J|A|C|S
ARTICLES

Published on Web 07/22/2010

Folding of a Helix at Room Temperature Is Critically Aided by Electrostatic Polarization of Intraprotein Hydrogen Bonds

Li L. Duan,[†,‡] Ye Mei,[*,‡] Dawei Zhang,[§] Qing G. Zhang,[†] and John Z. H. Zhang[*,‡,||]

College of Physics and Electronics, Shandong Normal University, Jinan 250014, China, State Key Laboratory of Precision Spectroscopy and Department of Physics, Institute of Theoretical and Computational Science, East China Normal University, Shanghai 200062, China, Division of Chemistry and Biological Chemistry, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371, and Department of Chemistry, New York University, New York, New York 10003

Received April 1, 2010; E-mail: ymei@phy.ecnu.edu.cn; john.zhang@nyu.edu

Abstract: We report direct folding of a 17-residue helix protein (pdb:2I9M) by standard molecular dynamics simulation (single trajectory) at room temperature with implicit solvent. Starting from a fully extended structure, 2I9M successfully folds into the native conformation within 16 ns using adaptive hydrogen bond-specific charges to take into account the electrostatic polarization effect. Cluster analysis shows that conformations in the native state cluster have the highest population (78.4%) among all sampled conformations. Folding snapshots and the secondary structure analysis demonstrate that the folding of 2I9M begins at terminals and progresses toward the center. A plot of the free energy landscape indicates that there is no significant free energy barrier during folding, which explains the observed fast folding speed. For comparison, exactly the same molecular dynamics simulation but carried out under existing AMBER charges failed to fold 2I9M into native-like structures. The current study demonstrates that electrostatic polarization of intraprotein hydrogen bonding, which stabilizes the helix, is critical to the successful folding of 2I9m.

Correct prediction of protein structure from a sequence is one of the most challenging scientific problems in the 21st century. Despite extensive effort that has been devoted to protein folding, it remains one of the most difficult problems in physics, chemistry, and biology for decades. In addition to its intrinsic scientific value to the understanding of the fundamental biology, there is also significant biological importance in protein folding. For example, many diseases are related to protein misfolding, including senile dementia, Parkinson's disease, and mad cow disease.[1−3] Thus, understanding the pathway and mechanism of protein folding is of both fundamental importance and practical utility. Thanks to increasing computer power and improved theoretical methods, correct predictions of folded structures of a few "fast folders"[4−29] and also some relatively large proteins[30] have been reported. In most of these folding studies, complete phase space sampling is performed in order to find the native structure, which limits the applications to very small proteins. In addition, the success of folding is often

[†] Shandong Normal University.
[‡] East China Normal University.
[§] Nanyang Technological University.
[||] New York University.

(1) Dobson, C. M. *Trends Biochem. Sci.* **1999**, *24*, 329–332.
(2) Muchowski, P. J. *Neuron* **2002**, *35*, 9–12.
(3) Slepoy, A.; Singh, R. R. P.; Pazmandi, F.; Kulkarni, R. V.; Cox, D. L. *Phys. Rev. Lett.* **2001**, *87*, 058101−104.
(4) Baumketner, A.; Shea, J. E. *J. Mol. Biol.* **2007**, *366*, 275–285.
(5) Berg, B. A.; Neuhaus, T. *Phys. Lett. B* **1991**, *267*, 249.
(6) Berg, B. A.; Neuhaus, T. *Phys. Rev. Lett.* **1992**, *68*, 9.
(7) Brooks, C. I. *Acc. Chem. Res.* **2002**, *35*, 447–454.
(8) Chowdhury, S.; Lei, H. X.; Duan, Y. *J. Phys. Chem. B* **2005**, *109*, 9073–9081.
(9) Daggett, V. *Acc. Chem. Res.* **2002**, *35*, 422–449.
(10) Duan, Y.; Kollman, P. A. *Science* **1998**, *282*, 740–744.
(11) Duan, Y.; Wang, L.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 9897–9902.
(12) Eleftheriou, M.; Germain, R. S.; Royyuru, A. K.; Zhou, R. *J. Am. Chem. Soc.* **2006**, *128*, 13388–13395.
(13) Hansmann, U. H. E. *Chem. Phys. Lett.* **1997**, *281*, 140–150.
(14) Juraszek, J.; Bolhuis, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 15859–15864.
(15) Lei, H.; Wang, Z.; Wu, C.; Duan, Y. *J. Chem. Phys.* **2009**, *131*, 165105−1165105−7.
(16) Lei, H.; Wu, C.; Liu, H.; Duan, Y. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 4925–4930.
(17) Mitsutake, A.; Sugita, Y.; Okamoto, Y. *Biopolymers* **2001**, *60*, 96–123.
(18) Nelson, E. D.; Grishin, N. V. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 1489.
(19) Pande, V. S.; Baker, I.; Chapman, J.; Elmer, S. P.; Khaliq, S.; Larson, S. M.; Rhee, Y. M.; Shirts, M. R.; Snow, C. D.; Sorin, E. J.; Zagrovic, B. *Biopolymers* **2003**, *68*, 91–109.
(20) Paschek, D.; Nymeyer, H.; Garcia, A. E. *J. Struct. Biol.* **2007**, *157*, 524–533.
(21) Periole, X.; Mark, A. E. *J. Chem. Phys.* **2007**, *126*, 014903.
(22) Scheraga, H. A.; Khalili, M.; Liwo, A. *Annu. Rev. Phys. Chem.* **2007**, *58*, 57–83.
(23) Simmerling, C.; Strockbine, B.; Roitberg, A. *J. Am. Chem. Soc.* **2002**, *124*, 11258–11259.
(24) Swendsen, R. H.; Wang, J. S. *Phys. Rev. Lett.* **1986**, *57*, 2607–2609.
(25) Wickstrom, L.; Okur, A.; Song, K.; Hornak, V.; Raleigh, D. P.; Simmerling, C. L. *J. Mol. Biol.* **2006**, *360*, 1094–1107.
(26) Xu, W. X.; Lai, T. F.; Yang, Y.; Mu, Y. G. *J. Chem. Phys.* **2008**, *128*, 175105.
(27) Yoda, T.; Sugita, Y.; Okamoto, Y. *Proteins: Struct., Funct., Bioinf.* **2007**, *66*, 846–859.
(28) Zhou, R. H. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13280–13285.
(29) Zhou, R. H.; Berne, B. J.; Germain, R. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 14931–14936.
(30) Liwo, A.; Khalili, M.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 2362–2367.

 © 2010 American Chemical Society

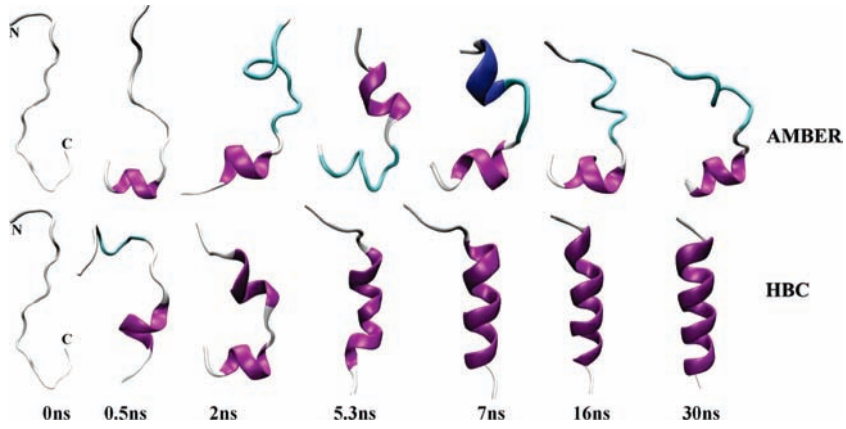

Figure 1. Snapshots of intermediate structures of the peptide at different simulation times using AMBER (upper) and adaptive HBC (lower). Here the N-terminal is always on the top. α helix: purple, coil: white, turn: cyan.

unpredictable due to various issues such as the inaccuracy of the existing force field used, the length of simulation, etc. By nature, the folding of protein does not sample the entire phase space to "determine" where the native state is. Instead, protein is guided by physical forces to efficiently find its native structure without exploring the entire phase space. Thus studying the direct folding process of protein at room temperature without phase space sampling should help provide much needed insight into the folding dynamics and elucidate the mechanism of folding such as the formation of secondary structures. Recent ultrafast spectroscopic experiments showed that the formation of an α helix occurs in 200 ns[31] and β hairpins in 1−10 μs.[32] Thus small proteins become attractive targets for studying the folding mechanism.[33]

The hydrogen bond network plays an essential role in secondary structures of proteins. Recent work has demonstrated that electrostatic polarization stabilizes intraprotein hydrogen bonds and proteins' local structures near the native state of proteins.[34−37] It is therefore conjectured that electrostatic polarization in hydrogen bonding may also play an important role in folding dynamics. In this study we carry out a direct folding study for a small protein (PDB 2I9M) by standard MD simulation at room temperature (a single trajectory study) with a generalized Born model for implicit solvent. To build into simulation the correct polarization effect in hydrogen bonding, we employ a discrete on-the-fly charge fitting scheme in the simulation, in which the corresponding atomic charges of hydrogen bond donors and acceptors are refitted from fragment quantum mechanical calculation in the instantaneous environment whenever the corresponding hydrogen bonds are either newly formed or broken. 2I9M[38] is a recently designed 17-residue α-helix (residues 1−14) peptide and folds itself without the assistance of disulfide bonds. Due to its small size and well-defined secondary structure, it has been an attractive system for studying protein folding. Recently, replica exchange molecular dynamics (REMD) simulations have been carried out by Lin et al.[39] and Kim et al.,[40] each performed as an ∼60 ns simulation per replica to find the folded structure by phase space sampling. However, there is no report of a direct folding study of this protein based on a single room temperature MD trajectory. In the present work, single MD trajectory at room temperature is performed under both a standard AMBER force field and the discrete on-the-fly charge fitting scheme to study the natural folding process, in particular to investigate the effect of electrostatic polarization of hydrogen bonds on folding dynamics.

Folding of 2I9M at Room Temperature

In order to closely mimic the natural folding process as much as possible, we started from a fully extended structure of 2I9M and performed MD simulation at room temperature with a generalized Born solvation model (igb5) of Onufriev, Bashford, and Case.[41] The salt concentration is set to 0.2 M. After energy minimization, the system is heated up to 300 K in 100 ps, followed by a 30 ns production MD run with a time step of 2 fs. The SHAKE algorithm[42] is employed to fix all bonds involving hydrogen atoms. The Langevin dynamics[43] with a collision frequency of 1.0 ps^{-1} is applied to regulate the temperature. The dielectric constants of the protein interior and of the solvent are set to, respectively, 1.0 and 78.5. The trajectory is saved every 1 ps, and a total of 30 000 snapshots are taken in a production run for detailed analysis. The above production MD simulation is performed using two different techniques, one with the straightforward AMBER force field and the other with atomic charges replaced by adaptive hydrogen bond-specific charge (HBC) for hydrogen bonds discussed below.

In order to gain insight into folding dynamics, we monitored snapshots of the intermediate states along with simulation time. First we examine the folding dynamics in MD simulation under the standard AMBER force field. It is apparent from Figure 1

(31) Williams, S.; Causgrove, T. P.; Gilmanshin, R.; Fang, K. S.; Callender, R. H.; Dyer, W. H. W. R. B. *Biochemistry* **1996**, *35*, 691–697.
(32) Munoz, V.; Thompson, P. A.; Hofrichter, J.; Eaton, W. A. *Nature* **1997**, *390*, 196–199.
(33) Kubelka, J.; Hofrichter, J.; Eaton, W. A. *Curr. Opin. Struct. Biol.* **2004**, *14*, 76–88.
(34) Duan, L. L.; Mei, Y.; Zhang, Q. G.; Zhang, J. Z. H. *J. Chem. Phys.* **2009**, *130*, 115102.
(35) Ji, C. G.; Mei, Y.; Zhang, J. Z. H. *Biophys. J.* **2008**, *95*, 1080–1088.
(36) Ji, C. G.; Zhang, J. Z. H. *J. Am. Chem. Soc.* **2008**, *130*, 17129.
(37) Tong, Y.; Ji, C. G.; Mei, Y.; Zhang, J. Z. H. *J. Am. Chem. Soc.* **2009**, *131*, 8636–8641.
(38) Pantoja, U. D.; Pastor, M. T.; Salgado, J.; Pineda, L. A.; Perez, P. E. *J. Pept. Sci.* **2008**, *14*, 845–854.
(39) Lin, E.; Shell, M. S. *J. Chem. Theory Comput.* **2009**, *5*, 2062–2073.
(40) Kim, E.; Jang, S.; Pak, Y. *J. Chem. Phys.* **2007**, *127*, 145104.
(41) Onufriev, A.; Bashford, D.; Case, D. A. *Proteins: Struct., Funct., Bioinf.* **2004**, *55*, 383–394.
(42) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
(43) Pastor, R. W.; Brooks, B. R.; Szabo, A. *Mol. Phys.* **1988**, *65*, 1409.

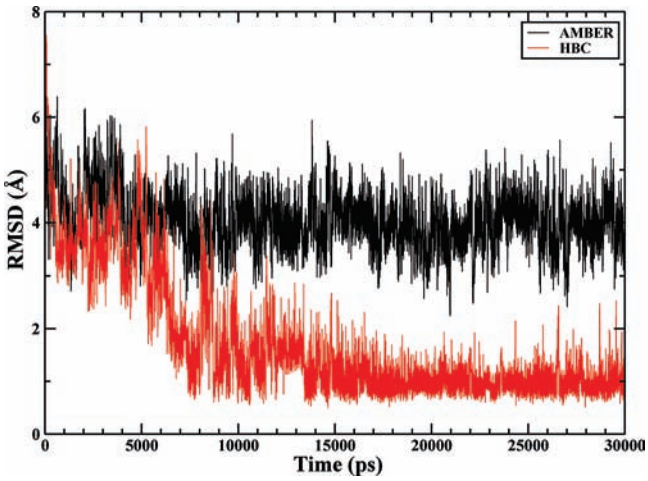

Figure 2. RMSD of backbone atoms of the peptide as a function of MD simulation time using AMBER 03 (black curve) and the adaptive HBC (red curve).

that simulation under the AMBER force field could not fold 2I9M within our simulation time of 30 ns. Although there are some partially folded helices, they do not lead to folded states as simulation progresses. We observed from these snapshots that although helices are partially formed at two terminals, the helix in the middle of protein is hard to form due to the lack of intraprotein hydrogen bonds. Before the helix is formed near the middle, the partially formed helices at the terminals are already broken due to unstable intraprotein hydrogen bonds there. It is hard to maintain stable hydrogen bonds at both terminals long enough for the middle helix to form. The root-mean-square deviation from the native structure is shown in Figure 2. Our result shows that, within a 30 ns simulation under the AMBER force field, the protein did not correctly fold into its native structure.

Next we build the polarization effect into MD simulation. In our previous studies with polarized protein-specific charges (PPCs), we were only interested in the energetics and dynamic properties of proteins in the vicinity of native states. Thus, we applied PPC based on the native structure and fixed them throughout simulation. However, when protein undergoes a large conformation change as in folding, applying a fixed PPC from a prefixed structure will certainly bias the simulation. Also, the native structure is unknown a priori. In this work we applied on-the-fly charge fitting based on the formation or cleavage of main chain hydrogen bonds. In the standard PPC scheme,[35] however, the atomic charges of entire protein residues are fitted by fragment quantum mechanical calculation at a given structure (typically the native structure). However, in protein folding, atomic charges vary considerably when large conformation change occurs. Updating charges of all atoms at every MD step is indispensable but is very demanding. In the HBC scheme, only residues involved in the formation or breaking of hydrogen bonds are subject to quantum calculation to refit their atomic charges. Atomic charges of the rest of residues are kept to their previous values. It is important to note that the current HBC is dynamic, i.e., the HBC are updated during MD simulation whenever hydrogen bonds are formed or broken. This is critical to the current folding study.

Force field parameters other than atomic charges, such as bond, angle, dihedral and vdW parameters, are taken from Amber force field and are kept intact during MD simulation. Atomic charges are initially taken from Amber force field, but are continuously updated using the HBC scheme described above. During MD simulation, a check is performed periodically to inspect if any hydrogen bonds are formed or broken. If yes, the above-described HBC scheme is performed to update the atomic charges of the corresponding residues and MD simulation is continued using the updated atomic charges. In practice, hydrogen bonds are detected during MD simulation using HBplus.[44] For the sake of computational efficiency and accuracy, the adaptive procedure is performed at discrete times, set at 20 ps in the present simulation, i.e., every 20 ps, at which point a check of the hydrogen bonds is performed. The same MD simulation, but under adaptive HBC, of the protein begins to form native-like structures in ~7 ns, and longer simulation leads to a well folded native helix as shown clearly in both Figures 1 and 2. The final folded structure has an rmsd from the experimental data of ~1.0 Å.

Cluster Analysis

At this stage, it is interesting to study the folding process in more detail by examining some representative intermediate structures under adaptive HBC simulation. As shown in Figure 1, after ~500 ps, a helix starts to form near the C-terminal and 40% of the native contacts are formed. In ~2 ns, a helix near the N terminal emerges. As the simulation time progresses, helix content gradually grows toward the middle of the peptide. At ~7 ns, a helix forms in the middle area and the entire peptide becomes a folded helix structure, essentially a fully folded native helix. Over the entire 30 ns simulation, the 2I9M successfully folds to the native conformation as judged by the low backbone rmsd which is used to evaluate the protein folding and the simulation structures are defined as the folded state if the backbone rmsd is within 1.5 Å. Here we use the first native structure of 2I9M determined by NMR as the reference structure in the calculation of the backbone rmsd. Figure 2 shows the comparison of the backbone rmsd of MD generated structures as a function of time using AMBER and HBC. In the AMBER simulation, the rmsd fluctuations are dominated from 3.5 to 5.5 Å. However, in the adaptive HBC simulation, the rmsd undergoes a rapid fall at the early folding stage until 6.3 ns when the first folding state appears with an rmsd of 1.49 Å. Then rmsd decreases continuously toward the native state, and after 16 ns, the system is already in a stable structure with the lowest backbone rmsd measured at 0.50 Å. Since the rmsd is the most commonly used measure of and most valuable tool in protein folding simulation to check the similarity between snapshots and native structure, backbone rmsd distributions during the entire simulation are further analyzed in Figure 3. The most populated state has an rmsd of 1.0 Å, while, under AMBER simulation, it is as large as 4.0 Å.

The snapshots of 2I9M extracted from the MD simulation are analyzed by K-means clustering,[45] and the three most populated clusters along with the representative structures are shown in Figure 4. The central structure among the most populated cluster (with a population of 68.2%) has a backbone rmsd of 0.69 Å, which corresponds to the native state. The second most populated cluster has a population of 14.1% with a backbone rmsd of 2.16 Å, which corresponds to a partially folded cluster. The third most populated conformation is also well folded with an rmsd of 1.27 Å and a population of 10.2%.

(44) McDonald, I. K.; Thornton, J. M. *J. Mol. Biol.* **1994**, *238*, 777–793.
(45) Shao, J.; Tanner, S. W.; Thompson, N.; Cheatham, T. E., III. *J. Chem. Theory Comput.* **2007**, *3*, 2312–2334.

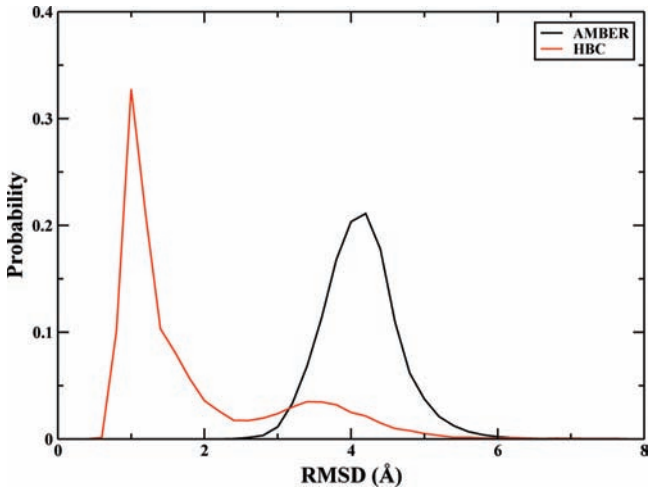

Figure 3. Distribution of backbone rmsd using AMBER (black curve) and PPC (red curve).

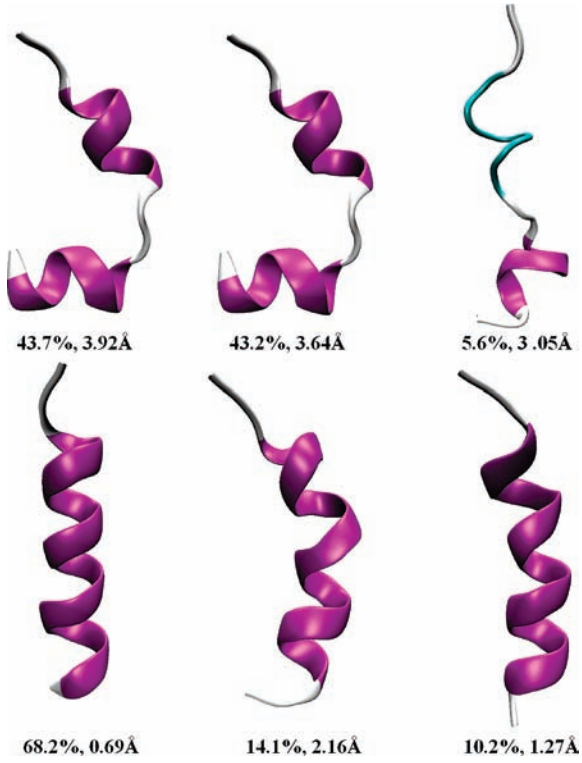

Figure 4. Representative structures of 2I9M conformations selected from the most populated clusters using AMEBR (top) and adaptive HBC (low). The population of clusters and the backbone rmsd of the cluster centers are indicated.

Therefore the folded conformations are dominant from our simulation under adaptive HBC. For comparison, in the simulation under the AMBER force field, the most populated cluster (43.7%) has an rmsd of 3.92 Å, and the second and third largest cluster, which is 43.2% and 5.6% populated, have rmsd values of 3.64 and 3.05 Å, respectively. Therefore no native structure is found in the simulation under the AMBER force field.

Native Contacts and Helix Content

A native contact is defined as a $C_\alpha - C_\alpha$ distance <7.0 Å for nonadjacent residues, and there are a total of 39 native contacts in the first NMR structure of 2I9M. The fractional native contact

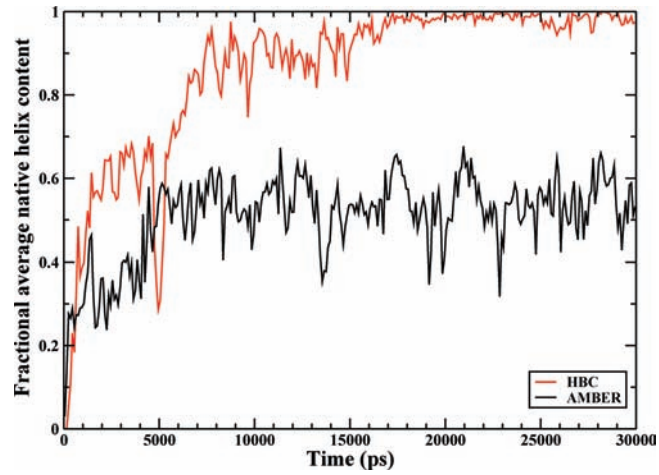

Figure 5. Fractional native helix content averaged every 100 snapshots during MD simulation as a function of simulation time using AMBER (black) and adaptive HBC (red).

is the number of total native contacts presented in the simulation structure divided by the number of native contacts in the NMR structure. In simulations under adaptive HBC, the native contacts start from 23% and quickly reach 80% within 7 ns, to a nearly folded state. Within 16 ns, the value reaches a level of 90% and fluctuates between 90% and 100% for the remaining simulation time, demonstrating that those simulated structures are correctly folded structures.

The fractional native helix content (number of residues in the helical conformation) averaged every 100 snapshots over the course of the simulation is also illustrated in Figure 5. The fraction of native helix content is those presented in both the NMR and the simulated structures divided by the total in the NMR structure. There are a total of 14 residues (1−14) in the helix of the NMR structure. In our analysis, residues 1 and 2 are removed and the defined secondary structure of proteins (DSSP)[46] is used to identify the secondary structure motif. Figure 5 shows that the helix starts to form immediately and the number of residues in helix is increased to the 60% level in just 1.5 ns. The helix content steadily increases as the simulation time increases. The overall fractional native helix content reaches near 100% after 16 ns which is in excellent agreement with the NMR result. In contrast, the helix is completed by only ~50% in the simulation under the AMBER force field as shown in Figure 5.

Free Energy Landscape

We constructed two-dimensional free energy landscapes from MD simulations under both the AMBER force field and adaptive HBC using the rmsd and radius of gyration (R_g) of the backbone as the reaction coordinates. The free energy landscape is determined using a weighted histogram analysis method (WHAM)[47−49] from state density $P(X) = \exp[-\beta W(X)]/Z$ where X is any set of reaction coordinates, $P(X)$ is the probability at X, and Z is the equipartition function. The relative free energy can then be easily expressed as $G(X_2) - G(X_1) = -RT \ln[P(X_2)/P(X_1)]$. Because the folding of 2I9M is almost barrierless, it

(46) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577–2637.
(47) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011–1021.
(48) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1995**, *16*, 1339–1350.
(49) Roux, B. *Comput. Phys. Commun.* **1995**, *91*, 275–282.

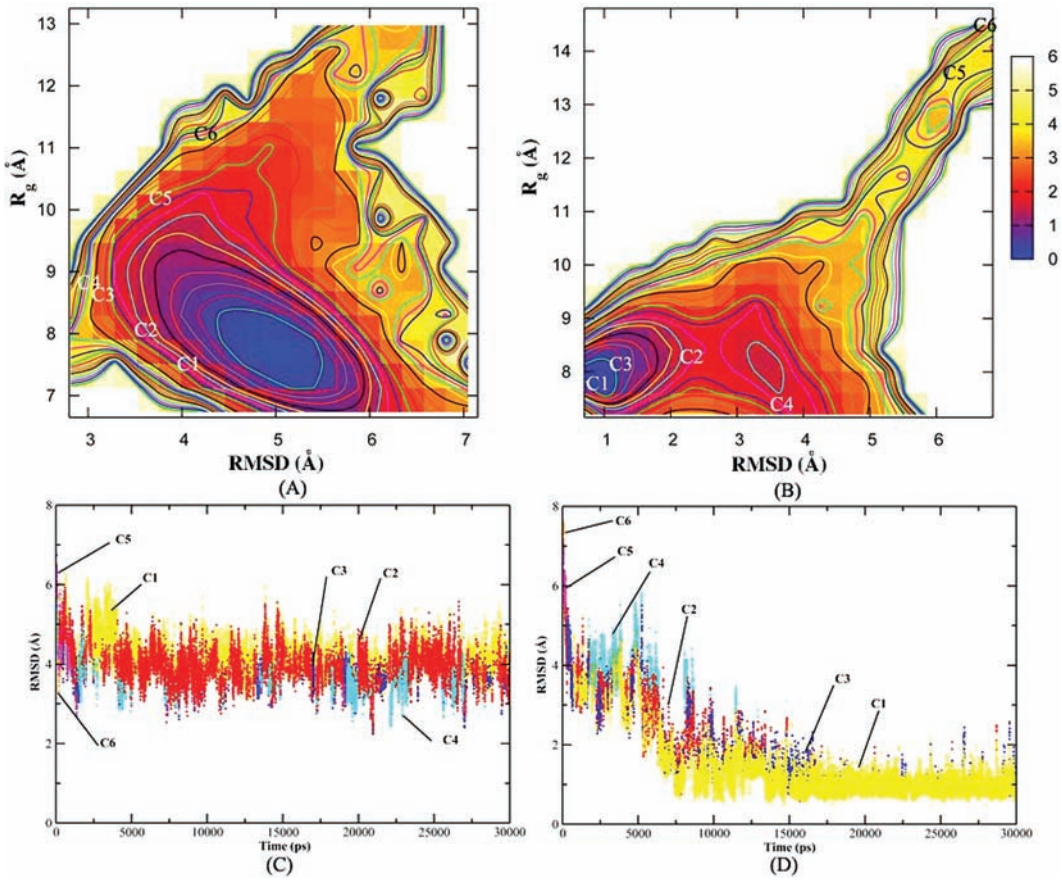

Figure 6. Free energy contour maps as a function of rmsd and radius of gyration using AMBER (A) and adaptive HBC (B) with the responding clusters (C, D). C1−C6 ranks from the most populated cluster to the least populated one. C1: yellow, C2: red, C3: blue, C4: cyan, C5: magenta, C6: orange.

transitions rapidly between folded and unfolded states. A biasing potential is unnecessary. Thus, we count the population of states along each reaction coordinate from the equilibrium simulation. For reference, the radius of gyration of the native structure averaged over 20 NMR structures is 7.98 Å.

The free energy contour maps as a function of rmsd and R_g are shown in Figure 6A and B, and they reveal some important features. First, the contour maps from AMBER and adaptive HBC show entirely different characteristics. Second, in the AMBER simulation, the native state is not the lowest free energy state (LFES). The LFES is near rmsd = 4.93 Å and R_g = 7.90 Å, which is far away from the native state. Finally, in the adaptive HBC simulation, the LFES falls within the approximal region of the native state (rmsd = 1.03 Å and R_g = 8.01 Å). The free energy contour maps indicate downhill folding without noticeable folding barriers and only the folded state is significantly populated in an apparently single free energy basin. This theoertical result is supported by a recent experimental observation of downhill folding of small, fast-folding proteins.[50−56]

The corresponding clusters are shown in Figure 6C, D, in which the abscissa denotes the MD simulation time and the ordinate denotes those conformations in different clusters relative to the rmsd of the NMR structure. In the simulation under the AMBER force field (Figure 6C), conformations in the same cluster are dispersive suggesting that the simulation does not reach convergence, and the LFES is not in the most populated cluster. Yet, in the simulation under the adaptive HBC (Figure 6D), the LFES is located at the native clusters (the most populated conformations and the third most populated conformations) indicating a good sampling of the conformation space. Those conformations in native clusters are almost continuous over the simulation time, and a majority of native conformations converge in 16 ns to 30 ns, demonstrating the native clusters are the most stable and our simulation has converged.

Discussion and Conclusion

In this work, the combined use of the AMBER force field and adaptive HBC with GB solvation model has been shown to successfully fold 2I9M into the correct α-helix from a standard MD simulation at room temperature. HBC is derived from quantum mechanical calculation for a protein in solution in which the atomic charges of relevant residues are fitted at regular intervals based on the variation of hydrogen bonds in the MD simulation. Starting from an extended conformation, 2I9M is found to correctly fold to the native state within 16 ns with a low rmsd, high fractional native contacts (100%), and high native helix content (100%). The folding process of the MD trajectory shows that the folding pathway initiates from

(50) Li, P.; Oliva, F. Y.; Naganathan, A. N.; Munoz, V. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 103–108.
(51) Sadqi, M.; Fushman, D.; Munoz, V. *Nature* **2006**, *442*, 317–321.
(52) Ma, H.; Gruebele, M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 2283–2287.
(53) Garcia-Mira, M. M.; Sadqi, M.; Fischer, N.; Sanchez-Ruiz, J. M.; Munoz, V. *Science* **2002**, *298*, 2191–2195.
(54) Liu, F.; Du, D.; Fuller, A. A.; Davoren, J. E.; Wipf, P.; Kelly, J. W.; Gruebele, M. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 2369–2374.
(55) Cho, S. S.; Weinkam, P.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 118–123.
(56) Liu, F.; Gruebele, M. *J. Mol. Biol.* **2007**, *370*, 574–584.

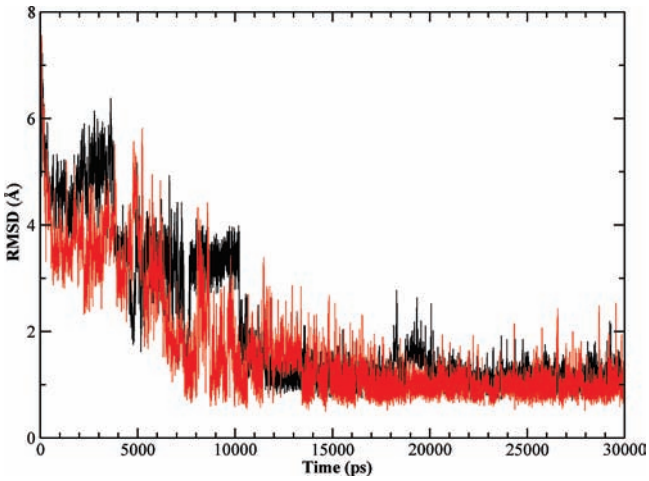

Figure 7. RMSD of backbone atoms of the peptide as a function of MD simulation time from two MD trajectories. The red curve denotes the trajectory discussed in the current paper and the black curve denotes another trajectory with the same starting structure but different random seed for momentum.

the C-terminal followed by the N-terminal and then the middle portion of 2I9M. The inability to fold 2I9M in a standard MD simulation under the AMBER force field is due to the lack of polarization that stabilizes intraprotein hydrogen bonds, which seems critical to the successful folding of a helix.

It should be mentioned that if different MD trajectories are run, specific details of the folding dynamics such as intermediate structures and their distributions may show variations.[57] For the present study, however, these specific and detailed variations of the intermediate structures from different MD trajectories do not have a significant effect on the overall folding dynamics, converged structures, and their distributions. This is shown in Figure 7 in which the backbone rmsd is plotted as a function of the simulation time from two MD trajectories starting from the same initial structure but with a different random seed. As is shown, there are variations especially in the intermediate structures before the folded structure is reached. But both trajectories reached the folded structure in approximately similar time frames.

Recently, Xie et al. employed an X-Pol potential from quantum calculations that is dynamically adjusted for MD simulation of a small protein in water.[58] Thus, employing a quantum-based electrostatic potential for MD simulation of biomolecules is finally coming of age.

Acknowledgment. Y.M. is supported by the National Natural Science Foundation of China (Grant No. 20803034). Q.G.Z. thanks the National Natural Science Foundation of China (Grant No. 10874104) and the Natural Science Foundation of Shandong Province (Grant No. Z2007A05). J.Z.H.Z. acknowledges financial support from the National Natural Science Foundation of China (Grant Nos. 20773060 and 20933002) and Shanghai PuJiang program (09PJ1404000).

JA102735G

(57) Kazmirski, S. L.; Li, A. J.; Daggett, V. *J. Mol. Biol.* **1999**, *290*, 283–304.
(58) Xie, W. S.; Orozco, M.; Truhlar, D. G.; Gao, J. L. *J. Chem. Theory Comput.* **2009**, *5*, 459–467.