

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/6184029>

Absolute Free-Energy Calculations of Liquids Using a Harmonic Reference State

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · SEPTEMBER 2007

Impact Factor: 3.3 · DOI: 10.1021/jp072357w · Source: PubMed

CITATIONS

16

READS

12

3 AUTHORS, INCLUDING:



[Richard Barry Sessions](#)

University of Bristol

182 PUBLICATIONS 4,260 CITATIONS

SEE PROFILE

Absolute Free-Energy Calculations of Liquids Using a Harmonic Reference State

Michael D. Tyka,* Richard B. Sessions, and Anthony R. Clarke

Department of Biochemistry, School of Medical Sciences, University of Bristol, Bristol BS8 1TD, UK

Received: March 25, 2007; In Final Form: May 18, 2007

Absolute free-energy methods provide a potential solution to the overlap problem in free-energy calculations. In this paper, we report an extension of the previously published confinement method (*J. Phys. Chem. B* **2006**, *110*, 17212–20) to fluid simulations. Absolute free energies of liquid argon and liquid water are obtained accurately and compared with results from thermodynamic integration. The method works by transforming the liquid state into a harmonic, solid reference state. This is achieved using a special restraint potential that allows molecules to change their restraint position during the simulation, which circumvents the need for the molecules to sample the full extent of their translational freedom. The absolute free energy of the completely restrained reference state is obtained from a normal mode calculation. Because of the generic reference state used, the method is applicable to nonhomogeneous, diffusive systems and could provide an alternative method in situations in which solute annihilation fails due to the size of the solute. Potential applications include calculation of solvation energies of large molecules and free energies of peptide conformational changes in explicit solvent.

1. Introduction

Calculation of free energies of solvation,^{1–3} ligand binding,^{4–8} and small conformational transitions^{9–11} have become commonplace in the computer simulation field. Most current methods rely on a path along which the system is transformed from one state to another. This path may be the annihilation of a molecule in a box of solvent to obtain the hydration free energy¹ or the path describing a conformational change.¹¹ The calculation of the free energy is broken down into individual simulations, each taking only a small step along the path. When larger, more complex changes occur, the mutual overlap of the end states worsens, and a longer path is required to connect them. This requirement makes the calculation of free energies of such processes prohibitively expensive in terms of computing effort. Examples include the calculation of solvation energies of large solutes or the calculation of the free energy of complex conformational changes.

Calculation of absolute free energies circumvents this problem because the states no longer require any overlap in phase space. For some systems, the absolute free energy can be obtained by transforming them into a reference state of known free energy. An Einstein solid can be used for solid states^{12,13} and even simple peptides.¹⁴ For liquids, an ideal fluid is suitable,^{15–19} although problems arise in the presence of bulky solutes as the overlap between the reference fluid and the real fluid worsens as the solute becomes larger and more complex. An interesting method was published recently in which a custom reference state is constructed to obtain the absolute free energy of a small peptide.²⁰ Homogeneous liquids also allow application of Gaussian²¹ and quasi-Gaussian approximations²² to extract free energies.

Other methods include “minima mining” approaches^{23–25} in which all the relevant energy minima of a system are found and the free energy is determined for each minimum in different ways, depending on the exact methodology. The data is then

combined into an absolute free energy of the entire state. Hypothetical scanning is another method for the calculation of absolute free energies and has been tested on systems as complex as peptides in implicit solvent and liquid argon and water.^{26–29}

Previously, we have developed a method for the calculation of absolute free energies³⁰ called the confinement method and applied it to a peptide in implicit solvent. As with other approaches^{12,14,15,18,20} it relies on transforming the system of interest into a state whose absolute free energy is known. In this case, the reference state is a highly restrained state in which all nonharmonic motion has been eliminated. The free energy of this reference state can be accurately obtained from a normal mode calculation, and thus, the absolute free energy of the unrestrained state can be determined. The confinement method performed well on peptide systems but was limited to implicit solvent models. In this paper, we show how to modify the method to calculate absolute free energies of fluids under periodic boundary conditions. This is a critical step toward the ultimate goal of calculating absolute free energies of peptides and other solutes in explicit solvent.

2. Methods

2.1. Theoretical Development. The confinement method³⁰ operates by freezing out all nonharmonic motions of the system of interest by applying a restraint and calculating the corresponding free-energy change, ΔF_{conf} . The absolute free energy of the restrained reference state, F_{ref} is then calculated using a normal-mode calculation, allowing the absolute free energy of the free system to be obtained.

In the general case, we can write the potential energy of a restrained system as the sum of the force field energy $\mathcal{V}_{\text{ff}}(\mathbf{r})$ and a restraint term,

$$\mathcal{V}(\mathbf{r}) = \mathcal{V}_{\text{ff}}(\mathbf{r}) + \frac{1}{2}k\omega(\mathbf{r}) \quad (1)$$

* Corresponding author. E-mail: mtyka@u.washington.edu.

where k is the restraint constant and $w(\mathbf{r})$ is some function of the system configuration \mathbf{r} . The free energy of turning up the restraint from 0 to k_f can be determined using thermodynamic integration³⁰ and is

$$\Delta F_{\text{conf}} = \frac{1}{2} \int_0^{k_f} \langle w(\mathbf{r}) \rangle_k dk \quad (2)$$

In the simplest case, the restraint function is simply a Cartesian restraint to a set of positions $\hat{\mathbf{r}}$.

$$w(\mathbf{r}) = |\mathbf{r} - \hat{\mathbf{r}}|^2 \quad (3)$$

The same process can, in principle, be applied to a liquid or a gas. This yields the absolute free energy of the system, F_{abs} , which is the sum of the excess free energy, F_e , and the ideal gas contribution, F_{id} , given by

$$F_{\text{id}} = -kT \ln \left[\frac{V^N (2\pi m k_B T)^{3N/2}}{N! h^2} \right] \quad (4)$$

Here, V is the volume of the simulation box, N is the number of particles, and m is their mass. k_B is the Boltzmann constant, T is the temperature, and h is Planck's constant.

The excess free energy is obtained by subtracting the reversible work, ΔF_{conf} , needed to apply the restraint and the ideal gas contribution, F_{id} , from the free energy of the reference state, F_{ref} .

$$F_e = F_{\text{abs}} - F_{\text{id}} = F_{\text{ref}} - \Delta F_{\text{conf}} - F_{\text{id}} + k_B T \ln N! \quad (5)$$

The last term accounts for the fact that applying the confinement restraint effectively labels each atom, but the ideal gas contribution is unlabeled. The labeling stems from the fact that each atom is assigned a unique restraint position, rendering each particle distinguishable from every other particle.

Let us consider the practicality of this straightforward application of the theory to the solvent simulation. When the confinement force constant, k , is close to 0, the solvent atoms are free to diffuse through the simulation box. During the simulation, we must calculate the canonical average of the restraint function, $\langle w(\mathbf{r}) \rangle_k$. In practice, this quantity would converge extremely slowly, because each atom would need to diffuse throughout the simulation box, sampling all possible positions in the simulation volume during the simulation. Simulations of impractical length would be required to obtain converged results because the diffusion of molecules is relatively slow compared to the timescales of a typical molecular simulation. Thus, we cannot expect this naive, straightforward application of the confinement method to work in practice.

The configurational integral for a set of N identical atoms is given by

$$Z_N = \int_{\Omega} d\mathbf{q} \exp(-\mathcal{V}(\mathbf{q})/k_B T) \quad (6)$$

where the integration is over the whole of the configurational phase space, Ω . Consider a given coordinate in configurational phase space $\mathbf{q} = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$. It is part of a set $P_{\mathbf{q}}$ of $N!$ coordinates $P_{\mathbf{q}} = \{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_{N!}\}$. The coordinates in the set $P_{\mathbf{q}}$ differ *only* by the fact that the coordinates of the atoms have been permuted around, yielding the $N!$ possibilities of arranging N atoms on a set of N unique positions in the simulation volume. Most importantly, all $N!$ states have the same energy; that is,

$$\mathcal{V}(\mathbf{q}_1) = \mathcal{V}(\mathbf{q}_2) = \dots = \mathcal{V}(\mathbf{q}_{N!})$$

because all the atoms in the set are identical. Thus, swapping any two atom positions does not affect the potential energy. Because all the coordinates in a set are equivalent in this way, we can represent the entire set of coordinates by one unique, representative coordinate, say, \mathbf{w} . Consider the hyperspace \mathbf{D} of *unique* coordinates \mathbf{w} that is a subvolume of the total phase space, Ω . Note that \mathbf{D} will cover exactly $1/N!$ of the volume of Ω and should thus be considerably easier to sample! If we perform the integration in eq 6 over only that subspace \mathbf{D} , each point of which really represents a set of $N!$ real coordinates, the configurational integral becomes³²

$$Z_N = N! \int_{\mathbf{D}} d\mathbf{q} \exp(-\mathcal{V}(\mathbf{q})/k_B T) = N! Z_{\mathbf{D}} \quad (7)$$

It is important to note that there are a large number of possible subspaces, all of which will yield the same result for Z_N because of the equivalence of the states in each set, as discussed above. This result will hold as long as only one representative member from each set $P_{\mathbf{q}}$ is included in the integral in eq 7.

The importance of the result is that if we can obtain the partition function $Z_{\mathbf{D}}$ (or the associated free energy, $F_{\mathbf{D}}$), then we can obtain the full partition function. To achieve this, we must restrict the simulation to the subspace of unique coordinates, where the choice of unique coordinates is arbitrary in principle and can be chosen appropriately to ensure fast convergence.

How can we use this transformation to implement the confinement method on an explicit fluid simulation? Let us consider a given step, τ , during the confinement simulation of a homogeneous monoatomic fluid of N atoms. We have a set of N atom positions,

$$P = \{\mathbf{r}_0, \mathbf{r}_1, \dots, \mathbf{r}_N\}$$

and a set of N restraint positions,

$$R = \{\hat{\mathbf{r}}_0, \hat{\mathbf{r}}_1, \dots, \hat{\mathbf{r}}_N\}$$

There is a one-to-one correspondence (bijection) between the two sets P and R , also known as an *assignment*. For a Cartesian restraint, the restraint function is the total squared displacement from the restraint positions,

$$w(\mathbf{r}) = \sum_i^N |\mathbf{r}_i - \hat{\mathbf{r}}_i|^2$$

Imagine swapping the positions of a pair of atoms, leaving their respective restraint positions where they are; the swap changes the total restraint energy. Among all the possible $N!$ permutations of atom swaps, there will be one arrangement in which the restraint energy will be at a minimum. For any set of atom positions, such a unique and optimal permutation will exist. If, during a sampling procedure, we determined the optimal permutation and reset the atom positions to that particular permutation after every change to the system, we cause all possible permutations of atoms swaps to be mapped back onto one *unique* representative structure. This satisfies the criterion established in eq 7 above for sampling from a subspace \mathbf{D} of the total configurational phase space. A similar permutation reduction has been used recently for quasiharmonic analysis of liquids.³¹

The problem of finding the optimal permutation above is a straightforward example of a problem known as the *linear assignment problem*, which can be stated in a general form as the following situation: Imagine a number of tasks and a number

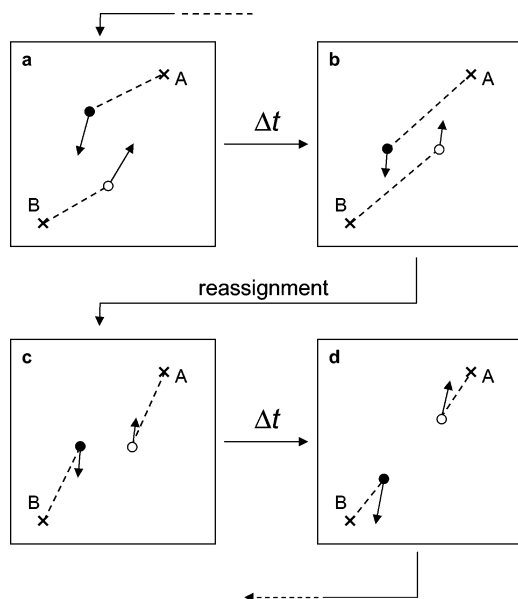


Figure 1. A schematic illustration of the reassignment of the mobile particle positions to the static restraint positions. Two integration time steps and one reassignment cycle are shown. See text for details.

of workers to perform the tasks. Any task can be performed by any worker and each possible task-to-worker assignment will incur a certain cost. Requiring that each task be assigned only one worker, the goal is to find the assignment of tasks to workers that will minimize the total cost. A typical example may be a taxi company with a set of taxis and a number of people waiting to be picked up. How do you choose which taxi to send to which person to minimize the total amount of fuel used?

Mathematically, the linear assignment can be defined as the following:

Given two sets, P and R , of equal size, find a bijection $f: P \rightarrow R$ such that the cost function,

$$\sum_{p \in P} C(p, f(p))$$

is minimized. In our case, we set the cost function, C , to the displacement of a given atom,

$$C(i, j) \rightarrow |\mathbf{r}_i - \mathbf{r}_j|^2$$

The total cost then equals the total squared displacement given a certain assignment of atoms to restraint positions. The solution to this classic problem was first devised by Harold Kuhn in 1955^{33,34} and later improved by James Munkres³⁵ and is known as the Kuhn–Munkres algorithm. Faster algorithms have since been developed, and we will not dwell here on the details of these procedures except to say that we have implemented the Jonker–Volgenant–Castanon (JVC) algorithm³⁶ into our molecular mechanics software.³⁷

Figure 1 illustrates the overall algorithm strategy schematically. Figure 1a shows a given point during a molecular dynamics simulation. Two particles (filled and open dots) are present inside a square, 2-dimensional box. Their current attachment to a restraint position (marked by the black crosses labeled A and B) is indicated by the dotted line along which the restraint force is exerted on each particle (in addition to any other forces arising from the normal force field). Their current velocities are shown as black arrows and indicate the approximate direction of travel. After an integration time step (Δt), the two particles will advance their positions, as shown

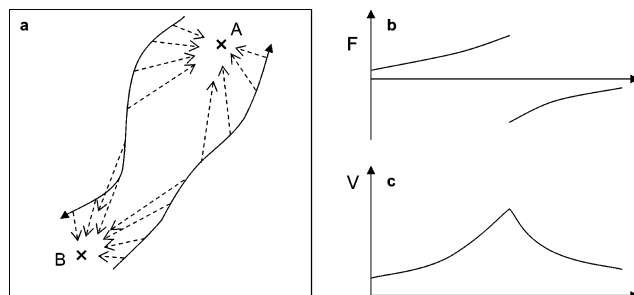


Figure 2. (a) Direction of the restraint force for two particles passing each other with the trajectories indicated by the solid, curved arrows. The direction of the restraint force at different time points along each trajectory is indicated by the dashed arrows. Roughly halfway, the assignment will change and the force direction will change abruptly. (b) The evolution of a force component (F_y) of one of the particles and the potential energy (V) of the entire system. The total potential energy is conserved during the switch (or minimally affected in practice due to a finite time step) because the cost function is set proportional to the restraint energy.

in Figure 1b. Next, as in any ordinary MD algorithm, the forces and energies are recalculated including the restraint forces. As part of this step, the assignment is updated. Because the particles have strayed quite far from their current restraint positions, the assignment algorithm would reassign their restraint positions (i.e., in this case, the restraint positions are swapped), as shown in Figure 1c. This is because the restraint energy with the old assignment in b is higher than with the new assignment shown in Figure 1c. The assignment algorithm always finds the precise arrangement that gives the lowest restraint energy. This process is iterated as in ordinary MD to obtain an NVT ensemble and, most importantly, the average restraint function that is used as the raw data to calculate the free energy of applying the restraint using eq 2.

2.2. Force Discontinuity. The swapping of restraint positions is an unusual force-field feature leading to severe discontinuities in the forces. Consider the two particles flying past each other as shown in Figure 2. Their restraint “attachment”, that is, the direction of the restoring force, is indicated by the thin black arrows. In their trajectory, the particles approach a critical point roughly halfway between their respective restraint positions. The critical point is where the two possible assignments of particle positions to restraint positions yield exactly the same value for the total cost. In this particular example, as they approach this critical point, they will initially be pulled backward, that is, the force vector will be in the direction opposite to the velocity vector. Assuming their kinetic energy is sufficient to carry them onward, they will pass the critical point and swap their restraint positions due to the reassignment. Suddenly, the restraint forces are pointing roughly in the same direction as the velocities. Figure 2b plots the force in the y direction and of the particle A as a function of time (particle B would display an analogous plot). Also plotted is the energy of the whole system. Because of the equivalence of assignment at the critical point, the total energy is continuous. In other words, in theory, assuming an infinitely short integration time step, the system has exactly the same amount of potential energy before and after the assignment, although the distribution of energies among the particles will have changed. Note that this will be true only when the cost function is proportional to the restraint function. Although in principle *any* assignment cost function could be used to calculate the correct unlabeled free energy, provided it generates a unique assignment for a given configuration, in practice, it is advantageous to make the cost function proportional to the total restraint energy. However, in contrast to the energy, the derivative of

the energy with respect to the particle positions (i.e., the *total* force vector) will have changed its direction after the assignment, but not its magnitude. If one considers a single particle, however, the magnitude of the atomic force changes also. The fact that the forces are discontinuous when assignments occur is a potential problem for molecular dynamics simulations, which rely on continuous forces (and continuous second derivatives, which in the present case suffer a singularity at the assignment point). It should be noted, however, that although the discontinuities become more severe with increasing restraint constant, the particles become far less likely to swap their assignment due to their stronger restraint. At very strong restraints, no reassignment will happen at all, completely eliminating the discontinuities. We thus expect the problem to be worst in the intermediate region where the liquid is turning into a solid, that is, near the forced phase transition.

It is critical to ensure that molecular dynamics will sample from a correct canonical ensemble, despite the force discontinuities. To investigate the effect of the discontinuity problem, we performed preliminary experiments to test the energy conservation in NVE trajectories. Satisfactory energy conservation in microcanonical sampling would strongly indicate that the impact of the discontinuities is not prohibitive to the extraction of correct thermodynamic data from trajectories sampling from the canonical ensemble. Self-consistent agreement of calculated free energies with those obtained using standard methods will further strengthen this assumption.

It is important to note that the problems discussed above do not apply for a Monte Carlo implementation of this algorithm, since Monte Carlo does not rely on the derivatives of the potential energy. Despite the assignment, every arrangement of particles has a well-defined, unique energy on which the Metropolis criterion can be applied, leading to a normal canonical ensemble. This may mean that Monte Carlo is inherently better suited to the proposed method. This, however, is not explored in this paper and will be the subject of future work. We will show in the following that stochastic MD simulation (Langevin dynamics) is, indeed, sufficient to obtain accurate free energies, despite the problems discussed above.

2.3. Molecular Fluids. In a molecular fluid, the lack of spherical symmetry of the molecules requires the confinement to restrain their orientations in addition to their positions. The simplest way to achieve this is to apply the restraints to every atom in the molecule. The reassignment, which removes the need to explore the entire translational freedom of each molecule, now occurs over the entire molecule, that is, molecule identities are swapped, as opposed to atom identities. The cost function for a given molecule and its restraint position is now the sum over the individual atom contributions.

As before, the ideal gas contributions need to be removed to obtain the excess free energy. In a molecular fluid, the ideal gas contribution now also contains rotational and vibrational parts; thus, the ideal gas part is now given as

$$F_{\text{id}} = -k_{\text{B}}T \ln(q_{\text{tra}}q_{\text{rot}}q_{\text{vib}}) \quad (8)$$

The translational part is calculated as for an ideal gas using eq 4, where the mass, m , is the total mass of the molecule. The rotational part is calculated from the moments of inertia of the molecule in its ground state. The expression for a nonlinear molecule is

$$q_{\text{rot}} = \frac{8\pi^2}{\sigma h^3} (2\pi k_{\text{B}}T)^{3/2} (I_{\text{A}}I_{\text{B}}I_{\text{C}})^{1/2} \quad (9)$$

where I_{A} , I_{B} , and I_{C} are the principal moments of inertia, that is, the eigenvalues of the inertia tensor. σ is the internal symmetry number, that is, the number of unique rotations that will restore the molecule into an indistinguishable state (for example H_2O has a symmetry number of 2). The symmetry number essentially takes the corresponding role to the $1/N!$ factor in the translational partition function. Analogously, if the symmetric atoms (in the above cases, the hydrogens) were labeled, the symmetry number would become 1, because all previously indistinguishable rotations are now distinguishable.

Finally, the vibrational part of the partition function is obtained from the normal modes of vibration and the classical expression for a harmonic oscillator,

$$q_{\text{vib}} = \prod_i^{\kappa} \frac{k_{\text{B}}T}{h\nu_i} \quad (10)$$

where ν_i is the i th normal-mode frequency and κ is the number of internal degrees of freedom: $\kappa = 3N - 6$, where N is the number of atoms in the molecule.

Internal Symmetries. Analogous to swapping the molecule identities, one can use the intramolecular symmetry to perform an identity reassignment within each molecule. By always choosing the atom-to-restraint assignment that gives the lower value of the restraint function, $w(\mathbf{r}) = \sum |\mathbf{r}_i - \hat{\mathbf{r}}_i|$, for every molecule in turn, we unlabel the atoms within each molecule. This must consequently be taken into account as a symmetry factor in the partition function. This intramolecular reassignment is easily extendable to molecules with symmetries >2 ; for example, NH_4^+ . Despite the fact that this reassignment must be done for each molecule in turn, it is computationally very efficient, since it is an operation of complexity $O(N)$. By performing the intramolecular assignment, we circumvent the need for the molecule to sample multiple orientations, essentially cutting the size of conformational space to be sampled by σ^N . This saving could be significant in a highly polar molecule, such as H_2O , because the hydrogen-bonded nature of liquid water slows rotational motion. When a restraint constant is applied and the overall structure of the liquid is further rigidified, molecules become easily trapped in local minima. Simulations with and without this symmetry switch were conducted.

2.4. Simulation Details. We applied the method to two systems. The first comprises 64 atoms of argon and the second comprises 100 molecules of TIP3P water. For each system, we investigated both a normal version and an “ideal” version, in which all *intermolecular* interactions were turned off to represent an ideal gas or an ideal molecular fluid, respectively.

Argon System. A cubic periodic boundary box with an edge length of 14.4 Å was filled with $N = 64$ argon atoms of 39.948 atomic mass units. The simulation was carried out using Langevin integration,³⁸ generating an NVT ensemble at $T = 96.53$ K. The restraint protocol used in the ideal-gas simulation is shown in Table 1. For some values of the restraint constant, we increased the simulation length as indicated in the table.

In the nonideal simulation, the argon atoms interacted via a 12–6 Lennard–Jones potential with the typical parameters^{15,29,39,40,41} $\epsilon/k_{\text{B}} = 119.8$ K and $\sigma = 3.405$ Å. The reduced density was $\rho^* = N\sigma^3/V = 0.846$. Interactions were cut off at 7.15 Å applying a force switch between 6 and 7.15 Å to prevent discontinuities in the potential and derivatives.⁴² No long-range corrections were included at this stage. The restraint structure was prepared by equilibrating the system by MD with a linear ramp from 300 to 10 K over 50 000 steps and followed by minimization until the energy gradient was negligible. The

TABLE 1: Confinement Protocol Used for Ideal Gas of Argon Atoms

k (kcal/mol/Å ²)	sampling ^a
0.001	1.0
0.01	1.0
0.03	1.0
0.055	1.0
0.1	1.0
0.17	1.0
0.3	2.0
0.55	2.0
1.0	2.0
1.7	1.0

^a Relative amount of sampling used for a given restraint force constant.

TABLE 2: Confinement Protocol for the Argon System

k (kcal/mol/Å ²)	sampling ^b	γ (ps ⁻¹) ^a
0.01	1.0	1
0.03	1.0	1
0.055	2.0	1
0.1	6.0	1
0.17	6.0	1
0.3	3.0	1
0.55	1.0	1
1.0	1.0	1
3.0	1.0	10
10.0	1.0	10
30.0	1.0	10
100.0	1.0	10

^a The Langevin friction coefficient γ was increased at higher restraint constants. ^b Relative amount of sampling used for a given restraint force constant.

restraint protocol used for the fully interacting system is shown in Table 2. Note that a stronger restraint was necessary before convergence was reached, as compared to the ideal gas simulations. Parameters for the relative length of simulation and friction coefficients at different values of confinement force constant were determined by short preliminary runs and chosen to minimize the uncertainties in the raw data obtained. We found that smaller time steps and higher friction coefficients at higher restraint constants gave more precise and convergent results. The linear assignment of atom-to-restraint positions was updated every 10 steps during the simulation.

To compare and validate the results, we also carried out a classic determination of the excess free energy F_e using thermodynamic integration. We used the soft-core protocol of Zacharias et al.,⁴³ used successfully by other workers on argon,^{29,41} to annihilate the argon atoms in up to 44 steps of λ between $\lambda = 0$ and $\lambda = 1$,

$$\phi(r_{ij}, \lambda) = \lambda 4\epsilon \left[\frac{\sigma^{12}}{(r_{ij}^2 + \delta(1 - \lambda))^6} - \frac{\sigma^6}{(r_{ij}^2 + \delta(1 - \lambda))^3} \right] \quad (11)$$

where δ is a shifting parameter that prevents overlap problems at very small values of λ . See refs 43 and 29 for details. Here, we used $\delta = 8$ Å², which gave consistent results and a relatively smooth function for $\partial F/\partial \lambda$ reducing errors during the numerical integration. Various numbers of intermediates along λ (23 and 44) and simulation lengths (20–500 ps per intermediate) were used to determine the computational efficiency and compare it to the method presented here. The data were numerically integrated using four-point quadrature.

Water System. A cubic solvent box of 100 molecules of TIP3P⁴⁴ was created with a box length of 14.410 Å (resulting in a density of 1.00 g/cm³). The temperature was set to 298 K,

TABLE 3: Configurational Excess Free Energies in Reduced Units ($A_c/\epsilon N$) from the Confinement Method Applied to a Periodic Boundary Box of 64 Argon Atoms

$N_{\text{total}} (\times 10^6)^a$	$A_c/\epsilon N$
1.25	-3.0909 ± 0.017 (0.54%)
2.50	-3.0889 ± 0.011 (0.37%)
6.25	-3.0725 ± 0.007 (0.24%)
12.50	-3.0892 ± 0.009 (0.29%)
25.00	-3.0847 ± 0.003 (0.09%)

^a N_{total} designates the total number of simulation steps

TABLE 4: Configurational Excess Free Energies in Reduced Units ($A_c/\epsilon N$) from Thermodynamic Integration Applied to a Box of 64 Argon Atoms

$N_{\text{total}} (\times 10^6)^a$	$A_c/\epsilon N$
0.46	-3.0894 ± 0.032 (1.04%)
1.15	-3.0875 ± 0.016 (0.51%)
2.30	-3.0838 ± 0.011 (0.35%)
4.60	-3.0907 ± 0.007 (0.24%)
8.80 ^b	-3.0869 ± 0.006 (0.20%)
11.50	-3.0860 ± 0.004 (0.14%)
22.00 ^b	-3.0869 ± 0.003 (0.10%)

^a N_{total} designates the total number of simulation steps. ^b Run in 44 steps of λ instead of 23.

TABLE 5: Comparison of Fitting Algorithms for the Confinement of 64 Ideal Gas Particles (F_{id}) and of 64 Liquid Argon Atoms (F_{argon})^a

method	F_{id}/N	F_{argon}/N
linear fit	-1.7497 (6.11%)	-0.5686 (9.41%)
double log/trapezoidal	-1.6432 (0.35%)	-0.5155 (0.81%)
double log/polynomial	-1.6489 (<0.01%)	-0.5191 (0.11%)
analytical result	-1.6490	
thermodynamic integration		-0.5197

^a All figures in kcal/mol; the bias compared to the respective reference result is shown in parentheses.

and simulations were run using Langevin dynamics.³⁸ We used a flexible version of TIP3P, with the bond stretching parameter $k_{\text{OH}} = 1106.0$ kcal/mol/Å² and the angle bending term $k_{\theta} = 200$ kcal/mol/rad².

In the fully interacting system, nonbonded interactions were cut off at 7.0 Å, and a force-switching function⁴⁵ was applied from 5.0 to 7.0 Å. No long-range corrections were applied. The electrostatic cutoffs are very short, but the purpose of the simulations here was to examine the methodology rather than obtain results that match experimental data.

Preceding the confinement, the water box was equilibrated by 150 ps steps of Langevin dynamics with exponential cooling from 500 to 300 K, followed by minimization until the energy gradient was negligible. This procedure allowed the unequilibrated system to obtain a reasonably equilibrated water structure and generated a well-minimized system to act as the restraint reference structure.

Two types of runs were conducted. In the first ($\eta = 1.0$), restraints on the hydrogens were not attenuated relative to those on the oxygen atoms, and the assignment was updated every 10 steps. In the second type of run ($\eta = 0.2$), we reduced the restraint on the hydrogens by a factor of 5 and updated the assignment more frequently (every 2 steps). Preliminary experiments showed that the latter simulation parameters led to a considerably better energy conservation while affecting computational efficiency minimally. For each of the simulation conditions, we ran two simulations, one with symmetry swapping (sym) enabled and one without (nosym), giving a total of four run types. Two different protocols were used for the

TABLE 6: Confinement Protocols for Restraining 100 TIP3P Molecules

k (kcal/mol/Å ²)	t (fs) ^a	protocol A		protocol B	
		sample ^b	γ (ps ⁻¹)	sample ^b	γ (ps ⁻¹)
00.03	1.0	1.0	10		
0.055	1.0	1.0	10	0.5	10
0.1	1.0	3.0	10	1.0	10
0.17	1.0	5.0	10	2.0	10
0.3	1.0	8.0	10	4.0	10
0.55	1.0	5.0	10	6.0	10
1.0	1.0	3.0	10	6.0	10
1.7	1.0	1.0	10	4.0	10
3.0	1.0	1.0	10	2.0	10
5.5	1.0	1.0	10	1.0	10
10.0	0.5	1.0	10	1.0	10
17.0	0.5	1.0	20	1.0	20
30.0	0.5	1.0	20	1.0	20
55.0	0.5	1.0	30	1.0	50
100.0	0.5	1.0	50	1.0	50
200.0	0.5	1.0	50	1.0	100

^a The integration time step t and the Langevin friction coefficient γ were adjusted at higher restraint constants. ^b Relative amount of sampling used for a given restraint force constant.

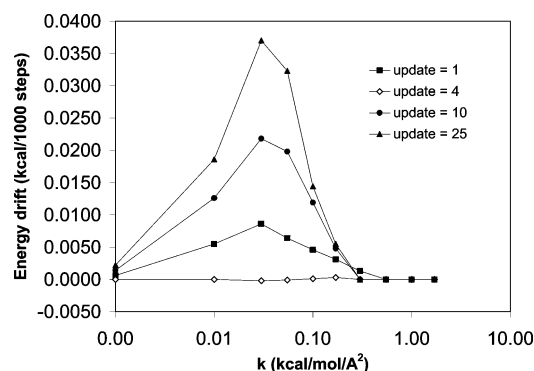


Figure 3. Potential energy drift in a constant energy simulation of the ideal gas system at different restraint constants and different restraint update frequencies.

simulations with and without hydrogen attenuation (see Table 6), because the restraint reduction causes the system to solidify later in the protocol. Adjusting the protocols accordingly reduces the uncertainty in the results. Runs were repeated 15 times to

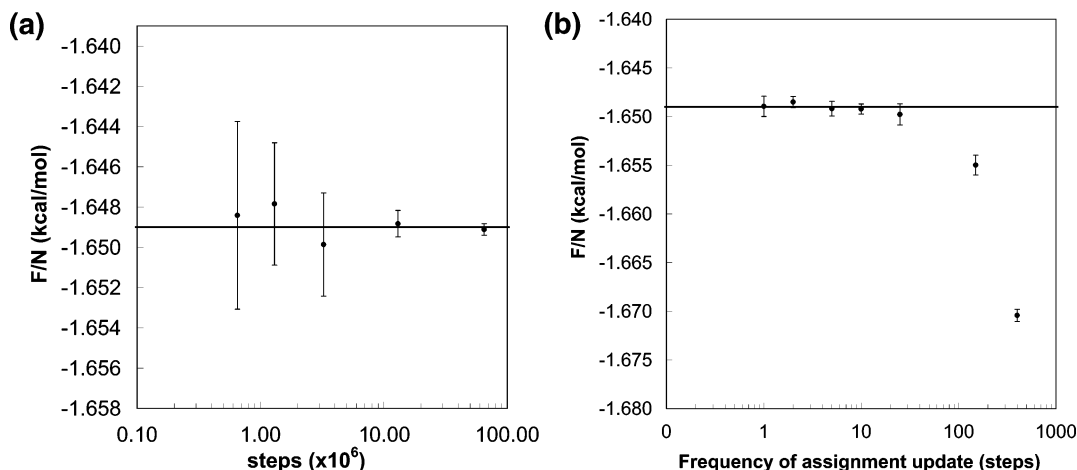


Figure 4. (a) Absolute Helmholtz free energies, F/N , calculated by confining 64 ideal-gas particles with different lengths of simulations. The fluctuation in the results observed (the standard deviation) over 10 repeat runs is indicated by the error bars. The correct analytical result is indicated by the thick horizontal line at $F/N = -1.649$ kcal/mol. The result from the confinement method can be seen to converge very precisely to the correct theoretical value with increasing simulation lengths. (b) Analysis of the bias introduced by not updating the assignment at every step. Each run is equivalent to that of the right-most point in part a, but run at various intervals between assignment updates. The results indicate that updating as rarely as once every 25 steps introduces negligible bias for this idealized system.

obtain the statistical uncertainty. Throughout this work, the first 40% of the data from each simulation was ignored to allow for equilibration.

As in the argon system, we also set up thermodynamic integration calculations to obtain the excess free energy. Here, we used the protocol of Quintana et al.,¹⁸ which annihilates the nonbonded interactions as

$$\phi(r_{ij}, \lambda) = \lambda^5 4\epsilon \left[\frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^6}{r_{ij}^6} \right] + \frac{\lambda^3}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}} \quad (12)$$

The thermodynamic integration data were integrated using a four-point quadrature.

2.5. Numerical Integration. In previous work,³⁰ we showed that the linear interpolation of the data transformed into double log space gives considerably more accurate results than applying a trapezoidal integration procedure directly to the untransformed data. This works extremely well at higher restraint constants, but at lower restraint strengths, the transformed data has curvature, and thus, simply connecting successive data points linearly introduces some error (usually underestimating the free energy).

Here, we improve the interpolation procedure further by using third-order Lagrange interpolating polynomials. The procedure works as follows. The raw data to be numerically integrated ($y = \langle w(\mathbf{r}) \rangle$, $x = k$, see eq 2) is transformed into double log space. Third-order Lagrange interpolating polynomials are then calculated to consecutive sets of four data points. The interpolation between the central two points is used to estimate the curve. This is then transformed back into the original space to perform the integration. We found this gives excellent results with fewer integration points, improving the convergence of the free energy at high restraint constants and accounting for the curvature at low force constants. Table 5 compares the accuracy of the three different fitting approaches: direct trapezoidal interpolation, trapezoidal integration in double-log space,³⁰ and the polynomial integration in double-log space.

3. Results and Discussion

3.1. Ideal Argon Gas. Prior to the determination of the free energy, we tested energy conservation in light of the method's

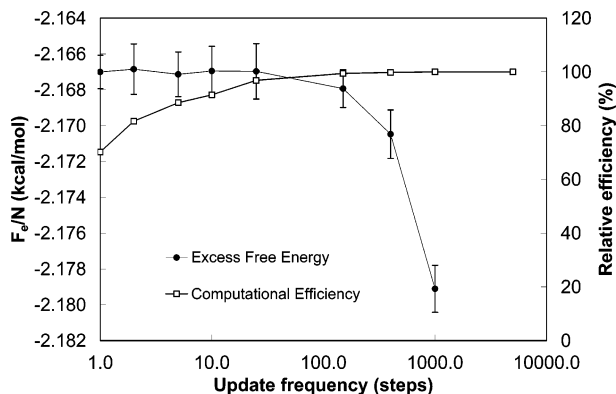


Figure 5. Bias of the calculated free energy of 64 argon atoms and efficiency of calculation (relative to normal MD) for different assignment updates. More frequent updates cause a smaller bias but are computationally more expensive. Updating every 10–25 steps appears to introduce a minimal bias while affecting the efficiency only slightly. Note that with molecular fluids such as water, the relative efficiency improves markedly as the computational load shifts toward the nonbonded calculation. In the TIP3P system examined in this paper, updating the assignment as often as every two steps incurred a mere 5% performance loss.

force discontinuities, discussed in the Methods. The drift in total energy over a 20 ps NVE trajectory (using Beeman integration⁴⁶) is plotted in Figure 3 for different assignment update frequencies. Although some drift is observed at intermediate force constants as expected, the drifts are very small and should not cause problems during a stochastic simulation. Energy conservation improves with more frequent updating of the assignment, although interestingly, the optimal energy conservation is observed around an update frequency of 4. During the stochastic simulation (especially in the absence of intermolecular forces), the optimal assignment may fluctuate rapidly; hence, updating the force directions too frequently may unnecessarily exacerbate the force discontinuities.

The absolute free energy of the system of 64 ideal argon atoms without intermolecular potentials can be obtained from eq 4 and is $F = -105.536$ kcal/mol for the entire system, or $F/N = -1.649$ kcal/mol per atom. The results from the confinement simulations are shown in Figure 4a with the uncertainty indicated by the error bars. Shown as a thick line is the analytical result. The results indicate an excellent convergence toward the correct answer with increasing simulation length with no apparent bias in the result.

Figure 4b shows the dependence of the calculated free energy F/N on the frequency of the assignment update. As expected, a negative bias is introduced when the assignment is updated less frequently than every 25 steps. At any time step when the assignment is out of date, the restraint function must always be larger than it would be if the assignment were updated (since by definition, the linear assignment algorithm always gives the arrangement of atom-to-restraint positions with the lowest possible restraint function). Thus, less frequent assignment updates lead, on average, to an overestimation of ΔF_{conf} and, thus, an underestimation of F . However, the results show that for update frequencies of more than every 25 steps, the results are only very marginally biased and, thus, acceptable in practice. In the following simulations, we used an update frequency of 10, giving a good tradeoff between computational speed and accuracy for real systems (see Figure 5).

3.2. Full Argon System. The results for the calculation of the excess free energy of argon are presented in Tables 3 and 4 and show that the results from the confinement method are

in excellent agreement with those from thermodynamic integration. For consistency and comparison with previous studies,^{27,29,47} the results are presented in the form of configurational free energy, A_c .

$$A_c = F_e - k_B T \ln \left(\frac{V^N}{N! \sigma^{3N}} \right)$$

As mentioned in the Methods section, the fluctuations were obtained by repeating the calculations with 10 different random number seeds and 10 different restraint positions. The fluctuations (variance) observed in the final free energy result are much smaller than the variations in the potential energy of the restraint structure, that is, the excess free energy obtained is independent of the chosen precise structure the free system is confined to. We generally observed however, that the restraint structure should be well-minimized (i.e., be at a local minimum) for the results to converge more quickly.

We found that the confinement method and a traditional annihilation approach give results of comparable uncertainty for a given amount of computer time. This is highly encouraging, since the direct absolute free-energy approach will have advantages over an annihilation approach in more complex situations, such as explicit simulations of large solutes.

It appears that in the argon case, the system does not become trapped in local minima, and thus, straightforward MD is sufficient to obtain highly accurate results. Because of the assignment “trick”, each atom essentially only has to explore its immediately surrounding space,³² in which there are hardly any barriers to cross.

As with the ideal gas system, we checked the energy conservation of the argon system to evaluate the impact of the force discontinuity in a real system. We found that the drift in total energy during an NVE simulation (using the Beeman algorithm⁴⁶) was less than 0.0003 kcal/1000 steps for all simulations, and assignment update frequencies were < 150 . This is *much lower* than in the ideal gas system, and the improvement can be rationalized by the fact that in the argon simulation, due to the excluded volume of each particle, the assignment changes much more slowly, and thus, the impact on energy conservation is minimized, even at relatively infrequent assignment updates. In summary, in both the ideal gas system and the argon system, the force discontinuities are completely negligible, leading to no significant bias in the results, as can be seen from the excellent agreement to analytical calculations and thermodynamic integration, respectively.

Because we were primarily interested in demonstrating self-consistency of our method rather than obtaining the excess free energy of argon as such, we chose different simulation parameters for our simulation as compared to other studies.^{27,29,47} Instead of a sharp cutoff, we used a switching function to taper out the van der Waals interactions from 6.0 to 7.15 Å, and we did not add any long-range corrections. Long-range corrections (E_{LRC}) can be added by analytically integrating the potential to an infinite distance,⁴⁸

$$E_{\text{LRC}} = 2N\pi\rho \int_{6.0}^{\infty} E_{\text{vdw}}(r)(1 - S(r))r^2 dr \quad (13)$$

where N is the number of particles, ρ is the particle density, $S(r)$ is the switching function, and $E_{\text{vdw}}(r)$ is the usual van der Waals pair energy function with parameters for argon. The correction for the argon system calculated using this method is $A_c^{\text{LRC}}/\epsilon N = -1.036$, and adding this to our result gives $A_c/\epsilon N =$

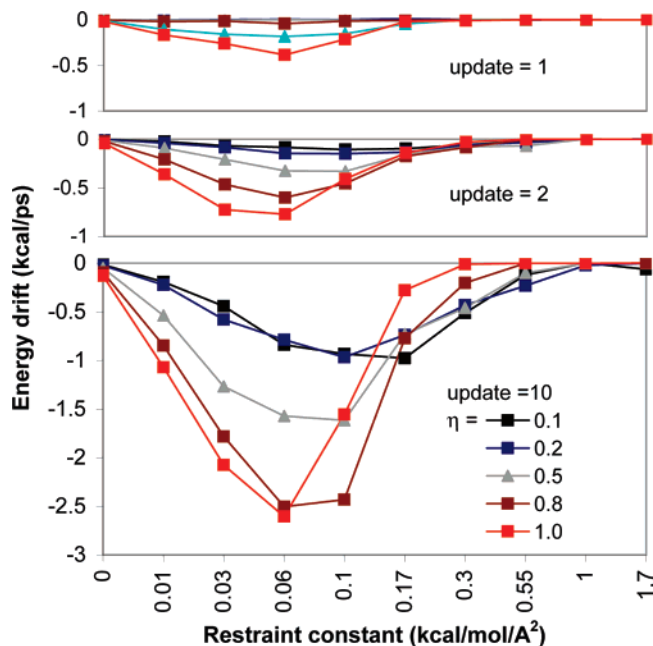


Figure 6. Energy drift of the ideal TIP3P system as a function of restraint force constant at different update frequencies (update) and hydrogen attenuations (η). For all update frequencies (update $\in \{1, 2, 10\}$), the attenuation does to some extent reduce the energy drift, in some cases substantially or even completely. It is interesting that when the reassignment is very frequent (on every step), an attenuation of merely 0.7 is sufficient to essentially remove any energy drift. In contrast, when the assignment is less frequent, even complete attenuation ($\eta = 0.0$, data not shown) does not completely remove the energy drift.

-4.11 ± 0.02 , which agrees well with published results on very similar setups ($A_c/\epsilon N \approx 4.1$).^{27,29,47}

Figure 5 shows the bias in the excess free energy, A_c , introduced by less frequent updates of the assignment as well as the associated computational cost relative to an ordinary MD simulation. The results show that updating every 5 or 25 steps introduces negligible bias while reducing the computational efficiency only marginally. Note that in practice, the reassignment only needs to be carried out when the restraint force constant, k , is relatively low, that is, when solvent molecules have enough thermal energy to escape from their restraint wells and diffuse around the simulation box. At higher restraint constants, the molecules are so tightly restrained to their positions that they never acquire enough energy to jump into the next well, and thus, the assignment never changes in practice. We observed that this situation occurs in simulations at $k > 0.55$ kcal/mol/Å² for the argon system. Because the assignment does not change, the reassignment is not required for these simulations (comprising 25% of the total simulation), and an ordinary harmonic restraint can be used instead, effectively further reducing the additional cost. Thus, in total, at an update frequency of 10 steps, the overall efficiency was close to 93%. This will, of course, further improve with molecular fluids, as opposed to atomic fluids, since more computational time will be spent on the nonbonded interactions, as compared with the reassignment as the ratio of $N_{\text{atoms}}/N_{\text{molecules}}$ increases.

3.3. Ideal Water System. Before testing our method on an interacting water system, we tested energy conservation in an ideal, noninteracting water system during a short NVE trajectory to assess the severity of the force discontinuities. It was found that the energy conservation is considerably worse than in the monoatomic fluids. The assignment is sensitive to the orientation of the molecules, which changes much more rapidly than their

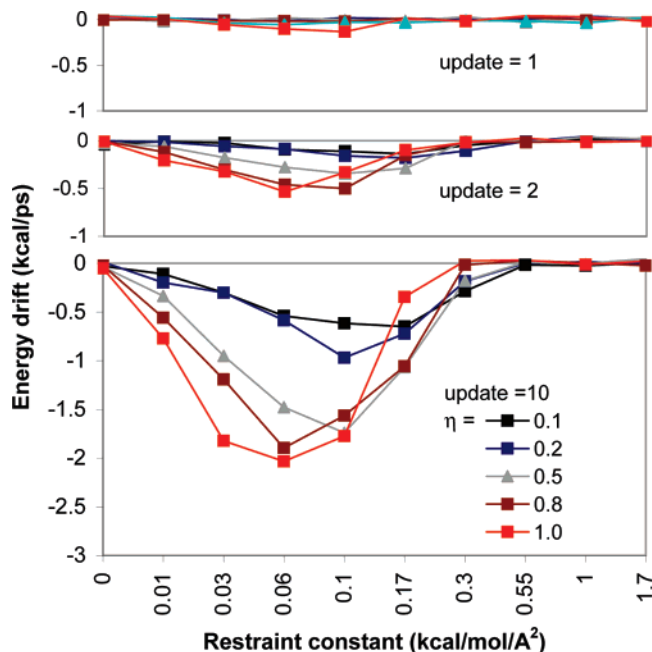


Figure 7. Energy drift of the real TIP3P system as a function of restraint force constant at different update frequencies (update) and hydrogen attenuations (η), analogous to Figure 6. See text for further discussion.

positions. This causes the assignment to change more frequently, therefore exacerbating the effects of the force discontinuity. Interestingly, we found that reducing the restraint strength on the hydrogens relative to the oxygen atom by the factor $\eta < 1.0$ improves the energy conservation markedly (Figure 6). This attenuation delays removal of rotational freedom, as compared to translational freedom, such that they no longer occur concurrently. This effect reduces the dependence of the assignment on the orientation of the molecules. In other words, the molecules first feel merely a restraint on the oxygen restraining their position, and thus, reassignment during the translationally free phase has a reduced effect on the hydrogens, hence, minimizing the discontinuity problem. Later, the restraint on the hydrogens becomes significant, and the molecules are prevented from rotating freely, but by this stage, the molecules have already forfeited their freedom to diffuse, and reassignment no longer occurs.

The absolute free energy of the ideal water system can be calculated from the ideal gas and intermolecular contributions to the partition function. The analytical value of the absolute free energy, F , of the idealized water system was calculated from eq 8. The principal moments of inertia and normal modes were obtained from the ground state structure of TIP3P, giving $I_A I_B I_C = 5.785\,207 \times 10^{-141}$ kg³ m⁶ and three eigen-frequencies at 2339.353 cm⁻¹, 3682.799 cm⁻¹, and 3733.828 cm⁻¹. The contributions to the partition functions at $T = 298$ K are, thus,

$$\ln(q_{\text{trans}}) = 12.307\,17 \quad (\text{see eq 4})$$

$$\ln(q_{\text{rot}}) = 4.452\,06 \quad (\text{see eq 9})$$

$$\ln(q_{\text{vib}}) = -8.194\,96 \quad (\text{see eq 10})$$

Note that the negative value of the logarithm of the vibrational partition function stems from the fact that a classical partition function is used here. Because the simulation is classical (molecular dynamics), the corresponding thermodynamic expressions must also be classical. Including the correction of $1/N!$

TABLE 7: Free Energies from Thermodynamic Integration for 100 TIP3P Molecules

$N_{\text{total}} (\times 10^6)$	F_e/N (kcal/mol)
1.1	-6.259 ± 0.0214 (0.34%)
2.2	-6.263 ± 0.0064 (0.10%)
4.4	-6.259 ± 0.0073 (0.12%)
11.0	-6.261 ± 0.0024 (0.04%)

TABLE 8: Results from the Confinement Calculations on a Box of 100 TIP3P Molecules

simulation	$N_{\text{total}} (\times 10^6)$	F_e/N (kcal/mol)
nosym, $\eta = 1.0$	1.625	-6.206 ± 0.027 (0.4%)
	3.250	-6.251 ± 0.018 (0.3%)
	8.125	-6.248 ± 0.020 (0.3%)
sym, $\eta = 1.0$	1.625	-6.235 ± 0.045 (0.7%)
	3.250	-6.265 ± 0.021 (0.3%)
	8.125	-6.255 ± 0.019 (0.3%)
nosym, $\eta = 0.2$	1.625	-6.156 ± 0.033 (0.5%)
	3.250	-6.183 ± 0.021 (0.3%)
	8.125	-6.211 ± 0.021 (0.3%)
sym, $\eta = 0.2$	1.625	-6.224 ± 0.022 (0.4%)
	3.250	-6.236 ± 0.017 (0.3%)
	8.125	-6.233 ± 0.012 (0.2%)

to account for the fact that the molecules as a whole are unlabeled, we obtain a free energy of 100 idealized TIP3P molecules in the above volume of 14.41 \AA^3 from $F_{\text{id}}^{\text{nosym}} = -Nk_B T \ln((1/N!)q_{\text{trans}}q_{\text{rot}}q_{\text{vib}})$. When the symmetry switch is used, then a factor of 1/2 needs to be included in the rotational partition function due to the removal of the label on the hydrogens. The results obtained from the confinement method of the normal ($F_{\text{cm}}^{\text{nosym}}$) and symmetry-swapped system ($F_{\text{cm}}^{\text{sym}}$), as well as their analytical values, are shown below.

type	numerical	analytical
$F_{\text{cm}}^{\text{nosym}}/N$	-2.907 ± 0.003	-2.9155
$F_{\text{cm}}^{\text{sym}}/N$	-2.501 ± 0.002	-2.5053

all figures in kcal/mol

The results obtained are very close to their expected analytical values. However, in both cases, we observe a very small bias, $\approx 0.3\%$ and $\approx 0.1\%$ for $F_{\text{cm}}^{\text{nosym}}$ and $F_{\text{cm}}^{\text{sym}}$, respectively. Although the obvious candidate for the source of the bias is the discontinuous nature of the forces, we found that the bias is quite independent of the attenuation factor, which improves the energy conservation considerably.

3.4. Full Water System. The energy conservation for the fully interacting water system gave results similar to those obtained on the ideal water system, and we observed that energy conservation is restored when the assignment is updated frequently and the restraint on the hydrogens is reduced. Some preliminary experimentation showed that the energy conservation is worsened by the low mass of the hydrogens and the lack of van der Waals forces on them.

The excess free energies from from the traditional thermodynamic integration calculations and the confinement simulations and their uncertainties are presented in Table 7 and Table 8, respectively. Excellent agreement is achieved between the two computational methods.

Although attenuation of the restraint strength on the hydrogens improves energy conservation, the respective free-energy estimates are slightly lower than the results from thermodynamic integration. This effect is most likely due to remaining anharmonicities in the final restraint state, since the final restraint constant on the hydrogen atoms is only $40 \text{ kcal/mol/\AA}^2$. The swapping of the hydrogen identities (sym/nosym) appears to improve the sampling marginally, with the effect considerably

pronounced in the simulations with $\eta = 0.2$. Here, the simulations with symmetry swap (sym) appear to converge much faster to the correct answer than those without. In these simulations, the molecules surrender their translational freedom before they are rotationally restrained. This can lead to sampling problems for rotational states when the system has already entered a solid state due to the translational restraint. The effect can be expected to be particularly pronounced in a highly polar liquid, such as water. The symmetry swap removes much of the need for sampling rotational states, thereby improving performance. In contrast, when translational and rotational freedoms are restrained concurrently, the effect of the symmetry swap is smaller.

The TIP3P simulations were benchmarked, and the performance loss due to the reassignment, compared with ordinary MD simulation, was smaller than in the argon system. In the TIP3P system, there are around 6 times more particles in the nonbonded calculation but only 1.5 times as many molecules to be reassigned, shifting the computational burden on the former. This means that updating the assignment as frequently as every two steps incurs only a 5% reduction in performance. This figure would improve even further if a large solute, such as a protein were included in the system.

The computational efficiency compared to TI was somewhat reduced, with results of comparable uncertainty requiring 2–4 times more simulation time. Analysis of the simulations showed that the data between $k = 0.3 \text{ kcal/mol/\AA}^2$ and $k = 1.7 \text{ kcal/mol/\AA}^2$ is contributing more than 90% of the uncertainty in the final result, despite the fact that 60% of the simulation time is spent here. In this part of the protocol, the liquid is forced into a phase transition due to the restraint; the relaxation time becomes very large; and thus, sampling becomes inefficient due to the glassy behavior of the fluid. One possible solution would be to use an improved sampling method in the critical part of the restraint protocol. For example, replica exchange MD could be used in place of normal MD, as was done previously for peptide systems.³⁰ We have not yet pursued this possibility, but there is evidence that REMD simulation does not improve sampling on glassy systems.⁴⁹ An alternative solution would be to take a different path from liquid to solid in which the transition problem is reduced. For example, during the argon simulation, although a phase transition was evident, as well, it was possible to sample the ensemble averages much more accurately. This indicates that it is the strong electrostatics that are responsible for the sampling problem in the water system. It is possible that a two-stage protocol could be superior, in which the electrostatics of the system are first removed using TI and the remaining van der Waals liquid are confined to obtain an absolute free energy.

4. Conclusions

We have demonstrated how the confinement method can be applied to liquids. This is an important extension of the method, since it makes it applicable to explicitly solvated protein or nucleic acid simulations of central importance in biochemical simulation. The method was applied successfully to monoatomic fluids with extremely good precision. Encouragingly, the confinement method appears to give results of roughly the same precision as thermodynamic integration for a comparable amount of computing effort. This compares favorably with recent studies on absolute free-energy calculations which report being computationally less efficient than TI by 3–8 fold, depending on the system.^{27,28} The excess free energy of water was also successfully calculated, but here, the method proved less

efficient than TI by a factor of around 2–4. Problems were encountered with the phase transition, causing sampling problems, and possible solutions have been identified that will be pursued in future work. Overall, the extension of the method to explicitly solvated, multi-molecular systems has been surprisingly successful, given the concerns about the radical nature of the new self-reassigning restraint potential.

In contrast to other methods for the calculation of absolute free energies,^{20,27–29} the confinement method is relatively straightforward to implement and could be integrated easily into other existing molecular mechanics packages. Further, the confinement approach does not involve any approximations to the energy surface^{21,22,31,50,51} and, thus, gives the exact free energy in the sampling limit. Because of the staged, independent simulations, it is amenable to trivial parallelization across computer clusters and additionally to any further parallelization within the simulation.

We regard this work on fluids to be an important milestone on the road to calculating absolute free energies of explicitly solvated polypeptides and other solutes.

Acknowledgment. MDT gratefully acknowledges funding support from the Department of Biochemistry, Bristol, and the authors thank Rhiju Das for helpful comments on the manuscript. We also thank Graeme Cappi for helping to install our distributed computing resources used for this work and all the voluntary contributors.

References and Notes

- (1) Jorgensen, W. L.; Ravimohan, C. *J. Chem. Phys.* **1985**, *83*, 3050–3054.
- (2) Miller, J. L.; Kollman, P. A. *J. Phys. Chem. A* **1996**, *100*, 8587–8594.
- (3) Shirts, M. R.; Pande, V. S. *J. Chem. Phys.* **2005**, *122*, 134508.
- (4) Simonson, T.; Archontis, G.; Karplus, M. *Acc. Chem. Res.* **2002**, *35*, 430–437.
- (5) Bash, P. A.; Singh, U. C.; Brown, F. K.; Langridge, R.; Kollman, P. A. *Science* **1987**, *235*, 574–576.
- (6) Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. *Biophys. J.* **1997**, *72*, 1047–1069.
- (7) Archontis, G.; Simonson, T.; Karplus, M. *J. Mol. Biol.* **2001**, *306*, 307–327.
- (8) Oostenbrink, C.; van Gunsteren, W. F. *Proteins* **2004**, *54*, 237.
- (9) Boczeko, E. M.; Brooks, C. L., III. *J. Phys. Chem.* **1993**, *97*, 4509–4513.
- (10) Woo, H. J.; Roux, B. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6825–6830.
- (11) Ravindranathan, K. P.; Gallicchio, E.; Levy, R. M. *J. Mol. Biol.* **2005**, *353* (1), 196–210.
- (12) Hoover, G.; Gray, S. G.; Johnson, K. W. *J. Chem. Phys.* **1971**, *55*, 1128.
- (13) Frenkel, D.; Ladd, A. J. C. *J. Chem. Phys.* **1984**, *81*, 3188.
- (14) Stoessel, J. P.; Nowak, P. *Macromolecules* **1990**, *23*, 1961–1965.
- (15) Levesque, D.; Verlet, L. *Phys. Rev.* **1969**, *182*, 307–316.
- (16) Hoover, W. G.; Ross, M.; Johnson, K. W.; Henderson, D.; Barker, J. A.; Brown, B. C. *J. Chem. Phys.* **1970**, *52*, 4931–4941.
- (17) Patey, G. N.; Valleau, J. *Chem. Phys. Lett.* **1973**, *21*, 297–300.
- (18) Quintana, J.; Haymet, A. *Chem. Phys. Lett.* **1992**, *189*, 273–278.
- (19) Hermans, J.; Pathiaseril, A.; Anderson, A. *J. Am. Chem. Soc.* **1988**, *110*, 5982.
- (20) Ytreberg, F. M.; Zuckerman, D. M. *J. Chem. Phys.* **2006**, *124*, 104105.
- (21) Henchman, R. H. *J. Chem. Phys.* **2003**, *119*, 400–406.
- (22) Roccatano, D.; Amadei, A.; Apol, M. F. E.; Di Nola, A.; Berendsen, H. J. C. *J. Chem. Phys.* **1998**, *109*, 6358–6363.
- (23) Head, M. S.; Given, J. A.; Gilson, M. K. *J. Phys. Chem. A* **1997**, *101*, 1609–1618.
- (24) Krivov, S. V.; Karplus, M. *J. Chem. Phys.* **2002**, *117* (23), 10894.
- (25) Evans, D. A.; Wales, D. J. *J. Chem. Phys.* **2003**, *119*, 9947–9955.
- (26) Cheluvraja, S.; Meirovitch, H. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 9241–9246.
- (27) White, R.; Meirovitch, H. *J. Chem. Phys.* **2006**, *124*, 204108.
- (28) White, R.; Meirovitch, H. *J. Chem. Phys.* **2004**, *121*, 10889–10904.
- (29) White, R.; Meirovitch, H. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 9235–9240.
- (30) Tyka, M. D.; Clarke, A. R.; Sessions, R. B. *J. Phys. Chem. B* **2006**, *110*, 17212–17220.
- (31) Reinhard, F.; Grubmüller, H. *J. Chem. Phys.* **2007**, *126*, 014102.
- (32) Stillinger, F. H.; Weber, T. A. *Phys. Rev. A: At., Mol., Opt. Phys.* **1982**, *25*, 978–989.
- (33) Kuhn, H. W. *Nav. Res. Logistics Q.* **1955**, *2*, 83–87.
- (34) Kuhn, H. W. *Nav. Res. Logistics Q.* **1956**, *3*, 253–258.
- (35) Munkres, J. J. *Soc. Ind. Appl. Math.* **1957**, *5*, 32–38.
- (36) Jonker, R.; Volgenant, A. *Computing* **1987**, *38*, 325–340.
- (37) Tyka, M. D.; Rea, J.; Clarke, A. R.; Sessions, R. B. Manuscript in preparation.
- (38) Paterlini, M. G.; Ferguson, D. M. *Chem. Phys.* **1998**, *236*, 243–252.
- (39) Michels, A.; Wijker, H.; Wijker, H. *Physica (Amsterdam)* **1949**, *15*, 627–633.
- (40) Hansen, J. P.; Verlet, L. *Phys. Rev.* **1969**, *184*, 151.
- (41) Szarecka, A.; White, R.; Meirovitch, H. *J. Chem. Phys.* **2003**, *119*, 12084–12095.
- (42) Steinbach, P.; Brooks, B. *J. Comput. Chem.* **1994**, *15*, 667–683.
- (43) Zacharias, M.; Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1994**, *100*, 9025–9031.
- (44) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. *J. Chem. Phys.* **1983**, *79*, 926.
- (45) Steinbach, P. J.; Brooks, B. R. *J. Comput. Chem.* **1994**, *15*, 667–683.
- (46) Beeman, D. *J. Comput. Phys.* **1976**, *20*, 130–139.
- (47) Li, Z.; Scheraga, H. A. *J. Phys. Chem.* **1988**, *92*, 2633–2636.
- (48) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press: New York, 1987.
- (49) De Michele, C.; Sciortino, F. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **2002**, *65*, 051202.
- (50) Karplus, M.; Kushick, J. *Macromolecules* **1981**, *14*, 325–332.
- (51) Levy, R. M.; Karplus, M.; Kushick, J.; Perahia, D. *Macromolecules* **1984**, *17*, 1370.