J|A|C|S
COMMUNICATIONS

Published on Web 04/17/2002

Molecular Dynamics Study of Peptides in Implicit Water: Ab Initio Folding of β-Hairpin, β-Sheet, and ββα-motif

Soonmin Jang,[†] Seokmin Shin,*[,†] and Youngshang Pak*[,‡]

School of Chemistry, Seoul National University, Seoul 151-747, Korea, and Department of Chemistry, Pusan National University, Pusan 609-735, Korea

Received January 23, 2002

During the past several years various de novo simulation methods based on molecular dynamics (MD) or Monte Carlo (MC) schemes have been proposed in an attempt to predict native structures of proteins.[1,2] Usually most protein molecules are in aqueous environments. Thus, the simulations need to take into account the solvation effect accordingly. Unfortunately, using explicit water models in any simulation requires a considerable amount of computational time and thus prohibits sufficiently long time simulations for protein-folding studies. Recently, a fast numerical approximation scheme to incorporate the solvation energy of an arbitrarily shaped solute has been developed for various empirical all-atom force fields using the generalized born (GB) model concept.[3] It has been shown that the GB model may recover most of the important solvation effects on small peptides.[4] In this communication, we demonstrate that one can successfully predict native structures of peptides with various structural motifs (β-hairpin, β-sheet, and ββα-moiety), using the all-atom based force field (CHARMM19)[5] in conjunction with the GB solvation model.[6,3f] Our MD simulations employed a total of four different peptides: 17-residue β-hairpin (PDB code: 1E0Q),[7] 16-residue segment of 56-residue peptide (PDB code: 1GB1)[8a,b] (we denote this segment as gb1 here), 20-residue three-strand antiparallel β-sheet (we denote it as b3s here),[9] and 28-residue peptide with a ββα-motif (PDB code: 1FSD).[10] The N and C termini of all the peptides were patched with the standard CO_2^- and NH_3^+ groups, respectively. All the MD simulations were started from the fully extended conformations of the peptides. The SHAKE algorithm is used to fix the bond distances consisting of heavy and hydrogen atoms, and the Berendsen thermostat is used for the temperature control. In this study, no nonbond energy cutoffs were employed to calculate the full GB solvation energy. Performing simulations at elevated temperatures can keep the system from being trapped in local minima, leading to a faster folding pathway. We showed that at moderately elevated temperatures complete folding events of all the peptides were observed within several tens of nanoseconds.

In aqueous solution, the peptide 1E0Q (MQIFVKTLDKT ITLEV) is known to exist as a β-hairpin.[7] In the present study, a total of six independent trajectories were obtained at 360 K with a time step of 0.15 fs for 15.0 ns. Out of the six trajectories, five were shown to form the β-hairpin conformation within time scales in the range of 0.4−7.0 ns. The predicted lowest-energy conformation of 1E0Q is in excellent agreement with the solution NMR structure[7] [Note that the root-mean-square deviation (RMSD) of the predicted structure is 1.36 Å with respect to the NMR one.] Snapshots of one of the hairpin-forming trajectories are shown in Figure 1.

The 16-residue segment GEWTYDDATKTFTVTE taken from the peptide 1GB1 also forms a β-hairpin in solution and has been extensively studied both theoretically and experimentally.[2g,8,11] Zagrovic et al.[11a] have performed MD simulation at 300 K for several micro-seconds, using the OPLS potential parameters[12] and a GB solvation model using a distributed computing technique.[13] On the basis of their simulation studies, it is concluded that the folding pathway involves the existence of a hydrophobically stable intermediate, suggesting a possible three-state folding mechanism. The computational study of Lee et al.[8g] also demonstrated the importance of hydrophobic interactions and the hydrogen bonding contributions during the folding/unfolding. We have also performed six MD simulations for 15 ns at 360 K with time step 0.15 fs. Two trajectories lead to "β-hairpin like" conformations which are close to the nativelike structure at the initial stage of the simulations, but we also observed the subsequent formation of semi-α-helical structures, which constituted a prevailing conformation during the MD runs. To sample more diverse conformations, the same MD simulation was performed at 380 K. The MD simulation located the semi-α-helical conformer at 3.7 ns and also found the stable "nativelike" β-hairpin structure at 14 ns. (Figure 1). Similar semihelical intermediates were observed in previous studies.[8f,11a] Surprisingly, our local energy minimization studies showed that the semi-α-helical structure is more stable than the "nativelike" one by 11 kcal/mol. It can be argued that the discrepancy is mainly due to the nature of the potential energy function employed in the present study.

De Alba et al.[9] have designed a synthetic peptide that forms a three strand antiparallel β-sheet in aqueous solution. There have been some folding simulations[14] on this peptide (TWIQNGSTK-WYQNGSTKLYT) using rather simple solvation models, such as solvent-referenced potential[15] and solvent-accessible surface model,[16] with all the charged residues artificially neutralized. Our eight independent simulations at 400 K with time step 0.2 fs for 30 ns all found a stable antiparallel β-sheet conformation. Furthermore, we performed two more independent MD simulations at 420 K for 30 ns and observed a reversible folding event. One of the trajectories at 420 K is displayed in Figure 2.

The peptide of the ββα-motif 1FSD (QQYTAKIKGRTFRNE KELRDFIEKFKGR) has also been investigated for the folding simulations using a coarse-grained force field.[17] We obtained six independent trajectories at 440 K with time step 0.1 fs. Each MD run was performed for 15 ns, and two trajectories resulted in successful location of the ββα-like native structure. Another simulation at 430 K for 19 ns also located the native structure, and the snapshots of this trajectory are shown in Figure 2. The folding event of this peptide is always initiated by the formation of the α-helix propagating from the middle of the backbone toward the N terminus and then followed by the formation of the small loop and the β-hairpin in the other region toward the C terminus. It is noted that the β-hairpin region is not fully stabilized at both of the temperatures.[10] We believe that the simulation temperatures may

* Corresponding authors. Email: sshin@snu.ac.kr or ypak@pusan.ac.kr.
[†] Seoul National University.
[‡] Pusan National University.

10.1021/ja025675b CCC: $22.00

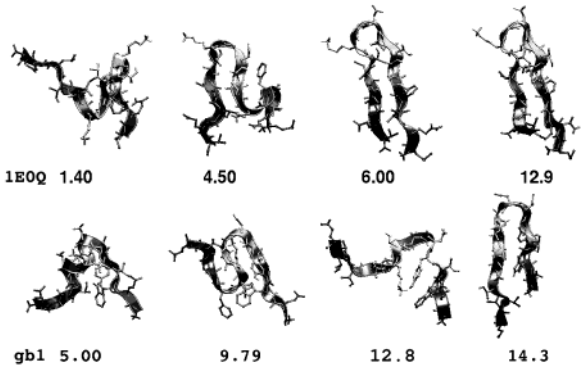

Figure 1. Snapshots of representative folding trajectories for 1E0Q (top) and gb1 (bottom). The backbones of the peptides and the side chains are shown with the C terminus on the left-hand side.The numbers represent simulation times in ns.

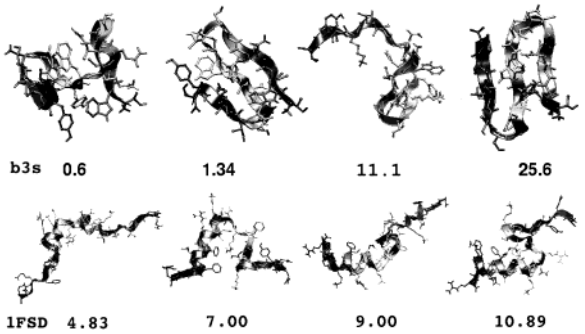

Figure 2. Snapshots of representative folding trajectories for the b3s (top) and 1FSD (bottom).

be too high to stabilize the β-hairpin moiety, although the α-helix part seems quite stable. The RMSD for one of the lowest-energy structures (the nativelike one) was 2.56 Å relative to the NMR result.[10] In addition, our simulations indicate that the β-hairpin moiety can undergo a slow conformational transition to another strand of α-helix for the whole peptide to form a purely α-helix bundle. Energetically the minimized energy of this conformer is even lower than that of the nativelike structure by 0.3 kcal/mol.

In conclusion, we have demonstrated that MD simulations using a GB implicit solvent model with an all-atom based force field can describe the spontaneous folding of small peptides in aqueous solution. It is generally viewed that the prediction of β-hairpins and β-sheets is computationally more challenging than that of α-helices. In the present MD study, the native structures of β-hairpin, antiparallel β-sheet, and $\beta\beta\alpha$-motif were successfully predicted within reasonable time scales at moderately elevated temperatures. The folding time scales observed in our simulations with the implicit solvation model may be different from the real folding times. One cannot expect the energy landscape at a moderately high temperature to be identical with that at physiological temperature. Nonetheless, the GB solvent model combined with high-temperature MD makes it possible to reduce the computational requirement tremendously, which can be crucial in studying the folding process of medium or large proteins. To accelerate conformational searches during folding, we have performed MD simulations at high temperatures, which may not be the most effective method for that purpose. Recently, "*q*-jumping MD" via a simple transformation of the potential energy function was shown to provide a reasonable tool for accelerating conformational searches in the study of α-helical folding.[18] In another approach, the SGMD method,[2a] combined with the efficient numerical solution of the Poisson−Boltzmann (PB) equation for the implicit solvent model, has been applied to the fast folding of several peptides by Luo and Kollman.[19] Despite their usefulness in predicting the native structures of peptides, the fictitious nature of dynamics resulting from either the "*q*-jumping MD" or SGMD imposes a fundamental limitation in describing real folding pathways. The present simulations are expected to provide further insight into the folding mechanism, as well as the efficient ab initio prediction of the native structures of the peptides. The detailed analysis of the folding trajectories will be reported elsewhere.

Acknowledgment. S.S. and S.J. acknowledge supports from the KISTI and the BK21 Division of Chemistry and Molecular Engineering. Y.P. acknowledges support from the Basic Research Program of the Korea Science and Engineering Foundation (Grant No. 2001-2-12100-001-2).

Supporting Information Available: Energy profiles of the MD simulations corresponding to Figures 1 and 2 and the superimposed images of predicted and native structures for 1E0Q and 1FSD (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Berne, B. J.; Straub, J. E. *Curr. Opin. Struct. Biol.* **1997**, *7*, 181. (b) Shea, J.-E.; Brooks, C. L., III. *Annu. Rev. Phys. Chem.* **2001**, *52*, 499. (c) Baker, D.; Sali, A. *Science* **2001**, *294*, 93.
(2) (a) Wu, X.; Wang, S. *J. Phys. Chem. B* **1998**, *102*, 7238. (b) Samudrala, R.; Xia, Y.; Huang, E.; Levitt. M. *Proteins (Suppl.)* **1999**, *3*, 93. (c) Ortiz, A. R.; Kolinski, A.; Rotkiewicz, P.; Ilkowski, B.; Skolnick, J. *Proteins* **1999**, *37*, 117. (d) Pillardy, J.; Czaplewski, C.; Liwo, A.; Lee, J.; Ripoll, D. R.; Kazmierkiewicz, R.; Oldziej, S.; Wedemeyer, W. J.; Gibson, K. D.; Arnautova, Y. A.; Saunders, J.; Ye, Y.-J.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 2329. (e) Jones, D. T. *Proteins (Suppl.)* **1997**, *1*, 185. (f) Simons, K.; Strauss, C.; Baker, D. *J. Mol. Biol.* **2000**, *306*, 1191. (g) Dinner, A.; Lazaridis, T.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *9068*, 96. (h) Hansman, L.; Okamoto, Y. *Curr. Opin. Struct. Biol*. **1999**, *9*, 177. (i) Bartels, C.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 865. (j) Hukushima, K.; Nemoto, K. *J. Phys. Soc. Jpn*. **1996**, *65*, 1604. (k) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett*. **1999**, *314*, 141. (l) Mitsutake, A.; Sugita, Y.; Okamoto, Y. *Biopolymers* **2001**, *60*, 96. (m) Stolovitzky, G.; Berne, B. J. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 11164.
(3) (a) Constanciel, R.; Contreras, R. *Theor. Chim. Acta* **1984**, *65*, 1. (b) Qui, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. *J. Phys. Chem. A* **1997**, *101*, 3005. (c) Still, W. C.; Tempczyk, A.; Hawley R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127. (d) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *Chem. Phys. Lett.* **1995**, *246*, 122. (e) Edinger, S. R.; Cortis, C.; Shenkin, P. S.; Friesner, R. A. *J. Phys. Chem. B* **1997**, *101*, 1190. (f) Dominy, B. H.; Brooks, C. L., III. *J. Phys. Chem. B* **1999**, *103*, 3765. (g) Srinivasan, J.; Trevathan, M. W.; Beroza, P.; Case, D. A. *Theor. Chem. Acc.* **1999**, *101*, 426.
(4) Bursulaya, B. D.; Brooks, C. L., III. *J. Am. Chem. Soc*. **1999**, *121*, 9947.
(5) MacKerell, A. D., Jr.; Brooks, B.; Brooks, C. L., III; Nilsson, L.; Roux, B.; Won, Y,; Karplus, M. *CHARMM: The Encyclopedia of Computational Chemistry*; Schleyer, P. R., Eds.; John Wiley & Sons: New York, 1998; Vol. 1, p 271.
(6) (a) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161. (b) Bashford, D.; Case, D. *Annu. Rev. Phys. Chem*. **2000**, *51*, 129.
(7) Zerella, R.; Chen, Y. P.; Evans, P. A.; Raine, A.; William, D. H. *Protein Sci.* **2000**, *9*, 2142.
(8) (a) Blanco, F.; Rivas, G.; Serrano, L. *Nat. Struct. Biol*. **1994**, *1*, 584. (b) Munoz, V.; Thompson, P. A.; Hofrichter, J.; Eaton, W. A. *Nature* **1997**, *390*, 196. (c) Pande, V.; Rokhsar, D. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *97*, 9062. (d) Klimov, D.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 2544. (e) Roccatano, D.; Amadei, A.; Nola, A. D.; Berendsen, H. J. *Protein Sci*. **1999**, *10*, 2130. (f) Garcia, A. E.; Sanbonmatsu, K. Y. *Proteins: Struct., Funct. Genet.* **2001**, *42*, 345. (g) Lee, J.; Shin, S. *Biophys. J.* **2001**, *81*, 2507.
(9) de Alba, E.; Santoro, J.; Rico, M.; Jimenez, M. A. *Protein Sci*. **1999**, *8*, 854.
(10) (a) Dahiyat, B. I.: Mayo, S. L. *Science* **1997**, *287*, 82. (b) DeGrado, W. F. *Science* **1997**, *287*, 80.
(11) (a) Zagrovic, B.; Sorin, E. J.; Pande, V. *J. Mol. Biol*. **2001**, *313*, 151. (b) Zhou, R.; Berne, B. J.: Germain, R. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 1491.
(12) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc*. **1988**, *110*, 1666.
(13) Shirts, M. R.; Pande, V. *Science* **2001**, *290*, 1903.
(14) (a) Ferrara, P.; Catflisch, A. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 10780. (b) Wang, H.; Sung, S. *J. Am. Chem. Soc.* **2000**, *122*, 1999.
(15) (a) Sung, S. S. *Biophys. J.* **1994**, *23*, 847 (b) Sung, S. S.; Wu, X. W. *Proteins* **1996**, *25*, 202.
(16) Eisenberg, D.; MaLachlan, A. D. *Nature* **1986**, *319*, 199.
(17) Derreumaux, P. *J. Chem. Phys*. **1999**, *111*, 2301.
(18) Pak, Y.; Wang, S. *J. Chem. Phys.* **1999**, *111*, 4359.
(19) *Chem. Eng. News* **2001**, *79*, Oct 8, 41.

JA025675B