to the reality.

Comparing With the Protein G β-Hairpin

The folding mechanism of trpzip2 is similar to that of the protein G β-hairpin in many aspects. For trpzip2, the early folding undergoes a rapid collapse to a compact H state, which is characterized by the random packing patterns of hydrophobic residues and small probabilities of backbone hydrogen bonds. This H state is reminiscent of the H state in the early folding stage of protein G β-hairpin.[7–9,13] The early collapse of both hairpins is determined by hydrophobic interactions. Furthermore, in the next folding stages, the hydrophobic interactions also play similar roles for two hairpins. According to previous experiments and simulations on protein G β-hairpin,[1,8–12,15] the hydrophobic interactions are of significant importance in that they help

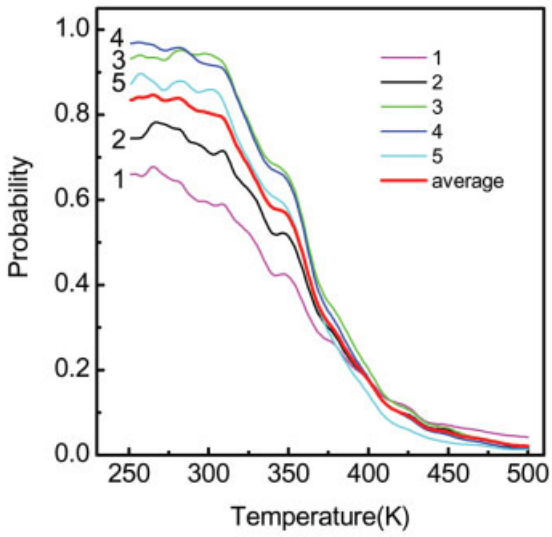

Fig. 18. The formation probability of each hydrogen bonds and their average as a function of temperature. [Color figure can be viewed on the online issue, which is available at www.interscience.wiley.com.]

stabilize the bent precursor of the hairpin. This is also observed in trpzip2 folding.

The folding process of trpzip2 and protein G β-hairpin are striking similar to each other when taking a close look at the formation order of the hydrophobic interactions. Based on the multicanonical MC simulations, Karplus and coworkers suggested that the hydrophobic assembly, in particular the Y45 and F52interaction, forms early.[7] This hydrophobic contact between Y45 and F52 is also found to be one of two nucleation contacts by Zhou et al.[15] (the other nucleation contact is between hydrophilic residues D46 and T49). Pande and coworkers also observed similar aspects.[8,9] Besides, the experimental observation that the Y45 and F52 are crucial for the hairpin folding demonstrates different relative importance of core residues.[12] Similarly, in case of trpzip2, the TS of trpzip2 is characterized by the native-like packing of the inner pair of core residues, whereas the outer pair is in great fluctuation, that is, the hydrophobic interaction between the inner pair of core residues forms earlier. The second pair of core residues packs together later, followed by the formation of the outmost hydrogen bonds. This picture is extremely similar to that of the protein G β-hairpin, demonstrating that they fold by almost the same mechanism. This conclusion was not proved previously but has been used implicitly in many works.

The Stability of the Hydrogen Bond

To study the temperature-dependent stability of hydrogen bonds, the formation probability of each hydrogen bond and their average are calculated over a wide range of temperatures, as shown in Figure 18. Generally speaking, the hydrogen-bond probabilities decrease with temperature, indicating a decrease of β-hairpin content. The averaged probability drops below 0.1 at about 420 K, for comparison, that of the protein G β-hairpin drops below

0.1 at about 570 K.[13] The sharp decrease in the former case indicates higher folding cooperativity in case of trpzip2 β-hairpin. Besides, the average hydrogen bond probability is also higher than that of the protein G β-hairpin, demonstrating its higher stability.[21]

Among the five backbone hydrogen bonds, the H-bonds 3 and 4 are the most stable below 400 K, whereas the first two are much easier to break. This feature could be attributed to two factors: (1) the first two hydrogen bonds are near the turn and it is hard to achieve an ideal N—H···O angle because of geometrical constraints; (2) the existence and high population of P state, where the ψ of residue Glu5 and the ϕ of residue Asn6 are incorrect, leading to the decrease of the stability of the first two hydrogen bonds.

It is interesting to investigate the relationship between the stability of hydrogen bonds and their relative importance in the folding process. Sometimes, the stability of each hydrogen bond as a function of temperature is used as a criterion of their relative importance in the folding process. The hydrogen bonds with high stability are assumed to be important ones in folding. However, this hypothesis may not be the case, at least for trpzip2 β-hairpin. According to our analysis of the transition state, the first three hydrogen bonds are the most important ones in the folding process. This conflicts with their stability as a function of temperature in Figure 18: whereas H-bonds 3 and 4 are the most stable ones, H-bonds 1 and 2 are unstable. However, the relative low stabilities of H-bonds 1 and 2 do not prevent them from playing relatively important roles in the folding process. Therefore, we conclude that the stabilities of the hydrogen bonds are not always a correct criterion in judging their relative importance in folding. This feature is also observed in peptide G β-hairpin. Dinner and Karplus have calculated the temperature dependence of formation probability of each hydrogen bond for this peptide.[7] Although the experiments have demonstrated the significant importance of hydrogen bonds in the turn region for both folding rate and stability,[12,22] their stabilities calculated by computer simulation are the lowest.

To confirm the above simulation results, we also performed the REMD simulations of trpzip2 with the GB/SA water model. It is also found that the hydrogen bonds 3 and 4 are the most stable ones whereas the H-bonds 1 and 2 are most unstable, although the stabilities of all hydrogen bonds are lower than that calculated by using the explicit water model (see Fig. 19). Moreover, for the calculations with the GB/SA model, all the intermediate states, that is, the H, P, and M are found, and the structures of these states are similar to that calculated by the explicit water model. These observations confirm the convergence of our calculations as well as the robustness of our conclusions.

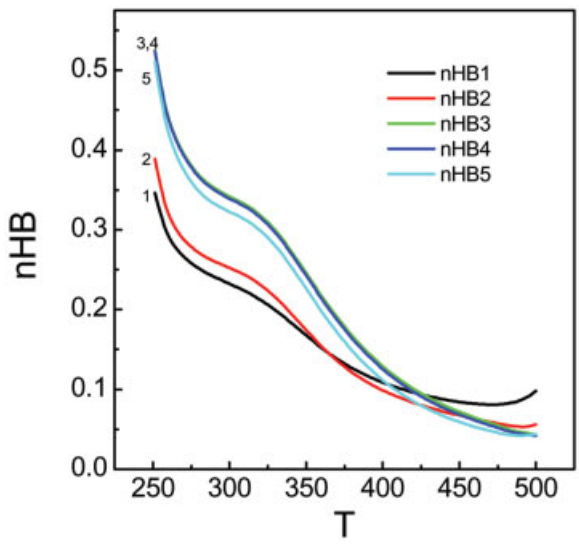

Fig. 19. The formation probability of each hydrogen bond and their average as a function of temperature, calculated by GB/SA model and REMD simulations. The relative stability of each hydrogen bond is as same as that calculated by explicit water model, although the absolute value is not. [Color figure can be viewed on the online issue, which is available at www.interscience.wiley.com.]

CONCLUSIONS

Understanding the folding of β-hairpins can provide glimpses into the folding process of large proteins and the mechanism of protein disease. Although extensive studies have been done, their folding mechanism remains elusive. By using a recently developed replica-exchange MD method, the trpzip2 is studied as a model β-hairpin. Several conclusions can be made: (1) The very early folding process is initiated by the nonspecific hydrophobic collapse; (2) The TS is characterized by a largely formed turn, high probabilities of the three inner hydrogen bonds, and the hydrophobic interaction between the inner pair of core residues. Moreover, the stability of the nascent turn is strongly related to the partly formed hydrophobic core. The turn formation and establishment of hydrophobic interactions are cooperative processes and occur almost simultaneously: It is not appropriate to say which one occurs earlier; (3) The roles played by turn and hydrophobic interactions in folding process are different. The turn sequence determines the size of conformation space of the unfolded states that the peptides have to search within to find the bottleneck of the TS, thus determining the folding rates. The hydrophobic interactions provide a compact ensemble to facilitate hydrogen-bond registration and stabilize the nascent turn; (4) The folding mainly follows a zipper mechanism, however, the hydrophobic interactions are also critical. Our folding picture is a blend of hydrogen bond-centric and hydrophobic core-centric mechanism.

Although the folding is mainly a zipping up process, we emphasize that the hairpin can also fold by the other mechanism which starts the folding from the terminal, as revealed by our direct MD simulations. Besides, our simulations cannot exclude the other possible mechanisms since the kinetic analysis may suffer from small number of trajectories and the thermodynamic analysis can only provide an approximation of the TS. For example, a recent simulation proposed a folding mechanism that can be described as a reptation move of one strand with respect to

the other.[16] To achieve a complete understanding of the hairpin folding, further massive computations are need and more accurate method to characterize the energy landscape and TS have to be developed.

ACKNOWLEDGMENTS

We thank two anonymous referees for their critical reading of the manuscript and very helpful comments. This work is supported by the National Natural Science Foundation of China (No. 10504012, 90403120, 10021001, and 10474041) and the Nonlinear Project (973) of the NSM. The computation was done at Shanghai Supercomputer Center. The author thanks Ting Zhang and Jun Yuan for their help at parallel computing.

REFERENCES

1. Munoz V, Thompson PA, Hofrichter J, Eaton WA. Folding dynamics and mechanism of β-hairpin formation. Nature 2003;390:196–199.
2. Munoz V, Henry ER, Hofrichter J, Eaton WA. A statistical model for β-hairpin kinetics. Proc Natl Acad Sci USA 1998;95:5872–5879.
3. Snow CD, Nguyen H, Pande VS, Gruebele M. Absolute comparison of simulated and experimental protein-folding dynamics. Nature 2002;420:102–106.
4. Snow CD, Qiu L, Du D, Gai F, Hagen SJ, Pande VS. Trp zipper folding kinetics by molecular dynamics and temperarture-jump spectroscopy. Proc Natl Acad Sci USA 2004;101:4077–4082.
5. Klimov K, Thirumalai D. Mechanisms and kinetics of β-hairpin formation. Proc Natl Acad Sci USA 2000;97:2544–2549.
6. Kolinski A, IIkowski B, Skolnick J. Dynamics and thermodynamics of β-hairpin assembly: insights from various simulation techniques. Biophys J 1999;77:2942–2952.
7. Dinner AR, Lazaridis T, Karplus M. Understanding β-hairpin formation. Proc Natl Acad Sci USA 1999;96:9068–9037.
8. Pande VS, Rokhsar DS. Molecular dynamics simulations of unfolding and refolding of a β-hairpin fragment of protein G. Proc Natl Acad Sci USA 1999;96:9062–9067.
9. Zagrovic B, Sorin EJ, Pande VS. β-hairpin folding simulations in atomistic detail using an implicit solvent model. J Mol Biol 2001;313:151–169.
10. García AE, Sanbonmatsu KY. Exploring the energy landscape of a β-hairpin in explicit solvent. Proteins 2001;42:345–354.
11. Bolhuis PG. Transition-path sampling of β-hairpin folding. Proc Natl Acad Sci USA 2003;100:12129–12134.
12. Kobayashi N, Honda S, Yoshii H, Munekata E. Role of side chains in the cooperative β-hairpin folding of the short C-terminal fragment derived from streptococcal protein G. Biochemistry 2000;39:6564–6571.
13. Zhou R, Berne BJ, Germain R. The free energy landscape for β-hairpin folding in explicit water. Proc Natl Acad Sci USA 2001;98:14931–14936.
14. Tsai J, Levitt M. Evidence of turn and salt bridge contributions to β-hairpin stability: MD simulations of C-terminal fragment from the B1 domain of protein G. Biophysical Chemistry 2002;101:187–201.
15. Zhou YQ, Linhananta A. Role of hydrophilic and hydrophobic contacts in folding of the second β-hairpin fragment of protein G: molecular dynamics simulation studies of an all-atom model. Proteins 2002;47:154–162.
16. Wei GH, Mousseau N, Derreumaux P. Complex folding pathways in a simple β-hairpin. Proteins 2004;56:464–474.
17. Du DG, Zhou YJ, Huang CY, Gai F. Understanding the key factors that control the rate of β-hairpin folding. Proc Natl Acad Sci USA 2004;101:15915–15920.
18. Sugita Y, Okamoto Y. Replica-exchange molecular dynamics method for protein folding. Chem Phys Lett 1999;314:141–151.
19. Ferrenberg AM, Swendsen RH. Optimized Monte Carlo data anaylsis. Phys Rev Lett 1998;63:1195–1198.
20. Zhou RH. Trp-cage: folding free energy landscape in explicit water. Proc Natl Acad Sci USA 2003;100:13280–13285.
21. Cochran AG, Skelton NJ, Starovasnik MA. Tryptophan zippers: stable, monomeric β-hairpins. Proc Natl Acad Sci USA 2001;98:5578–5583.
22. McCallister EL, Alm E, Baker D. Critical role of β-hairpin formation in protein G folding. Nat Struct Biol 2000;7:669–673.
23. Dyer RB, Maness SJ, Peterson ES, Franzen S, Fesinmeyer RM, Andersen NH. The mechanism of β-hairpin formation. Biochemistry 2004;43:11560–11566.
24. Fesinmeyer RM, Hudson FM, Andersen NH. Enhanced hairpin stability through loop design: the case of the protein G B1 domain hairpin. J Am Chem Soc 2004;126:7238–7243.
25. Yang WY, Pitera JW, Swope WC, Gruebele M. Heterogeous folding of trpzip hairpin: Full atom simulation and experiment. J Mol Biol 2004;336:241–251.
26. Sanbonmatsu KY, García AE. Structure of met-enkephalin in explicit aqueous solution using replica exchange molecular dynamics. Proteins 2002;46:225–234.
27. Dokholyan NV, Buldyrev SV, Stanley HE, Shakhnovich EI. Identifying the protein folding nucleus using molecular dynamics. J Mol Biol 2000;296:1183–1188.