Absolute free energy calculations by thermodynamic integration in four spatial dimensions

Tomas Rodinger, P. Lynne Howell, and Régis Pomès

Citation: J. Chem. Phys. **123**, 034104 (2005); doi: 10.1063/1.1946750
View online: http://dx.doi.org/10.1063/1.1946750
View Table of Contents: http://jcp.aip.org/resource/1/JCPSA6/v123/i3
Published by the American Institute of Physics.

Absolute free energy calculations by thermodynamic integration in four spatial dimensions

Tomas Rodinger
Structural Biology and Biochemistry, The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada; Department of Biochemistry, University of Toronto, Ontario, Canada; and Institute of Biomaterials and Biomedical Engineering, University of Toronto, Ontario, Canada

P. Lynne Howell and Régis Pomès[a)]
Structural Biology and Biochemistry, The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada; and Department of Biochemistry, University of Toronto, Ontario, Canada

(Received 8 March 2004; accepted 10 May 2005; published online 25 July 2005)

An optimized technique for calculating the excess chemical potential of small molecules in dense liquids and the binding affinity of molecular ligands to biomolecules is reported. In this method, a molecular species is coupled to the system of interest via a nonphysical fourth spatial dimension w through which insertion or extraction can be carried out [R. Pomès, E. Eisenmesser, C. B. Post *et al.*, J. Chem. Phys. **111**, 3387 (1999)]. Molecular simulations are used to compute the potential of mean force (PMF) acting on the solute molecule in the fourth dimension. The excess chemical potential of that molecule is obtained as the difference in the PMF between fully coupled and fully decoupled systems. The simplicity, efficiency, and generality of the method are demonstrated for the calculation of the hydration free energies of water and methanol as well as sodium, cesium, and chloride ions. A significant advantage over other methods is that the 4D-PMF approach provides a single effective and general route for decoupling all nonbonded interactions (i.e., both Lennard-Jones and Coulombic) at once for both neutral and charged solutes. Direct calculation of the mean force from thermodynamic integration is shown to be more computationally efficient than calculating the PMF from umbrella sampling. Statistical error analysis suggests a simple strategy for optimizing sampling. The detailed analysis of systematic errors arising from the truncation of Coulombic interactions in a solvent droplet of finite size leads to straightforward corrections to ionic hydration free energies. © *2005 American Institute of Physics*. [DOI: 10.1063/1.1946750]

I. INTRODUCTION

The rapid advancement of computer technology has stimulated the development of simulation tools designed to calculate the excess chemical potentials of small molecules in dense fluids[1] and the binding affinities of molecular ligands to biomolecules.[2–6] The free energy difference between two states of a system can be obtained by calculating the reversible thermodynamic work in a transformation from one state to the other. The choice of a suitable pathway along which the transformation can be made has long been debated.[7] Because free energy is a state function, this pathway need not be, and for practical reasons, generally is not chosen to be physical. In the realm of biomolecular simulations, the two most widely used methods by which free energy changes have traditionally been evaluated are free energy perturbation[8] and thermodynamic integration.[9] In both cases, the Hamiltonian (or potential energy function of the system) H is altered to include a dependence on a coupling parameter λ; by forcing λ from 0 to 1, the initial Hamiltonian governing the state of the system is transformed into the final Hamiltonian. In free energy perturbation, the free energy difference between two endpoint states is evaluated as

$$\Delta G = -\sum_{i=0}^{n-1} k_B T \ln\left\langle \exp\frac{-[H(\lambda_{i+1}) - H(\lambda_i)]}{k_B T}\right\rangle_{\lambda_i}, \tag{1}$$

and in thermodynamic integration, the equivalent calculation can be expressed as

$$\Delta G = \sum_{i=0}^{n-1} \left\langle \frac{\partial H}{\partial \lambda}\right\rangle_{\lambda_i} (\lambda_{i+1} - \lambda_i), \tag{2}$$

where k_B and T are the Boltzmann constant and the absolute temperature, respectively, and i represents an index for n discretely chosen values of λ in the range of 0–1. In this formalism, the overall free energy change is evaluated as a sum of individual free energy contributions from separate simulations carried out at each i. The nature of the dependence of the Hamiltonian on λ defines the path of the transformation and has been a subject of much research as described below.

Transforming the Hamiltonian through scaling by the coupling parameter λ, or by some function of it,[10,11] suffers from a long-recognized statistical problem associated with the repulsive part of the Lennard-Jones potential, namely, the numerical instability due to the presence of an infinitely small and infinitely repulsive body at small values of λ. This is sometimes referred to as the endpoint catastrophe. The

[a)]Author to whom correspondence should be addressed; FAX: 416-813-5022. Electronic mail: pomes@sickkids.ca

simplest method of addressing this problem is by not simulating at λ values near the endpoint states, but rather attaining the results at the endpoints through extrapolation.[12] Another approach is to embed the repulsive core of an appearing (or disappearing) particle inside the radius of an adjacent atom by "sprouting" (or "desprouting") an associated bond.[13] However, this becomes nontrivial for complex molecules. Alternatively, the instability problem can be completely bypassed by incorporating a λ-dependent "shifting" or "translation" term to the effective interatomic distance of any Lennard-Jones or Coulombic interaction that is being annihilated or created in the transformation.[11,14] In that scheme, the infinitely repulsive body (or "hard core") is replaced by a potential energy barrier of finite magnitude (or "soft core") which smoothly diminishes to zero in the limit of complete decoupling.

Other notable means of overcoming the singularity problem include the particle-insertion method by Widom[15] (which was designed to calculate the solvation free energies of small particles), and the one-step perturbation method[8] (which is limited to evaluating the free energy of conservative changes). For a detailed introduction to the modern techniques used for calculating free energies by molecular simulation, see Ref. 5.

An advantage afforded by soft-core potentials over conventional methods[4,5] is that they are not limited to conservative changes between similar solutes. A systematic comparison of six scaling schemes has shown that the only reliable and efficient protocols, which do not involve shrinking/expanding bonds, are those employing soft-core scaling.[7] Accordingly, soft-core potentials have become widely used in absolute free energy calculations. In early applications, the approach has been applied to the calculation of the hydration free energy of neon and ethanol,[11] to an atomic fluid,[16] and to the calculation of the binding free energy of camphor to cytochrome P450.[17] Examples of recent applications include the calculation of the free energy of association of a cyclophane-pyrene complex,[18] the binding affinity of both natural ligands and xenoestrogens to the estrogen receptor binding domain,[19] and the absolute hydration free energies of 15 amino acid side chain analogs.[20] Although soft-core methods effectively overcome the singularity problem, the fact that they rely on scaling of the Hamiltonian means that consideration must be given to the selection of appropriate scaling constants for both Lennard-Jones and Coulombic nonbonded interactions as well as the functional form of the scaling (linear, quadratic, or otherwise).[7,11,21]

A method departing from the coupling parameter approach was proposed by Pomès *et al.* whereby an unphysical fourth spatial dimension is added to the conformational space of the system.[22] The extension of physical space by an extra dimension was originally proposed for efficient conformational sampling of proteins[23] and atomic liquids.[16] This idea was combined with an approach introduced by Kong and Brooks,[24] where the coupling parameter is treated as a degree of freedom of the system rather than as a preset parameter, to yield the four-dimensional potential of mean force (4D-PMF) method.[22] In the 4D-PMF approach, the coupling between a molecular species (the "solute") and its environment (the "solvent") is modulated by a nonphysical degree of freedom: their spatial separation in the fourth dimension, w_u. The free energy change for the extraction (or insertion) of a molecular solute from (or into) a system of interest corresponds to the difference between fully coupled ($w_u=0$) and fully decoupled ($w_u=$infinity) states, and can be obtained from the reversible thermodynamic work or PMF along w_u. The PMF can be computed by umbrella sampling simulations, where biasing potentials [or "umbrellas" (Ref. 25)] are used to enforce the desired sampling profile along the transformation coordinate.[22]

The 4D-PMF method, like other soft-core methods, is not limited to applications involving conservative changes between similar solutes. Instead, the approach makes it possible to compute the absolute solvation free energy of an entire molecule. In addition, an up to twofold improvement in efficiency is possible as all nonbonded interactions are treated at once, without resorting to the separation of Lennard-Jones and Coulombic interactions as is commonly done.[26] Finally, in contrast to conventional soft-core scaling, there are no scaling parameters to be optimized, since no atoms are created or annihilated in the process, but rather atoms are separated in space. Groundwork results indicated that the method is well suited for the calculation of the absolute hydration free energy of molecules with a variety of sizes, shapes, and polarities: Lennard-Jones spheres, water, and camphor.[22]

The 4D-PMF approach presents practical advantages in both short- and long-range limits. In the short-range (fully coupled) limit, the method is formally analogous to soft-core coupling schemes and effectively overcomes the problems associated with the repulsive part of the Lennard-Jones potential. In addition, it was shown that the shape of the PMF profile at large solute-solvent separation distances can be derived from the attractive part of the Lennard-Jones potential and from Coulombic solvent-solute interactions in the limit of a solvent continuum.[22] This approximation, which was verified in calculations of the hydration free energy of nonpolar and polar neutral solutes, can be used reliably to extrapolate the PMF to infinite solute-solvent separation and thus to the limit of complete decoupling, further improving the efficiency of the approach.

In the current work, we present methodological advancements designed to optimize sampling efficiency and we improve upon the generality of the 4D-PMF algorithm by extending its application to charged solutes. Results are presented for the absolute hydration free energies of both charged and uncharged molecular species, together with extensive error analysis. We consider the advantages afforded by calculating the PMF by thermodynamic integration in the fourth dimension. The direct calculation of the mean force from thermodynamic integration is shown to be more computationally efficient and better suited for the 4D-PMF method as compared to umbrella sampling.

Furthermore, we address the systematic errors arising from the truncation of Coulombic interactions in free energy calculations of charged solutes.[27–32] A systematic analysis of finite-size effects in a spherical water droplet containing an ionic solute at its center leads to simple corrections to the

hydration free energy. The spherical solvent boundary potential[33] (SSBP) accounts for the reaction field outside of the explicit water droplet, which is reasonably well approximated by the Born energy in sufficiently large droplets. In addition, a straightforward correction to account for the artificial polarization of water molecules at the droplet surface is introduced. Together, these corrections are shown to lead to consistent estimates of ionic free energies.

II. THEORY

Here we describe the theoretical foundations of thermodynamic integration in four dimensions and its implementation in molecular-dynamics simulations. A single molecular solute immersed in a bulk solvent is considered. While both the solute and solvent atoms are allowed to evolve in physical three-dimensional space (with Cartesian coordinates x, y, and z) as governed by a standard molecular mechanics force field, the Hamiltonian of the system is extended such that the solute atoms can be transported through an unphysical fourth dimension (along the w axis). The total Hamiltonian of the extended system can be written as

$$H = V_{vv} + V_{uu} + V_{vu}, \tag{3}$$

where V_{vv}, V_{uu}, and V_{uv} are the solvent-solvent, solute-solute, and solvent-solute potential energies, respectively. Since the solute is not covalently bonded to the solvent, the only forces acting between atoms of the solute and solvent are those resulting from nonbonded interactions, which in a typical force field would consist of the Coulombic and Lennard-Jones energy terms. Therefore, V_{vu} can be written as

$$V_{vu} = \sum_i \sum_j 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \sum_i \sum_j \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}, \tag{4}$$

where i and j are the indices of the solvent and solute atoms, respectively, ε_{ij} and σ_{ij} are the solvent-solute Lennard-Jones parameters, q_i and q_j are the solvent and solute atomic charges, respectively, and ε_0 is the permittivity of free space. In this equation, the only dynamic parameters on which V_{vu} depends are the interatomic distances between the solute and solvent atoms, r_{ij}. In four-dimensional Cartesian space (or $\mathbb{R}^4$), r_{ij} is computed as follows:

$$r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2 + w_u^2}, \tag{5}$$

where x_i, y_i, and z_i are the coordinates of solvent atom i; x_j, y_j, and z_j are the coordinates of solute atom j; and w_u is a parameter that specifies the distance, along the w axis, between the solute and solvent. For simplicity, the discussion will assume that the four-dimensional position of the solvent is fixed at zero and only the solute will travel through the fourth dimension. Therefore, w_u can be thought of as the absolute position of the solute on the w axis. Also note that V_{vv} and V_{uu} do not depend on w_u, since all the atoms that comprise the solvent, or equivalently the solute, share the same position along the w axis. The PMF is determined from the following calculation:

$$W(s) = \int_0^s \left\langle \frac{\partial H}{\partial w_u} \right\rangle dw_u, \quad 0 \leqslant s \leqslant \infty. \tag{6}$$

Here the fourth dimension component of the mean force, which is the expectation value of the derivative of the Hamiltonian with respect to w_u, is integrated over the path of travel along w. This yields the work or free energy change associated with the transport of the solute. This calculation is formally analogous to the thermodynamic integration technique using λ as the coupling parameter, where the PMF can be expressed as

$$W(s) = \int_0^s \left\langle \frac{\partial H}{\partial \lambda} \right\rangle d\lambda, \quad 0 \leqslant s \leqslant 1. \tag{7}$$

In practice, the mean force is acquired by averaging the fourth dimension component of the force (or $\partial H/\partial w_u$) between solvent and solute over the sampling run of a molecular dynamics simulation. A separate simulation is carried out for each discrete step of w_u; the distance between two adjacent steps needs to be small enough such that the numerically integrated data accurately reconstructs the PMF.

Alternatively, the PMF can be calculated[22] using the umbrella sampling method where the underlying relationship is as follows:

$$W(w_u) = -k_B T \ln\left(\frac{\langle \rho(w_u) \rangle}{\langle \rho(0) \rangle}\right), \tag{8}$$

where ρ is a probability function of w_u attained from a molecular dynamics simulation in which w_u is allowed to evolve freely as governed by the Hamiltonian.[25] In practice, a series of simulations are run, each biased with a harmonic potential such that a different portion of the PMF is sampled efficiently. Raw data from this "umbrella sampling" scheme are subsequently debiased, resulting in an overall PMF for the abstraction (or insertion) of the molecule from (or into) the system of interest.

Based on Eqs. (4) and (5), we note that V_{vu} goes to zero when $w_u \to \infty$, and that V_{vu} is equal to the physical (three-dimensional) solvent-solute interaction potential when $w_u = 0$. These two states correspond to noninteracting solvent and solute, and to a fully interacting, physical three-dimensional system where the solute is completely immersed in the solvent, respectively. The excess chemical potential is calculated as the difference between the PMF at $w_u = 0$ and $w_u = \infty$,

$$\mu = W(\infty) - W(0). \tag{9}$$

The usefulness of any integration pathway depends on the sampling efficiency throughout the entire thermodynamic transformation. The advantages of the 4D-PMF approach stem from the effective treatment of both short- and long-range interactions. The largest forces acting on the solute occur when r_{ij} is small enough that solvent-solute interactions are dominated by the repulsive part of the Lennard-Jones potential [i.e., the r_{ij}^{-12} term in Eq. (4)]. The endpoint catastrophe encountered when inserting or extracting solutes with simple scaling of solvent-solute interactions is due to the singularity in V_{vu} in the $r_{ij} = 0$ limit. The 4D-PMF

approach avoids the endpoint catastrophe because finite values of w_u allow solvent-solute overlap. Thus, by analogy with shifting or other soft-core approaches, the extra degree of freedom (w_u) allows solvent-solute interactions to be coupled or decoupled without the need to create or annihilate atoms. However, unlike parametric approaches (such as scaling or shifting), which may be described as alchemical transformations[34] because they modify the nature of the potential energy hypersurface during the course of the free energy calculation, the 4D-PMF method relies entirely on spatial coordinates.

When w_u is large enough, V_{vu} becomes dominated by long-range interactions and the potential of mean force can be derived by integrating solvent-solute interactions in the limit of a solvent continuum.[22] In particular, it was shown that the long-range PMF profile in water obeys an asymptotic form given by

$$W(w_u) = -\frac{k_{vdw}}{w_u^3} - \frac{k_{dd}}{w_u^3} - \frac{k_{cd}}{w_u} + W(\infty), \tag{10}$$

where k_{vdw}, k_{dd}, and k_{cd} are constants governing the strength of solvent-solute van der Waals dispersion, dipole-dipole, and charge-dipole interactions, respectively. In the case of neutral solutes, the charge-dipole term drops out while in the case of nonpolar molecules, $k_{dd}=0$. Using Eq. (10), it is possible to determine the asymptotic limit of the PMF, $W(\infty)$, by extrapolation.[22]

III. METHOD

The PMF for the abstraction of a (single) TIP3P (Ref. 35) water molecule, a methanol molecule, and three ions from bulk water were calculated at 300 K using Langevin molecular-dynamics simulations and thermodynamic integration in four spatial dimensions. Version c28b1 of the CHARMM program[36] was utilized and parameters for all solutes were taken, without modification, from the CHARMM force field, versions 22 and 27.[37] Simulations were carried out in a 20 Å radius sphere containing between 992 and 1041 TIP3P water molecules. In addition, several simulations were designed with droplets varying in size between 14 and 1032 water molecules in order to study the effect of the system size on the free energy of hydration of an ion (see Table II). In all simulations of charged systems, the SSBP (Ref. 33) method was utilized to account for the reaction field due to the electrostatic interactions extending beyond the limits of the explicit system. Furthermore, simulations (of 100 ps in length) of the same droplets (containing between 14 and 1032 water molecules) but lacking a solute were carried out to examine the effect of the SSBP boundary on the charge density of the solvent. The observed boundary polarization effect was corrected by determining the work required to extract, along the 4D axis, a positive unit test charge, given the averaged radial charge density. Calculating this correction was computationally very inexpensive (ranging from 0.1 to 35 CPU hours for the various system sizes). A spherical quartic potential boundary was used for all simulations containing neutral solutes. In all cases, a weak harmonic restraint, which acts solely on the x, y, and z coordinates and ignores the w coordinate, was imposed on the solute to keep it at the center of the water droplet. No cutoffs for nonbonded interactions were imposed. All simulations were carried out using a time step of 2 fs and a friction coefficient of 5 ps^{-1} applied to all nonhydrogen atoms.

The CHARMM program,[36] which had already been previously modified to include 4D capability,[22] was refined further to record the w-axis projection of the net force acting on the solute at every time step. In all 4D-PMF simulations, the w coordinates of all solvent atoms were fixed at 0, while the w coordinates of the solute atoms were fixed at a preset w_u value. Separate simulations were executed for discrete preset w-axis positions of the solute ranging from $w_u=0$ Å to $w_u=20$ Å. In all cases, the spacing between adjacent windows was the smallest in places where the function to be integrated, the force, can potentially vary most rapidly, namely, at small values of w_u.[5,26] Step intervals of $\Delta w_u=0.1$ Å between $w_u=0$ Å and $w_u=4$ Å; then $\Delta w_u=0.5$ Å between $w_u=4$ Å and $w_u=10$ Å; and finally $\Delta w_u=1$ Å beyond $w_u=10$ Å were used (63 sampling runs in total per solute). Furthermore, a second PMF for a TIP3P water molecule was generated using five times as much sampling. In this control calculation, the interval between sampling runs in the range $w_u=0.1$ Å to $w_u=4$ Å was $\Delta w_u=0.05$ Å (103 sampling runs in total).

Each simulation consisted of a short equilibration run followed by a longer sampling run. For methanol, the equilibration time (2 ps) and sampling time (20 ps) were the same for all simulations. In the case of a water solute, an equilibration time of 0.5 ps was used for all simulations and a more efficient sampling scheme was used in which CPU resources were directed to where the uncertainty in the force was the greatest, near $w_u=2$ Å. Upwards of 30 ps of sampling per simulation (60 ps for the "control" simulation) were performed in this vicinity of w_u. Simulations of the ions were carried out using a similarly optimized scheme and consisted of 2 ps equilibration and up to 100 ps (in the vicinity of $w_u=2$ Å) sampling runs. Furthermore, a correction free energy was calculated for turning off the charge of the ion in the presence of the SSBP, to account for the interaction between the charge and the reaction field outside the explicit system. This was done using the PERT module of CHARMM (using 11 windows and thermodynamic integration). Each window consisted of a 0.4 ps equilibration and 1.6 ps sampling run.

For comparison, the 4D-PMF for TIP3P water was also calculated using umbrella sampling.[25] Harmonic biasing potentials of 30 kcal/mol/Å^2 were set up at intervals of $\Delta w_u=0.1$ Å over the range from $w_u=0$ Å to $w_u=10$ Å. Only minor adjustments in the duration of the sampling runs were made in order to keep the total CPU time identical for both the thermodynamic integration and the umbrella sampling simulations. See Table I for the relevant simulation times and CPU requirements for all the simulations.

The equilibration part of a simulation was started as a continuation of the equilibration portion of the simulation preceding it and was followed by the production run. Cascading the simulations in this way leads to a trivially parallelizable scheme for performing the calculations. Since each

TABLE I. Details of the free energy simulations.

Solute	Number of water molecules in droplet	Method	Total equilibration time (ps)	Total sampling time (ns)	Total CPU time (h)
TIP3P water	1041	TI[a]	31.5	0.89	330
TIP3P water (control)	1041	TI	51	4.8	1700
TIP3P water	1041	U[b]	50.5	0.87	330
Methanol	1041	TI	112	1.1	449
Sodium ion	14	TI	124	4.9	5
Sodium ion	22	TI	124	4.9	11
Sodium ion	37	TI	124	4.9	19
Sodium ion	58	TI	124	4.9	43
Sodium ion	124	TI	124	4.9	119
Sodium ion	449	TI	124	4.9	1162
Sodium ion	1032	TI	124	4.9	1760
Cesium ion	992	TI	124	4.9	1760
Chloride ion	1041	TI	124	4.9	1760

[a]Thermodynamic integration.
[b]Umbrella sampling.

step in w_u was small (in most cases 0.1 Å), the time required for equilibration after each perturbation was inherently short. For thermodynamic integration, the total force acting on the solute in the w direction was recorded at each time step at each discrete w_u value. The PMF was then calculated by numerically integrating the mean force with respect to w_u over the defined range [as per Eq. (6)]. The weighted histogram analysis method algorithm[38,39] was used to debias the data obtained from umbrella sampling and construct the PMF.

The error in the thermodynamic integration data was analyzed as per Allen and Tildesley[40] and is summarized here. The statistical property of interest is the force F, which was sampled n times in a given run. The mean force $\langle F \rangle$ was calculated, along with its standard deviation σ. If each sample of F was assumed to be statistically independent of the others then the standard deviation in the mean would simply be

$$\sigma_{\text{mean}} = \frac{\sigma}{\sqrt{n}}. \tag{11}$$

However, since samples of F were recorded at every time step, they are naturally highly correlated. It is therefore necessary to determine the correlation time or statistical inefficiency, a value reflecting the time that needs to elapse in a simulation before a new statistically independent sample of F can be acquired. The standard deviation in the mean is then adjusted accordingly,

$$\sigma_{\text{mean}} = \frac{\sigma}{\sqrt{n}}\sqrt{s}, \tag{12}$$

where s is a statistical inefficiency value calculated by dividing the entire data set into blocks each containing n_b samples. This is repeated for progressively larger blocks and the trend is extrapolated to block sizes of infinite size,

$$s = \lim_{n_b \to \infty} n_b \frac{\sigma_{\text{blocks}}^2}{\sigma^2}, \tag{13}$$

where σ_{blocks} is calculated by finding the mean of the data in each block and then determining the standard deviation of these mean values. The statistical errors associated with the PMFs are determined by choosing an error of zero for the first data point at $w_u = 0$ and accumulating the error over the integration using standard propagation rules.

IV. RESULTS AND DISCUSSION

A. Hydration free energy of a TIP3P water molecule

The results of the extraction of a single TIP3P water molecule from bulk water are shown in Fig. 1. Figure 1(a) depicts the mean force acting on the solute versus the w coordinate of the solute as derived from the "control" calculation. A one standard deviation error in the mean is also shown as the shaded area.

The shape of the force curve in the large-w limit is determined by differentiating equation (10). Since a water molecule is neutral, the charge-dipole interaction drops out ($k_{cd}=0$):

$$F(w_u) = 3\frac{k_{vdw} + k_{dd}}{w_u^4}. \tag{14}$$

The long-range fit, with $k_{vdw}+k_{dd}=62$ kcal Å^3/mol, was determined using least squares analysis on the PMF between $w_u=6$ Å and 20 Å. The black line in Fig. 1(b) shows the PMF resulting from the integration of the mean force data. This control curve appears smooth and noise-free as it was derived from a substantial 4.8 ns of sampling. The dotted line is a curve fit to this PMF and takes the form of Eq. (10). An optimal $W(\infty)$ was chosen as 6.70 kcal/mol. Since $W(0)$ was arbitrarily chosen as 0, the overall free energy change for the extraction of a TIP3P water molecule from bulk was determined to be 6.70 ± 0.04 kcal/mol. The corresponding

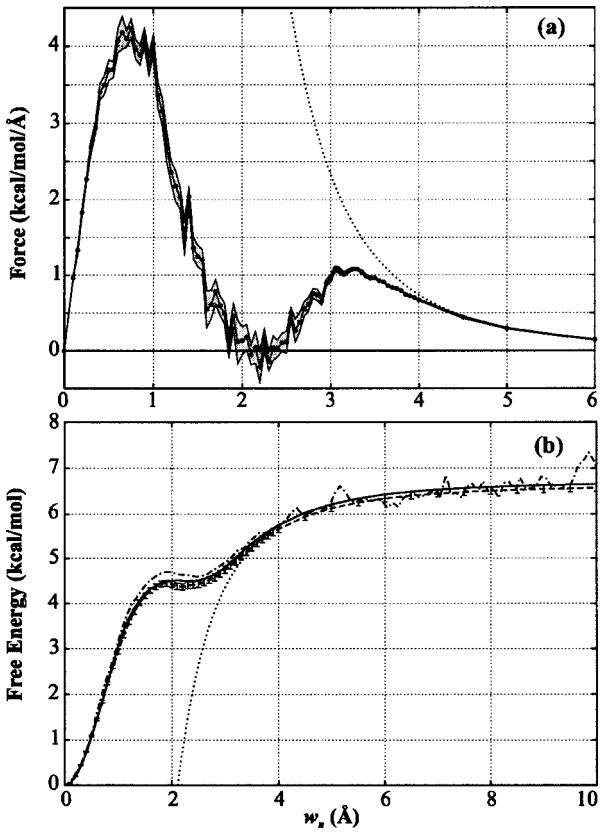

FIG. 1. Extraction of a single TIP3P water molecule from bulk water along a path through the fourth dimension. (a) Mean force acting on the molecule as it is removed into the fourth dimension. A positive force indicates attraction between solute and solvent while a negative force represents repulsion. The vertical width of the shaded region indicates a statistical error of one standard deviation. The dotted line is a large-w fit to the force curve and is given by the equation $F=186/w_u^4$. (b) Corresponding potential of mean force (PMF) profile arising from the extraction. The three plots show the PMF attained after 4.8 ns of sampling with thermodynamic integration (solid line; error bars are roughly the thickness of the line and are omitted for clarity), 0.89 ns of sampling with thermodynamic integration (dashed line with error bars), and 0.87 ns of umbrella sampling followed by weighted histogram analysis method analysis using 0.1 Å wide bins (dot-dashed line without error bars). The dotted line is a large-w fit to the black line and is given by the equation $W=-62/w_u^3+6.70$.

hydration free energy is -6.70 ± 0.04 kcal/mol. This result has an error of 6.3%, as compared to the experimental value of -6.3 kcal/mol.[41] The result obtained in the present study is also in good agreement with values ranging from -5.5 to -6.4 kcal/mol (Refs. 20, 22, 33, 42, and 43) reported from other computer simulations. The control calculation reported here was executed over $\approx$4 d on about 20 R3800 SGI processors with a total CPU time of 1700 h.

Figure 1(b) also compares the PMF computed from less than one-fifth of the sampling of the control. From this PMF and following the error analysis prescription by Allen and Tildesley[40] [see Eqs. (12) and (13)], we report an extraction free energy of 6.61 ± 0.08 kcal/mol. The fivefold decrease in sampling has resulted in roughly a doubling of the magnitude of the error. This is in line with previous work that has shown the statistical error to be proportional to the inverse of the square root of the amount of sampling.[44]

Finally, the PMF determined by umbrella sampling is shown for comparison. With an equivalent amount of sampling, the calculation yielded a result qualitatively in agreement, although the PMF is plagued with severe statistical noise. A PMF attained using umbrella sampling inherently suffers from statistical noise arising from nonuniform sampling within each umbrella and from cumulative error introduced by the uncertainty in matching adjacent umbrellas. This problem has been examined elsewhere.[45,46] By contrast, thermodynamic integration is easier to optimize in that the amount of sampling performed at any w_u is the only parameter of concern and this parameter can be adjusted trivially. Optimizing a series of umbrella sampling runs is more difficult because it involves the interplay between three parameters, the umbrella separation distances, umbrella stiffness, and the amount of sampling in each umbrella.[5,47] One way to avoid the need to specify these parameters in advance is to build up a nearly optimal biasing function as the simulation progresses. In such adaptive biasing force methods,[45,46,48] free energy barriers are effectively removed; however, uniform sampling is limited by diffusion and can only be realistically attained in very long simulations. Furthermore, it has been argued that the use of shorter sampling simulations with a larger number of narrow umbrellas rather than wider umbrellas with more sampling in each[39] is more efficient because sampling uniformity can be guaranteed in simulations of any length. In the limit of infinitely stiff umbrellas, the calculation becomes equivalent to thermodynamic integration, which partly explains the gain in efficiency afforded by thermodynamic integration in the 4D-PMF method.

However, the greatest pitfall of umbrella sampling in the present application resides in its inherent inefficiency at sampling regions of the PMF at large solvent-solute separations (i.e., beyond $w_u=4$ Å). As the solute moves away from the solvent, both the magnitude and fluctuations of the w projection of the force acting on the solute quickly diminish. Once the force is small enough, the solute is not compelled to move appreciably within the umbrella (and would not move at all if it were not for the random force of the Langevin integrator). As a result, it is necessary to space the umbrellas very densely to maintain a sufficient overlap between them. In adding more umbrellas, it is inevitable that the time spent sampling a region where the fluctuations are small, and thus the mean force is already quite certain, is not put to good use.

The control over the sampling profile afforded by the thermodynamic integration technique is well suited to improving statistical efficiency in the 4D-PMF method because it leads naturally to a strategy for optimizing the distribution of sampling over the relevant w_u range. Consider the plot of the force acting on a water molecule through the fourth dimension versus its position along the w axis [Fig. 1(a)]. The vertical width of the shaded region on this graph is an indicator of the uncertainty in the mean force (see Methods). To minimize the overall uncertainty in the PMF, it is logical to sample more thoroughly regions of w_u where the force fluctuates the most. In the calculation reported here, almost all of the sampling was directed to the region between $w_u=0$ Å and $w_u=4$ Å. Although the amount of sampling done was decided *a priori*, the process can be truly optimized by tai-

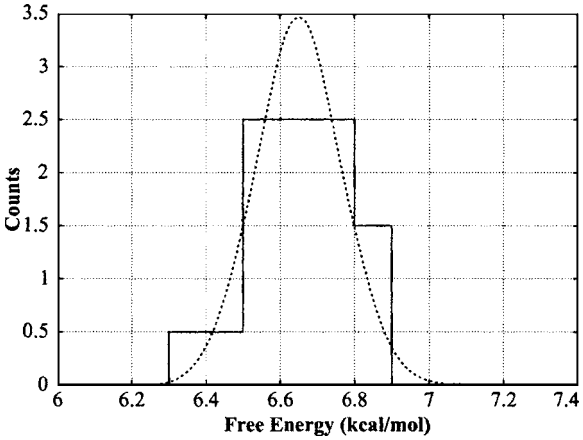

FIG. 2. Distribution of the free energy of dehydration of a TIP3P water molecule from a set of 20 calculations, each starting from a system of randomly placed water molecules. The solid line shows the actual histogram created from the data (with 0.1 kcal/mol wide bins). The dotted line is a Gaussian distribution fit to the data with a mean of 6.65 kcal/mol and a standard deviation of 0.12 kcal/mol.

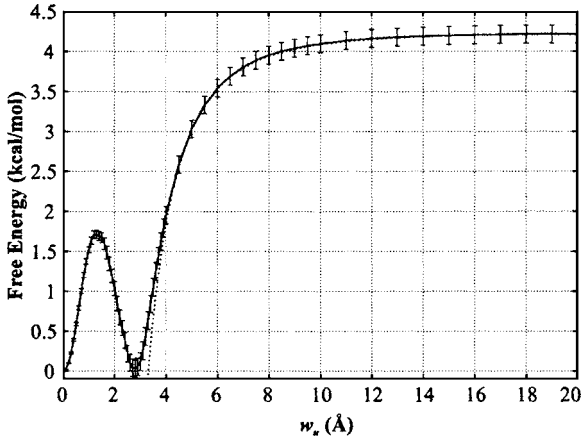

FIG. 3. PMF associated with the extraction of a methanol molecule from bulk water along a path through the fourth dimension. The dotted line represents the large-w fit to the PMF and is given by the equation $W=-153/w_u^3+4.24$.

loring the sampling distribution to the system at hand in real time. This can be done easily as it only requires that the error in the mean force be recomputed frequently, as new data are being generated, and available CPU resources be focused subsequently to the regions where the uncertainty is the greatest.[49]

The error calculated using the method described by Allen and Tildesley[40] was checked against an alternate means of error analysis, in which the standard deviation was attained from 20 independent trials of the same water PMF calculation. The simulation setup for each of these was identical to the 330 CPU-hour hydration free energy calculation of a TIP3P water molecule discussed above. However, each calculation was started with a different set of initial conditions where the positions, orientations, and velocities of the solvent molecules were assigned at random. The distribution of the final results is shown in Fig. 2. The fit Gaussian (also shown) has a mean of 6.65 kcal/mol and a standard deviation of 0.12 kcal/mol. The value attained using the Allen and Tildesley treatment[40] (6.61±0.08 kcal/mol), when compared to this result, shows an underestimate of the statistical error, in this case by 33%, and there is evidence suggesting that this gets worse for larger, more complex solutes (data not shown). However, attaining pinpoint accurate estimations of the statistical error is not critical, since errors as high as 0.12 kcal/mol or 2% are still likely to be much smaller than the errors introduced by the approximations in the force field in most cases (which are expected to be on the order of 1 kcal/mol, see Ref. 20).

B. Hydration free energy of a methanol molecule

The PMF for the extraction of methanol from bulk water (Fig. 3) is qualitatively similar to that of water. The profile reflects the balance of Coulombic and Lennard-Jones forces, at both short and long ranges. Both the shape and the magnitude of the PMF for the extraction of methanol from water can be rationalized in terms of solvent-solute interactions. When the methanol molecule is fully inserted in the system ($w_u=0$ Å), its hydroxyl group interacts favorably with the solvent by forming hydrogen bonds with water molecules. Because of this, there is an initial increase in free energy as the solute is forcefully extracted into the fourth dimension. At a point around $w_u=1.3$ Å, attractive Coulombic forces between water and the hydrophilic moiety of methanol weaken sufficiently and can no longer compensate for the entropic drive to fill the cavity with solvent and expel the solute. Because methanol is larger than water and has a hydrophobic moiety, the "squeezing out" effect is significantly more pronounced. When the solute has moved beyond the Lennard-Jones contact radius, the cavity becomes filled completely and the squeezing out effect ends. Beyond this point, the attractive force acting on the solute results from a combination of dispersion and dipole-dipole interactions. The large-w portion of the PMF expectedly follows a w_u^{-3} dependence [Eq. (10)] and asymptotically approaches a limit as the distance between the solute and solvent approaches infinity and the force between them tapers to zero. The expressions derived earlier[22] for a solvent continuum predict that $k_{vdw}=137$ Å^3 kcal/mol and $k_{dd}=0.69$ Å^3 kcal/mol, which is in fair agreement with the fit curve in Fig. 3, where $k_{vdw}+k_{dd}=153$ Å^3 kcal/mol. Thus, dipole-dipole interactions contribute to a much smaller extent to the long-range part of the PMF than do dispersion forces, consistently with the results obtained for water hydration, where the ratio of dispersion to dipolar interactions was shown to be 41.[22] Based on the long-range fit to the PMF and on error analysis, we predict a hydration free energy of −4.24±0.11 kcal/mol. This result differs by 17% from the experimental value of −5.1 kcal/mol.[41] The asymptotic limit of a PMF for an inhomogeneous system such as one containing a protein or other biological molecule may not be predictable from a continuum approximation; however, it is expected to follow a well-behaved long-range profile and thus be predictable through extrapolation.

C. Boundary polarization and reaction field corrections for charged solutes

Because an accurate treatment of long-range electrostatics is critical for charged solutes,[27–32] we have considered

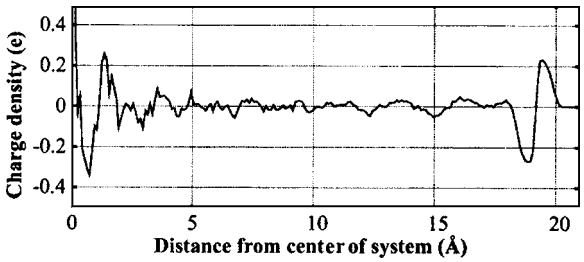

FIG. 4. Charge density as a function of radial distance from the center of a water droplet with radius 20 Å. Note that the peaks observed at a small radial distance from the center of the sphere are not statistically significant due to the fact that the integration took place over a small volume.

two possible difficulties arising from the fact that any simulated system containing explicit atoms must inevitably be finite in size.

The first of these is the effect of the boundary on the polarization of water molecules. Water molecules are preferentially polarized at vacuum-water interfaces in finite-size systems (see, for example, Refs. 33 and 50). While this is an effect of using a finite-sized system, it should be noted that artifacts resulting from the truncation of Coulombic interactions around charged solutes also result in systematic errors in simulations employing periodic conditions.[28] Our results indicate that surface water polarization, which exists both in the presence and in the absence of the SSBP (Ref. 33) boundary potential, has a significant effect on the extraction free energy of charged species. Figure 4 shows the average charge density as a function of the distance from the center of a spherical water droplet (no solute present) bound by an SSBP boundary. At the vacuum-water interface, the solvent molecules preferentially point their hydrogen atoms outward. The effect of this polarization is propagated towards the interior of the water sphere and the structure is evident as far in as 10 Å. This gives rise to an artificial static field that favors the extraction of negatively charged ions, and opposes the extraction of positively charged ions into the fourth dimension. The degree to which this boundary polarization biases the extraction free energy of a charged ion can be measured with a simple calculation of the work required to move a positive unit test charge from the center of the charge density to a distance infinitely far away via the fourth dimension. The force acting on the test charge and the accumulated work is plotted in Fig. 5. This interaction energy is later used as a correction to the extraction free energy of the three ions (see below).

Second, we have examined the effect of including a solvent continuum outside the fixed boundary of the explicit system (i.e., a dielectric medium with $\varepsilon=80$). The interaction between the charged solute and resulting reaction field was included by way of the SSBP algorithm.[33] Because the SSBP algorithm works in three-dimensional space and ignores the fourth dimension, a 4D extraction by itself does not include the effect of the reaction field. Accordingly, whether SSBP or a simple quartic boundary potential is used does not change the extraction PMF to any appreciable extent (data not shown). The reaction field contribution is, however, easy to calculate, by alchemically removing the charge on the ion once the ion is located far in the fourth dimension

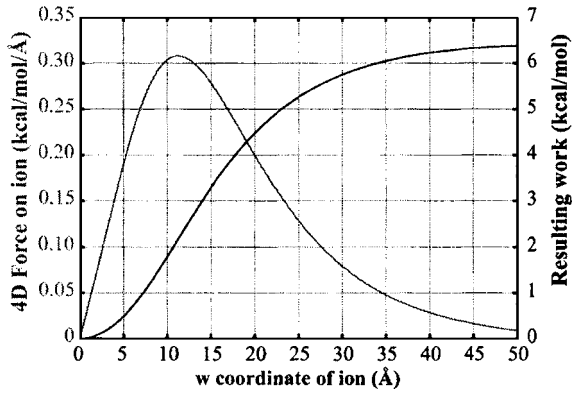

FIG. 5. Force (gray line; left-hand scale) acting and work done (black line; right-hand scale) in the 4D extraction of a test positive unit charge from a water droplet with radius 20 Å. Note that the results shown are calculated directly from the charge density data in Fig. 4.

(at $w_u=1000$ Å) (see Methods). Because the ion is at a large w_u, its interaction with the explicit system is essentially turned off and it interacts only with the reaction field, via the SSBP algorithm.

Figure 6 demonstrates the system size dependence of the extraction free energy of a sodium ion, both before and after corrections for the reaction field and boundary polarization are applied. The number of water molecules in the system was varied from 14 to 1032, resulting in a range of systems with radii varying from 4.6 to 19.5 Å. Results are summarized in Table II. The uncorrected data demonstrates a strong system-size dependence throughout the examined range. Both corrections decrease with increasing system size and presumably go to zero in a system approaching infinite size. A system size much greater than that studied here would be needed to bring the systematic error due to the finite-size boundary to the same order as the magnitude of the statistical error (simulation of such a large system would be

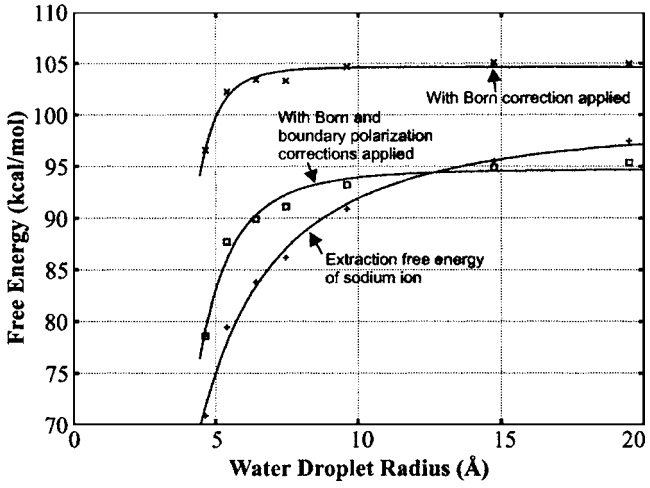

FIG. 6. Free energy of dehydration of a sodium ion vs the size of the explicit water droplet. The three lines from bottom to top represent the free energy of extraction as calculated using the 4D-PMF method, the free energy of extraction with Born energy correction applied, and the free energy of extraction with both the Born energy and boundary polarization corrections applied. The corresponding analytical fits are given by $W=-365/r^{1.68}+99.6$, $W=-192\,000/r^{6.58}+104.6$, and $W=-5850/r^{3.87}+94.7$. Also see Table II.

TABLE II. Dependence of sodium hydration free energy on system size.

Number of water molecules in the system	Water droplet radius (Å)	Extrection free energy (kcal/mol)	Reaction field energy (kcal/mol)	Born energy (kcal/mol)	Boundary polarization energy (kcal/mol)	Free energy of dehydration (kcal/mol)
14	4.6	70.8±7.1	25.7	26.9	18.0	78.6
22	5.4	79.4±3.1	22.8	23.9	14.5	87.7
37	6.4	83.8±1.1	19.6	20.8	13.5	89.9
58	7.5	86.1±8.1	17.1	18.4	12.2	91.1
124	9.6	90.8±6.1	13.8	14.8	11.5	93.2
449	14.7	95.5±1.1	9.6	10.1	10.2	94.9
1032	19.5	97.5±0.1	7.5	7.8	9.6	95.4

impractical). More practically, the hydration free energy of an ion can be obtained via an extrapolation based on several extraction free energy computations done with varying system sizes as was done here for sodium (Fig. 6). Extrapolating this curve to an infinite radius yields a dehydration free energy of 99.6 kcal/mol, although this may be inaccurate because of the inherent sensitivity of the extrapolated result on small errors in the data points. Furthermore, this is a computationally expensive procedure, since multiple systems of varying sizes need to be simulated.

Figure 6 also demonstrates that upon correction of both the boundary polarization and the reaction field errors, the calculated extraction free energy of the sodium ion reaches a plateau of 94.7 kcal/mol and becomes independent of system size at a radius greater than 10 Å. Therefore, with the corrections taken into account, the hydration free energy of an ion can be feasibly calculated from a simulation of a system as small as 10 Å in radius and would require roughly 120 h of CPU time. This amounts to an ≈26-fold reduction in CPU requirements in comparison to the 3119 CPU hours that were necessary for the extrapolation method discussed above.

The calculated reaction field energies are also in very good agreement with the respective Born energies[51] (listed in Table II). According to the Born model, the energy (ΔG_{Born}) should depend only on the radius (a) of the explicit system, the charge on the solute (q), and the dielectric constant of the medium (ε):

$$\Delta G_{\mathrm{Born}} = -\frac{q^2}{8\pi\varepsilon_0 a}\left(1-\frac{1}{\varepsilon}\right). \tag{15}$$

Thus, in a homogeneous medium, the explicit calculation of the reaction field may be replaced with a Born energy correction.

D. Extraction of ions from water

The artificial polarization of surface water molecules leads to, in terms of extraction free energy, an underestimate for negative ions and overestimate for positive ions. The boundary polarization correction was applied appropriately and the resulting PMFs for the extraction of sodium, chloride, and cesium ions from bulk water are shown in Fig. 7. Asymptotic limits yield extraction free energies of 88.6±0.1, 80.5±0.1, and 53.9±0.1 kcal/mol, respectively (note that the reaction field error correction is not accounted for in this data).

The extraction PMF profiles for the ionic solutes depicted in Fig. 7 are qualitatively different from those of water and methanol. Because they are charged and considerably more hydrophilic, cavity explusion is more than offset by the attractive Coulombic interactions acting between the charged solute and the polarized water molecules. The radial distribution function shown in Fig. 8(a) demonstrates a clear preference for water molecules to point their negatively charged oxygen atom towards the positive sodium ion, thus forming an interaction that strongly disfavors the extraction of the solute into the fourth dimension.

As early in the extraction process as w_u=2.2 Å [Fig. 8(b)], overlap of the water molecules with the sodium ion in the physical three-dimensional space becomes energetically feasible. Note that there is still local ordering in the structure of the first hydration shell. As demonstrated in an earlier study of simple Lennard-Jones solutes, the cavity expulsion only amounts to a few kcal/mol for small solutes but increases with the solute size.[22] On the other hand, the

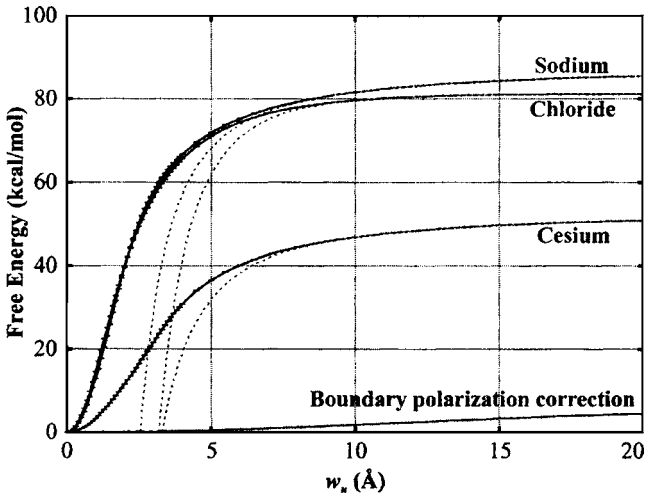

FIG. 7. PMF associated with the extraction of three ions, sodium, chloride, and cesium, from bulk water along a path through the fourth dimension. The PMFs depicted contain a correction for boundary polarization (see Fig. 5). Large-w fits to the PMFs (dotted lines) are given by $W=-59.1/w_u - 1073/w_u^3+88.6$, $W=19.7/w_u-2821/w_u^3+80.5$, and $W=-57.5/w_u - 1316/w_u^3+53.9$ kcal/mol for sodium, chloride, and cesium, respectively.

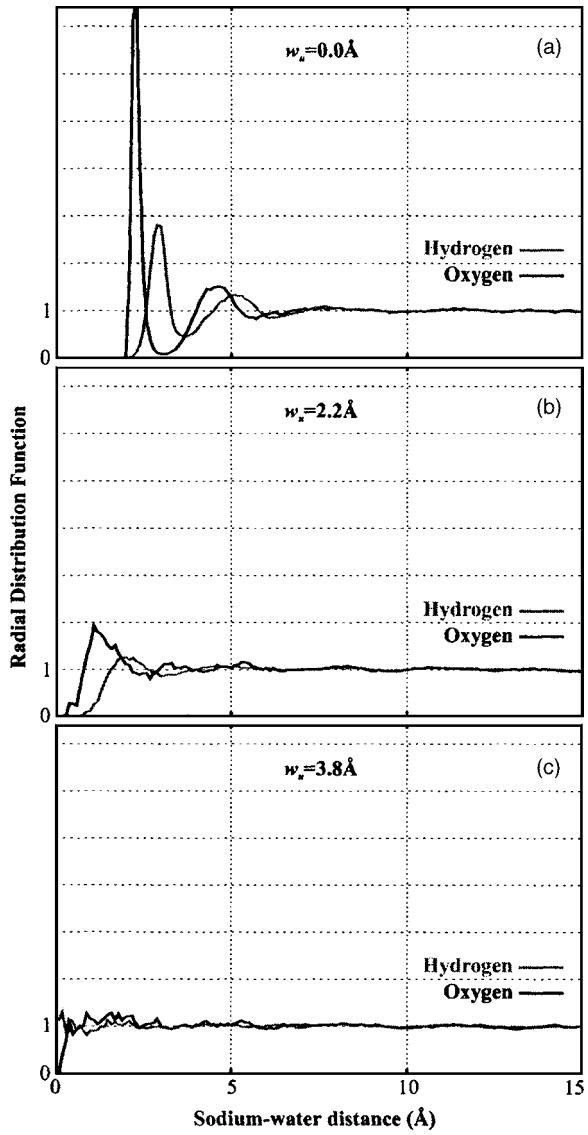

FIG. 8. Sodium-hydrogen and sodium-oxygen radial distribution functions at various stages of progression as a sodium ion is transported into the fourth dimension. Simulations are carried out in a system consisting of a 20 Å radius sphere of water and a centered sodium ion. Functions corresponding to four-dimensional positions of the sodium at w_u=0.0 Å (a), 2.2 Å (b), and 3.8 Å (c) are shown. Note that the radial distribution functions illustrated here [especially (c)] suffer from statistical noise at small separation distances due to an artifact of integrating over a small volume.

electrostatic attraction between the sodium ion and the polarized solvent amounts to a dominating 40 kcal/mol when the ion is transported to a distance of only w_u=2 Å. At w_u=3.8 [Fig. 8(c)] and beyond, the cavity is completely filled with water and although local water structure has visibly vanished at this point in the extraction, the water molecules remain polarized on average by the ion's presence. A final extraction free energy of 88.6 kcal/mol is attained upon completion of the extraction. Integration of Coulomb forces in the limit of a solvent continuum predict that the long-range tail of the PMF for a charged solute would follow a w_u^{-1} dependence.[22] This prediction could not be verified here because the reaction field correction, a substantial contribution to the overall free energy of hydration calculation, was not determined as a function of the four-dimensional coordinate of the solute, and thus cannot be added to the 4D PMFs. Instead, the correction was measured in a separate step where the charge of the solute was scaled to zero (see Methods) and separately added to the final extraction free energy values. However, the missing corrections to the PMFs in Fig. 7 would necessarily be identical for all three ions considered here, since such corrections would only depend on the system sizes and magnitudes of the ionic charges (both of which are identical). The intricacies of interactions involving explicit solvent do not come into play in this correction. Although the cations and anions investigated show a mixture of w_u^{-1} and w_u^{-3} dependence, the coefficients differ significantly between positive and negative ions (equations of the fit curves are given in the figure caption of Fig. 7). A reaction field correction applied identically to the PMFs of all three ions would not correct this discrepancy. Simple molecular mechanics force fields such as the one used in this study may be too crude to capture all the details pertaining to the structure and thermodynamics of ionic hydration. Recent studies have shown that nonrigid and polarizable models of water considerably improve the treatment of ionic hydration.[52,53]

Based on the results depicted in Fig. 7, and after separately adding the correction for the reaction field energy (of 7.5 kcal/mol), the final hydration free energies of sodium, chloride, and cesium are reported as -96.1 ± 0.1, -88.0 ± 0.1, and -61.4 ± 0.2 kcal/mol, respectively, which is in fair agreement with the experimental values of −98.4 (Refs. 54 and 55), −76.1 (Refs. 54 and 55), and −67.8 kcal/mol (Ref. 54). More importantly, our data indicate that combining reaction field and boundary polarization corrections with a sufficiently large spherical nonperiodic system can accurately represent an infinite solvent system, and that such a setup can be used for calculating hydration free energies of charged species in dense media.

V. CONCLUSIONS

We have built upon the 4D PMF method by incorporating thermodynamic integration as a means to calculate efficiently the excess chemical potential of a small molecule in a dense liquid. Direct accumulation of the mean force was shown to improve sampling efficiency over umbrella sampling and to be inherently better suited to the 4D method. The tractability, simplicity, effectiveness, and generality of the method were demonstrated through the calculations of the hydration free energy of water, methanol, and charged ions. Statistical error analysis was performed to gauge the magnitude of the sampling uncertainty resulting from these types of calculations. The detailed analysis of systematic errors arising from the truncation of Coulombic interactions in a solvent droplet of finite size leads to straightforward corrections to ionic hydration free energies. In turn, these corrections yield consistent results for droplets greater than 10 Å in radius (i.e., containing as little as 124 water molecules). Although there are many examples of PMF calculations for transformations in three-dimensional space,[45,46,56] to our knowledge the 4D-PMF approach is the only method that allows the free energy difference between two chemically distinct systems to be obtained without the need for creating

or annihilating atoms in the transformation process but instead by calculating a PMF along a spatial coordinate. An inherent advantage not offered by conventional soft-core methods is that it is practical to decouple both Lennard-Jones and Coulombic interactions at once, thus saving computational resources. Computer efficiency is further enhanced by the fact that the mean force converges to asymptotic behavior at relatively short 4D separations as the solute molecule is extracted from the solvent. Furthermore, the methodology offers a straightforward prescription for optimizing sampling at short 4D separations, where the fluctuations in the mean force are greatest. Hydration free energy calculations of small solutes, which represent an important step in the calculation of absolute protein-ligand binding affinities, have definitely entered the realm of practicability in the past several years. The present study demonstrates that the 4D-PMF method constitutes a viable alternative to conventional approaches[20,57] as results of comparable precision can be obtained within the limits of one work day on a modern but inexpensive cluster of computers. The advantages of statistical sampling efficiency afforded by the 4D-PMF method are expected to hold in molecular systems of greater complexity, such as proteins, and more generally, systems evolving on rugged energy landscapes. In such systems, systematic errors arising from slow relaxation of degrees of freedom perpendicular to the reaction coordinate present an additional challenge.[48] Forthcoming work will focus on applications of the method to solutes in inhomogeneous solutions and to systems of biological scale where the "solvent" degrees of freedom include those of a protein receptor or binding site.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the contribution of CPU resources and support from the Ontario Centre for Genomic Computing at the Hospital for Sick Children. This work was supported by a grant (Grant No. MT43998) from the Canadian Institutes of Health Research (CIHR). T.R. is a recipient of a Canada Graduate Scholarship from the Natural Sciences and Engineering Research Council of Canada (NSERC). P.L.H. is a CIHR Investigator. R.P. is a CRCP chairholder.

[1] D. Frenkel and B. Smith, *Understanding Molecular Simulation: From Algorithms to Applications* (Academic, San Diego, 1996).
[2] M. K. Gilson, J. A. Given, B. L. Bush, and J. A. McCammon, Biophys. J. **72**, 1047 (1997).
[3] B. Roux, M. Nina, R. Pomès, and J. C. Smith, Biophys. J. **71**, 670 (1996).
[4] C. Chipot and D. A. Pearlman, Mol. Simul. **28**, 1 (2002).
[5] W. F. van Gunsteren, X. Daura, and A. E. Mark, Helv. Chim. Acta **85**, 3113 (2002).
[6] T. Rodinger and R. Pomès, Curr. Opin. Struct. Biol. **15** 164 (2005).
[7] J. W. Pitera and W. F. van Gunsteren, Mol. Simul. **28**, 45 (2002).
[8] R. W. Zwanzig, J. Chem. Phys. **22**, 1420 (1954).
[9] T. P. Straatsma, H. J. C. Berendsen, and J. P. M. Postma, J. Chem. Phys. **85**, 6720 (1986).
[10] T. Simonson, Mol. Phys. **80**, 441 (1993).
[11] M. Zacharias, T. P. Straatsma, and J. A. McCammon, J. Chem. Phys. **100**, 9025 (1994).
[12] C. L. Lin and R. H. Wood, J. Comput. Chem. **15**, 149 (1994).
[13] T. P. Straatsma, M. Zacharias, and J. A. McCammon, Chem. Phys. Lett. **196**, 297 (1992).
[14] T. C. Beutler, A. E. Mark, R. C. van Schaik, P. R. Gerber, and W. F. van Gunsteren, Chem. Phys. Lett. **222**, 529 (1994).
[15] B. Widom, J. Chem. Phys. **39**, 2808 (1963).
[16] T. C. Beutler and W. F. van Gunsteren, J. Chem. Phys. **101**, 1417 (1994).
[17] V. Helms and R. C. Wade, J. Am. Chem. Soc. **120**, 2710 (1998).
[18] T. Z. M. Denti, W. F. van Gunsteren, and F. Diederich, J. Am. Chem. Soc. **118**, 6044 (1996).
[19] B. C. Oostenbrink, J. W. Pitera, M. M. H. van Lipzig, J. H. N. Meerman, and W. F. van Gunsteren, J. Med. Chem. **43**, 4594 (2000).
[20] M. R. Shirts, J. W. Pitera, W. C. Swope, and V. S. Pande, J. Chem. Phys. **119**, 5740 (2003).
[21] A. Dejaegere and M. Karplus, J. Phys. Chem. **100**, 11148 (1996).
[22] R. Pomès, E. Eisenmesser, C. B. Post, and B. Roux, J. Chem. Phys. **111**, 3387 (1999).
[23] R. C. van Schaik, H. J. C. Berendsen, A. E. Torda, and W. F. van Gunsteren, J. Mol. Biol. **234**, 751 (1993).
[24] X. J. Kong and C. L. Brooks, J. Chem. Phys. **105**, 2414 (1996).
[25] G. M. Torrie and J. P. Valleau, Chem. Phys. Lett. **28**, 578 (1974).
[26] B. O. Brandsdal and A. O. Smalas, Protein Eng. **13**, 239 (2000).
[27] I. G. Tironi, R. Sperb, P. E. Smith, and W. F. van Gunsteren, J. Chem. Phys. **102**, 5451 (1995).
[28] M. Bergdorf, C. Peter, and P. H. Hunenberger, J. Chem. Phys. **119**, 9129 (2003).
[29] H. Resat and J. A. McCammon, J. Chem. Phys. **108**, 9617 (1998).
[30] H. Resat and J. A. McCammon, J. Chem. Phys. **104**, 7645 (1996).
[31] T. Hansson, C. Oostenbrink, and W. F. van Gunsteren, Curr. Opin. Struct. Biol. **12**, 190 (2002).
[32] P. H. Hunenberger and J. A. McCammon, J. Chem. Phys. **110**, 1856 (1999).
[33] D. Beglov and B. Roux, J. Chem. Phys. **100**, 9050 (1994).
[34] T. P. Straatsma and J. A. McCammon, Annu. Rev. Phys. Chem. **43**, 407 (1992).
[35] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[36] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, J. Comput. Chem. **4**, 187 (1983).
[37] A. D. MacKerell, D. Bashford, M. Bellott *et al.*, J. Phys. Chem. B **102**, 3586 (1998).
[38] S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman, and J. M. Rosenberg, J. Comput. Chem. **13**, 1011 (1992).
[39] B. Roux, Comput. Phys. Commun. **91**, 275 (1995).
[40] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, New York, 1987).
[41] A. Bennaim and Y. Marcus, J. Chem. Phys. **81**, 2016 (1984).
[42] V. Helms and R. C. Wade, Biophys. J. **69**, 810 (1995).
[43] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans, *Intermolecular Forces* (Reidel, Dordrecht, 1981).
[44] J. Hermans, R. H. Yun, and A. G. Anderson, J. Comput. Chem. **13**, 429 (1992).
[45] E. Darve and A. Pohorille, J. Chem. Phys. **115**, 9169 (2001).
[46] E. Darve, M. A. Wilson, and A. Pohorille, Mol. Simul. **28**, 113 (2002).
[47] J. S. van Duijneveldt and D. Frenkel, J. Chem. Phys. **96**, 4655 (1992).
[48] J. Henin and C. Chipot, J. Chem. Phys. **121**, 2904 (2004).
[49] T. P. Straatsma and J. A. McCammon, J. Chem. Phys. **95**, 1175 (1991).
[50] M. Mucha, T. Frigato, L. M. Levering, H. C. Allen, D. J. Tobias, L. X. Dang, and P. Jungwirth, J. Phys. Chem. B **109** 7617 (2005).
[51] M. Z. Born, Z. Phys. **1**, 45 (1920).
[52] A. Grossfield, P. Y. Ren, and J. W. Ponder, J. Am. Chem. Soc. **125**, 15671 (2003).
[53] Z. H. Duan and Z. G. Zhang, Mol. Phys. **101**, 1501 (2003).
[54] R. M. Noyes, J. Am. Chem. Soc. **84**, 513 (1962).
[55] T. P. Straatsma and H. J. C. Berendsen, J. Chem. Phys. **89**, 5876 (1988).
[56] S. Shimizu and H. S. Chan, Proteins **49**, 560 (2002).
[57] Y. Q. Deng and B. Roux, J. Phys. Chem. B **108**, 16567 (2004).