Biophysical Chemistry 137 (2008) 116–125

Contents lists available at ScienceDirect

Biophysical Chemistry

journal homepage: http://www.elsevier.com/locate/biophyschem

Ab initio folding simulation of Trpcage by replica exchange with hybrid Hamiltonian

Weixin Xu, Yuguang Mu *

School of Biological Sciences, Nanyang Technological University, Singapore 637551, Singapore

ARTICLE INFO

Article history:
Received 12 July 2008
Received in revised form 5 August 2008
Accepted 5 August 2008
Available online 13 August 2008

Keywords:
Ab initio protein folding
Trpcage
Replica-exchange simulation method with hybrid Hamiltonian

ABSTRACT

Replica-exchange molecular dynamics simulations with hybrid Hamiltonian in explicit solvent were performed to study the folding of a designed 20-residue miniprotein, Trpcage, from a fully extended structure. During the simulations several folding/unfolding events happened. In the folded states the majority of experimentally observed NMR NOE restraints are satisfied. The folded structures have root mean squared deviation of 2.0 Å with respect to the NMR structures considering all heavy atoms. The free-energy surface constructed by the hybrid Hamiltonian simulations is similar to the one built by a standard replica-exchange simulation which started from the native structure. Consistent with previous experimental observation, a pre-existing hydrophobic collapse in the unfolded state is detected by investigating the desolvation behavior of Trpcage. At room temperature, an intermediate state featured by a misfolded core, a nearly formed α-helix segment and an absence of 3_{10}-helix is found. The replica exchange with hybrid Hamiltonian method is shown here to be capable of resolving the folding picture of the miniprotein.

1. Introduction

All-atom molecular dynamics (MD) simulations with physically based energy functions provide abundant information about movements of each atom in a protein [1–6]. In principle the mystery of protein folding could be resolved by such simulations. But due to the fact that proteins have very rough free-energy landscapes conventional MD simulations at room temperature suffer from a problem known as kinetic traps. A possible solution of this problem is to perform MD simulations in a generalized ensemble [7]. The idea is to achieve a random walk on the potential-energy surface. As a result the system can easily jump out of local minima and sample a much broader phase space. There are three well-known approaches for carrying out generalized ensemble MD simulations: multicanonical algorithm [8], simulated tempering [9] and replica-exchange method (REM) [10]. Different from other two methods REM employs standard Boltzmann weight factors and thus is more easily accomplished. Recently REM has been increasingly employed to study the energy landscape of protein folding [11–14]. However for large systems, such as proteins in explicit solvent, REM will require a huge number of computational facilities because the number of replicas needed increases as the square root of the degree of freedom of the system [15]. Endeavors to overcome the shortage of REM have been making by several groups [15–22].

Recently, we presented replica-exchange molecular dynamics simulations with hybrid Hamiltonian (REMhH/REMshH) in explicit solvent for *ab initio* folding of small peptides [23,24]. Using OPLSAA force field [25], these methods were able to fold small β-hairpins such as C-terminus (residue 41–56) of protein G and trpzip2 (12 amino acids) from fully extended structures into native configurations. The number of replica required is fewer than those needed by standard replica-exchange molecular dynamics (REMD) simulation. And the sampling ability of REMhH/REMshH was found to be approximative to REMD as the free-energy landscapes constructed by these methods were similar to those obtained from REMD method.

Presently, due to the limit of computational ability, the proper objects for all-atom simulations are focused on small proteins. The discovery of small and fast folding proteins, such as Avian pancreatic polypeptide [26], Villin headpiece [27], WW domain [28], Trp zipper [29] and Trpcage [30] has contributed much to bridge the gap between experiments and computer simulations. Among these, the designed miniprotein Trpcage is very attractive due to its small size (20 residues), fast folding kinetics (4 μs) and high thermodynamic stability under the condition without crosslinking via disulfide formation or metal ion chelation. A few all-atom model simulations [31–33], transition path sampling method [34] and experiments [35–37] have exposed many interesting features about the folding of Trpcage. However the disputes on folding mechanisms of Trpcage are still not solved. Qiu et al. suggested a cooperatively two-state folding picture by laser temperature-jump spectroscopy experiments [35]. However, both UV-resonance Raman spectroscopy measurements by Ahmed et al. [36] and fluorescent correlation spectroscopy by Neuweiler et al. [37] showed that Trpcage was not a simple two-state folder. Additionally, some important aspects about Trpcage folding are also unclear: (1) Does collapse precede or follow the formation of α-helix? (2) At what stage of folding is the α-helix formed? The molecular dynamics simulation indicated that the α-helix is formed in the final stage [32]. However, the recent experimental observations showed evidences of a helical structure in the denatured

* Corresponding author.
E-mail address: ygmu@ntu.edu.sg (Y. Mu).

0301-4622/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2008.08.002

state of Trpcage, suggesting an early formation of the helical component [36]. (3) What role does the salt bridge between Asp9 and Arg16 play? Replica exchange MD simulations [32] suggested that salt bridge formed early and played an important role in the folding kinetics as speeding up folding. However, the fluorescence experiments by Sauer and coworkers did not observe significant change in folding time after elimination of the salt bridge.

In this article, the *ab initio* folding of Trpcage in explicit solvent was studied using REMshH strategy aiming to investigate its reliability and to provide more insights into the problems mentioned above. The folded structures obtained from the simulations match NMR structures [30] very well with root mean squared deviations (RMSD) less than 2.0 Å. At room temperature there exists an intermediate state featured by a misfolded cage, a nearly formed α-helix and an absence of 3_{10}-helix on the folding free-energy landscape, which indicates a non-two-state folding behavior. The detailed folding order of key secondary structures was probed. By studying the water dynamics, a high degree of dehydration of the hydrophobic core was detected during the folding process of Trpcage.

2. Model and methods

2.1. Simulation method

In our simulations, the all-atom represented Trpcage was solvated with explicit 800 SPC water molecules, resulting in a cubic box with length of 30 Å. The replica-exchange molecular dynamics simulations with scaled hybrid Hamiltonian [24] were implemented. The technical details can be found in our previous study [24]. The GROMACS program suite [38] and the full atomic OPLSAA force field [39] were used. The Trpcage system was simulated with 16 replicas, with temperature ranging from 300.0–767.1 K. The starting structures were of fully extended structures. Three simulations with different initial velocities ran for 142 ns, 50 ns and 71 ns, respectively. In the Results section, the displayed simulation results are mainly from the 142-ns simulation if not specially indicated. To check the quality of samplings obtained by the hybrid Hamiltonian method standard REMD simulation with 36 replicas covering the temperature range from 297–551 K was also performed. In the standard REMD simulations, the initial structures were the NMR structure [30].

2.2. Analysis method

A weighted histogram analysis method (WHAM) [40,41] was used to combine data of all replicas to calculate some properties presented in this article. The reaction coordinates used in analysis are listed as follows: number of native backbone hydrogen bonds (nHB), radius of gyration (R_g), radius of gyration of the hydrophobic core ($R_{g_{core}}$), fraction of native contacts (Q), fraction of native contacts for Trp6 (Q_{Trp6}), RMSD with respect to the NMR structure which is calculated based on all heavy atoms, RMSD of the hydrophobic core, the length of α-helix and 3_{10}-helix, the salt-bridge distance between Asp9 and Arg16, the solvent-accessible surface for whole Trpcage or Trp6 only and the number of water molecules in the vicinity of the surface of Trpcage or Trp6. The native contacts are atom-based which were defined as those atom-pairs within 5.0 Å of each other when two amino acid residues are separated by five or more amino acids. Here, the common 140 contacts of 38 NMR structures in the PDB file (PDB code: 1L2Y.ent) are considered as the total number of native contacts in the analysis. We find that the choice of threshold of 5.0 Å is not critical and tuning this value within the range 4.5–5.5 Å does not change the results qualitatively. A hydrogen bond is counted if the distance between two heavy atoms (N and O in this case) is less than 3.5 Å and the angle N–H O is larger than 120°. To characterize the structure of the folded species identified from the free-energy surface, we cluster the corresponding ensembles by using the algorithm proposed by Snow et al. [42]. The conformations are clustered by their structural similarity, measured by the pair-wise RMSD. The related RMSD cutoff is chosen as 2.0 Å. All the ensemble statistical analysis reported here is performed on the last 92 ns of the trajectory. While the first 50 ns is treated as an equilibration phase.

Replica-exchange simulations sample a wide range of conformation space at the cost of eliminating the information on dynamics. The trajectory at each temperature is structurally discontinuous due to exchange of conformations between neighboring replicas under different temperatures. In the following, such trajectories are named after "ensemble trajectory". However, if we tail after one conformation following the exchange information, a continuous folding trajectory can be obtained, which is called "replica trajectory". Recently, such kind of trajectory has been used by Sanbonmatsu and Garcia [43] and Zhang et al. [13] in studying the folding mechanism of small proteins.

3. Results and discussion

3.1. Successful folded replica trajectories

Trpcage is observed to fold well from fully extended conformations into its native structure within a 142-ns simulation using replica-exchange molecular dynamics with hybrid Hamiltonian (denoted by S1). For reproducibility check, two additional simulations were also performed with the same method and displayed in Fig. 1 (denoted by S2 and S3). Fig. 1 (S1) displays one successful folding trajectory (replica trajectory). The evolutions of some quantities, such as simulation temperature, Q and RMSD with time demonstrate the successful folding process. This trajectory is extracted from the 16 replicas by tracking down one conformation. In such a replica trajectory the Cartesian coordinates of all atoms are continuous and updated according to the physics law with velocities rescaled universally at the exchange points. Thus the orderings of events happening in such a trajectory still have clear physical meanings. After 23 ns, Q and RMSD fluctuate around 0.80 and 2.0, respectively. This is a good demonstration that REMshH has sampled the global minimum of Trpcage. It should be pointed out that the definition of atom-based Q is much stricter than previously used residue-based definition of Q. A structure with Q=0.80 is already much native-like (with RMSD for all heavy atoms less than 2 Å). Also seen from the time evolution of RMSD for S2 and S3, the RMSD in several trajectories can reach around and even below 2 Å. Thus, this miniprotein can be folded into nature structures in the presence of explicit water molecules reproducibly. One representative folded structure which is obtained in S1 is shown in Fig. 2, which will be detailedly discussed in the following.

3.2. Folded structure

In order to obtain the representative folded structures from the simulation trajectory, the conformations with RMSD less than 5 Å are picked out and clustered according to their structural similarity which is measured by the pair-wise RMSD. The related RMSD cutoff is chosen as 2.0 Å. After clustering analysis, the conformers in the largest cluster are taken out as the representative folded structures.

Fig. 2 shows the comparison of one well folded structure (RMSD=1.2 Å) with a native structure determined by NMR (Model 1 of the NMR structures). Such well folded structures can also be found in other two simulations. The NMR structure (the blue ribbon one) shows an α-helix from residues 2 to 8 and a 3_{10}-helix from residues 11 to 14. It also reveals a compact hydrophobic core where four proline residues (Pro12, Pro17, Pro18, and Pro19) and a leucine (Leu7) pack against the aromatic side chains of Tyr3 and Trp6. The sixth residue, Trp6, is caged by the 3_{10}-helix group, the side chains of residues Tyr3, Leu7, Arg16, together with the C-terminal polyproline II stretch.

The similarity between the folded structure from simulation and the one determined by NMR can be reflected by the overlapping of

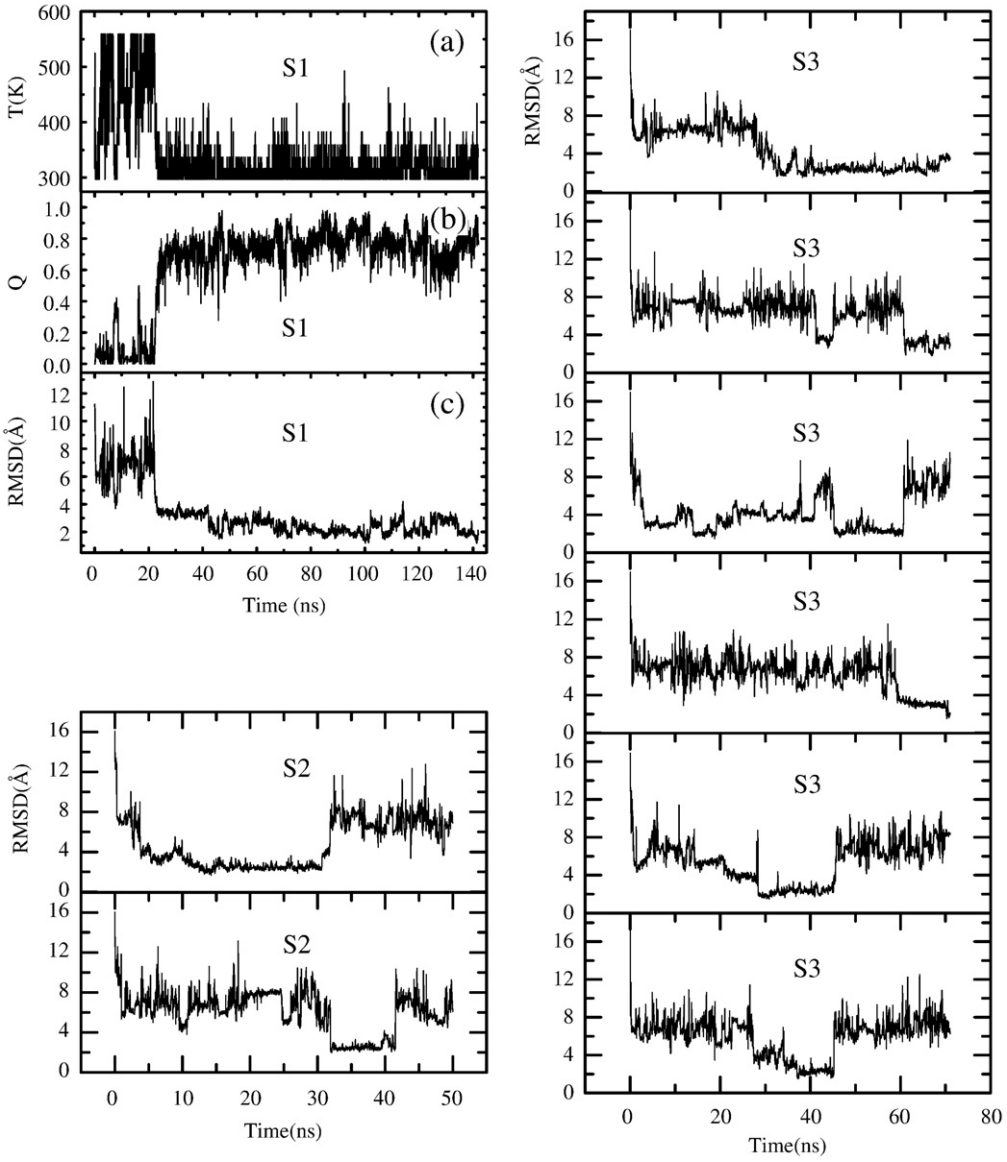

Fig. 1. Folding trajectories of three simulations using replica-exchange molecular dynamics with hybrid Hamiltonian in explicit solvent, denoted by S1, S2 and S3. The time evolution of simulation temperature (a), fraction of native contacts, Q (b), and RMSD (c) are shown for S1. Only RMSDs are displayed for S2 and S3.

their backbone conformations, as illustrated in Fig. 2. For the cluster of structures which is found to best reproduce the NMR structures the average backbone RMSD (based on the backbone atoms N, CA and C of each amino acid) is 1.42 Å, and the best backbone RMSD is 0.59 Å, which comes out as one of the good model compared with other predictions, listed in Table 1. The major deviations originate from N and C-terminal residues. As a matter of fact, the N and C-terminals of the 38 NMR structures are also quite flexible (data not shown). Interestingly, the α-helix and the 3_{10}-helix appearing in the folded structure consists of residues 2 to 8 and 11 to 14, respectively, exactly the same as the corresponding segments determined by NMR in Fig. 2. It is noted that both previous studies which were based on the OPLSAA force field with explicit solvent [32] (start from native structure) and based on ECEPP/3 force field in vacuo [44] found the segment of 11 to 14 does not show the short 3_{10}-helix conformation in the simulated structures. In addition, it is worth taking a look at the orientation and packing of the sidechains of Trpcage in the simulated structure. Except the sidechains of Asn1, Leu2, Lys8, Asp9, Arg16 and Ser20, sidechains of other residues in the simulated structure match well with the NMR structure. Among the most fluctuating sidechains, those of Asn1, Leu2 and Ser20 belong to the terminal residues. The long sidechain of Lys8 is charged and exposed to solvent which may result in fluctuating orientations. While for the sidechains of Asp9 and Arg16, they are nearer to each other than in the native state. The average distance in the folded structures is 4.3 ± 1.1 Å, and the native one is 7.6 ± 0.0 Å (see Table 2). Evidently a over-stabilized salt bridge is formed between Asp9 and Arg16 which may distort the orientations of the sidechains.

In Table 1, we summarize some computational studies that lead to structural predictions of Trpcage. The RMSDs of the obtained conformations with respect to the native structure determined by NMR are listed for comparison. Also listed in the table are the force fields, solvation models, and simulation methods used in these computations. The accuracy of the predictions depends on a number of factors, which include the force fields, the solvent models, and the computer simulation methods. From this table the simulations using the force fields Amber/Charmm/OPLSAA/PFF01 all succeeded in reproducing the experimentally determined structures within relatively small RMSDs. Obviously, comparing with others, our results on structural prediction of Trpcage are among the best.

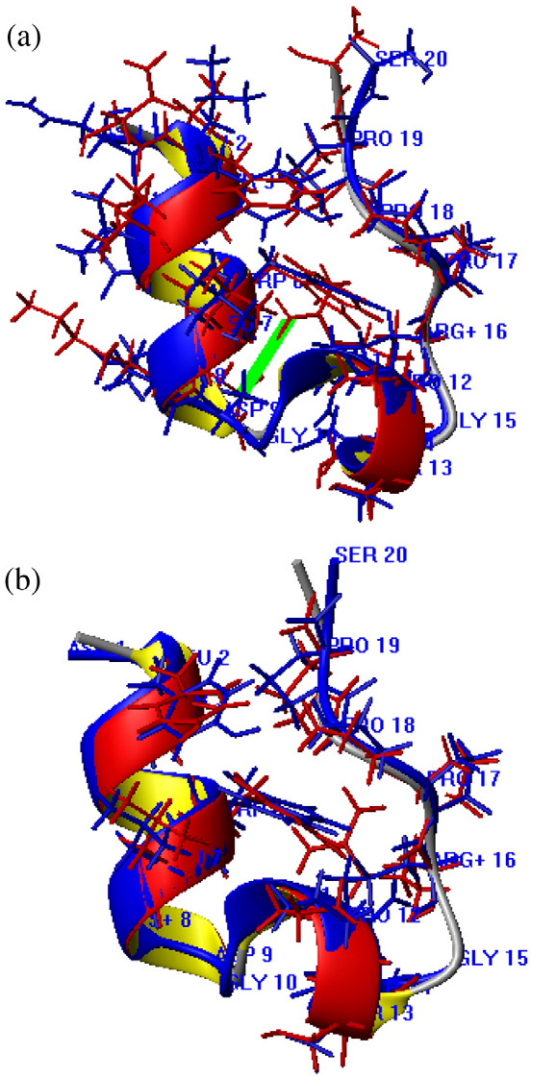

Fig. 2. The ribbon structures of native Trpcage obtained by NMR (blue color) and the representative folded one obtained by simulation (red color). (a) All sidechains are displayed. The green line denotes the salt bridge formed between Asp9 and Arg16. (b) Only sidechains of hydrophobic-core residues are shown for a clearer view. Shown by the program MOLMOL [69]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Comparison of recent works on Trpcage structure prediction

Author	RMSD (A)	Force field	Solvent model	Simulation method
1 Simmerling et al. [59]	0.97 (C_α)	Amber99	GB/SA	Molecular dynamics
2 Snow et al. [56]	2.1 (C_α)	OPLSua	GB/SA	Stochastic dynamics
3 Pitera and Swope [52]	< 1.0 (C_α)	Amber94	GB/SA	REMD
4 Chowdhury et al. [31]	< 1 (MC)	Amber03	GB/SA	Molecular dynamics
5 Carnevali et al. [60]	1.3 (C_α)	Amber94	GB/SA	Modal Monte Carlo
7 Chowdhury et al. [33]	1.0 (MC)	Amber03	GB/SA	Molecular dynamics
8 Chowdhury et al. [33]	1.05 (MC)	Amber03	GB/SA	Molecular dynamics
9 Ota et al. [61]	0.92 (MC)	Amber99	GB/SA	Molecular dynamics
10 Schug et al. [62]	2.83 (BB)	PFF01	SAS model	Stochastic tunneling
11 Schug et al. [63]	2.0 (BB)	PFF01	SAS model	Parallel tempering
12 Schug et al. [64]	3.2 (BB)	PFF01	SAS model	Energy landscape paving
13 Schug et al. [65]	3.19 (BB)	PFF01	SAS model	Basin hopping
14 Steinbach [66]	1.75 (MC) 2.39 (MC) 1.65 (MC) 0.82 (MC)	Charmm19 Charmm19 Charmm22 Charmm22-CMAP	EEF1 ACE ACE ACE	Monte Carlo minimization/ annealing
15 Ding et al. [67]	< 1 (C_α)	Simplified potential	In vacuo	Discrete MD
16 Irbäck et al. [68]	2.3 (BB)	Simplified potential	In vacuo	Monte Carlo
17 Zhan et al. [44]	2.24 (BB)	ECEPP/3	In vacuo	Basin paving
18 This work	1.42, 0.58 (BB) 2.34, 1.12 (HA) 1.48, 0.64 (MC) 1.62, 0.61 (CA)	OPLSAA	SPC explicit water	REMshH

RMSD, root mean square deviation; C_α, C_α-atom; BB, backbone; HA, heavy-atom; MC, main chain; GB/SA, generalized Born/solvent-accessible surface area implicit solvent model; SAS, solvent-accessible surface; EEF1, effective energy function; ACE, generalized Born with analytic treatment of continuum electrostatics; SPC, simple point charge. For our simulated results, the average values are shown, followed by minima.

3.3. NMR restraints

It is also of great interest to compare the simulated pair distances directly with the NMR nuclear Overhauser effect (NOE) data so that we can know whether the calculated proton pair distances fall within the distance range of the NOEs assigned in the NMR experiment. The NMR measurements by Neidigh et al. [30] have provided 169 NOE distance constraints, of which 43 are for intra-residue constraints, 62 for sequential residues, 36 for $i/i+n$ with n=2–4 residues, and 28 for $i/i+n$ with $n\geq 5$ residues. The last 28 constraints are the key long-range distance constraints which have attracted close investigations [32]. Here, we also use the last 28 constraints for direct comparison. As for the intra-residue and sequential residue NOE constraints, they are easily satisfied, even for structures significantly deviated from the native one. The calculating formula for distances is $R_{AVG}=\langle R_{HH}^{-6}\rangle^{-1/6}$.

Fig. 3 compares the detailed NOE pair distances of structures with RMSD ≤2.0 Å from the 142-ns simulation to the NMR data for the 28 key long-range ($i/i+n$ with $n\geq 5$) pairs. The agreement with the NMR NOEs is good that all of the 28 key proton pairs are found to be within 5 Å, the typical NOE signal distance. Actually, none of the simulation

Table 2
List of various statistical and folding features for Trpcage

Index		US	IS	TS	FS	NS
α-helix	L	6.4±1.2	6.0±1.5	4.0±1.4	6.3±1.0	6.2±0.1
3_{10}-helix	L	0.0±0.0	0.0±0.0	1.5±1.6	1.6±1.6	4.0±0.0
Core	R_g	6.5±0.2	5.9±0.1	5.4±0.1	5.3±0.1	5.5±0.0
	RMSD	2.9±0.1	2.8±0.1	4.8±0.2	4.4±0.3	0.0
HB		4.1±1.0	3.5±1.2	2.7±1.1	5.7±1.5	10.2±0.3
SB9-16	Dist (Å)	16.2±1.4	14.9±1.4	9.6±1.9	4.3±1.1	7.6±0.0
Q_{Trp6}		0.18±0.03	0.44±0.03	0.62±0.02	0.86±0.03	1.0±0.0
NC_{Trp6}		296±16	335±21	263±18	274±22	141±6
SAS_{Trp6}	$Å^2$	1.0±0.1	0.7±0.1	0.6±0.1	0.5±0.1	0.2±0.0
WAT	Num	92.1±4.7	88.0±5.7	87±5.9	85.0±5.9	NA.
WAT_{Trp6}	Num	4.9±1.3	3.8±1.3	2.5±1.1	2.2±1.3	NA.
Q		0.17±0.01	0.42±0.02	0.55±0.01	0.79±0.01	1.0
NC		471±284	512±289	529±332	501±286	268±13
R_g	Å	7.6±0.1	7.2±0.1	7.3±0.0	7.4±0.1	7.3±0.0
RMSD	Å	5.3±0.2	4.8±0.2	3.6±0.2	2.0±0.2	0.0

All values are averaged over the most populated clusters in corresponding regions. US, IS, TS, FS, NS represent the unfolded state, intermediate state, transition state, folded state and native state (38 NMR data), respectively. NC, SB, SAS and WAT Num mean the number of all contact, salt bridge, solvent-accessible surface and water number, respectively.

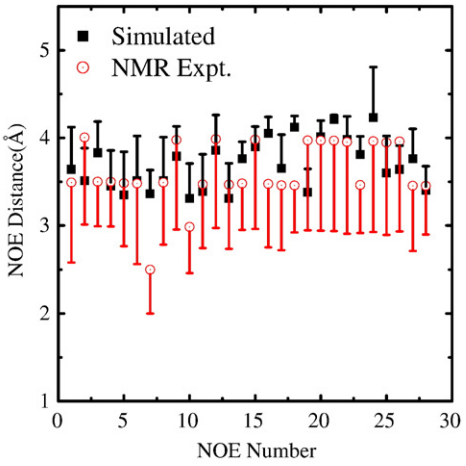

Fig. 3. Comparison of simulated nuclear Overhauser effect (NOE) with the NMR data. Only half error bars are shown for fear of overlapping of curves. 28 key long-range NOE pairs are displayed: 1, HE* Y3 and HB2 P18; 2, HA Y3 and HB2 P19; 3, HA Y3 and HD2 P19; 4, HA Y3 and HG* P19; 5, HB2 Y3 and HG* P19; 6, HD* Y3 and HAP12; 7, HZ2W6 and HAP12; 8, HH2W6 and HG* P12; 9, HH2W6 and HA P12; 10, HH2 W6 and HD1 P12; 11, HD1 W6 and HB* R16; 12, HE1 W6 and HE R16; 13, HE1 W6 and HB* R16; 14, HE1 W6 and HA P17; 15, HZ2 W6 and HA P17; 16, HD1 W6 and HD* R16; 17, HD1 W6 and HG* R16; 18, HZ2 W6 and HD1 P18; 19, HZ2 W6 and HD2 P18; 20, HZ2 W6 and HB1 P18; 21, HZ2 W6 and HG1 P18; 22, HH2 W6 and HB1 P18; 23, HE1 W6 and HA P18; 24, HH2 W6 and HG1 P18; 25, HD1 W6 and HA P18; 26, HD1 W6 and HD2 P19; 27, HD2* L7 and HD1 P12; 28, HB1 D9 and HB2 S14.

values exceeds 4.3 Å. The maximum one is 4.23±0.58 Å, pair 24 between HH2 W6 and HG1 P18. The minimum one is 3.31±0.40 Å, pair 10 between HH2 W6 and HD1 P12. Inspiringly, almost all of them are within the NOE distance ranges provided by Neidigh et al. [30], with the average deviation between simulated pair distances and NMR data is 0.26±0.22 Å.

3.4. Free-energy landscape

In the following section, the free-energy landscape (in units of k_BT) for the 142-ns simulation as well as the 100-ns standard REMD simulation is characterized as a function of R_g and Q at various temperatures: 300 K, 333 K and 400 K (Fig. 4) using the WHAM method [21,40,41]. The free-energy surface (Fig. 4a) constructed by REMshH is similar to the one built by standard REMD (Fig. 4d), both in the number and locations of local minima.

Surprisingly the free-energy landscapes are quite smooth, even more smooth than those of smaller peptides, β-hairpins [42,45–50]. It clearly discloses three states or basins of attractions on each free-energy surface: the unfolded state (US), intermediate state (IS) and folded state (FS). It should be pointed out here that the free-energy landscape for the 71-ns simulation is quite similar to that shown in Fig. 4. While the free-energy landscape for the 50-ns simulation is not characterized due to its short simulation time. According to the projected free-energy surfaces in Fig. 4, we define the unfolded state as $0.14<Q<0.19$ and 7.46 Å$<R_g<$7.71 Å, the intermediate state as $0.40<Q<0.45$ and 7.10 Å$<R_g<$7.35 Å, the transition state (between intermediate state and folded state) as $0.53<Q<0.58$ and 7.25 Å$<R_g<$7.51 Å and the folded state as $Q>0.77$ and $R_g<$7.43 Å. It is noted that the free-energy barrier between the intermediate state and the folded state is as low as $1.1k_BT$. However, this free-energy barrier can separate the intermediate state from the folded state since it is larger than $1k_BT$ [51].

In order to identify the representative structures in each state to provide glimpses into the underlying folding features, we performed cluster analysis of conformations in each basin of attraction. The representative structures are displayed in Fig. 5. In total 15 reaction coordinates are applied to characterize the most populated cluster, which are listed in Table 2. The statistics of various properties of native Trpcage is also listed in Table 2 for comparison.

Fig. 5 (US) represents the unfolded-state ensemble. The overall architecture was somewhat close to the native structure, especially the N-terminal portion. The α-helix is largely formed. However, the 3_{10}-helix as well as the caging of Trp6 is not observed. The salt-bridge distance (the green line) is very large which indicates that the salt bridge between Asp9 and Arg16 is absent. From Table 2, one can have a quantitative view: R_g is 7.6 Å which is close to the native value 7.3 Å; the length of α-helix and 3_{10}-helix is 6.4±1.2 and 0.0, respectively; $R_{g_{core}}$ is 6.5 Å significantly larger than the native one 5.5 Å, and the fraction of native contacts for Trp6 is as low as 0.18±0.03; also noted is the salt-bridge distance SB9-16 is 16.2±1.4 Å, much larger than the native value 7.6 Å.

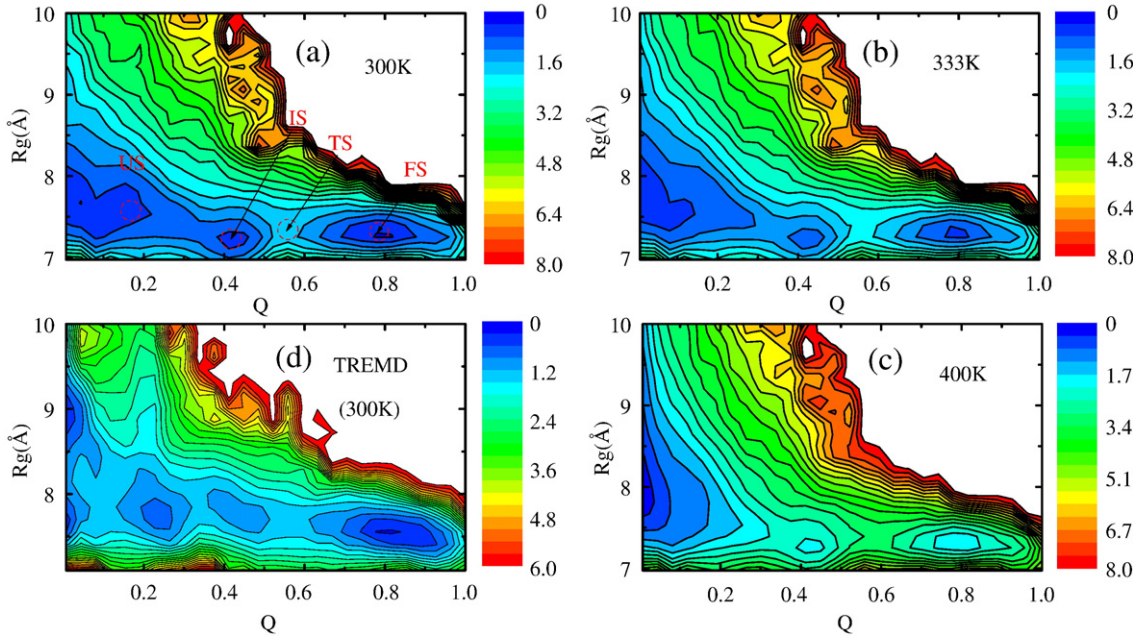

Fig. 4. The contour plots of the free-energy landscape at various temperatures projected onto the reaction coordinates Q, and radius of gyration R_g. US, IS, TS and FS represent the unfolded state, intermediate state, transition state and folded state, respectively. (a), (b) and (c) are constructed by REMshH data and (d) is built from standard REMD simulation.

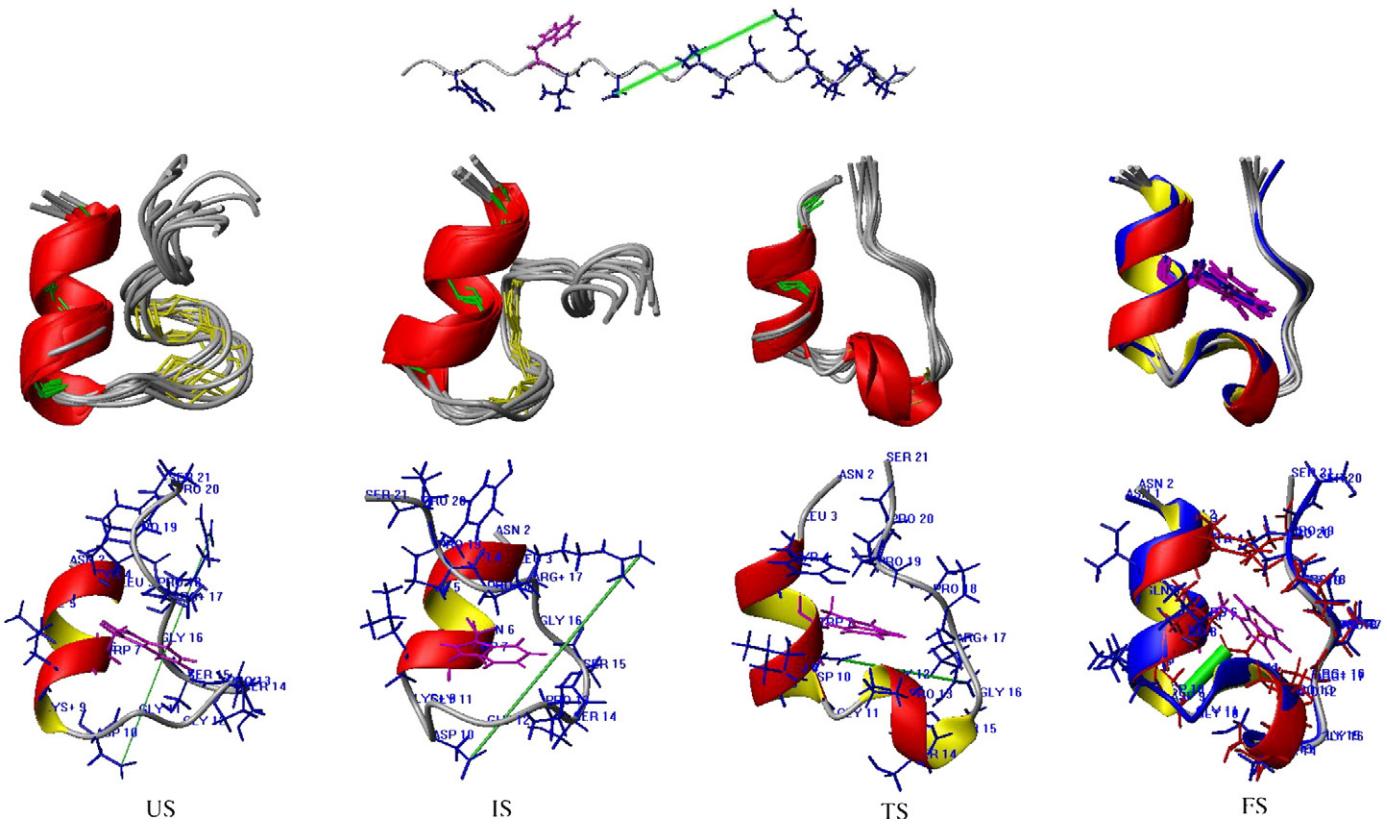

Fig. 5. Representative structures of the most populated clusters for unfolded state (US), intermediate state (IS), transition state (TS) and folded state (FS, here we only show those with $Q>0.80$). The main chains are shown in ribbon, sidechain of structurally important Tyr3, Leu7, Lys8, Ser13, Ser14, Gly15, Pro17, Pro18, Pro19 and Ser20 residues are shown in neon, as well as the sidechains of Asp9 and Arg16 which construct a salt bridge. The green line denotes the salt-bridge distance. The α-helix (3–9) is shown in red and the 3_{10}-helix (12–15) in blue (Here, the residue numbering is 2–21.). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 5 (IS) represents the populated intermediate-state (IS) ensemble. In the NMR structures, the hydrophobic core consists of residues Tyr3, Trp6, Leu7, Pro12, Pro17, Pro18 and Pro19. The N-terminus and the C-terminus are nearly antiparallel to each other. In IS ensemble the relative orientation of C-terminus to the N-terminus is different from an antiparallel configuration. The distorted conformation is caused mainly by the intimate mis-stacking of Trp6 with Pro17 and Leu7 with Pro18. The sidechains of Trp6, Leu7, Pro12, Pro17 and Pro18 form a misfolded hydrophobic core. At the same time, it is noted that the sidechain of Tyr3, which belongs to the native hydrophobic core, is oriented outside and exposed to the solvent. The α-helix segment, residues 2–8, is already formed. While the 3_{10}-helix is still absent and the salt-bridge distance is large. As quantitatively listed in Table 2, the length of α-helix and 3_{10}-helix is 6.0±1.5 and 0.0, respectively; SB9-16 is 14.9±1.4 Å; it is of interest to note that the total contact number (including non-native contacts) of Trp6 calculated from the representative structure is 335±21, significantly larger than other representative structures of other basins of attractions. This indicates that the misfolded cage around Trp6 is more compact than the native one. On the other hand the solvent-accessible surface of Trp6, SAS_{Trp6}, is larger than that of native state, which demonstrates more solvent exposure of Trp6 in IS than in NS. Thus our simulations provide a direct evidence for the recent experiments [36,37], which suggested that a molten globule-like intermediate existed in the folding process.

In the following, we take a look at the transition-state (TS) ensemble between the intermediate state and the folded state. The TS is located by spotting a small region with the highest free energy between IS and FS on the free-energy surface. Fig. 5 (TS) shows the representative structure taken out of the TS ensemble. In general, the architecture resembles the native structure with both the α-helix and 3_{10}-helix mostly folded. However, the N-terminal and C-terminal are too close to each other because of the strong interactions between sidechains, such as Leu2 and Pro19. Additionally, the salt-bridge distance decreases to 9.6±1.9 Å (Table 2) although is still larger than the native value.

Our detailed structural models indicate a synergetic formation of α-helical secondary structure and hydrophobic collapse. Moreover our model could provide a possible explanation of a recent experimental mutation result where a single point mutation, I4G, causes Trpcage to switch from a non-two-state folder to a two-state one [37]: although it does not involve in the misfolded core, replacement of I4 with a glycine will destabilize the α-helix, and in turn attenuate the IS. The α-helix formation could not be the rate-limiting step, instead the structural rearrangement from IS to FS in which the misfolded hydrophobic core centered on Trp6 is refolded turns out to be the limiting point. This finding is consistent with recent kinetic study using Amber03 force field in an implicit solvation environment [33].

3.5. Thermodynamic stability

The temperature dependence of the Trpcage unfolded/folded population at various temperatures has been determined by NMR CSD, CD, Tryptophan fluorescence and fluorescence correlation spectroscopy experiments. The melting transition temperature was found to be around 315 K [30,35,37]. It is noted that the temperature dependence of peptides or proteins deviates significantly from the experimental data in the previous simulation studies, no matter what force fields were used: OPLSAA, Amber or CHARMM. Here we would like to investigate the temperature dependence of Trpcage using the REMhH method to see whether our method can reproduce thermal unfolding feature of Trpcage. Fig. 6 shows the computational observables as a function of

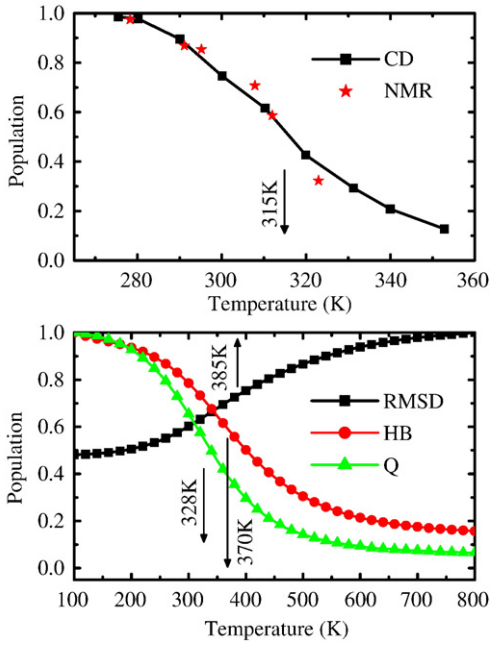

Fig. 6. Melting curves of simulated observables as well as experimental curves from NMR CSD and CD for comparison. The melting temperature T_m is marked by the vertical arrows. The populations of RMSD, HB and Q at various temperatures are scaled for convenience of comparisons.

temperature for the 142-ns simulation, which is scaled for convenience. The experimental unfolding molar fractions from NMR CSD and CD spectrum are also shown for comparison [30]. The theoretical melting point was estimated as the temperature where the simulated observable reaches the midpoint of those at low and high temperatures. It can be seen that when monitored by RMSD, HB and Q, the T_ms span a temperature range from 328 K to 385 K, which is larger than the experimental T_m=315 K, but smaller than the previous reported T_m that is larger than 400 K [32,52]. The difference between the present study and previous simulations may originate from the diverse simulation conditions. In the simulation study of Zhou [32], the initial structure was taken from NMR model and the simulation time was only 5 ns per replica. In Pitera and Swope's study [52], AMBER94 force field combined with implicit water model was used. Our simulation began with fully extended initial conformations in the presence of explicit water molecules and lasted for 142 ns. We found that the results obtained in the 71-ns simulation gave a similar melting temperatures comparing with the results shown in Fig. 6. It is noted that recent works employing modified AMBER force field and improved implicit water models [53,54] have demonstrated the simulated T_m (around 315 K) quite close to the experimental one.

3.6. Salt bridge forms early or late?

One of the most interesting aspects of molecular dynamics simulation is that the structural changes can be tracked with the highest tempo-space resolution, a feat difficult to achieve with current experimental methods. Here we explore the role of the salt bridge formed between Asp9 and Arg16 in the folding kinetics of Trpcage. A salt bridge is assumed to be formed if its distance is less than 5.2 Å. The choice of distance cutoff is based on the distribution plot of salt-bridge distances over the ensemble trajectory at room temperature in the 142-ns simulation. The salt-bridge distance as a function of Q obtained from the simulation trajectory at room temperature for three simulations (S1, S2 and S3) is shown in Fig. 7. For better discussion of the formation time of the salt bridge, the formation probability of α-helix is also shown for comparisons. Fig. 7 (S1) shows that the salt-bridge distance is larger than 6.6 Å as long as Q<0.50, which indicates that the salt bridge is not formed even the folding of Trpcage is 50% completed. While the α-helix is largely formed at this time. In fact, the salt bridge is formed until Q~0.6 as the distance is reduced to 5 Å (see Fig. 7S1). However, both simulation results of S2 and S3 show that the salt bridge between Asp9 and Arg16 is largely formed before the significant formation of α-helix. The different formation sequences of salt bridge found in simulations suggest that diverse folding pathways may exist in the folding of Trpcage.

Regarding the role of the salt bridge played in folding dynamics there exists disagreement between different theoretical studies. A recent study which eliminated the Asp-9–Arg-16 salt bridge by a theoretical mutation using all-atom Go model simulations found a significant decrease in the folding rate [55]. However, the simulation studies by Pande and co-workers through the Folding@Home distributed computing project [56], and by Duan and coworkers [31,33] suggested that salt bridge plays little roles in the folding kinetics of Trpcage since its formation is in the late stage. Recently, Juraszek and Bolhuis investigated the kinetic pathway of Trpcage with a novel transition path sampling method and found that Trpcage folds primarily via a pathway forming the salt bridge before helix formation, while the second pathway occurs by first forming the helix [34]. Our simulation results are consistent with what found by Juraszek and Bolhuis that alternative pathways exist in the folding of Trpcage.

3.7. Folding sequence

Fig. 8 shows the secondary structure content as a function of time for the replica folding trajectory. Secondary structure assignments are

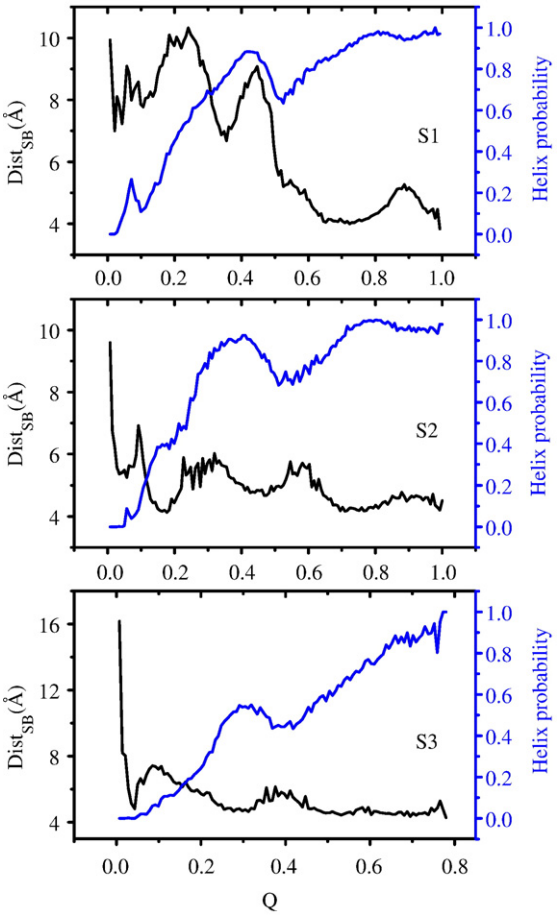

Fig. 7. The salt-bridge distance and formation probability of α-helix as a function of Q at room temperature for three simulations: S1, S2 and S3. 1 means the folding of helix is completed and 0 indicates no existence of helix.

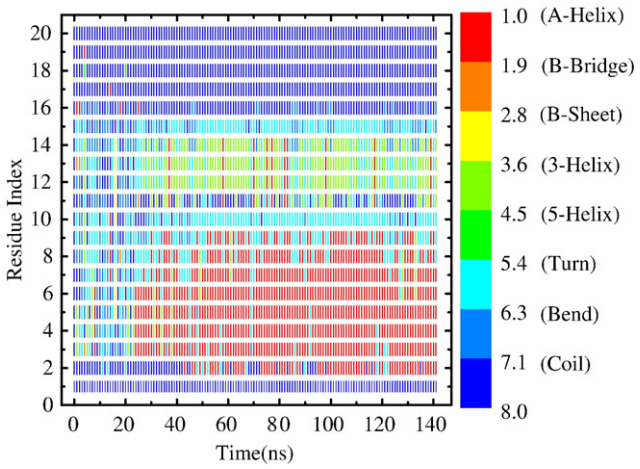

Fig. 8. The secondary structure content as a function of time for the replica folding trajectory. Secondary structure assignments were based on the DSSP algorithm [57]. The corresponding color bar denotes different secondary structures.

based on the DSSP algorithm [57]. The simulation was started from a fully extended conformation. So at t=0 ns all 20 residues are unstructured and no regular secondary structure exists. Such situation lasts until t=23 ns. After that the α-helical segment is largely formed; in particular, residues Y3, I4, Q5 and W6 retain the α-helical configuration throughout the remainder of the simulation. While L2, L7 and K9 convert to the native helical structure later and are less stable because transformation to random coils occurs more frequently. The 3_{10}-helix stably appears after t=27 ns. The residue G11 mostly adopts random coil structure although transformation into 3_{10}-helix structure does exist. Thus, from the folding trajectory of secondary structures shown in Fig. 8, one can find that formation of the 3_{10}-helix is initiated after the formation of α-helical segment which is in good agreement with Duan and coworkers' simulation results obtained by using AMBER03 force fields and GB/SA solvent model [31]. In addition, from Fig. 8, one can also see that G10 and G15 mostly adopt turn structure during the simulation, completely consistent with the NMR structures.

The above folding events about α-helix and 3_{10}-helix are observed in the replica folding trajectory. In the following, the formation sequence of α-helix and 3_{10}-helix is further studied by analyzing the ensemble trajectory at room temperature. Fig. 9 shows the probability of two helical structures as a function of Q for the 142-ns simulation.

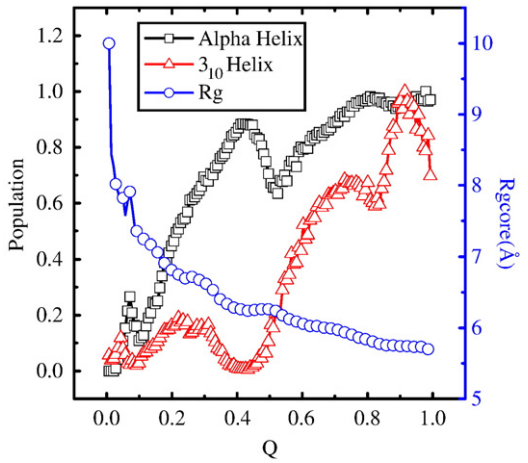

Fig. 9. The probability of α-helix (squares) and 3_{10}-helix (triangles) as a function of fraction of native contacts, Q. The curves are obtained by averaging over the whole ensemble at room temperature. 1 means the folding of helix is completed and 0 indicates no existence of helix. The radius of gyration for hydrophobic-core residues as a function of Q (circles).

The curves obtained in the two additional simulations are wholly similar and not shown. At Q~0.41, the probability of α-helix is as high as 0.9, while that of 3_{10}-helix is nearly equal to 0.0. Thus, one can clearly find that the α-helix forms early, followed by the 3_{10}-helix, totally consistent with the results obtained from analyzing the replica trajectory. The data for the ensemble analysis is taken from the room temperature trajectory with simulation t>50 ns. While the replica trajectory is analyzed before t=50 ns. The fact that the two independent data sets provide a consistent result is amazing which demonstrates that the 'continuous' replica trajectory can also be analyzed to obtain qualitative information about the conformational changes during folding process.

Also shown in this figure is the evolution of the radius of gyration of residues which consists of the hydrophobic core, $R_{g_{core}}$, as a function of Q. There is a sharp drop of $R_{g_{core}}$ as Q increases from 0 to 0.1. Then $R_{g_{core}}$ gradually decreases to 5.7 Å as Q increases to 1.0. In the meantime the α-helical segment also takes shape. Our data indicates that the formation of α-helix and the collapses of hydrophobic core are simultaneous.

3.8. Folding feature of α-helix

It is also of interest to take a closer look at the detailed folding feature of the α-helix since it is an important secondary structure and a key portion of Trpcage. The α-helix of Trpcage comprise 7 residues: L2, Y3, I4, Q5, W6, L7 and K8. Fig. 10 shows the α-helical probability of

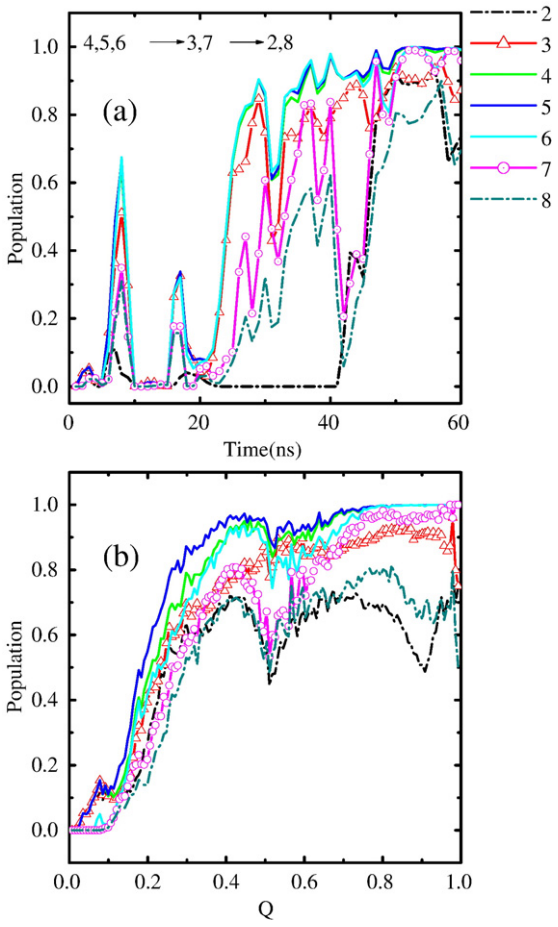

Fig. 10. (a) The helical probability of the 7 residues as a function of time. The probability is obtained from the replica folding trajectory, which is calculated averaging over every 2000 data points with 1000 data points overlapping. (b) The helical probability as a function of Q, which is derived from the last 92ns ensemble trajectory at room temperature. Numbers denote the index of 7 residues. The rough folding sequence is displayed on the plot.

these 7 residues for the 142-ns simulation. Value 1 means the residue adopts an α-helical structure. The assembly of I4, Q5 and W6 into α-helical structure preformed around 23 ns; followed by Y3 and L7. Finally, the α-helical structure of L2 and K8 was formed. Fig. 10b also displays the results of helical probability as a function of Q, which is derived from the ensemble trajectory at room temperature. At $Q\sim0.41$, I4, Q5 and W6 fold into helix structure; then at $Q\sim0.60$, the assembly of Y3 and L7 is completed; finally, L2 and K8 become helical. It is necessary to point out that the two additional simulations also gave similar results. Therefore, our simulations show that the formation of α-helix is initiated from the middle portion and extends outwards to two termini in the case of Trpcage.

3.9. Water dynamics

In this section, we investigate the behaviors of water molecules during the folding process of Trpcage to provide some insights into the role of water molecules played in protein folding. Here we focus on the water number in the vicinity of the Trpcage molecule. Only water molecules within a 3.0 Å shell around Trpcage is counted. It is noted that the folding process of Trpcage is accompanied by the repulsion of water molecules out of the miniprotein. Fig. 11a shows the change of the number of water molecules close to the protein surface with RMSD. One can see that about 100 water molecules are close to the unfolded Trpcage with RMSD larger than 12 Å. The number sharply drops from 100 to 70 as RMSD decreases from 12 Å to 9 Å. This corresponds to the cooperative and fast collapse process in the early folding stage of Trpcage, which is consistent with previous experimental observations [58]. During this process, the hydrophobic core is not formed as the number of water molecule around the key residue Trp6 of the hydrophobic core is large and has a negligible change (see Fig. 11b). Trp6 is the center of the hydrophobic core as it is caged by the 3_{10}-helix group, the side chains of residues Tyr3, Leu7, Arg16, together with the C-terminal polyproline II stretch. As RMSD further decreases from 9 Å to 1.4 Å, the number of water molecule around Trpcage gradually reduces to 60, while the number of water molecule located around Trp6 changes from 8 to 2. Thus, water molecule still exists inside the interior of Trpcage at the folded state.

Furthermore, the number of water molecules around each amino acid in Trpcage is calculated. Fig. 11c displays the number of water molecules as a function of residue index which is averaged over the whole ensemble at room temperature. One can clearly see that the number of water molecules around Lys8, Arg16 and Ser20 is the largest. The three resides are all polar amino acids: Lys8 and Arg16 are polar amino acids with charged sidechains, Ser20 is a polar amino acid with a non-charged sidechain. It is also noted that the numbers of water molecules around Gly10 and Gly11 are the smallest.

We also calculated the number of water molecules around each amino acid, which is only averaged over the unfolded and folded ensemble (shown in Fig. 11d). In an overall view, the number of water molecules around each amino acid at unfolded state (filled black symbol) is larger than that at folded state (holly symbol). However, the number around Lys8 and Ser20 nearly does not vary because the two residues are entirely exposed to the solvent. Among the 20 residues, the number of water molecules around Trp6 has the largest change comparing the value of 10 at the unfolded state with that of 2 at the folded state. This indicates a highest degree of dehydration in the hydrophobic core during the folding of Trpcage from unfolded structures into folded structures.

4. Conclusion

In this work the *ab initio* folding of a miniprotein, Trpcage in explicit solvent is studied by employing the replica-exchange molecular dynamics simulation with scaled hybrid Hamiltonian. The free-energy surface constructed by the REMshH simulations is similar to the one built by a standard REMD simulation. The folded structures of Trpcage have root mean squared deviations less than 2.0 Å with respective to NMR structures considering all heavy atoms. To our knowledge it is the first study to successfully fold Trpcage from fully extended structures using OPLSAA force field. It is also of interest to note that the 3_{10}-helix turn forms well in the folded structures which is difficult to achieve in previous simulation studies of Trpcage. In folded states the majority of experimentally observed NMR NOE restraints are satisfied. Thus, our method cannot only fold peptides with mainly β structures such as β-hairpins [23,24], but also the miniprotein with mainly α structures.

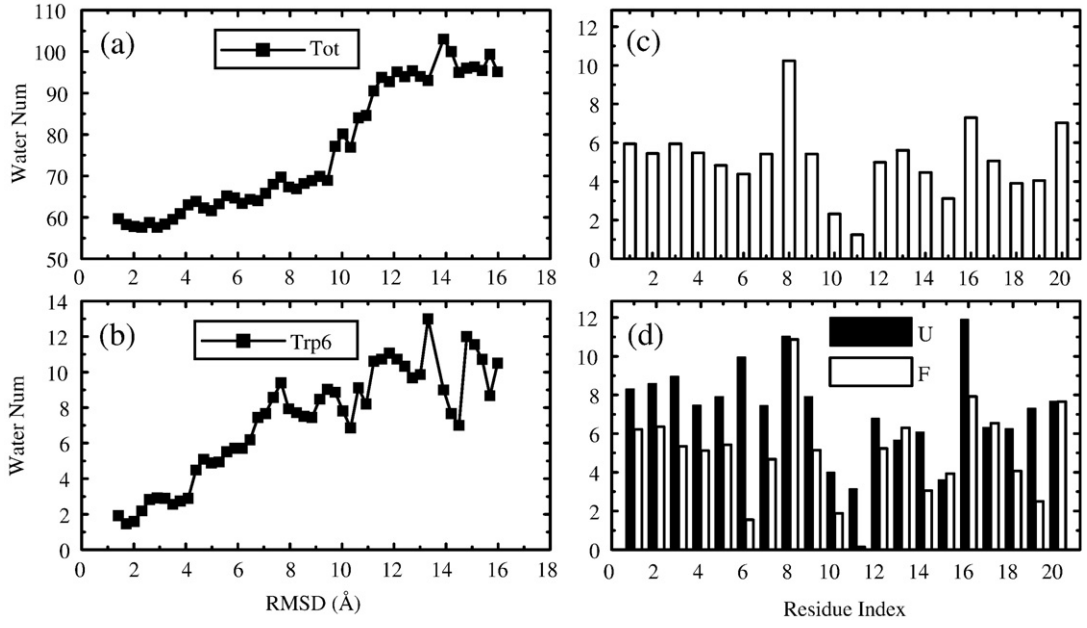

Fig. 11. The average number of water molecules in the vicinity of each amino acid in Trpcage, which is averaged over the entire ensemble (a), and average over the unfolded and folded ensemble (b) at room temperature. The average number of water molecules in the vicinity of Trpcage (c) and of the aromatic residue Trp6 (d) as a function of RMSD. Here, the water molecules within a 3.0 Å shell around Trpcage is counted.

In addition, detailed folding mechanisms of Trpcage is also explored. By investigating the desolvation behavior of Trpcage, a pre-existing hydrophobic collapse in the unfolded state is detected which is consistent with previous experimental observation. At room temperature, an intermediate state featured by a largely formed α-helix and an absence of 3_{10}-helix is obviously populated on the folding free-energy landscape, which indicates a non-two-state folding behavior. Our simulations provide a direct evidence for recent experimental results measured by UV-resonance Raman spectroscopy [36] and fluorescent correlation spectroscopy [37], which suggested that folding of Trpcage is not a simple two-state process. The detailed folding sequence of key secondary structures is also probed, revealing a clear folding pathway of Trpcage. Our results based on three simulations suggest two potential kinetic pathway of salt bridge: one pathway forming the salt bridge before α-helix formation, while the second pathway occurs by first forming the α-helix followed by salt bridge. In addition, the interesting property of water dynamics is investigated. The highest degree of dehydration occurs within the hydrophobic core as Trpcage folds from the unfolded state into the folded state. A correlation between water dynamics and protein residues is revealed that the water dynamics is dependent on the essential physicochemical property of residues.

Acknowledgments

The support of research grants, URC (RG65/06), from Nanyang Technological University and academic research fund (AcRF) Tier 2 by the MOE (T206B3210RS) is acknowledged. The simulations were performed on the supercomputer of Bioinformation Research Center in NTU and the Frankfurt Center of Scientific Computing, which are acknowledged by the generous allocation of CPU time. The help from Y. Yang for modifying GROMACS source code is highly acknowledged.

References

[1] Y. Duan, P.A. Kollman, Science 282 (1998) 740–744.
[2] J.M.e.a. Schmidt, J. Am. Chem. Soc. 115 (1993) 8747–8756.
[3] B.D. Bursulaya, C.L. Brooks, J. Phys. Chem. B. 104 (2000) 12378–12383.
[4] W.Y. Yang, J.W. Pitera, W.C. Swope, M. Gruebele, J. Mol. Biol. 336 (2004) 241–251.
[5] I.A. Hubner, E.J. Deeds, E.I. Shakhnovich, Proc. Natl. Acad. Sci. U. S. A. 103 (2006) 17747–17752.
[6] G. Jayachandran, V. Vishal, V.S. Pande, J. Chem. Phys. 124 (2006) 164902.
[7] A. Mitsutake, Y. Sugita, Y. Okamoto, Biopolymers 60 (2001) 96–123.
[8] B.A. Berg, T. Neuhaus, Phys. Rev. Lett. 68 (1992) 9–12.
[9] A.P. Lyubartsev, A.A. Martsinovski, S.V. Shevkunov, P.N. Vorontsov-Velyaminov, J. Chem. Phys. 96 (1992) 1776–1783.
[10] U.H.E. Hansmann, Chem. Phys. Lett. 281 (1997) 140–150.
[11] A.E. Garcia, J.N. Onuchic, Proc. Natl. Acad. Sci. U. S. A. 100 (2003) 13898–13903.
[12] R. Zhou, B.J. Germain, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 14931–14936.
[13] J. Zhang, M. Qin, W. Wang, Proteins 62 (2006) 672–685.
[14] Y.G. Mu, L. Nordenskiold, J.P. Tam, Biophys. J 90 (2006) 3983–3992.
[15] H. Fukunishi, O. Watanabe, S. Takada, J. Chem. Phys. 116 (2002) 9058–9067.
[16] P. Liu, B. Kim, R.A. Friesner, B.J. Berne, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 13749–13754.
[17] T.Z. Lwin, R. Luo, J. Chem. Phys. 123 (2005) 194904.1–194904.10.
[18] X. Cheng, G. Cui, V. Hornak, C. Simmerling, J. Phys. Chem. B 109 (2005) 8220–8230.
[19] E. Lyman, F.M. Ytreberg, D.M. Zuckerman, Phys. Rev. Lett. 96 (2006) 028105.1–028105.4.
[20] A. Okur, L. Wickstrom, M. Layten, R. Geney, K. Song, V. Hornak, C. Simmerling, J. Chem. Theory. Comput. 2 (2006) 420–433.
[21] Y.G. Mu, Y. Yang, W.X. Xu, J. Chem. Phys. 127 (2007) 084119.
[22] W.X. Xu, T.F. Lai, Y. Yang, Y.G. Mu, J. Chem. Phys. 128 (2008) 175105.
[23] Y.G. Mu, Y. Yang, W.X. Xu, J. Theor. Comput. Chem. 7 (2008) 177–187.
[24] W.X. Xu, Y. Yang, Y.G. Mu, L. Nordenskiold, Mol. Simul. 34 (2008) 575–590.
[25] G.A. Kaminski, R.A. Friesner, J. Tirado-Rives, W.L. Jorgensen, J. Phys. Chem. B 105 (2001) 6474.
[26] T.L.e.a. Blundell, Proc. Natl. Acad. Sci. U. S. A 78 (1981) 41754179.
[27] C.J. McKnight, P.T. Matsudaira, P.S. Kim, Nat. Struct. Biol. 4 (1997) 180184.
[28] A. Zarrinpar, W.A. Lim, Nat. Struct. Biol. 7 (2000) 611613.
[29] A.G. Cochran, N.J. Skelton, M.A. Starovasnik, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 55785583.
[30] J.W. Neidigh, R.M. Fesinmeyer, N.H. Andersen, Nat. Struct. Biol. 9 (2002) 425–430.
[31] S. Chowdhury, M.C. Lee, G.M. Xiong, Y. Duan, J. Mol. Biol. 327 (2003) 711–717.
[32] R.H. Zhou, Proc. Natl. Acad. Sci. U. S. A. 100 (2003) 13280–13285.
[33] S. Chowdhury, M.C. Lee, Y. Duan, J. Phys. Chem. B 108 (2004) 13855–13865.
[34] J. Juraszek, P.G. Bolhuis, Proc. Natl. Acad. Sci. U. S. A. 103 (2006) 15859–15864.
[35] L.L. Qiu, S.A. Pabit, A.E. Roitberg, S.J. Hagen, J. Am. Chem. Soc. 124 (2002) 12952–12953.
[36] Z. Ahmed, I.A. Beta, A.V. Mikhonin, S.A. Asher, J. Am. Chem. Soc. 127 (2005) 10943–10950.
[37] H. Neuweiler, S. Doose, M. Sauer, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 16650–16655.
[38] H.J.C. Berendsen, D. van der Spoel, R. van Drunen, Comput. Phys. Commun. 91 (1995) 43–56.
[39] W.L. Jorgensen, D.S. Maxwell, J. Tirado-Rives, J. Am. Chem. Soc. 118 (1996) 11225–11236.
[40] A.M. Ferrenberg, R.H. Swendsen, Phys. Rev. Lett. 63 (1998) 1195–1198.
[41] W.X. Xu, J. Wang, W. Wang, Proteins 61 (2005) 777–794.
[42] C.D. Snow, L. Qiu, D. Du, F. Gai, S.J. Hagen, V.S. Pande, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 4077–4082.
[43] K.Y. Sanbonmatsu, A.E. García, Proteins 46 (2002) 225–234.
[44] L. Zhan, J. Chen, W.K. Liu, Proteins 66 (2007) 436.
[45] A.R. Dinner, T. Lazaridis, M. Karplus, Proc. Natl. Acad. Sci. U. S. A. 96 (1999) 9068–9073.
[46] A.E. García, K.Y. Sanbonmatsu, Proteins 42 (2001) 345–354.
[47] B. Zagrovic, E.J. Sorin, V.S. Pande, J. Mol. Biol. 313 (2001) 151–169.
[48] R. Zhou, B.J. Berne, R. Germain, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 14931–14936.
[49] R. Zhou, B.J. Berne, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 12777–12782.
[50] P.H. Nguyen, Y.G. Mu, G. Stock, Proteins 60 (2005) 485–494.
[51] M. Gruebele, C. R. Biologies 328 (2005) 701–712.
[52] J.W. Pitera, W. Swope, Proc. Natl. Acad. Sci. U. S. A 100 (2003) 7587–7592.
[53] R. Geney, M. Layten, R. Gomperts, V. Hornak, C. Simmerling, J. Chem. Theory Comput. 2 (2006) 115–127.
[54] E. Kim, S. Jang, Y. Pak, J. Chem. Phys. 127 (2007) 145104–145109.
[55] A. Linhananta, J. Boer, M.L.J. Chem, Phys. 122 (2005) 114901.
[56] C.D. Snow, B. Zagrovic, V.S. Pande, J. Am. Chem. Soc. 124 (2002) 14548–14549.
[57] W. Kabsch, C. Sander, Biopolymers 22 (1983) 2577–2637.
[58] K.H. Mok, L.T. Kuhn, M. Goez, I.J. Day, J.C. Lin, N.H. Andersen, P.J. Hore, Nature 447 (2007) 106–109.
[59] C. Simmerling, B. Strockbine, A.E. Roitberg, J. Am. Chem. Soc. 124 (2002) 11258–11259.
[60] P. Carnevali, G. Toth, G. Toubassi, S.N. Meshkat, J. Am. Chem. Soc. 125 (2003) 14244–14245.
[61] M. Ota, M. Ikeguchi, A. Kidera, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 17658–17663.
[62] A. Schug, T. Herges, W. Wenzel, Phys. Rev. Lett. 91 (2003) 158102.
[63] A. Schug, W. Wenzel, Europhys. Lett. 67 (2004) 303–313.
[64] A. Schug, W. Wenzel, J. Chem. Phys. 122 (2005) 194711.
[65] A. Schug, T. Herges, A. Verma, K.H. Lee, W. Wenzel, Chem. Phys. Chem. 6 (2005) 2640–2646.
[66] P.J. Steinbach, Proteins 57 (2004) 665–677.
[67] F. Ding, S.V. Buldyrev, N.V. Dkholyan, Biophys. J. 88 (2005) 147–155.
[68] A. Irbäck, S. Mohanty, Biophys. J. 88 (2005) 1560–1569.
[69] R. Koradi, M. Billeter, and K. Wuthrich. J. Mol. Graphics. 14(51–55) 66–93.