JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Simple Method for Simulating the Mixture of Atomistic and Coarse-Grained Molecular Systems

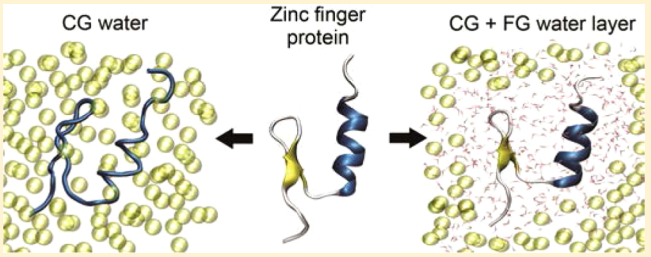
Pandian Sokkar,[†,§] Sun Mi Choi,[†,‡,§] and Young Min Rhee[†,‡,*]

[†]Center for Self-assembly and Complexity, Institute for Basic Science (IBS), Pohang 790-784, Korea

[‡]Department of Chemistry, Pohang University of Science and Technology (POSTECH), Pohang 790-784, Korea

Supporting Information

ABSTRACT: Combining fine-grained (FG) all-atom and coarse-grained (CG) systems in a single simulation in a hybrid manner is of immense interest in recent times, owing to the possibility of overcoming the limitations of both FG simulations as well as CG simulations. The existing methods for combining these two resolutions tend to require heavy parametrizations or sometimes lack in transferability to other systems of interest, and further developments toward such directions are highly required. We report here a simple protocol to combine CG and FG systems in a single simulation, using the standard FG and CG force field models by adopting a series of small proteins as test cases. Our method makes use of virtual sites as reported earlier for relatively simple butane and dialaine systems (Rzepiela et al. *Phys. Chem. Chem. Phys.* **2011**, *13*, 10437−10448), to bridge the interaction between FG protein atoms and CG water. We find that the conventional CG model (MARTINI potentials) couples too strongly with the FG model and that it leads to complete unfolding of a test protein within very short time. We find that reducing the Lennard-Jones potential between CG atoms and virtual site atoms stabilizes the secondary and tertiary structures, sometimes almost to a comparable level with the fully atomistic simulations. However, detailed inspection reveals that this reduction is not enough for satisfactory consistency of the hybrid scheme against the FG simulation. As a remedy, we observe that the addition of as small as 4 Å thick position-restrained FG water layer in the hybrid simulation can further improve the structural behaviors in many respects, with its results closely mimicking those of the FG-only simulations. However, free energy landscapes reveal that this agreement with a restrained solvent layer is still accompanied by the overstabilization of the protein native structure, which will likely pose limitations for studying protein dynamics with the scheme. We show various test results that we have tried in optimizing the FG-CG mixing scheme over the course and discuss future prospects as concluding remarks of the present work.



1. INTRODUCTION

Molecular dynamics (MD) simulation serves as an important tool for investigating the structural as well as functional properties of biological macromolecules that are otherwise difficult to obtain. The atomistic or fine-grained (FG) MD simulations have been applied to address a diverse set of problems and the results are often consistent with the experimental findings. However, the FG simulation can be cumbersome for larger molecular systems, as it requires long time computation despite the recent developments in computational resources. To reduce the computational cost associated with the FG simulations of larger systems, several minimalistic modeling approaches have been introduced. One such approach is the united-atoms method, in which all the aliphatic hydrogen atoms are merged with their parent carbon atoms.[1] The second and perhaps more important approach in terms of the efficiency is the use of coarse-grained (CG) resolution. In the extreme side, solvent molecules are often implicitly represented[2] as continuous medium without any explicit particles. Even with their excellent performances and overall reliabilities compared against explicit representation of solvent molecules,[3] the continuum models intrinsically cannot form empty spaces[4] and sometimes cause undesirable effects from erroneous estimation of the charge−charge interactions.[5,6] In such situations, the CG scheme can become more adequate, as it explicitly includes solvent molecules.

In CG simulations, typically three to four heavy atoms are lumped into a single atom-like particle or bead, thereby massively reducing the number of atoms as well as the degrees of freedom, which in turn smooths the potential energy surface.[7−11] Owing to the greater scalability and speed, the CG methodologies started to be widely adopted by many researchers in simulating very large systems (e.g., mesoscale polymeric systems and virus capsids).[12−16] Interestingly, the CG models are able to capture the molecular association processes with reasonable agreement with the experimental observation in spite of their low-resolution.[14,17−21] Currently, there are many CG force fields available. Among these, the MARTINI force field has gained considerable popularity due to

Received: February 4, 2013

dx.doi.org/10.1021/ct400091a | *J. Chem. Theory Comput.* XXXX, XXX, XXX−XXX

its simplicity, adaptability, and wide-coverage of molecular systems.[16,22−25]

However, one should be aware of the limitations of the CG method compared to the FG method. For instance, it is not always feasible to investigate the conformational transition of secondary structure in proteins or other delicate conformational changes by the CG methodology, due to the reduction in the torsional degrees of freedom.[26] Fortunately, the region of interest is usually a small segment in a larger system, such as ligand-binding site or interaction interface in protein-macromolecule complex. In such cases, a hybrid CG-FG simulation would be desirable, where both the FG system (describing the region of interest) and CG systems are simultaneously present, in a manner similar to quantum mechanics/molecular mechanics (QM/MM) simulation.[27,28] Nevertheless, the hybrid CG-FG simulation is challenging and is still at its infant state, mainly because of the complexities involved in the introduction of efficient coupling between two different resolutions. Several attempts have been made in this angle, and substantial amount of success has been achieved over the years. Shi et al. have employed force-matching technique to simulate FG ion channel protein in CG membrane and CG water.[29] This method requires extensive parametrization for the systems under study. Several other attempts to build hybrid CG-FG systems are limited to studying FG proteins in CG solvents, mainly to improve the scalability of atomistic simulations.[30−34] Recently, Rzepiela et al. have also described a scheme to combine CG MARTINI and FG GROMOS 53a6 force fields[35] by employing massless virtual interaction sites on relevant FG atom groups, which is conceptually quite similar to the center-of-mass particles employed in adaptive resolution hybrid schemes.[36−39] In hybrid simulations with the virtual sites, sometimes the incorrect electrostatic screening effect from CG water may induce abnormal behaviors. As a remedy, Wassenaar et al. attempted to mimic the electrostatic coupling using different relative dielectric constants with more advanced water models.[40]

In the method described by Rzepiela et al. with virtual sites,[35] specifically, there is no direct interaction between CG and FG atoms. Instead, the force acting on each virtual site is distributed to its constituent atoms weighted by their masses. Even though significant success has been reported, rather simple systems such as butane and dialanine in water were only adopted in demonstrating their approach. Indeed, a simple but universal scheme that utilizes the standard all-atom (AA) and CG force field parameters without extensive parametrizations and calibrations to model the hybrid CG-AA system is highly desired for routine applications. This has drawn our attention to focus on the suitability of this approach in investigating more complex systems.[35]

In this study, we have attempted to combine two different resolutions (CG and FG) in a single simulation based on virtual site-mediated coupling between the two. We have used a series of small proteins, a designed zinc finger protein (without Zn^{2+}), the tryptophan zipper2 (trpzip2), the subdomain of villin headpiece (HP36), and the tryptophan cage (trp-cage) as test cases. The selections were mainly based on their small sizes (no more than 36 residues) and the variety of their secondary structures. For example, the designed zinc finger protein has all the important secondary structure elements, namely an α-helix, a β-strand, a turn, and coils. More importantly, it is only marginally stable, which may provide very sensitive structural effects caused by the mixing of CG-FG systems.[41,42] The other three are relatively well-known fast folding proteins.[43] Structurally, trpzip2 represents a well-defined β-strand-rich family, HP36 is mainly α-helical, and trp-cage has both α and 3_{10} helices. Therefore, performing MD simulations with these additional proteins will reveal the degree of adequacy of applying any computational protocol for diverse protein systems.

2. METHODS

Interaction Scheme. The scheme for describing the interaction between CG and FG systems is adapted from Rzepiela et al.[35] In this approach, the interactions present in CG system are calculated by MARTINI force field and the interactions in FG system are by GROMOS 53a6 force field[1,22] without any additional direct interactions between CG and FG atoms. However, the CG atoms interact with virtual sites on FG atoms and the forces acting on virtual site is redistributed to its constituent heavy FG atoms, weighted by their masses. Virtual sites are chosen on the FG atoms in a manner similar to the MARTINI's mapping methodology (i.e., one virtual site is placed at the center of mass of 3−4 heavy atoms). In this way, the virtual sites act as virtual MARTINI CG beads and the interaction between a CG atom and a virtual site is calculated by the standard MARTINI LJ-potential.[22] On the contrary, the electrostatic interactions between CG and FG atoms are directly calculated and not mediated through the virtual sites. Mapping of virtual sites was performed in two steps. First, the atomistic protein model was converted into CG model using Martinize-1.2 script (available at http://md.chem.rug.nl/cgmartini/). Next, the coordinates of the CG atoms are appended to the atomistic protein coordinates as virtual site atoms and the LJ-potential terms from each CG atom are assigned to the corresponding virtual site atom. In other words, the virtual sites act as CG beads that transfer the force to its constituent atoms. It should also be noted that the virtual sites do not interact with each other and the FG-FG interactions are calculated by the GROMOS force field.

FG Protein in CG Water. Throughout this work, GROMACS package version 4.5 was used for all the simulations.[44,45] GROMOS 53a6 force field was used for the atomistic part of the simulation and MARTINI force field was used for CG part. The NMR structures of the adopted proteins were obtained from protein databank (PDB ID: 1FME, 1LE1, 1VII, 1L2Y).[46−49] The zinc finger protein is not chelated with a zinc ion, unlike its native structure. A representative structure was taken from the ensemble and was used for all the simulations. The atomistic representations of the proteins were used together with their virtual site atoms in all cases, except in the case of the fully atomistic simulations. Each protein was solvated by nonpolarizable CG water particle with the thickness of at least 30 Å. The simulation system was first energy minimized for 10000 steps with positional restraints on protein atoms, followed by unrestrained energy minimization for additional 2000 steps. The energy-minimized structures were subjected to 100 ps NPT simulation at 300 K with positional restraints on protein to equilibrate the water molecules.

The equilibrated structures were used for subsequent 10 ns simulations for all four proteins using the following settings. A time step of 2 fs was used to integrate Newton's equations of motion, and the trajectories were recorded at every 2 ps interval. All the bonds were constrained to their equilibrium bond length using LINCS algorithm with 6 matrices toward the expansion for the matrix inversion.[50] Periodic boundary

conditions were employed to eliminate the surface effects. The temperature was kept constant at 300 K with Parrinello's velocity rescaling scheme (time constant 1 ps),[51] and the Berendsen barostat was used to keep the pressure at 1 bar (time constant 2 ps).[52] The shift cutoff method was adopted to smoothly decay the van der Waals forces to zero between 9 to 14 Å. Particle-mesh Ewald (PME) summation was employed for long-range electrostatic interactions,[53,54] with the cutoff distance for short-range electrostatic interactions set to 16 Å. For electrostatic interactions, the relative dielectric constant (ε_r) of 1 was used as solvent molecules are explicitly represented in all cases.[1]

FG Protein in FG and CG Water. In a different simulation scheme, the FG protein was solvated by SPC atomistic water molecules[55] with a thickness ranging from 3 to 7 Å. This was further solvated by CG water with the combined thickness of FG and CG water of 30 Å. In another simulation scheme, the distance between the FG water and protein was restrained to avoid the diffusion of FG water into the bulk CG water. A flexible boundary that resembles the overall shape of the protein was used for FG water molecules (see the Results and Discussion). To ensure the uniform distribution of FG water molecules around the protein surface, the virtual site atoms, which are defined to mediate CG-FG interactions, were used to create distance restraints. This is based on reasoning that the virtual atoms better represent the approximate shape of the protein than just the center of mass of the entire protein. The distance between a virtual atom and a FG water molecule present within (d + 1) Å (where d is the thickness of the FG water layer) of the virtual atom is restrained. Energy penalty was applied to all the distance-restrained atoms that moved farther than ±1 Å from the original distance, with a harmonic force constant of 25 kJ mol^{-1} nm^{-2}. The interaction scheme and simulation settings for these systems are the same as detailed in previous parts for systems only with CG water, except the settings for the additional interactions between CG and FG water.

Parametrization of CG-FG Water Interaction. We adopted a simple method to parametrize the interaction between CG water and FG water. An additive combination of the σ values of CG water and FG water oxygen was used for van der Waals radius factor. For the van der Waals interaction strength $\varepsilon_{FG\text{-}CG}$, a multiplicative combination was first considered. However, obtaining the appropriate parameter for this interaction strength is rather complicated due to the very different nature of uncharged CG water model and charged FG water model. Thus, we have optimized it to improve the miscibility of these two water models. For this purpose, systems with different CG-FG water compositions (1:1, 4:1, and 8:1 in volume) were simulated using different $\varepsilon_{FG\text{-}CG}$ values. The finally obtained parameters were $\sigma_{FG\text{-}CG}$ = 3.933 Å and $\varepsilon_{FG\text{-}CG}$ = 4.1 kJ/mol, and these were used for all the hybrid simulations with mediating atomistic water layer. More detailed accounts on the CG-FG water−water interaction can be found in the Results and Discussion.

Fully Atomistic Simulation. The proteins were solvated by an at least 15 Å thick SPC water model. Counterions were added to neutralize the charges of the systems. The PME method was applied for long-range electrostatic interactions, using a short-range cutoff of 12 Å. For van der Waals forces, the shift cutoff method was adopted between 8 to 10 Å. LINCS algorithm was used to constrain all the bond lengths to their equilibrium value with the order of 4. Isothermal and isobaric conditions were maintained by velocity rescaling thermostat and Berendsen barostat similarly to the hybrid simulation cases. The time step of 2 fs was used and the system was subjected to NPT simulation at 300 K temperature and 1 bar pressure. For the purpose of testing various schemes with the four proteins, we performed MD simulations for 10 ns. Because this duration may not be long enough for reliably checking the agreement of the hybrid scheme against the fully atomistic simulation, at the last stage after reaching the most appropriate hybridizing scheme, we additionally performed both the hybrid and the all-atom simulations for 100 ns durations.

The adopted simulation protocols with various solvent settings that have been described in the above are summarized in Table 1.

Table 1. Summary of Simulation Protocols with Different Solvent Representations

		hybrid simulation		
methods	atomistic simulation	CG water	CG + FG[a] water	CG + restrained FG[a] water
protein representation	FG	FG	FG	FG
thermostat	v-rescale[b]	v-rescale[b]	v-rescale[b]	v-rescale[b]
barostat	Berendsen	Berendsen	Berendsen	Berendsen
electrostatic interactions	PME	PME	PME	PME
electrostatic cutoff[c]	12 Å	16 Å	16 Å	16 Å
van der Waals tapering	8−10 Å	9−14 Å	9−14 Å	9−14 Å

[a]FG water layer thicknesses adopted: 3, 4, 5, 6, and 7 Å. [b]Parinello's velocity rescaling method (ref 51). [c]PME real space cutoff. The neighbor search cutoffs were set to the same values.

Analysis of MD Simulation. The secondary structures were analyzed using STRIDE algorithm, as implemented in VMD.[56,57] In comparing the secondary structures between the FG and hybrid simulations, we have adopted the secondary structure propensity as obtained with the appearance probability of the given structure in a residue-by-residue manner. The number of hydrogen bonds (3.6 Å distance and 30° angle cutoffs) and the solvent distribution around the protein were additionally calculated for analysis. Formations of salt-bridges (3.2 Å cutoff distance) were also regularly checked to closely match with the secondary structure formations. The solvent accessible surface area was checked to supplement these analyses.

3. RESULTS AND DISCUSSION

Structural Behavior with Hybrid Simulations. Let us first validate our hybrid CG-FG model with the zinc finger protein (Figure 1a) by comparing its simulation results against the fully atomistic counterpart. For this, we first have to consider how to select the optimal simulation settings for the mixed CG-FG system. In fact, this is rather complicated because of the differences in settings recommended for MARTINI and GROMOS force fields. Recently, Rzepiela et al. pioneered successful applications of a hybrid model based on the test case of a butane system and a dialanine/water system.[35] However, when their simulation conditions were applied to our system (atomistic protein in CG water), it was observed that the protein unfolded completely within a few picoseconds

dx.doi.org/10.1021/ct400091a | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

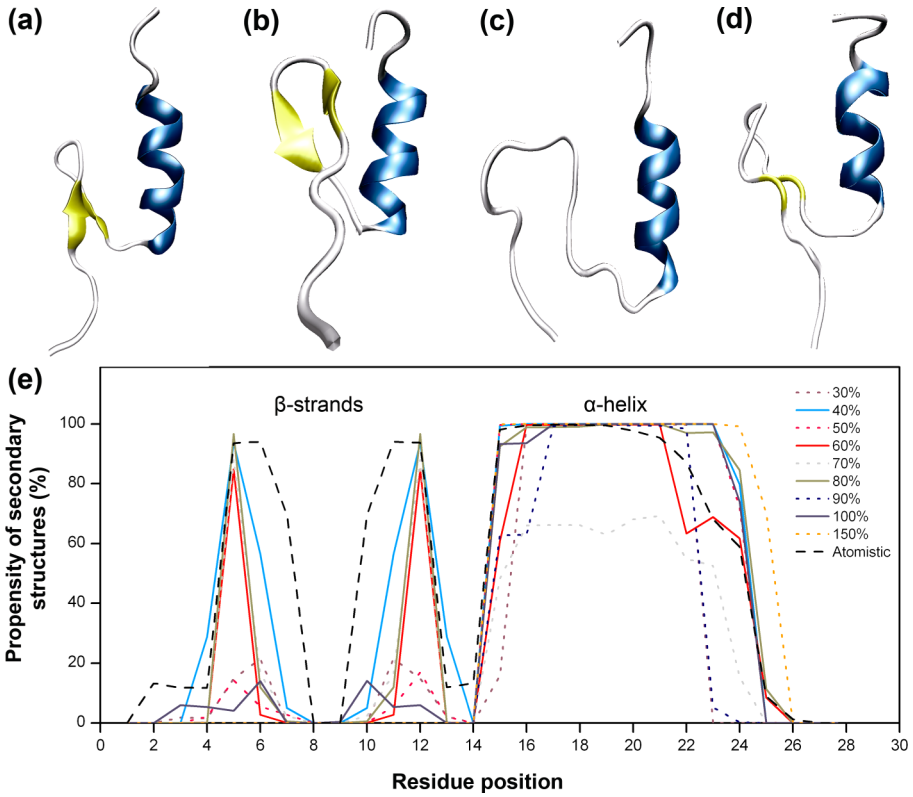


Figure 1. Analysis of the structural stability of the atomistic zinc finger protein in hybrid CG-FG simulation. The overall structures of the protein: (a) NMR structure (PDB ID: 1FME), (b) after 10 ns MD simulation in atomistic water, (c) after 10 ns MD simulation in CG water with 100% LJ potential, and (d) after 10 ns MD simulation with 60% LJ potential between CG and FG atoms. (e) Effects of scaling LJ interaction between CG and FG atoms on the populations of the β-strands and the α-helix. Fully atomistic simulation results are also shown for reference.

(Figure S1, Supporting Information). This suggests that we need to further improve this approach to apply it to larger systems. We speculate that this rapid unfolding happened due to a number of reasons. The relative dielectric constant ε_r value of 15 (used by Rzepiela et al.) is excessively high for the protein atoms with FG representation (ε_r = 1 with explicit water), which will severely weaken the electrostatic interactions in the FG system. In MARTINI force field, in fact, the CG water is uncharged and cannot screen electrostatic interactions between charged solute particles. This absence of electrostatic screening is compensated by the higher relative dielectric constant.[22] Even though this is more physical for a pure CG system, it may not be appropriate for particles with GROMOS 53a6 force field parameters. For example, the protein backbone−backbone hydrogen bonds are treated mostly by electrostatic interactions. Because the electrostatic screening effect of water would be minimal on the backbone hydrogen bonds due to the relatively low solvent accessibility, adopting a high value of ε_r will have a negative impact on the stabilization of secondary structures. Compared to the fully atomistic simulation (Figure 1b), the hybrid simulation yielded the loss of β-strands and noticeable distortions in helical segments (Figure 1c).

These structural distortions can also be readily explained by the inevitable limitations of CG water. Because CG water is composed of uncharged and relatively large particles, which cannot diffuse into smaller cavities of the protein, effective solvation is naturally unexpected. In addition, it is also possible that the forces acting on the FG atoms due to the interaction with CG system is slightly higher than what they would be with more physical FG counterparts. Such an artifact may also lead to unfolding. The relatively higher stability of helical segment can be due to the presence of oppositely charged residues in close proximity, namely in adjacent turns, in the helical conformation.

To verify whether the forces acting on FG atoms from CG water are indeed too high, we varied the LJ potential between CG water and virtual site atoms, ranging from 30% to 150% and compared the results with fully atomistic simulation (Figure 1). The results showed that as the interaction between CG and FG system is scaled down from the 100% CG-FG interaction potential, the stability of the secondary structural elements increases in many cases, as evident from the higher propensity of β-strands and α-helix observed during the simulation time.

Similar behaviors were observed with the other three proteins (Figures S2 and S3, Supporting Information). This suggests that the forces acting on the virtual site by CG atoms without any scaling are definitely higher in general than what it should be for correct behaviors of the protein. Therefore, LJ potential should be used with cautions when MD simulations with virtual sites for mixed atomistic and CG systems are performed. However, the optimal LJ potential scaling factors were different for each protein. In the case of the zinc finger protein, secondary structures at 40%, 60%, and 80% LJ potential were in good agreement with atomistic simulations (Figures 1e). In the case of trpzip2, 30% LJ potential was the best value when compared with all-atom simulations (Figure S2, Supporting Information). Reductions to 50% of regular LJ potential for HP36 and to 40% for trp-cage were in decent agreements with the results of all-atom simulations (Figure S3,

dx.doi.org/10.1021/ct400091a | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

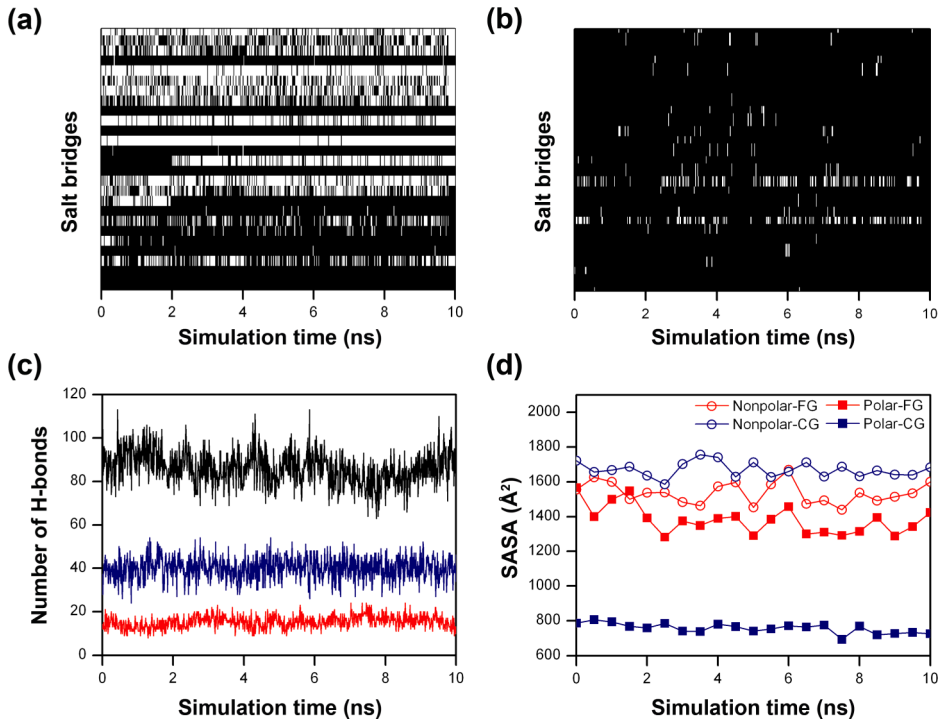


Figure 2. Comparisons between the time developments of various structural aspects from hybrid and atomistic simulations for the zinc finger protein. Evolution of salt-bridges (white streaks) in (a) hybrid and (b) atomistic simulations. (c) Occurrences of intraprotein hydrogen bonds during hybrid (blue) and in atomistic (red) simulations, together with the protein−water hydrogen bonds in atomistic simulation (black). (d) Solvent accessible surface area (SASA) around polar and nonpolar residues measured by a probe with 1.4 Å radius.

Supporting Information). Moreover, the behaviors with the scaling factor changes were rather erratic and did not exhibit any monotonic improvement or worsening.

Dynamical Behaviors of Structural Properties. As described in the above, the 30−80% reduction of LJ potential between CG and FG atoms resulted in better agreements between the hybrid and atomistic simulations in terms of stabilization of secondary structures, but the improvements exhibited somewhat erratic behavior. Thus, we can infer that scaling down LJ potential will be desirable but not fully satisfactory only by itself. In addition, having similar secondary structures does not yet mean that the protein dynamics in the two simulations are comparable. Hence, we have analyzed the differences between hybrid simulation and fully atomistic simulation in a more detailed manner. We will mention only the zinc finger protein in this section, because it will be sufficient for illustrating the problem of simple hybrid simulation. For simplicity, we will adopt only 60% level of LJ interaction, where the secondary structure behaviors were at least decent.

We first compared the occurrences of salt-bridge interactions in these two systems. As shown in Figure 2a and b, in contrast to atomistic simulation, a large number of salt-bridges were present in the hybrid simulation. This deviation is not an artifact of the choice of the salt-bridge cutoff distance, as adopting different cutoffs yield similarly large deviations (Figure S4, Supporting Information). A close inspection revealed that almost all the charged amino acids participated in the salt-bridge formations and that the salt bridges tended to persist for relatively long time. In fact, this aspect has induced the noisy appearance of the plot in Figure 2a. In contrast, in the case of atomistic simulation, only two significantly long-lived salt-bridges (both of them formed between ASP20 and LYS24) were observed during the simulation time (Figure 2b). In addition, the hybrid simulation generated a higher number of intraprotein hydrogen bonds than the atomistic simulation, as shown in Figure 2c. This figure also suggests that, in the case of fully atomistic simulation, the presence of many protein−solvent hydrogen bonds diminishes the intraprotein hydrogen bond formation. These results indicate that the intraprotein electrostatic contacts are exaggerated in the case of hybrid simulation. In fact, this electrostatic imbalance between FG-FG and FG-CG interactions is a well-known issue in FG/CG hybrid schemes.[33,34] This argument is further supported by the analysis of the solvent accessible surface area (SASA) shown in Figure 2d. In atomistic simulation, the SASAs of both nonpolar and polar residues were found to be similar (∼1500 Å^2), indicating that the protein possesses an amphipathic surface. The hydrophobic SASA from hybrid simulation was similar to the atomistic case, suggesting that the distribution of nonpolar residues at the protein surface is not affected by the model switch to a significant extent. However, the hydrophilic SASA from hybrid simulation is significantly lower (by ∼600 Å^2). This can be ascribed to the higher incidents of intraprotein polar interactions. Namely, because the polar residues tend to interact with each other more strongly in the hybrid simulation, compaction of the residues will inevitably occur.

All these results indicate that the hybrid simulation with the atomistic protein solvated by uncharged CG water unphysically promotes the intraprotein electrostatic interactions, even when the secondary structures and tertiary contacts are relatively well preserved during the simulation. Similar behaviors were found with the other three proteins. This limitation of our hybrid simulation is likely the result of the simplified CG water solvent that does not have any charge. Unless the dipolar character as in FG water is somehow embedded in CG water, it will be

dx.doi.org/10.1021/ct400091a | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

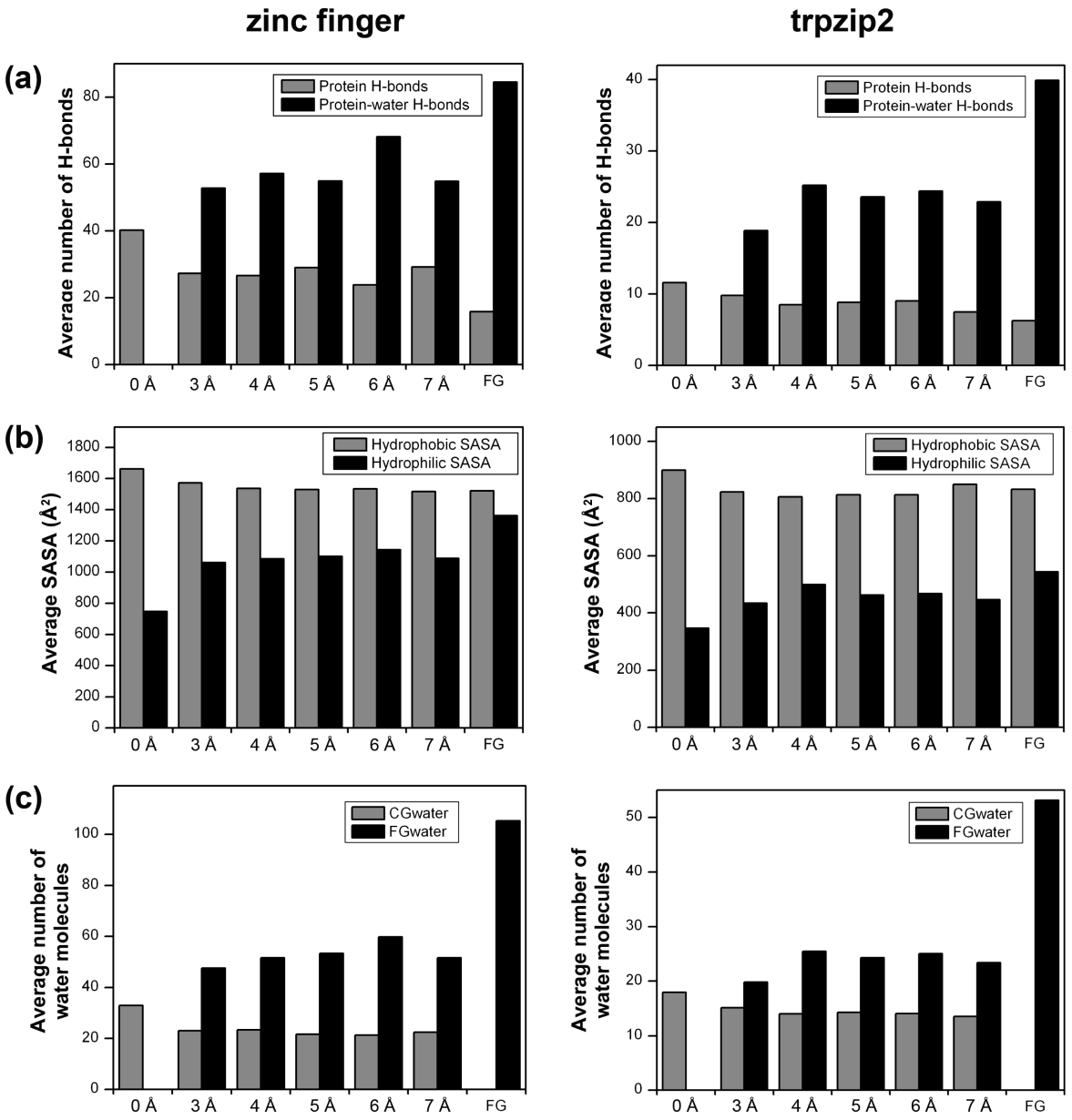


Figure 3. Structure quality analysis for the hybrid CG-FG simulation in the presence of FG water layer for zinc finger (left panels) and trpzip2 (right panels). (a) The number of intraprotein and protein−water hydrogen bonds, (b) the solvent accessible surface areas, and (c) the number of CG (within 4 Å from the protein) and FG water molecules (within 3 Å). These values are obtained from averages over the last 5 ns from hybrid simulations with FG water layers (noted with corresponding thicknesses below the horizontal axes) and from atomistic simulation (noted as FG).

difficult to eliminate such a limitation. We also repeated our hybrid simulations in the presence of CG ions or atomistic ions to check the potential improvements in the description of the electrostatic interactions, but the results were almost the same without any meaningful changes (data not shown).

Hybrid Simulation with Atomistic Water Layer. Very recently, Riniker et al. have reported a scheme to simulate atomistic protein solvated by a layer of atomistic water followed by an additional layer of supramolecular polarizable CG water.[33] The objective of their work was of course to increase the scalability of atomistic simulation by incorporating the polarizable CG water model without causing discrepancy with respect to corresponding fully atomistic simulation. With some tailored parametrization for building the CG water model, they have reported that this hybrid model with a buffer zone of 8 Å atomistic water layer can reliably reproduce the results of fully atomistic simulation. This work has motivated us to look into the possibility of using atomistic water layer to improve the structural features in our hybrid approach, with the purpose of establishing a simple and transferable approach for hybrid CG-FG simulations. Our main focuses are on the possibility of having a simple approach of adopting widely applied existing force field models without any necessity of new parametrizations that may have dependences on the systems of interest. It will also be beneficial if we can adopt thinner layer of the buffer zone.

With these focuses, we attempted to study the effect of adding a thin layer of atomistic water on the suitability of the hybrid CG-FG simulation. This, however, requires the interaction between CG water and FG water models. Unlike the protein with a fixed sequence, defining virtual atoms in freely diffusing FG water is not trivial. Therefore, we have taken

dx.doi.org/10.1021/ct400091a | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

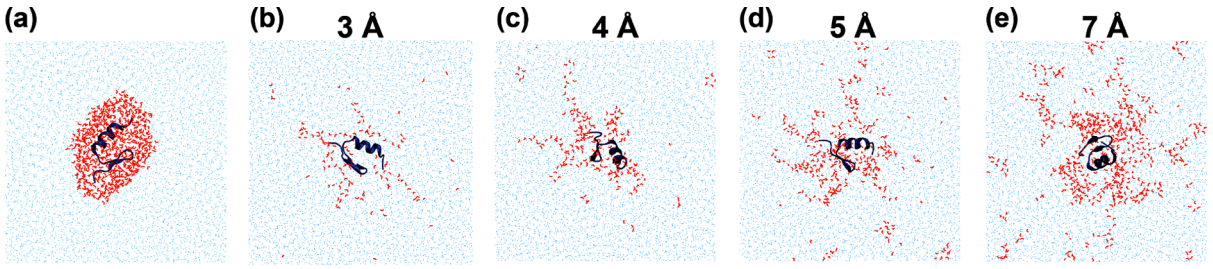


Figure 4. Solvent configuration around zinc finger protein (a) at the initial state and (b–e) after 10 ns of simulation. The thickness of FG water layer used for each simulation is given in the figure. The protein is shown with a cartoon representation (blue). The CG water atoms are displayed as cyan dots, while the FG water molecules are shown as red sticks.

a simple approach of directly adopting the LJ-interaction parameters of CG and FG water models. The σ value, which dictates the interaction distance between the two particles, was obtained by additively combining the self-interaction terms of CG and FG water. The interaction strength, ε, between CG and FG water was initially obtained through a multiplicative combination (ε = 2.6 kJ/mol). However, this was not large enough to induce diffusive mixing of the two different types of water into each other. Namely, when 1:1 FG-CG volume mixture of water was simulated, phase separation of the two types occurred (Figure S5, Supporting Information). Hence, we have performed a series of optimization in the ε value for the CG-FG water interaction, mainly to improve the miscibility of these two water models. When the ε value was increased to 4.1 kJ/mol, the two water models became freely miscible at three different CG-FG water compositions (Figure S6 and Table S1, Supporting Information), and this ε value was used for all hybrid simulations with atomistic water layer.

The reason for the poor miscibility of these two water models stems from the simple uncharged CG water. Recently, Riniker and van Gunsteren simulated the mixture of FG and polarizable CG water.[58] The use of polarizable CG-water model is slightly superior in terms of better electrostatic screening. It is important to note that the simulation of FG protein in polarizable CG water model is also known to introduce structural inaccuracies in protein,[34] similar to the results of FG protein in uncharged CG water simulation discussed in this study. This indicates that the larger size and not the electrostatics of the CG water models may be an important concern in FG-CG hybrid simulation, although further studies are required to confirm this. We also note that the simple CG water model has been widely employed in CG MD simulations and the results of which are in sufficient agreement with experiments or atomistic simulations. Further, because the majority of the computational cost in MD simulations is accounted for by the large number of solvent atoms, the use of simple uncharged CG water will contribute best for the scalability of the hybrid scheme.

In order to find out the minimal thickness of the FG water layer required to improve the protein structure quality in hybrid CG-FG simulation, the four test atomistic proteins were solvated with FG water with various thicknesses (3–7 Å). These were then further immersed in bulk CG water, and MD simulations were performed for 10 ns. The interaction scheme and the simulation setup were the same as mentioned before. As the secondary structural behaviors were decent to good with 30–80% scaling of CG-CG interaction toward the CG-virtual site interactions, as discussed previously, 60% down scaling was employed throughout this test. With the buffering water layer, the overall structural stabilities of zinc finger and trpzip2 improved with all thicknesses, comparable to that of fully atomistic simulation (Figure S7, Supporting Information).

However, average numbers of intraprotein hydrogen bonds were found to lie in-between the fully atomistic simulation and hybrid FG protein-CG water simulation (Figure 3a). Analysis of the average SASA indicated that the distribution of hydrophilic and hydrophobic residues at the protein surface lies intermediate to that of fully atomistic simulation and hybrid FG protein–CG water simulation (Figure 3b). Overall, the thickness of atomistic water layers had little influence on the protein structural qualities (SASA, intraprotein, and protein–water hydrogen bonds), although the simulations performed with 3 Å water layer exhibited poor structural quality. These intermediate results indicate that the presence of FG water layer did not fully improve the structure qualities yet.

To inspect the reason behind the limited improvement, we analyzed the distributions of CG (within 4 Å distance) as well as FG water (within 3 Å) around protein by calculating the average number of water molecules present in the last 5 ns of the simulations.[59] The results are shown in Figure 3c, and one can see that the numbers of CG and FG water molecules in the solvation shell around the protein do not change even when the FG water layer thicknesses are different. This is consistent with the observed trend of protein–water hydrogen bonds upon varying FG water layer size (Figure 3a). The results with the other proteins, HP36 and trp-cage, are also consistent with these findings and diplay that the application of the buffering FG water layer does not fully improve the structural behaviors (Figure S8, Supporting Information).

This behavior is ascribed to the distortions in the shapes of the FG water layer. In fact, as shown in Figure 4, diffusive redistribution of FG water molecules into the bulk CG water has occurred during the simulation time, leading to their depletion from the protein surface. Likely owing to the attraction between CG and FG water molecules together with the entropic factor, the FG water is dragged into the bulk CG water. In addition, the CG water can compete with FG water molecules to interact with protein atoms more closely. The redistribution phenomenon causes more CG water molecules to be present at the solvation shell of the protein, which in turn can affect the structural stability and intraprotein interactions in a similar manner as in the case with only CG water.

Restraining FG Water Layer. As mentioned with Figure 4, the expected improvement in the protein structural behaviors in the presence of FG water layer was hampered by the FG water "flying away" from the protein surface over the course of simulation. Indeed, when the FG-FG, CG-FG, and CG-CG water interactions are well balanced, complete mixing of the

dx.doi.org/10.1021/ct400091a | J. Chem. Theory Comput. XXXX, XXX, XXX–XXX

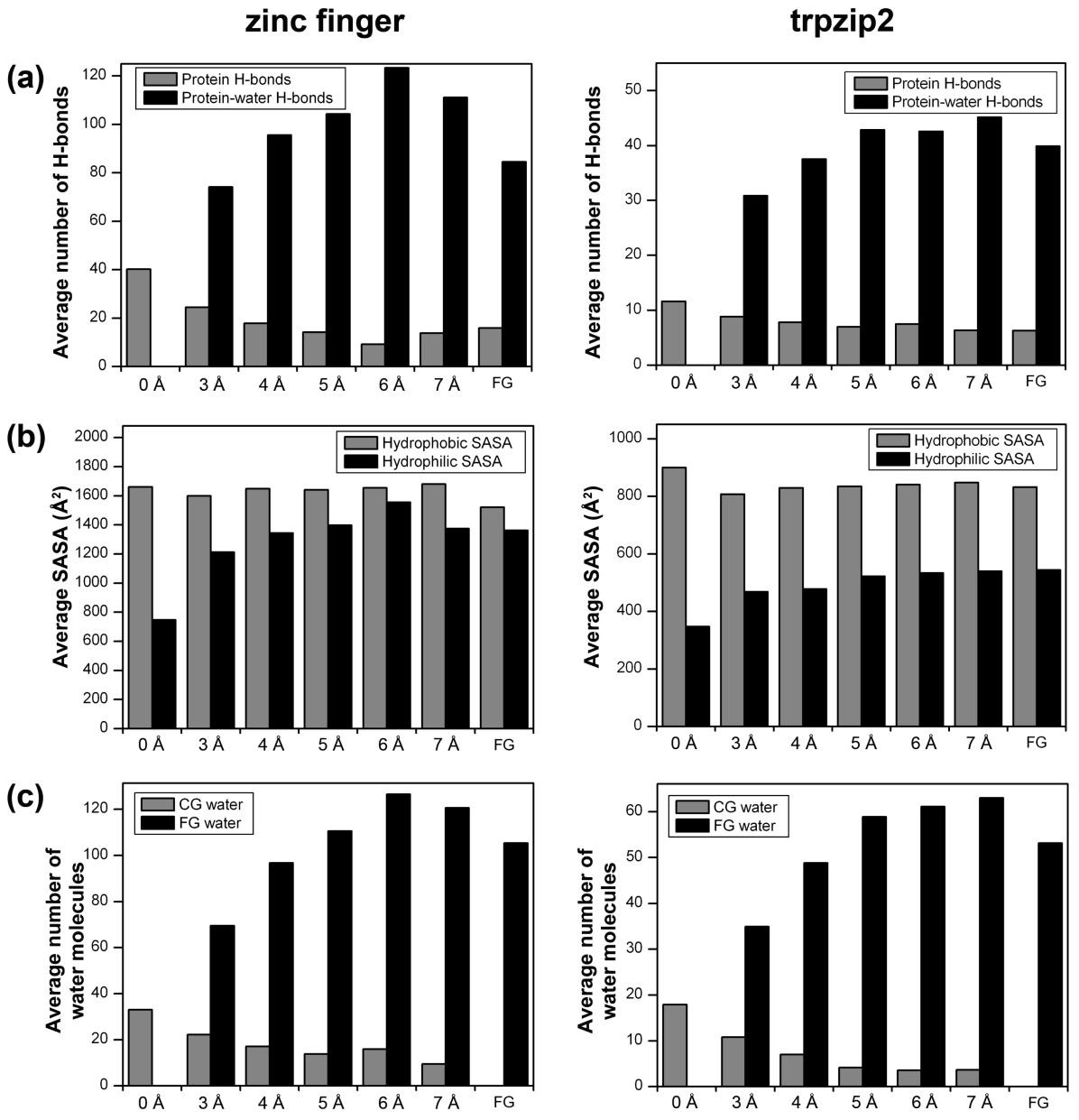


Figure 5. Structure quality analysis for the hybrid CG-FG simulation in the presence of distance restrained FG water layer with zinc finger (left panels) and trpzip2 (right panels). (a) The number of intraprotein and protein-water hydrogen bonds, (b) the solvent accessible surface areas, and (c) the number of CG and FG water molecules in the solvation shell of protein. These are obtained with averages over the last 5 ns from hybrid simulations with FG water layers (noted with corresponding thicknesses below the horizontal axes) and from atomistic simulation (noted as FG).

two different types of water should happen through self-diffusion. This should, of course, act as a major concern for applying the hybrid scheme simulations with FG water layer. Because the depletion of FG water molecules will cause an influx of CG water molecules in the solvation shell around the protein, some structural distortion in the protein will eventually happen as was already discussed in previous parts. In fact, we have seen that the hybrid solvation scheme involving explicit and implicit water models is not new to the modeling community,[60−63] and a similar issue does exist with that scheme. Several complex solutions have been introduced over the years to prevent this undesirable behavior of water molecules during the simulation. A relatively simple strategy against this problem will be to use attractive distance restraints between the FG water layer and the protein. Riniker et al. actually adopted this scheme by employing the center of mass (COM) of the protein for the restraining potential with a spherical boundary.[33] Recently, Lin et al. compared thermodynamic and dynamics properties between atomistic simulation and hybrid simulation with FG water layer and showed that these properties are very similar.[64] However, such a boundary for nonspherical protein structures will include unnecessarily large number of water molecules. Therefore, a flexible boundary for the FG water layer, which resembles the overall shape of the protein and also responds to its structural change, is highly desirable. The virtual atoms introduced for mediating FG and CG atoms[35] can indeed serve this purpose very well. Thus, in the present work, we achieved this flexible boundary of the FG water layer by introducing distance restraints between FG water and the virtual atoms that represent the approximate shape of protein surface.

dx.doi.org/10.1021/ct400091a | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

The results of 10 ns simulation of zinc finger and trpzip2 with distance restraints are shown in Figure 5. The average number of intraprotein hydrogen bonds was greatly reduced and the protein−water hydrogen bonds were found to be improved with the presence of distance-restrained FG water layers (4−7 Å), close to that of fully atomistic simulation (Figure 5a). The analysis of average SASA shows that the distribution of hydrophobic residues at the solvent accessible surface of the protein is not altered by the FG water layer of different thickness. On the other hand, the polar SASA was found to increase with increase in FG water layer thickness, as shown in Figure 5b. The average number of FG water molecules at the solvation shell of protein was found to increase significantly (Figure 5c) compared to the case with unrestrained-FG water simulation (Figure 3c), suggesting that the free diffusion of FG water molecules into the bulk CG water is eliminated. The simulation systems with 4−7 Å of FG water layers exhibit the protein hydration state that is well matching to that of fully atomistic simulation. Good agreements in the populations of the secondary structure elements were also observed (Figure S9, Supporting Information). The result also indicates that the 3 Å water layer is not sufficient to improve the hydration state of the protein and the electrostatic screening, meaning that its use will lead to increased intraprotein electrostatic interactions. HP36 and trp-cage with distance restrained water layers also exhibited improvements (Figure S10, Supporting Information), with some level of overestimation of protein−FG water contacts. Overall, the presence of 4−7 Å distance restrained water layer improves the structural quality of the protein, comparable to fully atomistic simulation. The occurrences of salt-bridges in the zinc finger protein (for the system with 4 Å FG water layer) show that only one pair of strong salt-bridge forming residues (ASP20 and LYS24) was observed during the simulation time and other nonspecific salt-bridges were largely suppressed as in Figure 6.

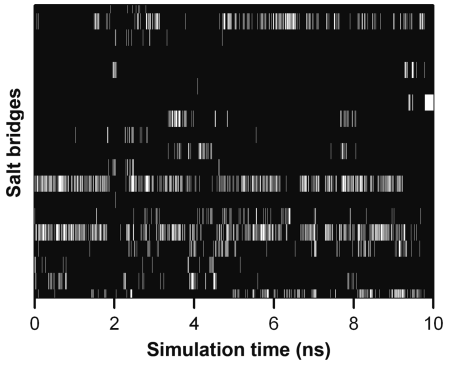


Figure 6. Occurrences of salt bridges of zinc finger protein during the simulation time for the hybrid CG-FG system with 4 Å FG water layer.

This is also in good accordance with the atomistic FG simulation shown in Figure 2b. Hence, the hybrid CG-FG simulation with a water layer even as small as 4 Å can reproduce much of the structural features of fully atomistic simulation.

Long Time Simulation. In previous parts, we have tried to reach a simple method for simulating the mixture of FG and CG systems by comparing rather short (10 ns) simulation results. To further demonstrate that this scheme is reliable for long time simulations, we have extended the trajectories to 100 ns durations. Here, for practical computational efficiency, we have adopted two proteins, zinc finger and trpzip2, and compared the results from 10 and 100 ns simulations. Figure 7 compares the structures of these two proteins obtained with the

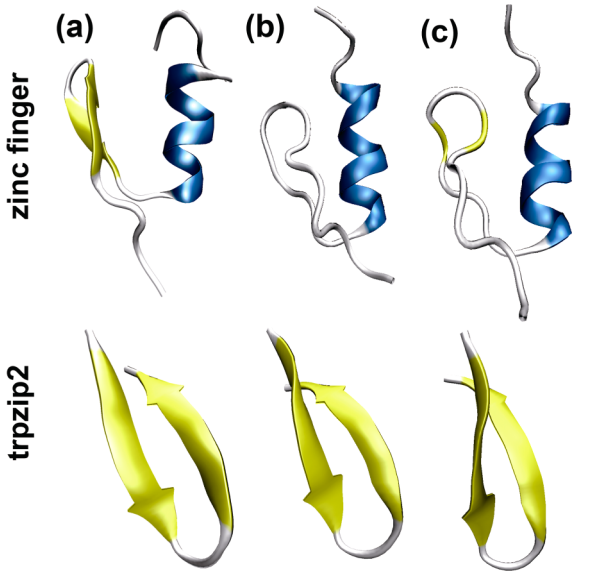


Figure 7. Overall structures of zinc finger (top panels) and trpzip2 (bottom panels) proteins after (a) 100 ns atomistic simulations together with the structures after (b) 10 ns and (c) 100 ns hybrid simulations with 4 Å buffer layer of position restrained atomistic water.

hybrid simulations at 10 and 100 ns time marks against the atomistic simulation results. One can clearly see that the structures are in good accord with each other. The hairpin structure within the zinc finger has some fluctuating characteristics, and depending on the criteria for registering a β-hairpin and depending on the individual trajectories, the ribbon representation, as in this figure, showed up somewhat differently. However, the overall structure was very similar in all inspected snapshots.

In fact, the structural similarity can be statistically assessed through the use of free energy landscapes. Figure 8 displays the free energy contours from the hybrid simulations as functions of α-carbon root-mean-squared deviations ($RMSD_\alpha$) from the experimental geometries and radii of gyration (R_g) for the two proteins, obtained by analyzing snapshots recorded during the last 10 ns of the 100 ns trajectories. Of course, care must be taken because 100 ns may not be long enough to reach convergence even for fast folding proteins. However, at least qualitative assessment can be deduced, as we have applied matching simulation conditions in both atomistic and hybrid simulations. In the figure, we can clearly see that $RMSD_\alpha$ from hybrid simulations are indeed quite close to the experimental values for both proteins. For trpzip2, the overall appearances of the free energy maps are close to each other with a slight distortion, suggesting that the hybrid simulation will likely have similar equilibrium behavior compared to the atomistic case. However, for the marginally stable zinc finger, the structural fluctuation represented by $RMSD_\alpha$ is somewhat too small compared to the FG simulation result. In fact, similar levels of agreement and disagreement have been observed by Riniker et al. with similar position restraints for FG water molecules.[33] Namely, a series of proteins with small $RMSD_\alpha$ deviations in atomistic simulations displayed excellent agreements with the hybrid scheme, while one protein with rather large $RMSD_\alpha$ (∼5 Å) in all-atom simulation exhibited noticeable inconsistency of

dx.doi.org/10.1021/ct400091a | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

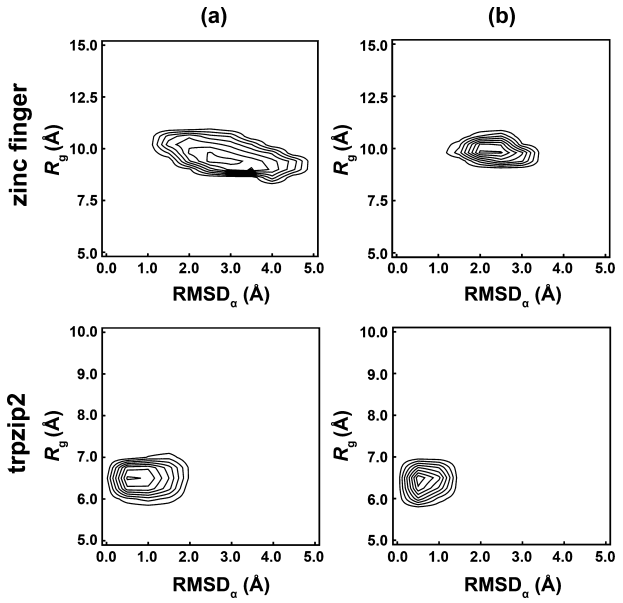


Figure 8. Free energy landscapes of zinc finger (top panels) and trpzip2 (bottom panels) proteins from (a) atomistic simulations and (b) hybrid simulations with 4 Å position restrained FG water layers. Contour lines are drawn at k_BT intervals, starting at $F = -k_BT$ at the outermost lines.

only ~2 Å deviation with the hybrid approach. Thus, we can infer that adding restraining potential for the buffering FG water layer is overstabilizing the protein native structure and/or slowing down the structural drift to a noticeable extent. Even trpzip2 is also showing this trend, and one can see that the hybrid simulation is slightly underestimating the RMSD$_\alpha$ fluctuation. This aspect will be a limitation for the present approach in studying relatively unstable and structurally fluctuating proteins.

4. SUMMARY AND CONCLUSION

In conclusion, we have presented a simple and straightforward strategy to combine FG and CG systems in a single simulation, using virtual sites as the mediator for FG and CG interactions and using standard GROMOS 53a6 (for FG systems) and MARTINI (for CG systems) force fields. We tested the approach with a series of miniproteins (zinc finger, trpzip2, HP36, and trp-cage protein) to validate its suitability. A simple approach of directly mixing the two models, which had been successful with rather simple systems such as solvated butane or dialanine,[35] was not appropriate for a more complex system adopted in our case (FG protein in CG water). Detailed analysis has suggested that the forces acting on FG atoms were generally too strong, which then led to extensive unfolding of the protein within a few picoseconds. When the LJ potential between CG and FG systems was weakened, the stability of the test protein increased. Reducing the LJ potential between CG water and FG protein to 30−80% level increased the stability of the protein secondary and tertiary structures, to a noticeable degree compared with the fully atomistic simulation results.

In the first attempt, however, the protein was simulated only with large and uncharged CG model water molecules. This promoted unnaturally exaggerated incidents of intraprotein electrostatic interactions, in a manner similar to the gas-phase situation. To further improve the reliability of the hybrid simulation, we included a small layer of FG water around the protein to more effectively solvate the protein.[33] We have tested various thicknesses of FG water layers ranging from 3 Å to 7 Å and investigated the structural stability and correspondences to FG-only simulations. The metrics that we have adopted include the secondary structure propensities, the CG and FG solvent distribution around the protein, the number of hydrogen bonds (intraprotein and protein-to-water), and the solvent accessible surface areas of polar and nonpolar residues. The presence of FG water layer improved the populations of secondary structures to an extent that was in good agreement with the fully atomistic simulation results. However, only marginal improvements in other protein structural parameters were observed even with a rather thick (7 Å) layer of FG water. This was mainly due to the diffusion of FG water molecules into the bulk CG water, which caused large influx of CG water molecules at the solvation shell of protein. As a remedy, distance restraints were applied on the FG water molecules to adopt a flexible boundary around the protein molecule.[33] Simulations with distance restrained FG water molecules showed large improvements in the hydration state and in the structural qualities of the protein. The intraprotein electrostatic interactions were also largely reduced and protein−water hydrogen bonds increased to a reliable level. In addition, the distributions of polar and nonpolar residues on the protein surface were significantly improved. Overall, the structural aspects from the hybrid simulations were in close accord with the atomistic simulation results. Good agreements were obtained by adding restraining potential to prevent FG water diffusion into CG water.

However, there are still some remaining issues in constructing a robust hybrid CG-FG simulation scheme. The most important one is related to the application of the restraining potential mentioned in the above. As shown in the previous section, the position restrained FG water layer exerts stabilizing forces to the protein, which, in turn, reduces thermal fluctuations around the native structure and the related disorder of the protein. This will of course impose deviating behaviors to protein folding and configurational dynamics. Therefore, a more advanced approach of eliminating FG water diffusion is highly required. At this point, it will be pertinent to comment that there are other possibilities for suppressing the FG water diffusion into CG water besides the application of the restraining forces. For example, a scheme that smoothly converts a CG water particle into FG water when it diffuses into the close-by region near the FG protein can be applied, similarly to the already established adaptive resolution hybridization schemes.[36−39] Such a development will be especially important based on the above-mentioned observation that the restraints may induce limited structural fluctuations for FG proteins, which will likely induce unphysical events in studying intrinsically flexible systems with the hybrid scheme. Also, adjusting the FG-CG water interaction in a more strict way to generate a properly mixing combination model will be important to supplement such a scheme. We hope to report on these with the continuation of our present work.

■ ASSOCIATED CONTENT

Supporting Information

Fast unfolding of zinc finger protein with an improper hybrid setting (Figure S1); secondary structure stabilities of trpzip2, HP36, and trp-cage with pure FG-CG combinations (Figures S2 and S3); effects of cutoff criteria for salt-bridge registering (Figure S4); mixing behaviors of CG-FG water mixtures

(Figures S5 and S6, Table S1); structural behaviors from hybrid simulations with buffering FG water layers with and without distance restraints (Figures S7–S10). This information is available free of charge via the Internet at http://pubs.acs.org/.

AUTHOR INFORMATION

Corresponding Author

*E-mail: ymrhee@postech.ac.kr.

Author Contributions

[§]These authors contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the Institute for Basic Science (IBS) in Korea. The supercomputer time from Korea Institute of Science and Technology Information (KISTI) under Grant No. KSC-2012-C3-04 is also gratefully acknowledged.

REFERENCES

(1) Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. A Biomolecular Force Field Based on the Free Enthalpy of Hydration and Solvation: The GROMOS Force-Field Parameter Sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25*, 1656–1676.

(2) Cramer, C. J.; Truhlar, D. G. Implicit Solvation Models: Equilibria, Structure, Spectra, and Dynamics. *Chem. Rev.* **1999**, *99*, 2161–2200.

(3) Bursulaya, B. D.; Brooks, C. L. Comparative Study of the Folding Free Energy Landscape of a Three-Stranded β-Sheet Protein with Explicit and Implicit Solvent Models. *J. Phys. Chem. B* **2000**, *104*, 12378–12383.

(4) Rhee, Y. M.; Pande, V. S. On the Role of Chemical Detail in Simulating Protein Folding Kinetics. *Chem. Phys.* **2006**, *323*, 66–77.

(5) Zhou, R.; Berne, B. J. Can a Continuum Solvent Model Reproduce the Free Energy Landscape of a β-Hairpin Folding in Water? *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12777–12782.

(6) Zhou, R. Free Energy Landscape of Protein Folding in Water: Explicit vs Implicit Solvent. *Proteins* **2003**, *53*, 148–161.

(7) Das, A.; Andersen, H. C. The Multiscale Coarse-Graining Method. IX. A General Method for Construction of Three Body Coarse-Grained Force Fields. *J. Chem. Phys.* **2012**, *136*, 194114.

(8) Maisuradze, G. G.; Senet, P.; Czaplewski, C.; Liwo, A.; Scheraga, H. A. Investigation of Protein Folding by Coarse-Grained Molecular Dynamics with the UNRES Force Field. *J. Phys. Chem. A* **2010**, *114*, 4471–4485.

(9) Smith, A. V.; Hall, C. K. Assembly of a Tetrameric α-Helical Bundle: Computer Simulations on an Intermediate-Resolution Protein Model. *Proteins* **2001**, *44*, 376–391.

(10) Takada, S. Coarse-Grained Molecular Simulations of Large Biomolecules. *Curr. Opin. Struct. Biol.* **2012**, *22*, 130–137.

(11) Tozzini, V. Coarse-Grained Models for Proteins. *Curr. Opin. Struct. Biol.* **2005**, *15*, 144–150.

(12) Arkhipov, A.; Freddolino, P. L.; Schulten, K. Stability and Dynamics of Virus Capsids Described by Coarse-Grained Modeling. *Structure* **2006**, *14*, 1767–1777.

(13) Depa, P.; Chen, C. X.; Maranas, J. K. Why are Coarse-Grained Force Fields Too Fast? A Look at Dynamics of Four Coarse-Grained Polymers. *J. Chem. Phys.* **2011**, *134*, 014903.

(14) Lee, O. S.; Cho, V.; Schatz, G. C. Modeling the Self-Assembly of Peptide Amphiphiles into Fibers Using Coarse-Grained Molecular Dynamics. *Nano Lett.* **2012**, *12*, 4907–4913.

(15) Li, X.; Caswell, B.; Karniadakis, G. E. Effect of Chain Chirality on the Self-Assembly of Sickle Hemoglobin. *Biophys. J.* **2012**, *103*, 1130–1140.

(16) Rossi, G.; Elliott, I. G.; Ala-Nissila, T.; Faller, R. Molecular Dynamics Study of a MARTINI Coarse-Grained Polystyrene Brush in Good Solvent: Structure and Dynamics. *Macromolecules* **2012**, *15*, 563–571.

(17) Bond, P. J.; Wee, C. L.; Sansom, M. S. P. Coarse-Grained Molecular Dynamics Simulations of the Energetics of Helix Insertion into a Lipid Bilayer. *Biochemistry* **2008**, *47*, 11321–11331.

(18) Kraft, J. F.; Vestergaard, M.; Schiøtt, B.; Thøgersen, L. Modeling the Self-Assembly and Stability of DHPC Micelles Using Atomic Resolution and Coarse Grained MD Simulations. *J. Chem. Theory Comput.* **2012**, *8*, 1556–1569.

(19) Li, D. C.; Liu, M. S.; Ji, B. H.; Hwang, K.; Huang, Y. G. Coarse-Grained Molecular Dynamics of Ligands Binding into Protein: The Case of HIV-1 Protease Inhibitors. *J. Chem. Phys.* **2009**, *130*, 215102.

(20) Setny, P.; Bahadur, R. P.; Zacharias, M. Protein-DNA Docking with a Coarse-Grained Force Field. *BMC bioinformatics* **2012**, *13*, 228.

(21) Urbanc, B.; Betnel, M.; Cruz, L.; Bitan, G.; Teplow, D. B. Elucidation of Amyloid β-Protein Oligomerization Mechanisms: Discrete Molecular Dynamics Study. *J. Am. Chem. Soc.* **2010**, *132*, 4266–4280.

(22) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007**, *111*, 7812–7824.

(23) Fegan, S. K.; Thachuk, M. Suitability of the MARTINI Force Field for Use with Gas-Phase Protein Complexes. *J. Chem. Theory Comput.* **2012**, *8*, 1304–1313.

(24) Lopez, C. A.; Rzepiela, A. J.; de Vries, A. H.; Dijkhuizen, L.; Hunenberger, P. H.; Marrink, S. J. MARTINI Coarse-Grained Force Field: Extension to Carbohydrates. *J. Chem. Theory Comput.* **2009**, *5*, 3195–3210.

(25) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S. J. The MARTINI Coarse-Grained Force Field: Extension to Proteins. *J. Chem. Theory Comput.* **2008**, *4*, 819–834.

(26) Bennett, W. F. D.; Tieleman, D. P. Water Defect and Pore Formation in Atomistic and Coarse-Grained Lipid Membranes: Pushing the Limits of Coarse Graining. *J. Chem. Theory Comput.* **2011**, *7*, 2981–2988.

(27) Acevedo, O.; Jorgensen, W. L. Advances in Quantum and Molecular Mechanical (QM/MM) Simulations for Organic and Enzymatic Reactions. *Acc. Chem. Res.* **2010**, *43*, 142–151.

(28) Sherwood, P.; Brooks, B. R.; Sansom, M. S. Multiscale Methods for Macromolecular Simulations. *Curr. Opin. Struct. Biol.* **2008**, *18*, 630–640.

(29) Shi, Q.; Izvekov, S.; Voth, G. A. Mixed Atomistic and Coarse-Grained Molecular Dynamics: Simulation of a Membrane-Bound Ion Channel. *J. Phys. Chem. B* **2006**, *110*, 15045–15048.

(30) Darré, L.; Tek, A.; Baaden, M.; Pantano, S. Mixing Atomistic and Coarse Grain Solvation Models for MD Simulations: Let WT4 Handle the Bulk. *J. Chem. Theory Comput.* **2012**, *8*, 3880–3894.

(31) Masella, M.; Borgis, D.; Cuniasse, P. Combining a Polarizable Force-Field and a Coarse-Grained Polarizable Solvent Model: Application to Long Dynamics Simulations of Bovine Pancreatic Trypsin Inhibitor. *J. Comput. Chem.* **2008**, *29*, 1707–1724.

(32) Masella, M.; Borgis, D.; Cuniasse, P. Combining a Polarizable Force-Field and a Coarse-Grained Polarizable Solvent Model. II. Accounting for Hydrophobic Effects. *J. Comput. Chem.* **2011**, *32*, 2664–2678.

(33) Riniker, S.; Eichenberger, A. P.; van Gunsteren, W. F. Structural Effects of an Atomic-Level Layer of Water Molecules Around Proteins Solvated in Supra-Molecular Coarse-Grained Water. *J. Phys. Chem. B* **2012**, *116*, 8873–8879.

(34) Riniker, S.; Eichenberger, A. P.; van Gunsteren, W. F. Solvating Atomic Level Fine-Grained Proteins in Supra-Molecular Level Coarse-Grained Water for Molecular Dynamics Simulations. *Eur. Biophys. J.* **2012**, *41*, 647–661.

(35) Rzepiela, A. J.; Louhivuori, M.; Peter, C.; Marrink, S. J. Hybrid Simulations: Combining Atomistic and Coarse-Grained Force Fields Using Virtual Sites. *Phys. Chem. Chem. Phys.* **2011**, *13*, 10437–10448.

(36) Praprotnik, M.; Site, L. D.; Kremer, K. Adaptive Resolution Molecular Dynamics Simulation: Changing the Degrees of Freedom on the Fly. *J. Chem. Phys.* **2005**, *123*, 224106.
(37) Ensing, B.; Nielsen, S. O.; Moore, P. B.; Klein, M. L.; Parrinello, M. Energy Conservation in Adaptive Hybrid Atomistic/Coarse-Grain Molecular Dynamics. *J. Chem. Theory Comput.* **2007**, *3*, 1100−1105.
(38) Heyden, A.; Truhlar, D. G. Conservative Algorithm for an Adaptive Change of Resolution in Mixed Atomistic/Coarse-Grained Multiscale Simulations. *J. Chem. Theory Comput.* **2008**, *4*, 217−221.
(39) Poblete, S.; Praprotnik, M.; Kremer, K.; Site, L. D. Coupling Different Levels of Resolution in Molecular Simulations. *J. Chem. Phys.* **2010**, *132*, 114101.
(40) Wassenaar, T. A.; Ingólfsson, H. I.; Prieß, M.; Marrink, S. J.; Schäfer, L. V. Mixing MARTINI: Electrostatic Coupling in Hybrid Atomistic-Coarse-Grained Biomolecular Simulations. *J. Phys. Chem. B* **2013**, *117*, 3516−3530.
(41) Kim, E.; Jang, S.; Pak, Y. All-Atom Ab Initio Native Structure Prediction of a Mixed Fold (1FME): A Comparison of Structural and Folding Characteristics of Various $\beta\beta\alpha$ Miniproteins. *J. Chem. Phys.* **2009**, *131*, 195102.
(42) Lei, H.; Duan, Y. The Role of Plastic β-Hairpin and Weak Hydrophobic Core in the Stability and Unfolding of a Full Sequence Design Protein. *J. Chem. Phys.* **2004**, *121*, 12104.
(43) Kubelka, J.; Hofrichter, J.; Eaton, W. A. The Protein Folding 'Speed Limit'. *Curr. Opin. Struct. Biol.* **2004**, *14*, 76−88.
(44) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(45) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. GROMACS: Fast, Flexible, and Free. *J. Comput. Chem.* **2005**, *26*, 1701−1718.
(46) Sarisky, C. A.; Mayo, S. L. The $\beta\beta\alpha$ Fold: Explorations in Sequence Space. *J. Mol. Biol.* **2001**, *307*, 1411−1418.
(47) Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. Tryptophan Zippers: Stable, Monomeric β-Hairpins. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 5578−5583.
(48) McKnight, C. J.; Matsudaira, P. T.; Kim, P. S. NMR Structure of the 35-Residue Villin Headpiece Subdomain. *Nat. Struct. Mol. Biol.* **1997**, *4*, 180−184.
(49) Neidigh, J. W.; Fesinmeyer, R. M.; Andersen, N. H. Designing a 20-Residue Protein. *Nat. Struct. Mol. Biol.* **2002**, *9*, 425−430.
(50) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18*, 1463−1472.
(51) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.
(52) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(53) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(54) Wong-ekkabut, J.; Karttunen, M. Assessment of Common Simulation Protocols for Simulations of Nanopores, Membrane Proteins, and Channels. *J. Chem. Theory Comput.* **2012**, *8*, 2905−2911.
(55) Wu, Y.; Tepper, H. L.; Voth, G. A. Flexible Simple Point-Charge Water Model with Improved Liquid State Properties. *J. Chem. Phys.* **2006**, *124*, 024503.
(56) Frishman, D.; Argos, P. Knowledge-Based Protein Secondary Structure Assignment. *Proteins* **1995**, *23*, 566−579.
(57) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 27−38.
(58) Riniker, S.; van Gunsteren, W. F. Mixing Coarse-Grained and Fine-Grained Water in Molecular Dynamics Simulations of a Single System. *J. Chem. Phys.* **2012**, *137*, 044120.
(59) Because the CG water beads are relatively large, no CG water was present within 3 Å distance from the protein surface in all cases. Thus, we have adopted a slightly larger value (4 Å) for the solvation shell thickness than the widely accepted one (3 Å) for the CG case.
(60) Brancato, G.; Rega, N.; Barone, V. A Hybrid Explicit/Implicit Solvation Method for First-Principle Molecular Dynamics Simulations. *J. Chem. Phys.* **2008**, *128*, 144501.
(61) Im, W.; Berneche, S.; Roux, B. Generalized Solvent Boundary Potential for Computer Simulations. *J. Chem. Phys.* **2001**, *114*, 2924−2937.
(62) Lee, M. S.; Salsbury, F. R., Jr.; Olson, M. A. An Efficient Hybrid Explicit/Implicit Solvent Method for Biomolecular Simulations. *J. Comput. Chem.* **2004**, *25*, 1967−1978.
(63) Shivakumar, D.; Deng, Y. Q.; Roux, B. Computations of Absolute Solvation Free Energies of Small Molecules Using Explicit and Implicit Solvent Model. *J. Chem. Theory Comput.* **2009**, *5*, 919−930.
(64) Lin, Z.; Riniker, S.; van Gunsteren, W. F. Free Enthalpy Differences between α-, π-, and 3_{10}-Helices of an Atomic Level Fine-Grained Alanine Deca-Peptide Solvated in Supramolecular Coarse-Grained Water. *J. Chem. Theory Comput.* **2013**, *9*, 1328−1333.