Using Interpolation for Fast and Accurate Calculation of Ion–Ion Interactions

Miha Lukšič, Christopher J. Fennell, and Ken A. Dill

ABSTRACT: We perform extensive molecular dynamics (MD) simulations between pairs of ions of various diameters (2−5.5 Å in increments of 0.5 Å) and charge (+1 or −1) interacting in explicit water (TIP3P) under ambient conditions. We extract their potentials of mean force (PMFs). We develop an interpolation scheme, called i-PMF, that is capable of capturing the full set of PMFs for arbitrary combinations of ion sizes ranging from 2 to 5.5 Å. The advantage of the interpolation process is computational cost. Whereas it can take 100 h to simulate each PMF by MD, we can compute an equivalently accurate i-PMF in seconds. This process may be useful for rapid and accurate calculation of the strengths of salt bridges and the effects of bridging waters in biomolecular simulations. We also find that our data is consistent with Collins’ “law of matching affinities” of ion solubilities: small−small or large−large ion pairs are poorly soluble in water, whereas small−large are highly soluble.

INTRODUCTION

We describe here a method that can rapidly and accurately compute the potentials of mean force (PMFs) between spherical univalent ions in water. The PMF between an ion A and ion B represents the reversible work (as a function of distance between the two ions, r), averaged over all the configurations of the surrounding solvent, needed to bring A and B from infinite separation to a distance r. PMFs are central to many solvation processes. For example, some pairs of ions form a stable contact pair or salt bridge; others form a stable solvent-separated pair and have intervening bridging water molecules. The PMF is relevant to understanding the solubilities of salts in water, the equilibria and dynamics when water molecules

Special Issue: James L. Skinner Festschrift

Received: January 31, 2014
Published: March 13, 2014
hydrogen atoms of charged protein groups, bringing the contact depth of GB PMFs in better agreement with explicit solvent for the tested salt bridges. Recently, Nguyen et al. did more systematic fits to improve the GB solvent model parameters for protein simulations. To account for the proper coupling between nonpolar and polar solvation energies in implicit-solvent models, Dzubiella et al. proposed a variational formalism where the Gibbs free energy of the system is expressed as a functional of the solvent volume exclusion function. The method captures the sensitivity of solvent expulsion in cases of simple spherical solutes. Morgan et al. proposed an extension of the GB model to correctly describe the solvent-excluded volume of each pair of atoms, and that improved the nonbonded PMFs. An empirical model of Chen and Brooks captures the context dependence of the effective surface tension, and the authors showed that it can resolve deficiencies in PMFs of nonpolar peptide side chain analogues.

Our goal here is an approach to computing ion–ion PMFs that is both fast and accurate. We first perform MD simulations in explicit solvent as our sort of gold standard for what we aim to capture in a fast computational method. We describe here a way—which we call interpolation PMF (i-PMF)—to determine any ion–ion PMF from this set of presimulated MD explicit-water PMF simulations. Thus, i-PMF gets its physical accuracy from this precomputation stage. i-PMF gets its speed because we simply capture the explicit results by simple interpolations of the precomputed results that we store in look-up tables. We show here that this simple approach results in fast and accurate PMF determinations. i-PMF is in the spirit of a recent solvation model called SEA (semi-implicit assembly) in which solvation free energies of single spheres are computed through an explicit-simulation precompute step, and then assembled into the solute of interest at run-time by summing regional free energies.

We describe below the two components of the i-PMF method: First, we performed extensive MD simulations of small spherical ions of different charges and radii, and harvested their PMFs. Second, we describe our interpolation algorithm, the i-PMF, that allows us to capture the presimulated MD results with interpolation fits. Next, we show that the i-PMF method gives good agreement with MD and experimental results.

### THEORETICAL METHODS

#### Details of the Ion–Ion PMF Simulations by MD in Explicit Solvent.

We performed MD simulations, mostly on the TITAN supercomputer at Oak Ridge National Laboratory, using version 4.6.2 of GROMACS. Each simulation consisted of two charged solute particles and 635 water molecules in a rhombic dodecahedron box. Periodic boundary conditions were applied. All simulations were performed in the isothermal–isobaric ensemble at 298.15 K and 1 atm by applying the Nosé–Hoover thermostat (coupling constant of 2 ps) and Parrinello–Rahman barostat (coupling constant of 10 ps). The equations of motion were integrated using the leapfrog algorithm with a time step of 0.2 ps. The smooth version of the particle mesh Ewald (PME) was used to treat electrostatics (grid spacing of 0.12, PME order of 4, real-space cutoff of 9 Å, and Ewald screening parameter of 0.347 Å⁻¹), and a 9 Å cutoff for Lennard-Jones (LJ) interactions was used along with long-ranged dispersion corrections for energy and pressure. The LJ size parameter (σ) of the solutes ranged from 2 to 5.5 Å in increments of 0.5 Å, while the LJ energy parameter (ε) was set to 0.1 kcal mol⁻¹ for all solutes. The choice for the σ range and for the ε value was based on the values for the alkali metal (Na, K, Rb, Cs) and halogenide (Cl⁻, Br⁻, I⁻) ion LJ parameters of the Dang force field. The solute particles were each given a formal charge of −1 or +1. We simulated all possible combinations of solute pairs (of different sizes and charges), a total of 136 combinations (+1:−1, +1:+1, and −1:−1 ion pairs) for a given water model. We performed simulations using the TIP3P water model. Lorentz–Berthelot mixing rules were used for σ and ε.

We obtained the PMFs from MD simulations in which constraints were applied to hold the solute pairs at a series of fixed interparticle separations during the course of the simulation. The SHAKE algorithm was used to apply the interparticle constraint. The closest distance between the solute particles was 1.6 Å, and the largest interparticle distance was 12 Å. In the interval from 1.6 to 5 Å, the separation step was 0.1 Å, from 5 to 8 Å it was 0.2 Å, and 0.4 Å increments were used for the rest (up to 12 Å). Therefore, to construct the PMF for a given solute pair, 60 independent simulations were performed where, after energy minimization and 0.1 ns equilibration, the trajectory over 3 ns was collected (in total, this amounted to 136 × 60 = 8160 independent MD runs). The individual PMF was calculated by integrating the average mean force over the solute separation distance, while the entropic force due to the increase in phase space with solute separation was added to the average mean force prior to integration. Uncertainties in the average mean force values were estimated from the limiting value of the block averages, and the errors in PMF were calculated by integrating the variances of the mean force.

In short, in the presimulation step of the i-PMF method, we performed explicit-solvent (TIP3P water) MD simulations to obtain the PMFs for different combinations of solute charges (Q, Q) and LJ size parameters (σ, σ). We call these the "presimulated PMFs". We tabulated the values of...
the presimulated PMFs for given \((Q_i, Q_j)\) and \((\sigma_{ij,1}, \sigma_{ij,2})\) at the end of the Supporting Information file. To validate that the i-PMF method is predictive, we performed additional test MD runs for selected charge and size combinations of ions with the same protocol. All results apply to systems at 298.15 K and 1 atm. The LJ energy parameter was the same for all ions \((\epsilon_{ij,\text{ion}} = 0.1 \text{ kcal mol}^{-1})\).

**Interpolation Procedure for Capturing the Presimulated MD PMFs.** Our procedure handles three independent variables. By definition, a PMF is the free energy as a function of distance \(r\) between the two ions. In addition, here we aim to account for two other variables at the same time, \(\sigma_{ij,1}\) and \(\sigma_{ij,2}\), the diameters of the two ions, each of which has a given fixed charge. Below is the procedure for performing this interpolation from our MD simulation data, and we use this process to determine PMFs for any given pair of charges on the ions. Here is a short summary of how interpolations are performed on two variables at the same time. Suppose that we are given a matrix of functional values, \(f(x_i, y_j)\), that correspond to points on a two-dimensional \((n \times m)\) grid of points \((x_i, y_j)\). Here, \(i\) goes from 1 to \(n\) and \(j\) goes from 1 to \(m\). In a two-dimensional interpolation, we want to estimate, by interpolation, the function \(f\) at some point not tabulated \((x, y)\)—we want to find \(f(x, y)\). The point \((x, y)\) falls into a grid square; i.e., four tabulated points surround the desired interior point. Say that \(x_1 \leq x \leq x_{i+1}\) and \(x_j \leq y \leq y_{j+1}\); let us denote these four functional points in the following way

\[
\begin{align*}
  f_1 &= f(x_1, y) \\
  f_2 &= f(x_{i+1}, y) \\
  f_3 &= f(x_{i+1}, y_{j+1}) \\
  f_4 &= f(x_i, y_{j+1})
\end{align*}
\]

In the simplest case of interpolation on the grid square—*bilinear interpolation*—the interpolation formula is\(^5\)

\[
f(x, y) = (1 - t)(1 - u)f_1 + t(1 - u)f_2 + tu_3 + (1 - t)u_4
\]

where \(t \in [0, 1]\) and \(u \in [0, 1]\) are

\[
\begin{align*}
  t &= (x - x_i)/(x_{i+1} - x_i) \\
  u &= (y - y_j)/(y_{j+1} - y_j)
\end{align*}
\]

In order to obtain a smoother interpolated surface, higher-order interpolation techniques are applied. For example, in *bicubic spline* interpolation, we enforce the smoothness of some of the derivatives as the interpolating point cross grid square boundaries. To do a bicubic interpolation within a grid square, one needs to know the function \(f\) and the derivatives \(f_x = df/dx\), \(f_y = df/dy\), \(f_{xx} = d^2f/dx^2\), \(f_{yy} = d^2f/dy^2\), \(f_{xy} = d^2f/dxdy\) at each of the four corners of the square. The values of the derivatives at the grid points are determined globally by one-dimensional splines. The interpolating function has the following form\(^5\)

\[
f(x, y) = \sum_{m=1}^{4} \sum_{n=1}^{4} \epsilon_{mn} x^{n-1} y^{m-1}
\]

where and we need to find 16 coefficients \(\epsilon_{mn}\) by first solving the equations for the square grid points. The function \(f\) was in our case the PMF, and \(x_i\) and \(y_j\) were the \(\sigma_{ij}\). Besides bilinear and cubic spline interpolation, we tested the performance of two additional interpolation techniques (nearest-neighbor and piecewise-cubic-Hermite-polynomial interpolations). For a more detailed discussion regarding the two-dimensional interpolation algorithms, see refs 52 and 53.

Now, we describe the i-PMF interpolation methodology for capturing the PMF for any particular pair of ion sizes and charges and water model. For a given pair of solutes with charges \(Q_i\) and \(Q_j\) \((Q = +1\) or \(-1)\) and for a given interparticle separation distance \(r\), a matrix of PMF values corresponding to different combinations of the solutes’ LJ size parameters \((\sigma_{ij,1} \text{ and } \sigma_{ij,2})\) was built from the presimulated PMFs. The span of \(\sigma_{ij,1} \text{ and } \sigma_{ij,2}\) was from 2 to 5.5 Å in increments of 0.5 Å, so the matrix dimension was \(8 \times 8\). In total, 60 such matrices were constructed, each belonging to a given interparticle separation distance \(r \in [1.6, 12] \text{ Å}\) (see section above). For every matrix, a two-dimensional interpolation fit through the grid points was performed. For that purpose, we used the function “interp2” implemented in the GNU Octave software (version 3.6.2). A matrix of interpolated values was constructed and used as a look-up table to determine the value of the PMF at a given \(r\) for an arbitrary pair of \(\sigma_{ij,1}, \sigma_{ij,2} \in [2, 5.5] \text{ Å}\) values. \(\sigma_{ij,1}\) and \(\sigma_{ij,2}\) are different from the size parameters \(\sigma_{ij,1}\) and \(\sigma_{ij,2}\) of the ions used in the grid. Figure 2 shows a cartoon representation of the proposed algorithm for PMF prediction.

We stored the interpolated values of the PMF for every \(\sigma_{ij,1}\) in the range from 2 to 5.5 Å in increments of \(\Delta \sigma_{ij,1} = 0.01 \text{ Å}\) (this means that the matrix of interpolated values was of dimension \(351 \times 351\) at each \(r\)). Although, for example, \(\sigma_{ij}\) values of ions in Dang force field are given to five significant digits (see Table 1), no significant differences in predicted PMFs were observed if smaller \(\Delta \sigma_{ij}\) were used to store the interpolated data.

We tested different interpolation algorithms. By comparing the PMF obtained via an interpolation scheme and a full explicit-solvent simulation, we concluded that the cubic spline interpolation method worked best in the majority of cases. Details of this comparison are shown below.

For our initial i-PMF implementation, it takes approximately 8 s on a personal computer (PC) to build tables of presimulated values, interpolating through the grid points, making the look-up tables of interpolated values, and extracting from them the PMF for a desired combination of charges and ion sizes. After the interpolation is performed a single time and the look-up tables are available, the PMF can be extracted in 2–3 s on a PC. This is a significant speed-up compared to MD simulations. Our MD PMFs required a...
cluster of computers with many processors (the value of the PMF at each \( r \) is its own simulation), and in total, it takes more than 100 h of CPU time to get a single PMF.

Comparison of Different Interpolation Schemes.

Figure 3 shows the results of the i-PMF method utilizing four different interpolation algorithms. A size asymmetric +1::−1 system mimicking potassium iodide (\( \sigma_{\text{Li}+} = 3.3345 \) Å, \( \sigma_{\text{I}−} = 5.1705 \) Å; see Table 1) at infinite dilution was selected as a representative case. A similar analysis of i-PMF predictions in other cases shown in this paper is provided as a representative case. A similar analysis of i-PMF predictions in other cases shown in this paper is provided as a representative case. A similar analysis of i-PMF predictions in other cases shown in this paper is provided as a representative case. A similar analysis of i-PMF predictions in other cases shown in this paper is provided as a representative case.

The bilinear interpolation underestimates the depth of the first minimum, corresponding to the contact ion pair (CP) peak and underestimates the values at short r (repulsive wall) are shifted to somewhat larger r. The value of the CP which follows from piecewise cubic Hermite polynomial interpolation lies on the border of the experimental error, but the rest of the PMF agrees within the error bars with the MD predictions in other cases shown in this paper is provided as a representative case. A similar analysis of i-PMF predictions in other cases shown in this paper is provided as a representative case. A similar analysis of i-PMF predictions in other cases shown in this paper is provided as a representative case. A similar analysis of i-PMF predictions in other cases shown in this paper is provided as a representative case.

Here, we want to find the best interpolation method. The goodness of the interpolation schemes for the i-PMF method can be expressed in terms of the reduced \( \chi^2 \) parameter (see the Supporting Information file for the definition). Smaller \( \chi^2 \) values indicate better agreement between the predicted and simulated PMF. The values of \( \chi^2 \) are 22.0 for bilinear interpolation, 8.2 for nearest-neighbor interpolation, 1.8 for piecewise cubic Hermite polynomial interpolation, and 0.4 for cubic spline interpolation algorithm. The large value of \( \chi^2 \) in the case of the bilinear interpolation algorithm originates in the differences between the predicted and simulated PMF at small interparticle separation distances (the repulsive wall of the PMF). The values of the PMF on the steep repulsive wall are on the order of 10^9 kcal mol⁻¹ (and larger), but the error in the case of +1::−1 ion pairs is small (~10⁻² kcal mol⁻¹). Even though the differences between the predicted and simulated PMFs are only around 1 kcal mol⁻¹, the \( \chi^2 \) becomes extremely large due to the small error in the
simulated PMF. This region of the PMF is not very important when one discusses the structural aspects of ion–ion and ion–water interactions (the $\chi_{\text{red}}^2$ was indeed defined in such a way as to avoid points on the repulsive wall; see the Supporting Information file). If two points on the steep repulsive wall, lying just before the CP minimum ($r < 3.6$ Å), are not taken into consideration when calculating $\chi_{\text{red}}$, then the value of $\chi_{\text{red}}^2$ for bilinear interpolation lowers to 6.1. This smaller value compared to the $\chi_{\text{red}}$ for the nearest-neighbor case agrees with the previous visual judgment of a better overall prediction of the bilinear interpolation compared to a nearest-neighbor fit. As a rule, a value of $\chi_{\text{red}}^2 = 1$ suggests that the extent of the agreement between observations and estimates is in accord with the error variance. $\chi_{\text{red}}^2 > 1$ indicates that the fit has not fully captured the data (or that the error estimate has been underestimated), while $\chi_{\text{red}}^2 < 1$ speaks in favor of an unexpectedly good fit. We have concluded that (in the majority of cases) the most satisfactory interpolation algorithm for predicting the PMF of opposite-charged as well as like-charged PMFs involves the cubic spline interpolation (see Tables S1–S4 of the Supporting Information file). The method gives quantitative agreement within the error bars in most of the relevant cases displayed in the rest of the figures and in data presented in the Supporting Information. In the following subsections, we show a few characteristic cases that further demonstrate the quality of the proposed PMF prediction scheme.

Here our studies are limited. We use only a single value of the LJ energy parameter ($\epsilon_{\text{1J}}$), regardless of the charge or size of the ions. This is true for the selected ions in the Dang ion force field (cf. Table 1), but in principle, each ion could require its own $\epsilon_{\text{1J}}$ value. The present approach could readily be generalized but at substantial computational expense. In that case, the presimulated PMFs would need to be collected for different combinations of $\epsilon_{\text{1J}}$. Instead of performing a two-dimensional interpolation, a four-dimensional interpolation through the PMF values corresponding to ($\sigma_{\text{1J}}, \epsilon_{\text{1J},1}, \sigma_{\text{1J},2}, \epsilon_{\text{1J},2}$) would need to be conducted.

## RESULTS AND DISCUSSION

Here, we compare the results of the i-PMF method to the results of the precomputed MD simulations. We find that i-PMF works quite well across the whole range of solutes, from small ions with high charge density to large ones having small charge density. It captures the position and the depth of the characteristic wells corresponding to the contact and solvent shared/-separated pairs and the height of the peaks of the solvation barriers. The method is even capable of capturing the complex behavior of the PMF in the case of small ions having a large charge density (<3 Å), where the path from the solvent shared to solvent separated pair involves multiple transition barriers.

Up to a distance of the CP minimum (i.e., for the values on the repulsive wall), the cubic Hermite polynomial interpolation performs slightly better than cubic spline interpolation (not shown). In an extreme case of $\sigma_{\text{1J}} = 2.25$ Å, the cubic Hermite polynomial interpolation works better up to the distance of the desolvation barrier. However, we need to stress that in these cases the selected size parameters of ions are small (2.25 Å) and do not correspond to practical model situations. We selected this example to demonstrate that the interpolation method works in “extreme” cases (inset in Figure 4). Improved performance of the interpolation could be achieved by providing a finer grid at the size edges, i.e., between 2 and 2.5 Å, and equivalently between 5 and 5.5 Å for larger solutes. This would, however, require an increase in the number of PMFs calculated in the presimulation step. Figure S1 of the Supporting Information shows how the interpolated PMF values change with increasing size of the ions (size symmetric case) for few selected interparticle separation distances. Comparison to the simulated values shows that the deviations between predicted and simulated values are the largest for solutes between 2 and 2.5 Å. The $\chi_{\text{red}}^2$ values for the cases displayed in Figure 4 are given in Table S2 of the Supporting Information file. In the cases where $\chi_{\text{red}}^2$ is smaller for the cubic Hermite polynomial interpolation compared to cubic spline interpolation, this originates from the data on the repulsive wall. For the relevant domains of the PMF, the cubic spline interpolation gives the best results, as already discussed.

We also find that, in addition to working for ions having identical sizes, i-PMF works for ions with different sizes. Figure 5 shows PMFs for combinations of small and large cations (Na$^+$ and Cs$^+$, respectively) with small and large anions (Cl$^-$ and I$^-$, respectively), using ion parameters listed in Table 1. Results of predicted PMFs for NaCl and NaI ion pairs in TIP3P water are given in panel a, and PMFs for CsCl and CsI, in panel b. Displayed MD simulation data (symbols) were obtained independently of the prediction scheme. In all cases, the predicted PMF (cubic spline interpolation through a presimulated grid of PMF values) agrees quantitatively with the reference MD simulations. Results for other combinations of ions given in Table 1 are shown in Figure S3 of the Supporting Information file. The method performs somewhat worse in the case of KCl, KBr, RbBr, and CsBr where some parts of the PMF are either slightly over- or underestimated compared to corresponding MD results. The differences are, however, small, and the agreement can still be considered very good. The quality of the prediction was estimated for all the shown cases in terms of the $\chi_{\text{red}}^2$ parameter (Table S1 of the Supporting Information file). Out of all four interpolation algorithms tested here, the cubic spline interpolation performs the best. We find that i-PMF gives more accurate PMFs.
more challenging than the previous size-symmetric solute pairings. Na+ and Cl−
dissociated states, one can calculate counting the time that the ions spend in the associated and
simulations where the free energy change is obtained by simulations.55
been studied by integral-equation theories and computer

\[ \Delta \ln K_{ie} = \ln \left( \frac{[X]}{[Y]} \right) \]

where \([X]\) denotes the molar concentration of species \(X\) and \(\Delta \)
attained to the equilibrium constant via a thermodynamic relation 

\[ K_{ie} = \exp \left( \frac{-\Delta G^\circ}{RT} \right) \]

\(A^+ + B^- \rightarrow AB\) de
de ned as 

\[ [AB] \frac{r}{R_{\text{max}}}^{2} g_{AB}(r) dr \]

\[ c_{\text{tot}} = \frac{3}{(N_{A}A_{\text{e}})} \]

\(A^+ = [A^+] = [B^-] = c_{\text{tot}} - [AB] \). \(R_{\text{max}}\) is the largest interparticle separation distance (12 Å in our case). The total concentration of the electrolyte is 

\[ c_{\text{tot}} = \frac{3}{(N_{A}A_{\text{e}})} \]

and the concentration of the free ions is \([A^+] = [B^-] = c_{\text{tot}} - [AB] \). \(R_{\text{max}}\) is the largest interparticle separation distance (12 Å in our case). The total concentration of the electrolyte is 

\[ c_{\text{tot}} = \frac{3}{(N_{A}A_{\text{e}})} \]

The apparent association constant for the process of forming an ion pair is in the case of univalent ions \((A^+ + B^- \Rightarrow AB)\) defined as 

\[ K_{i} = \left( [AB]/[A^+] \right) \left( [B^-]/[B^-] \right) = \frac{1}{c_{\text{tot}}} \]

where \([X]\) denotes the molar concentration of species \(X\) and \(c_{\text{tot}}\) is the standard molar concentration. The change in standard free energy of ion pair formation is related to the equilibrium constant via a thermodynamic relation 

\[ \Delta G^\circ = -RT \ln K_{i} \]

By conducting equilibrium MD simulations where the free energy change is obtained by counting the time that the ions spend in the associated and dissociated states, one can calculate \(K_{i}\). This, however, requires running long explicit-solvent MD simulations for a given ion pair in order to obtain reasonable statistics. An alternative is to calculate the association constants from the radial distribution function. Both approaches imply that a definition of associated/dissociated state is given. Here, we define an ion pair as a configuration of a cation and anion where the two are not separated more than a certain distance \(r_{CP}\). We take this distance to be the first minimum in the cation–anion radial distribution function \(g_{AB}(r)\), corresponding to the so-called contact pair distance. For all \(r > r_{CP}\), we treat ions as free. Concentrations of free ions and of the ion pair can be estimated by integrating the \(g_{AB}(r)\). The ratio of the concentration of the ion pair to the total concentration of the electrolyte \((c_{\text{tot}})\) is equal to

\[ c_{\text{tot}} = \frac{3}{(N_{A}A_{\text{e}})} \]

\(A^+ = [A^+] = [B^-] = c_{\text{tot}} - [AB] \). \(R_{\text{max}}\) is the largest interparticle separation distance (12 Å in our case). The total concentration of the electrolyte is 

\[ c_{\text{tot}} = \frac{3}{(N_{A}A_{\text{e}})} \]

and the concentration of the free ions is \([A^+] = [B^-] = c_{\text{tot}} - [AB] \). \(R_{\text{max}}\) is the largest interparticle separation distance (12 Å in our case). The total concentration of the electrolyte is 

\[ c_{\text{tot}} = \frac{3}{(N_{A}A_{\text{e}})} \]

and the concentration of the free ions is \([A^+] = [B^-] = c_{\text{tot}} - [AB] \). \(R_{\text{max}}\) is the largest interparticle separation distance (12 Å in our case). The total concentration of the electrolyte is 

\[ c_{\text{tot}} = \frac{3}{(N_{A}A_{\text{e}})} \]

and the concentration of the free ions is \([A^+] = [B^-] = c_{\text{tot}} - [AB] \). \(R_{\text{max}}\) is the largest interparticle separation distance (12 Å in our case). The total concentration of the electrolyte is 

\[ c_{\text{tot}} = \frac{3}{(N_{A}A_{\text{e}})} \]

and the concentration of the free ions is \([A^+] = [B^-] = c_{\text{tot}} - [AB] \). \(R_{\text{max}}\) is the largest interparticle separation distance (12 Å in our case). The total concentration of the electrolyte is 

\[ c_{\text{tot}} = \frac{3}{(N_{A}A_{\text{e}})} \]

and the concentration of the free ions is \([A^+] = [B^-] = c_{\text{tot}} - [AB] \). \(R_{\text{max}}\) is the largest interparticle separation distance (12 Å in our case). The total concentration of the electrolyte is 

\[ c_{\text{tot}} = \frac{3}{(N_{A}A_{\text{e}})} \]

and the concentration of the free ions is \([A^+] = [B^-] = c_{\text{tot}} - [AB] \). \(R_{\text{max}}\) is the largest interparticle separation distance (12 Å in our case). The total concentration of the electrolyte is 

\[ c_{\text{tot}} = \frac{3}{(N_{A}A_{\text{e}})} \]

and the concentration of the free ions is \([A^+] = [B^-] = c_{\text{tot}} - [AB] \). \(R_{\text{max}}\) is the largest interparticle separation distance (12 Å in our case). The total concentration of the electrolyte is 

\[ c_{\text{tot}} = \frac{3}{(N_{A}A_{\text{e}})} \]

and the concentration of the free ions is \([A^+] = [B^-] = c_{\text{tot}} - [AB] \). \(R_{\text{max}}\) is the largest interparticle separation distance (12 Å in our case). The total concentration of the electrolyte is 

\[ c_{\text{tot}} = \frac{3}{(N_{A}A_{\text{e}})} \]

and the concentration of the free ions is \([A^+] = [B^-] = c_{\text{tot}} - [AB] \). \(R_{\text{max}}\) is the largest interparticle separation distance (12 Å in our case). The total concentration of the electrolyte is 

\[ c_{\text{tot}} = \frac{3}{(N_{A}A_{\text{e}})} \]

and the concentration of the free ions is \([A^+] = [B^-] = c_{\text{tot}} - [AB] \). \(R_{\text{max}}\) is the largest interparticle separation distance (12 Å in our case). The total concentration of the electrolyte is 

\[ c_{\text{tot}} = \frac{3}{(N_{A}A_{\text{e}})} \]

and the concentration of the free ions is \([A^+] = [B^-] = c_{\text{tot}} - [AB] \). \(R_{\text{max}}\) is the largest interparticle separation distance (12 Å in our case). The total concentration of the electrolyte is 

\[ c_{\text{tot}} = \frac{3}{(N_{A}A_{\text{e}})} \]

and the concentration of the free ions is \([A^+] = [B^-] = c_{\text{tot}} - [AB] \). \(R_{\text{max}}\) is the largest interparticle separation distance (12 Å in our case). The total concentration of the electrolyte is 

\[ c_{\text{tot}} = \frac{3}{(N_{A}A_{\text{e}})} \]

and the concentration of the free ions is \([A^+] = [B^-] = c_{\text{tot}} - [AB] \). \(R_{\text{max}}\) is the largest interparticle separation distance (12 Å in our case). The total concentration of the electrolyte is 

\[ c_{\text{tot}} = \frac{3}{(N_{A}A_{\text{e}})} \]

and the concentration of the free ions is \([A^+] = [B^-] = c_{\text{tot}} - [AB] \). \(R_{\text{max}}\) is the largest interparticle separation distance (12 Å in our case). The total concentration of the electrolyte is 

\[ c_{\text{tot}} = \frac{3}{(N_{A}A_{\text{e}})} \]

and the concentration of the free ions is \([A^+] = [B^-] = c_{\text{tot}} - [AB] \). \(R_{\text{max}}\) is the largest interparticle separation distance (12 Å in our case). The total concentration of the electrolyte is 

\[ c_{\text{tot}} = \frac{3}{(N_{A}A_{\text{e}})} \]

and the concentration of the free ions is \([A^+] = [B^-] = c_{\text{tot}} - [AB] \). \(R_{\text{max}}\) is the largest interparticle separation distance (12 Å in our case). The total concentration of the electrolyte is 

\[ c_{\text{tot}} = \frac{3}{(N_{A}A_{\text{e}})} \]

and the concentration of the free ions is \([A^+] = [B^-] = c_{\text{tot}} - [AB] \). \(R_{\text{max}}\) is the largest interparticle separation distance (12 Å in our case). The total concentration of the electrolyte is 

\[ c_{\text{tot}} = \frac{3}{(N_{A}A_{\text{e}})} \]
limits of the estimated errors. The predicted $K_{a}$ is systematically smaller than the reference one in the cases of KCl, KBr, and RbBr, and larger in the case of RbI and CsI. Small differences in predicted and simulated PMFs accumulate in the integration process, resulting in larger $K_{a}$ differences than one might expect through visually comparing both PMFs. Also, the $R_{\text{max}}$ in eq 5 should in principle be $\infty$, but we were restricted by the dimensions of the simulation box. Differences in PMFs at larger interparticle separation distances amplify by integration and are reflected in estimated $K_{a}$. In Figure S6 of the Supporting Information, we show predicted association constants for size symmetric +1−1 ion pairs as a function of ion sizes. We see that the predicted $K_{a}$ agree well with constants determined from reference MD simulation PMFs.

Finally, we explore the “law of matching water affinities” of K. D. Collins.72 In short, this is the idea that small−small ion pairs are insoluble in water (because of strong electrostatic attractions), that large−large ion pairs are poorly soluble in water (because they act largely like hydrophobic spheres), but that large−small ion pairs are very soluble in water. Here, we make a simple test using our interpolated PMFs. Since we can consider the CP state as a precipitated form of the salt, the $K_{a}$ correlates with the solubility of the salt in question. Figure 7 shows inverse values of the experimentally determined solubilities,33 $S^{-1}$ (panel a), and ion-association constants from predicted PMFs, $K_{a}$ (panel b), for selected alkali metal halides. The associated content provides further details.

![Figure 7](image-url)
the end of the Supporting Information file, we give tabulated values of the precalculated PMFs. All results are given for \( T = 298.15 \) K and \( p = 1 \) atm. The ions were immersed in the TIP3P water model, and Dang LJ parameters were used to describe the ions (\( \varepsilon_{1\text{ion}} = 0.1 \) kcal mol\(^{-1}\) for all ions). This material is available free of charge via the Internet at http://pubs.acs.org.

**AUTHOR INFORMATION**

**Corresponding Author**

*E-mail: christopher.fennell@okstate.edu*

**Notes**

The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

M.L. thanks Dr. Evangelos Coutsias for discussions regarding the interpolation techniques. We acknowledge the ALCC project BIO-023 which enabled us to perform simulations on the TITAN supercomputer. M.L. appreciates the support of the Stony Brook Laufer Center and partial support of the Slovenian Research Agency (ARRS) through Program P1-0201 and Project J1-4148. The authors appreciate the support of National Institutes of Health Grant GM063592.

**REFERENCES**

(1) Kirkwood, J. G. Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys.* 1935, 3, 300.
(2) Hansen, J.-P.; McDonald, I. R. *Theory of Simple Liquids*, 3rd ed.; Elsevier: Amsterdam, The Netherlands, 2006.
(3) Roux, B. The Calculation of the Potential of Mean Force using Computer Simulations. *Comput. Phys. Commun.* 1995, 91, 275–282.
(4) Trzesniew, D.; Kunz, A.-P. E.; van Gunsteren, W. F. A Comparison of Methods to Compute the Potential of Mean Force. *ChemPhysChem* 2007, 8, 162–169.
(5) Hess, B.; Holm, C.; van der Vegt, N. Osmotic Coefficients of Aromatic NaCl (aq) Force Fields. *J. Chem. Phys.* 2006, 124, 164509.
(6) Trudinger, S.; Padró, J. A. Potential of Mean Force by Constrained Molecular Dynamics: A Sodium Chloride Ion-Pair in Water. *Chem. Phys. Phys.* 1991, 155, 187–195.
(7) Darve, E.; Pohorille, A. Calculating Free Energies Using Average Force. *J. Chem. Phys.* 2001, 115, 9169–9183.
(8) Beutler, T. C.; van Gunsteren, W. F. The Computation of a Potential of Mean Force: Choice of the Biasing Potential in the Umbrella Sampling Technique. *J. Chem. Phys.* 1994, 100, 1492–1497.
(9) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* 1992, 13, 1011–1021.
(10) Hirata, F.; Rossky, P. J.; M. P. B. The Interionic Potential of Mean Force in a Molecular Polar Solvent From an Extended RISM Equation. *J. Chem. Phys.* 1983, 78, 4133.
(11) Kovalenko, A.; Hirata, F. Potential of Mean Force between Two Molecular Ions in a Polar Molecular Solvent: A Study by the Three-Dimensional Reference Interaction Site Model. *J. Phys. Chem. B* 1999, 103, 7942–7957.
(12) Kovalenko, A.; Hirata, F. Potentials of Mean Force of Simple Ions in Ambient Aqueous Solution. I. Three-Dimensional Reference Interaction Site Model Approach. *J. Chem. Phys.* 2000, 112, 10391.
(13) Kovalenko, A.; Hirata, F. Potentials of Mean Force of Simple Ions in Ambient Aqueous Solution. II. Solvation Structure from the Three-Dimensional Reference Interaction Site Model Approach, and Comparison with Simulations. *J. Chem. Phys.* 2000, 112, 10403.
(14) Roux, B.; Simonson, T. Implicit Solvent Models. *Biophys. Chem.* 1999, 78, 1–20.
(15) Kovalenko, A.; Hirata, F. Potentials of Mean Force of Simple Ions in Ambient Aqueous Solution. *Biophys. Chem.* 2000, 78, 1–20.
(16) Roux, B.; Simonson, T. Implicit Solvent Models. *Comput. Phys. Commun.* 2000, 112, 10403.
(17) Simonson, T. Electrostatics and Dynamics of Proteins. *Rep. Prog. Phys.* 2003, 66, 737–787.
(18) Feig, M.; Brooks, C. L. I. Recent Advances in the Development and Application of Implicit Solvent Models in Biomolecule Simulations. *Curr. Opin. Struct. Biol.* 2004, 14, 217–224.
(19) Baker, N. A. Improving Implicit Solvent Simulations: a Poisson-Centric View. *Curr. Opin. Struct. Biol.* 2005, 15, 137–143.
(20) Gilson, M. K.; Davis, M. E.; Luty, B. A.; McCammon, J. A. Computer of Electrostatic Forces on Solvated Molecules Using the Poisson-Boltzmann Equation. *J. Phys. Chem.* 1993, 97, 3591–3600.
(21) Im, W.; Beglov, D.; Roux, B. Continuum Solvation Model: Computation of Electrostatic Forces from Numerical Solutions to the Poisson-Boltzmann Equation. *Comput. Phys. Commun.* 1998, 111, 59–75.
(22) Lam, G. In *Reviews in Computational Chemistry*: Lipkowski, K. B.; Larter, R., Candari, T. B., Eds.; John Wiley & Sons, Inc.: Hoboken, New Jersey, 2003; pp 147–365.
(23) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. Semi-analytical Treatment of Solvation for Molecular Mechanics and Dynamics. *J. Am. Chem. Soc.* 1990, 112, 6127–6129.
(24) Bashford, D.; Case, D. A. Generalized Born Models of Macromolecular Solvation Effects. *Annu. Rev. Phys. Chem.* 2000, 51, 129–152.
(25) Tsui, V.; Case, D. A. Theory and Applications of the Generalized Born Solvation Model in Macromolecular Simulations. *Biopolymers* 2000, 56, 275–291.
(26) Onufriev, A.; Bashford, D.; Case, D. A. Modification of the Generalized Born Model Suitable for Macromolecules. *J. Phys. Chem. B* 2000, 104, 3712–3720.
(27) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, L. M. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* 1983, 79, 926–933.
(28) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. Electrostatics of Nanosystems: Application to Microtubules and the Ribosome. *Proc. Natl. Acad. Sci. U.S.A.* 2001, 98, 10037–10041.
(29) Cramer, C. J.; Truhlar, D. G. Implicit Solvation Models: Equilibria, Structure, Spectra, and Dynamics. *Chem. Rev.* 1999, 99, 2161–2200.
(30) Marten, B.; Kim, K.; Cortis, C.; Friesner, R. A.; Murphy, R. B.; Ringnalda, M. N.; Sitkoff, D.; Honig, B. New Model for Calculation of Solvation Free Energies: Correction of Self-Consistent Reaction Field Continuum Dielectric Theory for Short-Range Hydrogen-Bonding Effects. *J. Phys. Chem.* 1996, 100, 11775–11788.
(31) Swanson, J. M. J.; Mongan, J.; McCammon, J. A. Limitations of Atom-Centered Dielectric Functions in Implicit Solvent Models. *J. Phys. Chem. B* 2005, 109, 14769–14772.
(32) Zhou, R.; Berne, B. J. Can a Continuum Solvent Model Reproduce the Free Energy Landscape of a β-Hairpin Folding in Water? *Proc. Natl. Acad. Sci. U.S.A.* 2002, 99, 12777–12782.
(33) Yu, Z.; Jacobson, M. P.; Josovitz, J.; Rapp, C. S.; Friesner, R. A. First-Shelf Solvation of Ion Pairs: Correction of Systematic Errors in Implicit Solvent Models. *J. Phys. Chem. B* 2004, 108, 6643–6654.
(34) Geney, R.; Layten, M.; Gomperts, R.; Hornak, V.; Simmerling, C. Investigation of Salt Bridge Stability in a Generalized Born Solvent Model. *J. Chem. Theory Comput.* 2006, 2, 115–127.
(35) Nguyen, H.; Roe, B. D.; Simmerling, C. Improved Generalized Born Solvent Model Parameters for Protein Simulations. *J. Chem. Theory Comput.* 2013, 9, 2020–2034.
(36) Dzubiella, J.; Swanson, J. M. J.; McCammon, J. A. Coupling Nonpolar and Polar Solvation Free Energies in Implicit Solvent Models. *J. Chem. Phys.* 2006, 124, 084905–1–084905–12.
(37) Mongan, J.; Simmerling, C.; McCammon, J. A.; Case, D. A.; Onufriev, A. Generalized Born Model with a Simple, Robust Molecular Volume Correction. *J. Chem. Theory Comput.* 2007, 3, 156–169.

(38) Chen, J.; Brooks, C. L. I. Critical Importance of Length-Scale Dependence in Implicit Modeling of Hydrophobic Interactions. *J. Am. Chem. Soc.* 2007, 129, 2444–2445.

(39) Fennell, C. J.; Keohoe, C.; Dill, K. A. Oil/Water Transfer Is Partly Driven by Molecular Shape, Not Just Size. *J. Am. Chem. Soc.* 2010, 132, 234–240.

(40) Fennell, C. J.; Dill, K. A. Physical Modeling of Aqueous Solution. *J. Stat. Phys.* 2011, 145, 209–226.

(41) Fennell, C. J.; Keohoe, C. W.; Dill, K. A. Modeling Aqueous Solution with Semi-Explicit Assembly. *Proc. Natl. Acad. Sci.* U.S.A. 2011, 108, 3234–3239.

(42) Li, L.; Fennell, C. J.; Dill, K. A. Field-SEA: A Model for Computing the Solvation Free Energies of Nonpolar, Polar, and Charged Solutes in Water. *J. Phys. Chem. B* 2013, DOI: 10.1021/jp4115139.

(43) http://www.olcf.ornl.gov/titan/, 2013.

(44) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* 2008, 4, 435–447.

(45) Essman, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* 1995, 103, 8577–8592.

(46) Dang, L. X.; Garrett, B. C. Photoelectron Spectra of the Hydrated Iodine Anion from Molecular Dynamics Simulations. *J. Chem. Phys.* 1993, 99, 6950–6956.

(47) Smith, D. E.; Dang, L. X. Computer Simulations of NaCl Association in Polarizable Water. *J. Chem. Phys.* 1994, 100, 3757–3766.

(48) Smith, D. E. Mechanism and Thermodynamics of Ion Selectivity in Aqueous Solutions of 18-Crown-6 Ether: A Molecular Dynamics Study. *J. Am. Chem. Soc.* 1995, 117, 6954–6960.

(49) Koneshan, S.; Rasaiah, J. C.; Lynden-Bell, R. M.; Lee, S. H. Solvent Structure, Dynamics, and Ion Mobility in Aqueous Solutions at 25°C. *J. Phys. Chem. B* 1998, 102, 4193–4204.

(50) Fennell, C. J.; Bizjak, A.; Vlachy, V.; Dill, K. A. Ion Pairing in Molecular Simulations of Aqueous Alkali Halide Solutions. *J. Phys. Chem. B* 2009, 113, 6782–6791.

(51) Hess, B. Determining the Shear Viscosity of Model Liquids from Molecular Dynamics Simulations. *J. Chem. Phys.* 2002, 116, 209–217.

(52) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical recipes: the art of scientific computing*, 3rd ed.; Cambridge University Press: Cambridge, U.K., 2007.

(53) Eaton, J. W.; Bateman, D.; Hauberg, S. *GNU Octave Manual Version 3*; Network Theory Limited: 2008.

(54) www.gnu.org/software/octave/; 2013.

(55) Dang, L. X.; Pettitt, B. M. Chloride Ion Pairs in Water. *J. Am. Chem. Soc.* 1987, 109, 5531–5532.

(56) Dang, L. X.; Pettitt, B. M. A Theoretical Study of Like Ion Pairs in Solution. *J. Phys. Chem.* 1990, 94, 4303–4308.

(57) Buckner, J. K.; Jorgensen, W. L. Energetics and Hydration of the Constituent Ion Pairs of Tetrathylammonium Chloride. *J. Am. Chem. Soc.* 1989, 111, 2507–2516.

(58) Karim, O. A. Potential of Mean Force for an Aqueous Chloride Ion Pair: Simulation with a Polarizable Model. *J. Chem. Phys.* 1992, 96, 9237–9238.

(59) Dang, L. X. Fluoride-Fluoride Association in Water from Molecular Dynamics Simulations. *Chem. Phys. Lett.* 1992, 200, 211–25.

(60) Guàrdia, E.; Rey, R.; Padró, J. A. Na⁺-Na⁺ and Cl⁻-Cl⁻ Ion Pairs in Water: Mean Force Potentials by Constrained Molecular Dynamics. *J. Chem. Phys.* 1991, 95, 2823–2831.