THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Systematic Improvement of a Classical Molecular Model of Water

Lee-Ping Wang,[†] Teresa Head-Gordon,[‡] Jay W. Ponder,[§] Pengyu Ren,[∥] John D. Chodera,[⊥] Peter K. Eastman,[†] Todd J. Martinez,[†] and Vijay S. Pande*,[†]

[†]Department of Chemistry, Stanford University, Stanford, California 94305, United States

[‡]Department of Chemistry, Bioengineering, Chemical & Biomolecular Engineering, University of California, Berkeley, California 94720, United States

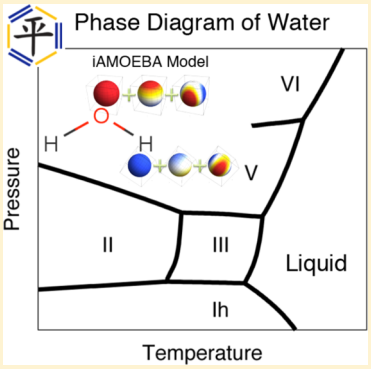
[§]Department of Chemistry, Washington University, St. Louis, Missouri 63130, United States

[∥]Department of Biomedical Engineering, University of Texas, Austin, Texas 78712, United States

[⊥]Computational Biology Program, Sloan-Kettering Institute, New York, New York 10065, United States

Supporting Information

ABSTRACT: We report the iAMOEBA ("inexpensive AMOEBA") classical polarizable water model. The iAMOEBA model uses a direct approximation to describe electronic polarizability, in which the induced dipoles are determined directly from the permanent multipole electric fields and do not interact with one another. The direct approximation reduces the computational cost relative to a fully self-consistent polarizable model such as AMOEBA. The model is parameterized using ForceBalance, a systematic optimization method that simultaneously utilizes training data from experimental measurements and high-level *ab initio* calculations. We show that iAMOEBA is a highly accurate model for water in the solid, liquid, and gas phases, with the ability to fully capture the effects of electronic polarization and predict a comprehensive set of water properties beyond the training data set including the phase diagram. The increased accuracy of iAMOEBA over the fully polarizable AMOEBA model demonstrates ForceBalance as a method that allows the researcher to systematically improve empirical models by efficiently utilizing the available data.



INTRODUCTION

Water is a fascinating liquid that possesses many anomalous physical and chemical properties, including the temperature of maximum density, expansion on freezing, unusually large heat capacity, compressibility minimum, and unique solvation properties. The study of how the molecular interactions in water are related to its unique properties is a fundamental and important topic in physical chemistry. A related essential question is how water interacts with solutes such as ions, organic molecules, and proteins to form the foundation of biomolecular structure and function, including studies of how proteins fold,[1−3] misfold,[4−6] undergo conformational change,[7−9] and interact with their environments.[10−14] For all of these reasons, the study of water interactions has been a highly active field for decades in experimental and theoretical chemistry.[15]

The theoretical and computational modeling of water allows us to investigate water and aqueous solutions with precise spatial and temporal resolution, providing a helpful complement to experiment. The application of these models has contributed to important progress in our understanding of pure water properties, such as the nature of the hydrogen bonding network and its structural organization,[15−18] the transport of hydronium and hydroxide ions,[19,20] the mechanism of freezing,[21] the surface properties of ice,[22] the geometries and binding energies of water clusters in the gas phase,[23−26] and the dual phases of liquid water at extremely low temperatures.[27−29] However, no current single model is able to capture the full complexity of water and its properties, due to the trade-off between model complexity and the ability to sample the condensed phase.

For example, some of the most widely used biomolecular force fields—including the popular AMBER[30−33] and CHARMM models[13,34]—are based on the three-site, pairwise-additive TIP3P water model[35] developed for use with Monte Carlo simulations with finite-range cutoffs for electrostatic interactions, nearly 30 years ago. The computationally tractable TIP3P model allows for long simulation time scales to be reached that are important for extreme regions of water's phase diagram, or solvation of large biomolecules, even though it was not parameterized for these types of simulations. The TIP3P model was fitted to the experimental density and heat of vaporization at room temperature and 1 atm pressure, and its validation showed it could predict the isobaric heat capacity quite well at this temperature and pressure—to within 1 J/g °C.[36] However, this model often fails on properties outside the

Received: April 17, 2013
Revised: June 7, 2013
Published: June 11, 2013

 © 2013 American Chemical Society dx.doi.org/10.1021/jp403802c | J. Phys. Chem. B 2013, 117, 9956−9972

ambient conditions for which it was parameterized; for example, the TIP3P model yields a simulated freezing point value of −91 °C, far below the experimental value of 0 °C, and incorrectly predicts the solid phase to be ice II rather than ordinary ice Ih.[37−39]

The development of water models using only pairwise-additive interactions (such as TIP3P) has significantly benefited from more sophisticated treatments of long-ranged electrostatics, extending the experimental data sets used for parameterization, and adding more interaction sites. For example, the TIP4P/Ew[40] and TIP4P/2005 models[41] were developed for use with Ewald summation techniques and parameterized using the temperature dependence of the density and heat of vaporization. These models provide agreement with a much broader range of experimental reference data outside of the parameterization data used in model refinement.

There is also a more sophisticated class of water models that move beyond point charges and the pairwise additive approximation by including anisotropic electrostatic multipole interactions and N-body polarization. For example, polarizable models such as SWM-4DP,[42] TIP4P-FQ,[43] DPP2,[23] TTM3-F,[44] and AMOEBA[45,46] further improve on the description of many water properties, especially gas-phase properties and the dielectric constant—whereas additive force fields typically overestimate binding energies of gas-phase clusters and underestimate the dielectric constant in order to capture a larger subset of condensed phase properties. Given the increase in model complexity, polarizable force fields are also more expensive to evaluate, often by a factor of 2−3 or greater.

The parameterization and validation of a water model is typically both difficult and time-consuming. Improved methods for building force fields are therefore desirable, for they can be used to improve models in a systematic and reproducible way, and potentially be applied more generally for parameterizing models of other solutes such as drugs, ligands, proteins, lipids, or nucleic acids that are compatible with the water model.

In this Article, we report the parameterization and validation of a new classical, flexible, direct polarization water model that is a simplified version of the fully polarizable AMOEBA water model. The direct polarization method determines induced dipole moments directly, based only on the electric field due to static multipoles.[47−50] The direct polarization approximation eliminates the need for an iterative solution to the self-consistent field in a fully interacting (mutual) model of polarization and reduces the computational cost; hence, we refer to the new model as inexpensive AMOEBA or iAMOEBA. However, neglecting mutual polarization results in a loss of ∼20% of the polarization energy, and that must be recovered through optimization of the model parameters.

Here, we use the ForceBalance method[51,52] for parameterizing the iAMOEBA model using a combination of experimental data and high-level *ab initio* calculations. We will describe our choice of reference data and provide a brief overview of the parameterization method. We demonstrate the accuracy of iAMOEBA using a published, comprehensive benchmark of water properties developed by Vega and co-workers[36] that covers a wide range of phases and thermodynamic conditions going far beyond the parameterization data set. We provide further discussion on the generality of the parameterization approach explored here for iAMOEBA and its potential for force field development for other liquids, biomolecular solutes, and other materials, or for potentials of arbitrary functional form.

THEORY AND METHODS

A. The iAMOEBA Model. The iAMOEBA model described in this work is a direct polarization approximation of the AMOEBA water model developed by Ren and Ponder, and we refer the reader to ref 45 for a complete description of the mutually polarizable AMOEBA water model. The iAMOEBA and AMOEBA functional form for water is briefly recapitulated here:

$$E = E_{\mathrm{bond}} + E_{\mathrm{angle}} + E_{\mathrm{b}\theta} + E_{\mathrm{vdW}} + E_{\mathrm{ele}}^{\mathrm{perm}} + E_{\mathrm{ele}}^{\mathrm{ind}} \tag{1}$$

where the first three terms describe the short-range valence interactions (bond stretching, angle bending, Urey−Bradley bond-angle cross-term) and the last three terms are the nonbonded van der Waals (vdW) and electrostatic contributions from permanent and induced dipoles. More specifically, the nuclei are described using classical point particles and the molecular connectivity is fixed; the water molecule is allowed to undergo classical vibrations represented using anharmonic potential functions in the O−H bond length, the H−O−H angle, and the H−H distance (Urey−Bradley interaction). The van der Waals interactions are described using a Halgren buffered 14−7 function that describes dispersion interactions at long-range and exchange repulsion at short-range. The permanent electrostatic interactions are represented as atomic multipole moments through the quadrupole, requiring the definition of a local coordinate system for each atom.

The main difference between the iAMOEBA and AMOEBA functional forms is that we evaluate only the *direct polarization*,[47−49,53] and not the full self-consistent mutual polarization. In this approximation, the polarizable dipoles are induced solely by the electric fields from the permanent multipoles. The dipoles are given by

$$\mu_{i,\alpha}^{\mathrm{ind}} = \alpha_i\Big(\sum_{\{j\}} T_{\alpha}^{ij} M_j\Big) \quad \text{for} \quad \alpha = 1, 2, 3 \tag{2}$$

where α_i is the atomic polarizability on site i, $M_j = [q_j, \mu_{j,1}, \mu_{j,2}\mu_{j,3}, ...]^T$ contains the permanent multipole components on site j, and $T_{\alpha}^{ij} = [T_{\alpha}, T_{\alpha 1}, T_{\alpha 2}, T_{\alpha 3}, ...]$ is the interaction matrix element between sites i and j following Stone's notation.[54] This stands in contrast to the expression for full self-consistent polarizability:

$$\mu_{i,\alpha}^{\mathrm{ind}} = \alpha_i\Big(\sum_{\{j\}} T_{\alpha}^{ij} M_j + \sum_{\{j'\}} T_{\alpha\beta}^{ij'} \mu_{j'\beta}^{\mathrm{ind}}\Big) \quad \text{for} \quad \alpha, \beta = 1, 2, 3 \tag{3}$$

where the additional term $\sum_{\{j'\}} T_{\alpha\beta}^{ij'} \mu_{j'\beta}^{\mathrm{ind}}$ represents the electric fields from all other induced dipoles and the index j' runs over all atomic sites not including i itself. The full expressions for the T matrix elements are the same as eqs 2−5 in ref 45.

Within the framework of the many-body expansion, direct polarization gives rise to at most three-body terms in which a polarizable site couples with the permanent moments of two other sites. For example, in a direct polarization interaction between an inducible dipole and two permanent multipoles, the induced dipole moment and the direct polarization energy are given by

dx.doi.org/10.1021/jp403802c | J. Phys. Chem. B 2013, 117, 9956−9972

$$\mu_{1,\alpha}^{\text{ind}} = \alpha_1(T_\alpha^{12}M_2 + T_\alpha^{13}M_3) \quad \text{for} \quad \alpha = 1, 2, 3$$

$$E_{\text{direct}} = \frac{1}{2}(\mu_1^{\text{ind}}\cdot M_2 + \mu_1^{\text{ind}}\cdot M_3)$$

$$= \frac{1}{2}\alpha_1 \sum_{\alpha=1}^{3} (M_2T_\alpha^{12}M_2 + M_3T_\alpha^{12}M_3 + M_3T_\alpha^{12}M_2 + M_2T_\alpha^{13}M_3) \tag{4}$$

where μ_1^{ind} is the induced dipole on site 1 with polarizability α_1 and $T_\alpha^{12}M_2$ and $T_\alpha^{13}M_3$ represent the electric fields from the permanent multipoles on sites 2 and 3. The first and second terms in E_{direct} are two-body terms, whereas the third and fourth terms are three-body terms representing the influence of M_3 on the interaction between μ_1^{ind} and M_2 (and vice versa). The extension to systems containing more than three particles is straightforward and gives rise to more three-body terms involving one induced dipole and two permanent multipoles. By contrast, the mutual scheme in AMOEBA represents the full N-body polarization interaction, in which the polarizable dipoles further induce one another and must be converged self-consistently. Direct polarization provides a model that still captures much of the polarization effects important in the condensed phase at greatly reduced computational cost, as the iterative evaluation of electrostatic energies to achieve self-consistency in polarizable dipoles is avoided entirely. Furthermore, the resulting model no longer requires a convergence tolerance, eliminating the possibility that incompletely converged dipoles might result in nonconservative forces.

By way of the direct polarization, the iAMOEBA model speeds up the calculation of energies and gradients by a factor of 1.5−6 over the mutual AMOEBA model. These estimates were derived from benchmark calculations using the TINKER and OpenMM software packages (see Table S2, Supporting Information); the precise value depends on the system and simulation settings, notably the self-consistent convergence tolerance for the AMOEBA induced dipoles. Furthermore, the use of extended Lagrangian formalisms[55−58] is known to greatly reduce the computational cost of mutual polarization methods, but it requires a careful choice of the fictitious mass parameter for the electronic degrees of freedom;[59−61] we did not consider the impact of these methods in these timing comparisons.

The more important scientific question is how effectively we can recapture the physical interactions using the approximate polarizable form. In order to address this question, we must reoptimize the 19 independent parameters of the iAMOEBA model: five due to intramolecular vibrations, two for van der Waals interactions, nine for permanent multipoles, and three for electronic polarization. This requires a discussion of the reference data set and the optimization method used to parameterize the model.

B. Reference Data. A significant challenge for model parameterization and validation is the choice of suitable reference data, which can be both experimental and theoretical in origin. For water, we are fortunate that both experimental and theoretical reference data are plentiful,[62,63] although this is often not the case for less studied materials and compounds. In this work, we apply a systematic optimization method (ForceBalance),[51,52] which allows us to efficiently utilize combinations of experimental and theoretical data. Here, we describe the different types of data used to parameterize iAMOEBA, which we summarize in Table 1. Due to the large size of the data set, the complete reference data set is provided electronically.[64]

Table 1. Data References for Parameterization of iAMOEBA (Full Tables Online, ref 64)[a]

Reference Data	Scaling Factor
Density ρ	2 kg m^{-3}
Heat of Vaporization ΔH_{vap}	5 kJ mol^{-1}
Thermal Expansion Coefficient α	10^{-4} K^{-1}
Isothermal Compressibility κ_T	10^{-5} bar^{-1}
Isobaric Heat Capacity c_P	2 kg mol^{-1} K^{-1}
Dielectric Constant $\varepsilon(0)$	2
Gas Phase Dipole	0.2 Debye
Gas Phase Quadrupole	1.0 Debye Å
Gas Phase Vibrational Modes	30 cm^{-1}
Water Dimer CCSD(T)/CBS Binding Energy, RMSD	1.0 kcal mol^{-1}, 0.1 Å
21 Small Gas Phase Clusters (size 3-6) CCSD(T)/CBS Binding Energy, RMSD	10.0 kcal mol^{-1}, 2.0 Å
18 Large Gas Phase Clusters (size 8-20) MP2/CBS Binding Energy, RMSD	100.0 kcal mol^{-1}, 2.0 Å
42,000 MP2/aug-cc-pVTZ Potential Energies and Atomistic Forces	stdev(E_{QM}), RMS($\lvert\mathbf{F}_{QM}\rvert$)

[a]Orange: Condensed phase experimental data, 249−373 K (1 atm), 1−8000 bar (298 K). Green: Gas phase experimental data. Blue: Theoretical reference data. The scaling factors for potential energies and atomistic forces are given by the standard deviation of the potential energies and the RMS norm of the forces in the ab initio reference calculations. The objective function is a scaled sum of squared differences between the simulation results and the reference data; the scaling factors (equivalent to inverse weights) are given in the right column.

Experimental Measurements. The dominant paradigm in water model development is to fit the parameters to reproduce a set of experimentally measured condensed phase properties. Generally speaking, a diverse data set over a wide range of thermodynamic conditions improves the domain of applicability of the model, but it also increases the practical difficulty of the optimization problem. In this work, we extend the parameterization data sets used in past studies to include six key experimental properties over a wide range of thermodynamic conditions: density, heat of vaporization, thermal expansion coefficient, isothermal compressibility, isobaric heat capacity, and dielectric constant. These properties are evaluated at 32 temperatures spanning a range of 249.15−373.15 K at atmospheric pressure and 10 pressures from 1 to 8000 bar at 298.15 K (Table S4, Supporting Information). Our liquid simulations for the heat of vaporization were performed at atmospheric pressure rather than the vapor pressure, which introduces a negligible correction term at the temperature range studied. We also include experimentally known properties of the water monomer in the gas phase, in particular its dipole moment, quadrupole moment, and vibrational frequencies. All of the experimental reference data, along with the theoretical reference data, is summarized in Table 1; the complete experimental data is provided in Tables S3 and S4 (Supporting Information), as well as on the Web.[64]

Theoretical Reference Data. In addition to experimental data, highly detailed theoretical calculations also provide valuable reference data for model development. Methods

following this approach include potential fitting,[65−67] force matching,[51,68−70] and relative entropy[71,72] methods, and they have found widespread success in building both atomistic and mesoscale (coarse grained) models. The general goal is to reproduce the accuracy of the highly detailed reference levels of theory with a comparatively simple and inexpensive classical model. Quantum chemistry methods such as Møller−Plesset perturbation theory,[73] coupled-cluster theory,[73] and density functional theory[74] are examples of reference theories used to develop atomistic models; they are impractical for many condensed-phase applications due to their high cost but provide valuable information on the electronic potential energy surface for smaller systems such as water clusters.

In this work, the theoretical reference data includes energies and gradients calculated at the dual basis RI-MP2[75−77]/heavy-aug-cc-pVTZ[78] level of theory for over 42 000 cluster geometries extracted randomly from simulations of liquid water. These geometries are taken from condensed-phase simulations performed using the AMOEBA model at temperatures ranging from 249.15 to 373.15 K and including cluster sizes ranging from 2 to 22 to minimize finite size effects. The calculations were performed using Q-Chem 4.0;[79] the RI approximation and the dual-basis approximation used their respective auxiliary basis sets corresponding to heavy-aug-cc-pVTZ as implemented in Q-Chem. We also include the optimal geometries and binding energies of 40 small water clusters ranging from 2 to 20 molecules. We have used the highest available level of theory for each cluster, with some of the largest calculations taken from literature benchmarks,[80−85] and include multiple key geometries for the dimer, hexamer, and larger clusters where available. Most of the reference calculations of the binding energy allow the cluster and monomer geometries to fully relax, so our calculations using the model follow the same approach; we include the RMSD of the minimized cluster into the objective function in order to ensure that the iAMOEBA model provides energy-minimized clusters with the correct geometry. We also include 10 dimer poses (the "Smith" dimer set)[23] where the geometries are not relaxed, and only the interaction energies enter into the objective function. All of the data used in the parameterization of iAMOEBA is available on the Web.[64]

C. Optimization. Least squares optimization of force fields first began with the consistent force field proposed by Lifson and Warshel in the 1960s.[86] Hagler and co-workers first proposed removal of hydrogen vdW sites in order to improve the description of hydrogen bonding.[87] Other early efforts extended formal least-squares optimization through use of *ab initio* calculations[88] and application to bulk phase crystal modeling.[89] The AMOEBA water model was parameterized by hand to fit results from *ab initio* calculations on gas phase clusters.

ForceBalance[51,52] (Figure 1) extends this prior work in several directions, including the ability to use a much larger and more diverse data set which includes experimental liquid phase measurements and *ab initio* calculations. Here we use ForceBalance to optimize the iAMOEBA parameters. The overall objective function is expressed as a weighted sum of squared residuals over the experimental and theoretical target data sets (weights and data types in Table 1). The exact gradient and approximate Hessian matrix of the objective function is derived from the first derivatives of the properties using the Gauss−Newton approximation. The Levenberg−Marquardt algorithm[90,91] with an adaptive trust radius[92,93] is used to perform an iterative minimization of the nonlinear least-squares objective function.

There are significant challenges involved in using such a large experimental reference data set in an optimization scheme, because many of these properties are difficult to simulate to convergence for a single set of parameters, let alone fit via a parameter optimization method. Furthermore, computer simulations of condensed-phase properties suffer from statistical errors, and the errors are compounded when estimating the dependence of a particular property on the force field parameters. These difficulties can relegate the researcher to performing manual parameter searches guided by insight and evaluating models by inspection, which gives force field model parameterization some of its reputation as an onerous task or "black art".

ForceBalance attempts to address these challenges in optimizing parameters to fit complex condensed phase properties. A key aspect of this approach is that we calculate analytic derivatives of the simulated properties with respect to the force field parameters using a new fluctuation formula similar to Hamiltonian Gibbs−Duhem integration.[41] In the past, these derivatives have been evaluated by running multiple simulations with different parameter values,[40,41] but statistical error in finite difference gradients from independent simulations is a major problem. Here we recognize our properties of interest originate from averages and fluctuations in the isothermal−isobaric (NPT) ensemble. For instance, the ensemble average of a generic observable A that does not depend explicitly on the force field parameters (for example, the density or an order parameter) can be expressed as follows:

$$\langle A\rangle_\lambda = \frac{1}{Q(\lambda)}\int A(\mathbf{r}, V)\exp(-\beta(E(\mathbf{r}, V; \lambda) + PV))\,\mathrm{d}\mathbf{r}\,\mathrm{d}V$$

$$Q(\lambda) = \int \exp(-\beta(E(\mathbf{r}, V; \lambda) + PV))\,\mathrm{d}\mathbf{r}\,\mathrm{d}V \tag{5}$$

where A is the observable, $\mathbf{r}$ a given molecular configuration in a periodic simulation cell, λ the force field parameter, E the potential energy, $\beta \equiv (k_B T^{-1})$ the inverse temperature, k_B the Boltzmann constant, T the temperature, P the pressure, V the volume, Q the isothermal−isobaric partition function, and the angle brackets with a λ subscript represent an ensemble average in the thermodynamic ensemble of the force field parameterized by λ. In practice, this integral is evaluated numerically using molecular dynamics or Monte Carlo simulation in the NPT ensemble.

Since the expression for A depends on λ only through the potential energy E, we can differentiate eq 5 and obtain the analytic derivative (eq 6):

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\lambda}\langle A\rangle_\lambda &= \frac{1}{Q(\lambda)}\int A(\mathbf{r}, V)\exp(-\beta(E(\mathbf{r}, V; \lambda) + PV))\\ &\quad\left(-\beta\frac{\mathrm{d}E(\mathbf{r}, V)}{\mathrm{d}\lambda}\right)\mathrm{d}\mathbf{r}\,\mathrm{d}V - \frac{1}{Q(\lambda)^2}\frac{\mathrm{d}Q}{\mathrm{d}\lambda}\\ &\quad\int A(\mathbf{r}, V)\exp(-\beta(E(\mathbf{r}, V; \lambda) + PV))\,\mathrm{d}\mathbf{r}\,\mathrm{d}V\\ &= -\beta\left(\left\langle A\frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle_\lambda - \langle A\rangle_\lambda\left\langle\frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle_\lambda\right)\end{aligned} \tag{6}$$

In practice, these calculations require the potential energy derivative, which we evaluate by postprocessing the simulation

dx.doi.org/10.1021/jp403802c | J. Phys. Chem. B 2013, 117, 9956−9972

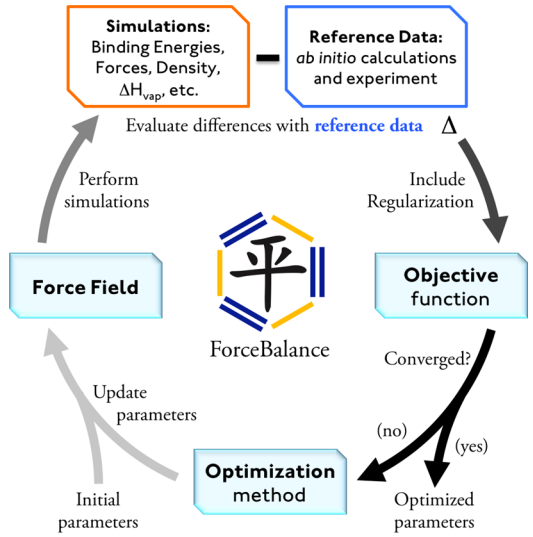


Figure 1. Illustration of ForceBalance procedure. The calculation begins with an initial set of parameters (lower left), which is used to generate a force field and perform simulations. The objective function is a weighted sum of squared differences between the simulation results and the reference data, plus a regularization term that prevents overfitting. The optimization method updates the parameters in order to minimize the objective function.

trajectory. This equation may be directly applied to obtain derivatives of ensemble-averaged observables that do not contain explicit force field parameter dependence, such as the density ρ.

By contrast, the derivative of the enthalpy contains an extra partial derivative term, because the potential energy appears in the integrand:

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\langle H\rangle = \left\langle \frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle - \beta\left(\left\langle H\frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle - \langle H\rangle\left\langle \frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle\right) \tag{7}$$

where for brevity we shall assume an implicit λ subscript for all quantities in all subsequent angle brackets. The derivative of the heat of vaporization ΔH_{vap} is given by subtracting the formulas for the liquid phase and gas phase enthalpies:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\lambda}\langle \Delta H_{\mathrm{vap}}\rangle &= \frac{\mathrm{d}}{\mathrm{d}\lambda}\left[(\langle E_{\mathrm{g}}\rangle + P\langle V_{\mathrm{g}}\rangle) - \frac{1}{N_{\mathrm{l}}}(\langle E_{\mathrm{l}}\rangle + P\langle V_{\mathrm{l}}\rangle)\right] \\ &= \left\langle \frac{\mathrm{d}E_{\mathrm{g}}}{\mathrm{d}\lambda}\right\rangle - \beta\left(\left\langle E_{\mathrm{g}}\frac{\mathrm{d}E_{\mathrm{g}}}{\mathrm{d}\lambda}\right\rangle - \langle E_{\mathrm{g}}\rangle\left\langle \frac{\mathrm{d}E_{\mathrm{g}}}{\mathrm{d}\lambda}\right\rangle\right) \\ &\quad - \frac{1}{N_{\mathrm{l}}}\left\langle \frac{\mathrm{d}E_{\mathrm{l}}}{\mathrm{d}\lambda}\right\rangle + \frac{\beta}{N_{\mathrm{l}}}\left(\left\langle E_{\mathrm{g}}\frac{\mathrm{d}E_{\mathrm{g}}}{\mathrm{d}\lambda}\right\rangle \right. \\ &\quad \left. - \langle E_{\mathrm{g}}\rangle\left\langle \frac{\mathrm{d}E_{\mathrm{g}}}{\mathrm{d}\lambda}\right\rangle\right)\end{aligned} \tag{8}$$

where $H \equiv E + PV$ is the enthalpy and the subscripts g and l denote the gas and liquid phases. Here, we have invoked the ideal gas law so that the molecular volume of water vapor does not appear in the expression.

We note the derivative of an ensemble average property resembles a fluctuation property or second-order correlation function; the above equation is derived in a similar manner to the fluctuation formulas for second-order thermodynamic properties like the thermal expansion coefficient, isothermal compressibility, etc. The derivative of nearly any ensemble average property can be evaluated in this fashion, and we also have derived formulas for differentiating fluctuation properties. For example, the derivatives of the thermodynamic fluctuation formulas for the thermal expansion coefficient α, isothermal compressibility κ_{T}, isobaric heat capacity c_P, and dielectric constant $\varepsilon(0)$ resemble third-order fluctuations (eqs 9−12):

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\lambda}\langle\alpha\rangle &= \frac{1}{k_{\mathrm{B}}T^2}\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{\langle HV\rangle - \langle H\rangle\langle V\rangle}{\langle V\rangle}\right) \\ &= \frac{1}{k_{\mathrm{B}}T^2}\left[-\beta\frac{\left\langle HV\frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle\langle V\rangle - \langle HV\rangle\left\langle V\frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle}{\langle V\rangle^2}\right. \\ &\quad + \frac{\left\langle V\frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle}{\langle V\rangle} - \left\langle \frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle + \beta\left(\left\langle H\frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle \right. \\ &\quad \left.\left. - \langle H\rangle\left\langle \frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle\right)\right]\end{aligned} \tag{9}$$

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\lambda}\langle\kappa_{\mathrm{T}}\rangle &= \beta\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{\langle V^2\rangle - \langle V\rangle^2}{\langle V\rangle}\right) \\ &= \beta^2\left(\frac{\langle V^2\rangle\left\langle V\frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle - \langle V\rangle\left\langle V^2\frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle}{\langle V\rangle^2} + \left\langle V\frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle \right. \\ &\quad \left. - \langle V\rangle\left\langle \frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle\right)\end{aligned} \tag{10}$$

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\lambda}\langle c_P\rangle &= \frac{1}{Nk_{\mathrm{B}}T}\frac{\mathrm{d}}{\mathrm{d}\lambda}(\langle H^2\rangle - \langle H\rangle^2) \\ &= \frac{1}{Nk_{\mathrm{B}}T}\left[2\left(\left\langle H\frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle - \langle H\rangle\left\langle \frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle\right)\right. \\ &\quad - \beta\left(\left\langle H^2\frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle - \langle H^2\rangle\left\langle \frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle\right) \\ &\quad \left. + 2\beta\langle H\rangle\left(\left\langle H\frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle - \langle H\rangle\left\langle \frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle\right)\right]\end{aligned} \tag{11}$$

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\lambda}\langle\varepsilon(0)\rangle &= \frac{\mathrm{d}}{\mathrm{d}\lambda}\left(1 + \frac{4\pi}{3k_{\mathrm{B}}T}\frac{\langle M^2\rangle}{\langle V\rangle}\right) \\ &= \frac{4\pi}{3k_{\mathrm{B}}T\langle V\rangle}\left[\sum_{i\in x,y,z}\left(\left\langle M_i\frac{\mathrm{d}M_i}{\mathrm{d}\lambda}\right\rangle - \langle M_i\rangle\left\langle \frac{\mathrm{d}M_i}{\mathrm{d}\lambda}\right\rangle\right.\right. \\ &\quad - \beta\left(\left\langle M_i{}^2\frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle - \langle M_i{}^2\rangle\left\langle \frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle\right. \\ &\quad \left.\left.\left. - 2\langle M_i\rangle\left\langle M_i\frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle + 2\langle M_i\rangle^2\left\langle \frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle\right)\right) \right. \\ &\quad \left. - \beta\frac{\langle M^2\rangle}{\langle V\rangle}\left(\left\langle V\frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle - \langle V\rangle\left\langle \frac{\mathrm{d}E}{\mathrm{d}\lambda}\right\rangle\right)\right]\end{aligned} \tag{12}$$

Here, M is the total dipole moment of the simulation cell and the summation goes over the three Cartesian axes. We remark

dx.doi.org/10.1021/jp403802c | J. Phys. Chem. B 2013, 117, 9956−9972

Table 2. Potential Parameters for AMOEBA (ref 45) and iAMOEBA (This Work)[a]

parameter name	units	AMOEBA	iAMOEBA	prior width
O−H equilibrium bond length	Å	0.9572	0.9584	0.1
O−H bond force constant	kcal/mol/Å^2	529.6	557.63	50
H−O−H equilibrium angle	deg	108.5	106.48	5
H−O−H angle force constant	kcal/mol/Å^2	34.05	49.90	40
H−H Urey−Bradley length	Å	1.5537	1.5357	N/A[b]
H−H Urey−Bradley force constant	kcal/mol/Å^2	38.25	−10.31	25
oxygen vdW sigma	Å	3.405	3.6453	0.3
oxygen vdW epsilon	kcal/mol	0.11	0.19682	0.1
hydrogen vdW sigma	Å	2.655	0	N/A[c]
hydrogen vdW epsilon	kcal/mol	0.0135	0	N/A[c]
hydrogen vdW reduction factor	none	0.91	0	N/A[c]
oxygen charge	e	−0.51966	−0.59402	0.4
oxygen dipole *Z*-component	e bohr	0.14279	0.08848	0.1
oxygen quadrupole *XX*-component	e bohr^2	0.37928	0.22618	0.2
oxygen quadrupole *YY*-component	e bohr^2	−0.41809	−0.32244	0.2
oxygen quadrupole *ZZ*-component	e bohr^2	0.03881	0.09626	0.2
hydrogen charge	e	0.25983	0.29701	0.4
hydrogen dipole *X*-component	e bohr	−0.03859	−0.09391	0.1
hydrogen dipole *Z*-component	e bohr	−0.05818	−0.12560	0.1
hydrogen quadrupole *XX*-component	e bohr^2	−0.03673	0.18754	0.2
hydrogen quadrupole *YY*-component	e bohr^2	−0.10739	0.02174	0.2
hydrogen quadrupole *XZ*-component	e bohr^2	−0.00203	−0.03635	0.2
hydrogen quadrupole *ZZ*-component	e bohr^2	0.14412	−0.20928	0.2
oxygen polarizability	Å^3	0.837	0.80636	0.1/0.2[d]
hydrogen polarizability	Å^3	0.496	0.50484	0.1/0.2[d]
polarization damping factor	Å^{-1}	0.39	0.23616	0.2

[a]The prior widths are used in regularization. [b]The Urey−Bradley equilibrium length is determined from the O−H bond length and H−O−H angle parameters using the law of cosines. [c]iAMOEBA does not contain hydrogen vdW interactions. [d]The polarizability parameters were optimized in terms of their sum ($\alpha_O + 2\alpha_H$, prior width 0.1 Å^3) and their ratio (α_O/α_H, prior width 0.2).

that, although we have avoided the issues associated with finite-difference derivatives from independent simulations, the derivative of an ensemble average property is intrinsically more difficult to estimate precisely, since it manifests as a fluctuation property, and the derivative of a fluctuation property manifests as a third-order fluctuation.

Finally, the different kinds of experimental and theoretical reference data used for parameterization have different physical units, so they require rescaling factors and weights in order to incorporate all data into an optimization scheme that minimizes a single objective function. The rescaling factor for each property is based on its intrinsic size and uncertainty from the simulation, and it is equal to the inverse weight of that property in the objective function. The values of the rescaling factors are given in Table 1.

Regularization. Although we included an unprecedented amount of reference data in our optimization, there is still the danger of overfitting. This arises because all of the simulated quantities emerge from the interactions in the model, which can easily have linear dependencies. Overfitting is treated by regularization, in which parameter values are penalized if they stray too far from their original values. Penalty functions have a natural interpretation in Bayesian statistics because they correspond to the negative logarithm of a *prior* distribution, analogous to how a potential energy function is the negative logarithm of a Boltzmann distribution.[94] For example, a harmonic penalty function corresponds to a Gaussian prior distribution:

$$P(\lambda) \propto e^{-\lambda^2/\alpha^2} \leftrightarrow R(\lambda) = \frac{\lambda^2}{\alpha^2} \tag{13}$$

where $P(\lambda)$ represents the prior probability distribution of the parameter λ and $R(\lambda)$ is the harmonic penalty function. The width of the prior distribution α (and corresponding inverse squared strength of the penalty function) represents our expectation of possible parameter values before introducing the parameterization data. The regularized objective function then corresponds to the *posterior* distribution.

The regularized optimization is more accurately described as performing maximum *a posteriori* estimation, instead of a formally Bayesian method which samples from the posterior distribution. Another important difference is that the "empirical Bayesian" methods treat the weights for reference data and the prior widths as nuisance parameters and samples over them, whereas we chose the weights and prior widths in this work by examining the physical scale and variability of each quantity. The reason for our approach is twofold. Given the high complexity of evaluating the objective function, a derivative-based optimization with predetermined weights and prior widths was the only feasible option with the current method and available resources. Furthermore, we found that perturbing the weights and prior widths by large amounts (∼50%) had minor effects on the behavior of the final model, but the same could not be said for perturbing force field parameters by a similar amount. In summary, the regularized optimization alleviates the task of manually selecting the highly sensitive force field parameters but still requires the researcher to

dx.doi.org/10.1021/jp403802c | J. Phys. Chem. B 2013, 117, 9956−9972

qualitatively specify the problem using his/her physical knowledge.

Our optimization was regularized using a Gaussian prior specified in Table 2. This corresponds to a parabolic penalty function in parameter space centered at the original AMOEBA parameter values with the chosen force constants. Since the various iAMOEBA parameters have different physical meanings (e.g., vdW well depth, O−H bond length), each parameter type was assigned its own prior width. We performed the optimization by first fitting only the theoretical data, because this was computationally less expensive. We then included the condensed phase properties and optimized the full objective function until fluctuations from numerical noise prevented further improvement. The full optimization converged to within the statistical error after about 10 nonlinear iterations, though we performed several optimizations with different choices of weights for the reference data and prior widths before arriving at the final answer.

In order to perform the simulations of condensed phase properties, ForceBalance interfaces with existing simulation software. To meet the high requirements for accuracy in the simulated properties, we used a combination of powerful and complementary methods in simulation software and distributed computing. The condensed phase simulations in the optimization used OpenMM 5.1,[95−97] a GPU-accelerated molecular dynamics software package with an extensively validated implementation of AMOEBA, which provides a speedup of an order of magnitude over the reference implementation in TINKER 6.1.[98] At each optimization step, the set of 42 simulations at different phase points (32 temperatures at 1.0 atm pressure plus 10 pressures at 298.15 K temperature, given in the Supporting Information) is performed simultaneously on multiple GPU clusters; the Work Queue library[99−101] allows ForceBalance to act as a distributed computing server and coordinate many OpenMM simulations running on multiple compute nodes in different physical locations. Finally, the data from the finished simulations was analyzed using the multistate Bennett acceptance ratio estimator (MBAR),[102,103] which allows each simulation to contribute to the estimated property of each other simulation. This combination of methods allowed us to optimize the condensed phase properties very accurately. Due to the nonoverlapping features of the simulation codes, we combined OpenMM 5.1 and TINKER 6.1 during the optimization to evaluate quantities for comparison with the *ab initio* and gas phase reference data, using OpenMM to evaluate the potential energies and forces and TINKER to evaluate the binding energies and monomer properties. ForceBalance,[51] TINKER,[98] OpenMM,[97] and Work Queue[101] are freely available on the Web.

RESULTS

A. iAMOEBA Parameters. Table 2 provides the optimized parameters for the iAMOEBA water model. Parameter files in the TINKER and OpenMM formats are provided on the Web.[64] The Gaussian prior widths in Table 2 are given to ForceBalance as inputs for the optimization, and they enter into the objective function as harmonic restraints in the parameter space. Our choices for the prior widths represent our expectations (from experience and physical intuition) that the parameters should not deviate from the initial AMOEBA values by more than these amounts, though ultimately the deviation in the parameter is also strongly affected by its importance in minimizing the objective function.

The first six rows contain the intramolecular parameters for iAMOEBA. The equilibrium bond length is essentially unchanged from the initial AMOEBA value. The force constants are changed more significantly; the original AMOEBA model reverses the symmetric and antisymmetric vibrational frequency order, and the iAMOEBA parameters restore the correct order (Table S3, Supporting Information). The equilibrium HOH angle parameter is reduced by 2°, which gives energies and atomistic forces in better agreement with *ab initio* calculations (Table S5, Supporting Information); however, as with all flexible models that lack intramolecular charge transfer, the angle contracts in the liquid phase relative to the gas phase value, in contrast to experimental measurements which indicate an expansion of the angle.[104,105] The liquid phase angle of iAMOEBA at room temperature and atmospheric pressure is 103° (standard error <1°), which is smaller than the value of 105−106° inferred from experiment[106] and *ab initio* molecular dynamics simulation.[107] The rationale for choosing an artificially large angle parameter of 108.5° in the original AMOEBA model was to reproduce the correct angle in the liquid phase, which aided in reproducing experimental liquid properties such as the dielectric constant; however, as we will see, iAMOEBA is able to accurately predict a broad range of condensed phase properties despite having a reduced bond angle value in the liquid phase.

The next 2 rows contain the vdW parameters, followed by 12 rows containing the permanent multipole parameters (only 9 of which are independent, due to the constraint of charge neutrality and use of traceless quadrupoles). iAMOEBA has a larger vdW radius and well depth compared to AMOEBA, which is largely due to the vdW interaction sites being removed from hydrogen (see the section on liquid structure for more details). The charges are increased in magnitude compared to AMOEBA, but the deviations in the higher-order electrostatic parameters show no clear pattern. The hydrogen quadrupole parameters have the most significant deviations with three out of four parameters changing sign from the initial AMOEBA values; this is probably related to alterations in the hydrogen interactions due to removing the vdW interactions, and/or the reduced angle in the condensed phase. Despite the large changes from the AMOEBA parameters, the iAMOEBA model gives very good agreement with the molecular multipole moments at the gas phase optimal geometry (Table S3, Supporting Information).

The polarizability parameters are perhaps the most interesting due to the direct polarization approximation. Direct polarization implies that atomic polarizabilities are additive (i.e., the molecular polarizability is equal to the sum of the atomic polarizabilities) and isotropic unless atomic polarizability tensors are used. While the atomic polarizabilities are additive with respect to an externally applied electric field, polarization effects become nonadditive for multipole interactions at short-range due to iAMOEBA's use of Thole damping. If the AMOEBA parameters were used to initialize the iAMOEBA parameterization, the initial molecular polarizability of iAMOEBA water would be 1.83 Å^3. Thus, we rescaled our initial parameters to 80.4% of the AMOEBA values, such that their sum was equal to the molecular polarizability of the AMOEBA model (1.47 Å^3). The optimization increased the polarizability parameters, resulting in a final value of 1.82 Å^3, an increase of 23%. This is in agreement with our intuition that the

dx.doi.org/10.1021/jp403802c | J. Phys. Chem. B 2013, 117, 9956−9972

iAMOEBA model should have a higher molecular polarizability than the AMOEBA model, because the entire electronic polarization must come from the permanent multipoles without any further mutual induction. To compensate for this stronger response, the Thole damping factor is significantly reduced from the value used by AMOEBA, resulting in increased damping at short-range.

B. Quality of Fit to Parameterization Data. In this section, we analyze the quality of fit that the iAMOEBA model obtains for the parameterization data listed in Table 1. Figure 2

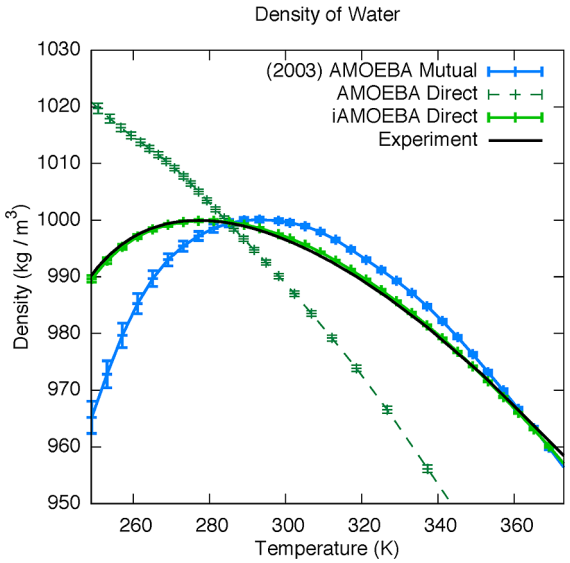


Figure 2. Density of liquid water over the temperature range 249−373 K at atmospheric pressure. The AMOEBA model from 2003 (blue line) uses mutual polarization. The direct polarization model gives a qualitatively incorrect result if the AMOEBA parameters are used (AMOEBA direct, dashed green line). Optimizing the parameters of the direct polarization model leads to iAMOEBA (solid green line), which reproduces the experimental density and TMD for the whole temperature range. Error bars represent one standard error.

shows the iAMOEBA, AMOEBA, and experimental density as a function of temperature over the temperature range 245−373 K at atmospheric pressure. The AMOEBA water model was only fit to the density and heat of vaporization at 25 °C but yields a temperature of maximum density (TMD) of 20 ± 2 °C, with the overall temperature dependence being approximately correct (blue line). If we use a direct polarization model with the AMOEBA parameters (i.e., turning off the mutual polarization), the result is qualitatively incorrect (dashed green line); the density decreases monotonically with temperature and the heat of vaporization is 5 kJ/mol too low (Supporting Figure S1, Supporting Information). Optimizing the parameters of the direct polarization iAMOEBA model using ForceBalance (solid green line) gives a TMD of 4 ± 2 °C that is closer to experiment compared to the original AMOEBA model, and the accuracy of the simulated density agrees with experiment to within 0.3% over the entire temperature range. Numerical values are given in Table 3.

We also used the second-order thermodynamic properties of water to parameterize iAMOEBA, which include the thermal expansion coefficient, isothermal compressibility, isobaric heat capacity, and dielectric constant. Their temperature dependence is shown in Supporting Figures S2−S5 (Supporting Information). At 25 °C, the simulated values for second-order properties agree with experiment to within 10%, but the deviations become slightly more significant at the temperature extremes. The isothermal compressibility is too low across the entire temperature range, which we were not able to improve further. We discovered an error in the calculation of the isobaric heat capacity during the optimization, so that the experimental heat capacity was not effectively targeted; after correcting the error, the simulated heat capacity is provided in Figure S4 (Supporting Information), and is on average slightly higher than experiment (by 0.3−0.6 kcal/mol) with the qualitatively correct temperature trend.

Figure 3 illustrates the quality of fit for the *ab initio* data for a representative water cluster from the liquid; the quality of fit for all *ab initio* potential energies and forces is given in Table S5 (Supporting Information). The MP2 and iAMOEBA forces are shown using blue and gold vectors, respectively. The strongest components of the force are often in the intramolecular degrees of freedom, which reflects the fact that forces tend to be stronger along degrees of freedom with higher force constants. This demonstrates the importance of explicitly looking at intermolecular forces, which we have done here by equally partitioning the force contribution into an atomistic, net force, and net torque component following ref 68.

Looking at all 2400 clusters for a specific cluster size (7-mers), the RMS error in relative potential energies (Figure 3, inset) is 4.6 kJ/mol, or 16% (standard error <1%) in comparison to the standard deviation of the reference potential energies (29.2 kJ/mol). The RMS error of the iAMOEBA atomistic forces relative to the MP2 values is 22 kJ/mol/Å, or 26% (standard error <1%) in comparison to the standard deviation of the reference atomistic forces (83 kJ/mol/Å). The errors in the intermolecular forces and torques are 9.4 kJ/mol/Å and 6.4 kJ/mol/rad, respectively, or 32% (standard error <1%) in comparison to the standard deviation of the reference net forces and torques (29 kJ/mol/Å and 20 kJ/mol/rad). By comparison, the AMOEBA model has atomistic/net force/net torque errors of 37%/31%/23%, and the flexible TIP3P model has much larger errors of 43%/61%/66%. This comparison indicates that polarizable force fields may be more appropriate for describing intermolecular interactions, since the pairwise-additive force fields are mainly designed to recapitulate bulk properties. The AMOEBA and iAMOEBA models perform comparably at predicting the intermolecular forces from this data set; although AMOEBA performs slightly better for torques, the force errors of 24−37% for both models are still somewhat large. We remark that it is possible to match the *ab initio* forces to 10−15% error if we had solely focused on force matching without any other reference data, but this could worsen the overall accuracy of the model for describing the condensed phase properties of water (see the Discussion section).

Supporting Figures S6 and S7 (Supporting Information) show the quality of fit for the binding energies of geometry-optimized clusters on a linear and log scale. iAMOEBA finds a 5.09 kcal/mol binding energy for the dimer, while AMOEBA yields 4.98 kcal/mol relative to the experimental value of 4.99 kcal/mol. For larger clusters, iAMOEBA systematically underestimates the binding energies, slightly more than AMOEBA. This indicates that mutual polarization likely plays an important role in obtaining a size-consistent description of cluster binding energies. We remark that AMOEBA does a very good job at predicting the *ab initio* binding energies for all of the clusters, despite having been parameterized using only a subset of the

dx.doi.org/10.1021/jp403802c | J. Phys. Chem. B 2013, 117, 9956−9972

Table 3. Experimental and Simulation Data from Different Water Models[a]

Property	Experiment	% tol.	TIP3P		TIP4P/2005		AMOEBA		iAMOEBA	
			Quantity	Score	Quantity	Score	Quantity	Score	Quantity	Score
Enthalpy of phase change / kcal mol^{-1}										
ΔH_{melt}	1.44	5	0.3	0	1.16	6	*		1.19	7
ΔH_{vap}	10.52	2.5	10.05	8	11.99	4	10.48	10	10.94	8
Critical point properties										
T_c/K	647.1	2.5	578	6	640	10	581 (2)	6	622	8
ρ_c/g cm^{-3}	0.322	2.5	0.272	4	0.337	8	0.334	9	0.333	9
p_c/bar	220.64	5	126	1	146	3	183	7	201	8
Surface tension / mN m^{-1}										
σ_{300K}	71.73	2.5	52.3	0	69.3	9	64 (4)	6	68.3 (11)	8
σ_{450K}	42.88	2.5	24.7	0	41.8	9	32 (2)	0	38.4 (7)	6
Melting properties										
T_m/K	273.15	2.5	146	0	252	7	*		261 (2)	8
ρ_{liq}/g cm^{-3}	0.999	0.5	1.017	6	0.993	9	*		0.999	10
ρ_{solid}/g cm^{-3}	0.917	0.5	0.947	3	0.921	9	*		0.929	7
dp/dT (bar K^{-1})	-137	5	-66	0	-135	10	*		-141	10
Orthobaric densities and temperature of maximum density (TMD)										
TMD/K	277	2.5	182	0	278	10	292 (2)	8	277 (1)	10
ρ_{298K}/g cm^{-3}	0.997	0.5	0.98	7	0.993	9	1.000	9	0.997	10
ρ_{400K}/g cm^{-3}	0.9375	0.5	0.868	0	0.93	8	0.928	8	0.934	9
ρ_{450K}/g cm^{-3}	0.8903	0.5	0.791	0	0.879	7	0.858	3	0.883	8
Isothermal compressibility / 10^{-6} bar^{-1}										
κ_T [1 bar; 298 K]	45.3	5	57.4	5	46	10	66 (1)	1	41.1 (4)	8
κ_T [1 bar; 360 K]	47	5	79.2	0	50.9	8	65 (1)	2	45.4 (4)	9
Gas properties										
p_v [350K] (bar)	0.417	5	0.56	3	0.13	0	0.43 (5)	9	0.39 (5)	9
p_v [450K] (bar)	9.32	5	11.72	5	4.46	0	17.3 (1)	0	9.8 (1)	9
B_2 [450K] (cm^3 mol^{-1})	-238	5	-476	0	-635	0	-214	8	-269	7
Heat capacity at constant pressure / cal mol^{-1} K^{-1}										
C_p [liq 298 K; 1 bar]	18	5	18.74	9	21.1	7	21.3	6	18.5 (2)	9
C_p [ice 250 K; 1 bar]	8.3	5	*	0	14	0	9.4 (7)	7	8.5 (1)	10
Static dielectric constant										
ε [liq; 298 K]	78.5	5	94	6	58	5	81.4 (14)	9	80.7 (11)	9
ε [I_h; 240 K]	107	5	19	0	53	0	*		96 (3)	8
Ratio	1.36	5	0.2	0	0.91	3	*		1.19 (4)	8
T_m-TMD-T_c ratios										
$T_m[I_h]/T_c$	0.422	5	0.253	2	0.394	9	*		0.416	10
TMD/T_c	0.428	5	0.315	5	0.434	10	0.503	6	0.445	9
TMD-T_m (K)	3.85	5	36	4	26	6	*		18	7
Densities of ice polymorphs / g cm^{-3}										
ρ[I_h 250 K; 1 bar]	0.92	0.5	*	0	0.921	10	0.894	4	0.930	8
ρ[II 123 K; 1 bar]	1.19	0.5	1.219	5	1.199	8	1.211	6	1.180	8
ρ[V 223 K; 5.3 kbar]	1.283	0.5	*	0	1.272	8	1.277	9	1.263	7
ρ[VI 225 K; 11 kbar]	1.373	0.5	1.366	9	1.38	9	1.384	8	1.364	9
EOS high pressure										
ρ[373 K; 10 kbar]	1.201	0.5	1.211	8	1.204	10	1.246	3	1.189	8
ρ[373 K; 20 kbar]	1.322	0.5	1.34	7	1.321	10	1.373	2	1.303	7
Self-diffusion coefficient / cm^2 s^{-1}										
ln D_{278K}	-11.24	0.5	-10.2	0	-11.27	9	-11.68	2	-11.15	8
ln D_{298K}	-10.68	0.5	-9.81	0	-10.79	8	-10.82	7	-10.58	8
ln D_{318K}	-10.24	0.5	-9.67	0	-10.39	7	-10.45	6	-10.19	9
E_a/kJ mol^{-1}	18.4	5	9.7	1	16.2	8	19.6	9	18.8	10
Shear viscosity / mPa s										
η[1 bar; 298 K]	0.896	5	0.321	0	0.855	9	1.08 (5)	6	0.85 (2)	9
η[1 bar; 373 K]	0.284	5	0.165	2	0.289	10	0.25 (2)	8	0.28 (2)	10
Orientational relaxation time / ps										
τ_2^{HH}[1 bar; 298 K]	2.36	5	0.8	0	2.3	9	1.94 (2)	6	2.41	10
Structure										
$\chi^2(F(Q))$	0	5	*	4	8.5	8	*		*	
χ^2(*overall*)	0	5	*	4	14.8	7	*		*	
Phase diagram	N/A	N/A		2		8				8
Overall score (out of 10)				2.6		7.1		5.9		8.4

[a]TIP3P and TIP4P/2005 data and percent tolerances for determining the numerical score (% tol. column) are reproduced from ref 39. All AMOEBA properties (except for melting properties) and iAMOEBA properties are calculated in this work. Cells with a green background indicate properties that are part of the parameterization data set; the green cells are excluded when calculating the average score for iAMOEBA. The liquid structure was not included in the score; see instead Figure 4. The phase diagram is scored qualitatively following ref 39.

dx.doi.org/10.1021/jp403802c | J. Phys. Chem. B 2013, 117, 9956−9972

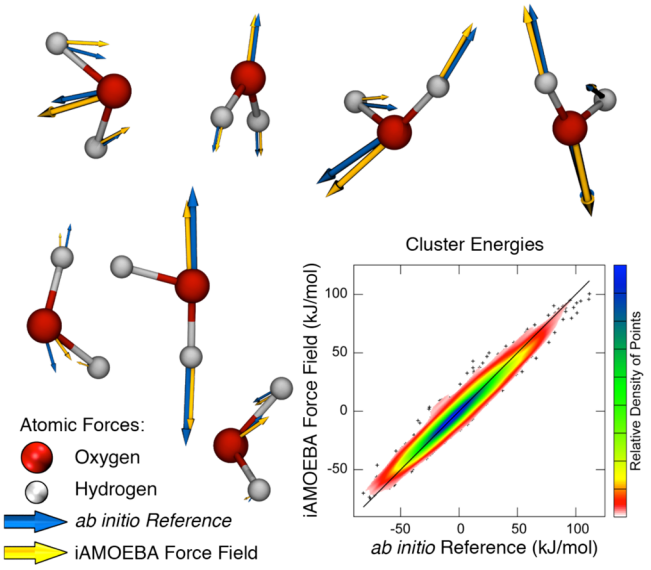


Figure 3. Illustration of iAMOEBA fit to theoretical data. The atomistic forces from the QM reference method (blue) and iAMOEBA model (yellow) are shown for a representative configuration of a water 7-mer extracted from the liquid. The relative RMS force error is 24% over all configurations. Inset: Scatter plot of cluster relative potential energies. Blue color indicates a higher density of points; the diagonal line indicates zero error.

dimer and hexamer conformations. iAMOEBA and AMOEBA are both able to accurately reproduce the optimal structures of the water clusters. The RMSDs to the reference QM-optimized structures are below 0.15 Å for all of the clusters in the parameterization data set; in particular, the RMSD of the iAMOEBA-minimized water dimer is 0.043 Å. The O−O distance of the water dimer is 2.836 Å in iAMOEBA, which is shorter than the values of 2.912 and 2.892 Å from the QM-optimized structure and AMOEBA. The flap angle between the O−O vector and the H−O−H bisector vector of the acceptor molecule is 54.9° from iAMOEBA, compared to 55.2° from the QM-optimized structure and 57.2° from AMOEBA.

While it is possible to obtain a more precise fit to the *ab initio* data, the utility of this data set is to ensure that the potential energy surface is qualitatively correct; the main goal is to accurately reproduce the experimental condensed phase properties. We return to this parameterization decision in the Discussion section.

C. Model Testing and Validation. The predictive power of the iAMOEBA model is tested using properties outside of the fitting set (Table 3, uncolored cells). All validation properties were computed using the TINKER 6.1 molecular dynamics software,[98] which provides a reference implementation of iAMOEBA. Here we include a comprehensive test set of properties taken from ref 36, which provides a scoring system on a 10-point scale derived from the accuracy of the model prediction relative to specified tolerance thresholds. The properties include thermodynamic and kinetic properties at a wide range of conditions, most of which are outside our parameterization data set; they cover the solid, liquid, and gas phases of water, focusing mainly on the liquid but also emphasizing reproducibility of the ice phases. The average of all property scores gives an overall score of the model from 0 to 10, which gives a general sense of the water model quality at a glance. To ensure a fair comparison, our average score excludes all properties that were used in the parameterization (highlighted in green in Table 3). We calculated an overall score of 8.4 out of 10 for iAMOEBA because of its excellent agreement for all of the tested properties. Furthermore, the liquid properties calculated using iAMOEBA compare favorably with a collection of well-established[42,44,113] and newly developed[108,112] polarizable water models in the literature (Table 4). The favorable comparison of iAMOEBA to existing water models with full self-consistent polarization gives credence to the idea that the direct polarization approximation may effectively capture polarization effects in water.

Liquid Phase Thermodynamic Properties. These properties include the density, thermal expansion coefficient, isothermal compressibility, and isobaric heat capacity; they are so categorized by their derivation from the thermodynamic ensemble as equilibrium averages and fluctuations from equilibrium. We expect iAMOEBA to reproduce these properties reasonably well because they were already included as part of the parameterization data set (though not at the same thermodynamic conditions). The density of liquid iAMOEBA water agrees with experiment to within 0.3% for thermodynamic conditions ranging from 249.15 to 373.15 K and up to 1 kbar. Small deviations of up to 1% are observed at very high

Table 4. Properties of Water Calculated Using Several Polarizable Models and Compared to Experimental Measurements[a]

property	experiment	AMOEBA	SWM4-NDP	TTM3-F	GCPM	SWM6	BK3	iAMOEBA
ρ (g cm^{-3})	0.997	1.000	0.994 (2)	0.994	1.007	0.996 (2)	0.9974 (2)	0.997
ΔH_{vap} (kcal mol^{-1})	10.52	10.48	10.44	11.4	11.30	10.52	10.94	10.94
α (10^{-4} K^{-1})	2.56	1.9 (6)			4.2		3.01 (8)	2.5 (1)
κ_T (10^{-6} bar^{-1})	45.3	66 (1)					44.4 (7)	41.1 (4)
C_p (cal mol^{-1} K^{-1})	18.0	21.3 (5)			22.5		22.0 (2)	18.5 (2)
$\varepsilon(0)$	78.5	81.4 (14)	78.0 (14)	67.7	84	78.1 (28)	79 (3)	80.7 (11)
D_0 (10^{-5} cm^2 s^{-1})	2.29	2.0	2.85 (28)	2.37	2.26	2.14 (19)	2.28 (4)	2.54 (2)
η (mPa s)	0.896	1.08 (5)	0.66 (9)			0.87 (12)	0.95 (1)	0.85 (2)
TMD (K)	277	292 (2)	<220		255	235	275 (3)	277 (1)
T_m (K)	273.15		<120	248 (2)			250 (3)	261 (2)
T_c (K)	647.1	581 (2)	576		642		629 (5)	622

[a]Liquid bulk properties are measured at 298 K, 1 bar; TMD and T_m are measured at 1 bar; and T_c is determined for the critical pressure of the model. Numbers in parentheses indicate one standard error in terms of the least significant digit, where available. AMOEBA and iAMOEBA properties were calculated in this work; standard errors are given in parentheses, except when the standard error is smaller than the number of significant figures given. SWM4-DP and SWM6 properties are from ref 108, except for the melting point[37] and critical point[109] of SWM4-DP. TTM3-F properties are from refs 44, 110, and 111. GCPM properties are from refs 112 and 113. BK3 properties are from ref 112.

dx.doi.org/10.1021/jp403802c | J. Phys. Chem. B 2013, 117, 9956−9972

temperatures (450 K) and pressures (10 kbar). The heat of vaporization is 10.94(1) kcal mol^{-1} at 298.15 K, which is higher than the experimental value of 10.52 kcal mol^{-1}; the temperature trend also has a larger slope compared to experiment. This insufficiency is due to well-known quantum nuclear effects on the heat capacity of the high-frequency vibrational modes, since at lower temperatures the faster vibrational degrees of freedom are frozen out.

The fluctuation properties also show good trends with experiment though they are less accurate, similar to what we found for the parameterization data set for these same properties. The isothermal compressibility is 10% too low at 298 K, and the error increases with lower temperatures; there exists a shallow minimum in the compressibility at 301−317 K, which is near the experimental value of 319 K. The simulated isobaric heat capacity agrees well with experiment at 298 K after including a quantum correction for the high frequency vibrational modes (see Table S4, Supporting Information).

Liquid Phase Kinetic Properties. Kinetic properties constitute an important test for iAMOEBA, particularly because no kinetic properties were included in the parameterization data set. We calculated the self-diffusion constant corrected for finite size effects,[114] the shear viscosity, and the orientational relaxation time at different temperatures. The self-diffusion constant for iAMOEBA at 298 K, 1 atm is 2.54(2) × 10^{-9} m^2 s^{-1} (experiment: 2.30 × 10^{-9} m^2 s^{-1}), and the temperature trend shows excellent agreement with experiment, as shown by the activation energy from the Arrhenius rate law (Supporting Figure S8, Supporting Information). The shear viscosity and orientational relaxation times also show very good agreement with experiment.

The infrared (IR) spectrum of water contains a wealth of information on the kinetic properties. In Figure 4, we show the calculated IR spectrum from iAMOEBA and the experimental IR spectrum on an arbitrary intensity scale. We also included the predictions from the TTM3-F model,[44] which was parameterized to reproduce the IR spectrum of water, and the flexible SPC/Fw model.[115] The spectra were generated by applying the quantum harmonic approximation to the Fourier-transformed dipole autocorrelation function of a classical MD trajectory. The peaks around 3500 and 1650 cm^{-1} correspond to vibrations of the O−H bond and the H−O−H angle, respectively; both TTM3-F and iAMOEBA predict the correct frequency shift from the gas phase values, but iAMOEBA and SPC/Fw do not predict the correct relative intensities. This is almost certainly due to the absence of intramolecular charge transfer[116] (also known as charge flux[53,117]) in the functional forms of iAMOEBA and SPC/Fw, which is required for a proper description of the dipole moment surface and which is included in the TTM3-F model. iAMOEBA and SPC/Fw both predict a slight splitting of the 3500 cm^{-1} peak corresponding to the symmetric and antisymmetric stretch frequencies; we postulate that including intramolecular charge transfer terms would broaden these peaks and lead to better agreement with experiment. Another promising route toward quantitative reproduction of the IR spectrum would be to account for nuclear quantum effects explicitly by reparameterizing the model for path integral molecular dynamics simulations.[118,119] The peaks below 1000 cm^{-1} correspond to librational and slower degrees of freedom. Here, both iAMOEBA and TTM3-F present a low-frequency shoulder below 300 cm^{-1}, whereas the SPC/Fw spectrum does not present this feature; this is thought to be a characteristic of polarizable water models, but more analysis is needed in order to establish this link.

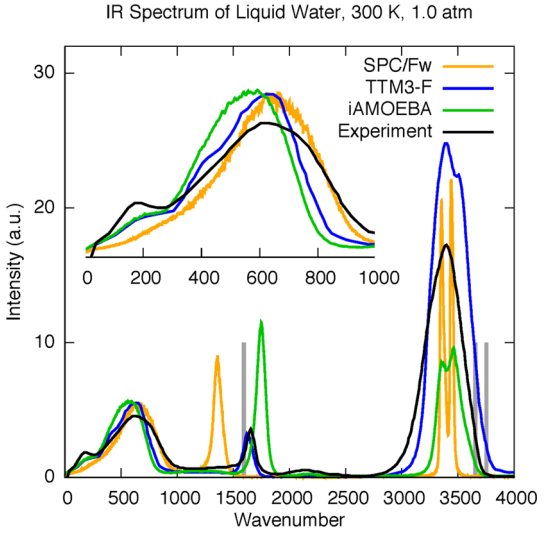


Figure 4. IR spectra of liquid water, measured using experiment and calculated using the SPC/Fw, TTM3-F, and iAMOEBA models. Experimental and TTM3-F data taken from ref 120. Gray bars represent gas phase vibrational frequencies from experiment. Inset: Magnification of the far IR region (<1000 wavenumber).

Liquid Structure. The partial radial distribution functions (RDFs) are an important indicator of liquid structure and connections to X-ray and neutron liquid diffraction.[28,121−124] For water, the oxygen−oxygen RDF, $gOO(r)$, is especially well captured in an X-ray scattering experiment due to oxygen's high atomic number relative to hydrogen. Figure 5, left panel, shows the $gOO(r)$ of water from iAMOEBA (green) along with experimentally derived RDFs from X-ray scattering data taken by Hura and co-workers at the Advanced Light Source[121] (ALS, black) and more recently by Skinner and co-workers at the Advanced Photon Source[125] (APS, orange). The agreement with experiment is good to within the distribution of past experimental derivations from X-ray and neutron diffraction data (Supporting Figure S9, Supporting Information). The O−H and H−H RDFs are provided in Supporting Figures S10 and S11 (Supporting Information) and compared to experimentally derived $gOH(r)$ and $gHH(r)$ from neutron scattering data.[123] Figure 5, right panel, compares the simulated X-ray scattering intensity from iAMOEBA and the experimental data from ALS and APS; the comparison is again favorable across all momentum transfer vectors, under assumptions of the electron density derived from the modified atomic form factors developed in ref 121. Both real-space and momentum-space comparisons indicate that iAMOEBA reproduces the structure of water at room temperature very accurately. Since we did not follow the exact procedure in ref 126 for quantifying the accuracy of the structure, this property was not included in the overall numerical score.

During the optimization, we monitored the $gOO(r)$ and X-ray intensity plot for agreement with experiment, though we did not explicitly include it in the optimization. In an intermediate stage of the optimization, we found that our model was performing poorly for the liquid structure prediction (blue dotted line) with significant deficiencies in the low-Q region of the X-ray intensity plot. After trying a number of modifications, we found that the agreement with experiment was completely recovered by eliminating the vdW interactions

dx.doi.org/10.1021/jp403802c | J. Phys. Chem. B 2013, 117, 9956−9972

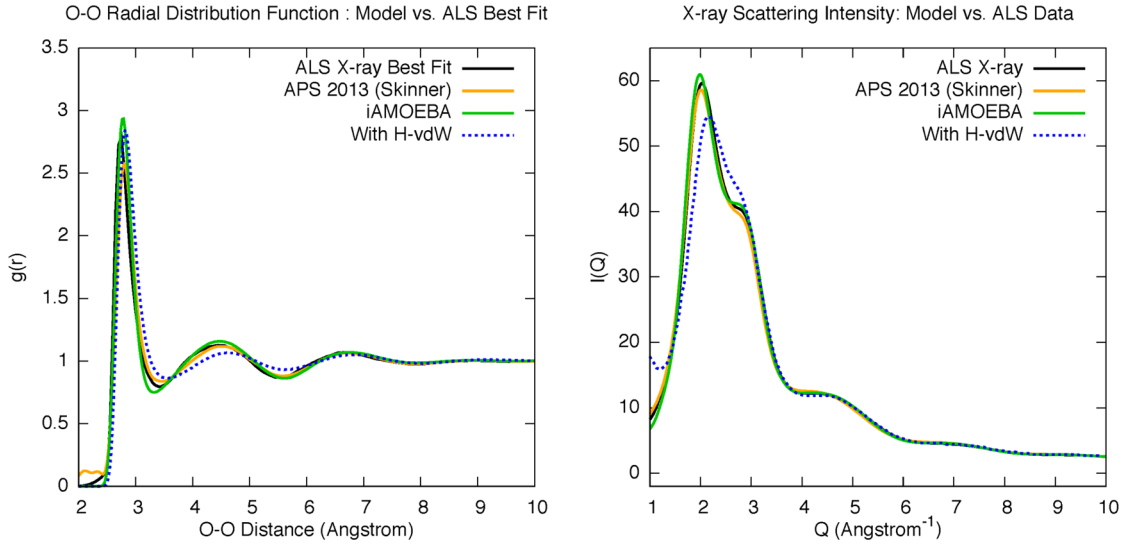


Figure 5. O−O radial distribution function (left) and X-ray scattering intensity (right) from ALS X-ray data (ref 121, black), APS X-ray data (ref 125, orange), the final iAMOEBA model (green), and an intermediate version of the model with hydrogen vdW interaction sites (blue dotted line). The model with hydrogen vdW interactions deviates significantly from experiment in the low-Q region of the intensity plot.

involving hydrogen; thus, the final iAMOEBA model has spherically symmetric vdW interactions centered on oxygen. Interestingly, while the choice of vdW functional form has a significant impact on the agreement with liquid structure, other properties were found to be in equally good agreement with experimental data whether hydrogen van der Waals interactions were included or not included.

Vapor and Critical Properties. The vapor and critical properties include the heat of vaporization, the vapor pressure, the surface tension, and the critical point; none of these properties were included in the optimization. These properties are very important tests for a polarizable model because the dipole moment of water molecules differs significantly between the vapor and liquid phases.[107,127,128] By contrast, nonpolarizable water models have a fixed molecular dipole moment that is parameterized to the liquid phase value in order to reasonably reproduce liquid properties. One consequence is that the fixed charge models overestimate the attraction between water molecules in the gas phase. These errors are demonstrated by the second virial coefficient and the vapor pressure, where nonpolarizable models significantly underestimate the experimental value.

We calculated the second virial coefficient (Table 3) and the liquid−vapor coexistence line of iAMOEBA up to the critical point (Supporting Figures S12−S13, Supporting Information). The calculated critical properties are in good agreement with experiment; the critical temperature and pressure are underestimated by a few percent. The vapor pressures and second virial coefficient also have good agreement with experiment. The accuracy of iAMOEBA exceeds AMOEBA in the neighborhood of the critical point; this is because iAMOEBA was explicitly optimized to reproduce densities at higher temperatures.[109] In general, the vapor and critical properties confirm that iAMOEBA with its direct polarization approximation is a highly viable model for water.

Ice, Melting Properties, and Phase Diagram. The melting point of ordinary ice (ice Ih) in the iAMOEBA model is calculated using direct coexistence simulation[38] to be 261 ± 2 K at 1 atm, 12° below the experimental value of 273.15 K (Figure 6). The enthalpy of melting is about 15% lower than the

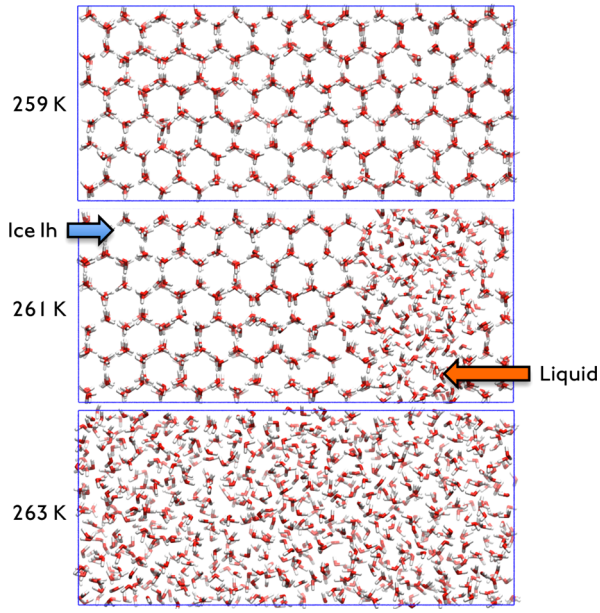


Figure 6. Determination of the iAMOEBA melting point by direct coexistence simulation. The initial configuration contains equal parts ice Ih and liquid water. Simulations run at temperatures below the freezing point of the model turned to ice (top), while simulations above the freezing point turned to liquid (bottom).

experimental value. The agreement with experiment is very good in the context of comparing with other models; however, there is a temperature gap of 16 K between the freezing point and the TMD (experimental gap 4 K). The overestimation of the TMD − T_m difference is a consistent trend across classical water models,[37−39] although promising results have been obtained using force-matching on configurations sampled from both water and ice.[129] Since isotope effects are known to significantly affect the freezing point of water, it remains to be seen whether nuclear quantum effects can provide a further improved description of this important property.

Water also forms many high-density ice polymorphs at high pressure; they constitute an interesting test for iAMOEBA

dx.doi.org/10.1021/jp403802c | J. Phys. Chem. B 2013, 117, 9956−9972

because we did not include any ice properties in the reference data. However, the diverse geometries of hydrogen bonding networks in different phases of ice show the versatility of water hydrogen bonding networks in different configurations, and we expect the *ab initio* calculations helped to serve as a guide in parameter space for describing these interactions correctly. The tests focus on the stable, crystalline, proton-disordered ice polymorphs; this rules out metastable ice phases (IV, XII) and extremely high-pressure phases (>100 kbar) where the individual water molecules dissociate. This leaves us with ices II, III, V, VI, and VII. iAMOEBA generally predicts the densities of the ice phases to be in generally good agreement with experiment, although there is a small systematic underestimation of roughly 1%.

Figure 7 shows the phase diagram of the iAMOEBA water model compared to experiment, which represents the first time

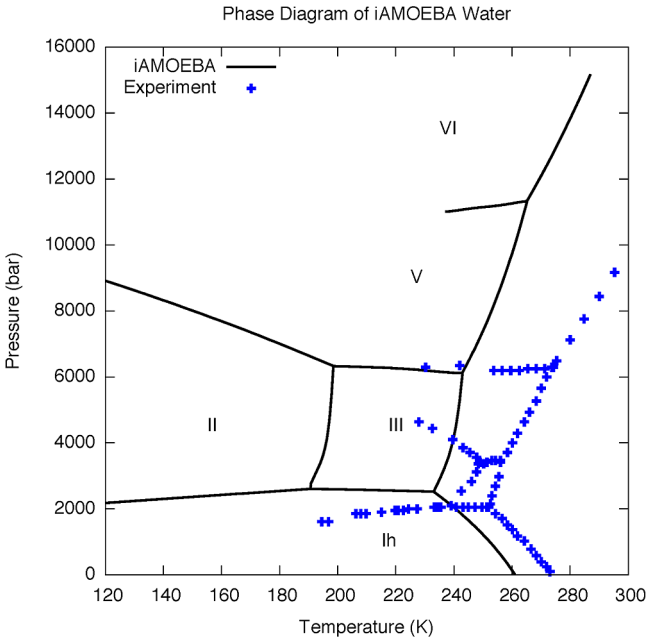


Figure 7. Phase diagram of the iAMOEBA water model compared to experiment (blue crosses). Experimental data adapted from ref 38.

the calculated phase diagram has been shown for any polarizable water model. The phase diagram shows the phase with the lowest free energy at different thermodynamic conditions, with each line on the diagram indicating a coexistence line between two phases. To calculate the phase diagram, we first determined the melting point of each phase of ice using direct coexistence simulations with the liquid[39] (Supporting Figure S14, Supporting Information). Following this, we numerically integrated the Clapeyron equation to provide the melting curves.[36,38,130] Two melting curves meet at a triple point, from which we integrated the Clapeyron equation to obtain the solid−solid coexistence lines; we continued to propagate solid−solid coexistence lines from triple points and arrived at the phase diagram shown in Figure 7. The qualitative structure of the phase diagram is correct, with ice phases Ih, II, III, V, and VI all appearing in the correct relative positions. This is a surprisingly difficult test for water models; to date, the only models that are known to reproduce the qualitative structure of the phase diagram are those that adopt the TIP4P functional form. By contrast, the TIP3P and TIP5P models predict qualitatively incorrect phase diagrams, in which ordinary ice only exists under large tensions (negative pressures).[38] In the low pressure region (<2500 bar), our calculated phase diagram is quantitatively accurate; the melting point is consistently underestimated by 12−15°, and the slope of the melting curve and the Ih−III transition pressure are correct to within 200 bar (10% of the total pressure). However, at higher pressures, the slopes of the melting curves are too high, and the solid−solid coexistence lines are also predicted to be too high. Following the scoring system in ref 36, we assigned 8 points out of 10 for the phase diagram, based on the observation that four phases of ice appeared in the correct relative positions.

Dielectric Properties. The dielectric constant should be a property for which polarizable models perform well due to their focus on properly describing the molecular electrostatics. Our simulated value for the dielectric constant agrees with experiment with a value of 81 ± 1 at 298 K (experiment: 78.3), and the trend is correct across the entire temperature range, although we overestimate the experimental value by 2−3. Polarizable models also allow us to calculate the dielectric constant of ice, where the water molecules are expected to have a larger dipole moment than the liquid. We calculated the dielectric constant of ice Ih using the electrostatic switching method[131] and found that iAMOEBA correctly predicts a higher dielectric constant in ice Ih compared to the liquid, although the ratio is a bit lower than the experimental value (Table 3); this is where full mutual polarization may be important. Nonpolarizable models, on the other hand, incorrectly predict the ratio to be less than 1, largely due to their fixed dipole moment.[38]

The gas phase dipole moment of the iAMOEBA water molecule is 1.86 D; the dipole increases to 2.78 D (standard error <0.01 D) in the liquid at 298 K with a full width at half-maximum (fwhm) of 0.25 D and increases even further to 2.90 D in ice at 240 K with a somewhat narrower fwhm (0.16 D). Although the molecular dipole moment in the condensed phase cannot be directly measured, literature estimates based on ab initio calculations[107,128,132] and experimental measurements[127,133] suggest values of 2.3−3.1 D, indicating that the direct polarization approximation correctly describes the average dipole moment in the various phases. Moreover, the iAMOEBA liquid phase dipole moment agrees very closely with AMOEBA, which also predicts a value of 2.78 D (standard error <0.01 D); this hints that the agreement may go beyond simple coincidence, especially given the important relationships between the liquid phase dipole moment and other, experimentally observable liquid properties.

■ DISCUSSION

Water models can be derived from experimental data, *ab initio* quantum chemistry, or a combination of the two. Since the condensed phase is an emergent property of the microscopic interactions, it is certainly desirable to derive a model entirely based on *ab initio* quantum chemistry, and indeed water models have been successfully developed following this approach.[52,115,129,134] The main drawback is that agreement with experimental condensed phase values is not guaranteed due to assumptions in the functional form of the classical model, integrating the equations of motion using Newtonian dynamics, approximations in the reference *ab initio* theory, or incomplete sampling of the thermodynamic ensemble of either the model or the reference theory.[135] In a compromise approach, the model parameters can be further modified to

dx.doi.org/10.1021/jp403802c | J. Phys. Chem. B 2013, 117, 9956−9972

match the experimental properties, but this must be done carefully in order to preserve agreement with the theoretical data, which may be representative of important molecular interactions when solutes are introduced. By contrast, water models that are solely fitted to a subset of experimental properties can have significant predictive power. Although most ab initio results are poorly predicted, some condensed phase properties are predicted more accurately than others, and the models can sometimes fail outside their parameterization range. Here, we addressed this challenge by including a wide range of experimental and *ab initio* reference data simultaneously using the ForceBalance method to optimize the iAMOEBA model.

We validated the iAMOEBA model using a scoring scheme developed by Vega and co-workers that evaluates the predictive power of a model against an experimental data set that covers thermodynamic and kinetic data over much of water's phase diagram. iAMOEBA based on the direct polarization approximation receives an overall score of 8.4 out of 10 because of its excellent agreement for all of the tested properties. We caution that a single numerical score is not sufficient to fully assess a model for a particular application; rather, the primary utility of the undertaking advocated in ref 36 is the wide range of properties investigated to provide a complete characterization of the iAMOEBA model.

Fixed charge pairwise additive models such as TIP3P earned a score of 2.7, while TIP4P/2005 earned a score of 7.2. TIP4P/2005, a model by Vega and co-workers using Ewald long-range electrostatics in a similar fashion to the TIP4P-Ew model, achieves remarkable accuracy for many properties (e.g., the critical temperature) and sacrifices accuracy for other properties (e.g., the critical pressure). One possible interpretation is that the properties for which TIP4P/2005 gives poor agreement are fundamental limitations of nonpolarizable models. Therefore, properties that are thought to require an explicit polarization treatment would include the dielectric constant of ice, the IR spectrum, the vapor pressure, and the critical pressure. The iAMOEBA model in fact is qualitatively superior to the fixed charge models on these properties, showing the importance of polarization.

While vapor properties are an important test of polarization effects where iAMOEBA performs well, these effects can have far-reaching implications beyond the gas phase. For instance, water molecules in biological settings such as protein active sites, membrane channels, and the coordination spheres of ions experience significantly different electric fields compared to the bulk liquid. While direct polarization clearly provides a better description compared to nonpolarizable models, its appropriateness for describing interactions with solute molecules, in particular ionic species, is an important open question; answering this question would require the development of force fields with direct and mutual polarization for these solute species with the parameterization conditions held constant.

CONCLUSION

In this work, we presented the iAMOEBA polarizable water model, which uses the direct polarization approximation. Compared to the AMOEBA model, iAMOEBA is simpler and computationally more efficient because it does not require a self-consistent solution for the induced dipoles. We confirmed the viability of the direct polarization approximation and established iAMOEBA as a highly accurate model with broad predictive power and applicability. We used the ForceBalance method to optimize the parameters, which allowed us to utilize a diverse data set with elements from experimental measurements and high-level theoretical calculations. The iAMOEBA model was tested against an extensive benchmark set of water properties and found to accurately describe many aspects of the vapor, liquid, and ice phases.

The improved accuracy of iAMOEBA over AMOEBA for many properties (Table 3) illustrates the advances in methodology and computation over the past 10 years; ForceBalance allows us to systematically optimize many parameters using a much larger and more diverse set of gas and condensed phase properties compared to previous methods. By combining GPU acceleration with distributed computing tools, we are able to perform simulations over a wide range of conditions rather than a single thermodynamic phase point. In the future, we postulate that kinetic properties could be differentiated in our optimization framework using a path integral approach in a similar manner to how thermodynamic properties are differentiated through the use of configuration integrals.

Our work has important implications for the fundamental understanding of molecular interactions in water and for the construction of molecular models in general. By building realistic but approximate models, we learn about which microscopic interactions are truly important for describing the properties of interest, and which other neglected interactions (in this case mutual polarization) can be effectively captured by other terms. For example, the main lesson from this work is that an appropriately parameterized direct polarization model is capable of capturing the effect of mutual polarization for a very wide range of water properties. Another major lesson is the role of model parameterization in establishing this understanding. For most practical problems, it is impossible to explore the entire parameter space, so we can only provide a lower bound on the accuracy of the empirical model using the most accurate parameters we can find. Thus, the model must be parameterized carefully before drawing conclusions about the *upper bound* of its descriptive capability. Thinking along these lines, the optimization approach outlined here can be easily applied to AMOEBA to further increase its accuracy for water properties and explore the aspects of mutual polarization which cannot be recovered through the direct approximation.

In terms of practical application, iAMOEBA is faster than AMOEBA (Table S2, Supporting Information) and does not compromise on the accuracy of water properties. The behavior of water at interfaces and in confined spaces is a highly active area of research where iAMOEBA can generate quantitatively accurate predictions at a reduced computational cost. Furthermore, the iAMOEBA water model lays the groundwork for analogous inexpensive polarizable models for solutes and biomolecules; while iAMOEBA is still more expensive than simple fixed point charge models (roughly 3−6 times the computational cost of TIP3P), it represents a significant increase in efficiency over AMOEBA. The accuracy of a future iAMOEBA model for solutes remains to be seen, but the success of the present work is encouraging. The possibility of developing a hybrid model using a combination of iAMOEBA water and mutually polarizable solutes is also promising, since most of the computational cost in typical biomolecular simulations comes from the solvent degrees of freedom.

■ ASSOCIATED CONTENT

Supporting Information

Computational details, further characterization of the iAMOEBA model, and experimental data used in parameterization. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: (650) 723-3660. Fax: (650) 725-0259. E-mail: pande@stanford.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are thankful for the assistance of Douglas Thain, Dinesh Rajan, and Li Yu in their dependable support and improvement of the Work Queue library. T.H.-G. thanks Martin Head-Gordon for insightful discussions that helped to inspire the idea for this project. L.-P.W. thanks William Swope, Julia Rice, and Sotiris Xantheas for helpful discussions in building the theoretical reference data set for this project. L.-P.W. thanks Carlos Vega for guidance in interpreting the solid−liquid coexistence simulations and for helpful discussions that led up to the phase diagram calculation. L.-P.W. thanks Ivan Ufimtsev and Lu Wang for assistance in calculating the IR spectrum. We also thank the anonymous referees for their constructive comments. L.-P.W. gratefully acknowledges support from the SimBios program, funded through the NIH Roadmap for Medical Research Grant U54 GM072970, and from the Department of Defense (Office of the Director of Defense Research and Engineering) through the National Security Science and Engineering Fellowship program. T.H.-G. thanks the NSF for support under grant CHE-1265731. T.H.-G. also acknowledges computational resources obtained under NSF grant CHE-1048789. J.W.P. would like to acknowledge support from NSF CHE-1152823 from the National Science Foundation Division of Chemistry. P.R. acknowledges support by the Robert A. Welch Foundation (F-1691). This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number OCI-1053575; the majority of GPU hours were provided by the *Stampede* supercomputer under XSEDE grant MCB100057. This research used resources from the Keeneland Computing Facility at the Georgia Institute of Technology, which is supported by the National Science Foundation under contract OCI-0910735. The authors acknowledge the following award for providing computing resources that have contributed to the research results reported within this paper: *MRI-R2: Acquisition of a Hybrid CPU/GPU and Visualization Cluster for Multidisciplinary Studies in Transport Physics with Uncertainty Quantification (NSF Award Number 0960306)*. This award is funded under the American Recovery and Reinvestment Act of 2009 (Public Law 111-5).

■ REFERENCES

(1) Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Shaw, D. E. *Science* **2011**, *334*, 517.

(2) Simmerling, C.; Strockbine, B.; Roitberg, A. E. *J. Am. Chem. Soc.* **2002**, *124*, 11258.

(3) Snow, C. D.; Nguyen, N.; Pande, V. S.; Gruebele, M. *Nature* **2002**, *420*, 102.

(4) Baumketner, A.; Shea, J.-E. *J. Mol. Biol.* **2006**, *362*, 567.

(5) Kelley, N. W.; Huang, X.; Tam, S.; Spiess, C.; Frydman, J.; Pande, V. S. *J. Mol. Biol.* **2009**, *388*, 919.

(6) Lin, Y.-S.; Bowman, G. R.; Beauchamp, K. A.; Pande, V. S. *Biophys. J.* **2012**, *102*, 315.

(7) Ball, K. A.; Phillips, A. H.; Nerenberg, P. S.; Fawzi, N. L.; Wemmer, D. E.; Head-Gordon, T. *Biochemistry* **2011**, *50*, 7612.

(8) Fawzi, N. L.; Phillips, A. H.; Ruscio, J. Z.; Doucleff, M.; Wemmer, D. E.; Head-Gordon, T. *J. Am. Chem. Soc.* **2008**, *130*, 6145.

(9) Gnanakaran, S.; Garcia, A. E. *Proteins: Struct., Funct., Bioinf.* **2005**, *59*, 773.

(10) Bizzarri, A. R.; Cannistraro, S. *J. Phys. Chem. B* **2002**, *106*, 6617.

(11) Fujitani, H.; Tanida, Y.; Ito, M.; Jayachandran, G.; Snow, C. D.; Shirts, M. R.; Sorin, E. J.; Pande, V. S. *J. Chem. Phys.* **2005**, *123*, 084108.

(12) Grubmuller, H.; Heymann, B.; Tavan, P. *Science* **1996**, *271*, 997.

(13) Pastor, R. W.; MacKerell, A. D., Jr. *J. Phys. Chem. Lett.* **2011**, *2*, 1526.

(14) Wang, L.; Berne, B. J.; Friesner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 1326.

(15) Clark, G. N. I.; Cappa, C. D.; Smith, J. D.; Saykally, R. J.; Head-Gordon, T. *Mol. Phys.* **2010**, *108*, 1415.

(16) Eaves, J. D.; Loparo, J. J.; Fecko, C. J.; Roberts, S. T.; Tokmakoff, A.; Geissler, P. L. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13019.

(17) Luzar, A.; Chandler, D. *Phys. Rev. Lett.* **1996**, *76*, 928.

(18) Rahman, A.; Stillinger, F. H. *J. Chem. Phys.* **1971**, *55*, 3336.

(19) Roberts, S. T.; Petersen, P. B.; Ramasesha, K.; Tokmakoff, A.; Ufimtsev, I. S.; Martinez, T. J. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 15154.

(20) Ufimtsev, I. S.; Kalinichev, A. G.; Martinez, T. J.; Kirkpatrick, R. J. *Chem. Phys. Lett.* **2007**, *442*, 128.

(21) Matsumoto, M.; Saito, S.; Ohmine, I. *Nature* **2002**, *416*, 409.

(22) Materer, N.; Starke, U.; Barbieri, A.; VanHove, M. A.; Somorjai, G. A.; Kroes, G. J.; Minot, C. *Surf. Sci.* **1997**, *381*, 190.

(23) Kumar, R.; Wang, F.-F.; Jenness, G. R.; Jordan, K. D. *J. Chem. Phys.* **2010**, *132*, 014309.

(24) Lambrecht, D. S.; Clark, G. N. I.; Head-Gordon, T.; Head-Gordon, M. *J. Phys. Chem. A* **2011**, *115*, 5928.

(25) Liu, K.; Cruzan, J. D.; Saykally, R. J. *Science* **1996**, *271*, 929.

(26) Xantheas, S. S.; Burnham, C. J.; Harrison, R. J. *J. Chem. Phys.* **2002**, *116*, 1493.

(27) Abascal, J. L. F.; Vega, C. *J. Chem. Phys.* **2010**, *133*, 234502.

(28) Clark, G. N. I.; Hura, G. L.; Teixeira, J.; Soper, A. K.; Head-Gordon, T. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 14003.

(29) Soper, A. K.; Teixeira, J.; Head-Gordon, T. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, E44.

(30) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712.

(31) Kollman, P. A. *Acc. Chem. Res.* **1996**, *29*, 461.

(32) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; shaw, D. E. *Proteins: Struct., Funct., Bioinf.* **2010**, *78*, 1950.

(33) Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157.

(34) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D., Jr. *J. Chem. Theory Comput.* **2012**, *8*, 3257.

(35) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.

(36) Vega, C.; Abascal, J. L. F. *Phys. Chem. Chem. Phys.* **2011**, *13*, 19663.

(37) Kiss, P. T.; Bertsyk, P.; Baranyai, A. *J. Chem. Phys.* **2012**, *137*, 194102.

(38) Vega, C.; Abascal, J. L. F.; Conde, M. M.; Aragones, J. L. *Faraday Discuss.* **2009**, *141*, 251.

(39) Vega, C.; Sanz, E.; Abascal, J. L. F. *J. Chem. Phys.* **2005**, *122*, 114507.

(40) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. *J. Chem. Phys.* **2004**, *120*, 9665.

(41) Abascal, J. L. F.; Vega, C. *J. Chem. Phys.* **2005**, *123*, 234505.

(42) Lamoureux, G.; MacKerell, A. D.; Roux, B. *J. Chem. Phys.* **2003**, *119*, 5185.
(43) Rick, S. W.; Stuart, S. J.; Berne, B. J. *J. Chem. Phys.* **1994**, *101*, 6141.
(44) Fanourgakis, G. S.; Xantheas, S. S. *J. Chem. Phys.* **2008**, *128*, 074506.
(45) Ren, P. Y.; Ponder, J. W. *J. Phys. Chem. B* **2003**, *107*, 5933.
(46) Ren, P. Y.; Ponder, J. W. *J. Phys. Chem. B* **2004**, *108*, 13427.
(47) Ferenczy, G. G.; Reynolds, C. A. *J. Phys. Chem. A* **2001**, *105*, 11470.
(48) Ren, P. Y.; Ponder, J. W. *J. Comput. Chem.* **2002**, *23*, 1497.
(49) Straatsma, T. P.; McCammon, J. A. *Mol. Simul.* **1990**, *5*, 181.
(50) Dykstra, C. E. *Chem. Rev.* **1993**, *93*, 2339.
(51) Wang, L.-P. ForceBalance: Systematic Force Field Optimization; accessed March 24, 2013. Available from: https://simtk.org/home/forcebalance/.
(52) Wang, L.-P.; Chen, J.; Van Voorhis, T. *J. Chem. Theory Comput.* **2013**, *9*, 452.
(53) Palmo, K.; Krimm, S. *Chem. Phys. Lett.* **2004**, *395*, 133.
(54) Stone, A. J. *The Theory of Intermolecular Forces*; Oxford University Press: New York, 1996.
(55) Toukmaji, A.; Sagui, C.; Board, J.; Darden, T. *J. Chem. Phys.* **2000**, *113*, 10913.
(56) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471.
(57) Lamoureux, G.; Roux, B. *J. Chem. Phys.* **2003**, *119*, 3025.
(58) Vanbelle, D.; Froeyen, M.; Lippens, G.; Wodak, S. J. *Mol. Phys.* **1992**, *77*, 239.
(59) Genzer, J.; Kolafa, J. *J. Mol. Liq.* **2004**, *109*, 63.
(60) Kuo, I. F. W.; Mundy, C. J.; McGrath, M. J.; Siepmann, J. I. *J. Chem. Theory Comput.* **2006**, *2*, 1274.
(61) Ong, S. W.; Tok, E. S.; Kang, H. C. *Phys. Chem. Chem. Phys.* **2010**, *12*, 14960.
(62) Wagner, W.; Pruss, A. *J. Phys. Chem. Ref. Data* **2002**, *31*, 387.
(63) Kell, G. S. *J. Chem. Eng. Data* **1975**, *20*, 97.
(64) Wang, L.-P. iAMOEBA parameter file and data repository; accessed March 24, 2013. Available from: https://simtk.org/home/iamoeba.
(65) Banks, J. L.; Kaminski, G. A.; Zhou, R. H.; Mainz, D. T.; Berne, B. J.; Friesner, R. A. *J. Chem. Phys.* **1999**, *110*, 741.
(66) Brommer, P.; Gaehler, F. *Modell. Simul. Mater. Sci. Eng.* **2007**, *15*, 295.
(67) Liu, Y. P.; Kim, K.; Berne, B. J.; Friesner, R. A.; Rick, S. W. *J. Chem. Phys.* **1998**, *108*, 4739.
(68) Akin-Ojo, O.; Song, Y.; Wang, F. *J. Chem. Phys.* **2008**, *129*, 4739−4755.
(69) Ercolessi, F.; Adams, J. B. *Europhys. Lett.* **1994**, *26*, 583.
(70) Izvekov, S.; Parrinello, M.; Burnham, C. J.; Voth, G. A. *J. Chem. Phys.* **2004**, *120*, 10896.
(71) Shell, M. S. *J. Chem. Phys.* **2008**, *129*, 144108.
(72) Wu, D.; Kofke, D. A. *J. Chem. Phys.* **2005**, *123*, 054103.
(73) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry*; Dover: Mineola, NY, USA, 1996.
(74) Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; Oxford University Press: 1989.
(75) Distasio, R. A., Jr.; Steele, R. P.; Head-Gordon, M. *Mol. Phys.* **2007**, *105*, 2731.
(76) Steele, R. P.; DiStasio, R. A., Jr.; Head-Gordon, M. *J. Chem. Theory Comput.* **2009**, *5*, 1560.
(77) Steele, R. P.; DiStasio, R. A.; Shao, Y.; Kong, J.; Head-Gordon, M. *J. Chem. Phys.* **2006**, *125*, 074108.
(78) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.
(79) Shao, Y.; Molnar, L. F.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S. T.; Gilbert, A. T. B.; Slipchenko, L. V.; Levchenko, S. V.; O'Neill, D. P.; DiStasio, R. A., Jr.; Lochan, R. C.; Wang, T.; Beran, G. J. O.; Besley, N. A.; Herbert, J. M.; Lin, C. Y.; Van Voorhis, T.; Chien, S. H.; Sodt, A.; Steele, R. P.; Rassolov, V. A.; Maslen, P. E.; Korambath, P. P.; Adamson, R. D.; Austin, B.; Baker, J.; Byrd, E. F. C.; Dachsel, H.; Doerksen, R. J.; Dreuw, A.; Dunietz, B. D.; Dutoi, A. D.; Furlani, T. R.; Gwaltney, S. R.; Heyden, A.; Hirata, S.; Hsu, C.-P.; Kedziora, G.; Khalliulin, R. Z.; Klunzinger, P.; Lee, A. M.; Lee, M. S.; Liang, W.; Lotan, I.; Nair, N.; Peters, B.; Proynov, E. I.; Pieniazek, P. A.; Rhee, Y. M.; Ritchie, J.; Rosta, E.; Sherrill, C. D.; Simmonett, A. C.; Subotnik, J. E.; Woodcock, H. L., III; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, F. J.; Warshel, A.; Hehre, W. J.; Schaefer, H. F., III; Kong, J.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172.
(80) Yoo, S.; Apra, E.; Zeng, X. C.; Xantheas, S. S. *J. Phys. Chem. Lett.* **2010**, *1*, 3122.
(81) Bates, D. M.; Tschumper, G. S. *J. Phys. Chem. A* **2009**, *113*, 3555.
(82) Bulusu, S.; Yoo, S.; Apra, E.; Xantheas, S.; Zeng, X. C. *J. Phys. Chem. A* **2006**, *110*, 11781.
(83) Fanourgakis, G. S.; Apra, E.; Xantheas, S. S. *J. Chem. Phys.* **2004**, *121*, 2655.
(84) Lagutschenkov, A.; Fanourgakis, G. S.; Niedner-Schatteburg, G.; Xantheas, S. S. *J. Chem. Phys.* **2005**, *122*, 194340.
(85) Xantheas, S. S.; Apra, E. *J. Chem. Phys.* **2004**, *120*, 823.
(86) Lifson, S.; Warshel, A. *J. Chem. Phys.* **1968**, *49*, 5116.
(87) Hagler, A. T.; Huler, E.; Lifson, S. *J. Am. Chem. Soc.* **1974**, *96*, 5319.
(88) Maple, J. R.; Hwang, M. J.; Stockfisch, T. P.; Dinur, U.; Waldman, M.; Ewig, C. S.; Hagler, A. T. *J. Comput. Chem.* **1994**, *15*, 162.
(89) Williams, D. E. *Top. Curr. Phys.* **1981**, *26*, 3.
(90) Levenberg, K. *Q. Appl. Math.* **1944**, *2*, 164.
(91) Marquardt, D. W. *J. Soc. Ind. Appl. Math.* **1963**, *11*, 431.
(92) Dennis, J. E.; Gay, D. M.; Welsch, R. E. *ACM Trans. Math. Software* **1981**, *7*, 348.
(93) More, J. J.; Sorensen, D. C. *SIAM J. Sci. Stat. Comput.* **1983**, *4*, 553.
(94) Liu, P.; Shi, Q.; Daume, H.; Voth, G. A. *J. Chem. Phys.* **2008**, *129*, 214114.
(95) Eastman, P.; Friedrichs, M. S.; Chodera, J. D.; Radmer, R. J.; Bruns, C. M.; Ku, J. P.; Beauchamp, K. A.; Lane, T. J.; Wang, L.-P.; Shukla, D.; Tye, T.; Houston, M.; Stich, T.; Klein, C.; Shirts, M. R.; Pande, V. S. *J. Chem. Theory Comput.* **2013**, *9*, 461.
(96) Eastman, P.; Pande, V. S. *Comput. Sci. Eng.* **2010**, *12*, 34.
(97) Pande, V. S.; Eastman, P. OpenMM; accessed March 24, 2013. Available from: https://simtk.org/home/openmm.
(98) Ponder, J. W. TINKER Molecular Modeling Package; accessed March 24, 2013. Available from: http://dasher.wustl.edu/ffe/.
(99) Abdul-Wahid, B.; Yu, L.; Rajan, D.; Feng, H.; Darve, E.; Thain, D.; Izaguirre, J. A. IEEE International Conference on e-Science, 2012.
(100) Bui, P.; Rajan, D.; Abdul-Wahid, B.; Izaguirre, J. A.; Thain, D. Workshop on Python for High Performance and Scientific Computing (PyHPC), 2011.
(101) Thain, D. Work Queue: A Scalable Master/Worker Framework; accessed March 24, 2013. Available from: http://www3.nd.edu/~ccl/software/workqueue/.
(102) Shirts, M. R.; Chodera, J. D. *J. Chem. Phys.* **2008**, *129*, 124105.
(103) Shirts, M. R.; Chodera, J. D. A Python implementation of the multistate Bennett acceptance ratio (MBAR); accessed March 24, 2013. Available from: https://simtk.org/home/pymbar.
(104) Burnham, C. J.; Xantheas, S. S. *J. Chem. Phys.* **2002**, *116*, 5115.
(105) Fanourgakis, G. S.; Xantheas, S. S. *J. Chem. Phys.* **2006**, *124*, 174504.
(106) Ichikawa, K.; Kameda, Y.; Yamaguchi, T.; Wakita, H.; Misawa, M. *Mol. Phys.* **1991**, *73*, 79.
(107) Silvestrelli, P. L.; Parrinello, M. *Phys. Rev. Lett.* **1999**, *82*, 3308.
(108) Yu, W.; Lopes, P. E. M.; Roux, B.; MacKerell, A. D. *J. Chem. Phys.* **2013**, *138*, 034508.
(109) Kiss, P. T.; Baranyai, A. *J. Chem. Phys.* **2012**, *137*, 194103.
(110) Imoto, S.; Xantheas, S. S.; Saito, S. *J. Chem. Phys.* **2013**, *138*, 054506.
(111) Yoo, S.; Xantheas, S. S. *J. Chem. Phys.* **2011**, *134*, 121105.
(112) Kiss, P. T.; Baranyai, A. *J. Chem. Phys.* **2013**, *138*, 204507.
(113) Paricaud, P.; Predota, M.; Chialvo, A. A.; Cummings, P. T. *J. Chem. Phys.* **2005**, *122*, 244511.

(114) Yeh, I. C.; Hummer, G. *J. Phys. Chem. B* **2004**, *108*, 15873.
(115) Wu, Y. J.; Tepper, H. L.; Voth, G. A. *J. Chem. Phys.* **2006**, *124*, 024503.
(116) Chen, J.; Martinez, T. J. *Chem. Phys. Lett.* **2007**, *438*, 315.
(117) Mannfors, B.; Palmo, K.; Krimm, S. *J. Phys. Chem. A* **2008**, *112*, 12667.
(118) Liu, J.; Miller, W. H.; Paesani, F.; Zhang, W.; Case, D. A. *J. Chem. Phys.* **2009**, *131*, 164509.
(119) Paesani, F.; Iuchi, S.; Voth, G. A. *J. Chem. Phys.* **2007**, *127*, 074506.
(120) Liu, J.; Miller, W. H.; Fanourgakis, G. S.; Xantheas, S. S.; Imoto, S.; Saito, S. *J. Chem. Phys.* **2011**, *135*, 244503.
(121) Hura, G.; Sorenson, J. M.; Glaeser, R. M.; Head-Gordon, T. *J. Chem. Phys.* **2000**, *113*, 9140.
(122) Soper, A. K. *Chem. Phys.* **2000**, *258*, 121.
(123) Soper, A. K. *J. Phys.: Condens. Matter* **2007**, *19*, 335206.
(124) Soper, A. K.; Phillips, M. G. *Chem. Phys.* **1986**, *107*, 47.
(125) Skinner, L. B.; Huang, C.; Schlesinger, D.; Pettersson, L. G. M.; Nilsson, A.; Benmore, C. J. *J. Chem. Phys.* **2013**, *138*, 074506.
(126) Pusztai, L.; Pizio, O.; Sokolowski, S. *J. Chem. Phys.* **2008**, *129*, 184103.
(127) Badyal, Y. S.; Saboungi, M. L.; Price, D. L.; Shastri, S. D.; Haeffner, D. R.; Soper, A. K. *J. Chem. Phys.* **2000**, *112*, 9206.
(128) Gregory, J. K.; Clary, D. C.; Liu, K.; Brown, M. G.; Saykally, R. J. *Science* **1997**, *275*, 814.
(129) Pinnick, E. R.; Erramilli, S.; Wang, F. *J. Chem. Phys.* **2012**, *137*, 014510.
(130) Habershon, S.; Manolopoulos, D. E. *Phys. Chem. Chem. Phys.* **2011**, *13*, 19714.
(131) Lindberg, G. E.; Wang, F. *J. Phys. Chem. B* **2008**, *112*, 6436.
(132) Batista, E. R.; Xantheas, S. S.; Jonsson, H. *J. Chem. Phys.* **1999**, *111*, 6011.
(133) Gubskaya, A. V.; Kusalik, P. G. *J. Chem. Phys.* **2002**, *117*, 5290.
(134) Akin-Ojo, O.; Wang, F. *J. Comput. Chem.* **2011**, *32*, 453.
(135) Wang, L.-P.; Van Voorhis, T. *J. Chem. Phys.* **2010**, *133*, 231101.