THE JOURNAL OF PHYSICAL CHEMISTRY B LETTERS

**2006,** *110,* 15045−15048
Published on Web 07/20/2006



# Mixed Atomistic and Coarse-Grained Molecular Dynamics: Simulation of a Membrane-Bound Ion Channel

**Qiang Shi, Sergei Izvekov, and Gregory A. Voth***

*Center for Biophysical Modeling and Simulation and Department of Chemistry, University of Utah, 315 S. 1400 E. Rm 2020, Salt Lake City, Utah 84112-0850*

*Received: May 2, 2006; In Final Form: July 3, 2006*

The recently developed multiscale coarse-graining (MS-CG) method (Izvekov, S.; Voth, G. A. *J. Phys. Chem. B* **2005**, *109*, 2469; *J. Chem. Phys.* **2005**, *123*, 134105) is used to build a mixed all-atom and coarse-grained (AA-CG) model of the gramicidin A (gA) ion channel embedded in a dimyristoylphosphatidylcholine (DMPC) lipid bilayer and water environment. In this model, the gA peptide was described in full atomistic detail, while the lipid and water molecules were described using coarse-grained representations. The atom-CG and CG-CG interactions in the mixed AA-CG model were determined using the MS-CG method. Molecular dynamics (MD) simulations were performed using the resulting AA-CG model. The results from simulations of the AA-CG model compare very favorably to those from all-atom MD simulations of the entire system. Since the MS-CG method employs a general and systematic approach to obtain effective interactions from the underlying all-atom models, the present approach to rigorously develop mixed AA-CG models has the potential to be extended to many other systems.

Computer simulations play an increasingly important role in understanding the structure and dynamics of biomolecular systems. However, molecular dynamics (MD) simulations of large systems (hundreds of thousands of atoms) for long time scales (micro- to milliseconds) using all-atom models are extremely expensive, if they are even possible at the present time. An alternative to all-atom models is to employ coarse-grained (CG) models which reduce computational costs by simplifying the description of the system, while, at the same time, retaining the ability to predict the properties of interest (for recent reviews, see, e.g., refs 1−3). Such CG approaches have been applied to a wide range of problems, including lipid bilayer structure and dynamics,[4−11] protein folding and design,[12−17] protein−protein interactions,[18,19] and nanotube or protein interactions with lipid bilayers.[20−22] Although CG models have been successfully used to achieve fast simulations for large systems, the lack of detailed atomistic information in CG simulations still limits the systems and processes that can be studied using these models. A very attractive approach to incorporate atomistic details into the CG models is therefore to create a mixed all-atom and coarse-grained (AA-CG) model,[23] in which the most interesting part of the system is represented in full atomistic detail, and the remaining parts are modeled at the CG level. In fact, the idea of using a mixed description for large systems is not unfamiliar to the molecular modeling community: recent years have seen considerable success in the combined quantum mechanical and molecular mechanical (QM/MM) approaches (see, e.g., refs 24 and 25 for recent reviews). Similar to the QM/MM approaches where the central problem is to construct the QM and MM interface, a major difficulty in building a consistent mixed AA-CG model is to define and accurately parametrize the interactions between the all-atom and CG subsystems. This is not just a technical issue but instead one involving the fundamental statistical mechanics of defining *N*-body mean interactions in complex systems. The present work therefore has broad and important implications for the field of molecular simulation beyond the specific application system (gramicidin) chosen to highlight the features of our particular approach.

There already exist widely used all-atom models for biomolecules.[26−29] The strategy to develop mixed AA-CG models thus depends greatly upon how the effective interactions describing the CG part of the model are built. For reasons of consistency and convenience in constructing the AA-CG coupling, it would be ideal to derive the effective potential for the CG part of system systematically from the same atomistic force field used to describe the all-atom part of the system. However, this is not the case for most existing CG approaches. Many CG models for proteins are structure-based Gō-like models[30] or elastic network models (ENMs).[31,32] A broad range of CG models for liquids and lipid bilayers are parametrized to reproduce thermodynamic properties.[6,7,33] There are also CG models that employ all-atom MD simulation data to obtain effective CG force fields. But rather than directly utilizing the atomic forces from all-atom MD simulations, these approaches focus on reproducing average structural properties such as radial distribution functions (RDFs)[6,8,34] or, similarly, the restricted free energies.[35] The multiscale coarse-graining (MS-CG) method[9,11,36,37] recently developed in our group represents a significantly different approach, which uses a new force-matching (FM) procedure to determine effective pairwise CG force field directly from trajectory and force data generated by all-atom MD simulations. As a general and systematic way to derive CG potentials from the underlying atomistic interactions,

10.1021/jp062700h CCC: $33.50 

the MS-CG method has recently been applied to various biomolecular[9,11] and liquid systems.[36,37] The fact that the MS-CG force field is derived directly from the underlying atomistic interactions, as well as the capability of the MS-CG method to treat complex CG schemes, makes it a powerful tool for building mixed AA-CG models systematically.

As an example, a mixed AA-CG model for gramicidin A (gA) polypeptide ion channel in a dimyristoylphosphatidylcholine (DMPC) lipid bilayer is developed in the present work. In this mixed model, the gA polypeptide, modeled in full atomistic detail, is embedded in a CG lipid bilayer and water environment. The peptide antibiotic gramicidin A (gA) is a small and well-characterized ion channel[38] which has served as a prototypical model for both experimental[39−41] and theoretical studies[42−44] of more complex ion channels. In this work, the gA/DMPC system was used as a model system to test the capability of the MS-CG method for developing mixed AA-CG models.

The mathematical algorithm employed by the MS-CG method has been presented elsewhere.[9,36] The reader is referred to the Supporting Information for details of the methodology. Since the MS-CG method is general and flexible even for complex CG schemes, an AA-CG model for a given system can be obtained in a rather straightforward way: All-atom MD simulations of the whole system are first carried out; the system is then divided into atomistic and CG parts by retaining the all-atom description for the most interesting part of the system and then mapping the remaining parts into CG descriptions. The effective AA-CG force field is subsequently obtained by treating all the atomistic and CG sites equally in the FM procedure. However, to reduce the computational cost, simplifications and approximations have been made in the current study. The interactions in the mixed AA-CG system are first classified into three different categories: atom−atom, CG−CG, and atom−CG interactions. Forces on all-atom and CG sites are then given by

$$\mathrm{F}_i^{\mathrm{atom}} = \sum_{j\in\mathrm{atom}, j\neq i} \mathrm{F}_{ij}^{\mathrm{atom-atom}} + \sum_{j\in\mathrm{CG}} \mathrm{F}_{ij}^{\mathrm{atom-CG}}$$

$$\mathrm{F}_i^{\mathrm{CG}} = \sum_{j\in\mathrm{atom}} \mathrm{F}_{ij}^{\mathrm{atom-CG}} + \sum_{j\in\mathrm{CG}, j\neq i} \mathrm{F}_{ij}^{\mathrm{CG-CG}} \qquad (1)$$

The atom−atom interactions in this mixed AA-CG model are assumed to be the same as in the all-atom model used to generate the trajectory and force data for the MS-CG method. This assumption greatly reduced the number of parameters to be determined.

The mixed AA-CG model is next developed in two stages. In the first stage, CG-CG interactions for the lipid bilayer and water are obtained by applying the MS-CG approach to the trajectory and force data obtained from all-atom MD simulation of pure lipid bilayer embedded in water. In the second stage, after the trajectory from the all-atom simulations of the whole system has been mapped to the mixed AA-CG representation, atom−atom and CG−CG forces are subtracted from the reference forces and the atom−CG interactions are obtained by fitting the residual forces using the MS-CG method. In principle, the two-stage approach described above is not as rigorous as the straightforward application of the MS-CG method mentioned early, although it was found in this work that the two-stage strategy works quite well for the current system.

Details of the all-atom simulations to generate input data for the MS-CG method are summarized in the Supporting Information. The CG scheme for the DMPC molecule is illustrated in Figure 1 (see also ref 9). The CG sites of the DMPC molecule are defined as the center of mass of the atomic groups they represent. The water molecules were mapped into single CG sites at their geometric centers, as this gave a better structure when compared to the all-atom simulation. Despite the fact that some CG sites of the DMPC molecule are charged in the underlying all-atom model, it was found that the CG force field without explicit Coulomb interactions was able to effectively account for the missing Coulomb terms, at least when fairly large cutoffs were used for the short-range interactions. The mesh spacings used in the FM procedure for the bonded and nonbonded forces were approximately 0.0025 and 0.05 nm, respectively. The CG force field obtained differs subtly from the one in ref 9, since a new version of our FM code which treats the bonded and nonbonded interactions separately[36] was used. The cutoff for the nonbonded interactions was 1.3 nm. A minimal bonding scheme was used in the CG lipid model where only bond and angle interactions are included. For simplicity, analytic forms were used to describe all the bonded interactions. Parameters for the bonded force were obtained using the method described in ref 37.

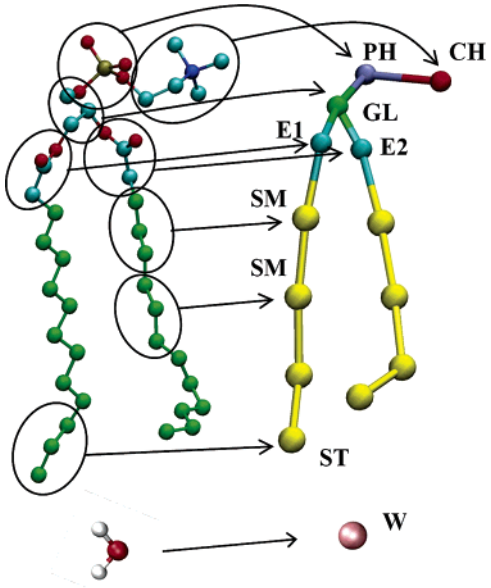


**Figure 1.** Coarse-grained representation of the DMPC and water molecules. The all-atom model for DMPC is coarse-grained to 13 interaction sites consist of 7 different types: the choline (CH), phosphate (PH), glycerol backbone (GL), ester (E1 and E2), alkane chain (SM), and tail of alkane chain (ST) groups. The water molecule is coarse-grained into one site (W).

The MD trajectory of the peptide−lipid bilayer system was reduced to the mixed AA-CG representation. For simplicity, water molecules inside the gA channel were also coarse-grained. The treatment of the water in the channel as CG sites is obviously a significant approximation if one is interested in the atomistic-scale details of the water in the channel (i.e., the hydrogen-bonding interactions in the water chain and between the water chain and the channel). However, in this work, the focus is on the structural integrity of the AA polypeptide in the CG bilayer environment. The resolution of the AA-CG modeling can be systematically refined to include explicit water in the channel if desired. The forces arising from atom−atom and CG−CG interactions were subtracted from the reference forces in the mixed AA-CG representation using the previously determined AA and CG force fields, and the parameters for the atom−CG interactions were fitted to the residual reference forces. Two different schemes to treat the Coulomb interactions

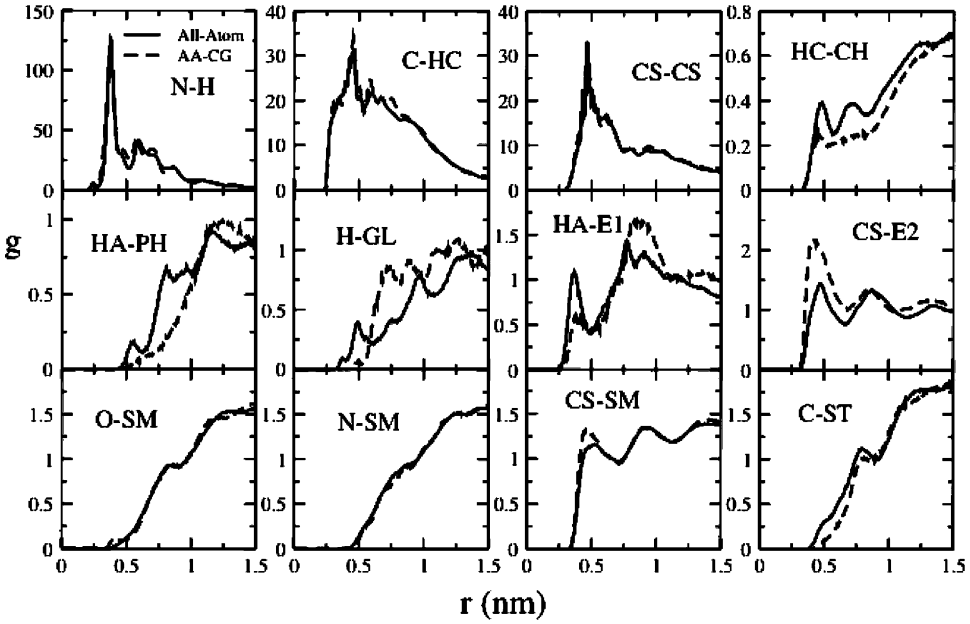


**Figure 2.** Comparison of the AA-CG and all-atom gA-gA and gA-CG lipid rdfs. C, O, N, H: peptide backbone atoms. CS: side chain carbon. HC: nonpolar and nonaromatic side chain hydrogen. HA: aromatic side chain hydrogen. Notations for the CG lipid are the same as in Figure 1.

were used in subtracting the atom−atom forces: PME and the truncated and shifted potential (with a cutoff radius of 1.3 nm). The atom−CG interactions differ slightly between the two schemes, but subsequent MD simulations using the AA-CG force field obtained from the two schemes showed no significant differences. Thus, the MS-CG method is able to pick up the difference in the atom−atom interactions and generate the appropriate atom−CG interactions accordingly. Results reported below were obtained using the truncated and shifted method. To accelerate the fitting of the atom−CG interactions, only the (residual) reference forces on the atom sites (i.e., AA part of the system) were included. This allowed the inclusion of more configurations in the FM procedure, and smoother force curves were obtained. Tabulated forces and potential energies of the atom−CG and nonbonded CG−CG interactions were then interpolated to a fine mesh for subsequent MD simulations. During the following AA-CG simulation, the lipid headgroups were found to come too close to the gA peptide. This artifact is mainly caused by the interaction of lipid headgroups with the TRP side chains. It is suspected that the attractive interactions between the two were slightly overestimated in the MS-CG procedure, and the problem became more severe because of the large number of atoms involved (several atoms on the side chain belong to the same atom type with the overestimated attractive interactions) and the rigid structure of the rings. Two interactions between the carbon atoms on the TRP side chain and the CH headgroup were manually adjusted to correct this artifact. In the future, we intend to develop a more systematic and theoretically justified approach to refining the accuracy of the AA-CG effective potential. Selected force profiles of the atom−CG interactions are plotted in Figure S1 (Supporting Information). The forces approach zero at long distance but show more complex structures at intermediate distances.

The AA-CG force field was used to perform a 10 ns constant NVT simulation of the peptide−lipid bilayer system. The initial configuration of the AA-CG MD simulation was adapted from one of the equilibrated all-atom configurations. A time step of 2 fs was used in the AA-CG simulation, which was the same as those used in the all-atom simulations. Because of the larger cutoff radius used for the nonbonded interactions, the multiple time step algorithm in *DL_POLY*[45] was used with inner and outer cutoffs set to 0.8 nm and 1.3 nm, respectively. The AA-CG simulation is approximately 7 times faster than the all-atom MD simulation. It is expected with further algorithmic improvements specifically targeted to the AA-CG approach that this speed-up can be significantly increased.

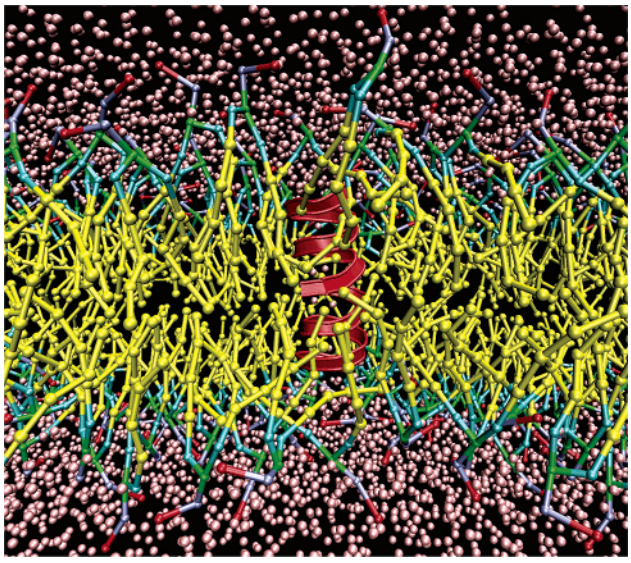

**Figure 3.** Configuration of the atomistic gA peptide in CG DMPC lipid and water, after 10 ns AA-CG MD simulation. The color scheme is the same as used in Figure 1. The image was prepared using VMD.[46]

Both the gA helical dimer and the CG lipid bilayer remained stable during the 10 ns AA-CG MD simulation. The gA−gA and gA−lipid radial distribution functions (RDFs) are compared to the all-atom MD result in Figure 2. It can be seen that the gA−gA RDFs from AA-CG simulation agree very well with the all-atom ones. The gA−lipid tail (SM, ST) RDFs are also very good, suggesting that the AA-CG model successfully reproduces the interactions between the peptide and the hydrophobic lipid tails. Since the gA polypeptide is shorter than the bilayer thickness, RDFs for gA−headgroups (CH, PH) are harder to model, and larger deviations are observed.

The final configuration of the 10 ns AA-CG trajectory is shown in Figure 3. The RMSD of the peptide backbone atoms

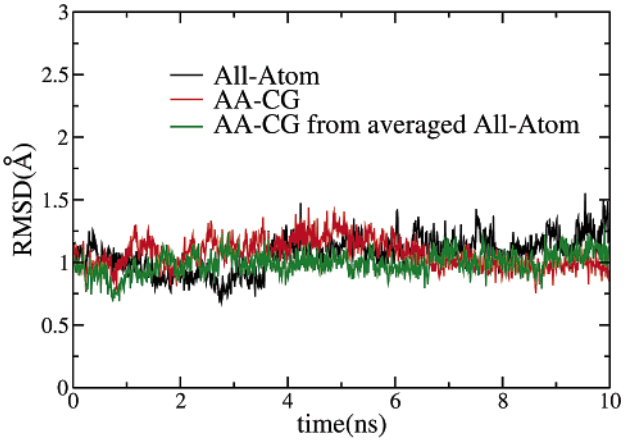


**Figure 4.** RMSD of gA polypeptide from the NMR structure (protein data bank entry 1MAG),[47] calculated over the backbone atoms: all-atom (black) and AA-CG (red) MD simulation. Also shown in green is the RMSD of gA in the AA-CG simulation from the averaged all-atom structure, which was calculated using the first 1 ns trajectory of the all-atom simulation.

in the AA-CG simulation was also calculated and compared to that of a 10 ns all-atom MD simulation in Figure 4. Also plotted in Figure 4 is the RMSD of AA-CG simulation from the averaged all-atom structure calculated using the first 1 ns trajectory of the all-atom MD simulation. The AA-CG RMSD from the NMR reference structure is seen to be similar to the all-atom MD case and the AA-CG RMSD from the averaged all-atom structure. The side-chain RMSD was also calculated after fitting the AA-CG backbone atoms to the averaged all-atom ones; the result is 1.6 Å, which is somewhat larger than the backbone RMSD. In general, the RMSD values are quite good considering the degree to which the lipid and solvent molecules have been coarse-grained.

In summary, it has been demonstrated in this work that the MS-CG method is sufficiently flexible and accurate to develop a mixed AA-CG model for a gA peptide channel embedded in a DMPC lipid bilayer and water environment, in which the all-atom description was retained for the gA, while CG descriptions were used for lipid and solvent molecules. The present approach is to be contrasted to other methods for developing AA-CG models[23] in which the AA-CG interface was constructed according to structural considerations. Since the MS-CG method is systematic and self-consistent, it should also be readily extended to many different kinds of systems, and such studies will be the focus of future research.

**Acknowledgment.** This research was supported by a grant from the National Institutes of Health (R01 GM63796). Computational resources were provided in part by the National Institutes of Health (grant NCRR 1 S10 RR17214-01) on the Arches Metacluster, administered by the University of Utah Center for High Performance Computing. Some of the computations were also performed on the National Science Foundation Teragrid system at the National Center for Supercomputing Applications and the Texas Advanced Computing Center (TG-MCA94P017), which are also gratefully acknowledged.

**Supporting Information Available:** Details of the MS-CG methodology and the all-atoms simulations to generate the data for the force-matching are provided in the Supporting Information. This material is available free of charge via the Internet at http://pubs.acs.org.

## References and Notes

(1) Nielsen, S. O.; Lopez, C. F.; Srinivas, G.; Klein, M. L. *J. Phys.: Condens. Matter* **2004**, *16*, R481.
(2) Tozzini, V. *Curr. Opin. Struct. Biol.* **2005**, *15*, 144.
(3) Chu, J.-W.; Izvekov, S.; Voth, G. A. *Mol. Simul.* In press, 2006.
(4) Smit, B.; Hilbers, P. A. J.; Esselink, K.; Rupert, L. A. M.; van Os, N. M.; Schlijper, A. G. *Nature (London)* **1990**, *348*, 624.
(5) Goetz, R.; Lipowsky, R. *J. Chem. Phys.* **1998**, *108*, 7397.
(6) Shelley, J. C.; Shelley, M. Y.; Reeder, R. C.; Bandyopadhyay, S.; Klein, M. L. *J. Phys. Chem. B* **2001**, *105*, 4464.
(7) Marrink, S. J.; de Vries, A. H.; Mark, A. E. *J. Phys. Chem. B* **2004**, *108*, 750.
(8) Murtola, T.; Falck, E.; Patra, M.; Karttunen, M.; Vattulainen, I. *J. Chem. Phys.* **2004**, *121*, 9156.
(9) Izvekov, S.; Voth, G. A. *J. Phys. Chem. B* **2005**, *109*, 2469.
(10) Stevens, M. J. *J. Chem. Phys.* **2004**, *121*, 11942.
(11) Izvekov, S.; Voth, G. A. *J. Chem. Theor. Comput.* **2006**, *2*, 637.
(12) Friedel, M.; Shea, J. E. *J. Chem. Phys.* **2004**, *120*, 5809.
(13) Brown, S.; Fawzi, N. J.; Head-Gordon, T. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 10712.
(14) Fujitsuka, Y.; Takada, S.; Luthey-Schulten, Z. A.; Wolynes, P. G. *Proteins* **2004**, *54*, 88.
(15) Smith, A. V.; Hall, C. K. *Proteins* **2001**, *44*, 344.
(16) Buchete, N. V.; Straub, J. E.; Thirumalai, D. *Protein Sci.* **2004**, *13*, 862.
(17) Onuchic, J. N.; Wolynes, P. G. *Curr. Opin. Struct. Biol.* **2004**, *14*, 70.
(18) Nguyen, H. D.; Hall, C. K. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 16180.
(19) Zacharias, M. *Protein Sci.* **2003**, *12*, 1271.
(20) Shih, A. Y.; Arkhipov, A.; Freddolino, P. L.; Schulten, K. *J. Phys. Chem. B* **2006**, *110*, 3674.
(21) Bond, P. J.; Sansom, M. S. P. *J. Am. Chem. Soc.* **2006**, *128*, 2697.
(22) Lopez, C. F.; Nielsen, S. O.; Moore, P. B.; Klein, M. L. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 4431.
(23) Neri, M.; Anselmi, C.; Cascella, M.; Maritan, A.; Carloni, P. *Phys. Rev. Lett.* **2005**, *95*, 218102.
(24) Gao, J. L.; Truhlar, D. G. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467.
(25) Friesner, R. A.; Guallar, V. *Annu. Rev. Phys. Chem.* **2005**, *56*, 389.
(26) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.
(27) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
(28) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657.
(29) Scott, W. R. P.; Hunenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Kruger, P.; van Gunsteren, W. F. *J. Phys. Chem. A* **1999**, *103*, 3596.
(30) Gō, N. *J. Stat. Phys.* **1983**, *30*, 413.
(31) Tirion, M. M. *Phys. Rev. Lett.* **1996**, *77*, 1905.
(32) Bahar, I.; Rader, A. J. *Curr. Opin. Struct. Biol.* **2005**, *15*, 586.
(33) Brannigan, G.; Brown, F. L. H. *J. Chem. Phys.* **2004**, *120*, 1059.
(34) Meyer, H.; Biermann, O.; Faller, R.; Reith, D.; Muller-Plathe, F. *J. Chem. Phys.* **2000**, *113*, 6264.
(35) Liwo, A.; Czaplewski, C.; Pillardy, J.; Scheraga, H. A. *J. Chem. Phys.* **2001**, *115*, 2323.
(36) Izvekov, S.; Voth, G. A. *J. Chem. Phys.* **2005**, *123*, 134105.
(37) Wang, Y.; Izvekov, S.; Yan, T.; Voth, G. A. *J. Phys. Chem. B* **2006**, *110*, 3564.
(38) Wallace, B. A. *J. Struct. Biol.* **1998**, *121*, 123.
(39) Koeppe, R. E., II; Andersen, O. S. *Annu. Rev. Biophys. Biomech.* **1996**, *25*, 231.
(40) Anderson, O. S. *Annu. Rev. Physiol.* **1984**, *46*, 531.
(41) Wallace, B. A. *Annu. Rev. Biophys. Chem.* **1990**, *19*, 127.
(42) Tieleman, D. P.; Biggin, P. C.; Smith, G. R.; Sansom, M. S. P. *Q. Rev. Biophys.* **2001**, *34*, 473.
(43) Roux, B.; Karplus, M. *Annu. Rev. Biolphys. Biomol. Struct.* **1994**, *23*, 731.
(44) Woolf, T. B.; Roux, B. *Proteins* **1996**, *24*, 92.
(45) Forester, T. R.; Smith, W. *DL_POLY user manual*; CCLRC, Daresbury Laboratory, Daresbury, Warrington, U.K., 1995.
(46) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33.
(47) Ketchem, R. R.; Roux, B.; Cross, T. A. *Structure* **1997**, *5*, 1655.