JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Small Molecule Solvation Free Energy: Enhanced Conformational Sampling Using Expanded Ensemble Molecular Dynamics Simulation

Andrew S. Paluch,[†] David L. Mobley,[‡] and Edward J. Maginn*[,†]

[†]Department of Chemical and Biomolecular Engineering, University of Notre Dame, Notre Dame, Indiana 46556, United States

[‡]Department of Chemistry, University of New Orleans, New Orleans, Louisiana 70148, United States

Supporting Information

ABSTRACT: We present an efficient expanded ensemble molecular dynamics method to calculate the solvation free energy (or residual chemical potential) of small molecules with complex topologies. The methodology is validated by computing the solvation free energy of ibuprofen in water, methanol, and ethanol at 300 K and 1 bar and comparing to reference simulation results using Bennett's acceptance ratio method. Difficulties with ibuprofen using conventional molecular dynamics methods stem from an inadequate sampling of the carboxylic acid functional group, which, for the present study, is subject to free energy barriers of rotation of 14−20 k_BT. While several advances have been made to overcome such weaknesses, we demonstrate how this shortcoming is easily overcome by using an expanded ensemble methodology to facilitate conformational sampling. Not only does the method enhance conformational sampling but it also boosts the rate of exploration of the configurational phase space and requires only a single simulation to calculate the solvation free energy. Agreement between the expanded ensemble and the reference calculations is good for all three solvents, with the reported uncertainties of the expanded ensemble being comparable to the uncertainties of the reference calculations, while requiring less simulation time; the reduced simulation time demonstrates the improved performance gained from the expanded ensemble method.

1. INTRODUCTION AND MOTIVATION

Knowledge of the underlying free energy (or chemical potential) of biologically active species is key to understanding their thermodynamic phase behavior and is crucial for rational drug design.[1] For example, it is well-known that the solubility of a drug is dictated by its chemical potential in a pure crystalline phase relative to the solution phase. Additionally, the ability of a drug to partition between cell membranes is governed by the relative chemical potential[2,3] and its binding with proteins by the relative free energy.[4,5] The crucial role that free energy plays in the drug discovery process is emphasized by the fact that entire monographs have been devoted to the topic.[6,7] Likewise, insight into the native structure and folding mechanism of proteins in solution may be obtained by examining the solvation free energy of individual constituent amino acid analogs.[8,9]

The free energy of a given species may be understood in terms of the fundamental molecular level details via molecular simulation. Several recent studies have highlighted the ability of molecular simulations to precisely and accurately compute the solvation free energy of amino acid analogs[10−12] and small drug and drug-like molecules.[13−15] However, the work of Hodel et al.[16] warned that even for small polypeptides, inadequate conformational sampling with molecular dynamics (MD) could lead to errors in free energy calculations in much greater excess of any other source of error. As the size and complexity of the molecule increases, the problem becomes even worse. This warning was reiterated in the recent work of Leitgeb et al.[17] for amino acids and by Klimovich and Mobley[15] for the case of small drug molecules.

It follows that extreme care must be taken to ensure that systems studied using molecular simulation are ergodic.[18,19] For complex systems, it is common that ergodicity is "broken" over the time scale of the simulation.[20,21] That is, the important regions of the configurational phase space may be separated by large free energy barriers which are not overcome during the course of conventional simulations, yielding erroneous results. While this challenge can be overcome with advanced Monte Carlo (MC) methods,[22] it is often difficult to formulate a general protocol to ensure adequate sampling by MD. Many highly efficient MD algorithms have been developed to enhance the sampling of configurational phase space;[23−30] these methods may be applied to enhance sampling when performing free energy calculations.[17,25,31,32] However, given the diversity of methods available, proper method selection poses a huge challenge when computational efficiency is important.

Recently, Paluch et al.[33] demonstrated the use of expanded ensemble (EE)[34−36] methods to increase the rate of exploration of the important regions of configurational phase space when performing free energy calculations, when MD is used to sample configurational phase space within each subensemble. Not only does EE improve the rate of exploration of configurational phase space, but the method allows for the calculation of the free energy in an efficient manner, requiring the use of a *single* simulation. Moreover, the method may be readily implemented in existing MD simulation software. The present study extends our previous work[33] by showing how EE may additionally be used to seamlessly increase the rate of conformational sampling of molecules of complex topologies within an MD framework.

Received: June 6, 2011
Published: July 27, 2011

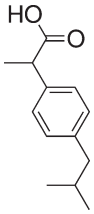

Figure 1. The chemical structure of ibuprofen.

Klimovich and Mobley[15] illustrated the inability of conventional MD-based free energy calculations to observe conformational changes of the H−O−C−O dihedral angle of carboxylic acid functional groups of small drug molecules in water and the impact this had on the computed free energy. In the present study, we have repeated the calculations of Klimovich and Mobley[15] for ibuprofen (Figure 1), which contains a carboxylic acid functional group, in water, methanol, and ethanol using MD-based EE. We will demonstrate how this technique overcomes H−O−C−O barriers of rotation of approximately 14−20 k_BT unhindered. Furthermore, akin to replica exchange (or parallel tempering) simulations,[26] ensemble average configurational properties may be calculated within each subensemble, allowing the use of EE whenever advanced conformational sampling is required. We anticipate that the technique presented in the present study, combined with previous free energy methodologies,[2,4,37,38] will aid in the rational drug design process.[1] In Section 2 we will present an overview of the EE methodology, followed by the relevant computational details in Section 3. Results and discussion are given in Section 4, followed by a summary of our findings in Section 5.

2. METHODOLOGY

An EE procedure was used to compute the free energy of solvation. The method is described in detail elsewhere,[33] so only a brief summary of the essential concepts and modification from our previous work is provided here. The basic idea behind the EE method[34−36] is to construct an augmented ensemble as a sum of $M_{\text{Total}} + 1$ subensembles. This series of subensembles connects two systems of interest by gradually performing transitions between the two systems. In the present work, the systems of interest are a noninteracting solute (ibuprofen) molecule in a pure solvent (i.e., the solute is in an ideal gas state) and a single fully interacting solute molecule in the solvent, with both states at the same temperature and pressure. The free energy difference between these two systems gives the free energy of solvation. Typically, the intermediate EE subensembles between the noninteracting and fully interacting solute subensembles serve to scale only the *intermolecular* interaction potential of the solute.[33,38−41] To improve the conformational sampling of the solute, we have additionally introduced intermediate subensembles to scale the *intramolecular* interaction potential of the solute; this addition is similar to the torsion angle potential fluctuation method of Liu and Berne[24] and the simulated scaling method of Hongzhi et al.[25] A specific subensemble is designated by index m, where the solute intermolecular Lennard-Jones (LJ) and the electrostatic interactions are regulated by the subensemble-dependent coupling parameters λ_m^{LJ} and λ_m^{elec}, respectively, and the intramolecular torsional potential of the solute is regulated by λ_m^{tors}. These coupling parameters vary from $0 \le \lambda_m^{\text{LJ}} \le 1$, $0 \le \lambda_m^{\text{elec}} \le 1$, and, in the present study, $0.1 \le \lambda_m^{\text{tors}} \le 1$.

While within a given subensemble, configurational phase space is sampled using MD within an isothermal−isobaric (NPT) ensemble. Periodically, a stochastic transition to an adjacent subensemble is attempted. The transitions are accepted using an appropriate acceptance rule.[33] However, as the free energy between subensembles increases, the probability of accepting a move decreases exponentially. This adversity is overcome by employing a biasing scheme that utilizes a combined Wang−Landau (WL)[42−44] and Bennett's acceptance ratio (BAR)[45−48] method.[33] The difference in free energy between the two end states and hence the free energy of solvation is calculated using BAR. Complete details can be found elsewhere.[33]

3. COMPUTATIONAL DETAILS

Molecular Models. For the solvents and solute (ibuprofen), nonbonded intermolecular interactions were treated using a combined LJ and fixed point charge model of the form:

$$U_{nb}(r_{ij}) = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{1}{4\pi\varepsilon_0}\frac{q_i q_j}{r_{ij}} \qquad (1)$$

where r_{ij} is the site separation distance between atoms i and j, ε_{ij} and σ_{ij} are LJ parameters, and q_i and q_j are the partial charge values of atoms i and j, respectively. For interactions between unlike LJ sites, Lorentz−Berthelot[49] combining rules were employed.

To prevent instabilities in the trajectory when the solute was nearly decoupled from the system, that is when $\lambda_m^{\text{LJ}} \approx 0$, solute−solvent intermolecular nonbonded LJ interactions were modeled with a modified "soft-core" potential, $U_{\text{LJ}}^{\text{sc}}$, of the form:[11,50,51]

$$U_{LJ}^{sc}(r_{ij};m) = 4\lambda_m^{LJ}\varepsilon_{ij}\left\{\frac{\sigma_{ij}^{12}}{\left[(1-\lambda_m^{LJ})\alpha_{LJ}\sigma_{ij}^{6} + r_{ij}^{6}\right]^2} - \frac{\sigma_{ij}^{6}}{\left[(1-\lambda_m^{LJ})\alpha_{LJ}\sigma_{ij}^{6} + r_{ij}^{6}\right]}\right\} \qquad (2)$$

where r_{ij}, ε_{ij}, and σ_{ij} are the same LJ parameters as in eq 1, λ_m^{LJ} is the subensemble-dependent coupling strength of the LJ potential, and α_{LJ} is a constant, taken in this study to be 1/2. When the solute is fully coupled to the system, $\lambda_m^{\text{LJ}} = 1$, and eq 2 reduces to the normal LJ potential given by eq 1. When the solute is nearly decoupled, λ_m^{LJ} approaches 0, and eq 2 becomes a smooth interaction function that allows solvent molecules to overlap the solute with finite energy. When the solute is decoupled from the system, $\lambda_m^{\text{LJ}} = 0$, and the solute has no interaction with the solvent. Thus, the potential form in eq 2 correctly represents the limiting behavior of the solute−solvent interactions, while eliminating instabilities when $\lambda_m^{\text{LJ}} \to 0$.

Moreover, the use of eq 2 improves the configurational phase space overlap between the decoupled and nearly decoupled solute subensembles as compared to linearly scaling eq 1. This is clearly illustrated by the fact that when the solute is decoupled from the system, $\lambda_m^{\text{LJ}} = 0$, the solute will explore all regions of configurational phase space with equal probability. When the solute is nearly decoupled, $\lambda_m^{\text{LJ}} \approx 0$, a scaled form of eq 1 would perfectly exclude configurations in which the solute overlaps the solvent from the available phase space. On the other hand, use of eq 2 allows overlapping configurations to be observed with finite energy (and hence finite probability) for the nearly decoupled states, thereby increasing the regions of configurational phase

space mutual to both the decoupled and nearly decoupled subensembles. An increase in the configurational phase space overlap ultimately leads to a decrease in the bias of the calculated free energy between the decoupled and nearly decoupled subensembles.[52,53]

Solute intermolecular electrostatic interactions were decoupled in a linear fashion via the coupling parameter λ_m^{elec}; a detailed description regarding the decoupling of intermolecular interactions with Ewald summation may be found elsewhere.[33]

The same standard LJ and electrostatic interaction potential (eq 1) and combining rules were used for all intramolecular nonbonded interactions by all pairs of atoms separated by four or more bonds. For the case in which the intramolecular sites were separated by exactly three bonds, the LJ and electrostatic interactions were scaled by factors of 1/2 and 5/6,[54,55] respectively, for ibuprofen. These interactions were not necessary for the solvents studied here.

All of the molecules were modeled with flexible bonds, angles, and dihedral angles. The bond stretching and angle bending intramolecular interactions between sites separated by one and two bonds, respectively, were modeled by simple harmonic potentials of the form:

$$U_{bond}(r_{ij}) = k_{ij}(r_{ij} - r_{ij}^0)^2 \tag{3}$$

and

$$U_{angle}(\theta_{ijk}) = k_{ijk}(\theta_{ijk} - \theta_{ijk}^0)^2 \tag{4}$$

where k_{ij}, r_{ij}, and r_{ij}^0 are the force constant, distance between sites i and j, and the corresponding nominal bond length, respectively. Likewise, k_{ijk}, θ_{ijk}, and θ_{ijk}^0 are the force constant, angle between sites i, j, and k, and the corresponding nominal bond angle, respectively. The torsional potential describing the intramolecular interaction between sites separated by three bonds was modeled by a potential of the form:

$$U_{tors}(\phi_{ijkl}) = \sum_{n=0}^{5} K_n \cos^n(\phi_{ijkl} - 180^\circ) \tag{5}$$

where ϕ_{ijkl} is the dihedral angle between sites i, j, k, and l, and the K_n coefficients are constants. The same torsional potential was used to describe improper dihedral angles, meant to keep planar groups planar and to prevent unrealistic chiral inversions. The torsional potential of the solute molecule was scaled linearly via the coupling parameter λ_m^{tors}.

The molecular models for methanol and ethanol were taken from the united-atom transferable potential for phase equilibria (TraPPE-UA) force field of Siepmann and co-workers.[56,57] The TraPPE-UA models have rigid bond lengths; to avoid the use of constraints during the MD simulations, missing harmonic bond parameters were taken from the AMBER Parm94 force field.[58] For water, we employed the simple point charge[59] flexible water (SPC/Fw) model of Wu et al.[60] Parameters for ibuprofen are from the general AMBER force field (GAFF)[54,55] and were taken directly from Klimovich and Mobley.[15]

Further, to increase the MD time step of the EE simulations, we redistributed the mass of the studied molecules as suggested by Feenstra et al.;[61] the procedure involved redistributing a small amount of the mass of carbon and oxygen atoms bonded to a hydrogen to slow the corresponding vibrational frequency and to increase the simulation stability. Additionally, to slow the C═O carbonyl vibrational frequency, the mass of oxygen was redistributed such that both atoms had an atomic mass of 14 amu. Note that while redistributing the mass alters the dynamics of the system, it has no effect on the solvation free energy.[61,62] All of the force field files (which include atomic masses) used in the present study are provided in the Supporting Information.

Simulation Details. *EE.* All of the EE calculations were performed with a modified version of the MD simulation package MDynaMix 5.2.[63,64] For these simulations, LJ interactions were truncated at a distance of r_{cut} = 14 Å, and standard uniform fluid tail corrections were applied to both the energy and the pressure, assuming $g(r)$ = 1 beyond the cutoff.[22,49] Electrostatic interactions were evaluated with an Ewald summation with tin foil boundary conditions,[22,49] with real space interactions truncated at r_{cut}. A damping parameter of αr_{cut} = 3.14 was used, and the maximum number of reciprocal space lattice vectors was set by $K_{\max}$ = 7.0 for water and $K_{\max}$ = 7.3 for methanol and ethanol. Integration of the equations of motion was performed with the multiple-time step method of Tuckerman et al.[65] in Cartesian coordinates. A short time step of 0.2 fs was used for fast intramolecular degrees of freedom and nonbonded interactions within a cutoff of r_{short} = 8 Å, and a time step of 4 fs was used for all other interactions. An Andersen thermostat[66] and Andersen−Hoover barostat[66−68] were used with the collision time for the thermostat set to 0.5 ps, and the time constant for the barostat set to 1.5 ps. Modifications to MDynaMix include implementation of the Andersen thermostat, the "soft-core" potential (eq 2), separate decoupling of LJ and electrostatic solute intermolecular interactions for EE calculations, scaling of the solute torsional potential for EE calculations, WL-BAR, modification of the Ewald summation with EE fractional particles,[33] and other minor additions.

For each ibuprofen−solvent system studied, five independent simulations were performed. The systems were set up by randomly inserting a gas-phase minimized ibuprofen molecule into five previously equilibrated pure solvent boxes for each of the solvents studied. The boxes contained 1000 water, 500 methanol, or 350 ethanol molecules, which gave cubic box lengths of approximately 31−33 Å (water−ethanol). The velocities of each of the systems were initialized from a Maxwell−Boltzmann distribution with a unique seed to the random number generator. Production runs were carried out in an EE-NPT ensemble at 300 K and 1 bar for a total of 24 ns. Each of the five independent systems for each solvent was initialized with a unique random seed for the random number generator used by the thermostat and for the MC random walk. The system began in the subensemble with a noninteracting solute molecule, and attempts to change subensembles were made every 32 fs. Over the first 0.5 ns, the random walk was carried out with WL biasing, in which the WL weight factor was initially taken to be υ_{WL} = 0.5 and reduced as $\upsilon_{\text{WL}}^{\text{new}} = 0.25\upsilon_{\text{WL}}^{\text{old}}$ every 0.1 ns. During the entire course of the simulation, transition energies (in both directions) were computed each time a transition between subensembles was attempted/proposed, and new subensemble weights were computed from BAR every 0.5 ns.[33] To access the ability of EE to sample the H−O−C−O carboxylic acid dihedral angle of ibuprofen, after every proposed subensemble transition (32 fs), the current subensemble and atomic positions of ibuprofen were saved for postsimulation analysis.

In the reference subensemble (i.e., ideal gas state), m = 0, the solute is noninteracting with the rest of the system (λ_0^{LJ} = 0 and λ_0^{elec} = 0), but the intramolecular interactions are full (λ_0^{tors} = 1). The solute was taken from the reference subensemble to the

target subensemble, $m = M_{\text{Total}}$, with a fully interacting solute molecule ($\lambda^{\text{LJ}}_{M_{\text{Total}}} = 1$, $\lambda^{\text{elec}}_{M_{\text{Total}}} = 1$, and $\lambda^{\text{tors}}_{M_{\text{Total}}} = 1$) by first *reducing* the solute intramolecular torsional potential over $M^{\text{ig}}_{\text{tors}} = 3$ subensembles. For these $M^{\text{ig}}_{\text{tors}}$ subensembles, the intermolecular interactions remained off, and the torsional potential was reduced as $\lambda^{\text{tors}}_m = \{0.7, 0.3, \text{and } 0.1\}$ over the range $1 \le m \le 3$. Next, while the torsional potential was reduced, the intermolecular interactions were brought to full strength by first bringing the intermolecular LJ interaction to full strength and then adding in the intermolecular electrostatic interactions. The addition of the LJ and electrostatic interactions were performed separately in $M_{\text{LJ}} = 15$ and $M_{\text{elec}} = 4$ steps, respectively, for a total of 19 steps. First, for the M_{LJ} steps, the intermolecular electrostatic interactions were turned off, and the intermolecular LJ interactions were strengthened as $\lambda^{\text{LJ}}_m = \{0.05, 0.10, 0.20, 0.30, 0.40, 0.50, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95, \text{and } 1.0\}$ over the range $4 \le m \le 18$. Next, the intermolecular electrostatic interactions were strengthened as $\lambda^{\text{elec}}_m = \{0.25, 0.50, 0.75, \text{and } 1.0\}$ over the range $19 \le m \le 22$; details regarding the coupling/decoupling of intermolecular interactions with Ewald summation may be found in our previous work.[33] Lastly, with the intermolecular interactions at full strength, the intramolecular torsional potential of the solute was restored (or restrengthened) over the last $M^{\text{full}}_{\text{tors}} = 3$ subensembles. For these $M^{\text{full}}_{\text{tors}}$ subensembles, the torsional potential was strengthened as $\lambda^{\text{tors}}_m = \{0.3, 0.7, \text{and } 1.0\}$ over the range $23 \le m \le 25$. Therefore there were $M_{\text{Total}} = 25$ subensembles. The intermolecular subensembles were chosen to agree with previous work of Mobley et al.[13−15] No attempt was made to optimize the intramolecular scaling; however, the employed intramolecular scaling scheme was found to be adequate to obtain full rotations of the problematic carboxylic acid dihedral angle of ibuprofen in both the reference and fully interacting (target) subensembles. All of the reported EE results are taken as the average value of the property computed by our five independent simulations, and the uncertainty is taken as one standard deviation.

Conventional MD. Reference calculations were performed using BAR with Gromacs 3.3.4[69] and directly followed the protocol of Klimovich and Mobley.[15] Generally, simulation parameters were similar for both the EE and reference calculations. Minor differences included the use of different MD integrators, a different thermostat and barostat, the use of constrained dynamics for bonds involving hydrogen for methanol, ethanol, and ibuprofen, the use of LJ and electrostatic cutoffs (r_{cut}) of 10.2, 11.25, and 11.85 Å for water, methanol, and ethanol, respectively, and a different treatment of the long-range electrostatic interactions. These differences resulted from the use of different MD simulation software and attempts to efficiently perform both the EE and reference calculations. We recently found that similar small discrepancies did not yield statistically significant differences when computing the hydration free energies of amino acid analogs.[33]

The BAR calculations were performed in 20 steps, composed of an ideal gas reference state and the same 19 (M_{LJ} and M_{elec}) steps to couple/decouple the intermolecular interactions of the solute as were used for the EE calculations. Note that the torsional potential was *not* perturbed for the reference calculations. An independent 5 ns simulation was performed for each state, for a total of 100 ns. The time step for integrating the equations of motion was 2 fs for simulations involving methanol and ethanol and 1 fs for water.

As mentioned previously, the reference simulations suffered from an inadequate sampling of the carboxylic acid dihedral angle H−O−C−O. As a result, a "confine-and-release" framework[70] was necessary to include the free energy of the conformational transformation. This was performed separately using umbrella sampling[71] to compute the potential of mean force (PMF) of the dihedral angle,[72] again following the work of Klimovich and Mobley[15] with Gromacs 3.3.4. Each PMF was constructed of 24 umbrellas. For each umbrella, a 0.5 ns simulation was conducted, for a total of 12 ns of simulation time per PMF.

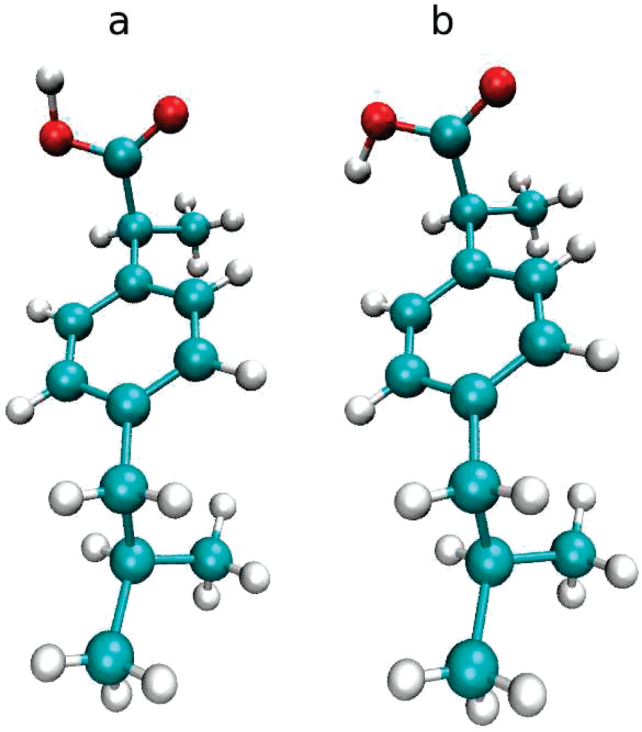

Figure 2. An illustration[81] of the conformational minima of the H−O−C−O dihedral angle of the carboxylic acid functional group of ibuprofen with: (a) the terminal H pointed toward the terminal O ($\phi = 0°$), in what we term conformation "a", and (b) the terminal H pointed away from the terminal O ($\phi = \pm 180°$), in what we term conformation "b". Note that the illustration shows only the rotation of the hydroxyl group; the rest of the molecule is left in the same conformation.

4. RESULTS AND DISCUSSION

As alluded to earlier, conventional MD simulations of ibuprofen in water, methanol, and ethanol are unable to effectively sample the H−O−C−O dihedral angle of the carboxylic acid functional group. For reference, an illustration of the conformational minima of the carboxylic acid dihedral angle is given in Figure 2. An understanding of the molecular level origin of this inefficiency may be understood by looking at the computed PMF of the dihedral angle in an ideal gas state and in solution; PMF calculations were performed using conventional MD as mentioned previously (Section 3). As illustrated in Figure 3, in the ideal gas state, the H−O−C−O dihedral angle has a strong preference to remain in the range of −90 to +90° (conformation "a"). This conformation corresponds to the terminal H pointed toward the terminal O, resulting in a strong intramolecular hydrogen bond. The strength of the interaction is emphasized by the fact that the barrier of rotation is on the order of 20 $k_{\text{B}}T$,

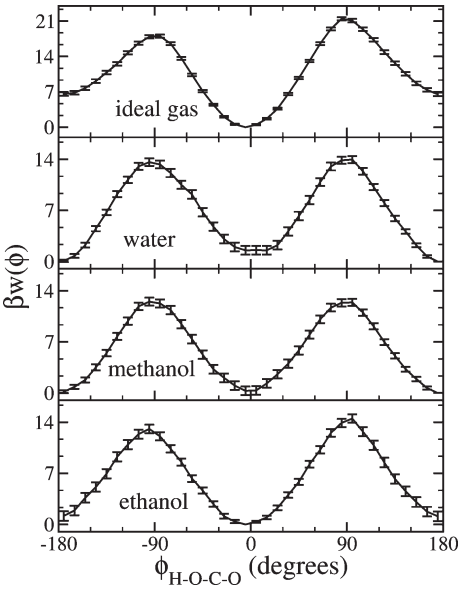

Figure 3. The PMF of rotation of the carboxylic acid dihedral angle of ibuprofen in the reference (ideal gas) subensemble and in water, methanol, and ethanol The *x*-axis corresponds to the H−O−C−O dihedral angle, and the *y*-axis is the dimensionless PMF (or relative free energy). The error bars correspond to the uncertainty of the umbrella sampling calculations.

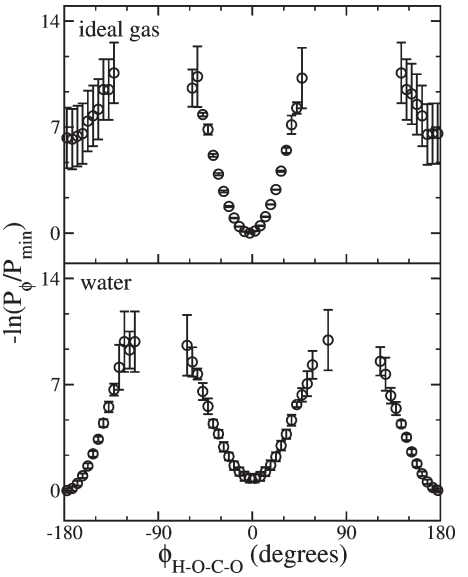

Figure 4. The visited states histogram of observing an H−O−C−O dihedral angle of ibuprofen with EE in the target (solution) and reference (ideal gas) subensemble for water. The *x*-axis corresponds to the H−O−C−O dihedral angle, and the *y*-axis is the negative logarithm of the visited states probability of observing an H−O−C−O dihedral angle relative to the minimum negative logarithm visited states probability. The error bars correspond to the uncertainty taken as the standard deviation of five independent simulations.

and the corresponding conformational minima differ by approximately 7 k_BT.

On the other hand, Figure 3 shows that for water and methanol, the H−O−C−O dihedral angle favors conformations in the range of −180 to −90 and +90 to +180° (conformation "b"). For the case of ethanol, conformation "a" is slightly favored over conformation "b". Moreover, the barrier of rotation in all of the solvents is reduced to approximately 14 k_BT, and the conformational minima differ by only approximately 1 k_BT. This change in behavior results from the presence of a hydrogen-bond accepting solvent. Whereas in the ideal gas state the strongest interaction present was the intramolecular hydrogen bond, in solution a competing hydrogen bonding scenario exists for the terminal H to hydrogen bond with the solvent. The strength of the intermolecular hydrogen bond directly influences the PMF minima of conformation "a" relative to "b" in solution. From Figure 3, we find that the difference in PMF minima between conformation "a" and "b" follows the trend: ethanol < methanol < water. As anticipated, this observation agrees with the trend of increasing strength of the hydrogen bond formed between the terminal H of the carboxylic acid group and the O of the solvent: ethanol < methanol < water.[73,74]

The inefficiencies of conventional MD simulations to compute the solvation free energy of the systems of interest are now evident. When performing free energy calculations, regardless of the employed technique, it is crucial that *all* of the important configurational phase space of the system be adequately sampled in order to obtain a meaningful result. In the present study, this condition demands that both conformations "a" and "b" of the carboxylic acid functional group be sampled in both the ideal gas (reference) and solution (target) states. This would require routinely overcoming barriers for rotation of approximately 14−20 k_BT to sample both conformations and is not realistic using conventional MD simulations. Corrections may be made to the computed free energy to compensate for the inadequate sampling postsimulation using a "confine-and-release" framework.[15,70] These corrections, however, require that the inefficiency and its cause are identified prior to calculating a solvation free energy, and subsequent umbrella sampling calculations be performed to compute the PMF along the relevant reaction coordinate. This requires human intervention and planning, so it would be highly advantageous to avoid such calculations altogether. Another key reason to solve this problem is that it may not always be clear when simulations suffer from inadequate conformational sampling. The large barrier of rotation encountered in the present study is a concerted effect of the intramolecular hydrogen bonding and the torsional potentials of both the carboxylic acid group (H−O−C−O) dihedral and the dihedral of the hydroxyl of the carboxylic acid group and the α carbon (H−O−C−C); the torsional potential of these two dihedral angles used in the present study is similar to other force fields commonly used for biological simulations.[75−77] A recent study of benzoic acid, acetylsalicylic acid, and ibuprofen (all of which contain carboxylic acid functional groups) in water assumed that a conventional NPT MD simulation of 5 ns was sufficient to sample all of the important configurational phase space;[78] the present results suggest that this may not be the case irrespective of the employed force field, motivating the present study.

Contrary to conventional MD simulations, the use of EE may alleviate the need to identify the inadequacy of sampling and the relevant reaction coordinate, dismissing the need to perform subsequent simulations to compute the relevant correction to the free energy. As seen in Figures 4−6, for EE calculations of ibuprofen in water, methanol, and ethanol, respectively, both conformations "a" and "b" of the problematic H−O−C−O dihedral angle of ibuprofen in solution are adequately sampled. This apparent gain in efficiency is a result of the introduction of

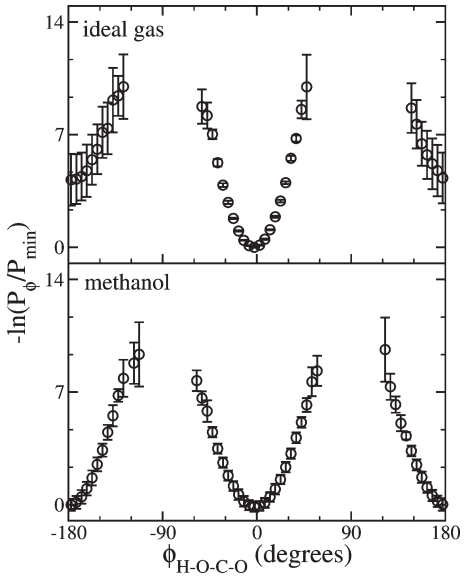

Figure 5. The visited states histogram of observing an H−O−C−O dihedral angle of ibuprofen with EE in the target (solution) and reference (ideal gas) subensemble for methanol. The *x*-axis corresponds to the H−O−C−O dihedral angle, and the *y*-axis is the negative logarithm of the visited states probability of observing an H−O−C−O dihedral angle relative to the minimum negative logarithm visited states probability. The error bars correspond to the uncertainty taken as the standard deviation of five independent simulations.

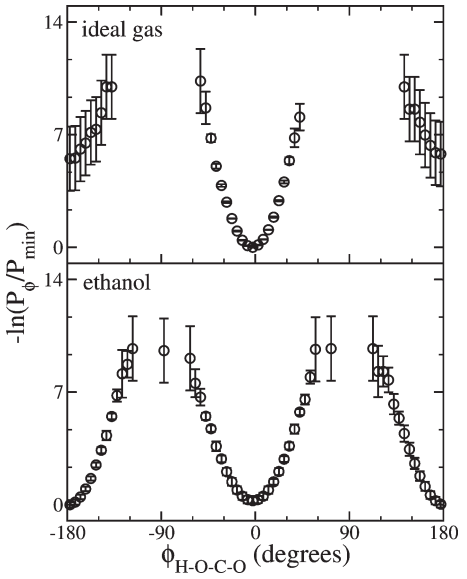

Figure 6. The visited states histogram of observing an H−O−C−O dihedral angle of ibuprofen with EE in the target (solution) and reference (ideal gas) subensemble for ethanol. The *x*-axis corresponds to the H−O−C−O dihedral angle, and the *y*-axis is the negative logarithm of the visited states probability of observing an H−O−C−O dihedral angle relative to the minimum negative logarithm visited states probability. The error bars correspond to the uncertainty taken as the standard deviation of five independent simulations.

intermediate subensembles in which the torsional barrier of rotation is reduced, greatly increasing the frequency of conformational changes in these subensembles. Analogous to replica exchange simulations, the conformations of the intermediate subensembles are then propagated between the target and reference subensembles, thereby effectively enhancing the conformational sampling. This characteristic behavior may be seen in Figure 7; we see in Figure 7 that when ibuprofen is interacting in solution with scaled torsional potentials, the reduced barrier of rotation allows for interconversion between conformation "a" and "b". These conformations are then propagated to our target subensemble. However, we should expect that the lowering of the torsional barrier of rotation in the intermediate subensembles will expand the important configurational phase space of the system. Nevertheless, as compared to conventional MD simulations, the use of EE naturally decreases the configurational correlation time, increasing the rate of exploration of phase space.[33] Ultimately, the significant increase in computational time necessary to sufficiently sample conformational changes of the H−O−C−O dihedral angle far outweighs the cost associated with the increase in the important configurational phase space of the system.

Note that in Figures 4−6, it appears that data are missing for conformations of the carboxylic acid dihedral angle near −90 and +90°, corresponding to the maximum of the PMF in the ideal gas (reference) and solution (target) states (Figure 3). However, this is an artifact of collecting data after every proposed subensemble transition (32 fs) and the propagation of conformations between subensembles. The time spent at these intermediate dihedral angles is short-lived in solution and in the ideal gas state, and as a result, they are not seen in our collected data for the corresponding subensembles. The observance of these dihedral angles is confirmed, however, by the fact that we observe transitions between conformation "a" and "b" during our EE simulations, as shown in Figure 7. Similarly, note that in Figures 4−6, while the observed distribution of conformations of the carboxylic acid dihedral angle in the ideal gas (reference) subensemble are statistically equivalent for all three systems studied, the uncertainties over the range of −90 to −180 and +90 to +180° are much larger than over than range of −90 to +90°. This results simply because these states are so improbable. The uncertainty could be reduced by running longer simulations; however, the impact of this on the computed free energy is negligible.

Having confirmed that enhanced conformation sampling is obtained with EE, we next draw our attention to the calculation of the solvation free energy.Table 1 compares the free energies computed using EE to reference calculations using BAR with additional conformational corrections.[15] The average absolute difference between EE and the reference calculations for the three systems studied is 0.5 ± 0.6 k_BT. For methanol and ethanol, agreement is excellent, with the EE and reference results being statistically equivalent. For water, the discrepancy is larger, 1.0 ± 0.7 k_BT, where the uncertainty is computed from propagation of errors and is taken as one standard deviation. This is still within two standard deviations, reaffirming confidence in the performance of the EE method.

Furthermore, Table 1 demonstrates how poor the uncorrected results are with conventional MD. If an unaware practitioner did not perform the necessary PMF corrections, then they would obtain nonsensical results. This is emphasized by the fact that the PMF correction is 10.5 and 5.6 k_BT for water and methanol, respectively; the magnitude of the PMF correction is 89 and 24% of the magnitude of the uncorrected (BAR) free energy in water and methanol, respectively. These absurd results would have an adverse effect on computed properties of biological interest.[2,4,37,38]

Additionally, two points may be made to emphasize the improved performance gained from employing EE over conventional

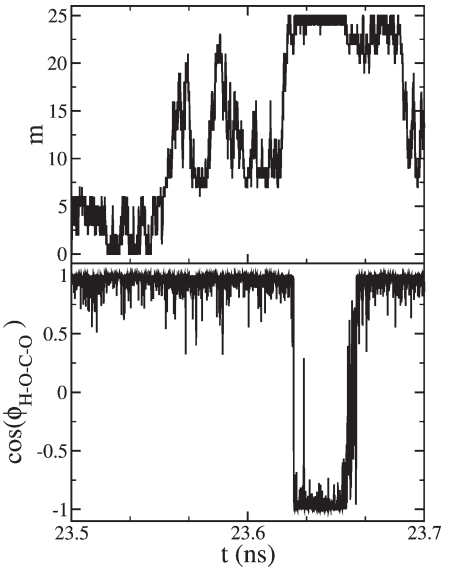

Figure 7. A sample plot of the time evolution of the EE subensemble and the cosine of the H−O−C−O carboxylic acid dihedral angle in water. The plot corresponds to 200 ps of simulation time from our first of five independent simulations of ibuprofen in water and is representative of the enhanced sampling behavior observed in each solvent.

Table 1. Comparison of the Computed Dimensionless Gibbs Free Energy of Solvation, $\Delta\beta G_{solv}$, using EE and Conventional Reference MD Simulations[a]

	EE	reference		
system	$\Delta\beta G_{solv}$	$\Delta\beta G_{solv}$	$\Delta\beta G_{solv}^{BAR}$	$\Delta\beta G_{solv}^{PMF}$
water	−12.8 ± 0.3	−11.8 ± 0.5	−22.3 ± 0.4	10.5 ± 0.3
methanol	−22.9 ± 0.3	−23.2 ± 0.3	−28.8 ± 0.1	5.6 ± 0.3
ethanol	−23.0 ± 0.2	−22.8 ± 0.3	−22.5 ± 0.1	−0.3 ± 0.3

[a] The reference results are composed of a BAR contribution ($\Delta\beta G_{solv}^{BAR}$) and a "confine-and-release" (PMF) correction ($\Delta\beta G_{solv}^{PMF}$) accounting for inadequate sampling of the ibuprofen H−O−C−O dihedral angle. Uncertainties of the EE calculations are taken as the standard deviation of the five independent simulations.

MD simulations. First, for each system studied using EE, we performed 5 independent 24 ns simulations, for a total of 120 ns of simulation time. For the conventional MD calculations, 20 independent 5 ns simulations were conducted to obtain the uncorrected solvation free energy, followed by 72 independent 0.5 ns simulations for the 3 necessary PMF corrections,[15] for a total of 136 ns of simulation time. Therefore, as seen in Table 1, for less simulation time, EE computes solvation free energies with comparable precision to the conventional MD simulations. The efficiency of free energy calculations using conventional MD is limited by the configurational correlation time of the system. On the other hand, for EE, subensemble transitions are periodically attempted. When a transition is accepted, the system begins sampling in a different subensemble. This transitioning reduces the configurational correlation time of the system and increases the rate of configurational phase space sampling, akin to replica exchange simulations, leading to the apparent improved efficiency of EE as compared to conventional MD. Note that both the EE and conventional MD simulations could have been run longer to improve the precision of the calculations; however, we anticipate that the EE calculations would remain more efficient.

However, although the net simulation time required by the EE simulations was less than the conventional MD simulations, it should be noted that with parallelization, the "wall-clock" time of the conventional MD simulations may be less than that of the EE simulations. This is because conventional MD uses more short simulations, while EE uses fewer long simulations. In practice, however, we believe the benefits of running a few simulations to get a solvation free energy directly as opposed to having to run many simulations with subsequent postprocessing corrections argues in favor of the EE method. In addition, implementation of the present EE method in a replica exchange[26] formalism would be a promising solution to reduce the "wall-clock" time.

Second, for the present study, we expect that the relative solvation free energies of ibuprofen in methanol and ethanol to be related to their relative solubility limits by the following relation:[38]

$$\Delta\beta G_{solv}^{methanol} - \Delta\beta G_{solv}^{ethanol} = \ln \frac{x_{ethanol}}{x_{methanol}} \quad (6)$$

where $\Delta\beta G_{solv}^{methanol}$ and $\Delta\beta G_{solv}^{ethanol}$ are the dimensionless Gibbs free energy of solvation of ibuprofen in methanol and ethanol, respectively, $x_{methanol}$ and $x_{ethanol}$ are the mole fraction solubility limit of ibuprofen in methanol and ethanol, respectively, and $\beta = 1/k_BT$. Experimentally,[79] we find that $\ln x_{ethanol}/x_{ethanol} = 0.33$. This is in good agreement with the EE results which give $\Delta\beta G_{solv}^{methanol} - \Delta\beta G_{solv}^{ethanol} = 0.1 \pm 0.4$ but is contrary to the results of the reference calculations which give $\Delta\beta G_{solv}^{methanol} - \Delta\beta G_{solv}^{ethanol} = -0.4 \pm 0.4$. That is, the reference calculations incorrectly predict that ibuprofen is more soluble in methanol relative to ethanol. While the results are statistically equivalent, the EE result is more accurate. Note that the Gibbs free energy of solvation for water predicted using the reference calculations is in excellent agreement with the reported experimental value of $-11.8\ k_BT$.[80] However, without further investigation, it is impossible to determine whether this is a success of the model or just a coincidence, especially since the present study does not account for dissociation of the carboxylic acid functional group in water.

5. SUMMARY AND CONCLUSION

Results have been presented that demonstrate that enhanced conformational sampling may be readily achieved via expanded ensemble molecular dynamics simulations. As a result of the ability of the carboxylic acid functional group of ibuprofen to form both intramolecular and intermolecular hydrogen bonds in the presence of hydrogen-bond accepting solvents, the H−O−C−O dihedral angle is subject to barriers of rotation of $14-20\ k_BT$. Consequently, conventional molecular dynamics simulations are unable to sufficiently observe conformational changes between conformation "a" and "b" during reasonable simulation time scales. This deficiency is particularly problematic when performing free energy calculations for which *all* of the important configurational phase space needs to be adequately sampled to obtain meaningful results. While the computed free energy may be corrected for inadequate sampling with subsequent calculations, this requires a careful practitioner to identify the problem and the relevant reaction coordinate necessary for the correction.

On the other hand, the sampling of the important regions of configurational phase space may easily be improved through use of an expanded ensemble methodology. The method not only boosts the rate of sampling of configurational phase space but also additional subensembles may be introduced to improve the conformational sampling of the solute. The additional subensembles used to enhance conformational sampling do not require an explicit definition of the relevant (burdensome) reaction coordinate. Rather, the torsional potential of all rotatable dihedral angles of the solute may be scaled without complication, while ensuring sufficient conformational sampling of the solute.

The gain obtained by using an expanded ensemble method was emphasized by calculating the solvation free energy of ibuprofen in water, methanol, and ethanol, and by comparing to reference simulation results; the reference simulation results required subsequent calculations to correct for insufficient sampling of the carboxylic acid dihedral angle. Agreement for methanol and ethanol was excellent. Deviations were larger for the results in water, but the expanded ensemble and reference results were within two standard deviations of each other. The agreement reinforced the effectiveness of the expanded ensemble method. Furthermore, the comparable precision obtained by the expanded ensemble and reference calculations, with the expanded ensemble requiring less simulation time, emphasized the enhanced computational efficiency gained. Also, the expanded ensemble calculations require performing only a single simulation and, unlike the reference results, were able to properly rank the solubility of ibuprofen in the studied solvents.

The results are extremely promising and suggest that the expanded ensemble may be used to study the solvation free energies of a wide range of molecules having complicated intramolecular potential energy surfaces. Application of the method to more complicated problems, including protein−ligand binding, is currently being successfully applied in one of our laboratories and will be described in detail in a subsequent paper.

■ ASSOCIATED CONTENT

Supporting Information. MDynaMix 5.2 and Gromacs 3.3.4 force field and topology files for water, methanol, ethanol, and ibuprofen used in this work. This material is available free of charge via the Internet at http://pubs.acs.org/.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: ed@nd.edu. Telephone: (574) 631-5687.

■ ACKNOWLEDGMENT

A.S.P. gratefully acknowledges a fellowship from the Arthur J. Schmitt Foundation, additional funding provided by the Notre Dame Sustainable Energy Initiative, and computing support provided by Notre Dame's Center for Research Computing. D.L.M. acknowledges the Louisiana Board of Regents Research Competitiveness and Research Enhancement Subprograms as well as the Louisiana Optical Network Initiative (supported by the Louisiana Board of Regents Post-Katrina Support Fund Initiative) grant LEQSF(2007-12)-ENH-PKSFI-PRS-01, and the NSF (EPS-1003897).

■ REFERENCES

(1) Chodera, J. D.; Mobley, D. L.; Shirts, M. R.; Dixon, R. W.; Branson, K.; Pande, V. S. *Curr. Opin. Struct. Biol.* **2011**, *21*, 150–160.
(2) Jorgensen, W. L.; Briggs, J. M.; Contreras, M. L. *J. Phys. Chem.* **1990**, *94*, 1683–1686.
(3) Prausnitz, J. M.; Lichtenthaler, R. N.; de Azevedo, E. G. *Molecular Thermodynamics of Fluid-phase Equilibria*, 3rd ed.; Prentice-Hall PTR: Upper Saddle River, NJ, 1999.
(4) Tembe, B. L.; McCammon, J. A. *Comput. Chem.* **1984**, *8*, 281–283.
(5) Gilson, M. K.; Zhou, H. X. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 21–42.
(6) Connors, K. A.; Mecozzi, S. *Thermodynamics of Pharmaceutical Systems: An Introduction to Theory and Applications*, 2nd ed.; John Wiley and Sons, Inc.: Hoboken, NJ, 2010.
(7) *Water-Insoluble Drug Formulation*, 2nd ed.; Liu, R., Ed.; CRC Press: Boca Raton, FL, 2008.
(8) Anfinsen, C. B. *Science* **1973**, *181*, 223–230.
(9) Dill, K. A.; Ozkan, S. B.; Shell, M. S.; Weikl, T. R. *Annu. Rev. Biophys.* **2008**, *37*, 289–316.
(10) Shirts, M. R.; Pitera, J. W.; Swope, W. C.; Pande, V. S. *J. Chem. Phys.* **2003**, *119*, 5740–5761.
(11) Shirts, M. R.; Pande, V. S. *J. Chem. Phys.* **2005**, *122*, 134508.
(12) Hess, B.; van der Vegt, N. F. A. *J. Phys. Chem. B* **2006**, *110*, 17616–17626.
(13) Mobley, D. L.; Bayly, C. I.; Cooper, M. D.; Dill, K. A. *J. Phys. Chem. B* **2009**, *113*, 4533–4537.
(14) Mobley, D. L.; Bayly, C. I.; Cooper, M. D.; Shirts, M. R.; Dill, K. A. *J. Chem. Theory Comput.* **2009**, *5*, 350–358.
(15) Klimovich, P. V.; Mobley, D. A. *J. Comput.-Aided Mol. Des.* **2010**, *24*, 307–316.
(16) Hodel, A.; Simonson, T.; Fox, R. O.; Brunger, A. T. *J. Phys. Chem.* **1993**, *97*, 3409–3417.
(17) Leitgeb, M.; Schroder, C.; Boresch, S. *J. Chem. Phys.* **2005**, *122*, 084109.
(18) McQuarrie, D. A. *Statistical Mechanics*; University Science Books: Sausalito, CA, 2000.
(19) Eckmann, J. P.; Ruelle, D. *Rev. Mod. Phys.* **1985**, *57*, 617–656.
(20) Palmer, R. G. *Adv. Phys.* **1982**, *31*, 669–735.
(21) Thirumalai, D.; Mountain, R. D.; Kirkpatrick, T. R. *Phys. Rev. A* **1989**, *39*, 3563–3574.
(22) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed.; Academic Press: San Diego, CA, 2002.
(23) Berne, B. J.; Straub, J. E. *Curr. Opin. Struct. Biol.* **1997**, *7*, 181–189.
(24) Liu, Z.; Berne, B. J. *J. Chem. Phys.* **1993**, *99*, 6071–6077.
(25) Hongzhi, L.; Fajer, M.; Yang, W. *J. Chem. Phys.* **2007**, *126*, 024106.
(26) Earl, D. J.; Deem, M. W. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3910–3916.
(27) Rosso, L.; Minary, P.; Zhu, Z.; Tuckerman, M. E. *J. Chem. Phys.* **2002**, *116*, 4389–4402.
(28) Zhu, Z.; Tuckerman, M. E.; Samuelson, S. O.; Martyna, G. J. *Phys. Rev. Lett.* **2002**, *88*, 100201.
(29) Lin, I. C.; Tuckerman, M. E. *J. Phys. Chem. B* **2010**, *114*, 15935–15940.
(30) Hansen, H. S.; Hunenberger, P. H. *J. Comput. Chem.* **2010**, *31*, 1–23.
(31) Ferguson, A. L.; Debenedetti, P. G.; Panagiotopoulos, A. Z. *J. Phys. Chem. B* **2009**, *113*, 6405–6414.
(32) Khavrutskii, I. V.; Wallqvist, A. *J. Chem. Theory Comput.* **2010**, *6*, 3427–3441.
(33) Paluch, A. S.; Shah, J. K.; Maginn, E. J. *J. Chem. Theory Comput.* **2011**, *7*, 1394–1403.
(34) Lyubartsev, A. P.; Martsinovski, A. A.; Shevkunov, S. V.; Vorontsov-Velyaminov, P. N. *J. Chem. Phys.* **1992**, *96*, 1776–1783.
(35) Lyubartsev, A. P.; Laaksonen, A.; Vorontsov-Velyaminov, P. N. *Mol. Phys.* **1994**, *82*, 455–471.
(36) Wilding, N. B.; Muller, M. *J. Chem. Phys.* **1994**, *101*, 4324–4330.
(37) Paluch, A. S.; Jayaraman, S.; Shah, J. K.; Maginn, E. J. *J. Chem. Phys.* **2010**, *133*, 124504.

(38) Paluch, A. S.; Cryan, D. D.; Maginn, E. J. *J. Chem. Eng. Data* **2011**, *56*, 1587–1595.
(39) Lyubartsev, A. P.; Jacobsson, S. P.; Sundholm, G.; Laaksonen, A. *J. Phys. Chem. B* **2001**, *105*, 7775–7782.
(40) Chang, J. *J. Chem. Phys.* **2009**, *131*, 074103.
(41) Boulougouris, G. C.; Errington, J. R.; Economou, I. G.; Panagiotopoulos, A. Z.; Theodorou, D. N. *J. Phys. Chem. B* **2000**, *104*, 4958–4963.
(42) Wang, F.; Landau, D. P. *Phys. Rev. Lett.* **2001**, *86*, 2050–2053.
(43) Yan, Q.; Faller, R.; de Pablo, J. J. *J. Chem. Phys.* **2002**, *116*, 8745–8749.
(44) Shell, M. S.; Debenedetti, P. G.; Panagiotopoulos, A. Z. *Phys. Rev. E* **2002**, *66*, 056703.
(45) Bennett, C. H. *J. Comput. Phys.* **1976**, *22*, 245–268.
(46) Lu, N.; Singh, J. K.; Kofke, D. A. *J. Chem. Phys.* **2003**, *118*, 2977–2984.
(47) Shirts, M. R.; Bair, E.; Hooker, G.; Pande, V. S. *Phys. Rev. Lett.* **2003**, *91*, 140601.
(48) Fenwick, M. K.; Escobedo, F. A. *J. Chem. Phys.* **2004**, *120*, 3066–3074.
(49) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press Inc.: New York, 1987.
(50) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1994**, *222*, 529–539.
(51) Steinbrecher, T.; Mobley, D. L.; Case, D. A. *J. Chem. Phys.* **2007**, *127*, 214108.
(52) Wu, D.; Kofke, D. A. *J. Chem. Phys.* **2005**, *123*, 054103.
(53) Wu, D.; Kofke, D. A. *J. Chem. Phys.* **2005**, *123*, 084109.
(54) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157–1174.
(55) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. *J. Mol. Graphics Modell.* **2006**, *25*, 247–260.
(56) Martin, M. G.; Siepmann, J. I. *J. Phys. Chem. B* **1998**, *102*, 2569–2577.
(57) Chen, B.; Potoff, J. J.; Siepmann, J. I. *J. Phys. Chem. B* **2001**, *105*, 3093–3104.
(58) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T. F.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
(59) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269–6271.
(60) Wu, Y.; Tepper, H. L.; Voth, G. A. *J. Chem. Phys.* **2006**, *124*, 024503.
(61) Feenstra, K. A.; Hess, B.; Berendsen, H. J. C. *J. Comput. Chem.* **1999**, *20*, 786–798.
(62) Bennett, C. H. *J. Comput. Phys.* **1975**, *19*, 267–279.
(63) Lyubartsev, A. P.; Laaksonen, A. *Comput. Phys. Commun.* **2000**, *128*, 565–589.
(64) Lyubartsev, A. P.; Laaksonen, A. *MDynaMix: a Molecular Dynamics Program*; Stockholm University: Stockholm, Sweden; http://www.fos.su.se/~sasha/mdynamix/. Accessed February 1, 2010).
(65) Tuckerman, M.; Berne, B. J.; Martyna, G. J. *J. Chem. Phys.* **1992**, *97*, 1990–2001.
(66) Andersen, H. C. *J. Chem. Phys.* **1980**, *72*, 2384–2393.
(67) Martyna, G. J.; Tobias, D. J.; Klein, M. L. *J. Chem. Phys.* **1994**, *101*, 4177–4189.
(68) Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. *Mol. Phys.* **1996**, *87*, 1117–1157.
(69) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comput. Chem.* **2005**, *26*, 1701–1718.
(70) Mobley, D. L.; Chodera, J. D.; Dill, K. A. *J. Chem. Phys.* **2006**, *125*, 084902.
(71) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187–199.
(72) Shirts, M. R.; Chodera, J. D. *J. Chem. Phys.* **2008**, *129*, 124105.
(73) Dill, K. A.; Bromberg, S. *Molecular Driving Forces: Statistical Thermodynamics in Chemistry and Biology*; Garland Science: New York, 2003.
(74) Israelachvili, J. *Intermolecular and Surface Forces*, 2nd ed.; Academic Press: San Diego, CA, 1991.
(75) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049–1074.
(76) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
(77) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
(78) Garrido, N. M.; Queimada, A. J.; Jorge, M.; Economou, I. G.; Macedo, E. A. *Fluid Phase Equilib.* **2010**, *296*, 110–115.
(79) Stovall, D. M.; Givens, C.; Keown, S.; Hoover, K. R.; Rodriquez, E.; Acree, W. E.; Abraham, M. H. *Phys. Chem. Liq.* **2005**, *43*, 261–268.
(80) Gaballe, M. T.; Skillman, A. G.; Nicholls, A.; Guthrie, J. P.; Taylor, P. J. *J. Comput.-Aided Mol. Des.* **2010**, *24*, 259–279.
(81) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.