Direct folding simulation of a long helix in explicit water

Zhang, John Z. H.; Gao, Ya; Lu, Xiaoliang; Duan, Lili; Zhang, Dawei; Mei, Ye

2013

Gao, Y., Lu, X., Duan, L., Zhang, D., Mei, Y., & Zhang, J. Z. H. (2013). Direct folding simulation of a long helix in explicit water. *Applied Physics Letters*, 102(19), 193706.

<https://hdl.handle.net/10356/98147>

<https://doi.org/10.1063/1.4807145>

© 2013 AIP Publishing LLC. This paper was published in *Applied Physics Letters* and is made available as an electronic reprint (preprint) with permission of AIP Publishing LLC. The paper can be found at the following official DOI: [<http://dx.doi.org/10.1063/1.4807145>]. One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution to multiple locations via electronic or other means, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper is prohibited and is subject to penalties under law.

Downloaded on 26 Apr 2023 08:42:30 SGT

Direct folding simulation of a long helix in explicit water

Ya Gao, Xiaoliang Lu, Lili Duan, Dawei Zhang, Ye Mei et al.

Citation: *Appl. Phys. Lett.* **102**, 193706 (2013); doi: 10.1063/1.4807145

View online: <http://dx.doi.org/10.1063/1.4807145>

View Table of Contents: <http://apl.aip.org/resource/1/APPLAB/v102/i19>

Published by the AIP Publishing LLC.

Additional information on Appl. Phys. Lett.

Journal Homepage: <http://apl.aip.org/>

Journal Information: http://apl.aip.org/about/about_the_journal

Top downloads: http://apl.aip.org/features/most_downloaded

Information for Authors: <http://apl.aip.org/authors>

ADVERTISEMENT

Direct folding simulation of a long helix in explicit water

Ya Gao,¹ Xiaoliang Lu,¹ Lili Duan,² Dawei Zhang,³ Ye Mei,^{1,a)} and John Z. H. Zhang^{1,4,a)}

¹State Key Laboratory of Precision Spectroscopy, Institute of Theoretical and Computational Science, East China Normal University, Shanghai 200062, China

²College of Physics and Electronics, Shandong Normal University, Jinan 250014, China

³Division of Chemistry and Biological Chemistry, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371, Singapore

⁴Department of Chemistry, New York University, New York, New York 10003, USA

(Received 14 January 2013; accepted 3 May 2013; published online 17 May 2013)

A recently proposed Polarizable Hydrogen Bond (PHB) method has been employed to simulate the folding of a 53 amino acid helix (PDB ID 2KHK) in explicit water. Under PHB simulation, starting from a fully extended structure, the peptide folds into the native state as confirmed by measured time evolutions of radius of gyration, root mean square deviation (RMSD), and native hydrogen bond. Free energy and cluster analysis show that the folded helix is thermally stable under the PHB model. Comparison of simulation results under, respectively, PHB and standard nonpolarizable force field demonstrates that polarization is critical for stable folding of this long α -helix. © 2013 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4807145>]

Elucidation of the kinetics and thermodynamics of protein folding has wide-ranging implications in biochemistry, genetics, and pharmaceutical chemistry.^{1,2} Currently, all atom molecular dynamics (MD) simulation can be carried out to simulate folding of small proteins³ and even some relatively large proteins.^{4,5} Yet despite significant progress, it is still difficult to correctly predict folding and establish detailed description of the folding mechanism.⁶ In addition to difficulties in accessible time scale,⁷ a major challenge lies in the accuracy of the force field used. An accurate (or reliable) force field must correctly describe the relative energies of folded, unfolded, and perhaps the intermediate states in free energy landscape. Although widely used force fields such as AMBER,⁸ CHARMM,⁹ GROMOS,¹⁰ and OPLS¹¹ have enjoyed great success in biomolecular simulations,¹² a serious deficiency of these force fields is the lack of polarization effect. For example, secondary structure elements in proteins rely in large part on the formation of hydrogen bonds, which have strong polarization effect.¹³ Thus, including polarization in the force field for protein dynamics simulation is extremely important.

Modeling force fields with directional hydrogen bonding have frequently shown improved accuracy.¹⁴ In our previous studies, electrostatic polarization has been demonstrated to be very important in stabilizing the native structure of protein and intra-protein hydrogen bonds.¹⁵ It is noted that without the polarization effect, hydrogen bonds are found to be “too fragile,” which leads to less stable local structure of protein.¹⁶ By including the polarization of hydrogen bond, we folded a short helix to the native structure using the polarizable hydrogen bond (PHB) at room temperature.^{17,18} However, these studies are based on short helical peptide. Whether it is still effective when applying to long helical peptide is still unknown. In this work, we performed folding simulation of a long helical peptide in explicit water by

employing the PHB model, and compare the result with that using standard nonpolarizable force field.

The basic idea of PHB model can be found in a previous paper.¹⁸ Here we just give a brief introduction. When hydrogen bond forms or breaks, the electron distribution will be distorted due to mutual electrostatic polarization between the hydrogen bond donor and acceptor. Thus, electrostatic polarization effect plays an important role in folding energetics. By using a pair of di-alanine in helical conformation as a model system and changing the donor-acceptor distance from 2.5 to 6.5 Å while fixing other degrees of freedom, we calculated the distance-dependent atomic charges by using the restrained electrostatic potential (RESP)¹⁹ charge fitting method. Electrostatic potential on grids around the residue was calculated based on the polarized wave function, and new atomic charges were obtained. Iterations were conducted until convergence was reached. Quantum mechanical calculations were carried out using Gaussian 09 (Ref. 20) at HF/6-31 G* level in order to be consistent with AMBER charges.

The change of atomic charge from the reference value in a backbone hydrogen bonding pair –CO and –NH is fitted to an exponential function as

$$\Delta q_O = -0.334 \times \exp(-0.466 \times R) \quad (1)$$

and

$$\Delta q_N = -0.493 \times \exp(-0.455 \times R), \quad (2)$$

where R is the distance between donor and acceptor. The values of transferred charges for H and C atoms were exactly the negative of that for N and O atoms.

2KHK is a 53 amino acid straight long helix belonging to b30 to 82 domain of F1Fo adenosine-5'-triphosphate synthesis in *Escherichia coli*. In view of its large size, a short MD simulation (10 ns) in implicit water was performed first to make the peptide somewhat more compact in order to minimize the overall size of the water box in explicit water

^{a)}Authors to whom correspondence should be addressed. Electronic addresses: ymei@phy.ecnu.edu.cn and john.zhang@nyu.edu

FIG. 1. The last snapshots of implicit water simulations (a) under AMBER03 force field and (b) PHB model.

MD simulation that follows. The last snapshot from this short MD simulation was extracted as the initial structure (shown in Figure 1) for explicit water simulation. Simulations were carried out using AMBER10 package with some in-house modifications. The structure was immersed in the center of a truncated octahedral box of TIP3P water molecules, and all of the peptide atoms were no less than 12 Å from the edge of the water box. Two chloride ions were added to neutralize the charge of the system. To remove bad contact before the simulation, we ran 5000 steps of steepest descent followed by 5000 steps of conjugate gradient energy minimizations. The system was heated up from 0 to 288 K in 100 ps, which was the experimental temperature of 2KHK, during which all atoms in the protein were restrained with a force constant of 10 kcal/mol·Å². Particle Mesh Ewald²¹ method with a 10 Å cutoff in real space was used to calculate electrostatic interaction. Langevin thermostat with collision frequency 1.0 ps⁻¹ was used to regulate the temperature.²² Isotropic pressure coupling with a relaxation time of 2 ps was used to maintain the

pressure to 1 atm.²³ Integration time step was set to 2 fs. Trajectories were saved every 1 ps, and 150 000 snapshots were saved in production run for further analysis.

The production runs were performed using both the standard AMBER03 force field and the PHB, respectively. The only difference between PHB and standard AMBER03 is that atomic charges in hydrogen bond donor and acceptor, while other parameters of the force field are exactly the same. During MD simulation, hydrogen bonds were checked periodically using SIMULAIID program.²⁴ The PHB scheme was employed to update the atomic charges of C, O, N, and H for those hydrogen bonded residues, and MD simulations were continued using the updated atomic charges. To balance the efficiency and accuracy, the periodic check was performed every 100 ps in this study, which was found to be reasonable in terms of accuracy.

The variations of Root Mean Square Deviation (RMSD) from the native structure and radius of gyration (Rg) along the simulation time are shown in Figure 2. The overall fluctuations of RMSD and Rg using AMBER03 charges are both larger than those using PHB model. The RMSD distribution and representative structures are also shown in Figure 2. The most populated state under PHB model has an RMSD of 2.1 Å, while under standard AMBER simulation, it is as large as 4.0 Å. The result indicates that hydrogen bond polarization helps the folding of long helix. This is also reflected from the fluctuation of Rg along the simulation time. Under standard AMBER03 force field, Rg fluctuates widely and after about 120 ns, the Rg gradually deviated from the native value, indicating that the helix conformation is deformed. While under PHB simulation, initially the peptide adopts a more compact structure than the native structure. After 75 ns, the Rg wanders around the experimental value with small fluctuation.

In order to gain insight into the folding dynamics, we monitored snapshots of the intermediate states along the simulation time under PHB model as plotted in Figure 3. It is noted that the structure near the N-terminal of the peptide is flexible while the middle long helix is becoming more and more stable as simulation goes on.

FIG. 2. Backbone RMSD from the native structure and radius of gyration as a function of simulation time using AMBER03 charge (black) and the PHB (red) model in explicit water. The representative structures and their relative distributions from these two simulations are also plotted (right panels).

FIG. 3. Snapshots of intermediate structures of the peptide at different simulation time under PHB model (compared to the NMR structure). The C-terminal is always on the top.

The driving force for protein folding, or even more simply the driving force for the formation of helix, in water is still controversial.²⁵ The polarization effect along the hydrogen bonds is essential for helix formation.²⁶ This statement is also supported by the hydrogen bond cooperativity.²⁷ It is counted that the final number of hydrogen bond is 21.6 using AMBER charge but is 24.6 using PHB charge, indicating that hydrogen bonds are strengthened when polarization effect is included.

We constructed two-dimensional free energy landscapes from MD simulations under both the AMBER03 force field and PHB model using the RMSD and Rg of the backbone as the reaction coordinates. Free energy was calculated by Weighted Histogram Analysis Method (WHAM).²⁸ Figure 4 plots the 2D free energy map at 288 K from simulations using, respectively, standard AMBER03 and PHB model. The free energy map under AMBER03 shows a global minimum with RMSD of around 4.1 Å and Rg around 17.2 Å (versus the value of 20.3 Å calculated from the nuclear magnetic resonance (NMR) structure). The structure is obviously twisted from the middle of the peptide as shown in Figure 4 (top panel). So, no stable native structure is founded in standard AMBER simulation. A local minimum with Rg around 26 Å is obtained under AMBER03 force field and this is not the global minimum.

For comparison, the free energy landscape under the PHB model displays a global minimum at RMSD around 2.0 Å (folded state) as shown in Figure 4 (bottom panel). The global minimum obtained in the PHB model has a Rg value of about 20.0 Å, which is very close to the NMR value of 20.3 Å. There is also a local minimum at RMSD around 3.1 Å, which is a partially unfolded state. At this local minimum, the N terminal of the peptide is unfolded, but the

FIG. 4. Free energy landscapes at 288 K mapped to RMSD from the native structure and radius of gyration obtained from simulations employing (top) AMBER03 charge and (bottom) PHB model. The representative structures are also shown in the figure.

middle helical fragment is almost intact. This free energy analysis shows that the helix is thermally more stable under the PHB model.

In this work, we utilized a recently proposed PHB method to simulate the folding of a long (53 amino acid) helical peptide (PDB ID 2KHK) in explicit water. Under PHB simulation, the peptide folds into the native state and remains stably folded. Comparison of simulation results under, respectively, PHB and standard AMBER force field demonstrates that explicit polarization effect is essential for folding of this long α -helix. This work demonstrates that polarization of hydrogen bond should be taken into account in protein folding.

It should be noted that folding simulations of peptides with β -sheet or mixed helix-sheet structures are very challenging. The difficulty lies in not only the lack of polarization effect, but also the unbalanced secondary structure propensity in widely used force field. For instance, AMBER03 force field biases the helical conformation, which may

impede the folding of β structure. Refinement of backbone torsion parameters is an ongoing research project in our group and we hope to be able to obtain good result for folding of β structure soon.

This work was supported by the National Science Foundation of China (Grant Nos. 20933002, 21173082, and 11147026), Shanghai PuJiang program (09PJ1404000), Shanghai Rising-Star program (Grant No. 11QA1402000) and Tier 1 RG58/11 of Nanyang Technological University. We also thank the High Performance Computer Center of East China Normal University for CPU time support.

¹Y. Duan and P. A. Kollman, *Science* **282**, 740 (1998).

²R. N. Sifers, *Nat. Struct. Biol.* **2**, 355 (1995).

³T. Fox and P. A. Kollman, *Proteins* **25**, 315 (1996).

⁴C. L. Brooks, *Acc. Chem. Res.* **35**, 447 (2002).

⁵R. H. Zhou, B. J. Berne, and R. Germain, *Proc. Natl. Acad. Sci. U. S. A.* **98**, 14931 (2001).

⁶K. A. Dill and H. S. Chan, *Nat. Struct. Biol.* **4**, 10 (1997).

⁷C. Simmerling, B. Strockbine, and A. E. Roitberg, *J. Am. Chem. Soc.* **124**, 11258 (2002).

⁸V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling, *Proteins* **65**, 712 (2006).

⁹A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha *et al.*, *J. Phys. Chem. B* **102**, 3586 (1998).

¹⁰W. F. Van Gunsteren, C. Oostenbrink, A. Villa, and A. E. Mark, *J. Comput. Chem.* **25**, 1656 (2004).

¹¹R. A. Friesner, G. A. Kaminski, J. Tirado-Rives, and W. L. Jorgensen, *J. Phys. Chem. B* **105**, 6474 (2001).

¹²Y. Q. Deng and B. Roux, *J. Phys. Chem. B* **113**, 2234 (2009).

¹³R. B. Best and G. Hummer, *J. Phys. Chem. B* **113**, 9004 (2009).

¹⁴J. H. Lii and N. L. Allinger, *J. Comput. Chem.* **19**, 1001 (1998).

¹⁵C. G. Ji and J. Z. H. Zhang, *J. Phys. Chem. B* **113**, 16059 (2009).

¹⁶C. G. Ji and J. Z. H. Zhang, *J. Phys. Chem. B* **113**, 13898 (2009).

¹⁷L. L. Duan, Y. Mei, D. W. Zhang, Q. G. Zhang, and J. Z. H. Zhang, *J. Am. Chem. Soc.* **132**, 11159 (2010).

¹⁸Y. Gao, X. L. Lu, L. L. Duan, J. Z. H. Zhang, and Y. Mei, *J. Phys. Chem. B* **116**, 549 (2012).

¹⁹W. D. Cornell, P. Cieplak, C. I. Bayly, and P. A. Kollman, *J. Am. Chem. Soc.* **115**, 9620 (1993).

²⁰M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson *et al.*, Gaussian 09, Revision B.01, Gaussian, Inc., Wallingford CT, 2010.

²¹T. Darden, D. York, and L. Pedersen, *J. Chem. Phys.* **98**, 10089 (1993).

²²B. P. Uberuaga, M. Anghel, and A. F. Voter, *J. Chem. Phys.* **120**, 6363 (2004).

²³H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, *J. Chem. Phys.* **81**, 3684 (1984).

²⁴M. Mezei, *J. Comput. Chem.* **31**, 2658 (2010).

²⁵R. L. Baldwin, *J. Biol. Chem.* **278**, 17581 (2003).

²⁶C. Y. Wei, D. Tung, Y. M. Yip, Y. Mei, and D. W. Zhang, *J. Chem. Phys.* **134**, 171101 (2011).

²⁷Y. D. Wu and Y. L. Zhao, *J. Am. Chem. Soc.* **123**, 5313 (2001).

²⁸S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman, and J. M. Rosenberg, *J. Comput. Chem.* **13**, 1011 (1992).