JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

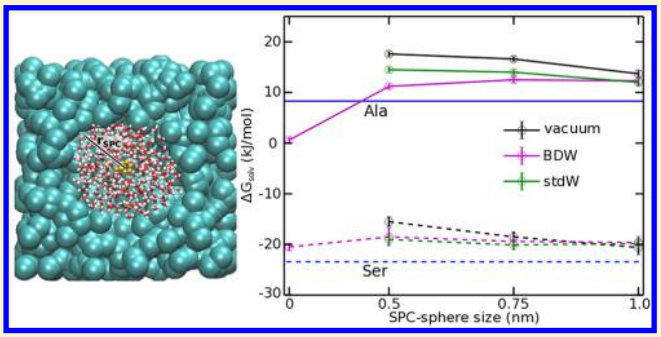
On Using Atomistic Solvent Layers in Hybrid All-Atom/Coarse-Grained Molecular Dynamics Simulations

Alexander B. Kuhn, Srinivasa M. Gopal, and Lars V. Schäfer*

Lehrstuhl für Theoretische Chemie, Ruhr-University Bochum, D-44780 Bochum, Germany

Supporting Information

ABSTRACT: Hybrid all-atom/coarse-grained (AA−CG) simulations in which AA solutes are embedded in a CG environment can provide a significant computational speed-up over conventional fully atomistic simulations and thus alleviate the current length and time scale limitations of molecular dynamics (MD) simulations of large biomolecular systems. On one hand, coarse graining the solvent is particularly appealing, since it typically constitutes the largest part of the simulation system and thus dominates computational cost. On the other hand, retaining atomic-level solvent layers around the solute is desirable for a realistic description of hydrogen bonds and other local solvation effects. Here, we devise and systematically validate fixed resolution AA−CG schemes, both with and without atomistic water layers. To quantify the accuracy and diagnose possible pitfalls, Gibbs free energies of solvation of amino acid side chain analogues were calculated, and the influence of the nature of the CG solvent surrounding (polarizable vs nonpolarizable CG water) and the size of the AA solvent region was investigated. We show that distance restraints to keep the AA solvent around the solute lead to too high of a density in the inner shell. Together with a long-ranged effect due to orientational ordering of water molecules at the AA−CG boundary, this affects solvation free energies. Shifting the onset of the distance restraints slightly away from the central solute significantly improves solvation free energies, down to mean unsigned errors with respect to experiment of 2.3 and 2.6 kJ/mol for the polarizable and nonpolarizable CG water surrounding, respectively. The speed-up of the nonpolarizable model renders it computationally more attractive. The present work thus highlights challenges, and outlines possible solutions, involved with modeling the boundary between different levels of resolution in hybrid AA−CG simulations.



1. INTRODUCTION

Computational cost is one of the major limitations of molecular dynamics (MD) simulations of condensed-phase systems at the fully atomistic (all-atom, AA) level. Despite tremendous progress in the field, the accessible length and time scales largely do not exceed 10^6 particles and several microseconds, respectively, even on modern supercomputers. Unfortunately, many biomolecular processes occur on slower time scales and, at the same time, they also involve larger system sizes. Coarse-grained (CG) models can speed up simulations significantly by combining several atoms into a single CG bead, thus extending the application range of MD simulations toward larger systems and longer time scales.[1−3] However, this gain in computational efficiency comes at the cost of reduced accuracy due to the approximations that are unavoidable upon coarse graining. In particular, flexible biological macromolecules such as proteins represent a major challenge for many CG force fields, especially the realistic description of folding/unfolding, secondary structure changes, conformational transitions, and the delicate balance between intra- and intermolecular interactions and changes in solvation that govern molecular recognition events.

The purpose of multiscale simulations is to overcome these drawbacks by combining models with different levels of resolution. In hybrid AA−CG methods, the two models are present simultaneously in the same simulation. This requires interactions between AA and CG particles. For typical biomolecular systems, solvent molecules (i.e., in most cases, water) represent the largest part of the simulation system and thus dominate the overall computational cost. Hence, to achieve a considerable computational speed-up, it is particularly rewarding to devise a hybrid model in which an AA force field is only used for a small subsystem, e.g., the biomolecular solute of interest, whereas a computationally cheap CG force field is used for the remainder, especially the bulk solvent. Various hybrid AA−CG methods have been developed, which can be categorized into two classes: fixed resolution and adaptive resolution methods. Adaptive resolution methods allow molecules to change their resolution on the fly upon diffusing from one spatial regime to another during the simulation.[4−9] A buffer (or healing) zone is introduced, in which the molecules gradually switch resolution between AA and CG.

By contrast, in fixed resolution methods, particles are *a priori* defined to be either AA or CG in nature, and this preassigned

Received: May 27, 2015
Published: July 28, 2015

DOI: 10.1021/acs.jctc.5b00499
J. Chem. Theory Comput. 2015, 11, 4460−4472

resolution does not change during the simulation. Such a setup is analogous to the well-established QM/MM methods and, similar to QM/MM, a major challenge lies in the coupling between the different levels of resolution, i.e., the description of the AA−CG interactions. These mixed AA−CG interactions can be obtained from corresponding all-atom simulations using force matching,[10] or derived from standard combination rules for the Lennard-Jones interactions.[11−13] In this latter approach of Essex and co-workers, additional system-dependent scaling parameters had to be introduced. This limitation is overcome in the recent AA−CG scheme of Orsi and co-workers,[14] which is based on the ELBA CG water model.[15,16] In all of the above AA−CG schemes, CG water models in which a single atomistic water molecule is mapped to a single CG molecule were applied, thus limiting the achievable computational speed-up. In the AA−CG scheme proposed by Rzepiela and co-workers,[17] the Martini CG water model[18] was used, in which four atomistic water molecules are represented by a single CG site. For the CG−CG and AA−AA interactions (except for the AA−AA electrostatics; see below), the respective pure force fields were used without further alterations. The mixed AA−CG interactions were described at the level of the CG force field, i.e., there were no *direct* AA−CG interactions. This was achieved by introducing CG virtual sites in the AA subsystem and mapping the forces on these virtual sites (from the CG interactions) onto the corresponding atomistic particles. A drawback of this scheme is that standard Martini water interacts purely via Lennard-Jones interactions, i.e., does not explicitly include electrostatics and hence cannot mediate any dielectric screening of the electrostatic interactions between the (partially) charged AA particles. To account for this, Rzepiela and co-workers[17] uniformly screened the AA−AA Coulomb interactions by a constant relative dielectric permittivity ε_r. Different values for ε_r had to be used in apolar and polar solvents, limiting the applicability of such an approach for large biomolecular systems in which charges are heterogeneously distributed. In a similar approach, an united-atom protein model was coupled with the Martini CG water model.[19,20]

Recently, different hybrid AA−CG models that include *explicit* electrostatic coupling between AA and CG subsystems were proposed.[14,21−24] By using polarizable CG water models that include partial charges, direct Coulomb interactions between AA and CG particles enable an orientational response of the CG solvent molecules to charges in their surroundings. Hence, the AA−AA electrostatic interactions are explicitly modulated, and an *ad hoc* screening of the AA−AA electrostatics is no longer required. In the model of Riniker and co-workers,[21,22] a polarizable two-site CG dipole water model[25] that effectively represents five water molecules (referred to as Gromos CG water or big dipole water, BDW) was coupled with the Gromos protein and SPC water force fields. The dielectric permittivity for the direct AA−CG electrostatic interactions, ε_r(AA−CG) = 2.3, was parametrized based on simulations of AA−CG water mixtures, and standard combination rules were used for the mixed Lennard-Jones interactions.[21] Short MD simulations of four AA proteins in BDW CG water showed that the proteins did not deviate much from their starting structures, and various structural features matched those from fully atomistic simulations and experiments.[22] However, since hydrogen bonds with the CG solvent are impossible, an increased tendency to form intraprotein hydrogen bonds was noticed.[22] Previously, we devised a conceptually similar AA−CG coupling scheme[23] that uses the polarizable Martini CG water model.[26] Potentials of mean force (PMFs) of the dimerization of atomistic ions and amino acid side chains in different polarizable CG water models showed that even very small variations in ε_r(AA−CG) can have drastic effects and lead to the formation of too stable contact pairs, especially between charged species. This effect, which was shown to even exceed 100 kJ mol^{-1} in some cases,[23] can overly stabilize and rigidify the structures of atomistic proteins in CG solvent. One possible way to improve this is to introduce shells of atomistic solvent molecules around the biomolecular solute of interest, as suggested recently.[24,27−30] In these shell models, the AA−CG boundary is shifted away from the solute, which is located at the center of the AA solvent region. To inhibit breakdown of the atomistic solvation shell through diffusional mixing of AA and CG solvent molecules in the course of the simulation, attractive harmonic distance restraints beyond a certain distance cutoff were introduced between the atomistic solvent molecules and the protein.[27−30]

In the present study, we devise different fixed resolution hybrid AA−CG schemes with distance-restrained atomic-level solvent layers, and we quantitatively examine their accuracy and performance. Our goal is to systematically and thoroughly validate the AA−CG schemes and to diagnose possible pitfalls. To that aim, we calculated Gibbs free energies of solvation, ΔG_{solv}, of amino acid side chain analogues in water. The agreement of hydration free energy with experiment and fully atomistic simulations was used as a proxy for accuracy since ΔG_{solv} is a challenging quantity to match: it involves specific interactions, such as solute−solvent hydrogen bonds and the rearrangement of intrasolvent hydrogen bonds around the solute, in addition to nonspecific van der Waals interactions. Thus, a proper description of hydration thermodynamics critically determines the accuracy of any (biomolecular) simulation model. We calculated ΔG_{solv} of 12 uncharged amino acid side chain analogues using thermodynamic integration (TI). We varied (i) the size of the atomistic inner solvation shell, described by SPC water, and (ii) the nature of the surrounding (polarizable/nonpolarizable CG water, or vacuum). The results are compared to reference simulations at the fully atomistic level and experimental data. Our results show that water molecules at the AA−CG boundary undergo preferential orientational ordering, an effect that extends up to 1 nm into the atomistic region. In addition, the use of distance restraints to retain the atomistic water shell around the central solute molecule leads to too high of a density in the inner shell. These effects have an impact on solvation properties. We found that shifting the onset of the half-harmonic restraints, but not the number of restrained atomistic water molecules, slightly toward larger distances from the central solute significantly improves ΔG_{solv}. Although our findings are strictly valid only for the particular AA−CG schemes investigated here, they provide general insight into the effects of different descriptions of the AA−CG boundary.

The rest of the article is organized as follows. First, in Methods, we describe the hybrid AA−CG models and details of the simulations. In the Results and Discussion, we first discuss the solvation of AA amino acid side chain analogues in pure polarizable BDW CG water. Then, we consider the influence of SPC spheres of different size and compare polarizable and nonpolarizable water models for the CG surrounding. Finally, we characterize the different AA−CG interfaces and investigate in detail the influence of the distance restraints. A Summary and Conclusions section closes this article.

DOI: 10.1021/acs.jctc.5b00499
J. Chem. Theory Comput. 2015, 11, 4460−4472

Table 1. Nonbonded Interactions in Hybrid AA−CG Simulations

	AA−AA			AA−CG			CG−CG		
	stdW	BDW	polW	stdW	BDW	polW	stdW	BDW	polW
ε_r	1.0	1.0	1.0		2.3	1.45[b]		2.5	2.5
Coulomb cutoff[a] [nm]	1.4	2.0	1.4		2.0	1.4		2.0	1.4
Lennard-Jones cutoff [nm]	1.4	2.0	1.4	1.4	2.0	1.4	1.2[c]	2.0	1.2[c]

[a]Reaction-field (ε_{rf} = 78). [b]Wassenaar et al.[23] [c]Shifted 0.9−1.2 nm.

2. METHODS

2.1. Hybrid AA−CG Models. We devised a computationally efficient hybrid AA−CG method that couples an atomistic force field, including SPC water for the inner solvent region, with the nonpolarizable standard Lennard-Jones Martini[18] (stdW) water model for the CG part. For the AA−CG simulations with polarizable CG water, we employed the model proposed by Riniker and co-workers for the BDW model.[21,22,27] In addition, we used the polarizable Martini water model.[26] The stdW and polW models have previously been used by Zavadlav and co-workers in adaptive resolution AA−CG simulations.[31−33] In these simulations, a bundled SPC water model[34] was used for the inner (atomistic) region, which is not required in our setup. In our hybrid AA−CG simulations, the interactions between particles belonging to the same resolution (i.e., AA−AA and CG−CG) are described with the respective pure force fields without further alterations. In this work, we used the Gromos 53a6[35] force field for the atomistic part. However, as outlined in the following, the AA−CG scheme is modular and can be readily extended to other atomistic force fields as well. For some systems, additional AA−CG simulations with the polarizable Martini CG water model (polW)[26] were carried out. The polW model is a 3-site model that applies a 4-to-1 mapping, i.e., a single polW molecule effectively represents 4 atomistic water molecules by a central, uncharged Lennard-Jones bead to which two oppositely charged satellite beads (of charge ±0.46) are bonded at a fixed bond length of 0.14 nm.[26] The 3 beads are connected by a weak harmonic angle potential energy function. By contrast, a BDW CG water molecule represents 5 atomistic water molecules.[25] It is a dipole consisting of two oppositely charged particles (±0.575). Only the negatively charged CW particle carries Lennard-Jones C6 and C12 parameters, whereas the positively charged DP particle does not have Lennard-Jones interactions. The particles are connected by a quartic bond, $V_{bond}(r_{CW-DP}) = 0.5 \cdot k_{bond}(r_{CW-DP} - r_0)^4$, which acts beyond a distance of r_0 > 0.2 nm with a force constant of $k_{bond} = 2 \cdot 10^6$ kJ mol^{-1} nm^{-4}. The stdW standard Martini water model replaces 4 waters by a single Lennard-Jones bead and is thus devoid of any explicit electrostatics, unlike the polarizable CG models.

The cross-interactions between AA and CG particles are treated as follows. Standard combination rules (geometric averaging) were used to obtain the Lennard-Jones C6 and C12 parameters for the AA−CG interactions. Thus, no further parametrization of cross-interactions or the inclusion of CG virtual sites, as done before,[17,23] was needed for the Lennard-Jones interactions. An additional benefit of this choice is that no bundling potential energy functions have to be applied to groupwise restrict the motions of water molecules in the atomistic solvation shells.[34,36] For the AA−CG electrostatic interactions, which occur only in the simulations with the BDW and polW models, the approach suggested previously[21] was used. All electrostatic interactions, including the AA−AA and, in case of BDW and polW, CG−CG interactions, were described using reaction field potential energy functions and forces, $V_{rf} = q_i q_j (4\pi\varepsilon_0\varepsilon_r)^{-1}(r_{ij}^{-1} + k_{rf} r_{ij}^2 - c_{rf})$ and $F_{rf} = q_i q_j (4\pi\varepsilon_0\varepsilon_r)^{-1}(r_{ij}^{-2} - 2k_{rf} r_{ij})$, respectively, with $k_{rf} = r_c^{-3}(\varepsilon_{rf} - \varepsilon_r)(2\varepsilon_{rf} + \varepsilon_r)^{-1}$ and $c_{rf} = r_c^{-1} + k_{rf} r_c^2$. Here, q_i and q_j are the charges on particles i and j, respectively, r_{ij} is the distance between them, ε_0 and ε_r are the vacuum permittivity and the relative dielectric permittivity inside the cutoff sphere, respectively; ε_{rf} = 78 is the dielectric constant of the continuum and r_c, the cutoff radius. The BDW model was parametrized with reaction field,[25] and test simulations showed that the density and static dielectric permittivity of the polW model, which was originally parametrized to be used with shifted cutoff or PME,[26] are not affected by this choice. The above potential energies and forces were tabulated with a spacing of 0.002 nm. Different values of r_c and ε_r were used for the different interactions, as summarized in Table 1. Since the BDW model was developed to be used with a 2.0 nm nonbonded cutoff, this cutoff was also used for the AA−AA and AA−CG interactions in the hybrid simulations. Such a large cutoff was not necessary for the stdW and polW models, which were parametrized with shorter (and thus computationally more efficient) cut-offs.

In AA−CG simulations with the polarizable CG water models, the distance between the charged dummy particles of BDW or polW and charged atomistic particles can become very small, which can lead to numerical instabilities. To avoid this polarization catastrophe, a purely repulsive Lennard-Jones term with a C12-parameter of $7.7848 \cdot 10^{-10}$ kJ mol^{-1} nm^{12} was introduced for the interaction between the CG water dummy particles and AA atoms, as suggested previously.[21]

To construct atomistic solvent spheres around the solutes, equilibrated configurations were extracted from fully AA simulations (1 ns MD simulations, periodic boundary conditions, NpT ensemble). From these fully AA simulations, we determined the average number of water molecules $\langle N \rangle$ within a sphere of radius r_{SPC} around the center of mass (COM) of the central solute molecule. To set up the AA−CG simulation system, a sphere containing the solute and $\langle N \rangle$ SPC water molecules was extracted from the AA simulation, and the simulation box was filled with CG water such that the desired density was obtained, taking care that initially no CG water was placed inside the atomistic subsystem (Figure 1). The instantaneous number of water molecules, N, within a sphere of a certain radius around the COM of the solute fluctuates significantly and can easily vary by 5−10% for the investigated sphere sizes. During the subsequent AA−CG simulations, flat-bottom half-harmonic distance restraints beyond a certain distance, defined by r_{SPC}, were used to keep the SPC molecules around the solute, as suggested previously.[27,29] The contribution of the restraining forces to the virial was taken into account. In addition, we also carried out simulations of SPC spheres in a vacuum surrounding. In both cases, a restraining potential energy function with a force constant of k_{dr} = 300 kJ mol^{-1} nm^{-2} was defined between the oxygen atoms of the SPC

DOI: 10.1021/acs.jctc.5b00499
J. Chem. Theory Comput. 2015, 11, 4460−4472

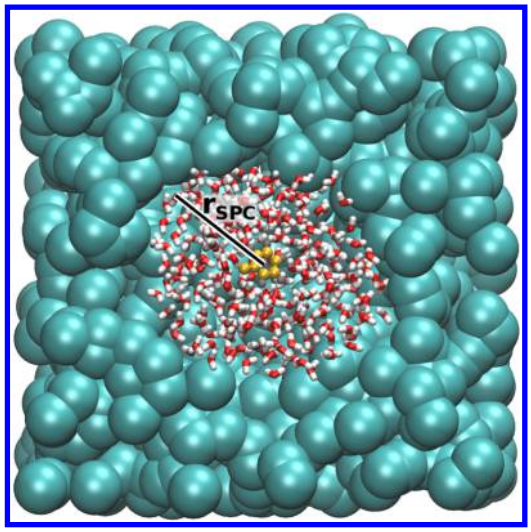


Figure 1. Hybrid AA−CG simulation. An atomistic solute surrounded by atomistic SPC water layers is embedded in CG water. Half-harmonic attractive distance restraints beyond r_{SPC} are applied between the solute and the atomistic water molecules to keep the atomistic solvent around the solute during the MD simulation.

molecules and the COM of the solute.[27] In the simulations of the atomistic spheres in vacuum, the water/vacuum surface tension stabilizes the spherical shape, and distance restraints are thus, in principle, not required. Likewise, in the hybrid AA−CG simulations with the stdW CG water surrounding, spontaneous mixing of SPC and stdW solvent molecules does not occur. Nevertheless, distance restraints were applied in these cases as well: (i) for consistency with the simulations in polarizable CG water surrounding and (ii) to keep the solutes close to the center of the atomistic solvent sphere. To investigate the effect of the thickness of the SPC water layer, we considered three different sphere sizes (0.5, 0.75, and 1.0 nm). To have a comparable number of water layers around the different solutes, the size of the atomistic hydration sphere was defined with respect to the outermost atom of the solute, not to its COM. For example, an atomistic layer of size 1.0 nm corresponds to r_{SPC} = 1.0, 1.13, and 1.36 nm for the Ala, Ser, and Trp side chains, respectively. Nevertheless, for simplicity, we refer to all of these as 1.0 nm sphere systems in the text.

2.2. Molecular Dynamics Simulations. All simulations were performed with the Gromacs MD simulation package (version 4.6.5).[37] The temperature was kept constant at 298 K using mild Langevin coupling (stochastic dynamics (SD) leapfrog integrator[38] in Gromacs) with a friction coefficient of 1 ps^{-1}. The pressure was kept constant at 1 bar with a Berendsen barostat,[39] with time constant τ_p = 1 ps and compressibility $4.5 \cdot 10^{-5}$ bar^{-1}. The bonds of the solutes and, in the AA−CG simulations, also the internal degrees of freedom of the SPC water molecules were constrained using LINCS,[40,41] with the number of iterations to correct for the rotational lengthening increased to 8. In the fully atomistic simulations, SETTLE[42] was to constrain the water molecules. An integration time step of 2 fs was used for all systems. The different cutoffs for the nonbonded interactions are listed in Table 1. Nonbonded interactions within the cutoff were calculated at every time step, and the charge group-based neighbor lists were updated every 20 fs.

Thermodynamic integration (TI) calculations were performed to determine the Gibbs free energies of solvation of those amino acid side chain analogues that do not carry a net charge at pH 7, as listed in Table 2 (His, Ile, and Pro are also uncharged at pH 7 but were not considered). For reference, in addition to the hybrid AA−CG systems with atomistic solvent shells of different size, ΔG_{solv} was also determined for the atomistic solutes embedded in pure BDW CG water. The corresponding reference values from fully atomistic simulations (i.e., solutes in pure SPC water) are taken from Hess and van der Vegt.[43] To set up the simulations, single uncharged amino acid side chain analogues were solvated with 4212 SPC water molecules for the fully AA simulations or 842 BDW (5:1 mapping) molecules for the fully CG simulations in a periodic cubic simulation box of size 5.0 nm. The compositions (in terms of numbers of particles) of the hybrid AA−CG systems depend on the size of the atomistic solvent sphere and are given in the Supporting Information. The simulation box was deliberately chosen to be rather large to avoid any possible finite size effects, which could be more pronounced in the CG and hybrid AA−CG simulations than in the fully atomistic systems due to the long-range structure in the pair correlation function of the CG solvents.

Table 2. Amino Acid Side Chain Analogues Studied in This Work

amino acid	3-letter code	1-letter code	analogue
alanine	Ala	A	methane
asparagine	Asn	N	acetamide
cysteine	Cys	C	methanethiol
glutamine	Gln	Q	propionamide
leucine	Leu	L	isobutane
methionine	Met	M	methyl ethyl sulfide
phenylalanine	Phe	F	toluene
serine	Ser	S	methanol
threonine	Thr	T	ethanol
tryptophan	Trp	W	3-methylindole
tyrosine	Tyr	Y	*p*-cresol
valine	Val	V	propane

In the TI, the nonbonded interactions between solute and solvent were turned off using a coupling parameter λ. Soft-core interaction energies were used, $V_{sc} = (1 - \lambda)V([\alpha\sigma^6\lambda + r^6]^{1/6})$, where $V(r)$ is the original Lennard-Jones potential energy function, α = 0.6 is the potential energy height, and σ = 0.28 nm is the interaction range.[44] The basic spacing between the λ-points was 0.025, i.e., 41 λ-points bridge between the fully coupled state (at λ = 0) and the fully decoupled state (at λ = 1). For the polar side chains, a closer λ-spacing of 0.005 was used between λ = 0 and 0.1 to capture the curvature of the $\langle \partial H/\partial\lambda \rangle$-over-$\lambda$ curve in the early transition region. Thus, in these cases, 57 λ-points were used. To generate starting structures for the simulations, the solutes were solvated at the correct density (as described above) and energy minimized at each λ-point (1000 steps steepest descent). Every λ-point was sampled for 1 ns, and the derivative of the Hamiltonian $\partial H/\partial\lambda$ was gathered, discarding the first 100 ps of the trajectories. The $\langle \partial H/\partial\lambda \rangle$-over-$\lambda$ curves were integrated using the trapezoidal rule to obtain ΔG. Statistical errors at each λ-point were estimated from the limiting values of the block averages.[45] These errors were integrated to estimate upper bounds for the errors in ΔG_{solv}, which are below a maximum error of 2.7 kJ mol^{-1} in all cases, often about 1.0 kJ mol^{-1} (see Supporting Information).

DOI: 10.1021/acs.jctc.5b00499
J. Chem. Theory Comput. 2015, 11, 4460−4472

To obtain the enthalpies of solvation ΔH_{solv}, 100 ns MD simulations of (i) the solute in solvent, (ii) the pure solvent (consisting of the same number of solvent molecules as used before), and (iii) the solute molecule in vacuo were performed. For the vacuum simulations, no cutoffs were used for the nonbonded interactions, and the overall translation and rotation was removed at every time step. The other simulation parameters were the same as described above. The hydration enthalpy ΔH_{solv} was calculated by subtracting the sum of the potential energies of the pure solvent and the solute in vacuum from the system potential energy, $\Delta H_{solv} \approx \langle U_{system} \rangle - (\langle U_{solvent} \rangle + \langle U_{solute} \rangle)$, where the brackets indicate averages over the 100 ns simulations. Statistical errors were estimated using block averaging, as described above. The $p\Delta V$ contribution is smaller than the statistical error and was thus neglected. The hydration entropy $T\Delta S_{solv}$ was obtained from the difference $\Delta H_{solv} - \Delta G_{solv}$. The maximum statistical error estimates in ΔH_{solv} and $T\Delta S_{solv}$ are 3.3 and 3.4 kJ mol^{-1}, respectively; for many side chain analogues, the errors are much smaller (see Supporting Information).

To characterize the properties of the SPC/CG and SPC/vacuum interfaces, we took the atomistic simulation box containing 4212 SPC molecules (pre-equilibrated in the NpT ensemble) and elongated it in the z-direction, yielding a simulation box with xyz box vectors of 5.06, 5.06, 10.12 nm. The empty half of this simulation box was either left empty (i.e., SPC/vacuum interface) or filled with CG water molecules such that an overall density of 975 kg m^{-3} was obtained (SPC/CG interfaces). To prevent the SPC water from entering into the CG subsystem during the subsequent 25 ns NVT simulations, external half-harmonic potential energy functions along the z-dimension with force constant 300 kJ mol^{-1} nm^{-2} (identical to the TI simulations) were applied on the SPC oxygens if the z-coordinate of the oxygen was smaller than 2.5 nm or larger than 7.5 nm. These simulations, which were carried out with a version of Gromacs[46,47] that includes an implementation of mean-field force approximation boundary potentials,[48] allowed us to study flat interfaces that are more straightforward to analyze.

3. RESULTS AND DISCUSSION

3.1. Solvation of Atomistic Solutes in Pure Coarse-Grained Water. First, before turning to the AA−CG simulations with atomistic solvent shells, we discuss the solvation of atomistic amino acid side chain analogues in pure BDW CG water. Figure 2 compares the solvation free energies obtained from our simulations with the experimental data and with the values obtained from fully atomistic simulations in pure SPC water,[43] for reference. The data are also summarized in Table 3.

To quantify the overall accuracy of the solvation free energies, the last column in Table 3 contains the deviations from the experimental values. For the fully atomistic simulations in SPC water, solvation free energies are very close to the experimental values.[49] For the hybrid simulations of atomistic solutes directly embedded in pure BDW CG water, the mean unsigned error (MUE) is 8.2 kJ mol^{-1}. This can be compared to other hybrid schemes in which AA solutes are directly embedded in polarizable CG water. Wassenaar and co-workers[23] reported MUEs in ΔG_{solv} of 6.3 and 7.3 kJ mol^{-1} for the polarizable Martini water model[26] and the big multipole water (BMW) model,[50] respectively. The recent AA−CG scheme of Orsi and co-workers[14] yields MUEs of 3.8 and 5.1 kJ mol^{-1} for atomistic side chains modeled with the CHARMM general force field and the general Amber force field, respectively. For BDW, Figure 2 shows that the solubility of the strongly polar side chains (Asn, Gln) is underestimated, whereas ΔG_{solv} of the apolar side chains Ala, Val, Leu, Phe, and Met is too negative, i.e., these side chains are too soluble in BDW. For the remaining side chains, the free energies of solvation are in good agreement with the experimental values and the fully atomistic simulations. Taken together, these results show that, for the hybrid AA−CG scheme employing the BDW CG water model, it is difficult to achieve a balanced and accurate treatment of both hydrophilic and hydrophobic side chains in terms of solvation free energy.

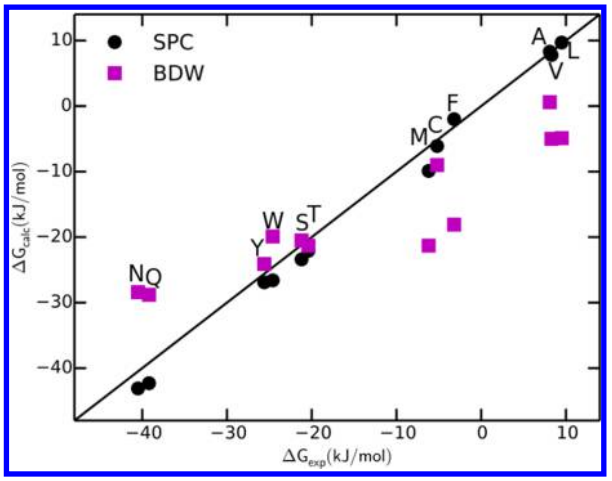


Figure 2. ΔG_{solv} of atomistic amino acid side chain analogs in pure BDW CG water (magenta squares) and pure SPC (black circles). The experimental values were taken from Wolfenden et al.[49]

Next, we decomposed ΔG_{solv} into the enthalpic and entropic contributions. As can be seen in Figure 3, both ΔH_{solv} and $T\Delta S_{solv}$ are overestimated (i.e., not negative enough), with the exception of the solvation enthalpies of Phe and Met, which are underestimated. These systematic deviations are also reflected in the fact that the mean signed errors (MSEs) are positive and (almost) equal to the mean unsigned errors. The MUEs in ΔH_{solv} and $T\Delta S_{solv}$ are 17.1 and 18.3 kJ mol^{-1}, respectively; the respective MSEs are 14.9 and 18.3 kJ mol^{-1} (all individual values with their statistical errors are listed in the Supporting Information). Hence, the MUE in ΔG_{solv} of 8.2 kJ mol^{-1} results from a significant degree of error compensation between ΔH_{solv} and $T\Delta S_{solv}$. This compensation of errors is expected because, by construction, the entropy of the CG water model is smaller than for a corresponding fully atomistic model due to the reduced number of degrees of freedom. Thus, to compensate for this effect and obtain a desired target free energy that is similar to the fully atomistic case, some of this missing entropy has to be implicitly included in the CG interaction energy. For the BDW model, the internal energy per CG bead is −23.3 kJ mol^{-1} at 298 K,[25] which is much less negative than for atomistic water (−42.0 kJ mol^{-1}). Since ΔH_{solv} is dominated by water−water interactions, this leads to the too positive hydration enthalpies. Likewise, the too small entropy of CG water leads to too positive hydration entropies, since, for the small solutes studied here, entropic contributions due to changes in intramolecular degrees of freedom are negligible and ΔS_{solv} is dominated by the solvent response to the presence of a solute.

Figure 3 shows that the polar Asn and Gln side chains most strongly disagree with the experimental values, with ΔH_{solv} and

DOI: 10.1021/acs.jctc.5b00499
J. Chem. Theory Comput. 2015, 11, 4460−4472

Table 3. Free Energies of Solvation ΔG_{solv} and Mean Unsigned Error MUE (Mean Signed Error (MSE) in Parentheses) with Respect to Experiment[a]

		Ala	Asn	Cys	Gln	Leu	Met	Phe	Ser	Thr	Trp	Tyr	Val	MUE (MSE)
SPC[b]		8.3	−43.1	−6.1	−42.3	9.7	−9.9	−2.0	−23.4	−22.1	−26.6	−26.9	7.8	1.6 (1.4)
experiment[c]		8.1	−40.5	−5.2	−39.2	9.5	−6.2	−3.2	−21.2	−20.4	−24.6	−25.6	8.3	
BDW		0.6	−28.4	−9.0	−28.8	−4.9	−21.3	−18.1	−20.5	−21.3	−19.9	−24.1	−5.0	8.2 (3.3)
SPC sphere in BDW	0.5 nm	11.2	−37.7	−1.8	−32.9	17.9	−3.6	2.8	−18.5	−15.1	−8.3	−11.4	14.1	6.4 (6.4)
	0.5 nm[d]	2.4	−38.9	−7.9	−38.4	0.5	−16.8	−11.7	−22.0	−21.2	−21.9	−24.5	0.2	4.4 (3.3)
	0.75 nm	12.5	−36.5	−0.5	−32.7	19.7	−0.1	5.5	−19.4	−14.9	−6.8	−10.4	16.0	7.6 (7.6)
	0.75 nm[d]	5.3	−40.9	−5.9	−38.0	6.4	−11.1	−5.8	−22.4	−20.1	−7.6	−21.0	5.3	3.5 (0.4)
	1.0 nm	12.3	−37.2	−1.3	−33.0	18.4	−1.4	3.9	−19.6	−15.7	−8.5	−11.9	14.8	6.5 (6.5)
	1.0 nm[d]	8.9	−40.2	−4.2	−37.6	12.0	−6.8	−1.8	−21.2	−19.1	−15.9	−16.9	9.7	2.3 (2.2)
	1.0 nm[e]	10.7	−37.5	−3.6	−37.5	14.1	−5.7	−1.4	−21.5	−18.8	−15.7	−17.0	12.0	3.2 (3.2)
SPC sphere in stdW	0.5 nm	14.5	−37.7	−0.9	−32.2	20.9	−8.2	4.1	−19.0	−14.2	−5.5	−8.8	17.2	7.9 (7.5)
	0.5 nm[d]	2.0	−49.5	−0.1	−47.0	−3.9	−21.8	−19.3	−22.1	−25.8	−28.3	−30.2	−3.5	8.3 (7.4)
	0.75 nm	14.0	−37.2	−2.2	−33.8	18.5	−1.7	1.5	−20.1	−16.0	−6.7	−10.2	15.5	6.8 (6.8)
	0.75 nm[d]	8.4	−42.5	−3.7	−37.6	11.0	−7.7	−2.9	−21.1	−18.2	−15.2	−15.2	8.7	2.6 (2.0)
	1.0 nm	12.0	−37.8	−2.3	−35.9	16.8	−4.4	0.6	−19.8	−16.8	−9.3	−12.8	13.3	5.3 (5.3)
	1.0 nm[d]	11.3	−39.3	−3.5	−36.5	14.0	−4.8	−2.8	−21.2	−18.2	−12.2	−14.1	11.8	3.7 (3.7)
SPC sphere in vacuum	0.5 nm	17.6	−31.1	6.8	−23.5	31.2	10.9	19.8	−15.5	−7.4	1.5	−4.6	26.7	16.0 (16.0)
	0.75 nm	16.6	−33.1	1.9	−29.1	25.0	4.5	13.1	−18.5	−12.2	−8.8	−12.3	20.9	10.7 (10.7)
	1.0 nm	13.7	−35.6	−0.1	−32.4	21.1	0.6	9.4	−20.6	−14.6	−13.0	−15.3	16.8	7.5 (7.5)
	1.0 nm[d]	12.9	−37.1	−1.4	−33.9	18.6	−1.7	7.7	−19.2	−16.2	−16.0	−18.0	15.6	5.9 (5.9)
SPC	ρ/ρ_{ref} = 1.06	13.7	−42.1	1.7	−36.0	21.2	2.7	12.8	−19.4	−12.7	−8.2	−9.9	17.2	8.7 (8.4)

[a]All values are in kJ mol^{-1}. [b]Data from Hess and van der Vegt.[43] [c]Data from Wolfeneden et al.[49] [d]Onset of the distance restraints on SPC increased to $r'_{SPC} = r_{SPC} + 0.3$ nm. [e]As in footnote (d) but with additional repulsive distance restraints on BDW (onset at $r_{BDW} = r_{SPC} - 0.1$ nm).

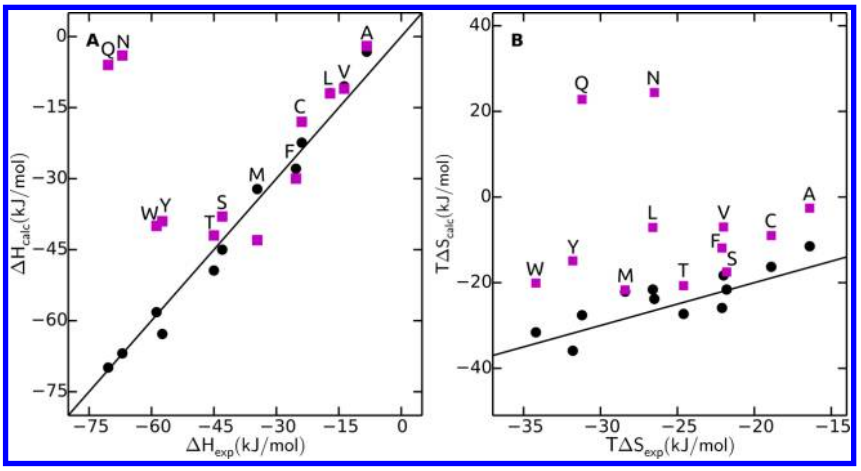


Figure 3. Calculated ΔH_{solv} and $T\Delta S_{solv}$ of atomistic amino acid side chain analogues in pure SPC water (black circles)[43] and in pure BDW CG water (magenta squares) are compared to experimental values.[49,51] (A) Solvation enthalpy ΔH_{solv}; (B) solvation entropy $T\Delta S_{solv}$ at 298 K.

$T\Delta S_{solv}$ deviating by more than 40 and 60 kJ mol^{-1}, respectively. Although exploring the reasons underlying these very large deviations is beyond the scope of this study, it is obvious that improving the solvation behavior of these two side chains would considerably improve the overall performance of the model. Leaving Asn and Gln out of consideration reduces the MUE in ΔH_{solv} and $T\Delta S_{solv}$ to 6.5 and 9.5 kJ mol^{-1}, respectively. Nevertheless, these errors are still significantly larger than for the fully atomistic simulations in SPC water, which yield MUEs in ΔH_{solv} and $T\Delta S_{solv}$ of 2.8 and 3.5 kJ mol^{-1}, respectively.[43]

3.2. Including Atomic-Level Solvent Shells. The previous section showed that obtaining a balanced and accurate description of the solvation of both hydrophobic and hydrophilic AA solutes in pure CG water is challenging. Thus, we next explored whether including atomistic solvent shells lead to an improvement. In these shell models, the AA solute is located at the center of an AA water sphere, which, in turn, is surrounded by CG solvent (Figure 1). We first discuss the influence of the nature of the surrounding (polarizable CG water model, nonpolarizable CG water model, or vacuum) before we study the effect of varying the size of the inner, atomic-level solvent region.

The results from the hybrid AA−CG simulations are summarized in Table 3. Overall, including atomistic water significantly improves the description of side chain solvation. The MUEs of ΔG_{solv} are 7.5, 6.5, and 5.3 kJ mol^{-1} for a 1.0 nm SPC sphere surrounded by vacuum, BDW, and stdW CG water, respectively (Table 3). These deviations are from the simulations in which the onset distance of the half-harmonic

DOI: 10.1021/acs.jctc.5b00499
J. Chem. Theory Comput. 2015, 11, 4460−4472

restraints exactly matches the volume occupied by the affected ⟨N⟩ SPC water molecules (see Methods). As discussed below, shifting the onset of the restraints to slightly larger distances reduces these MUEs to 5.9, 2.3, and 3.7 kJ mol^{-1}, respectively. Compared to the above simulations of AA solutes in pure BDW, for the apolar side chains (Ala, Leu, Phe, Val), including AA solvent layers leads to the correct (positive) sign of ΔG_{solv}, which arises from the entropically unfavorable orientational ordering of water molecules around the hydrophobe (hydrophobic effect) that obviously can be captured only at the AA level. For the more polar solutes, the improved description of local electrostatic interactions and hydrogen bonding improves the solvation free energies. However, ΔG_{solv} is now too positive (or, for the polar side chains, not negative enough). This trend is observed for all side chains, irrespective of their size and polarity. As a consequence, for some side chains whose solvation free energies in pure BDW were close to experimental values (partly due to fortuitous error compensation; see above), addition of an AA solvation sphere actually worsens the results, which is thus clearly the wrong trend. This is particularly the case for Thr, Trp, and Tyr. Finally, to investigate the role of the nature of the polarizable CG solvent surrounding, we calculated ΔG_{solv} for the Gln, Phe, Ser, and Val side chains in SPC spheres embedded in polarizable Martini CG water[26] instead of BDW. The deviations of the solvation free energies from the experimental values (see Supporting Information) are comparable to BDW, suggesting that the details of the polarizable CG solvent model do not play a major role. The similar behavior of the BDW and polW CG models is also supported by our simulations of flat AA−CG interfaces; see below.

Taken together, including atomistic solvent layers improves the description of side chain solvation. Surprisingly, embedding the atomistic subsystem in the computationally efficient standard Lennard-Jones Martini CG water model (stdW) yields comparable results to the computationally more expensive polarizable BDW CG water. Even neglecting the outermost layer entirely, as is the case for an SPC sphere in a vacuum surrounding, yields better results than solvating the AA solute directly in BDW CG water.

To investigate these effects more closely, in the following, we have a closer look on the influence of the size of SPC solvation shell. To that aim, in addition to the 1.0 nm SPC spheres discussed above, we have considered SPC spheres of radius 0.5 and 0.75 nm. The data is summarized in Table 3, and Figure 4 shows the results for Ala and Ser as two representative examples. Even a computationally efficient setup with a rather small atomistic solvent sphere seems to be sufficient, especially for BDW, as adding more AA solvent layers changes ΔG_{solv} only moderately. The agreement with the fully AA reference slightly improves with increasing size of the atomistic sphere, with the exception of the BDW model, for which this systematic improvement is not observed (Table 3). In summary, these results demonstrate the importance of an atomic-level description of the first solvation shells.

3.3. Reasons for the Observed Deviations. In the hybrid AA−CG simulations with an SPC sphere, the water solubility of all simulated side chain analogues is too low (ΔG_{solv} is too positive), suggesting an underlying general phenomenon. To unravel the reason behind this observation, we analyzed the number of solvent molecules around the central solute, as observed in the AA−CG simulations, and compared it to the fully atomistic reference simulations in pure SPC water. Figure 5 shows the results for the Val side chain as a representative

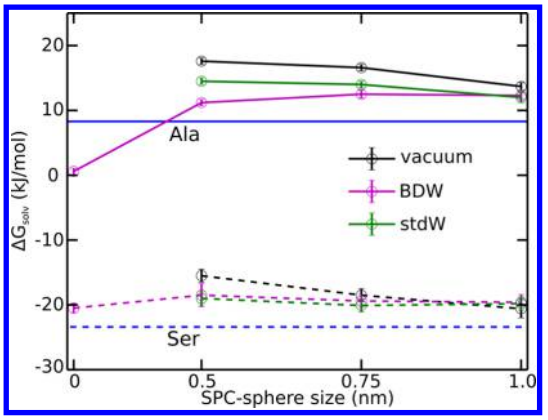


Figure 4. Free energy of solvation as a function of the size of the atomistic solvent sphere for Ala (solid lines) and Ser (dashed lines) in BDW (magenta), stdW (green), or in vacuum (black). Reference values from fully atomistic simulations in SPC[43] are shown as blue lines. A sphere size of zero corresponds to pure BDW CG water.

example; similar results were obtained for all solutes. In the AA−CG simulations, for both the BDW and stdW CG water surroundings, the local solvent density in the vicinity of the solute at the center of the SPC sphere is higher than in the reference simulation (black solid line). As shown in the inset of Figure 5a, this density increase is about 6%, averaged over a distance range from 0.7 to 1.5 nm. A similar effect was found in the simulations with polW water surrounding (results not shown). To isolate the effect of the distance restraints on the SPC water molecules, we carried out an additional simulation of a fully atomistic system in which we distance-restrained 190 SPC water molecules around the Val side chain, identical to the previous hybrid AA−CG simulations. Indeed, this simulation also yielded a density around the solute that was too high by about 6% (dashed black line in the inset of Figure 5a), confirming that the distance restraints are responsible for the artificial density increase. Additional TI simulations for this system yielded ΔG_{solv} = +16.6 kJ mol^{-1}, in agreement with the too positive solvation free energies observed in the SPC spheres in CG water surrounding (Table 3). We thus attribute this effect to the larger free energy penalty of creating a cavity around the solute in the too dense solvent. As a final verification of our hypothesis that the density increase significantly contributes to the observed deviations of the solvation free energies, we recalculated ΔG_{solv} of all 12 amino acid side chains in fully atomistic SPC water box with a constant density (NVT ensemble) that was 6% higher than that in our previous NpT simulations (Table 3, last row). As expected, these solvation free energies are in good agreement with those from the AA−CG simulations with the (too dense) SPC water spheres. Notably, the pressure was very high in these NVT simulations, about 1475 bar. A similar pressure of ca. 1 kbar is thus also expected within the SPC spheres in the AA−CG simulations.

The observed pronounced impact on the density may seem surprising at first, given that the distance restraint energies are rather small. Taking the Val side chain in a 1.0 nm SPC sphere (190 atomistic water molecules) again as a representative example system, the total average distance restraint energies are 33.7 kJ mol^{-1} (0.18 kJ mol^{-1} per SPC molecule) and 9.3 kJ mol^{-1} (0.05 kJ mol^{-1} per SPC) for the AA−CG simulations in BDW and stdW surrounding, respectively, which is negligible compared to the total potential energy of the systems (−26 378

DOI: 10.1021/acs.jctc.5b00499
J. Chem. Theory Comput. 2015, 11, 4460−4472

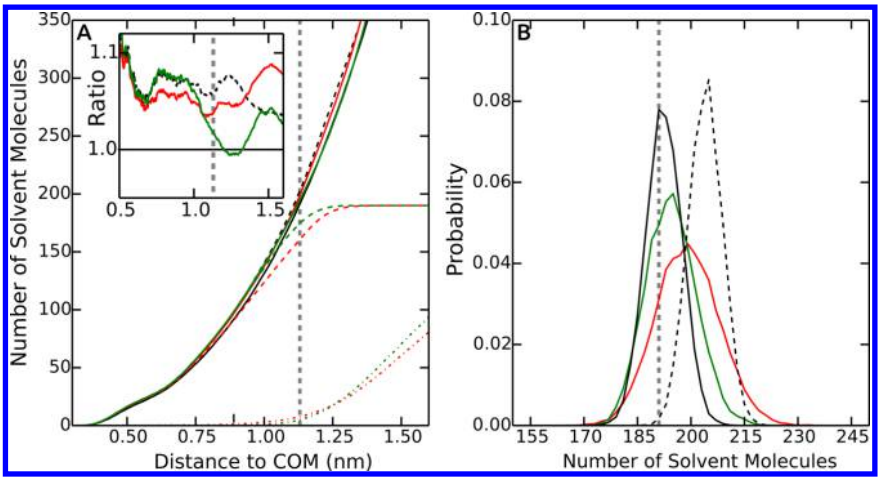


Figure 5. (A) Number of solvent molecules as a function of distance to the COM of the Val side chain located at the center of the 1.0 nm SPC sphere. Fully atomistic reference simulation in SPC water (black line), fully atomistic SPC water with 190 distance restrained SPC molecules (dashed black line), and hybrid AA−CG with 1 nm SPC sphere in BDW (red line) or in stdW (green line). Dashed red and dashed green lines: SPC molecules from SPC sphere in BDW and in stdW, respectively. Dotted red and dotted green lines: BDW and stdW CG water molecules, respectively. Single BDW and stdW molecules were counted as 5 and 4 individual solvent molecules, respectively, according to the AA-to-CG mapping. The vertical dashed gray line indicates r_{SPC} (1.13 nm in this case). Inset: Ratio of solvent molecules with respect to the unrestrained fully atomistic reference simulation in SPC. Dashed black line: Fully atomistic simulation with 190 distance-restrained SPC molecules. Red line: SPC sphere in BDW. Green line: SPC sphere in stdW. (B) Probability distribution of the number of water molecules within r_{SPC} = 1.13 nm of Val. The color scheme is as above. Dotted black line: Simulation at a constant 6% higher density. The vertical dashed gray line indicates the number of SPC molecules ($\langle N \rangle$ = 190) used in the hybrid AA−CG simulations.

and −34 189 kJ mol^{-1}, respectively). Similarly small distance restraint energies per SPC molecule were found for all investigated systems (see Supporting Information). However, if these energies are hardly sufficient to fully explain the observed density and pressure increase within the atomistic water spheres, then which additional factors are at play? The answer to this question is different for polarizable (BDW or polW) and nonpolarizable (stdW) water. BDW mixes with SPC.[21] Thus, BDW water molecules penetrate the SPC sphere; the same is observed for polW. Since the distance restraints energetically penalize SPC water molecules that leave the sphere, this leads to the observed density increase. This penetration of CG water will always occur, irrespective of the size of the SPC sphere. Thus, adding more atomistic water layers does not lead to an improvement. The same behavior is also observed in our reference simulation of a fully atomistic system with 190 SPC molecules affected by distance restraints (see above). By contrast, stdW Martini CG water is not miscible with SPC and hence does not enter the atomistic subsystem. Nevertheless, too high of a density is also observed in this case. Here, the density increase is due to the surface tension of the AA−CG interface, which leads to a pressure inside the sphere that can be described by the Young−Laplace equation, $\Delta p = 2\gamma/r$. As described below, we determined the surface tension of the SPC/stdW boundary as γ = 73 mN m^{-1}, which yields a pressure difference of 1500 bar for the spheres with 1.0 nm radius. Although, for such highly curved interfaces, corrections to the Young−Laplace equation for macroscopic systems have to be applied,[52] this rough estimate is in agreement with the pressure of 1475 bar observed in the NVT simulation of SPC at 6% too high density (see above). Likewise, for the SPC sphere in vacuum, the SPC/vacuum surface tension of γ = 53 mN m^{-1} causes a similar pressure difference. The surface tension of the SPC/vacuum interface is slightly lower than that for the SPC/stdW interface because, due to the depletion of water at the interface (see below), the latter resembles two interfaces, SPC/vacuum and vacuum/stdW, whose surface tensions partially add up. For both the SPC/vacuum and SPC/stdW interfaces, the solvation free energies are systematically improvable by increasing the size of the atomistic solvent sphere (Table 3), since the Laplace pressure decreases with increasing radius.

3.4. Possible Remedies. Given that the distance restraints lead to the density increase in the inner shell, one obvious way to improve the situation is to lower the distance restraint force constant. We carried out additional TI simulations of the Val side chain in a 1.0 nm SPC sphere surrounded by BDW with k_{dr} = 30 and 3 kJ mol^{-1} nm^{-2} (rather than 300 kJ mol^{-1} nm^{-2}), yielding ΔG_{solv} = +12.3 and +8.8 kJ mol^{-1}, respectively. Hence, the weaker the distance restraints, the closer the solvation free energy is to the reference value (+8.3 kJ mol^{-1} for Val), as expected. However, too low force constants are not sufficient to keep the restrained water molecules in the vicinity of the solute throughout the simulations and thus we do not consider them to be a suitable solution.

Instead, we decided to slightly increase the onset distance of the half-harmonic restraining potential energy functions but not the number of restrained atomistic water molecules. Thus, the onset distance for the restraints is now $r'_{SPC} = r_{SPC} + 0.3$ nm. In other words, the distance restraints now form an oversized suit for the atomistic water molecules. Table 3 and Figure 6 show that, indeed, this consistently improves ΔG_{solv}. For the BDW model, the MUE is significantly reduced from 7.6 to 3.5 kJ mol^{-1} and from 6.5 to 2.3 kJ mol^{-1} for the 0.75 and 1.0 nm SPC spheres, respectively. The respective improvements for stdW are similar, from 6.8 to 2.6 kJ mol^{-1} and from 5.3 to 3.7 kJ mol^{-1}. Thus, the MUE is now almost comparable to that of the fully atomistic reference simulations (1.6 kJ mol^{-1}). Interestingly, for the stdW model, the minimal MUE is not found for the large 1.0 nm SPC sphere but for the 0.75 nm one, suggesting that compensating effects are at play, as discussed in the following.

DOI: 10.1021/acs.jctc.5b00499
J. Chem. Theory Comput. 2015, 11, 4460−4472

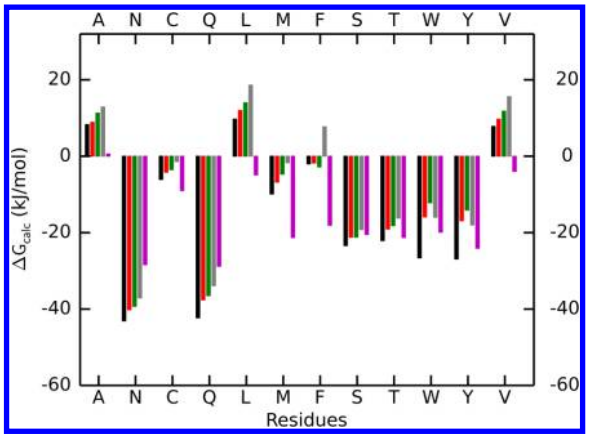


Figure 6. Free energy of solvation of atomistic amino acid side chain analogues from simulations in pure SPC (black) or in pure BDW CG water (magenta) and from hybrid AA−CG simulations of 1.0 nm SPC spheres surrounded by BDW (red), stdW (green), or vacuum (gray). The onset distance for the restraints was increased to $r'_{SPC} = r_{SPC} + 0.3$ nm in these simulations.

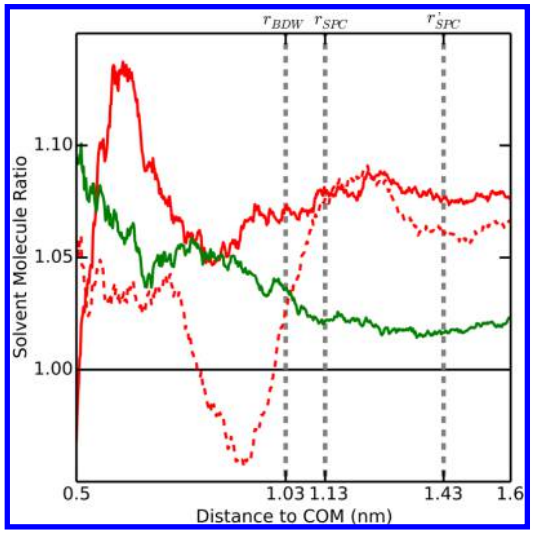


Figure 7. Ratio of solvent molecules with respect the unrestrained fully atomistic reference simulation in pure SPC. Colors as in Figure 5, inset (red, SPC sphere in BDW; green, SPC sphere in stdW). These data were obtained from the simulations of a Val side chain in a 1.0 nm SPC sphere, with an increased onset of the distance restraints on SPC ($r'_{SPC} = r_{SPC} + 0.3$ nm, i.e., from 1.13 to 1.43 nm for Val). Dashed red line: Additional repulsive restraints on BDW water for distances less than $r_{BDW} = r_{SPC} - 0.1$ nm.

Figure 7 shows that in these new simulations the density in the inner solvent shell is also too high and hence cannot account for the observed improvement. Instead, due to the outward-shifted distance restraints, a more pronounced penetration of CG water molecules into the inner shell is observed. As a consequence, the direct contribution of the CG solvent to ΔG_{solv} is increased, and the resulting solvation free energies are thus closer to the respective values in pure CG solvent. As can be seen from ΔG_{solv} in pure BDW (Table 3), this leads to a better overall agreement with experiment due to compensating effects. We note that these improved solvation free energies do not necessarily imply reasonable PMFs between ionic species, as our previous AA−CG simulations in pure CG solvent have shown.[23]

Next, we investigated whether additional, repulsive distance restraints on the BDW water molecules could help to more cleanly separate the AA from the CG solvent regions and possibly diminish the density variations. To that end, we applied repulsive half-harmonic distance restraints on the BDW molecules beyond $r_{BDW} = r_{SPC} - 0.1$ nm, i.e., we introduced a mixing (overlap) zone of width $r'_{SPC} - r_{BDW} = 0.4$ nm. As can be seen from the density profiles in Figure 7 (dashed red line), the average density in the AA−CG interface region is indeed lower in this case. However, the MUE in ΔG_{solv} increases from 2.3 to 3.2 kJ mol^{-1} because the CG solvent and density related effects do not, in this case, compensate to a similar extent as previously.

3.5. Characterization of the Interface. Our above results show that, even with large atomistic solvent regions, the solvation of the solutes is still affected by the surrounding CG solvent. Although this can be partly attributed to the described density mismatch, long-range effects due to orientational ordering of SPC water molecules at the AA−CG boundary may play an additional role. These arise from the absence of electrostatic interactions that the atomistic water molecules would normally have with other atomistic waters outside the sphere. Since the polarizable CG water models explicitly take electrostatic interactions with the SPC water molecules into account, intuitively, this latter effect could be expected to be less pronounced for the BDW and polW models than for stdW or vacuum surrounding.

Figure 8 shows the mass density and E-field profiles across the different interfaces. As mentioned above, stdW and SPC effectively repel each other, causing a depletion at the interface (Figure 8a, green curve). The density modulations within the bulk stdW CG region result from a long-range perturbation of the packing of the CG beads due to the presence of the interface, an overstructuring that is well-known also for bulk

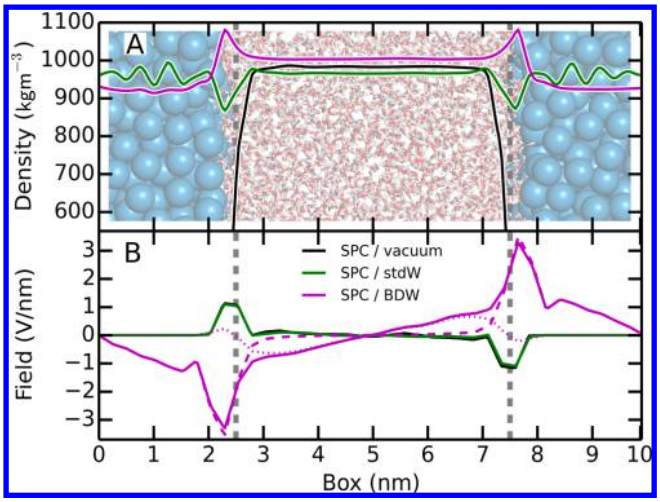


Figure 8. Density (top) and electrostatic field (bottom) profile across the interface (indicated by dashed gray lines). The SPC/vacuum interface is shown as black line, SPC/stdW, green, and SPC/BDW, magenta. The dotted and dashed magenta lines show the contributions of the SPC and BDW molecules, respectively, to the E-field. The density profile was calculated by dividing the simulation box along the z-axis into slabs of thickness 0.025 nm. The E-field was calculated by first summing the charges in the slabs and then integrating the charge distribution along the z-axis. Very similar results were obtained for the SPC/vacuum and SPC/stdW interfaces in the presence and absence of a boundary potential (see Supporting Information).

DOI: 10.1021/acs.jctc.5b00499
J. Chem. Theory Comput. 2015, 11, 4460−4472

stdW Martini CG water.[18] At the SPC/stdW interface, the density dip extends over an interface width of about 0.5 nm. By contrast, BDW interacts favorably with SPC and accumulates at the AA−CG boundary (Figure 8a, magenta curve). In addition, a significant fraction of the BDW molecules penetrate the atomistic subsystem, thus shifting the density to ca. 1030 kg m^{-3}. Since these interface simulations are carried out in the NVT ensemble (see Methods), this inevitably leads to a density decrease in the CG subsystem. Qualitatively similar results were obtained for the SPC/polW interface (see Supporting Information). The observed local density perturbations near the planar interface match those in the spherical setups described above (Figure 5). However, unlike in the simulations with an AA sphere, for the planar interfaces there is no Laplace pressure. Hence, for stdW CG water, the density in the atomistic subsystem corresponds to the density of SPC. This is not the case for polarizable CG water, which enters the atomistic subsystem irrespective of the curvature of the interface.

In addition to the density profiles, we analyzed the E-field across the simulation box (Figure 8b), which arises from preferred orientations of water dipoles at the interface. At the SPC/stdW and SPC/vacuum interfaces, orientational ordering of the SPC water molecules causes very similar E-fields of 1 V nm^{-1}, which decay over a distance range of about 0.5 nm. This is in agreement with our above results and underlines the similarity of the SPC/vacuum and SPC/stdW interfaces. In both cases, the SPC molecules are oriented such that O−H bonds are, on average, slightly more likely to point away from the bulk and toward the interface. This orientational preference is well-known from simulations[53] and spectroscopy of the water/air interface.[54]

In striking difference to the SPC/stdW and SPC/vacuum interfaces, the SPC/BDW interface produces a very strong E-field (3 V nm^{-1}) of opposite sign, which has a long-range component that extends over the entire simulation box (magenta line in Figure 8b). In contrast to the above interfaces, where only SPC molecules can contribute to the E-field, here both CG and SPC dipoles are altered due to the presence of the dividing surface. The individual contributions of the SPC and BDW molecules to the overall E-field are plotted as dotted and dashed magenta lines, respectively, showing that the presence of BDW molecules in the bulk SPC phase leads to a long-range orientational ordering of the SPC molecules. A similar effect is seen for the SPC/polW interface, although the sign of the E-field is reversed in this case (see Supporting Information). A complete characterization of the orientational preferences of AA and CG solvent molecules is beyond the scope of the present study, however. In summary, it is clear that both density modulations and orientational ordering at the interface affect the solvation properties.

3.6. Computational Efficiency. Table 4 quantifies the computational performance of the different models. Even for a large 1.0 nm SPC sphere, the stdW model provides a speed-up of a factor 10.6 with respect to the same simulation system treated at the fully atomistic level. For the BDW model, only a 2.2-fold difference is seen, owing to the rather long 2.0 nm cutoff used for the nonbonded interactions. Using a 2.0 nm cutoff also for the fully AA reference simulations would, of course, improve this relative speed-up; however, since such large cutoffs are usually not considered to be necessary in full AA simulations, we compare this to an AA simulation with a 1.4 nm cutoff (Table 4). Reducing the number of AA solvent layers reduces computer time requirements, with a difference between the large (1.0 nm) and small (0.5 nm) SPC spheres of about 50%. For interpreting the computational performance given in Table 4, it has to be kept in mind that we deliberately used rather large simulation boxes (edge length about 5.0 nm) to avoid any possible finite-size effects. Thus, the simulation box is much larger than one would normally choose for a small solute in water and contains a relatively large number of CG solvent molecules. When comparing to the same (large) simulation system described at the fully atomistic level, the speed-up of the hybrid AA−CG simulations given in Table 4 may thus be larger than in more realistic situations. To illustrate this point, we have carried out an additional AA−CG simulation of a 1.0 nm SPC sphere embedded in only 400 (instead of ca. 1000) stdW CG water beads, which yielded a smaller simulation box of 3.8 nm. With this smaller system, a 6.1-fold speed-up (instead of 10.6, see above) was observed, resulting from absolute performances of 28.4 ns/day at the fully AA and 173.0 ns/day at the hybrid AA−CG level. The computational efficiency of the hybrid AA−CG simulations could be further improved by using different time steps for integrating the equations of motion of the AA and CG particles because the CG degrees of freedom need to be updated up to 10 time less frequently than the AA ones. This would speed up the simulations, since the CG subsystem still dominates the computational costs. This is evident from the substantial gain that can be obtained by removing the CG particles altogether, i.e., using a vacuum surrounding (Table 4). Depending on the ratio of AA and CG particles, we expect that using larger time steps for the CG subsystem could further boost computational efficiency by another factor of 3−8.

Table 4. Computational Efficiency[a]

	abs. performance (ns/day)			speed-up factor		
	stdW	BDW	vacuum	stdW	BDW	vacuum
1.0 nm SPC sphere	110.9	22.6	430.4	10.6	2.2	41.4
0.75 nm SPC sphere	132.1	31.9	1171.0	12.7	3.1	112.6
0.5 nm SPC sphere	168.0	34.1	4102.7	16.1	3.3	394.5
CG		42.8			4.1	

[a]All values were obtained for the Ser side chain system on a single i7 3.4 GHz CPU core. The speed-up factor refers to a fully atomistic simulation of the same system with the same settings for the nonbonded interactions (i.e., 1.4 nm nonbonded cutoff, reaction field), with which the absolute performance is 10.4 ns/day.

4. SUMMARY AND CONCLUSIONS

In this work, we have critically evaluated the accuracy of different fixed resolution hybrid AA−CG simulations in terms of Gibbs free energies of solvation of atomistic amino acid side chain analogues (i) in polarizable coarse-grained water and (ii) solvated in a sphere of atomistic SPC water, which, in turn, was embedded in a CG water environment. For the latter, we compared polarizable CG water models (BDW and polW) with the computationally more efficient standard Martini water (stdW) model, which is a pure Lennard-Jones fluid without partial charges. In addition, we investigated the effect of completely neglecting the surrounding, i.e., an SPC sphere in vacuum. Our results show that the free energies of solvation from the AA−CG simulations of AA solutes solvated in pure BDW CG water deviate from experiment by a MUE in ΔG_{solv} of 8.2 kJ mol^{-1}. Adding atomic-level solvent shells leads to an improvement. For the stdW Martini CG water surrounding, the

DOI: 10.1021/acs.jctc.5b00499
J. Chem. Theory Comput. 2015, 11, 4460−4472

error with respect to experiment can be systematically decreased by increasing the size of the atomistic solvent sphere, down to a MUE in ΔG_{solv} of 2.6 kJ/mol for a 0.75 nm SPC sphere with slightly outward-shifted onset of the distance restraints on SPC. Completely neglecting the effect of the surrounding, as in the simulations of SPC spheres in vacuum, did not yield satisfactory results. The systematic improvement with increasing the size of the atomistic solvent region is not observed for the polarizable BDW water because, unlike stdW, it is miscible with SPC and thus enters the atomistic solvent sphere. Since the SPC molecules cannot diffuse out of the sphere due to the distance restraints, this penetration cannot be compensated and leads to too high of a density inside the inner shell (by about 6%). This effect is independent of the size of the AA solvent region. Nevertheless, for large SPC spheres in combination with outward-shifted distance restraints, the solvation free energies obtained with the AA−CG hybrid setup yield a MUE of only 2.3 kJ/mol. For the standard Martini CG water and the vacuum surrounding, the density inside the SPC sphere is also too high, but for a different reason. Here, the surface tension at the curved interface causes a substantial pressure difference (of about 1 kbar for the 1.0 nm spheres), which also increases the density by about 6%. In addition, irrespective of the CG water model used, a long-range effect due to orientational ordering of water molecules at the AA−CG interface is observed, affecting solvation properties. In the present study, we computed solvation free energies to quantify the influence of the AA−CG boundary and pinpoint possible pitfalls. However, we expect our results to be relevant also for the investigation of structural, dynamic, and thermodynamic properties of larger (biomolecular) solutes in similar AA−CG schemes as well as for related triple-scale QM/MM/CG approaches.[55,56] In this context, we would also like to note in passing that processes in which the AA subsystem of interest significantly changes volume, such as a protein that undergoes large-scale conformational transitions, e.g., upon folding, pose a tough challenge for hybrid AA−CG schemes with a fixed number of AA solvent molecules (or likewise, for adaptive schemes, a fixed volume of the AA region). For studying such cases, it would be beneficial to vary the number of AA solvent molecules (or volume of the AA region) adaptively as the simulation progresses.

Multiscale modeling always involves a trade-off between efficiency and accuracy. Thus, for judging the practical use of a particular model, it should be kept in mind that computational speed-up is one of the main motivations for multiscale modeling. Compared to adaptive resolution AA−CG methods, fixed resolution partitioning methods have, on one hand, the potential to be computationally more efficient because they do not require calculating AA interactions and forces also in the switching (or healing) region. On the other hand, the abrupt AA−CG boundary poses additional tough challenges linked to the thermodynamic mismatch between the AA and CG levels of description. In adaptive resolution schemes, this mismatch is balanced by a thermodynamic force, which smoothens out the density profile across the AA−CG interface. In addition, a sufficiently large healing region ensures that artifacts due to orientational ordering of solvent molecules at the AA−CG boundary are negligible inside the inner, fully atomistic region. Finally, by construction, penetration of CG solvent into the inner region is not possible in adaptive schemes.

Although, in this study, we largely restricted ourselves to the SPC/stdW hybrid system (for a nonpolarizable CG model) and the SPC/BDW model of Riniker and co-workers[21,22,27] (for a polarizable CG model) and did not test other CG water models[57] apart from the polarizable Martini (polW) model, qualitatively similar effects are expected for other CG models as well. Using a vacuum surrounding around the atomistic solvent region provides the most pronounced computational speed-up, but very large atomic-level solvent spheres are required. However, even for these sphere sizes, a speed-up of more than a factor 40 was achieved (Table 4). Nevertheless, this finding may provoke the question of why using a CG surrounding should be beneficial at all. Arguments in favor of a hybrid model with explicit AA and CG particles are, first, that energy flow across the interface through collisions between AA and CG particles thermalize the molecules at the AA−CG interface. Second, compressibility is, at least in principle, accounted for. Nevertheless, preferred orientation of water molecules and density deviations at the interface are observed. To improve on the latter issue, stochastic boundary potential methods have been introduced[48,58] that describe the missing surrounding through mean field boundary forces, also incorporating thermal fluctuations. In addition, methods to bridge explicit and implicit (e.g., of generalized Born or Poisson−Boltzmann type) descriptions of the solvent[59−61] as well as approaches that account for orientational polarization of solvent molecules at the interface[62,63] have been devised. If an explicit treatment of the surrounding solvent is desired while maintaining a fixed, predefined partitioning of the simulation system into regions of different resolution, additional potential energy terms to separate the AA and CG solvent ensembles may be hard to avoid. However, as the present study shows, half-harmonic distance restraints between the AA solvent and the solute can lead to undesirable edge effects. A more suitable alternative could be to apply the additional potential energy terms in such a way that they do not prevent the AA solvent from diffusing outward into the CG zone but *vice versa* prevent the CG solvent from entering the AA region. Defining the restraints with respect to the outermost AA solvent molecule, and not relative to the solute, could be beneficial because the position of the outermost AA solvent molecule fluctuates during the simulation, allowing the density of the inner region to adjust. One such method is the flexible inner region ensemble separator approach.[64] This method has been validated and compared to adaptive resolution methods in the context of QM/MM simulations,[65] but it remains to be tested in hybrid AA−CG simulations with a large number of solvent molecules in the inner shell.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.5b00499.

Simulation systems and free energies of solvation (Table S1); enthalpies and entropies of solvation of AA solutes in pure BDW CG water (Table S2); free energies of solvation in polarizable Martini water (Table S3); density profile across the AA−CG interface in the presence and absence of a boundary potential (Figure S1); corresponding density profile for a SPC/polW interface (Figure S2) (PDF).

DOI: 10.1021/acs.jctc.5b00499
J. Chem. Theory Comput. 2015, 11, 4460−4472

AUTHOR INFORMATION

Corresponding Author

*E-mail: lars.schaefer@ruhr-uni-bochum.de. Phone: +49 234 3221582. Fax: +49 234 3214045.

Funding

The Deutsche Forschungsgemeinschaft (DFG) supported this work (Cluster of Excellence RESOLV (EXC 1069), Emmy Noether grant to L.S.).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Sereina Riniker for insightful discussions. The Research Department Interfacial Systems Chemistry (IFSC) of the Ruhr-University Bochum is acknowledged.

REFERENCES

(1) *Coarse-Graining of Condensed Phase and Biomolecular Systems*; Voth, G. A., Ed.; CRC Press/Taylor and Francis: Boca Raton, FL, 2009.
(2) Noid, W. G. *J. Chem. Phys.* **2013**, *139*, 090901.
(3) Ingólfsson, H. I.; Lopez, C. A.; Uusitalo, J. J.; de Jong, D. H.; Gopal, S. M.; Periole, X.; Marrink, S. J. *WIRES Comput. Mol. Sci.* **2014**, *4*, 225−248.
(4) Praprotnik, M.; Delle Site, L.; Kremer, K. *J. Chem. Phys.* **2005**, *123*, 224106.
(5) Abrams, C. F. *J. Chem. Phys.* **2005**, *123*, 234101.
(6) Ensing, B.; Nielsen, S. O.; Moore, P. B.; Klein, M. L.; Parrinello, M. *J. Chem. Theory Comput.* **2007**, *3*, 1100−1105.
(7) Praprotnik, M.; Delle Site, L.; Kremer, K. *Annu. Rev. Phys. Chem.* **2008**, *59*, 545−571.
(8) Heyden, A.; Truhlar, D. G. *J. Chem. Theory Comput.* **2008**, *4*, 217−221.
(9) Kreis, K.; Fogarty, A. C.; Kremer, K.; Potestio, R. *Eur. Phys. J.: Spec. Top.* **2015**, 1−16.
(10) Shi, Q.; Izvekov, S.; Voth, G. A. *J. Phys. Chem. B* **2006**, *110*, 15045−15048.
(11) Michel, J.; Orsi, M.; Essex, J. W. *J. Phys. Chem. B* **2008**, *112*, 657−660.
(12) Orsi, M.; Sanderson, W. E.; Essex, J. W. *J. Phys. Chem. B* **2009**, *113*, 12019−12029.
(13) Orsi, M.; Noro, M. G.; Essex, J. W. *J. R. Soc., Interface* **2011**, *8*, 826−841.
(14) Orsi, M.; Ding, W.; Palaiokostas, M. *J. Chem. Theory Comput.* **2014**, *10*, 4684−4693.
(15) Orsi, M.; Essex, J. W. *PLoS One* **2011**, *6*, e28637.
(16) Orsi, M. *Mol. Phys.* **2014**, *112*, 1566−1576.
(17) Rzepiela, A. J.; Louhivuori, M.; Peter, C.; Marrink, S. J. *Phys. Chem. Chem. Phys.* **2011**, *13*, 10437−10448.
(18) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. *J. Phys. Chem. B* **2007**, *111*, 7812−7824.
(19) Han, W.; Wan, C.-K.; Wu, Y.-D. *J. Chem. Theory Comput.* **2008**, *4*, 1891−1901.
(20) Han, W.; Wan, C.-K.; Jiang, F.; Wu, Y.-D. *J. Chem. Theory Comput.* **2010**, *6*, 3373−3389.
(21) Riniker, S.; van Gunsteren, W. F. *J. Chem. Phys.* **2012**, *137*, 044120.
(22) Riniker, S.; Eichenberger, A. P.; van Gunsteren, W. F. *Eur. Biophys. J.* **2012**, *41*, 647−661.
(23) Wassenaar, T. A.; Ingólfsson, H. I.; Prieß, M.; Marrink, S. J.; Schäfer, L. V. *J. Phys. Chem. B* **2013**, *117*, 3516−3530.
(24) Gonzalez, H. C.; Darré, L.; Pantano, S. *J. Phys. Chem. B* **2013**, *117*, 14438−14448.
(25) Riniker, S.; van Gunsteren, W. F. *J. Chem. Phys.* **2011**, *134*, 084110.
(26) Yesylevskyy, S. O.; Schäfer, L. V.; Sengupta, D.; Marrink, S. J. *PLoS Comput. Biol.* **2010**, *6*, e1000810.
(27) Riniker, S.; Eichenberger, A. P.; van Gunsteren, W. F. *J. Phys. Chem. B* **2012**, *116*, 8873−8879.
(28) Lin, Z.; Riniker, S.; van Gunsteren, W. F. *J. Chem. Theory Comput.* **2013**, *9*, 1328−1333.
(29) Sokkar, P.; Choi, S. M.; Rhee, Y. M. *J. Chem. Theory Comput.* **2013**, *9*, 3728−3739.
(30) Huang, W.; Riniker, S.; van Gunsteren, W. F. *J. Chem. Theory Comput.* **2014**, *10*, 2213−2223.
(31) Zavadlav, J.; Melo, M. N.; Marrink, S. J.; Praprotnik, M. *J. Chem. Phys.* **2014**, *140*, 054114.
(32) Zavadlav, J.; Melo, M. N.; Cunha, A. V.; de Vries, A. H.; Marrink, S. J.; Praprotnik, M. *J. Chem. Theory Comput.* **2014**, *10*, 2591−2598.
(33) Zavadlav, J.; Melo, M. N.; Marrink, S. J.; Praprotnik, M. *J. Chem. Phys.* **2015**, *142*, 244118.
(34) Fuhrmans, M.; Sanders, B. P.; Marrink, S. J.; de Vries, A. H. *Theor. Chem. Acc.* **2010**, *125*, 335−344.
(35) Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **2004**, *25*, 1656−1676.
(36) Gopal, S. M.; Kuhn, A. B.; Schäfer, L. V. *Phys. Chem. Chem. Phys.* **2015**, *17*, 8393−8406.
(37) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(38) van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Simul.* **1988**, *1*, 173−185.
(39) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(40) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *J. Comput. Chem.* **1997**, *18*, 1463−1472.
(41) Hess, B. *J. Chem. Theory Comput.* **2008**, *4*, 116−122.
(42) Miyamoto, S.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 952−962.
(43) Hess, B.; van der Vegt, N. F. A. *J. Phys. Chem. B* **2006**, *110*, 17616−17626.
(44) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1994**, *222*, 529−539.
(45) Hess, B. *J. Chem. Phys.* **2002**, *116*, 209−217.
(46) Risselada, H. J.; Mark, A. E.; Marrink, S. J. *J. Phys. Chem. B* **2008**, *112*, 7438−7447.
(47) Louhivuori, M.; Risselada, H. J.; van der Giessen, E.; Marrink, S. J. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 19856−19860.
(48) Brooks, C. L.; Karplus, M. *J. Chem. Phys.* **1983**, *79*, 6312.
(49) Wolfenden, R.; Andersson, L.; Cullis, P. M.; Southgate, C. C. *Biochemistry* **1981**, *20*, 849−855.
(50) Wu, Z.; Cui, Q.; Yethiraj, A. *J. Chem. Theory Comput.* **2011**, *7*, 3793−3802.
(51) Makhatadze, G.; Privalov, P. L. *J. Mol. Biol.* **1993**, *232*, 639−659.
(52) Malijevský, A.; Jackson, G. *J. Phys.: Condens. Matter* **2012**, *24*, 464121.
(53) Sokhan, V. P.; Tildesley, D. J. *Mol. Phys.* **1997**, *92*, 625−640.
(54) Verreault, D.; Hua, W.; Allen, H. C. *J. Phys. Chem. Lett.* **2012**, *3*, 3012−3028.
(55) Meier, K.; Choutko, A.; Dolenc, J.; Eichenberger, A. P.; Riniker, S.; van Gunsteren, W. F. *Angew. Chem., Int. Ed.* **2013**, *52*, 2820−2834.
(56) Sokkar, P.; Boulanger, E.; Thiel, W.; Sanchez-Garcia, E. *J. Chem. Theory Comput.* **2015**, *11*, 1809−1818.
(57) Hadley, K. R.; McCabe, C. *Mol. Simul.* **2012**, *38*, 671−681.
(58) Berkowitz, M.; McCammon, J. A. *Chem. Phys. Lett.* **1982**, *90*, 215−217.
(59) Lin, J.-H.; Baker, N. A.; McCammon, J. A. *Biophys. J.* **2002**, *83*, 1374−1379.
(60) Lee, M. S.; Salsbury, F. R., Jr; Olson, M. A. *J. Comput. Chem.* **2004**, *25*, 1967−1978.
(61) Brancato, G.; Rega, N.; Barone, V. *J. Chem. Phys.* **2008**, *128*, 144501.
(62) King, G.; Warshel, A. *J. Chem. Phys.* **1989**, *91*, 3647−3661.
(63) Beglov, D.; Roux, B. *J. Chem. Phys.* **1994**, *100*, 9050−9063.

DOI: 10.1021/acs.jctc.5b00499
J. Chem. Theory Comput. 2015, 11, 4460−4472

(64) Rowley, C. N.; Roux, B. *J. Chem. Theory Comput.* **2012**, *8*, 3526−3535.
(65) Bulo, R. E.; Michel, C.; Fleurat-Lessard, P.; Sautet, P. *J. Chem. Theory Comput.* **2013**, *9*, 5567−5577.

DOI: 10.1021/acs.jctc.5b00499
J. Chem. Theory Comput. 2015, 11, 4460−4472