JCTC Journal of Chemical Theory and Computation

The Use of Anisotropic Potentials in Modeling Water and Free Energies of Hydration

Panagiotis G. Karamertzanis,*,† Paolo Raiteri,‡ and Amparo Galindo†

Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, London, SW7 2AZ, United Kingdom, and Department of Chemistry and Nanochemistry Research Institute, GPO Box U1987, 6845 Perth, Western Australia

Received December 26, 2009

Abstract: We propose a novel, anisotropic rigid-body intermolecular potential model to predict the properties of water and the hydration free energies of neutral organic solutes. The electrostatic interactions of water and the solutes are modeled using atomic multipole moments up to hexadecapole; these are obtained from distributed multipole analysis of the quantum mechanically computed charge densities and include average polarization effects in solution. The repulsion−dispersion water−water interactions are modeled with a three-site, exp-6 model fitted to the experimental liquid water density and oxygen−oxygen radial distribution function at ambient conditions. The proposed water model reproduces well several water properties not used in its parametrization, including vapor−liquid coexistence densities, the maximum in liquid water density at atmospheric pressure, the structure of ordered ice polymorphs, and the liquid water heat capacity. The model is used to compute the hydration free energy of 10 neutral organic solutes using explicit-solvent free energy perturbation. The solute−solute repulsion−dispersion intermolecular potential is obtained from previous parametrizations on organic crystal structures. In order to calculate the free energies of hydration, water−solute repulsion−dispersion interactions are modeled using Lorenz−Berthelot combining rules. The root-mean-square error of the predicted hydration free energies is 1.5 kcal mol^{-1}, which is comparable to the error found using a continuum mean-field quantum mechanical approach parametrized using experimental free energy of hydration data. The results are also contrasted with explicit-solvent hydration free energies obtained with an atomic charge representation of the solute's charge density computed at the same level of theory used to compute the distributed multipoles. Replacing the multipole description of the solute's charge density with an atomic charge model changes the free energy of hydration by as much as 3 kcal mol^{-1} and provides an estimate for the effect of the modeling quality of the intermolecular electrostatic forces in free energy of solvation calculations.

Introduction

Computational chemistry techniques are often used to model protein folding, ligand recognition and binding, partition coefficients, solubility, reaction rates, pK_a, and tautomer ratios, all of which depend crucially on the accuracy with which solvation effects can be modeled. Water is the most important and widely studied solvent due to its ubiquity in biological and industrial processes.[1] Many water models, differing in the way they treat intramolecular distortions, electrostatic interactions,[2] polarization effects,[3−5] and their parametrization strategy using experimental[6,7] or *ab initio*[8] data, have been developed. Although these water models are

* Corresponding author e-mail: p.karamertzanis@imperial.ac.uk.
† Imperial College London.
‡ Department of Chemistry and Nanochemistry Research Institute.

10.1021/ct900693q
Published on Web 04/02/2010

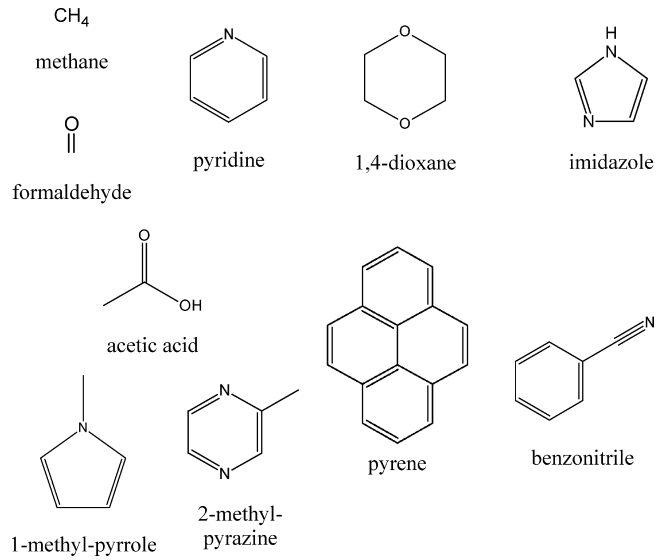
successful in predicting selected liquid water and ice polymorph properties, it is still debated what the indispensable aspects of an accurate water model are. This information is needed to construct a model of general applicability that is transferable and of tractable complexity to be of practical use in computationally intensive calculations, such as the hydration of biomolecules and modeling of nucleation.

The prediction of the free energy of hydration is a prime example of the use of water models and a stringent test of their accuracy. The free energy of hydration is associated with the tendency of the solute to leave the aqueous solution. Molecules with strong hydrogen bond donor and acceptor groups have a strongly negative ΔG_{hyd} (hydrophilic), while molecules that are poorly attracted to water (hydrophobic) have a positive ΔG_{hyd}, as the stabilization due to water−solute interactions is not sufficiently large to compensate the disruption of the energetically more favorable water−water interactions. The free energy of hydration can be experimentally determined by measuring the equilibrium constant for the solute transfer between vapor and aqueous solution, under experimental conditions and concentrations that eliminate solute self-association in both phases.[9] However, its computational prediction is of interest, as it can be used to compute partition coefficients[10] and solubility,[11,12] which are key quantities in pharmaceutical development. A wide range of methods to compute hydration free energies has been proposed.[13] The methods differ markedly in computational cost and extent of parametrization and include group additivity schemes,[14] continuum solvation models,[15] and explicit-solvent, free energy approaches[16] based on exhaustive sampling of all thermally accessible states using molecular dynamics or Monte Carlo simulations.

Explicit-solvent free energy methods have the potential to be systematically improved by minimizing the sources of statistical and systematic error until quantitative predictions of hydration free energies are obtained. Good progress has been achieved in designing better free energy methods and protocols[16−19] that minimize the error due to finite sampling. However, errors due to approximations in the intermolecular potential model have received less attention, although the accuracy of explicit-solvent hydration free energy calculations has been shown to be strongly dependent on the model for the water−water interactions[20,21] and the model for the solute's intermolecular electrostatic interactions.[22,23] Despite the known limitations in describing aqueous solutions, most water models used in hydration free energy calculations employ a point-charge approximation. Alternative representations include the use of smeared (Gaussian) charges[2,24,25] and multipoles,[3,26,27] which improve the representation of water's charge density, but their use in modeling hydration has thus far been limited.[28,29]

The modeling of the organic solid state has provided an impetus for developing accurate models for the intermolecular forces, with particular emphasis on the electrostatic[30] and induction contributions.[31] A key element of these models is the description of the intermolecular electrostatic forces with a distributed multipole model, which has been shown to improve accuracy when features such as lone-pair and π-electron densities[32,33] are present. These models have been used successfully in quantifying the small energy differences between polymorphs[33] and predicting bulk crystal properties[34,35] but have not been used to study the interaction of the crystal with the solvent in the context of nucleation and growth or solubility predictions, which would also require the development of an accurate model for the solvent.

Chart 1. Molecules Used in Hydration Free Energy Calculations



The objective of this work is to construct a high-rank multipole model for water and to evaluate its applicability in dynamic simulations that sample the whole range of water−water and solute−water molecular configurations. The model for the electrostatic forces for both water and the solute is derived from a distributed multipole analysis[36] of the quantum mechanically computed molecular charge densities and includes average (implicit) polarization effects in aqueous solutions. Hence, intermolecular electrostatic interactions are modeled accurately, avoiding the need for explicit polarization[37−39] that is computationally prohibitive in combination with a high rank, multipole model. Repulsion−dispersion parameters are fitted to liquid water experimental data in the case of water and taken from an earlier parametrization[30] for the solute. The proposed water model is found suitable for modeling the structure, and a range of properties of ordered ice polymorphs and liquid water not included in its parametrization.

We also use this anisotropic intermolecular potential model to compute the free energy of hydration of 10 rigid, uncharged organic molecules (see Chart 1) using an explicit-solvent free energy perturbation approach. The molecules are chosen so that their ΔG_{hyd} varies from −10 to +2 kcal mol^{-1}, practically covering the entire range of hydration free energies typically obtained for neutral organic solutes. We have deliberately not considered charged species because ionic hydration free energies are generally affected by strong electrostatic finite-size effects.[40] We contrast our hydration free energy calculations with predictions using a self-consistent reaction field quantum mechanical method and also with explicit-solvent free energy perturbation calculations using a point-charge model for the solute, computed at the same level of theory as used for the distributed multipole

expansions. The purpose of the latter comparison is to examine if an atomic charge representation is sufficiently accurate to describe the hydration of molecules that are strongly hydrophilic due to their aromatic character and/or hydrogen bond acceptor and donor groups. Such tests are instrumental in establishing whether discrepancies from experiment[13,22,41] in the predicted ΔG_{hyd} are due to the limited accuracy of the atomic-charge representation of the molecular charge density or, alternatively, due to the lack of explicit polarization and errors in the repulsion−dispersion parametrization.

Methodology

The solutes were optimized in isolation at the MP2(fc)/6-31G(d,p) level of theory and were treated as rigid in all subsequent calculations. Water was also kept rigid in its TIP4P conformation (OH = 0.957 Å, HOH = 104.52°), as in the TIP4P[42] and TIP4P/2005[6] water models. The charge densities used in the distributed multipole analysis[36] (DMA) and to compute the molecular electrostatic potentials (ESP) to fit atomic-charge models were computed at the MP2(fc)/aug-cc-pVTZ level of theory for both water and the solutes, apart from pyrene, for which we used PBE0/aug-cc-pVTZ due to computational limitations. All molecular optimizations, charge density, and electrostatic potential calculations were carried out in Gaussian 03.[43]

Model for Intermolecular Forces. *Water−Water Interactions.* In the seminal work of Bukowski et al.,[8] the model for the water−water interactions was computed entirely from first principles using perturbation theory and dimer CCSD(T) calculations. Unfortunately, such models inevitably include nonadditive terms that are impractical for use in the long simulation runs required for accurate hydration calculations. On the other hand, several water pair potentials have been successfully developed by fitting both repulsion−dispersion and electrostatic components to experimental data,[6,7,42,44,45] despite the difficulty in extracting details of the potential energy function from bulk property measurements that represent only averages over the potential surface.[46] The predictive ability of these models often deteriorates for properties and thermodynamic conditions not included in the parametrization. To some extent, this is due to the large number of variables that can be altered and, perhaps, to the existence of multiple optimal solutions that differ in the quality of reproduction of different properties. In this work, we combine both approaches in developing a novel, anisotropic water model: the electrostatic component is computed quantum-mechanically and includes average polarization effects in liquid water at ambient conditions (see the section Electrostatic and Polarization Interactions), while the repulsion−dispersion interaction potential is fitted to experimental liquid water data (see the section Repulsion−Dispersion Interactions).

Electrostatic and Polarization Interactions. From early on in this work, it became apparent that the structure of ice polymorphs and hydrates, as well as the properties of liquid water, water vapor−liquid equilibria, and hydration free energies, could not all be modeled accurately using a distributed multipole model derived from the isolated-water charge density. Ideally, the response of the charge density to its surroundings in condensed phases should be treated using an accurate, distributed polarizability model.[47] However, the computational cost of including explicit polarization in conjunction with a multipole representation[48] in molecular dynamics is computationally expensive and impractical for explicit-solvent free energy perturbation calculations. Hence, we opted for an implicitly polarized model that would, on average, reproduce water's charge density in condensed phases.[27]

The wave function of a water molecule A in the vicinity of water molecule B can be approximated[49] by solving Schrödinger's equation:

$$(H_A + V_{AB})|\Psi_A\rangle = E_A\Psi_A\rangle \quad (1)$$

where H_A is the Hamiltonian of the isolated molecule A and V_{AB} represents the electrostatic interaction between molecules A and B. This interaction term requires computing the electrostatic potential field generated by molecule B, which in this work was approximated by an atomic charge representation

$$V_{AB} = \sum_{b\in B} Q_{00}^{b}\left\{-\sum_{i\in A}\frac{1}{|\mathbf{r}_i - \mathbf{r}_b|} + \sum_{a\in A}\frac{Z_a}{|\mathbf{r}_a - \mathbf{r}_b|}\right\} \quad (2)$$

with the atomic charges Q_{00}^{b} of molecule B computed from its molecular electrostatic potential with the CHELPG scheme.[50] In eq 2, a and b refer to nuclei and i to electrons. Similarly, the charge density of molecule B will be perturbed due to molecule A, and hence eqs 1 and 2 need to be solved iteratively to self-consistency for both molecules.

We generated structurally uncorrelated, spherical water clusters by selecting 1996 configurations from a 298 K, 1 atm molecular dynamics run of an equilibrated 542-water molecule system with the TIP4P/2005[6] model. For each configuration, one water molecule (central) was arbitrarily chosen. All water molecules separated from the central molecule by an oxygen−oxygen distance of up to 12 Å were retained. These clusters of 225−250 water molecules were used to compute an average, distributed multipole moment model of water in the liquid state without structural relaxation of the cluster geometry.

For each cluster, eq 1 was solved for the central molecule. The central molecule was exposed to the field generated by the atomic charges of the surrounding molecules in the cluster. In the first iteration, the surrounding molecules were modeled with the CHELPG[50] ESP charges of isolated water. In the second iteration, the computed CHELPG ESP atomic charges of the central molecule were placed at the atomic position of the surrounding molecules, and eq 1 was solved repeatedly to self-convergence. Typically, this required fewer than six iterations for atomic charges to converge to 0.001e. The converged charge density of the central molecule was used in distributed multipole analysis[36] to compute water's atomic multipole moments up to hexadecapole. This method is limited in accuracy, because it does not include charge transfer effects, and additionally, the field that the central molecule is experiencing in the cluster is of limited accuracy due to its monopole representation. Nevertheless, it has been found to give comparable lattice energies to those obtained

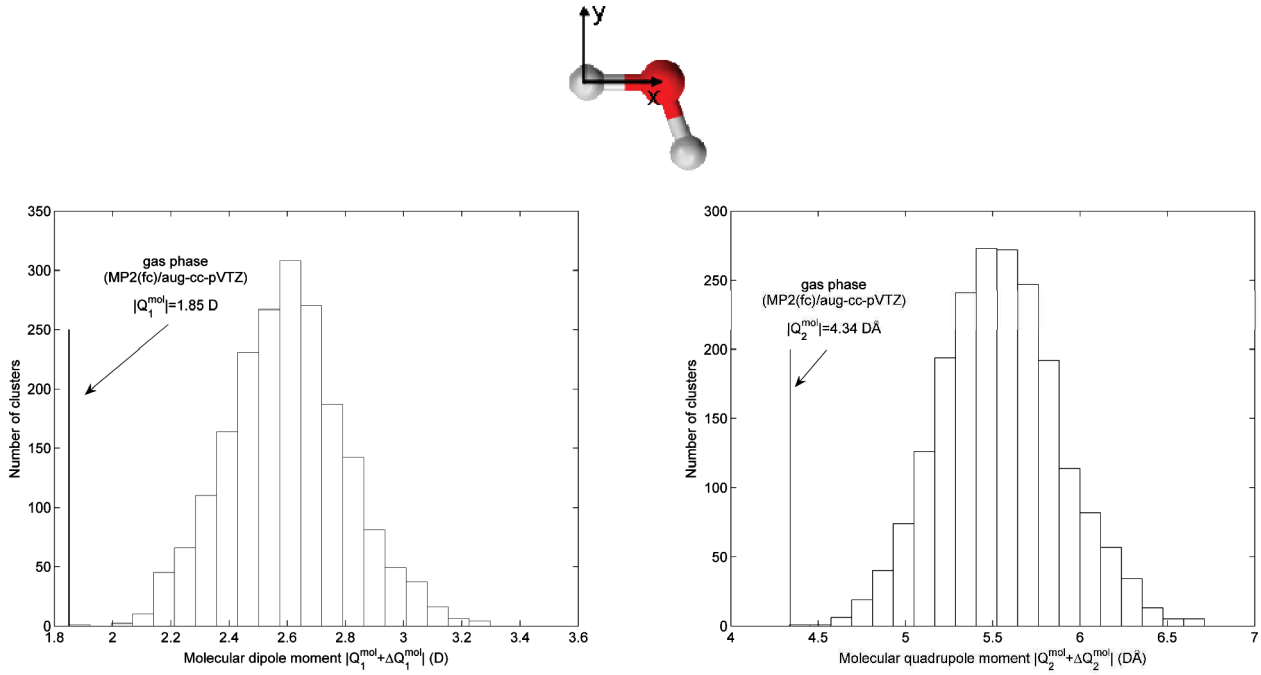


Figure 1. Distribution of magnitudes (defined as $|Q_l^{\mathrm{mol}} + \Delta Q_l^{\mathrm{mol}}| = \sqrt{[\sum_m (Q_{lm}^{\mathrm{mol}} + \Delta Q_{lm}^{\mathrm{mol}})^2)]}$ of converged molecular dipole (in D) and quadrupole moments (in DÅ) of the central water molecule in water clusters taken from a TIP4P/2005[6] liquid water simulation at ambient conditions. Molecular multipole moments Q_{lm}^{mol} were computed[54] from the distributed multipoles and refer to the molecular axes system shown (see also Supporting Information).

with an elaborate distributed-polarizability model for organic crystal structures exhibiting hydrogen bonding.[37] Once all clusters had been processed, the atomic multipole moments of the central water molecule were averaged (Figure 1). These average multipole moments can be thought of as being the sum of the static (isolated water molecule) Q_{lm} and average induced multipole moments ΔQ_{lm}, where indices l and m refer to the component of the multipole moments 00, 10, 11s, ..., 44s. The average induced moments ΔQ_{lm} are readily obtained by subtracting the static multipole moments Q_{lm} computed from the distributed multipole analysis of the isolated-water charge density. The obtained average water dipole moment in the liquid is 2.59 D and is in reasonable agreement with the value of 2.7 D obtained for water clusters in the seminal work of Gregory et al.,[51] although the enhancement in the dipole moment of water in condensed phases is still a matter of contention.[52,53]

The induced multipole moments of a polarizable molecule A are determined by the competition between the lowering of the intermolecular energy due to the interaction of the induced moments with the field created by surrounding molecules and the energy cost (internal energy) to distort the molecule's charge density in zero field to the charge density in solution. If we assume a bilinear dependence of the internal energy on the induced moments, it can be shown[55] that the lowering of the intermolecular energy is twice the internal energy (in absolute terms). Hence, the overall lowering of the system's energy due to the polarization of molecule A (called induction energy), without damping, is

$$E_{\mathrm{ind}}(A) = (1/2) \sum_{\substack{B \\ B \neq A}} \sum_{a \in A} \sum_{b \in B} \sum_{bm,l'm'} \Delta Q_{lm}^a T_{lm,l'm'}^{ab} Q_{l'm'}^b$$

where $T_{lm,l'm'}^{ab}$ is the interaction tensor[55] that depends on the relative position and orientation of sites a and b and the factor 1/2 accounts for the internal energy cost. All molecular dynamics runs were performed using DL_MULTI,[56] which allows the modeling of electrostatic interactions with standard Ewald summation for multipoles up to rank 4 (hexadecapole), although explicit induction is not included. However, the average effect of induction can be accounted for by performing simulations using fixed, effective moments $Q_{\mathrm{eff},lm} = Q_{lm} + (1/2)\Delta Q_{lm}$, which provides approximately the electrostatic and average induction energy of the system, so that

$$\begin{aligned}
&\frac{1}{2} \sum_{\substack{A,B \\ B \neq A}} \sum_{a \in A} \sum_{b \in B} \sum_{lm,l'm'} \left(Q_{lm}^a + \frac{\Delta Q_{lm}^a}{2} \right) T_{lm,l'm'}^{ab} \left(Q_{l'm'}^b + \frac{\Delta Q_{l'm'}^b}{2} \right) \\
&= \frac{1}{2} \sum_{\substack{A,B \\ B \neq A}} \sum_{a \in A} \sum_{b \in B} \sum_{lm,l'm'} \left(Q_{lm}^a T_{lm,l'm'}^{ab} Q_{l'm'}^b + \frac{\Delta Q_{lm}^a}{2} T_{lm,l'm'}^{ab} Q_{l'm'}^b + \right. \\
&\qquad \left. Q_{lm}^a T_{lm,l'm'}^{ab} \frac{\Delta Q_{l'm'}^b}{2} + \frac{\Delta Q_{lm}^a}{2} T_{lm,l'm'}^{ab} \frac{\Delta Q_{l'm'}^b}{2} \right) \\
&= E_{\mathrm{elec}} + E_{\mathrm{ind}} + \Delta E_{\mathrm{error}}
\end{aligned} \tag{3}$$

assuming that no damping is used. The last term in eq 3 can in principle be computed at each molecular dynamics time step using $\Delta Q_{lm}/2$ on each atom. However, this would practically double the computational cost and is not worthwhile given that the contribution of $\Delta E_{\mathrm{error}}$ to the lattice energy of organic hydrogen-bonded crystals was found to be less than a few percentage points of the lattice energy and around 8% of the induction energy. Hence, this term was omitted in this work, and its effects were absorbed in the water repulsion−dispersion parametrization (see the section "Repulsion−Dispersion Interactions") using the ef-

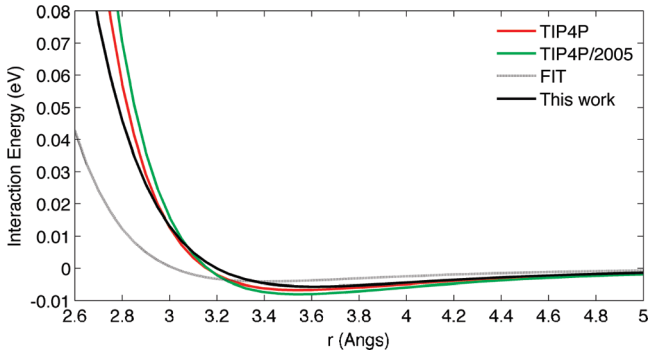


Figure 2. Fitted oxygen−oxygen exp-6 repulsion−dispersion interaction potential (black line) compared with TIP4P,[42] TIP4P/2005,[6] and FIT[30] parametrizations.

fective Q_{eff} multipole moments. In principle, the procedure outlined above can be repeated to compute new average induced multipole moments using the fitted repulsion−dispersion potential, and the whole scheme can be iterated to convergence. We did not, however, follow such an iterative scheme, since the TIP4P/2005 water model provides a sufficiently accurate reproduction of the liquid water structure.[6]

Repulsion−Dispersion Interactions. We considered the combination of the distributed-multipole description of water's electrostatic interactions obtained in this work with repulsion−dispersion parameters of well-established literature water models based on a monopole representation of the charge density,[6,42] but the resulting potential failed to reproduce liquid water properties in all cases. Hence, we carried out a parametrization of the repulsion−dispersion interactions of a three-site, exp-6 model for water using as a starting point the exp-6 repulsion−dispersion parameters of the oxygen and polar hydrogen atoms (hydrogen atoms connected to oxygen and nitrogen) of the FIT[30] empirical model, which was fitted to experimental structural data and sublimation energies of organic crystal structures in conjunction with distributed multipoles.

In order to keep the complexity of the parametrization manageable, the hydrogen parameters of the FIT potential were not altered, and the oxygen−hydrogen cross interaction parameters were computed using standard Lorentz−Berthelot combining rules. We fitted the oxygen parameters to reproduce the liquid water density and the oxygen−oxygen radial distribution function[57] at 298 K and 1 atm. Figure 2 shows a comparison of the obtained optimal oxygen−oxygen exp-6 potential with the starting FIT parametrization and the TIP4P[42] and TIP4P/2005[6] water models (see also the Supporting Information for a full set of repulsion−dispersion parameters and multipole moments for the proposed water model). We note that the water model obtained in this work appears less repulsive than the two potentials of the TIP4P family, but this is not the case, given that our potential also includes oxygen−hydrogen interactions that are strongly repulsive at typical hydrogen bonding distances.

In Figure 3, it can be seen that the proposed water model reproduces well the experimentally determined oxygen−oxygen radial distribution function at 298 K and 1 atm, given the experimental error[57] and the neglect of quantum effects[58,59] and molecular flexibility[60] in our classical calculations. The number of water molecules in the first hydration shell, obtained by integrating the oxygen−oxygen radial distribution function up to the distance of its first minimum, is 4.8 and also in good agreement with experimental results (Figure 3b). Moreover, the model reproduces well the second peak at 4.5 Å; this is indicative of the accuracy in modeling the hydrogen bond network in liquid water. It is encouraging that the water model also reproduces the oxygen−hydrogen and hydrogen−hydrogen radial distribution functions that were not included in the parametrization. The predicted liquid water density at 298 K and 1 atm is 0.994 g cm^{-3} compared to the experimental value[61] of 0.997 g cm^{-3}. The model was further tested by modeling several additional solid and liquid water properties as detailed in the section Testing of the Model for the Intermolecular Forces.

Solute−Water Interactions. For the solutes, we followed a simpler approach to derive the $Q_{lm} + \Delta Q_{lm}$ multipole moments, by computing the charge density for distributed multipole analysis using a polarizable continuum model[62] with the United Atom Topological Model (UA0) for the water cavity and the default parameters for water as a solvent in Gaussian 03.[43] As in the case of water, the solute's electrostatic intermolecular interactions were modeled using the effective multipole moments $Q_{eff,lm} = Q_{lm} + \Delta Q_{lm}/2$. Hydration free energy calculations were also performed using effective ESP atomic charges $Q_{eff,00} = Q_{00} + \Delta Q_{00}/2$ computed from the same dielectric continuum and gas-phase wave functions with the CHELPG scheme.[50]

The solute−water repulsion−dispersion interactions were obtained by applying standard Lorentz−Berthelot combining rules[63] to the original FIT parametrization[30] for the solute and the water model developed in this work (see the Supporting Information). In the case of the free energy perturbation calculations (see Repulsion−Dispersion Interactions), the exp-6 solute−water repulsion−dispersion potential was supplemented with a γ/r^{12} term to alleviate the divergence of the exp-6 potential to minus infinity at short distances. The coefficients γ were chosen for each interaction independently so as to make the added term negligible for all solute−water intermolecular distances sampled in a simulation with the unperturbed solute−water interaction potential.

Testing of the Model for the Intermolecular Forces. The developed water potential was first tested by modeling the five proton-ordered ice polymorphs XI,[64] II,[65] IX,[66] VIII,[67] and XIII.[68] Ice XIII is a hydrogen-ordered phase of disordered ice V prepared under high pressure but determined at ambient pressure by powder neutron diffraction.[68] In all ordered ice phases, the coordination of water molecules resembles the liquid phase with each water molecule tetrahedrally hydrogen bonded to four neighbors. The ice polymorphs were lattice energy minimized with respect to the relative position and orientation of the water molecules and the cell geometry using DMACRYS.[69] The minimizations were performed within the resulting space group symmetry constraints after lowering the symmetry so the asymmetric unit comprised complete molecules. Charge−charge, charge−dipole, and dipole−dipole interactions were calculated with Ewald summation, while repulsion−dispersion

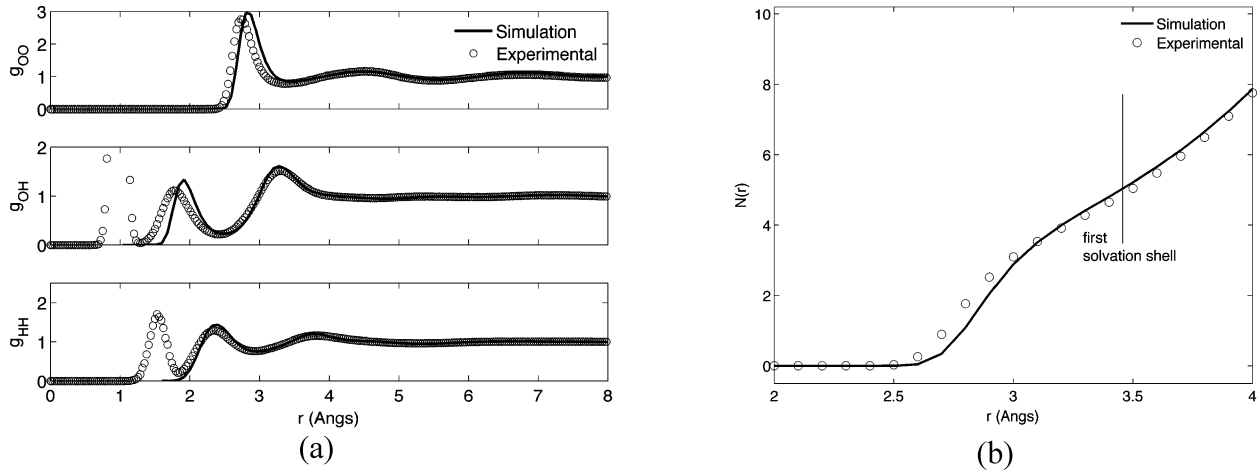


Figure 3. (a) Simulated site−site distribution functions for liquid water at 298 K and 1 atm compared with experimental data.[57] Simulation results for OH and HH do not include contributions from bonded atoms. (b) Average number of water oxygen atoms $N(r)$ within distance r from any given water oxygen atom under the same conditions.

and higher multipole contributions were evaluated in direct space up to a 15 Å cutoff. All ice lattice energy minimizations were performed at 0 Pa pressure, apart from ice VIII, which was modeled at the determination pressure of 2.4 GPa. The neglected thermal expansion changes the volume of organic crystals by approximately one percentage point per 100 K, which is within the error margin of the proposed water model. The reproduction of the five ordered ice polymorphs was contrasted with three different intermolecular potential models: TIP4P,[42] TIP4P/2005,[6] and the original FIT parametrization[30] combined with gas-phase MP2(fc)/aug-cc-pVTZ water multipole moments. For the two water models of the TIP4P family, we included in the lattice energy calculations the oxygen partial charge at the nonnuclear position as described in the original force field specifications. The quality of the modeling of the ice structures was evaluated by computing[70] the root-mean-square discrepancy in overlaying the oxygen positions of a 20-water-molecule cluster ($RMS_{cs\text{-}20}$).

The Cambridge Structural Database[71] contains a trihydrate crystal structure of pyridine that was determined at 223 K, that is, 20 K below its decomposition temperature.[72,73] The structural reproduction of this crystal structure was used as an additional test of the quality of the potential model, and in particular of the solute−water cross interactions, by examining its thermal stability in an isothermal, isobaric ensemble with fully flexible cell (hereafter referred to as $N\sigma T$) molecular dynamics simulations at its determination conditions. The solute−water interaction potential was further tested by assessing its ability to model hydration free energies as detailed in the section Free Energy of Hydration.

The liquid water density at 1 atm was computed in the temperature range 253−323 K in a series of isothermal, isobaric molecular dynamics simulations. The dependence of the simulated water density with temperature was fitted to a fourth-order polynomial that was subsequently used to compute the thermal expansion coefficient $\alpha_P = 1/V(\partial V/\partial T)_P = -1/\rho(\partial\rho/\partial T)_P$ at 298.15 K. The result was checked for consistency with the value obtained from the volume-enthalpy fluctuations[74] $\alpha_P = (\langle VH\rangle_{NPT} - \langle V\rangle_{NPT}\langle H\rangle_{NPT})/(kT^2\langle V\rangle_{NPT})$ in an isothermal, isobaric simulation at 298.15 K and 1 atm, where $\langle x\rangle$ denotes ensemble averages.

The experimental molar volume of water at 298.15 K depends linearly on pressure in the range 1−200 atm. Hence, the isothermal compressibility $k_T = -1/V(\partial V/\partial P)_T = 1/\rho(\partial\rho/\partial P)_T$ was obtained with a series of five isothermal isobaric simulations at 1, 10, 50, 100, and 200 atm at 298.15 K by computing the slope of a linear model fitted to the pressure dependence of the simulated molar volume. To check consistency, k_T was also computed from the volume fluctuations $K_T = (\langle V^2\rangle_{NPT} - \langle V\rangle_{NPT}{}^2)/(kT\langle V\rangle_{NPT})$ at 298.15 K and 1 atm.

The enthalpy of vaporization of liquid water was computed as the enthalpy difference $\Delta H_{vap} = H_{vap} - H_{liq} = U_{vap} - U_{liq} + P(V_{vap} - V_{liq})$ where U is the configurational energy, ignoring the kinetic energy, which will be equal in the two phases at the coexistence (vapor) pressure and temperature. If we assume that the vapor behaves ideally, this equation simplifies to

$$\Delta H_{vap} \approx -\langle U_{liq}\rangle_{NPT} + RT \tag{4}$$

which was used to compute ΔH_{vap} at 298.15 K from an isothermal, isobaric simulation of liquid water at 1 atm. This pressure is higher than the vapor pressure at 298.15 K, but the effect on ΔH_{vap} can be ignored since the liquid phase enthalpy change over this pressure range can be considered negligible. It has been proposed that ΔH_{vap} calculated with effective pair potentials should also be corrected for self-polarization,[22,75] so that the energy cost to distort the molecular charge density to its polarized state is taken into account. This positive correction can be approximated[75] from the difference in the dipole moment of water in the liquid and gas phases and the isotropic scalar polarizability from the relation $E_{cor} = 1/2(\mu_{liq} - \mu_{gas})/a$. However, this correction is not appropriate for our model, because the intermolecular electrostatic interactions are modeled using the effective multipole moments Q_{eff} that already account for the cost of distorting the charge density of water from its gas state to its average polarized state in liquid water.

The experimental liquid water heat capacity under constant pressure $C_P = (\partial H/\partial T)_p$ varies[61] in the temperature range 273−323 K by less than 1%; this is reflected in our model by an almost linear dependence of liquid water's enthalpy on temperature. Hence, C_P was computed from the slope of the liquid enthalpy with respect to temperature in six isothermal, isobaric simulations in the aforementioned temperate range. The liquid water enthalpy was computed as the sum of the average configurational energy, rotational and translational kinetic energy (difference from $3RT$ was within standard deviation), and the PV_{liq} term. To check for consistency, C_P was also computed from the enthalpy fluctuations $C_P = (\langle H^2\rangle_{NPT} - \langle H\rangle^2_{NPT})/(kT^2)$ in an isothermal, isobaric simulation at 298.15 K and 1 atm.

The heat capacity under constant volume $C_V = (\partial E/\partial T)_V$ at 298.15 K and 1 atm was computed from the almost perfectly linear dependence of the total energy E on temperature in a series of four constant volume, constant temperature simulations in the range 293.15−308.15 K in 5 K intervals. For these simulations, the density was constrained to the experimental density of water at 298.15 K. The heat capacity C_V was also computed from the liquid water's potential energy U fluctuations as

$$C_V = \frac{\langle E^2\rangle_{NVT} - \langle E\rangle^2_{NVT}}{kT^2} = \frac{\langle U^2\rangle_{NVT} - \langle U\rangle^2_{NVT}}{kT^2} + 3R \quad (5)$$

where the last term arises from the rotational and translational kinetic energy contribution of the rigid water molecules. We finally computed the self-diffusion coefficient using the Einstein relationship

$$D = \frac{1}{6(t - t_0)}\lim_{t\to\infty}\langle|\mathbf{r}(t) - \mathbf{r}(t_0)|^2\rangle$$

The water model was also tested by carrying out molecular simulations of direct liquid−vapor coexistence[76] to compute the saturated liquid and vapor densities as a function of temperature. A previously equilibrated cubic cell of 542 molecules was expanded 2.5 times in the z direction, and the additional volume was left empty. The system was allowed to evolve in the NVT ensemble for 300 ps following a 40 ps equilibration and for six temperatures spanning the range 300−550 K. The coexistence densities and interface thickness d were computed by fitting a hyperbolic tangent function[76] of the form

$$\rho(z) = \frac{1}{2}(\rho_{\text{liq}} + \rho_{\text{vap}}) - \frac{1}{2}(\rho_{\text{liq}} + \rho_{\text{vap}})\tanh[(z - z_0)/d] \quad (6)$$

where z_0 is Gibbs' dividing surface. For all temperatures studied, the two interfaces were symmetric and the liquid region was sufficiently wide to provide a reliable measure of the saturated liquid density. Longer simulation times and a larger number of water molecules gave statistically identical coexistence densities. The surface tension was estimated from the components of the pressure tensor.[76,77]

Free Energy of Hydration. Several explicit-solvent methods to compute the free energy of hydration have been proposed.[16,21,78] In the so-called free energy perturbation approach, the free energy difference[19] between two states A and B of a system defined with energy functions U_A and U_B can be computed as

$$\Delta G_{A,B} = G_B - G_A = -\frac{1}{\beta}\ln\langle\exp[-\beta\Delta U_{A,B}]\rangle_A = \frac{1}{\beta}\ln\langle\exp[\beta\Delta U_{A,B}]\rangle_B \quad (7)$$

i.e., by accumulating the energy differences $\Delta U_{\text{A,B}} = U_B - U_A$ (work distributions) in a simulation with either Hamiltonian (forward or reverse simulation). Exponential averaging in one direction does not necessarily give the minimum bias and variance of the free energy difference for a given set of $\Delta U_{A,B}$ measurements.[17] Bennet's acceptance ratio[79] is generally more efficient but requires sampling in both forward and reverse directions. As the purpose of our current work is not to identify the most efficient method to compute the free energy, but to test the effect of the accuracy of the intermolecular potential model in hydration calculations, all free energy estimates are based on forward exponential averaging (solute-creation only), apart from the free energy for the transfer of a single water molecule from the gas phase to the liquid state, which was independently computed with exponential averaging in both directions.

In hydration free energy calculations, state B in eq 7 corresponds to the solute interacting with the solvent, while in state A the solute−solvent interactions are fully annihilated. For practically all systems of interest, the overlap in phase space between the two states is sufficiently low that the free energy difference cannot be computed from data for only the end states A and B. Instead, the solute needs to be introduced into solution in stages by defining a series of intermediate states $i = 1, ..., N$ such that

$$\Delta G_{A,B} = \sum_{i=1}^{N-1}\Delta G_{i,i+1} = -\frac{1}{\beta}\sum_{i=1}^{N-1}\ln\langle\exp[-\beta\Delta U_{i,i+1}]\rangle_i \quad (8)$$

where states 1 and N are the end states A and B. In this work, the solute is introduced to a previously equilibrated water system, by first switching on the solute−solvent repulsion−dispersion interactions in a series of 29 simulations $U_{\text{solute-water}} = \lambda U^{\text{vdW}}_{\text{solute-water}}$, for λ = (0.0, 1e−9, 1e−8, 1e−7, 1e−6, 1e−5, 1e−4, 3e−4, 7e−4, 1e−3, 2e−3, 4e−3, 7e−3, 0.01, 0.015, 0.025, 0.04, 0.05, 0.07, 0.1, 0.15, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9). The uneven spacing in λ reflects the much larger curvature $dG/d\lambda = \langle dU/d\lambda\rangle_\lambda$ for small λ. We did not observe numerical instabilities at low λ values, where the λ-scaled potential increases abruptly at short separations. However, designing a suitable[20,21] λ-dependent (soft-core) functional form for exp-6 potentials would be beneficial in alleviating this source of potential instability and also reducing the number of intermediate states required. Preliminary simulations using a soft-core solute−water repulsion−dispersion interaction[21] gave free energies of hydration that were within standard deviation from our predictions, as discussed in the Supporting Information.

Once the solute's van der Waals interactions were fully switched on, the solute's effective multipole moments Q_{eff} were gradually switched on in a series of 10 simulations with

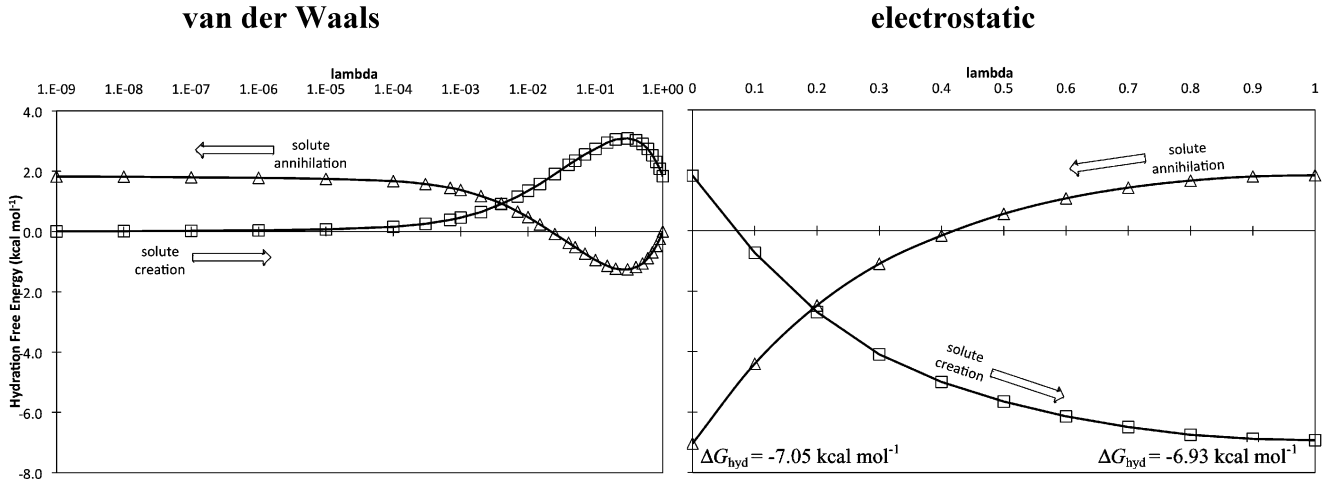


Figure 4. Repulsion−dispersion (left) and electrostatic (right) contributions to the free energy of hydration of water at ambient conditions as a function of the coupling parameter λ for solute-creation and solute-annihilation simulations.

the solute−water interaction potential being $U_{\text{solute-water}} = U^{\text{vdW}}_{\text{solute-water}} + \lambda U^{\text{elec+ind}}_{\text{solute-water}}$, for λ = (0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9). By comparing with results using larger perturbation steps, we conclude that the hydration free energies obtained with this set of intermediate states are converged to accuracy that is within the statistical uncertainty due to finite sampling of the work distributions. The isothermal, isobaric simulations for all intermediate states were performed in parallel starting from the same equilibrated configuration with the unperturbed solute−water $U_{\text{solute-water}}$ interaction potential. The infinitesimal correction in the predicted hydration free energy due to the reduction in system volume when the solute is fully decoupled was ignored.[16]

The bias in exponential averaging has opposite signs for the forward and reverse simulations and, generally, depends on the size of the perturbation step. In solute-creation simulations, large perturbation steps lead to less negative ΔG_{hyd} caused by the overlap of the solute and water. In annihilation simulations, large perturbation steps lead to more negative ΔG_{hyd} due to the loss of van der Waals and electrostatic attraction between the solute and water. Figure 4 shows that, for the free energy of transfer of a water molecule from a gas to a liquid at ambient conditions, the result from forward and reverse exponential averaging differs by only 0.1 kcal mol^{-1}, which is less than the 0.3−0.4 kcal mol^{-1} statistical uncertainty in determining ΔG_{hyd} for our method (c.f. the section Modeling of Hydration) and confirms that the $\Delta\lambda$ increments are sufficiently small.

For comparison purposes, the hydration free energies were also computed using the polarizable continuum model[62] (PCM) at the B3LYP/aug-cc-pVTZ level of theory. The solvent cavity was built using the United Atom Topological Model with the recommended atomic radii for solvation free energy calculations (UAHF), and the default parameters for water solvent as in Gaussian 03.[43] The hydration free energies with the polarizable conductor calculation model[80] (CPCM) differed by less than 0.1 kcal mol^{-1} and are not reported. We note that the use of the UFF force field atomic radii (UA0) gave large errors, of up to 5 kcal mol^{-1}, in the calculated hydration free energies; this shows the pronounced sensitivity[81,82] to the method used to construct the cavity surface. The atomic radii had a much smaller effect in computing the solute's distributed multipole moments.

All reported experimental and computed free energies of hydration refer to ambient conditions and correspond to the use of molar concentration units for the solute both in the gas phase and in solution.[82,83]

Molecular Dynamics Simulations. Molecular dynamics simulations were performed with a 1.2 fs time step, the leapfrog integrator, and the Nose−Hoover thermostat and barostat[84] as implemented in DL_MULTI.[56] The Ewald precision was set to 5×10^{-7} for all multipole orders in reciprocal space and charges in direct space. The precision of all other multipole orders in direct space was 5×10^{-8}. With these settings, the integration of the equations of motion of liquid water at ambient conditions for 1 ns gives a drift in the total Hamiltonian per degree of freedom that is more than an order of magnitude smaller than kT.

All pure water simulations were performed with 542 water molecules. For hydration free energy calculations, a solute molecule was inserted in a previously equilibrated cell that contained 542 water molecules at ambient conditions, and 2−4 water molecules were removed, depending on the size of the solute, to alleviate overlaps and short contacts. Work distributions were accumulated every 10 time steps over a 250 ps isothermal, isobaric simulation following 50 ps of equilibration. This scheme is sufficient to obtain converged free energy differences to within a few tenths of a kilocalorie per mole; further sampling of the work distributions was limited by the availability of computing resources. The simulations carried out to obtain water properties varied in length from 1 ns to several nanoseconds, depending on the property under consideration; the longest runs performed were those required to compute the density of supercooled liquid water. A 542-water-molecule, 100 ps simulation required approximately 8 h on an eight-core Intel Xeon E5462 2.8 GHz node, resulting in the equivalent of one CPU year to compute the hydration free energy per solute. Roughly, the use of a distributed multipole model up to hexadecapole is an order of magnitude more expensive compared with simpler monopole charge models.

Table 1. Lattice Energy Minimization of Five Ordered Ice Polymorphs[a]

model	% error lattice lengths a	b	c	density (g cm^{-3})	% error density	lattice energy (kcal mol^{-1})	RMS$_{cs\text{-}20}$[b] (Å)
ice XI,[64] *Cmc*2$_1$, *Z′* = 2, 5 K, ambient pressure, experimental density 0.930 g cm^{-3}							
TIP4P	−3.33	−0.35	−0.98	0.975	+4.84	−13.62	0.12
TIP4P/2005	−2.53	+0.33	−0.20	0.953	+2.46	−15.05	0.11
FIT+vacuo DMA[c]	+0.33	−11.15	−4.01	1.087	+16.87	−14.96	0.29
this work	**+2.74**	**−2.58**	**+0.95**	**0.921**	**−1.02**	**−15.79**	**0.12**
ice II,[65] *R*$\bar{3}$, *Z′* = 2, 110 K, ambient pressure, experimental density 1.180 g cm^{-3}							
TIP4P	−1.98	−1.98	−2.18	1.255	+6.40	−13.38	0.10
TIP4P/2005	−1.24	−1.24	−1.58	1.229	+4.18	−14.83	0.09
FIT+vacuo DMA[c]	−5.24	−5.24	−0.91	1.326	+12.38	−13.46	0.22
this work	**−0.14**	**−0.14**	**+2.16**	**1.158**	**−1.83**	**−14.91**	**0.10**
ice IX,[66] *P*4$_1$2$_1$2, *Z′* = 1.5, 110 K, ambient pressure, experimental density 1.160 g cm^{-3}							
TIP4P	−1.93	−1.93	−1.46	1.224	+5.52	−13.52	0.07
TIP4P/2005	−1.25	−1.25	−1.01	1.202	+3.60	−14.97	0.05
FIT+vacuo DMA[c]	−6.67	−6.67	+6.32	1.253	+7.97	−13.80	0.27
this work	**−0.30**	**−0.30**	**+3.82**	**1.125**	**−3.09**	**−14.99**	**0.10**
ice VIII,[67] *I*4$_1$/*amd*, Z′ = 0.5, 10 K, 2.4 GPa, experimental density 1.629 g cm^{-3}							
TIP4P	−2.62	−2.62	+2.33	1.679	+3.03	−5.33	0.15
TIP4P/2005	−1.76	−1.76	+3.08	1.637	+0.50	−6.54	0.18
FIT+vacuo DMA[c]	−3.09	−3.09	−12.46	1.982	+21.62	−9.07	0.27
this work	**+0.04**	**+0.04**	**−3.30**	**1.684**	**+3.34**	**−8.99**	**0.10**
ice XIII,[68] *P*2$_1$/*a*, Z′ = 7, 80 K, ambient pressure, experimental density 1.251 g cm^{-3}							
TIP4P	+0.80	−2.83	−2.22	1.309	+4.63	−13.20	0.16
TIP4P/2005	+1.60	−2.18	−1.52	1.282	+2.48	−14.64	0.15
FIT+vacuo DMA[c]	−3.85	−2.72	−3.65	1.441	+15.14	−13.59	0.26
this work	**+0.36**	**+0.94**	**+1.27**	**1.241**	**−0.80**	**−14.86**	**0.12**

[a] All minimizations were performed at 0 K and 0 Pa with water held rigid to the TIP4P conformation, apart from ice VIII, which was minimized at the experimental pressure of 2.4 GPa. [b] RMS overlay[70] of the oxygen positions of a 20-molecule water cluster. [c] Repulsion−dispersion interactions computed using FIT[30] repulsion−dispersion potential combined with multipole moments derived from the DMA of the isolated-water MP2(fc)/aug-cc-pVTZ charge density.

The hydration free energy depends on the exact protocol with which the solute creation is effected, which includes the treatment of long-range repulsion−dispersion interactions.[85] In this work, repulsion−dispersion interactions were computed up to a 10 Å cutoff, which corresponds to three times the oxygen−oxygen distance in liquid water's first hydration shell (Figure 3b). Long-range correction[84] to energy and pressure was only applied in pure water calculations. In the hydration free energy calculations, water−water and solute−water repulsion−dispersion interactions were smoothly switched off between 9 and 10 Å using a cubic spline. This alleviates the need to apply different long-range corrections to the two configurational energies when computing the work distributions, which we estimated to change the predicted free energies of hydration by less than the statistical uncertainty (Supporting Information) due to finite sampling of the work distributions.

Results

Accuracy of Water Model. *Modeling of Ordered Ice Polymorphs and Pyridine Hydrate.* Table 1 contrasts the lattice energies and lattice parameters for five proton-ordered ice polymorphs with the proposed water model and with other popular potentials for water. It can be seen that the proposed water model achieves the smallest overall RMS error in the reproduction of the 20-molecule coordination sphere of the five ordered ice polymorphs considered, which was in all cases smaller than 0.12 Å. The reproduction of hydrogen bonding geometries is also satisfactory. However, the differences in reproduction with the proposed water model compared to TIP4P[42] and TIP4P/2005[6] are small, and comparable to the effect of neglected thermal expansion and molecular distortions. Using the FIT[30] repulsion−dispersion potential with the isolated-water MP2(fc)/aug-cc-PVTZ multipole moments largely overestimates the ices' densities. This contrasts the successful modeling of four hydrogen-ordered ice polymorphs with FIT and multipole moments computed from the isolated-water MP2/6-31G(d,p) charge density by Hulme and Price.[86] The agreement in this case can be attributed to cancellation of errors due to the limited size of the 6-31G(d,p) basis set in computing the water's charge density. It is encouraging that the low-temperature ice XI corresponds to the most stable form with all the models in Table 1, and that the energy differences between all polymorphs modeled at 0 Pa lie in a narrow energy range of a couple of kilocalories per mole.

The isothermal bulk modulus of ice II at 0 K and 0.35 GPa with our water potential is approximately 20 GPa, compared with 23 GPa for TIP4P/2005 and 14 at 0.35 GPa and 225 K, experimentally.[87]

The pyridine trihydrate crystal structure is found to be stable in a *NσT* molecular dynamics simulation at the experimental determination conditions of 223 K and 1 atm. Pyridine molecules remain enclosed between water layers with limited $\pi-\pi$ stacking despite the almost parallel arrangement of their molecular planes, in good agreement with the experimental crystal structure that is reflected in cell length errors of less than 2.5%. The first peak in the

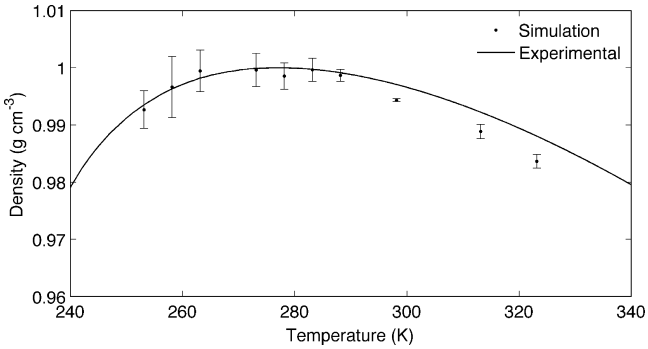


Figure 5. Simulated vs experimental[61,91] liquid water densities at 1 atm.

site−site correlation function between pyridine's nitrogen and water's oxygen atoms at these conditions is at 2.65 Å compared with the 2.79 Å N···O hydrogen bond length in the experimentally determined crystal. This underestimation of hydrogen bond lengths is due to the use of implicit induction in conjunction with a repulsion−dispersion potential for pyridine that was parametrized using isolated-molecule multipole moments.

Modeling of Liquid Water and Water Vapor−Liquid Equilibria. In Figure 5, it can be seen that the proposed water model successfully predicts the density variation along the ambient pressure isobar including the temperature of maximum density. The thermal expansion coefficient computed from the volume-enthalpy fluctuations changes sign at approximately 279 K (6 °C), which is in good agreement with the 273 K temperature of maximum density obtained from a fourth-order polynomial fitted to density−temperature data and the experimental temperature of maximum density of 277 K. For comparison, from the TIP family of water models, TIP4P/2005,[6] TIP4P/Ew,[44] and TIP5P[7] (without Ewald summation) predict[88] this water density anomaly. It should be noted, however, that these models were fitted to do so, while the parametrization of our model only used the liquid water density at 298.15 K. From the slope of the fourth-order polynomial, we compute a thermal expansion coefficient α_P of 3.4×10^{-4} K^{-1} at 298.15 K compared with 3.1×10^{-4} K^{-1} calculated from the volume-enthalpy fluctuations and 2.6×10^{-4} K^{-1} obtained experimentally.[89] Despite predicting the temperature of maximum density, TIP5P overestimates[90] α_T by more than 100%, while TIP4P/2005[6] is only slightly better than our model with α_T equal to 2.8×10^{-4} K^{-1}. The average simulated isothermal compressibility in the pressure range 1−200 atm computed by fitting a linear model for the dependence of the simulated molar volume on pressure is 5.4×10^{-5} atm^{-1} compared with 5.6×10^{-5} atm^{-1} from volume fluctuations. The experimental isothermal compressibility in the same pressure range is 4.5×10^{-5} atm^{-1}, while the values for TIP5P[90] and TIP4P/2005[6] are 4.1×10^{-5} and 4.6×10^{-5} atm^{-1}, respectively.

The calculated enthalpy of vaporization at 298.15 K and 1 atm overestimates the experimental value by approximately 1.6 kcal mol^{-1}. For comparison, TIP4P/2005 and TIP4P/Ew overestimate[6] the experimental value by 1.5 and 1.2 kcal mol^{-1}, although this discrepancy is mainly because these models were parametrized by fitting to the experimental vaporization enthalpy after it was corrected for the polarization energy. In all these calculations, ΔH_{vap} is computed from eq 4 by treating the system classically. Strictly, the energy of a quantum-mechanical oscillator depends on the frequency, and hence there is a contribution to ΔH_{vap} due to the quantum-mechanical character of the intermolecular modes of liquid water and also due to the shifting of the intramolecular frequencies when a water molecule goes from the gas to the liquid phase. The vibrational corrections to ΔH_{vap} have been estimated[44] to be approximately −0.07 kcal mol^{-1} at 298 K. Moreover, the ΔH_{vap} increases by approximately 0.5 kcal mol^{-1} on going from H_2O to T_2O,[92] and the configurational energy of liquid water calculated with path-integral simulations is on the order of 1 kcal mol^{-1} less negative compared with classical water at ambient conditions.[59,93] All this suggests that inclusion of quantum effects would further reduce the deviation of the predicted ΔH_{vap} from experimental values.

The simulated enthalpy of liquid water in the temperature range 273.15−323.15 K is a linear function of temperature to a good approximation. From the slope of the fitted line, we compute that C_P is 101.4 J mol^{-1} K^{-1} compared to 75.3 J mol^{-1} K^{-1} experimentally and 100.8 J mol^{-1} K^{-1} obtained from the enthalpy fluctuations. For comparison, TIP4P/2005 and TIP4P/Ew predict C_P to be 88.3 and 89.5 J mol^{-1} K^{-1}, respectively, although once more these values do not include the polarization energy that, nevertheless, was included in their parametrization. Although the vibrational correction[44] to the enthalpy of vaporization is small, its temperature dependence is significant and results in a −9.4 J mol^{-1} K^{-1} correction to C_P at 298.15 K that significantly reduces the discrepancy of all three rigid-water models from experimental results. Quantization of classical models is also known to reduce the value of the heat capacity.[59] From the linear dependence of the simulated total energy with temperature in the range 293.15−308.15 K, we compute that the heat capacity under constant volume without any corrections is 101.2 J mol^{-1} K^{-1} compared to 74.5 J mol^{-1} K^{-1} obtained experimentally. The value obtained from the configurational energy fluctuations in a canonical ensemble simulation at 298.15 K and 1 atm is in excellent agreement and equal to 101.1 J mol^{-1} K^{-1}. The predicted self-diffusion coefficient is 1.4×10^{-9} m^2 s^{-1} and hence severely underestimates the experimental value (see Table 2) and the predictions of the two best performing TIP water models, TIP4P/2005 and TIP4P/Ew, that predict D to be 2.1 and 2.4×10^{-9} m^2 s^{-1}, respectively.[44,88] Nevertheless, the agreement of these models may be fortuitous given that quantum effects have been shown to increase[58] D by 50%.

A summary of the calculated properties of liquid water at 298.15 K and 1 atm with the proposed model is presented in Table 2. The computed liquid water properties are self-consistent, in that the difference in heat capacities obtained from the molar volume, thermal expansion coefficient, and isothermal compressibility satisfy $C_P - C_V = T\hat{V}\alpha_P^2/k_T$.

In Figure 6, the simulated water vapor−liquid coexistence envelope in the temperature range 298−550 K is presented compared to experimental data. The vapor density is

Table 2. Simulated and Experimental Properties of Liquid Water at 298.15 K and 1 atm[a]

property	simulated	experimental[b]
density (g cm^{-3})	0.994	0.997
$10^4\alpha_P$ (K^{-1})	3.4[c]	2.6
$10^5\kappa_T$ (atm^{-1})	5.4[d]	4.5
C_V (J mol^{-1} K^{-1})	101.2[e,f]	74.5
C_P (J mol^{-1} K^{-1})	101.4[e,f]	75.3
ΔH_{vap} (kcal mol^{-1})	12.1[f]	10.5
10^9D (m^2 s^{-1})	1.4	2.3

[a] α_P is the thermal expansion coefficient; κ_T, the isothermal compressibility; C_v and C_P, the heat capacities at constant volume and pressure, respectively; ΔH_{vap}, the enthalpy of vaporization; and D, the self-diffusion coefficient. [b] Experimental values from refs 61, 89, 91, 94, and 95. [c] Computed from the slope of a fourth-order polynomial fitted to simulated density vs temperature data. [d] Computed from the slope of a linear model fitted to simulated molar volume vs pressure data. [e] Computed assuming a linear dependence of enthalpy and total energy on the temperature in NPT and NVT simulations, respectively. [f] Computed without any dipole moment, vibrational, and quantum corrections.

reproduced excellently, while the saturated-liquid density is moderately underestimated at high temperatures. This is consistent with the slightly overestimated thermal expansion coefficient, given that the liquid density at coexistence is almost equal to the liquid density at ambient pressure at a given temperature. From the normal and tangential components of the pressure tensor, we found that the surface tension of our water model weakly underestimates the experimental values by 4−6 mJ m^{-2} in the temperature range 350−550 K and shows the experimental decreasing trend with increasing temperature. We note that the statistical uncertainty of the surface tension in our 300 ps simulations is on the same order of magnitude as this discrepancy (see the Supporting Information). The direct coexistence method is not sufficiently accurate at higher temperatures to determine the critical temperature T_c. However, from the temperature dependence[78] of the surface tension, we estimated the critical temperature to be 637 K, which compares favorably with the experimental value of 647 K. We note that our early attempts to model water using gas-phase multipole moments led to a severe underestimation of the critical temperature despite reproducing the liquid density and oxygen−oxygen radial distribution function at ambient conditions.

When the solute is the same as the solvent, the free energy of solvation (or free energy of vaporization) can be computed from[96] $\Delta G_{vap} = \Delta G_{solv} = -kT \ln (\hat{V}_{vap}/\hat{V}_{liq})$, where $\hat{V}$ is the molar volume of the two phases in equilibrium. Hence, the experimental free energy of hydration for water at 298.15 K and 1 atm is estimated to be −6.3 kcal mol^{-1}, in good agreement with the −6.9 kcal mol^{-1} value obtained from the free energy perturbation calculations (see Figure 4). The entropy of liquid water computed from $S_{liq} = -(\Delta H_{vap} + \Delta G_{vap})/T + S_{vap}$ is approximately 3 cal mol^{-1} K^{-1} lower than the experimental at these conditions. This is consistent with the simulated site−site distribution functions (Figure 3) being somewhat more structured compared with the neutron experimental results, although any such qualitative comparison is also subject to errors due to the neglect of molecular flexibility and quantum effects in our simulations.

Modeling of Hydration. In Figure 7, the predicted hydration free energies for the 10 organic solutes are compared with the experimental values.[97] The root-mean-square (RMS) error using the distributed multipole model is 1.50 kcal mol^{-1}. The maximum error of 2.49 kcal mol^{-1} is obtained for imidazole, which is the molecule with the most negative free energy of hydration in our set. For comparison, the polarizable continuum model, which has been parametrized using experimentally determined free energies of hydration, gives a root-mean-square error of 1.53 kcal mol^{-1} with a maximum error of 3.24 kcal mol^{-1} for pyrene.

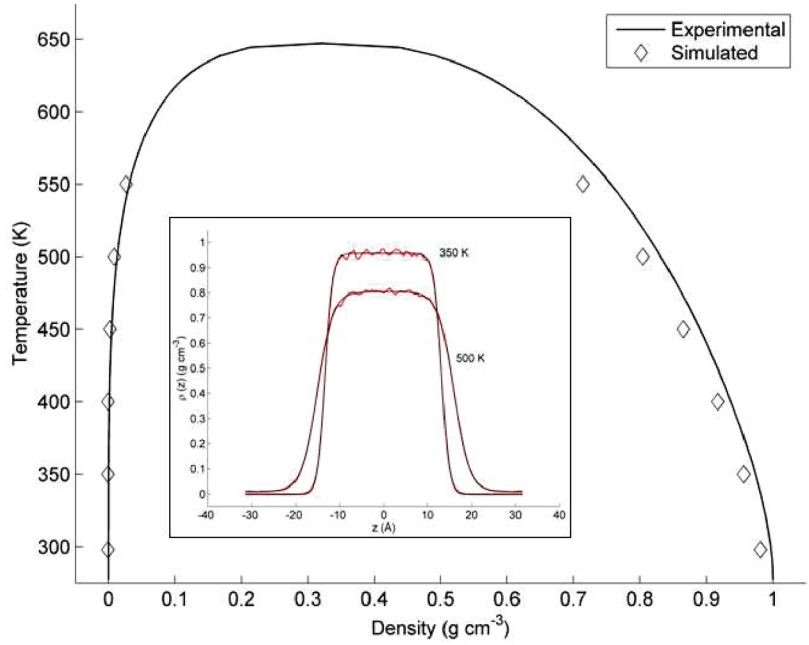


Figure 6. Water vapor−liquid equilibria. Experimental values from ref 61. The inset shows the z-density profile for 350 and 500 K; the molecular dynamics results are shown in red and the fitted tangent hyperbolic function in black (see also the Supporting Information).

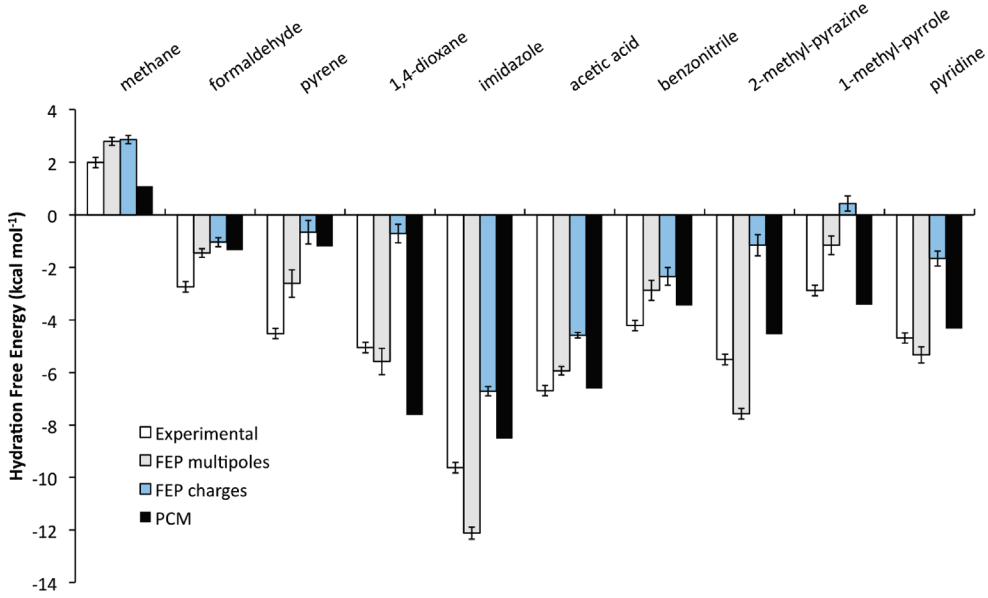


Figure 7. Predicted vs experimental[97] hydration free energies at ambient conditions using free energy perturbation (FEP) and a self-consistent reaction field method using the polarizable continuum model (PCM). Error bars in the FEP calculations correspond to one standard deviation computed by splitting the simulation into five equal parts; the error in experimental measurements was set to a nominal 0.2 kcal mol^{-1} value.[13]

Changing the charge density representation of the solute from a distributed multipole to an atomic charge model and keeping all other potential and simulation parameters identical leads to more positive free energies of hydration for all molecules. However, the effect is very molecule-dependent and varies from +0.07 kcal mol^{-1} for methane to +6.42 kcal mol^{-1} for 2-methyl-pyrazine. With the atomic charge model, the RMS error of the hydration free energies increases to 3.05 kcal mol^{-1}, while the free energy of hydration of 1-methyl-pyrrole is predicted to have the wrong sign.

The quality of the solute's charge density representation has a profound effect on the calculated free energies of hydration that should be reflected in qualitative differences in the hydrogen bonding between the water and the solute's hydrogen bond donors and acceptors. Figure 8a,b shows the site−site radial distribution functions for water oxygen with imidazole's nitrogen and 1,4-dioxane's oxygen atoms as obtained in a 1 ns isothermal, isobaric molecular dynamics run at ambient conditions with the unperturbed water−solute interactions. The use of distributed multipoles results in more pronounced first peaks and hence more directional, spatially confined hydrogen bonding. These differences in solute−water hydrogen bonding propagate to the second hydration shell before leveling off. The gray regions in Figure 8c,d show the areas of the first hydration shell of 1,4-dioxane and imidazole in which the number density of water oxygen atoms is equal to 3.3 times the average number density in liquid water at ambient conditions. It is clear that, when 1,4-dioxane's electrostatic interactions are modeled with multipole moments, hydrogen bonding is found more localized at the oxygen lone pairs, compared with the more scattered and less directional hydrogen bonding contacts with atomic charges. On the other hand, the isodensity surface of imidazole's first hydration shell appears to depend only weakly on the electrostatic model. However, for both solutes, the use of a distributed multipole model results in significantly wider ranges of number densities for the water oxygen atoms in the first hydration shell compared to the overall more isotropic distribution of surrounding water with atomic charges.

We have finally computed the probability distribution of water-exchange times in the first hydration shell and in the vicinity of the solute's hydrogen bond donors and acceptors. By fitting the obtained function with an exponentially decaying function $p(t) = A \exp(-t/\tau)$, we estimate the residence time of water molecules in the first hydration shell. When multipole moments are used, τ is found to be 1.7 ± 0.2 ps for 1,4-dioxane's oxygen acceptor and 1.4 ± 0.2 ps for imidazole's nitrogen acceptor. The corresponding residence times when atomic charges are used are 3−4 times shorter, which suggests that the dynamic properties of the solute's hydration also depend strongly on the approach with which the electrostatic interactions are modeled.

Discussion

The first part of this study is concerned with the development of a rigid-body water model that comprises a quantum-mechanically derived distributed multipole representation of the charge density, which includes average polarization effects in liquid water at ambient conditions. Parameterization was limited to the repulsion−dispersion potential, which is fitted to the experimental liquid water density and oxygen−oxygen radial distribution function at ambient conditions. Despite restricting the fitting to a very narrow set of experimental data and to only one temperature, the model is found to be successful in modeling a wide range of liquid water and ordered ice properties, including the notably elusive temperature of maximum density at 1 atm and the vapor−liquid equilibrium densities. This contrasts simpler water models that employ a monopole representation of the charge density,[7,44,88] the computational efficiency of which

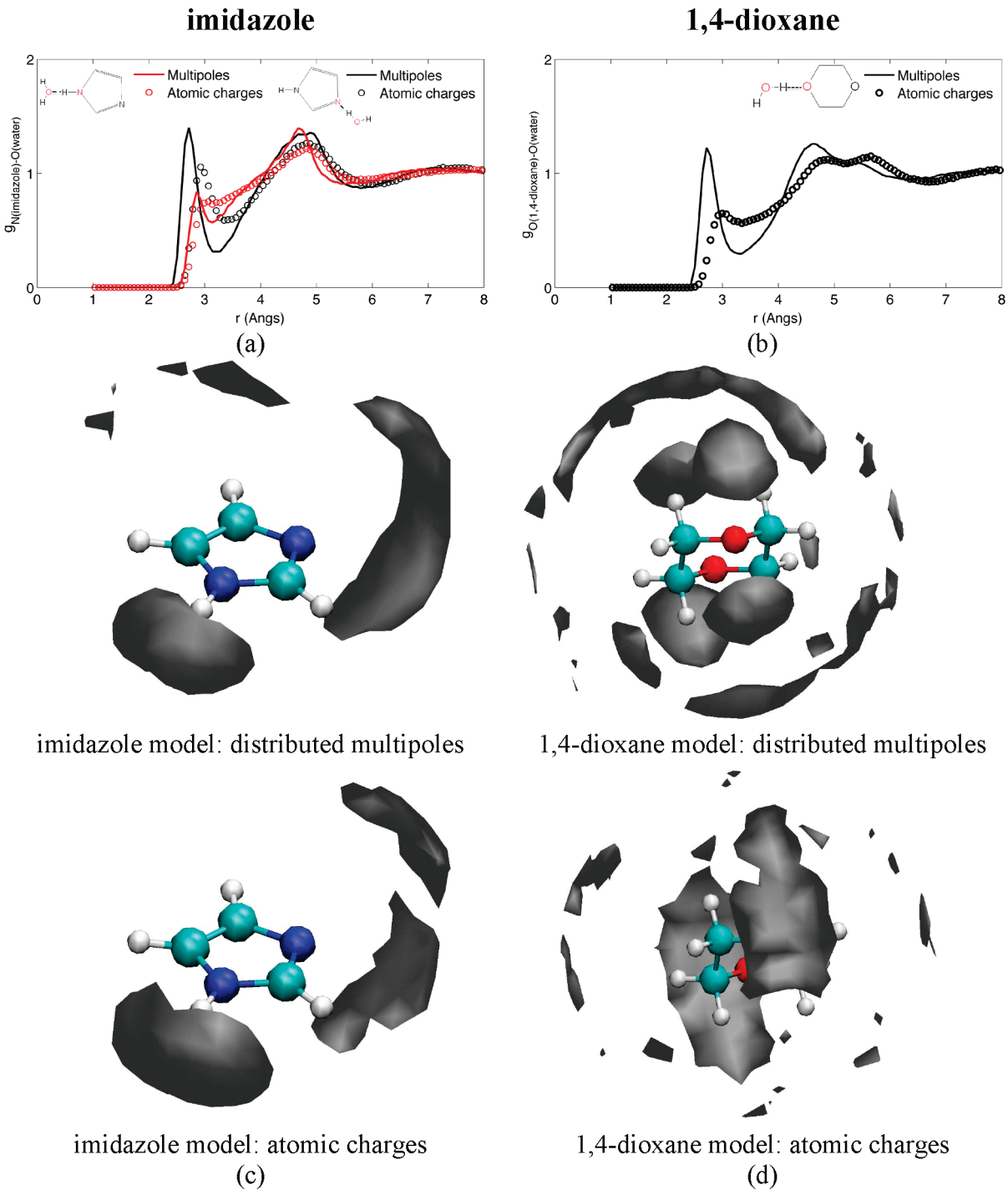


Figure 8. Site−site correlation functions demonstrating the differences in hydration environment of (a) imidazole's and (b) 1,4-dioxane's hydrogen bond donors and acceptors when the solute's charge density is modeled with distributed multipoles up to rank 4 (continuous lines) and atomic charges (open circles). (c and d) Isodensity surfaces corresponding to 3.3 times the average number density of water oxygen atoms in liquid water at ambient conditions for the first hydration shell of imidazole and 1,4-dioxane, respectively.

allows the parametrization of all potential parameters, including the charges and the position of off-nuclei interaction sites, using a much wider range of experimentally measured quantities. Despite its greater computational cost in molecular simulation, the derivation of a multipole model for the dominant electrostatic forces from first principles simplifies the paramerization procedure by reducing the number of independent variables that need to be optimized. Finally, the partial reliance on quantum mechanical calculations, instead of fitting all water potential terms to bulk water properties, is likely to increase the transferability[98] of the water model in hydration simulations.

The dipole moment of the proposed water model is 2.22 D compared with the 1.85 D dipole moment in the gas phase and our computed 2.59 D average dipole moment of liquid water at ambient conditions. By using the effective multipole moments $Q_{\text{eff}} = Q + \Delta Q/2$ we account for the cost of polarizing the water charge density from the gas to the liquid

state and alleviate the need to include a self-polarization correction in the predicted vaporization enthalpy. The inclusion of a self-polarization correction in the vaporization enthalpy has been an important aspect in developing water models, such as in the derivation of SPC/E from SPC[76] and TIP4P/2005[6] from TIP4P.[42] TIP4P and TIP4P/2005 have a dipole moment[88] of 2.18 and 2.31 D, respectively, that is significantly smaller than the liquid water dipole moment in our calculations and other studies,[51,52] and despite the uncertainty in the average induced dipole moment in liquid water $\Delta\mu$,[53] they most likely approximate the $\mu + \Delta\mu/2$ value. This suggests that the self-polarization correction to the predicted vaporization enthalpy with these models may also be not applicable. This reflects the difficulty in developing a nonexplicitly polarizable, classical water model to predict accurately water properties without overestimating the vaporization enthalpy. Similarly, it has been shown that the melting temperature of ice Ih and the vaporization enthalpy and temperature of maximum density cannot be simultaneously predicted,[88] although, once more, no physical explanation is obvious apart from the limitations of the underlying water model. It has been suggested[21] that hydration free energies predicted with fixed (prepolarized) electrostatic models should also be corrected to account for the energy cost associated with the polarization of the solute by the electric field of the solvent. In this work, we do not, however, correct the calculated free energies of hydration, because the solute was also modeled using the effective $Q_{\text{eff}} = Q + \Delta Q/2$ multipole moments.

Using a high-rank, multipole description of the electrostatic forces, the hydration energies of 10 solutes with diverse chemistries are predicted with an RMS error of 1.50 kcal mol^{-1}, which is similar to the error with the self-consistent reaction field method that has been parametrized for this task. A comparative test[13] of implicit- and explicit-solvent free energy approaches for 17 small solutes gave an RMS error in predicted ΔG_{hyd} that ranged between 1.3 and 2.6 kcal mol^{-1}. These results are consistent with a recent blind test,[99] which showed that the solvation free energy of complex, drug-like molecules can at present be predicted with an RMS error of 2.5−3.5 kcal mol^{-1}. In light of these comparative tests, and despite the limited size of the molecules, our predictions are encouraging given that the solute's repulsion−dispersion interactions have not been parametrized for hydration calculations,[98,100] nor in conjunction with explicit or implicit modeling of induction, which leaves a lot of scope for improvement. Indeed, using the effective multipole moments to account for polarization, we predict that the free energies of hydration of imidazole and 1,4-dioxane are 2.5 and 0.6 kcal mol^{-1} more negative compared to the experimental values, respectively. On the other hand, when the solutes are modeled with their gas phase multipole moments, the predicted hydration free energies are 1.8 and 1.7 kcal mol^{-1} more positive compared to experimental results. Hence, the errors of two predictions have opposite sign and show the sensitivity of the hydration free energies to the model for the intermolecular forces. Despite the challenge in predicting ΔG_{hyd} to a target accuracy of 1 $kT = 0.5$ kcal mol^{-1} or better, explicit-solvent free energy methods have the advantage that they can be systematically improved by employing theoretically justified models for the intermolecular forces derived from first principles.[101] In contrast, dielectric continuum methods depend strongly on parameters that have little physical meaning, such as the model for the solute's cavity, the permittivity at the cavity boundary, and the modeling of the nonelectrostatic contributions to ΔG_{hyd},[15] that limit their predictive power for molecules dissimilar from the training set. Unfortunately, the scope for more extensive parametrizations of such methods is limited because experimental data for hydration free energies are sparse,[82] especially for complex, polyfunctional molecules for which predictions are mostly needed.

A comparative, explicit-solvent hydration free energy study with a diverse set of charge models showed that agreement with experimental results does not improve with an increasing level of quantum theory to compute the atomic charges.[22] Our study demonstrates that representing the solute's charge density with an atomic charge model changes the hydration free energy by as much as 6 kcal mol^{-1} compared with a distributed multipole expansion, despite the two models being derived from the same wave function calculation. The discrepancy is greater for very polar molecules that have the most negative free energy of hydration. Hence, the inherent limitations of the monopole model are sufficiently large to suggest that the accurate modeling of electrostatic interactions in dynamic simulations should be a higher priority research goal and perhaps precede the development of isotropic polarizable models. We accept the view that the effect of modeling the electrostatic interactions with an atomic charge model can be partially absorbed in the repulsion−dispersion parametrization. However, the success of this strategy is likely to be limited for a diverse set of chemical functionalities, because the mathematical form of the models for the two intermolecular energy contributions is very different: electrostatic interactions are long-ranged and highly dependent on the relative molecular orientation. We note that the effect of including higher multipole moments may not always be evident in structural reproduction: the isotropic models TIP4P and TIP4P/2005 achieve comparable accuracy with our model in reproducing the structure of ice polymorphs and liquid water. The difference between a multipole and monopole representation of the charge density is more likely to be manifested in evaluating the relative stability of different molecular arrangements, including the relative energy and dynamics of a solute in a solvent and in a vacuum, as shown by the sensitivity of the predicted hydration free energy to the electrostatic model. Similar strong dependence on the electrostatic model has been well established in predicting the relative stability of different packing arrangements of organic molecules.[102,103]

Our ultimate goal is the prediction of the aqueous solubility of crystalline materials, which is defined as the solution concentration for which the chemical potential of the solute in solution and in the solid state are equal. The solubility depends on the relative strength of the solute's intermolecular interactions in the solid state and in solution.[104] Hence, a computationally viable route for computing solubility is through the thermodynamic cycle *crystal structure* → *gas*

phase → solution that involves the calculation of the free energy of solvation and sublimation free energy of the crystal.[11] The latter can be computed using anisotropic model potentials within the harmonic approximation[35] and using the Einstein crystal methodology[84] at elevated temperatures compared with the melting point. In this first publication, we investigated the possibility of computing the hydration free energy using explicit-solvent free energy perturbation and the same anisotropic potentials we have been developing for modeling the crystal structure of organic molecules.[33] It is expected that computing the hydration free energy and sublimation free energy of the crystal using the same accurate model for the intermolecular forces would be advantageous compared with approaches that mix classical lattice energy with self-consistent reaction field quantum mechanical calculations.[11] Our results show that the prediction of solubility using such thermodynamic cycles would require further improvements in the models for the intermolecular forces, given that an error of 1.5 kcal mol^{-1} in hydration free energy alone would cause a discrepancy in the predicted solubility that is comparable with the accuracy of statistical QSPR methods[105,106] to predict solubility, which have negligible computational cost.

Conclusions

We present a rigid-body, implicitly polarized water model based on a high-rank, distributed multipole representation of the quantum-mechanically computed water charge density, which was computed to include average polarization effects in liquid water. The repulsion−dispersion water−water interactions are modeled with an exp-6 potential fitted only to the experimental density and oxygen−oxygen site correlation function of liquid water at ambient conditions. The model performs well in modeling a wide range of water properties not used in its parametrization, including the heat capacity, diffusion coefficient, density maximum of liquid water and vapor−liquid phase equilibria data. This water model was used in explicit-solvent free energy perturbation calculations to compute the hydration free energy of 10 organic solutes. The solute−water interactions were also modeled with an implicitly polarized, distributed multipole model and an empirical exp-6 repulsion−dispersion potential parametrized for organic crystal structures in conjunction with distributed multipoles. The root-mean-square error of the predicted hydration free energies is 1.50 kcal mol^{-1}, which is comparable with the accuracy of a self-consistent reaction field model that had been parametrized explicitly for this task. The free energy of hydration was found to be particularly sensitive to the accuracy of the model for the intermolecular electrostatic forces. Representing the solute's charge density using an atomic charge model changes the predicted hydration free energy by up to 6 kcal mol^{-1} compared with a distributed multipole model computed at the same level of theory. The discrepancy between the two models is very molecule-dependent and provides an estimate for the effect of the modeling quality of the intermolecular electrostatic forces in hydration free energy calculations.

Acknowledgment. P.G.K. thanks Dr. Carlos Avendaño (Imperial College) for useful discussions and Mr. Simon Burbidge for assistance with high-performance computing. Funding to the Molecular Systems Engineering group from the Engineering and Physical Sciences Research Council (EPSRC) of the U.K. [EP/E016340] is gratefully acknowledged. The majority of calculations were performed on the High Performance Computing Cluster of Imperial College. P.R. thanks the Australian Research Council for financial support through a Discovery grant (DP0986999) and NCI and iVEC for the provision of computing resources.

Supporting Information Available: The proposed model for the water−water intermolecular electrostatic and repulsion−dispersion interactions, detailed water vapor−liquid equilibria results, the hydration free energy of 1,4-dioxane and pyrene using a soft-core interaction potential for the solute−water repulsion−dispersion interactions, and the effect of long-range corrections in modeling the hydration free energy of 1,4-dioxane. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Guillot, B. A Reappraisal of What We Have Learnt During Three Decades of Computer Simulations on Water. *J. Mol. Liq.* **2002**, *101*, 219–260.

(2) Paricaud, P.; Predota, M.; Chialvo, A.; Cummings, P. From Dimer to Condensed Phases at Extreme Conditions: Accurate Predictions of the Properties of Water by a Gaussian Charge Polarizable Model. *J. Chem. Phys.* **2005**, *122*, art-244511.

(3) Ren, P.; Ponder, J. Polarizable Atomic Multipole Water Model for Molecular Mechanics Simulation. *J. Phys. Chem. B* **2003**, *107*, 5933–5947.

(4) Lamoureux, G.; MacKerell, A.; Roux, B. A Simple Polarizable Model of Water Based on Classical Drude Oscillators. *J. Chem. Phys.* **2003**, *119*, 5185–5197.

(5) Yu, H.; van Gunsteren, W. Charge-on-Spring Polarizable Water Models Revisited: From Water Clusters to Liquid Water to Ice. *J. Chem. Phys.* **2004**, *121*, 9549–9564.

(6) Abascal, J.; Vega, C. A General Purpose Model for the Condensed Phases of Water: Tip4p/2005. *J. Chem. Phys.* **2005**, *123*, art-234505.

(7) Mahoney, M.; Jorgensen, W. A Five-Site Model for Liquid Water and the Reproduction of the Density Anomaly by Rigid, Nonpolarizable Potential Functions. *J. Chem. Phys.* **2000**, *112*, 8910–8922.

(8) Bukowski, R.; Szalewicz, K.; Groenenboom, G. C.; van der Avoird, A. Predictions of the Properties of Water from First Principles. *Science* **2007**, *315*, 1249–1252.

(9) Wolfenden, R. Waterlogged Molecules. *Science* **1983**, *222*, 1087–1093.

(10) Garrido, N. M.; Queimada, A. J.; Jorge, M.; Macedo, E. A.; Economou, I. G. 1-Octanol/Water Partition Coefficients of N-Alkanes from Molecular Simulations of Absolute Solvation Free Energies. *J. Chem. Theory Comput.* **2009**, *5*, 2436–2446.

(11) Palmer, D. S.; Llinas, A.; Morao, I.; Day, G. M.; Goodman, J. M.; Glen, R. C.; Mitchell, J. B. O. Predicting Intrinsic Aqueous Solubility by a Thermodynamic Cycle. *Mol. Pharm.* **2008**, *5*, 266–279.

(12) Westergren, J.; Lindfors, L.; Hoglund, T.; Luder, K.; Nordholm, S.; Kjellander, R. In Silico Prediction of Drug

Solubility: 1. Free Energy of Hydration. *J. Phys. Chem. B* **2007**, *111*, 1872–1882.

(13) Nicholls, A.; Mobley, D. L.; Guthrie, J. P.; Chodera, J. D.; Bayly, C. I.; Cooper, M. D.; Pande, V. S. Predicting Small-Molecule Solvation Free Energies: An Informal Blind Test for Computational Chemistry. *J. Med. Chem.* **2008**, *51*, 769–779.

(14) Hine, J.; Mookerjee, P. Intrinsic Hydrophilic Character of Organic Compounds - Correlations in Terms of Structural Contributions. *J. Org. Chem.* **1975**, *40*, 292–298.

(15) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Perspective on Foundations of Solvation Modeling: The Electrostatic Contribution to the Free Energy of Solvation. *J. Chem. Theory Comput.* **2008**, *4*, 877–887.

(16) Shirts, M.; Pitera, J.; Swope, W.; Pande, V. Extremely Precise Free Energy Calculations of Amino Acid Side Chain Analogs: Comparison of Common Molecular Mechanics Force Fields for Proteins. *J. Chem. Phys.* **2003**, *119*, 5740–5761.

(17) Shirts, M.; Pande, V. Comparison of Efficiency and Bias of Free Energies Computed by Exponential Averaging, the Bennett Acceptance Ratio, and Thermodynamic Integration. *J. Chem. Phys.* **2005**, *122*, art-144107.

(18) Jorgensen, W. L.; Thomas, L. L. Perspective on Free-Energy Perturbation Calculations for Chemical Equilibria. *J. Chem. Theory Comput.* **2008**, *4*, 869–876.

(19) Rodinger, T.; Pomes, R. Enhancing the Accuracy, the Efficiency and the Scope of Free Energy Simulations. *Curr. Opin. Struct. Biol.* **2005**, *15*, 164–170.

(20) Shirts, M.; Pande, V. Solvation Free Energies of Amino Acid Side Chain Analogs for Common Molecular Mechanics Water Models. *J. Chem. Phys.* **2005**, *122*, art-134508.

(21) Hess, B.; van der Vegt, N. F. A. Hydration Thermodynamic Properties of Amino Acid Analogues: A Systematic Comparison of Biomolecular Force Fields and Water Models. *J. Phys. Chem. B* **2006**, *110*, 17616–17626.

(22) Mobley, D. L.; Dumont, E.; Chodera, J. D.; Dill, K. A. Comparison of Charge Models for Fixed-Charge Force Fields: Small-Molecule Hydration Free Energies in Explicit Solvent. *J. Phys. Chem. B* **2007**, *111*, 2242–2254.

(23) Shivakumar, D.; Deng, Y.; Roux, B. Computations of Absolute Solvation Free Energies of Small Molecules Using Explicit and Implicit Solvent Model. *J. Chem. Theory Comput.* **2009**, *5*, 919–930.

(24) Chialvo, A.; Cummings, P. Simple Transferable Intermolecular Potential for the Molecular Simulation of Water over Wide Ranges of State Conditions. *Fluid Phase Equilib.* **1998**, *150*, 73–81.

(25) Burnham, C.; Xantheas, S. Development of Transferable Interaction Models for Water. III. Reparametrization of an All-Atom Polarizable Rigid Model (Ttm2-R) from First Principles. *J. Chem. Phys.* **2002**, *116*, 1500–1510.

(26) Liem, S. Y.; Popelier, P. L. A. Properties and 3d Structure of Liquid Water: A Perspective from a High-Rank Multipolar Electrostatic Potential. *J. Chem. Theory Comput.* **2008**, *4*, 353–365.

(27) Walsh, T. R.; Liang, T. A Multipole-Based Water Potential with Implicit Polarization for Biomolecular Simulations. *J. Comput. Chem.* **2009**, *30*, 893–899.

(28) Jiang, H.; Jordan, K. D.; Taylor, C. E. Molecular Dynamics Simulations of Methane Hydrate Using Polarizable Force Fields. *J. Phys. Chem. B* **2007**, *111*, 6486–6492.

(29) Liang, T.; Walsh, T. R. Simulation of the Hydration Structure of Glycyl-Alanine. *Mol. Simulat.* **2007**, *33*, 337–342.

(30) Coombes, D.; Price, S.; Willock, D.; Leslie, M. Role of Electrostatic Interactions in Determining the Crystal Structures of Polar Organic Molecules. A Distributed Multipole Study. *J. Phys. Chem.* **1996**, *100*, 7352–7360.

(31) Price, S.; Leslie, M.; Welch, G.; Habgood, M.; Ls, P.; Karamertzanis, P.; Day, G. Modelling Organic Crystal Structures Using Distributed Multipole and Polarizability-Based Model Intermolecular Potentials. Submitted, 2010.

(32) Price, S.; Andrews, J.; Murray, C.; Amos, R. The Effect of Basis Set and Electron Correlation on the Predicted Electrostatic Interactions of Peptides. *J. Am. Chem. Soc.* **1992**, *114*, 8268–8276.

(33) Price, S. L. Computational Prediction of Organic Crystal Structures and Polymorphism. *Int. Rev. Phys. Chem.* **2008**, *27*, 541–568.

(34) Day, G.; Price, S.; Leslie, M. Elastic Constant Calculations for Molecular Organic Crystals. *Cryst. Growth Des.* **2001**, *1*, 13–26.

(35) Day, G.; Price, S.; Leslie, M. Atomistic Calculations of Phonon Frequencies and Thermodynamic Quantities for Crystals of Rigid Organic Molecules. *J. Phys. Chem. B* **2003**, *107*, 10919–10933.

(36) Stone, A. Distributed Multipole Analysis: Stability for Large Basis Sets. *J. Chem. Theory Comput.* **2005**, *1*, 1128–1132.

(37) Welch, G. W. A.; Karamertzanis, P. G.; Misquitta, A. J.; Stone, A. J.; Price, S. L. Is the Induction Energy Important for Modeling Organic Crystals. *J. Chem. Theory Comput.* **2008**, *4*, 522–532.

(38) Misquitta, A. J.; Stone, A. J.; Price, S. L. Accurate Induction Energies for Small Organic Molecules. 2. Development and Testing of Distributed Polarizability Models against Sapt(Dft) Energies. *J. Chem. Theory Comput.* **2008**, *4*, 19–32.

(39) Handley, C. M.; Hawe, G. I.; Kell, D. B.; Popelier, P. L. A. Optimal Construction of a Fast and Accurate Polarisable Water Potential Based on Multipole Moments Trained by Machine Learning. *Phys. Chem. Chem. Phys.* **2009**, *11*, 6365–6376.

(40) Kastenholz, M. A.; Huenenberger, P. H. Computation of Methodology-Independent Ionic Solvation Free Energies from Molecular Simulations. Ii. The Hydration Free Energy of the Sodium Cation. *J. Chem. Phys.* **2006**, *124*, art-224501.

(41) Mobley, D. L.; Bayly, C. I.; Cooper, M. D.; Dill, K. A. Predictions of Hydration Free Energies from All-Atom Molecular Dynamics Simulations. *J Phys Chem B.* **2009**, *113*, 4533–4537.

(42) Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(43) Frisch, M.; Trucks, G.; Schlegel, H.; Scuseria, G.; Robb, M.; Cheeseman, J.; Montgomery, J.; Vreven, T.; Kudin, K.; Burant, J.; Millam, J.; Iyengar, S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J.; Hratchian, H.; Cross, J.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R.; Yazyev, O.; Austin, A.; Cammi, R.; Pomelli, C.; Ochterski, J.; Ayala, P.; Morokuma, K.; Voth,

G.; Salvador, P.; Dannenberg, J.; Zakrzewski, V.; Dapprich, S.; Daniels, A.; Strain, M.; Farkas, O.; Malick, D.; Rabuck, A.; Raghavachari, K.; Foresman, J.; Ortiz, J.; Cui, Q.; Baboul, A.; Clifford, S.; Cioslowski, J.; Stefanov, B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R.; Fox, D.; Keith, T.; Al Laham, M.; Peng, C.; Nanayakkara, A.; Challacombe, M.; Gill, P.; Johnson, B.; Chen, W.; Wong, M.; Gonzalez, C.; Pople, J. *Gaussian 03*; Gaussian Inc.: Wallingford, CT, 2003.

(44) Horn, H.; Swope, W.; Pitera, J.; Madura, J.; Dick, T.; Hura, G.; Head-Gordon, T. Development of an Improved Four-Site Water Model for Biomolecular Simulations: Tip4p-Ew. *J. Chem. Phys.* **2004**, *120*, 9665–9678.

(45) Liem, S.; Popelier, P.; Leslie, M. Simulation of Liquid Water Using a High-Rank Quantum Topological Electrostatic Potential. *Int. J. Quantum Chem.* **2004**, *99*, 685–694.

(46) Stone, A. J. Water from First Principles. *Science* **2007**, *315*, 1228–1229.

(47) Misquitta, A.; Stone, A. Distributed Polarizabilities Obtained Using a Constrained Density-Fitting Algorithm. *J. Chem. Phys.* **2006**, *124*, art-024111.

(48) Millot, C.; Soetens, J.; Costa, M.; Hodges, M.; Stone, A. Revised Anisotropic Site Potentials for the Water Dimer and Calculated Properties. *J. Phys. Chem. A* **1998**, *102*, 754–770.

(49) Tu, Y.; Laaksonen, A. The Electronic Properties of Water Molecules in Water Clusters and Liquid Water. *Chem. Phys. Lett.* **2000**, *329*, 283–288.

(50) Breneman, C.; Wiberg, K. Determining Atom-Centered Monopoles from Molecular Electrostatic Potentials - the Need for High Sampling Density in Formamide Conformational-Analysis. *J. Comput. Chem.* **1990**, *11*, 361–373.

(51) Gregory, J.; Clary, D.; Liu, K.; Brown, M.; Saykally, R. The Water Dipole Moment in Water Clusters. *Science* **1997**, *275*, 814–817.

(52) Silvestrelli, P.; Parrinello, M. Water Molecule Dipole in the Gas and in the Liquid Phase. *Phys. Rev. Lett.* **1999**, *82*, 3308–3311.

(53) Handley, C. M.; Popelier, P. L. A. The Asymptotic Behavior of the Dipole and Quadrupole Moment of a Single Water Molecule from Gas Phase to Large Clusters: A Qct Analysis. *Synth. React. Inorg. Met.* **2008**, *38*, 91–99.

(54) Stone, A. J. *Orient*, 4.6; University of Cambridge: Cambridge, 2009. http://www-stone.ch.cam.ac.uk/programs.html (accessed Mar 2010).

(55) Stone, A. *The Theory of Intermolecular Forces*; Clarendon Press: Oxford, 1996.

(56) Leslie, M. Dl_Multi - a Molecular Dynamics Program to Use Distributed Multipole Electrostatic Models to Simulate the Dynamics of Organic Crystals. *Mol. Phys.* **2008**, *106*, 1567–1578.

(57) Soper, A. The Radial Distribution Functions of Water and Ice from 220 to 673 K and at Pressures up to 400 MPa. *Chem. Phys.* **2000**, *258*, 121–137.

(58) de la Pena, L.; Kusalik, P. Quantum Effects in Light and Heavy Liquid Water: A Rigid-Body Centroid Molecular Dynamics Study. *J. Chem. Phys.* **2004**, *121*, 5992–6002.

(59) Mahoney, M.; Jorgensen, W. Quantum, Intramolecular Flexibility, and Polarizability Effects on the Reproduction of the Density Anomaly of Liquid Water by Simple Potential Functions. *J. Chem. Phys.* **2001**, *115*, 10758–10768.

(60) Allesch, M.; Schwegler, E.; Gygi, F.; Galli, G. A First Principles Simulation of Rigid Water. *J. Chem. Phys.* **2004**, *120*, 5192–5198.

(61) Lemmon, E.; McLinden, M.; Friend, D. Thermophysical Properties of Fluid Systems. In *NIST Chemistry Webbook, NIST Standard Reference Database Number 69*; Linstrom, P., Mallard, W., Eds.; National Institute of Standars and Technology: Gaithersburg, MD. http://webbook.nist.gov (retrieved September 20, 2009).

(62) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, *105*, 2999–3093.

(63) Haslam, A. J.; Galindo, A.; Jackson, G. Prediction of Binary Intermolecular Potential Parameters for Use in Modelling Fluid Mixtures. *Fluid Phase Equilib.* **2008**, *266*, 105–128.

(64) Leadbetter, A.; Ward, R.; Clark, J.; Tucker, P.; Matsuo, T.; Suga, H. The Equilibrium Low-Temperature Structure of Ice. *J. Chem. Phys.* **1985**, *82*, 424–428.

(65) Kamb, B.; Hamilton, W.; Laplaca, S.; Prakash, A. Ordered Proton Configuration in Ice-Ii, from Single-Crystal Neutron Diffraction. *J. Chem. Phys.* **1971**, *55*, 1934–1945.

(66) Laplaca, S.; Hamilton, W.; Kamb, B.; Prakash, A. Nearly Proton-Ordered Structure for Ice Ix. *J. Chem. Phys.* **1973**, *58*, 567–580.

(67) Kuhs, W.; Finney, J.; Vettier, C.; Bliss, D. Structure and Hydrogen Ordering in Ice-Vi, Ice-Vii, and Ice-Viii by Neutron Powder Diffraction. *J. Chem. Phys.* **1984**, *81*, 3612–3623.

(68) Salzmann, C.; Radaelli, P.; Hallbrucker, A.; Mayer, E.; FINNEY, J. The Preparation and Structures of Hydrogen Ordered Phases of Ice. *Science* **2006**, *311*, 1758–1761.

(69) Willock, D.; Price, S.; Leslie, M.; Catlow, C. The Relaxation of Molecular-Crystal Structures Using a Distributed Multipole Electrostatic Model. *J. Comput. Chem.* **1995**, *16*, 628–647.

(70) Chisholm, J.; Motherwell, S. Compack: A Program for Identifying Crystal Structure Similarity Using Distances. *J. Appl. Crystallogr.* **2005**, *38*, 228–231.

(71) Allen, F.; Taylor, R. Research Applications of the Cambridge Structural Database (Csd). *Chem. Soc. Rev.* **2004**, *33*, 463–475.

(72) Mootz, D.; Wussow, H. A Novel Proton-Ordered, Two-Dimensional Cross-Linking of Water-Molecules in Pyridine Trihydrate. *Angew. Chem., Int. Ed.* **1980**, *19*, 552–553.

(73) Mootz, D.; Wussow, H. Crystal-Structures of Pyridine and Pyridine Trihydrate. *J. Chem. Phys.* **1981**, *75*, 1517–1522.

(74) Allen, M.; Tildesley, D. *Computer Simulation of Liquids*; Oxford University Press: New York, 1992.

(75) Berendsen, H.; Grigera, J.; Straatsma, T. The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* **1987**, *91*, 6269–6271.

(76) Alejandre, J.; Tildesley, D.; Chapela, G. Molecular-Dynamics Simulation of the Orthobaric Densities and Surface-Tension of Water. *J. Chem. Phys.* **1995**, *102*, 4574–4583.

(77) Vega, C.; de Miguel, E. Surface Tension of the Most Popular Models of Water by Using the Test-Area Simulation Method. *J. Chem. Phys.* **2007**, *126*, art-154707.

(78) Deng, Y.; Roux, B. Hydration of Amino Acid Side Chains: Nonpolar and Electrostatic Contributions Calculated from Staged Molecular Dynamics Free Energy Simulations with

Explicit Water Molecules. *J. Phys. Chem. B* **2004**, *108*, 16567–16576.

(79) Bennett, C. Efficient Estimation of Free-Energy Differences from Monte-Carlo Data. *J. Comput. Phys.* **1976**, *22*, 245–268.

(80) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, Structures, and Electronic Properties of Molecules in Solution with the C-Pcm Solvation Model. *J. Comput. Chem.* **2003**, *24*, 669–681.

(81) Cramer, C. J.; Truhlar, D. G. A Universal Approach to Solvation Modeling. *Acc. Chem. Res.* **2008**, *41*, 760–768.

(82) Guthrie, J.; Povar, I. A Test of Various Computational Solvation Models on a Set Of "Difficult" Organic Compounds. *Can. J. Chem.* **2009**, *87*, 1154–1162.

(83) Jorgensen, W.; Blake, J.; Buckner, J. Free-Energy of Tip4p Water and the Free-Energies of Hydration of Ch4 and Cl- from Statistical Perturbation-Theory. *Chem. Phys.* **1989**, *129*, 193–200.

(84) Frenkel, D.; B., S. *Understanding Molecular Simulation*; Academic Press: San Diego, 2002.

(85) Wescott, J.; Fisher, L.; Hanna, S. Use of Thermodynamic Integration to Calculate the Hydration Free Energies of N-Alkanes. *J. Chem. Phys.* **2002**, *116*, 2361–2369.

(86) Hulme, A. T.; Price, S. L. Toward the Prediction of Organic Hydrate Crystal Structures. *J. Chem. Theory Comput.* **2007**, *3*, 1597–1608.

(87) Fortes, A.; Wood, I.; Alfredsson, M.; Vocadlo, L.; Knight, K. The Incompressibility and Thermal Expansivity of D2o Ice Ii Determined by Powder Neutron Diffraction. *J. Appl. Crystallogr.* **2005**, *38*, 612–618.

(88) Vega, C.; Abascal, J. L. F.; Conde, M. M.; Aragones, J. L. What Ice Can Teach Us About Water Interactions: A Critical Comparison of the Performance of Different Water Models. *Faraday Discuss.* **2009**, *141*, 251–276.

(89) Kell, G. Precise Representation of Volume Properties of Water at One Atmosphere. *J. Chem. Eng. Data* **1967**, *12*, 66–69.

(90) Jorgensen, W.; Tirado-Rives, J. Potential Energy Functions for Atomic-Level Simulations of Water and Organic and Biomolecular Systems. *Proc. Natl. Acad. Sci U. S. A.* **2005**, *102*, 6665–6670.

(91) Hare, D.; Sorensen, C. The Density of Supercooled Water 0.2. Bulk Samples Cooled to the Homogeneous Nucleation Limit. *J. Chem. Phys.* **1987**, *87*, 4840–4845.

(92) Jancso, G.; Vanhook, W. Condensed Phase Isotope-Effects (Especially Vapor-Pressure Isotope-Effects). *Chem. Rev.* **1974**, *74*, 689–750.

(93) Billeter, S.; King, P.; Vangunsteren, W. Can the Density Maximum of Water Be Found by Computer-Simulation. *J. Chem. Phys.* **1994**, *100*, 6692–6699.

(94) Krynicki, K.; Green, C.; Sawyer, D. Pressure and Temperature-Dependence of Self-Diffusion in Water. *Faraday Discuss.* **1978**, *66*, 199–208.

(95) Mills, R. Self-Diffusion in Normal and Heavy-Water in Range 1−45 Degrees. *J. Phys. Chem.* **1973**, *77*, 685–688.

(96) Hermans, J.; Pathiaseril, A.; Anderson, A. Excess Free-Energy of Liquids from Molecular-Dynamics Simulations - Application to Water Models. *J. Am. Chem. Soc.* **1988**, *110*, 5982–5986.

(97) Rizzo, R.; Aynechi, T.; Case, D.; Kuntz, I. Estimation of Absolute Free Energies of Hydration Using Continuum Methods: Accuracy of Partial, Charge Models and Optimization of Nonpolar Contributions. *J. Chem. Theory Comput.* **2006**, *2*, 128–139.

(98) Docherty, H.; Galindo, A.; Vega, C.; Sanz, E. A Potential Model for Methane in Water Describing Correctly the Solubility of the Gas and the Properties of the Methane Hydrate. *J. Chem. Phys.* **2006**, *125*, art-074510.

(99) Guthrie, J. P. A Blind Challenge for Computational Solvation Free Energies: Introduction and Overview. *J. Phys. Chem. B* **2009**, *113*, 4501–4507.

(100) Oostenbrink, C.; Villa, A.; Mark, A.; van Gunsteren, W. A Biomolecular Force Field Based on the Free Enthalpy of Hydration and Solvation: The Gromos Force-Field Parameter Sets 53a5 and 53a6. *J. Comput. Chem.* **2004**, *25*, 1656–1676.

(101) Misquitta, A. J.; Welch, G. W. A.; Stone, A. J.; Price, S. L. A First Principles Prediction of the Crystal Structure of C6br2clfh2. *Chem. Phys. Lett.* **2008**, *456*, 105–109.

(102) Karamertzanis, P. G.; Kazantsev, A. V.; Issa, N.; Welch, G. W. A.; Adjiman, C. S.; Pantelides, C. C.; Price, S. L. Can the Formation of Pharmaceutical Cocrystals Be Computationally Predicted? 2. Crystal Structure Prediction. *J. Chem. Theory Comput.* **2009**, *5*, 1432–1448.

(103) Day, G.; Motherwell, W.; Jones, W. Beyond the Isotropic Atom Model in Crystal Structure Prediction of Rigid Molecules: Atomic Multipoles Versus Point Charges. *Cryst. Growth Des.* **2005**, *5*, 1023–1033.

(104) Huang, L.; Tong, W. Impact of Solid State Properties on Developability Assessment of Drug Candidates. *Adv. Drug Delivery Rev.* **2004**, *56*, 321–334.

(105) Johnson, S. R.; Chen, X.-Q.; Murphy, D.; Gudmundsson, O. A Computational Model for the Prediction of Aqueous Solubility That Includes Crystal Packing, Intrinsic Solubility, and Ionization Effects. *Mol. Pharm.* **2007**, *4*, 513–523.

(106) Delaney, J. Predicting Aqueous Solubility from Structure. *Drug Discovery Today.* **2005**, *10*, 289–295.

CT900693Q