J Mol Model (2013) 19:2605–2612
DOI 10.1007/s00894-013-1798-8

ORIGINAL PAPER

An implementation of hydrophobic force in implicit solvent molecular dynamics simulation for packed proteins

Li L. Duan · Tong Zhu · Ye Mei · Qing G. Zhang · Bo Tang · John Z. H. Zhang

Received: 9 October 2012 / Accepted: 5 February 2013 / Published online: 12 March 2013

Abstract MD simulations of five proteins in which helical chains are held together by hydrophobic packing were carried out to investigate the effect of hydrophobic force on simulated structures of these protein complexes in implicit generalized Born (GB) model. The simulation study employed three different methods to treat hydrophobic effect: the standard GB method that does not include explicit hydrophobic force, the LCPO method that includes explicit hydrophobic force based directly on solvent accessible surface area (SASA), and a proposed packing enforced GB (PEGB) method that includes explicit hydrophobic force based on the radius of gyration of the protein complex. Our simulation study showed that all five protein complexes were unpacked in the standard GB simulation (without explicit hydrophobic force). In the LCPO method, three of the five protein systems remained well packed during the simulation, indicating the need for an explicit hydrophobic force in GB model for these packed protein systems. However, two of the five systems were still unpacked during LCPO simulation. For comparison, all five protein systems remain well packed in simulation using the new PEGB method. Analysis shows that the failure of the LCPO method in two cases is related to the way that SASA changes during the unpacking process for these two systems. These examples showed that standard GB method without explicit hydrophobic force is not suitable for MD simulation of protein systems involving hydrophobic packing. A similar problem remains but to a much lesser extent in the LCPO method for some packed protein systems. The proposed PEGB method seems quite promising for MD simulation of large, multi-domain packed proteins in implicit solvent model.

Keywords GB model · Hydrophobic force · Packed proteins · Radius of gyration · Solvent accessible surface area

Electronic supplementary material The online version of this article (doi:10.1007/s00894-013-1798-8) contains supplementary material, which is available to authorized users.

L. L. Duan · Q. G. Zhang
College of Physics and Electronics, Shandong Normal University, Jinan 250014, China

T. Zhu · Y. Mei (✉) · J. Z. H. Zhang (✉)
State Key Laboratory of Precision Spectroscopy and Department of Physics, Institute of Theoretical and Computational Science, East China Normal University, Shanghai 200062, China
e-mail: ymei@phy.ecnu.edu.cn
e-mail: john.zhang@nyu.edu

L. L. Duan · B. Tang
College of Chemistry, Chemical Engineering and Materials Science, Engineering Research Center of Pesticide and Medicine Intermediate Clean Production, Ministry of Education, Key Laboratory of Molecular and Nano Probes, Shandong Normal University, Jinan 250014, China

J. Z. H. Zhang
Department of Chemistry, New York University, New York, NY 10003, USA

Introduction

The hydrophobic effect and hydrogen bonding are two main driving forces that stabilize proteins in solution [1–14]. The hydrophobic effect is the observed tendency of nonpolar groups of proteins or substances to aggregate in aqueous solution and exclude water molecules [7, 15]. Hydrophobic effect is neither due to a specific attractive physical force between nonpolar molecules nor a repulsive physical force between such molecules and water. It is mainly an entropy-driven free energy effect originating from the disruption of dynamic hydrogen bonds by the nonpolar solute, which causes reduced mobility of water molecules in the solvation shell of the non-polar solute [16]. The hydrophobic interaction is a major factor that drives the folding of globular protein with hydrophobic cores by the aggregation or cluster

of nonpolar amino acids including alanine, valine, leucine, isoleucine, phenylalanine, proline, tryptophan and methionine [7, 17–21]. Many self-assembly processes are also driven by the hydrophobic interaction, such as micelle formation, vesicles and bilayers [22]. For example, hydrophobic interaction is used to account for why water and oil do not mix, understand the structure of proteins that have hydrophobic amino acids clustered together, and why biomembranes hold together. Thus, in order to properly describe the structure and dynamics of proteins, especially those systems with hydrophobic packing or aggregation, it is critical to properly include hydrophobic interaction effectively in molecular dynamics simulation.

In molecular dynamics (MD) simulation with explicit water molecules, such hydrophobic effect is explicitly included. However, explicit water MD simulation is expensive, especially for large protein systems and for long simulation time. Implicit water MD simulation method such as generalized Born (GB) model provides an efficient means for such simulation studies. Because water molecules are not explicitly present in GB model, MD simulation is much faster and therefore can explore much longer time dynamics behavior of the protein systems than explicit water MD simulation could reach. Thus, implicit water MD simulation is a desired tool for studying large protein systems and longer time dynamics behavior. In the current implementation of the GB model such as in AMBER, the hydrophobic effect is taken into account by a free energy term determined by the solvent accessible surface area (SASA). However, the effect of this hydrophobic energy is not reflected in MD simulation, *i.e.*, there is no effective hydrophobic force that drives the dynamics of the proteins. Because the hydrophobic interaction energy is not an explicit function of atomic coordinate, it is not clear how to include it explicitly in implicit water MD simulation. This issue has been discussed by some researchers [23, 24]. As a result, MD trajectory in the current implementation of the GB model in the widely used AMBER program does not contain hydrophobic effect, which could be a serious problem for simulating certain protein systems as will be shown in this study. However, there is an implementation of hydrophobic force in the AMBER package using the LCPO method that calculates the SASA approximately [25]. In this method, the solvent accessible surface areas (SASAs) and the first and second derivative of it are with respect to atomic coordinates, so the hydrophobic force can be estimated by the derivative of SASA.

In this work, we propose a form of explicit hydrophobic force for MD simulation of protein systems in the GB model. This explicit hydrophobic force is based on the correlation between SASA and radius of gyration (R_g) in globular protein. Since the nonpolar solvation free energy is proportional to SASA which can also be expressed as a function of R_g, the derivative of this energy with respect to R_g can be used as an effective hydrophobic force. The present work is focused on studying the dynamical stability of five protein systems in which helix chains are held together by close packing. The five specific systems that we studied here are from Protein Data Bank: 1ZIK, 1COS, 2IPZ, 1CE9, 3KIK. These are multi-chain proteins that are held together by hydrophobic packing. Specifically, we performed three types of comparative MD simulations for these five proteins using the GB model. The first type of simulation employs the straightforward GB method (denoted gbsa=0) which does not include the hydrophobic force in AMBER, while the second and third type of simulation add explicitly the hydrophobic force. In the second type of simulation uses the LCPO method (denoted gbsa=1) in AMBER, and the third type of simulation uses the packing enforced GB (PEGB) method in which a specific hydrophobic force is added in MD using the method described in the Methods section of this paper. Our study shows clearly that the standard GB simulation unfolds the protein systems within a short simulation time. Although using the LCPO method three protein systems are stable, two other proteins still unfold in MD simulation. In contrast, the PEGB simulation correctly preserves the stable protein structures, demonstrating the critical importance of hydrophobic effect in maintaining the folded structure of these multi-chain, packed protein systems.

Methods

In the GB model, the nonpolar contribution to the solvation free energy in AMBER is approximated by the formula

$$G_{np} = \gamma SASA + \beta, \tag{1}$$

where SASA is the solvent accessible surface area, γ is the surface tension parameter and β is the offset. This energy term represents the hydrophobic energy in the GB model. Since in the implicit model, there is no explicit water molecules present in the system to exert an effective hydrophobic force on the protein in MD simulation, it is not clear how to implement this hydrophobic interaction in MD trajectory of GB model. In fact in the current implementation of standard GB model in AMBER package, this hydrophobic interaction is simply ignored in MD simulation, and Eq. 1 is only used when solvation energy calculation is needed. Thus current MD trajectory that determines the structure of the protein system is not affected by this hydrophobic interaction. As a result, there is actually no "hydrophobic force" in MD simulation using standard GB model. This lack of hydrophobic effect presents a serious problem for simulating systems with hydrophobic packings. As will be shown in the following study, the protein systems could rapidly unfold in MD simulation using the present standard GB model.

In order to overcome this "hydrophobic catastrophe" in current MD simulation using the GB method, we introduce an effective hydrophobic force which is consistent with the hydrophobic free energy defined in the GB model (Eq. 1). In a globular protein, *SASA* is approximately proportional to the square of the radius of gyration (R_g). Therefore, we can define a relation

$$\gamma SASA + \beta = G_{np} = CR_g^2, \tag{2}$$

where the coefficient C is easily determined from the calculated *SASA* and R_g. If we assume that C is a slowly varying function of R_g, then we can approximately evaluate an effective hydrophobic force acting on atom i by the relation

$$F_{np} = -\frac{\partial G_{np}}{\partial r_i} = -C\frac{\partial R_g^2}{\partial r_i} = -2C\frac{m_i(r_i - R_0)}{\sum_i m_i}, \tag{3}$$

where the definition for the radius of the gyration

$$R_g = \frac{m_i(r_i - R_0)^2}{\sum_i m_i}, \tag{4}$$

was used, and $\boldsymbol{R_0}$ is the center of mass of the protein system. By adding the force $\boldsymbol{F}_{np}$ on every atom, one effectively includes a hydrophobic force in MD simulation of GB method. In actual implementation, the force $\boldsymbol{F}_{np}$ is calculated at every MD step. We call this method "packing-enforced GB" or PEGB method.

Results and discussion

MD simulation

The native structures of five proteins (1ZIK, 1COS, 2IPZ, 1CE9, 3KIK) from Protein Data Bank (PDB) are taken as the starting structures. These five protein systems have something in common, *i.e.*, they are comprised of more than one helices packed by hydrophobic interaction, ranging from two-stranded helices to eight-stranded tetramer. In our study, three MD simulations, one with the standard GB model, one with the LCPO method and the other with PEGB method described in the previous section, are carried out for each of the five systems. The AMBER03 force field [26] is employed in the MD simulation. After an initial energy minimization, the protein system is heated to 300 K, followed by MD simulation with a time step of 2 fs. SHAKE algorithm [27] is used to fix all chemical bonds involving hydrogen atoms and the salt concentration is set to 0.2 M. The Langevin dynamics [28] with a collision frequency of 1.0 ps^{-1} is applied to regulate the temperature. The dielectric constant of the protein interior and of the solvent is set, respectively, to 1.0 and 78.5, respectively. The trajectory is saved every 1 ps and the corresponding snapshots are stored for further analysis. The GB model (igb5) of Onufriev, Bashford, and Case [29] is used in MD simulation. In PEGB method, the hydrophobic force given by Eq. 3 is updated at every time step.

Packed protein systems

Two, three, and four-stranded helices: 1ZIK, 1COS, and 2IPZ

The above three packed proteins systems are briefly described first. 1ZIK is a two-stranded, parallel coiled-coil leucine zipper core of the yeast transcriptional activator protein GCN4 [30, 31]. A single leucine zipper consists of multiple leucine residues at approximately 7-residue intervals and the helices are held together by hydrophobic interactions between leucine residues located on one side of each helix like a zipper [32]. 1COS, designed by Lovejoy *et al.*, [33] is a triple-stranded coiled coil formed by three α-helices, which are largely stabilized through hydrophobic interactions between leucine side chains. Lovejoy *et al.* believed that the hydrophobic interaction is a dominant factor in the stabilization of coiled coils [33]. 2IPZ is a four-stranded, parallel coiled coil in GCN4 leucine zipper [34].

In the present work, MD simulations were carried out for 150 ns/500 ns/500 ns, 150 ns/200 ns/200 ns and 100 ns/100 ns/100 ns using, respectively, GB, LCPO and PEGB methods for the three protein systems. The dynamics simulation showed that the 1ZIK, 1COS and 2IPZ systems quickly unpack in the standard GB simulation and these helical chains begin to separate at around 100 ns, 0.7 ns and 1.5 ns, respectively, obviously due to the lack of hydrophobic force as shown in Figs. 1, 2, and 3. However, by adding the hydrophobic force using either the LCPO and PEGB methods in MD simulation, these helical chains remain packed. Figures 1, 2, and 3 plots the comparison between the native structure and final structures of 1ZIK, 1COS and 2IPZ resulting from MD simulation using, respectively, the standard GB, LCPO and PEGB methods. In the standard GB simulation, those helical chains are completely separated as shown in Figs. 1, 2, and 3. The 3-strand 1COS separated into a 2-strand helices and a single strand helix (Fig. 2a) while the 4-strand 2IPZ separated into a 3-strand helix bundle and a single strand helix (Fig. 3a). In contrast, those helical chains remained compact in LCPO and PEGB simulations with the explicit hydrophobic force during the entire simulation time. The RMSD of backbone atom of 1ZIK, 1COS, 2IPZ as a function of the simulation time up to a total of 500 ns, 200 ns and 100 ns, respectively, are shown in Figs. 1, 2, and 3.

For 1ZIK system, the two trajectories generated by using LCPO and PEGB methods do not show much difference,

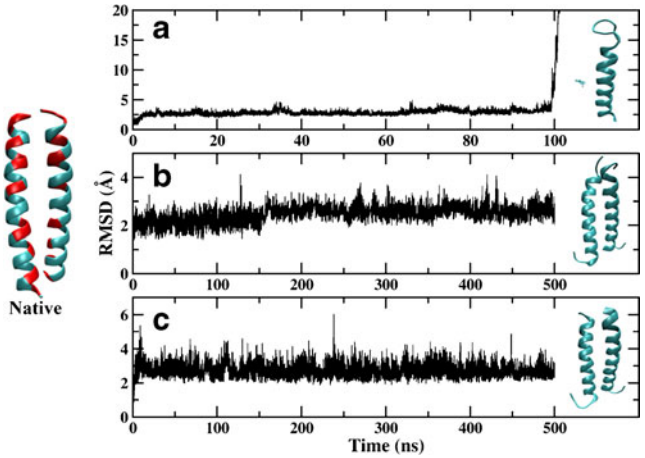

Fig. 1 The native structure of two-stranded helical system 1ZIK (*shown in cyan*) with the hydrophobic amino acids (*shown in red*). **a** RMSD of the backbone atoms as a function of simulation time and the simulated final structure using the standard GB method (without explicit hydrophobic force). **b** The same as (**a**) but using the LCPO method. **c** The same as (**a**) but using the PEGB method

they fluctuated with a RMSD of about 3.0 Å as shown in Fig. 1b–c. This confirms that the system remains compact throughout the simulation time by adding the hydrophobic force. For comparison, the RMSD in the standard GB simulation also fluctuated around 3.0 Å in the initial 100 ns and then it underwent a rapid rise to about 125 Å at 140 ns (Fig. 1a), indicating that the two chains have separated due to the lack of hydrophobic force in standard GB model. This hydrophobic catastrophe is more severe for 1COS system as shown in Fig. 2a. For example, the RMSD quickly rises from an initial value of about 2.0 Å all the way to around 170 Å after just 120 ns, indicating widely separated helices. For comparison, the RMSDs from LCPO and PEGB simulations remain compact throughout the entire MD simulation time of 200 ns (Fig. 2b–c). Similar to the three-stranded 1COS, 2IPZ quickly unpacks within a few ns of standard GB simulation as shown in Fig. 3a. The RMSD from

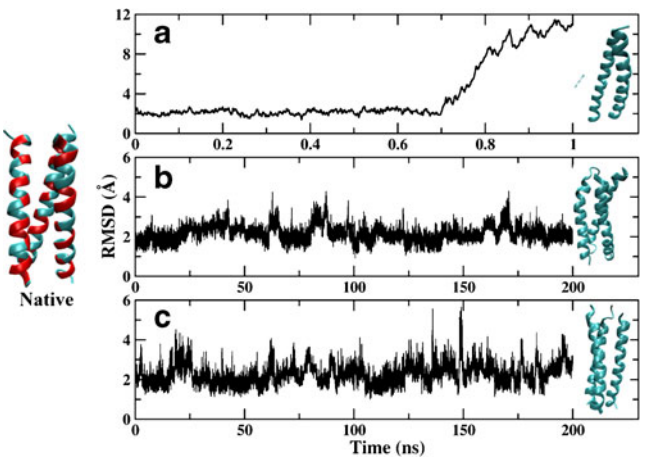

Fig. 2 The same as Fig. 1 but for 1COS

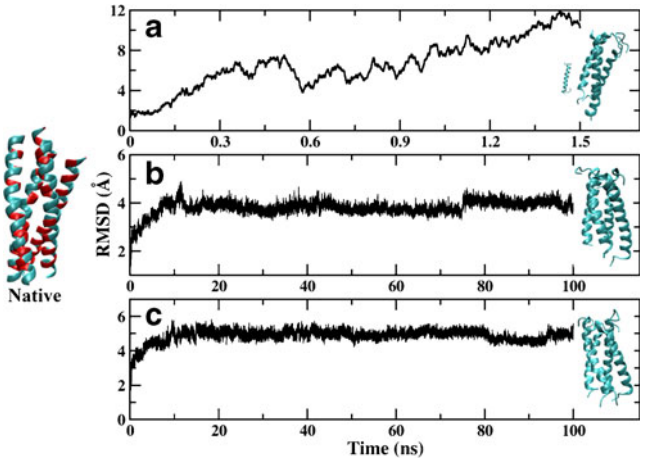

Fig. 3 The same as Fig. 1 but for 2IPZ

standard GB simulation increased to 68 Å demonstrating clearly that the packed four-strand helix bundles were well separated apart. However, 2IPZ remains well packed throughout MD simulation using both LCPO and PEGB method and the dynamic structures of the proteins are in close agreement with the corresponding native structure. The RMSDs are generally fluctuated near 4.0 Å and 5.0 Å over the entire simulation time of 100 ns respectively (Fig. 3b–c).

Solvent accessible surface area (SASA), which represents the hydrophobic energy and characterizes the compactness of the protein system, is often used as an analysis tool by structural biologists [35]. The concept of SASA defined by a probe rolling over the protein surface was first introduced by Lee and Richards [36] as a way of quantifying hydrophobic burial. The hydrophobic effect can cause nonpolar amino acids to aggregate together in the protein interior, whereas the polar amino acids tend to maximize the contacts with the outer solvent molecules when chains undergo dimerization, trimerization, tetramerization and *et al.* So the stability of protein structure is related to the burial of the nonpolar amino acids and can be measured by the loss of SASA of the protein [37]. Here, the SASA is calculated using the LCPO algorithm [25] by the "cpptraj" module in the AMBER package. The computed SASAs from the LCPO and PEGB simulation remain stable for 1ZIK system, indicating that the compactness of the protein system is maintained throughout MD simulation. In contrast, the SASA in the standard GB simulation is much larger than the corresponding trajectories of LCPO and PEGB, and starts to rise quickly at near 100 ns, indicating that the two helices begin to separate quickly (see Fig. S1(A1-A3) in the Supporting Information). The above result shows that the packed helical chains start to depart from each other quickly at around 100 ns in the standard GB simulation, but remain compact throughout the LCPO and PEGB simulation. Those

results of SASA for 1COS and 2IPZ are consistent with the two-strand 1ZIK system, showing the critical effect of hydrophobic force needed to stabilize the packing of this system. It is interesting to note that the computed SASA is slightly larger in simulation using PEGB method than using LCPO method for these three systems (see Fig. S1(A1-C3) in Supporting information for more details).

Four-stranded helices 1CE9

1) Standard GB result

1CE9 is another four-stranded, parallel coiled coil in GCN4 leucine zipper [38]. MD was carried out up to 40 ns using GB, LCPO and PEGB simulations. Similar to the previous three systems, the systems quickly unpacks in 2 ns during standard GB simulation as shown in Fig. 4a. The 4-strand 1CE9 separated into two 2-strand helices and the RMSD from standard GB simulation increased to 96 Å demonstrating clearly that the packed four-strand helix bundles are well separated apart duo to lack the hydrophobic force.

2) LCPO result

However, when adding the hydrophobic force using the LCPO method, starting from the native structure, the backbone RMSD of 1CE9 increased quickly to even 18 Å. Analyzing the MD trajectory finds that those helical chains still quickly unpack to two 2-strand helices and the beginning apart time are at around 0.5 ns which is earlier than in the standard GB simulation, as shown in Fig. 4b. It is known that the unpacking of the structure generally increases the SASA value. However, as seen from Fig. 6a, although the complex structure is separated, the SASA of this protein had been decreasing during the simulation. Further analysis finds that these separated 2-stranded monomers are over-packed in LCPO simulation. We know that adding hydrophobic force should help reduce the SASA value. This means that adding the hydrophobic force by LCPO method actually accelerated the unpacking of this protein.

3) PEGB result

On the contrary, 1CE9 remains well packed throughout MD simulation using the PEGB method (Fig. 4c) and the dynamic structure of the proteins are in close agreement with the corresponding native structures. Although in the beginning of simulation, the RMSD ascends to about 8.0 Å, the hydrophobic force puts the helical chains back very quickly and the RMSD were generally around 5.0 Å over the entire 40 ns simulation. It can be seen that the value of SASA using LCPO method is the smallest among these three models and the value of SASA using GB model is the largest (see Fig. S1(D1-D3) in Supporting information). Because the hydrophobic force is obtained by the derivative of the cavitation energy in terms of radius of gyration (R_g) in the PEGB method, it is necessary to analyze the variation of R_g along the MD trajectory. As shown in Fig. 7a, the general trend is a reduction in R_g due to the effect of hydrophobic force until the protein gets to a stable state which is a reasonable result.

Eight-stranded tetramer 3KIK

1) Standard GB result

3KIK is a tetramer and each monomer is comprised of a 2-stranded helix, with a total of eight helical chains bundled together. The stability of the structure is achieved by hydrophobic interactions between the helices [39]. Similarly, starting from the native structure, those helical chains begin to separate at 0.3 ns in standard GB simulation and the tetramer finally separated into two monomers and a dimer. The corresponding RMSD of the complex increased quickly to about 80 Å as shown in Fig. 5a.

2) LCPO result

MD simulation was carried out for 40 ns using LCPO method for 3KIK protein. The unpacking of this protein is

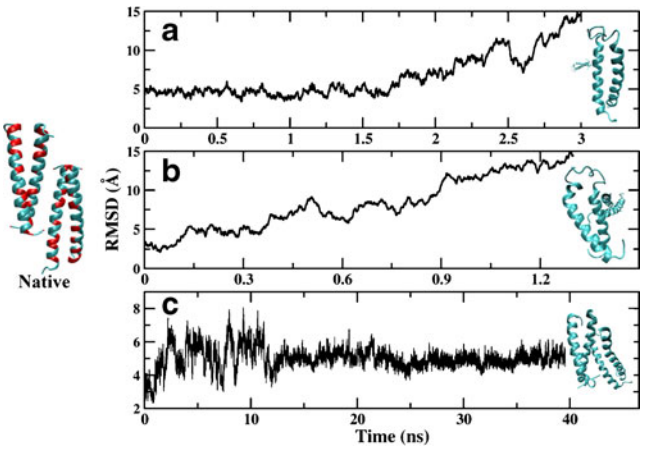

Fig. 4 The same as Fig. 1 but for 1CE9

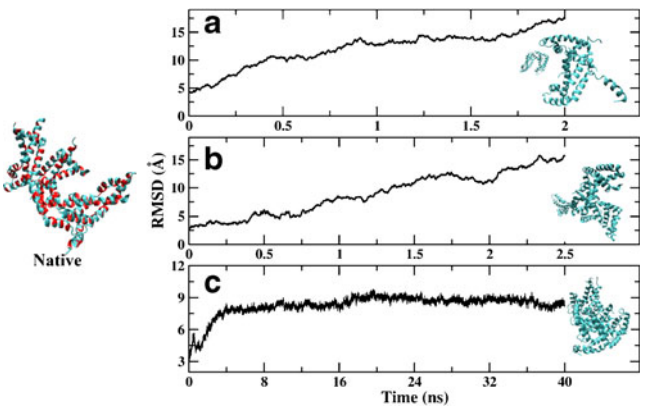

Fig. 5 The same as Fig. 1 but for 3KIK

slower than in the standard GB simulation as shown in Fig. 5b. The tetramer separated into a trimer and a monomer and the RMSD quickly rose to about 50 Å in the simulation. The computed SASA shows a rapid rise initially and then decreases as a function of simulation time as shown Fig. 6b. This initial rise of SASA delayed the unpacking of the system due to the hydrophobic force in LCPO simulation. This explained why the unpacking of the system is slower in LCPO than in standard GB simulations. However, the LCPO method still cannot maintain the packed structure of 3KIK during the simulation, the trimer shrinked and significantly decreased SASA while the monomer shows no structural change.

3) PEGB result

The dynamic RMSD in 40 ns PEGB simulation and the final structure are shown in Fig. 5c, the protein remains compact with the value of RMSD generally around 7.5 Å. And the computed SASAs based on the simulation using LCPO and PEGB model are obviously smaller than using GB model. The value of SASA is slightly smaller using PEGB model than using LCPO method (Fig. S1(E1-E3)). Last, we also investigate the R_g which is shown in Fig. 7b. The R_g gradually declines and it moves toward a stable state which is similar to that found in 1CE9 system. Those results demonstrated the critical effect of hydrophobic interaction that helps keep the system packed in solvent.

Conclusions

In this work, we presented a theoretical method (PEGB) to include an explicit hydrophobic force for MD simulation of packed protein systems in implicit water GB model. The method is based on a correlated relationship between SASA and the radius of gyration of the packed protein systems. The method has been applied to study the stability of five protein systems, representing, respectively, 2-stranded, 3-stranded, 4-stranded and 8-stranded helix bundles. These five protein systems are all packed together due to hydrophobic interaction. Our study shows that these packed protein systems begin to unfold or de-pack at 100 ns, 0.7 ns, 1.5 ns, 2 ns, 0.3 ns, respectively, in standard GB simulation without enforcement of hydrophobic force. In the AMBER package using the LCPO method can include explicit hydrophobic force based directly on solvent accessible surface area (SASA), while our study found that three of the five protein systems remained well packed during the simulation, and two other systems were still unpacked during LCPO simulation. In contrast, with the addition of an explicit hydrophobic force based on the radius of gyration of the protein complex in the packing enforced GB (PEGB) method, these protein systems remain compact throughout simulation time limit of up to 500 ns.

In the LCPO method, if the unpacking of the structure decreases the SASA value, the adding of the hydrophobic force will accelerate the unpacking, and we also find subunit shrinks after the unpacking of the protein. While, in the PEGB method, the hydrophobic force is determined by the R_g of the protein. When the R_g becomes larger with the protein unpacking, the hydrophobic force will draw the submit back and maintains the whole structure. No matter how the SASA changes, the hydrophobic force can constrain the protein in a reasonable position and prevent the protein unpacking in PEGB method. It can be noted that the computed R_g is smaller than that of native structure, the rationality of structure will be studied in the future.

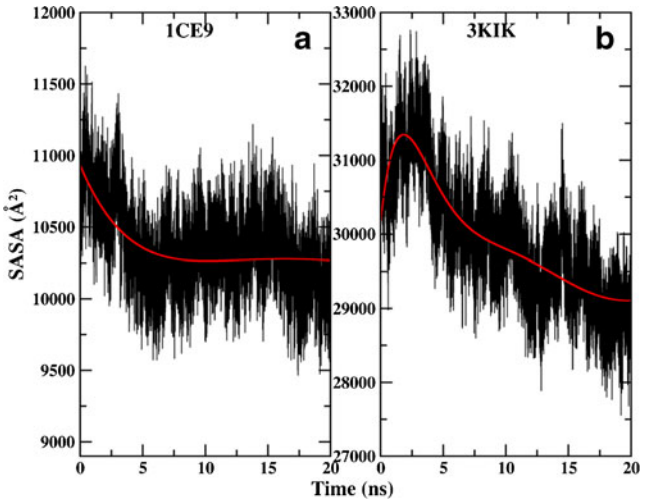

Fig. 6 The calculated solvent accessible surface area (SASA) as a function of MD simulation time using the LCPO method for 1CE9 (**a**) and 3KIK (**b**), respectively

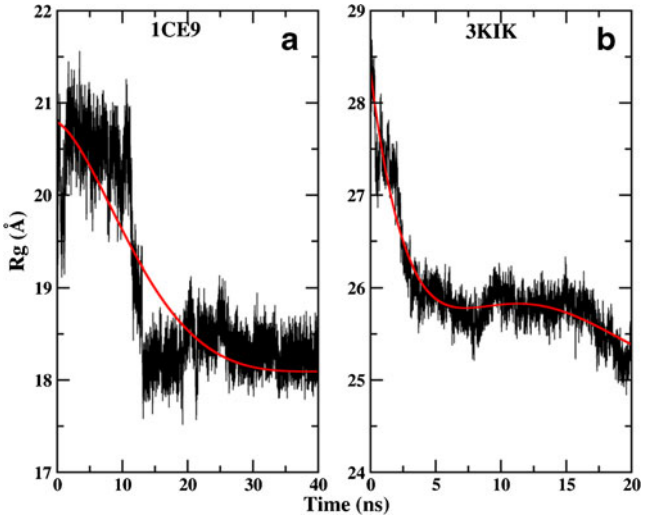

Fig. 7 The radius of gyration (R_g) of the whole protein as a function of MD simulation time using the PEGB models for 1CE9 (**a**) and 3KIK (**b**), respectively

The PEGB model is intended to be an efficient computational tool for dynamics study of large and/or complex protein systems with hydrophobic packing for which long time explicit water MD simulation is very expensive or too expensive to perform. It is envisioned to apply to large and more complex proteins systems in which hydrophobic packing interactions are important, such as the formation of lipid bilayers and micelles, insertion of membrane proteins into the nonpolar lipid environment.

To further support the conclusion of this work, two additional MD runs were carried out for the 1CE9 system. One is MD simulation of 1CE9 in explicit water (TIP3P) and the other is a separate trajectory simulation using PEGB method. The result from both simulation runs prove that the system is stable as shown in Figs. S3 and S4 in the Supporting information.

Acknowledgments L.L.D. is grateful to the National Natural Science Foundation of China (Grants 11147026 and 31200545) and the Scientific Research Award Fund for the Excellent Middle-Aged and Young Scientists of Shandong Province (Grant BS2011SW046) for support. Y.M. was supported by the Shanghai Rising-Star program (Grant11QA1402000). Q.G.Z. thanks the National Natural Science Foundation of China (Grant 11274206) for support. B.T. was supported by 973 Program (2013CB933800) and the National Key Natural Science Foundation of China (Grants 21035003 and 21227005). J.Z.H.Z. acknowledges financial support from the National Natural Science Foundation of China (Grants 10974054 and 20933002) and the Shanghai PuJiang program (Grant 09PJ1404000). We thank the Supercomputer Center of East China Normal University for CPU time support.

References

1. Pauling L, Corey RB, Branson HR (1951) Two hydrogen-bonded helical configurations of the polypeptide chain. Proc Natl Acad Sci U S A 37(4):205–211
2. Avbelj F, Baldwin RL (2006) Limited validity of group additivity for the folding energetic of the peptide group. Proteins 63(2):283–289
3. Avbelj F, Baldwin RL (2009) Origin of the change in solvation enthalpy of the peptide group when neighboring peptide groups are added. Proc Natl Acad Sci U S A 106(9):3137–3141
4. Bernal JD (1939) Structure of proteins. Nature 143(3625):663–667
5. Bolen DW, Rose GD (2008) Structure and energetics of the hydrogen-bonded backbone in protein folding. Annu Rev Biochem 77(1):339–362
6. Chen J, Stites WE (2001) Packing is a key selection factor in the evolution of protein hydrophobic cores. Biochemistry 40(50):15280–15289
7. Kauzmann W (1959) Some factors in the interpretation of protein denaturation. Adv Protein Chem 14:1–63
8. Pace C, Shirley B, McNutt M, Gajiwala K (1996) Forces contributing to the conformational stability of proteins. FASEB J 10(1):75–83
9. Rose G, Fleming P, Banavar J, Maritan A (2006) A backbone-based theory of protein folding. Proc Natl Acad Sci U S A 103(45):16623–16633
10. Tanford C (1962) Contribution of hydrophobic interactions to the stability of the globular conformation of proteins. J Am Chem Soc 84(22):4240–4247
11. Dill KA (1990) Dominant forces in protein folding. Biochemistry 29(31):7133–7155
12. Mirsky AE, Pauling L (1936) On the structure of native, denatured, and coagulated proteins. Proc Natl Acad Sci U S A 22(7):439–447
13. Pauling L, Corey RB (1951) A proposed structure for nucleic acids. Proc Natl Acad Sci U S A 37(11):729–740
14. Fersht AR (1987) The hydrogen-bond in molecular recognition. Trends Biochem Sci 12:301–304
15. Chandler D (2005) Interfaces and the driving force of hydrophobic assembly. Nature 437(7059):640–647
16. Silverstein TP (1998) The real reason why oil and water don't mix. J Chem Educ 75(1):116–346
17. Charton M, Charton BI (1982) The structural dependence of amino acid hydrophobicity parameters. J Theor Biol 99(4):629–644
18. Gilmanshin R, Dyer RB, Callender RH (1997) Structural heterogeneity of the various forms of apomyoglobin: implications for protein folding. Protein Sci 6(10):2134–2142
19. Arai M, Kuwajima K (1996) Rapid formation of a molten globule intermediate in refolding of alpha-lactalbumin. Fold Des 1(4):275–287
20. Agashe VR, Shastry MC, Udgaonkar JB (1995) Initial hydrophobic collapse in the folding of barstar. Nature 377(6551):754–757
21. Vidugiris GJ, Markley JL, Royer CA (1995) Evidence for a molten globule-like transition state in protein folding from determination of activation volumes. Biochemistry 34(15):4909–4912
22. Meyer EE, Rosenberg KJ, Israelachvili J (2006) Recent progress in understanding hydrophobic interactions. Proc Natl Acad Sci U S A 103(43):15739–15746
23. Brooks BR, Brooks CL III, Mackerell AD Jr, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M (2009) CHARMM: the biomolecular simulation program. J Comput Chem 30(10):1545–1614
24. Lee MS, Feig M, Salsbury FR Jr, Brooks CL III (2003) New analytic approximation to the standard molecular volume definition and its application to generalized born calculations. J Comput Chem 24(11):1348–1356
25. Weiser J, Shenkin PS, Still WC (1999) Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO). J Comput Chem 20(2):217–230
26. Pande VS, Baker I, Chapman J, Elmer SP, Khaliq S, Larson SM, Rhee YM, Shirts MR, Snow CD, Sorin EJ, Zagrovic B (2003) Atomistic protein folding simulations on the submillisecond time scale using worldwide distributed computing. Biopolymers 68(1):91–109
27. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical integration of the Cartesian equations of motion of a system with constrains: molecular dynamics of n-alkanes. J Comput Phys 23(3):327–341
28. Pastor RW, Brooks BR, Szabo A (1988) An analysis of the accuracy of langevin and molecular mynamics algorithm. Mol Phys 65(6):1409–1419
29. Onufriev A, Bashford D, Case DA (2004) Exploring protein native states and large-scale conformational changes with a modified generalized born model. Proteins Struct Funct Bioinforma 55(2):383–394
30. Harbury PB, Kim PS, Alber T (1994) Crystal structure of an isoleucine-zipper trimer. Nature 371(6492):80–83
31. Gonzalez L Jr, Woolfson DN, Alber T (1996) Buried polar residues and structural specificity in the GCN4 leucine zipper. Nat Struct Biol 3(12):1011–1018
32. Krylov D, Mikhailenko I, Vinson C (1994) A thermodynamic scale for leucine zipper stability and dimerization specificity: e and g interhelical interactions. EMBO J 13(12):2849–2861

33. Lovejoy B, Choe S, Cascio D, McRorie D, DeGrado W, Eisenberg D (1993) Crystal structure of a synthetic triple-stranded α-helical bundle. Science 259(5099):1288–1293
34. Liu J, Deng Y, Zheng Q, Cheng CS, Kallenbach NR, Lu M (2006) A parallel coiled-coil tetramer with offset helices. Biochemistry 45 (51):15224–15231
35. Cavallo L, Kleinjung J, Fraternali F (2003) A fast algorithm for solvent accessible surface areas at atomic and residue level. Nucleic Acids Res 31(13):3364–3366
36. Lee B, Richards FM (1971) The interpretation of protein structures: estimation of static accessibility. J Mol Biol 55(3):379–400
37. Becue A, Meurice N, Leherte L, Vercauteren DP (2004) Evaluation of the protein solvent-accessible surface using reduced representations in terms of critical points of the electron density. J Comput Chem 25(9):1117–1126
38. Lu M, Shu W, Ji H, Spek E, Wang LY, Kallenbach NR (1999) Helix capping in the GCN4 leucine zipper. J Mol Biol 288(4):743–752
39. Ellisdon AM, Jani D, Köhler A, Hurt E, Stewart M (2010) Structural basis for the interaction between yeast Spt-Ada-Gcn5 acetyltransferase (SAGA) complex components Sgf11 and Sus1. J Biol Chem 285(6):3850–3856