

## United polarizable multipole water model for molecular mechanics simulation

Rui Qi,<sup>1,a)</sup> Lee-Ping Wang,<sup>2,a)</sup> Qiantao Wang,<sup>1</sup> Vijay S. Pande,<sup>2</sup> and Pengyu Ren<sup>1,b)</sup>

<sup>1</sup>Department of Biomedical Engineering, The University of Texas at Austin, Austin, Texas 78712, USA

<sup>2</sup>Department of Chemistry, Stanford University, Stanford, California 94305, USA

(Received 5 February 2015; accepted 21 June 2015; published online 7 July 2015)

We report the development of a united AMOEBA (uAMOEBA) polarizable water model, which is computationally 3–5 times more efficient than the three-site AMOEBA03 model in molecular dynamics simulations while providing comparable accuracy for gas-phase and liquid properties. In this coarse-grained polarizable water model, both electrostatic (permanent and induced) and van der Waals representations have been reduced to a single site located at the oxygen atom. The permanent charge distribution is described via the molecular dipole and quadrupole moments and the many-body polarization via an isotropic molecular polarizability, all located at the oxygen center. Similarly, a single van der Waals interaction site is used for each water molecule. Hydrogen atoms are retained only for the purpose of defining local frames for the molecular multipole moments and intramolecular vibrational modes. The parameters have been derived based on a combination of *ab initio* quantum mechanical and experimental data set containing gas-phase cluster structures and energies, and liquid thermodynamic properties. For validation, additional properties including dimer interaction energy, liquid structures, self-diffusion coefficient, and shear viscosity have been evaluated. The results demonstrate good transferability from the gas to the liquid phase over a wide range of temperatures, and from nonpolar to polar environments, due to the presence of molecular polarizability. The water coordination, hydrogen-bonding structure, and dynamic properties given by uAMOEBA are similar to those derived from the all-atom AMOEBA03 model and experiments. Thus, the current model is an accurate and efficient alternative for modeling water. © 2015 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4923338>]

### INTRODUCTION

Water is an important solvent in living systems<sup>1,2</sup> and many industrial applications.<sup>3–7</sup> There are a number of molecular mechanics models, such as the three-site TIPS,<sup>8</sup> TIP3P,<sup>9</sup> SPC,<sup>10</sup> and SPC/E,<sup>11</sup> the four-site TIP4P,<sup>9</sup> TIP4P-Ew,<sup>12</sup> TIP4P/2005,<sup>13</sup> and the five site ST2,<sup>14</sup> TIP5P,<sup>15</sup> and TIP5P-E,<sup>16</sup> that are commonly used in molecular simulation of water. These models use fixed atomic partial charges, with electrostatic energy evaluated in pairwise-additive fashion. Some models incorporate explicit electronic polarization to allow the charge distribution to respond to electrostatic environment and to further improve the reproduction of many water properties; these include Dang-Chang,<sup>17</sup> Thole-Type-Model (TTM),<sup>18–22</sup> SWM-4DP,<sup>23</sup> DPP2,<sup>24</sup> and AMOEBA03.<sup>25</sup> The AMOEBA03 water model<sup>26</sup> was developed with a focus on capturing molecular polarizability, electrostatic potential, as well as the interaction energy from gas to condensed-phase, by utilizing permanent atomic monopole, dipole, and quadrupole moments and mutual atomic dipole-dipole induction.<sup>27,28</sup>

The inexpensive AMOEBA (iAMOEBA) model was introduced recently as a way to achieve improved computational efficiency.<sup>29</sup> In this model, the induced atomic dipoles

are determined directly from the permanent multipole electric fields without further interactions between induced dipoles. Thus, it captures some 3-body effects in polarization while reducing the computational cost relative to the fully self-consistent AMOEBA03 water model by a factor of 1.5–6. An alternative to improve the computational efficiency, without sacrificing the many-body effect, is to reduce the number of interaction sites within the model. Water models with single dipole moment representation have been developed since the 1980s.<sup>30</sup> Ichiye and her coworkers<sup>31,32</sup> introduced the “soft sticky dipole potential” for liquid water, with one spherical repulsive potential, a short-range tetrahedral “sticky” potential, and a point dipole at the center of mass. Compared with TIP3P and TIP4P models, it produced similar liquid water properties but with up to one order of magnitude speed-up. Later in 2010, the “soft-sticky dipole-quadrupole-octupole” water model was presented by the same group. It was suggested that the addition of octupole moments improved the dielectric constant (75 at 298 K).<sup>33</sup> Carnie and Patey reported a polarizable dipole-tetrahedral quadrupole water model and a self-consistent mean field theory was applied to account for molecular polarizability.<sup>34</sup> Later, Kusalik and Patey added the octupole moments to their water model discussed above and observed strong preferential solvation of anions at infinite dilution, suggesting an important role of octupole moments in ion solvation.<sup>35</sup> Jonsson *et al.*<sup>36</sup> introduced a one-site

<sup>a)</sup>R. Qi and L.-P. Wang contributed equally to this work.

<sup>b)</sup>Author to whom correspondence should be addressed. Electronic mail: [pren@mail.utexas.edu](mailto:pren@mail.utexas.edu). Telephone: (512) 567-1468.

water potential based on electrostatic, induction, dispersion, and short-range repulsion interactions. Also, multipoles up to quadrupole moment for polarizability and up to hexadecapoles for permanent electrostatics were included. Previously, we have also explored one-site non-polarizable models, based on permanent molecular multipoles and the Gay-Berne potential, for molecular liquids including water.<sup>37,38</sup> Molinero and Moore proposed a water model (mM) that further omitted long-range intermolecular interactions. This model generally reproduces the bulk properties of liquid water—except for the self-diffusion coefficient, which is too fast.<sup>39</sup> The faster diffusion rate is likely due to the lack of long-ranged electrostatic forces in the model.

In this paper, we describe a new (nonbonded) one-site polarizable water model, united AMOEBA (uAMOEBA). This model removes two nonbonded interaction sites on H atoms to speed up the energy and force calculations by a factor of 3 to 5 times over the previous three-site AMOEBA03 model in molecular dynamics when particle-mesh Ewald (PME) is used to treat long-range electrostatics.<sup>40</sup> Importantly, the full many-body polarization effects are retained via mutual induction of molecular dipoles. In this model, the permanent molecular multipole (dipole and quadrupole) moments, isotropic molecular dipole polarizability, and a single van der Waals (vdW) interaction site are placed on oxygen. The remaining hydrogen atoms only carry atomic masses to define the local coordinate frames and the intramolecular geometry. For computational efficiency, uAMOEBA does not contain any octupole moment, which is a potential limitation of this model. The hydrogen atoms experience forces due to valence (bond, angle, and Urey-Bradley) interactions. In addition, the torques experienced by the molecular dipole and quadrupole moments on oxygen are translated into forces on the hydrogen atoms in the same way as all-atom AMOEBA03.<sup>25</sup> The use of a single scalar polarizability is well justified, as the three components of water molecular polarizability (1.528, 1.415, 1.468) are indeed similar in magnitude. We demonstrate the accuracy of uAMOEBA water model by evaluating a range of gas-phase and liquid properties and comparing the results with the existing AMOEBA03/iAMOEBA models and experimental data. The current work aims to extend beyond the previous one-site water models by systematically examining gas-phase clusters of increasing size and important thermodynamic properties at a wide range of temperatures. In addition, the uAMOEBA model is applied to study interactions between water and other common organic molecules (modeled by all-atom AMOEBA03<sup>41</sup>) to investigate its transferability.

## METHODOLOGY

### The AMOEBA03 polarizable force field

The previously published AMOEBA03 functional form<sup>42</sup> has been applied in this study, except that each water molecule has a single electrostatic and vdW interaction site. The AMOEBA03 potential energy is written as a sum of valence and nonbonded contributions,

$$E = E_{\text{angle}} + E_{\text{bond}} + E_{b\theta} + E_{\text{vdw}} + E_{\text{ele}}^{\text{perm}} + E_{\text{ele}}^{\text{ind}}, \quad (1)$$

$$E_{ij}^{\text{buff}} = \varepsilon_{ij} \left( \frac{1 + \delta}{\rho_{ij} + \delta} \right)^{n-m} \left( \frac{1 + \gamma}{\rho_{ij}^m + \gamma} - 2 \right). \quad (2)$$

The intramolecular valence energy is described by the first three terms. The vdW interaction is described by the buffered-14-7 vdW formula (Eq. (2)).  $\varepsilon_{ij}$  is the potential well depth and  $\rho_{ij}$  is the ratio between the actual separation of *i*-*j* sites and the minimum energy distance, described as  $\rho_{ij} = R_{ij}/R_{ij}^0$ .<sup>43</sup> The electrostatic potential energy is evaluated from the permanent molecular dipole and quadrupole moments (Eq. (3)) and molecular isotropic dipole polarizability (Eq. (4)). All electrostatic energy and force terms, including polarization, are calculated using the PME approach,<sup>40,44</sup>

$$E_{\text{ele}}^{\text{perm}}(ij) = M_i^T T_{ij} M_j, \quad (3)$$

$$\mu_{i,\alpha}^{\text{ind}} = \alpha_i \left( \sum_{\{j\}} T_{\alpha}^{ij} M_j + \sum_{\{j'\}} T_{\alpha\beta}^{ij'} \mu_{j',\beta}^{\text{ind}} \right) \quad \text{for } \alpha, \beta = 1, 2, 3. \quad (4)$$

In Eq. (4), the first term represents the dipole induced at site *i* by the electric field generated from the permanent multipoles of the neighboring molecules. The second term represents the contributions due to the field of induced dipoles on all other sites. In order to avoid the polarization catastrophe, the polarization is modified at short-range with the Thole damping scheme.<sup>45</sup>

### Parameterization data set

The data utilized for fitting the parameters were composed of a combination of experimentally determined liquid properties as well as high-level *ab initio* quantum mechanics (QM)-derived properties. The liquid properties considered were density, enthalpy of vaporization, isothermal compressibility, isobaric heat capacity, thermal expansion coefficient, and dielectric constant. The temperature and pressure combinations were: 1 atm at temperatures ranging from 249 K to 373 K (32 total) and 298 K at pressures from 1 kbar to 4 kbar (4 total).

The *ab initio* QM reference data included properties for systems ranging in size from the monomer to clusters of 23 water molecules. For the monomer, the molecular dipole, quadrupole, polarizability, vibration, and geometry were considered. The *ab initio* QM interaction energy and geometries for the ground state dimer, Smith dimer set (10 total),<sup>46</sup> trimer, tetramer, pentamer, eight hexamers,<sup>47</sup> two octamers,<sup>48</sup> five 11-mers,<sup>49</sup> five 16-mers, two 17-mers, and four 20-mers<sup>50,51</sup> were utilized for calibration. In previous work,<sup>29</sup> over 42 000 cluster (ranging from 2 to 22 water molecules) geometries were obtained from AMOEBA03 simulations of liquid water for temperatures ranging from 249 K to 373 K. Energy and gradients for the clusters were determined via RIMP2<sup>52,53</sup>/heavy-aug-cc-pVTZ<sup>54</sup> as implemented in Q-Chem 4.0.<sup>55</sup> This large compilation of theoretical data was included in the fitting of the model parameters.

### Parameter optimization

We applied ForceBalance, an automatic optimization method, to parameterize the uAMOEBA water model using

the expanded data set described above. ForceBalance supports many different optimization algorithms, and the calculation in this work was carried out using the trust-radius Newton-Raphson (or Levenberg-Marquardt<sup>56,57</sup>) algorithm with an adaptive trust radius.<sup>58,59</sup> This algorithm requires the first and second derivatives of the objective function in the parameter space, which we derive from the first derivatives of the simulated properties using the Gauss-Newton approximation.

A major challenge in force field parameterization is the significant statistical noise in the objective function from the sampling of properties to be matched to experiment. The parametric derivatives are challenging to evaluate because numerical differentiation requires running multiple simulations and evaluating small differences between statistically noisy estimates. ForceBalance uses thermodynamic fluctuation formulas to calculate accurate parametric derivatives of simulated properties without running expensive multiple simulations.<sup>60,61</sup> 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:

$$\begin{aligned}\langle A \rangle_\lambda &= \frac{1}{Q(\lambda)} \int A(r, V) \exp(-\beta(E(r, V; \lambda) + PV)) dr dV, \\ Q(\lambda) &= \int \exp(-\beta(E(r, V; \lambda) + PV)) dr dV,\end{aligned}\quad (5)$$

where  $A$  is the observable,  $r$  a given molecular configuration in a periodic simulation cell,  $\lambda$  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  $\lambda$  subscript represent an ensemble average in the thermodynamic ensemble of the force field parameterized by  $\lambda$ . 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  $\lambda$  only through the potential energy  $E$ , we can differentiate Eq. (5) analytically,

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

The potential energy derivative  $\frac{dE}{d\lambda}$  is evaluated by numerically differentiating the potential energies at the sampled structures. Eq. (6) provides a way to evaluate the parametric derivative of thermodynamic properties without running additional sampling simulations, though the derivative of any observable always manifests as a higher order correlation function and has a larger statistical error than the observable itself. This equation may be directly applied to obtain derivatives of ensemble-averaged observables with implicit parametric dependence through the thermodynamic ensemble, such as the density. Eq. (6) is easily extensible to obtain derivatives of observables with explicit parameter dependence, such as the enthalpy; derivatives for higher-order thermodynamic response properties such as the dielectric constant are obtained using the chain rule.<sup>29</sup> We omit the calculation of second parametric derivatives for reasons of computational cost and statistical noise, relying instead on the least-squares form of the objective function and the Gauss-Newton approximation to obtain the Hessian.

The problem of overfitting is treated by regularization via a penalty function, which corresponds to imposing a *prior* distribution of parameter probabilities in a Bayesian interpretation. The prior widths reflect the expected variations of the parameters during the optimization. We used a Gaussian prior distribution, corresponding to a parabolic penalty function in parameter space centered at the original AMOEBA03 parameter values with the chosen force constants. Since the various parameters have different physical meanings

(e.g., vdW well depth, O-H bond length), each parameter type was assigned its own prior width.

We ran the optimization until fluctuations from numerical noise prevented further improvement. The calculation converged to within the statistical error after about 15 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.

## COMPUTATIONAL DETAILS

### Parameterization calculations

ForceBalance carried out the condensed phase simulations in the optimization by interfacing with OpenMM 6.1,<sup>62-64</sup> a graphics processing unit (GPU)-accelerated molecular dynamics software package with an extensively validated implementation of AMOEBA03, which provides a speedup of an order of magnitude over the reference implementation in TINKER 6.3.<sup>65</sup> At each optimization step, the set of 36 simulations at different phase points (32 temperatures at 1.0 atm pressure plus 4 pressures at 298.15 K temperature) is performed simultaneously on multiple nodes in a GPU cluster; the Work Queue library<sup>66-68</sup> 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 were analyzed using the multistate

Bennett acceptance ratio (MBAR) estimator,<sup>69,70</sup> which allows each simulation to contribute to the estimated properties of all other simulations. This combination of methods allowed us to optimize the condensed phase properties very accurately. Due to the non-overlapping features of the simulation codes, we combined OpenMM 6.1 and TINKER 6.3 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,<sup>71</sup> TINKER,<sup>65</sup> OpenMM,<sup>64</sup> and Work Queue<sup>68</sup> are freely available on the web.

### Validation calculations

The validation calculations were conducted using the TINKER 6.3 modeling package. PME summation was used to handle the electrostatic interactions (real-space cutoff at 7 Å) and the atom-based switching window was applied to restrict the vdW interactions with a cutoff of 12 Å. During the MD simulation of *NPT* ensembles, we utilized the Nose-Hoover algorithm to integrate the equation of motion and control pressure and temperature.<sup>72</sup>

### Diffusion

The diffusion coefficient is typically computed from the slope of the mean-square displacement as a function of time, averaged over the MD trajectories of individual particles,

$$D_{PBC} = \lim_{t \rightarrow \infty} \frac{1}{6} \frac{d \langle |r(t) - r(0)|^2 \rangle}{dt}. \quad (7)$$

Yeh and coworkers<sup>73</sup> showed that for a system of nearly 2000 water molecules in a cell with periodic boundary conditions (PBC), the diffusion coefficient could be underestimated by around 10%; correcting for such systematic errors is particularly crucial in comparisons of simulation to experiment when transport properties are used to assess interaction potentials.<sup>74-76</sup> A correction to the system-size dependence was proposed,

$$D_0 = D_{PBC} + \frac{2.837297k_B T}{6\pi\eta L}, \quad (8)$$

where  $L$  is the length of the cubic simulation box,  $D_{PBC}$  is the diffusion coefficient calculated in the simulation,  $k_B$  the Boltzmann constant,  $T$  the absolute temperature, and  $\eta$  the shear viscosity.

To obtain the system-size independent diffusion coefficient  $D_0$ , we calculated the diffusion coefficients  $D_{PBC}$  under periodic boundary conditions with  $N = 216, 343, 512, 1000, 1600,$  and  $2500$  water molecules. The production time of the simulations is 6 ns ( $N \leq 512$ ), 5 ns ( $N = 1000, 1600$ ), or 3 ns ( $N = 2500$ ). In practice, each MD trajectory was divided into 500 ps blocks from which the water diffusion coefficient was evaluated. The final  $D_{PBC}$  was computed as the average over these blocks (TABLE SXI).<sup>83</sup> The size-independent diffusion coefficient  $D_0$  was obtained by fitting a straight line to  $D_{PBC}$  vs.  $1/L$  and extrapolating to  $1/L = 0$ .

### Viscosity

No significant system-size dependence of viscosity was observed in computer simulations of Lennard-Jones fluids.<sup>77,78</sup> Here, we used the slope of Eq. (8) and data from Table SXI<sup>83</sup> to obtain the average viscosity. To estimate the error bar for size-independent viscosity and self-diffusion coefficient, the size-dependent diffusion coefficient for each box size was computed from a randomly selected 500 ps block of trajectory. These diffusion coefficients for the 6 box sizes were then combined to compute the size-independent viscosity and diffusion coefficient according to Eq. (8). The above process was repeated for 16 times to calculate the standard error.

## RESULTS AND DISCUSSION

### Optimized parameters

The final parameters for the new uAMOEBA, AMOEBA03, and iAMOEBA models are compared in Table I. The first six rows contain the intramolecular parameters for uAMOEBA. The equilibrium bond length was set to 0.9499 Å, which is slightly shorter than the experimental value of 0.9572 Å.<sup>79</sup> The ideal bond angle parameter of 105.9° is reduced by 1°-2° from the iAMOEBA and AMOEBA03 values. The slightly increased value from the experimental gas-phase angle of 104.52° is necessary to reproduce the experimental liquid properties such as the dielectric constant<sup>7,80,25</sup> (Fig. 3). The three force constants for the valence terms were fit to reproduce the experimental gas-phase vibrational frequencies of the water monomer.<sup>81</sup> The bond, angle, and the Urey-Bradley force constants for uAMOEBA are essentially unchanged from the iAMOEBA values.

In uAMOEBA, the non-bonded interaction sites on the hydrogen atoms are removed. The repulsion-dispersion parameters (vdW radius and well-depth), permanent molecular multipole (dipole and quadrupole) moments, isotropic molecular dipole polarizability, and a single vdW interaction site are assigned to oxygen, which is slightly shifted from the molecular center of mass.

The next five rows contain the vdW parameters, followed by twelve rows containing the permanent multipole parameters. uAMOEBA has a larger vdW radius and well depth compared to AMOEBA03, which is largely due to the removal of vdW interaction sites from hydrogen. With no partial charges in this model, the electrostatic representation relies on high order molecular moments, which are significantly different from the atomic multipole moments of AMOEBA03 or iAMOEBA. In this case, it is more meaningful to compare molecular properties as described below.

### Fitted gas phase water properties of uAMOEBA

Recently, Abascal and Vega pointed out that water multipole moments, specifically quadrupole moments, are crucial for capturing water properties from the vapour and liquid to solid phases accurately in multi-site models.<sup>82</sup> In Table II, a comparison of the experimental, *ab initio* QM, and calculated molecular dipole moments, quadrupole moments, and



TABLE I. Parameters for uAMOEBA water model.

Parameter name	Units	uAMOEBA	AMOEBA03 <sup>25</sup>	iAMOEBA <sup>29</sup>
O–H equilibrium bond length	Å	0.949 9	0.957 2	0.958 4
O–H bond force constant	kcal/mol Å <sup>-2</sup>	557.55	529.60	557.63
H–O–H equilibrium angle	deg	105.95	108.50	106.48
H–O–H angle force constant	kcal/mol Å <sup>-2</sup>	48.93	34.05	49.90
H–H Urey–Bradley length	Å	1.516 8	1.553 7	1.535 7
H–H Urey–Bradley force constant	kcal/mol Å <sup>-2</sup>	-9.74	38.25	-10.31
Oxygen vdW sigma	Å	3.755 3	3.405 0	3.645 3
Oxygen vdW epsilon	kcal/mol	0.142	0.110	0.197
Hydrogen vdW sigma	Å	...	2.655	...
Hydrogen vdW epsilon	kcal/mol	...	0.013 5	...
Hydrogen vdW reduction factor	none	...	0.91	...
Oxygen charge	<i>e</i>	...	-0.519 66	-0.594 02
Oxygen dipole Z-component	<i>e bohr</i>	0.708 89	0.142 79	0.088 48
Oxygen quadrupole XX-component	<i>e bohr</i> <sup>2</sup>	2.080 11	0.379 28	0.226 18
Oxygen quadrupole YY-component	<i>e bohr</i> <sup>2</sup>	-2.088 28	-0.418 09	-0.322 44
Oxygen quadrupole ZZ-component	<i>e bohr</i> <sup>2</sup>	0.008 17	0.038 81	0.096 26
Hydrogen charge	<i>e</i>	...	0.259 83	0.297 01
Hydrogen dipole X-component	<i>e bohr</i>	...	-0.038 59	-0.093 91
Hydrogen dipole Z-component	<i>e bohr</i>	...	-0.058 18	-0.125 60
Hydrogen quadrupole XX-component	<i>e bohr</i> <sup>2</sup>	...	-0.036 73	0.187 54
Hydrogen quadrupole YY-component	<i>e bohr</i> <sup>2</sup>	...	-0.107 39	0.021 74
Hydrogen quadrupole XZ-component	<i>e bohr</i> <sup>2</sup>	...	-0.002 03	-0.036 35
Hydrogen quadrupole ZZ-component	<i>e bohr</i> <sup>2</sup>	...	0.144 12	-0.209 28
Oxygen polarizability	Å <sup>3</sup>	1.720 9	0.837 0	0.806 4
Hydrogen polarizability	Å <sup>3</sup>	...	0.496 0	0.504 8
Polarization damping factor	Å <sup>-1</sup>	0.390	0.390	0.236

polarizability of an isolated water molecule at equilibrium geometry is given. The uAMOEBA water monomer possesses a molecular dipole of 1.801 D, similar to that of the all-atom AMOEBA14 water model with revised parameters (1.808 D). The *yy* and *zz* components of the uAMOEBA molecular quadrupole moments are in better agreement with experimental values than the previous AMOEBA14 or iAMOEBA model. The isotropic molecular polarizability of the uAMOEBA

model is noticeably greater than the AMOEBA14 or experimental value but slightly lower than that of iAMOEBA.

Figure 1 shows the plots of the electrostatic potential of different water models on the vdW surface. The water structure (O and H) used in calculations was based on the MP2/cc-pVTZ optimized geometry. In general, both uAMOEBA and AMOEBA03 compare well with the MP2/6-311++G(2d,2p) electrostatic potential (ESP). The similarity between the

TABLE II. Gas phase monomer properties of the uAMOEBA, AMOEBA14, and iAMOEBA models compared with experiment, evaluated at the energy-minimized geometry. The molecular multipole moments were evaluated at the center of mass of water molecule.

Property		Expt.	QM <sup>a</sup>	uAMOEBA	AMOEBA14 <sup>b</sup>	iAMOEBA
Vibrational frequencies (cm <sup>-1</sup> )	Antisymmetric stretch	3755 <sup>c</sup>	...	3754	3755	3755
	Symmetric stretch	3656 <sup>c</sup>	...	3658	3656	3656
	Bend	1594 <sup>c</sup>	...	1594	1594	1594
Dipole (D)	<i>d<sub>z</sub></i>	1.855 <sup>d</sup>	1.840	1.802	1.808	1.864
Quadrupole (D Å)	<i>Q<sub>xx</sub></i>	2.63 <sup>e</sup>	2.57	2.913	2.626	2.584
	<i>Q<sub>yy</sub></i>	-2.50 <sup>e</sup>	-2.42	-2.693	-2.178	-2.178
	<i>Q<sub>zz</sub></i>	-0.13 <sup>e</sup>	-0.14	-0.220	-0.045	-0.406
Polarizability (Å <sup>3</sup> )	<i>α<sub>xx</sub></i>	1.528 <sup>f</sup>	1.47	1.721	1.767	1.816
	<i>α<sub>yy</sub></i>	1.415 <sup>f</sup>	1.38	1.721	1.308	1.816
	<i>α<sub>zz</sub></i>	1.468 <sup>f</sup>	1.42	1.721	1.420	1.816

<sup>a</sup>Reference 89.<sup>b</sup>Reference 84.<sup>c</sup>Reference 88.<sup>d</sup>Reference 85.<sup>e</sup>Reference 86.<sup>f</sup>Reference 87.

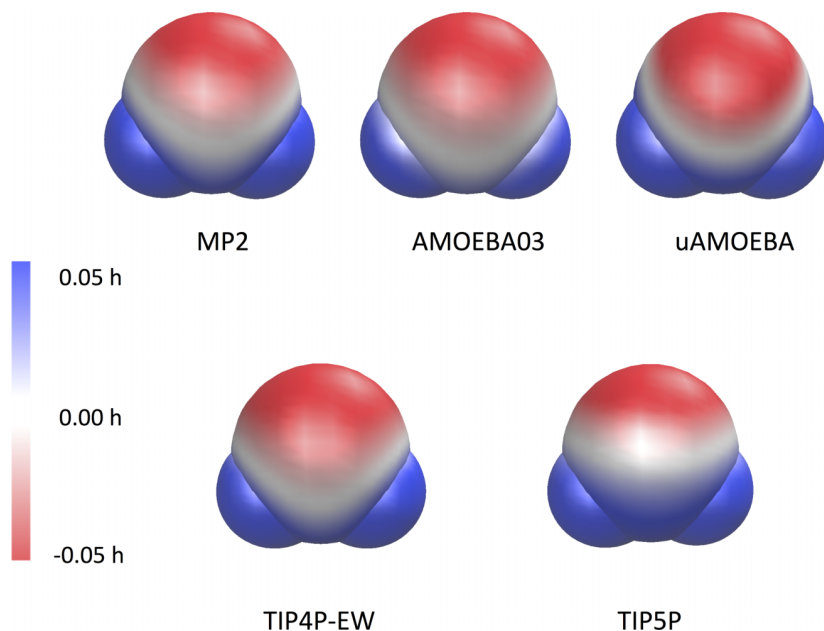


FIG. 1. Electrostatic potential plotted on the vdW surface, with blue representing 0.05 h and red  $-0.05$  h. The MP2 result was obtained using the 6-311++G (2d, 2p) basis set.

uAMOEBa and MP2 potentials around H is notable given that uAMOEBa has no electrostatic parameters on H. However, uAMOEBa is slightly more negative around the oxygen site; this is also possibly a consequence of the missing hydrogen sites, which leads to larger multipole moments on oxygen and stronger electrostatic interactions at short range. This may also explain the slightly over-structured radial distribution function (RDF) plot around the second solvation shell, due to the stronger electrostatic interaction. The fixed charge TIP4P-EW model displays a very reasonable ESP surface while the TIP5P model seems not negative enough around the O. Nonetheless, due to the lack of explicit polarization, both TIP4P-EW and TIP5P models give a water dipole moment of  $\sim 2.3$  D, about 30% higher than the experimental or uAMOEBa/AMOEBa03 value for a gas-phase water monomer (Table II).

The interaction energies of water clusters ranging from dimers to clusters of 20 water molecules<sup>83</sup> are shown in Figure 2. The predicted cluster energies are in generally good agreement with *ab initio* QM results, with a root-mean-

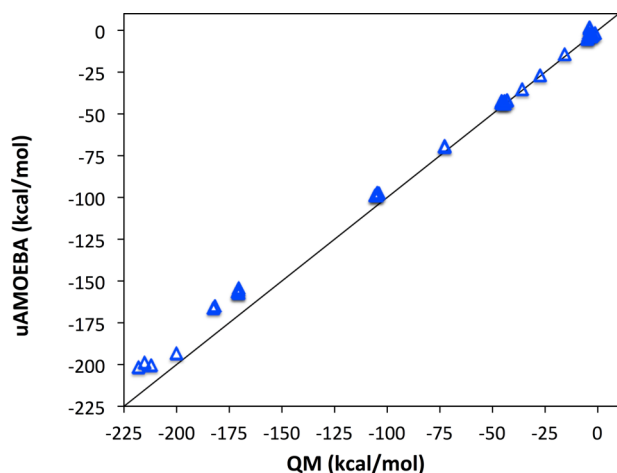


FIG. 2. Cluster energy of gas phase geometry-optimized clusters ranging from size 2 to 20.

square error (RMSE) value of 0.85 kcal/mol per molecule and a correlation coefficient  $r^2$  of 0.988. The error increases for larger clusters. Despite the overall success, uAMOEBa has trouble with certain molecular orientations due to its isotropic, spherical nature. For example, it finds 1.056 and 1.841 kcal/mol for the Smith5 and Smith6 dimer interaction energies, while the experimental value is  $\sim -4$  kcal/mol. Besides, clusters provide a critical calibration for the increasing importance of polarization as one move from the gas-phase toward bulk phases. uAMOEBa is able to accurately reproduce the optimal structures of the water clusters; the root-mean-square deviations (RMSDs) to the reference QM-optimized structures are around 0.15 Å for all of the clusters in the parameterization data set except a couple of hexamers and one eleven-mer (hexamer07, hexamer08, and n11\_434 in the Table SI).<sup>83</sup>

### Fitted liquid water properties of uAMOEBa

Figure 3 shows the temperature dependence of thermodynamic properties water simulated using uAMOEBa and compared to experiment: the density, enthalpy of vaporization, thermal expansion coefficient, isothermal compressibility, isobaric heat capacity, and dielectric constant, which are included as part of the parameterization data set. Overall, the coarse-grained uAMOEBa model is able to describe the liquid properties as well as the previous AMOEBa03 models and other non-polarizable and polarizable models in the literature (Table III).

The enthalpy of vaporization, thermal expansion coefficient, isothermal compressibility, and dielectric constant all show excellent agreement with experiment over a range of temperatures after the parameter fit. The density of liquid uAMOEBa shows a correct maximum at 277 K. Small deviations in density of up to 1.4% are observed at very high temperatures (373.15 K). As the AMOEBa03 model, uAMOEBa overestimates the water heat capacity at room temperature by 3 cal/mol K<sup>-1</sup>, similar to the other two

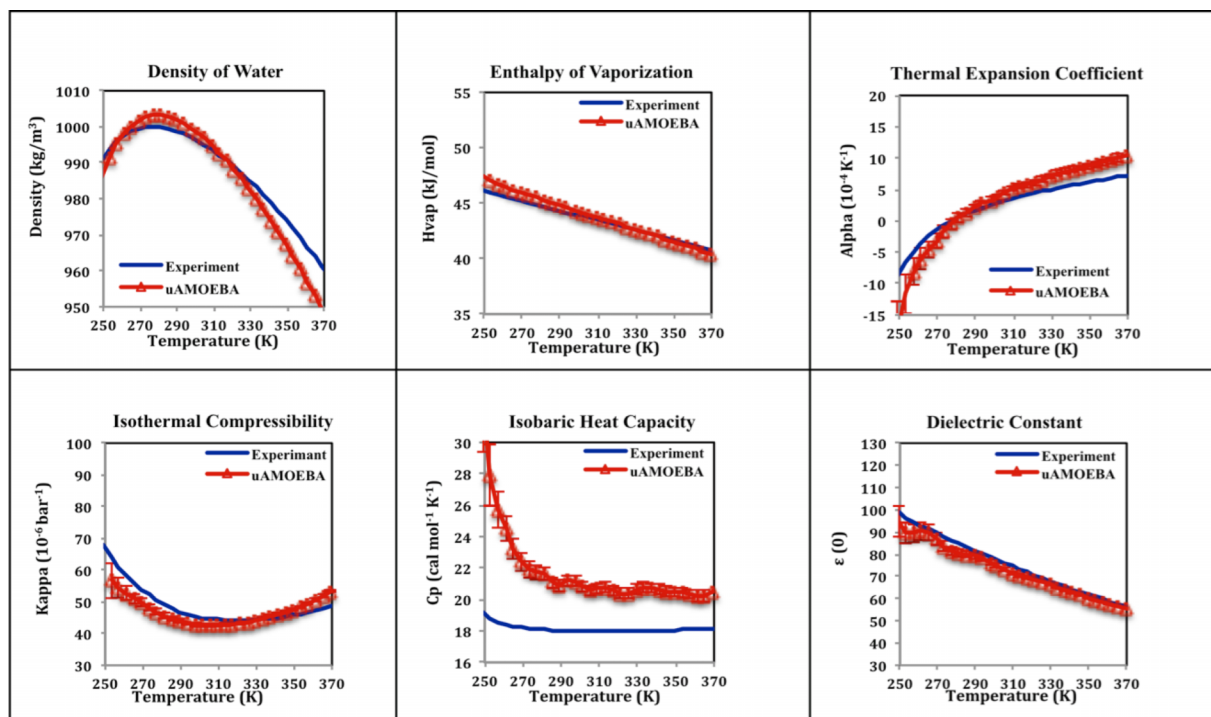


FIG. 3. Thermodynamic properties of uAMOEBA liquid water as a function of temperature. Error bars indicate one standard error.

polarizable models that have reported  $C_p$ , GCPM (22.5), and SWM6 (22.0). This deficiency is likely due to the approximated quantum correction to the heat capacity applied to the classical, flexible model.<sup>12,25,90</sup>

The dielectric constant of water is a critical property that is tightly coupled to the electrostatic model. However, the evaluation of dielectric constant by computer simulation is difficult due to the slow convergence near ambient conditions<sup>91</sup>

and its dependence on the long-range interactions,<sup>7</sup> as well as the H–O–H angle.<sup>80</sup> Previous model shows that the non-polarizable models tend to underestimate the dielectric constant (68 for SPC/E<sup>80</sup> and 62 for TIP4P-Ew<sup>92</sup>), likely due to the fixed atomic charges.<sup>93</sup> Our simulated value for the dielectric constant,  $78.4 \pm 1$ , matches perfectly with the experimental measurement of 78.5 at 298 K. The use of quadrupole moments and incorporation of many body

TABLE III. Comparison of experimental and simulated liquid data from different water models ( $T = 298.15$  K,  $P = 1$  atm). Numbers in parentheses include one standard error in terms of the least significant digit.

Property	Expt.	TIP3P	TIP4P-Ew	SWM4-NDP	TTM3-F	GCPM	SWM6	BK3	AMOEA03	uAMOEBA
Enthalpy of vaporization/kcal mol <sup>-1</sup>										
$\Delta H_{vap}$	10.52	10.45	10.583	10.44	11.40	11.30	10.52	10.94	10.48	10.599 (7)
Density of water/g cm <sup>-3</sup>										
$\rho$	0.997	0.982	0.995	0.994	0.994	1.007	0.996	0.997	1.000	0.999 (1)
Isothermal compressibility/10 <sup>-6</sup> bar <sup>-1</sup>										
$\kappa$	45.3	57.4	48.1	...	...	...	...	44.4	66.0	42.7 (5)
Isobaric heat capacity/cal/mol K <sup>-1</sup>										
$C_p$	18.0	16.8	19.2	...	...	22.5	...	22.0	21.3	21.2 (3)
Static dielectric constant										
$\epsilon$	78.5	94.0	63.90	78.0	67.7	84.0	78.1	79.0	81.4	76.3 (12)
Thermal expansion coefficient/10 <sup>-4</sup> K <sup>-1</sup>										
$\alpha$	2.56	4.10	3.40	...	...	4.20	...	3.01	1.90	3.38 (26)
Self-diffusion coefficient/10 cm <sup>2</sup> s <sup>-1</sup>										
$D_0$	2.30	6.14	2.70	2.85	2.37	2.26	2.14	2.28	2.00	2.41 ± 0.03
Viscosity/cP										
$\eta$	0.896	0.321	0.573	0.66	...	...	0.87	0.95	1.08	0.72 ± 0.05
References	102	9, 103, 104	92, 103, 105, 106	104	19, 107, 108	109, 110	104	109	25	

polarization, even though at a coarse-grained molecular level, seems sufficient to capture the dielectric response of water.

In liquid phase, the instantaneous water dipole moment according to density functional theory (DFT) simulations ranges between 2.6 and 3.0 D at room temperature.<sup>94</sup> The average molecular dipole moment predicted by uAMOEBA, including both the permanent and induced components, is  $2.80 \pm 0.19$  D, which is consistent with the 2.78 D given by the all-atom AMOEBA03 model. Due to the induced dipoles, the liquid phase principle molecular quadrupole moments, located at the water center of mass, changed slightly to (2.962, -2.645, and -0.317).

## Validation of uAMOEBA

In this section, we validate the uAMOEBA model by predicting several properties not used in parameterization, including the RDF, O–O–O angle distribution, self-diffusion coefficient, viscosity, and interaction energy with molecules other than water.

## Radial distribution function

To characterize the liquid structure, the O–O and O–H radial distribution functions were sampled from the *NPT* molecular dynamics simulations. As shown in Fig. 4, the O–O RDF displays two well-defined peaks, similar to the experimentally derived RDFs from X ray scattering data taken by Hura and co-workers using the Advanced Light Source (ALS),<sup>95</sup> more recently by Skinner and co-workers using the Advanced Photon Source (APS)<sup>96</sup> and simulations using the 2013 iAMOEBA model.<sup>29</sup> The positions of the first and second peaks of the uAMOEBA  $g_{OO}(r)$  agree very well with the experimental data, especially the ALS data. Like iAMOEBA, the position of the first trough is slightly shifted to the left. The first peak height is almost identical to that of iAMOEBA, both similar to the ALS RDF. The second peak height is notably higher than the rest, and accordingly the first trough is 0.1 lower than the experimental RDF, which suggests that the second shell of uAMOEBA water is somewhat over structured. The positions of the first and second peaks of uAMOEBA  $g_{OH}(r)$  show excellent agreement with the experimental curve

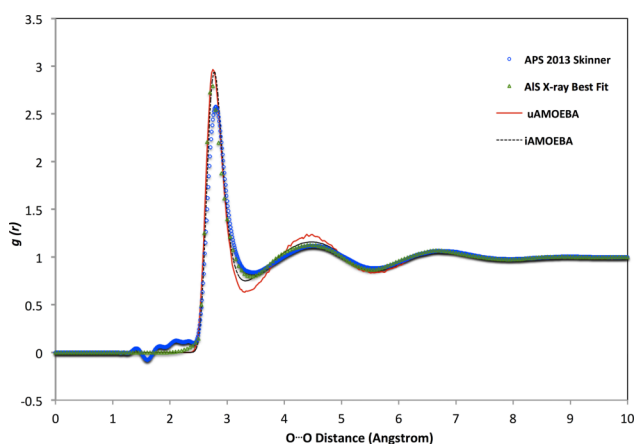


FIG. 4. The oxygen-oxygen RDF curves of the uAMOEBA water model, compared with experimentally derived RDFs.

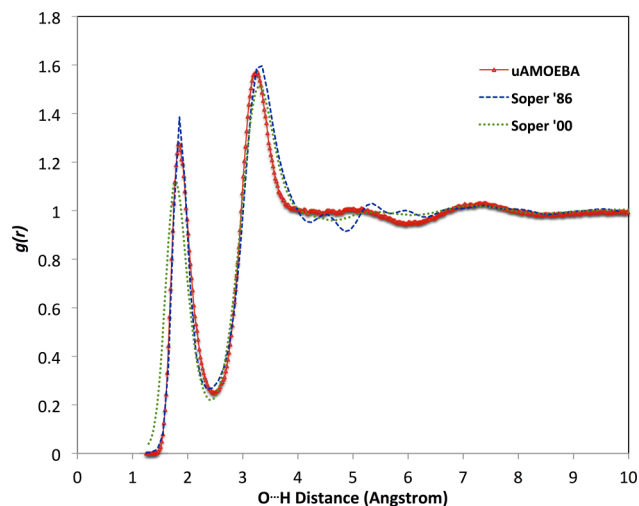


FIG. 5. The oxygen-hydrogen RDF curves of the uAMOEBA water model, compared with experimentally derived RDFs.

(Fig. 5). The first peak appears around 1.9 Å, which matches the reported hydrogen bond length (1.5–2.5 Å<sup>97</sup>). The first peak of the uAMOEBA  $g_{OH}(r)$  is lower than the Soper 1986 data<sup>98</sup> but higher than the Soper 2000 results.<sup>99</sup> Features in the first trough and second peak are similar to the Soper 1986 data.

The O···O···O angle distributions sampled using uAMOEBA, AMOEBA03, TIP4P-EW, and TIP5P models are shown in Fig. 6. From MD trajectories, we computed the O···O···O angle distribution within 3.4 Å of each oxygen atom. Overall, the O···O···O angle distribution in these models suggests a tetrahedron-like structure, as the maximum probability appears around 101°–105°. All models display a small shoulder at 55°–58°, indicating a fifth atom entering the first solvation shell. This is strong evidence that the uAMOEBA model can describe hydrogen bonding as well as the other all-atom models. Note that the uAMOEBA profile is particularly similar to that of TIP4P-Ew.

The results for the self-diffusion coefficient as a function of system size are reported in Table SXI<sup>83</sup> and the corresponding size-independent diffusion coefficient  $D_0$  and viscosities  $\eta$  are summarized in Table III. Polarizable water models

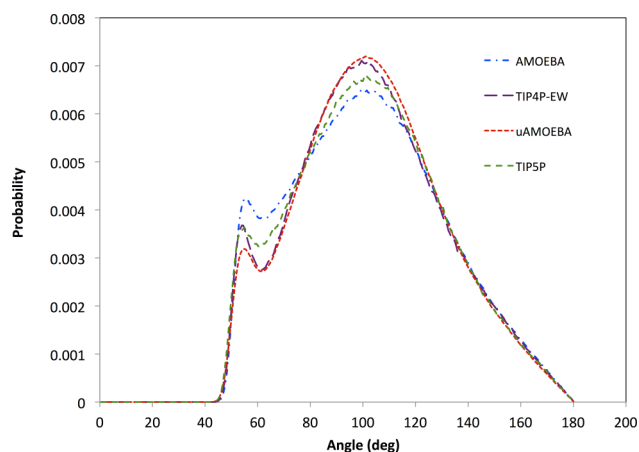
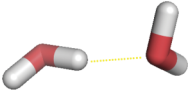
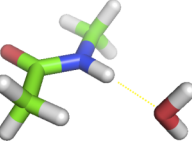
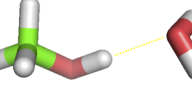
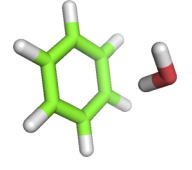
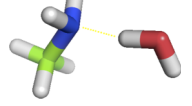
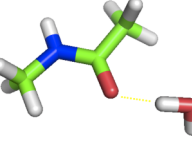
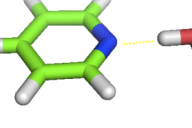
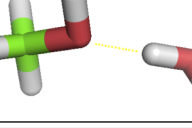


FIG. 6. The O···O···O angle distributions of uAMOEBA, AMOEBA03, TIP4P-Ew, and TIP5P water models.



TABLE IV. Dimer equilibrium interaction energy between water and small molecules. Compared results from uAMOEBa, AMOEBA03 water model, and SAPT2+/CBS (calculated using PSI4<sup>101</sup>). For the water-water dimer, AMOEBA03 or uAMOEBa was used for both molecules. For the other heterodimers, the water was modeled with either uAMOEBa or AMOEBA03 (as labeled in the 3rd and 4th column) while the other molecule was always modeled with AMOEBA03. The structural RMSD values in parentheses represent the structure different from the MP2/cc-pVTZ optimized structures.<sup>111</sup>

Dimer	Dimer structure	uAMOEBa $\Delta E_{\text{int}}$	AMOEBA03 $\Delta E_{\text{int}}$	SAPT2+/CBS $\Delta E_{\text{int}}$
(a) Water-water		-5.16 (0.13)	-4.99 (0.06)	-4.92
(b) Peptide-water		-5.35 (0.10)	-5.16 (0.13)	-5.14
(c) MeOH-water		-5.61 (0.07)	-4.83 (0.05)	-5.03
(d) Benzene-water		-3.14 (0.25)	-3.16 (0.23)	-3.25
(e) Water-MeNH <sub>2</sub>		-6.69 (0.13)	-8.53 (0.08)	-7.18
(f) Water-peptide		-6.48 (0.35)	-8.28 (0.26)	-8.04
(g) Water_pyridine		-6.46 (0.10)	-7.26 (0.05)	-7.15
(h) Water-MeOH		-4.84 (0.24)	-6.77 (0.19)	-5.61

generally produce reasonable diffusion constants compared to nonpolarizable ones,<sup>7</sup> except for SWM4-NDP which is slightly under-polarized. Fig. SX and Table SXI<sup>83</sup> illustrate the strong dependence of the calculated diffusion constant on the system size. With a small box of 216 water molecules, the simulated  $D$  using uAMOEBa is  $1.95 \times 10^{-5} \text{ cm}^2 \text{ s}^{-1}$ , which increases to  $2.21 \times 10^{-5} \text{ cm}^2 \text{ s}^{-1}$  when the simulation box contains 2500 molecules. By using extrapolation, the system-independent self-diffusion constant by uAMOEBa is  $2.41 \times 10^{-5} \text{ cm}^2 \text{ s}^{-1}$ , in excellent agreement with the experimental value.

For the viscosity, which is considered system size-independent, most models underestimate the viscosity except the BK3 and AMOEBA03 model. The deviation in viscosity given by the uAMOEBa model is similar to AMOEBA03 models and larger than those of SWM6, BK3, and iAMOEBa (0.85 cP).<sup>29</sup>

To investigate the transferability of the uAMOEBa model and its ability to interact with the all-atom AMOEBA03 model, we have computed the dimer equilibrium interaction energy between water and several small molecules (Table IV). For the water-water dimer, AMOEBA03 or uAMOEBa was used for both molecules. For the other heterodimers, the water was modeled with either uAMOEBa or AMOEBA03 (as labeled in the 3rd and 4th column in Table IV) while the other molecule was always modeled with AMOEBA03. These comparisons allow us to understand the potential differences between the AMOEBA03 and uAMOEBa water models as solvent. Starting from QM optimized structures,<sup>100</sup> each dimer was optimized using force fields to obtain the corresponding interaction energy. These molecules were chosen to test performance of uAMOEBa water model as the hydrogen bond donor, acceptor and interacting with aromatic benzene. The overall

TABLE V. Efficiency test for the uAMOEBA water model. For all simulations, the vdW cutoff was set to 12 Å and Ewald real-space cutoff was 7 Å. For the first 3 systems of 512 molecules or less, no neighbor-list was used for the vdW calculation. Computer hardware: Intel E5-2697 v2 @ 2.70 GHz, 8 threads.

No. of molecules	AMOEBA03 models/1000 force evaluations	uAMOEBA models/1000 force evaluations	Ratio
216	59.6	12.3	4.8
512	95.6	22.4	4.3
1000	97.3	35.2	2.8
1600	155.4	59.9	2.6
2500	245.2	88.2	2.8
4000	386.0	132.1	2.9

trend given by uAMOEBA model is in good agreement with the SAPT2+/CBS data (calculated using PSI4<sup>101</sup>). The correlation coefficient ( $R^2$ ) between uAMOEBA and SAPT2+/CBS results is 0.83 while the correlation coefficient between all-atom AMOEBA03 and SAPT2+/CBS is 0.92. As a hydrogen bond acceptor, uAMOEBA performs equally well compared to all-atom AMOEBA03, even in the more complicated peptide–water interaction (−5.35 for uAMOEBA, −5.16 for AMOEBA03, and −5.14 kcal/mol for SAPT2+/CBS). In addition, uAMOEBA accurately captures the OH- $\pi$  interaction when facing the aromatic molecules. We also tested the hybrid water dimer where one of the water is described AMOEBA03 while the other is uAMOEBA. The dimer interaction energy is −4.11 kcal/mol when uAMOEBA is the H-bond donor in the dimer, and −5.90 kcal/mol when uAMOEBA is the H-bond acceptor. The average of the two is −5.0 kcal/mol, matching very well with AMOEBA03 or QM values. Overall, the uAMOEBA water model performs reasonably when replacing the AMOEBA03 model in the hybrid uAMOEBA-AMOEBA03 application even though there are no electrostatic or vdW parameters on the hydrogen atoms at all.

### Computational efficiency

The main motivation for developing uAMOEBA is to improve computational efficiency by reducing the number of nonbonded interaction sites. We compared the simulation time between the uAMOEBA model and the all-atom AMOEBA03 model in 1000 steps of gradient evaluation (see Table V). The same simulation settings were used for both models and the only difference is the model parameters that distinguish the one-site model from the three-site model. The efficiency ratio is defined by the simulation time from the all-atom AMOEBA03 model divided by that of the one-site model. For relatively small systems containing less than 1000 atoms, the speed up is almost a factor of 5, while for large water box the improvement is about a factor of 3; this is expected as the computational cost of the PME method scales as  $N \log(N)$ .

### CONCLUSION

Advancement in molecular simulation relies on accurate potential models and efficient sampling methods.<sup>112</sup> In this

work, we present the development of a coarse-grained polarizable water model, “uAMOEBA,” where all nonbonded interactions, including the polarizability, are placed on the oxygen atom. The parameters of this model are determined from a wide range of *ab initio* and experimental data using the automated ForceBalance procedure. The model and parameters are validated by comparing with additional *ab initio* and experimental results, including liquid structural properties, self-diffusion coefficient, shear viscosity, and interaction energies with other small organic molecules. Overall, uAMOEBA shows good transferability between gas and liquid phases, polar and nonpolar environments, most likely because of the incorporation of molecular polarizability. The water structural and dynamic properties given by uAMOEBA are in very good agreement with those derived from all-atom AMOEBA03 model and experiments. The dimer interaction energy between AMOEBA03 small molecules and uAMOEBA water are mixed together also show a satisfying trend in comparison with all-AMOEBA03 and SAPT results. Meanwhile, the computational efficiency is improved by a factor of three compared to atomistic AMOEBA03. We expect uAMOEBA to be a useful solvent model in simulations of biological systems such as proteins and nucleic acids and it can be readily combined with the existing all-atom polarizable protein force field.<sup>113</sup>

### ACKNOWLEDGMENTS

The authors are grateful for support by the Robert A. Welch Foundation (Grant No. F-1691), the National Institutes of Health (Grant No. GM106137), the CPRIT (Grant Nos. R01GM106137 and R01GM114237) and the high performance computing resources provided by TACC and XSEDE (Grant No. TG-MCB100057). L.-P.W. and V.S.P. gratefully acknowledge support from the SimBios program, funded through the NIH Roadmap for Medical Research Grant No. U54 GM072970.

<sup>1</sup>M. Gerstein and M. Levitt, *Sci. Am.* **279**(5), 100–105 (1998).

<sup>2</sup>M. Chaplin, *Nat. Rev. Mol. Cell Biol.* **7**(11), 861–866 (2006).

<sup>3</sup>P. H. Hemmerich and A. H. von Mikecz, *PLoS One* **8**(4), e62018 (2013).

<sup>4</sup>T. Negami, K. Shimizu, and T. Terada, *J. Comput. Chem.* **35**(25), 1835–1845 (2014).

<sup>5</sup>H. Venthur, A. Mutis, J. J. Zhou, and A. Quiroz, *Physiol. Entomol.* **39**(3), 183–198 (2014).

<sup>6</sup>J. Israelachvili and H. Wennerstrom, *Nature* **379**(6562), 219–225 (1996).

<sup>7</sup>B. Guillot, *J. Mol. Liq.* **101**(1-3), 219–260 (2002).

<sup>8</sup>W. L. Jorgensen, *J. Am. Chem. Soc.* **103**(2), 335–340 (1981).

<sup>9</sup>W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, *J. Chem. Phys.* **79**(2), 926–935 (1983).

<sup>10</sup>K. Toukan and A. Rahman, *Phys. Rev. B* **31**(5), 2643–2648 (1985).

<sup>11</sup>H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, *J. Phys. Chem.* **91**(24), 6269–6271 (1987).

<sup>12</sup>H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura, and T. Head-Gordon, *J. Chem. Phys.* **120**(20), 9665–9678 (2004).

<sup>13</sup>J. L. F. Abascal and C. Vega, *J. Chem. Phys.* **123**(23), 234505 (2005).

<sup>14</sup>F. H. Stillinger and A. Rahman, *J. Chem. Phys.* **61**(12), 4973–4980 (1974).

<sup>15</sup>M. W. Mahoney and W. L. Jorgensen, *J. Chem. Phys.* **112**(20), 8910–8922 (2000).

<sup>16</sup>S. W. Rick, *J. Chem. Phys.* **120**(13), 6085–6093 (2004).

<sup>17</sup>L. X. Dang and T. M. Chang, *J. Chem. Phys.* **106**(19), 8149–8159 (1997).

<sup>18</sup>C. J. Burnham and S. S. Xantheas, *J. Chem. Phys.* **116**(12), 5115–5124 (2002).

<sup>19</sup>G. S. Fanourgakis and S. S. Xantheas, *J. Chem. Phys.* **128**(7), 074506 (2008).

- <sup>20</sup>C. J. Burnham and S. S. Xantheas, *J. Chem. Phys.* **116**(4), 1500–1510 (2002).
- <sup>21</sup>C. J. Burnham and S. S. Xantheas, *J. Chem. Phys.* **116**(4), 1479–1492 (2002).
- <sup>22</sup>S. S. Xantheas, C. J. Burnham, and R. J. Harrison, *J. Chem. Phys.* **116**(4), 1493–1499 (2002).
- <sup>23</sup>G. Lamoureux, A. D. MacKerell, and B. Roux, *J. Chem. Phys.* **119**(10), 5185–5197 (2003).
- <sup>24</sup>R. Kumar, F. F. Wang, G. R. Jenness, and K. D. Jordan, *J. Chem. Phys.* **132**(13), 139902 (2010).
- <sup>25</sup>P. Y. Ren and J. W. Ponder, *J. Phys. Chem. B* **107**(24), 5933–5947 (2003).
- <sup>26</sup>P. P. Ewald, *Ann. Phys.* **369**(3), 253–287 (1921).
- <sup>27</sup>P. Y. Ren and J. W. Ponder, *J. Comput. Chem.* **23**(16), 1497–1506 (2002).
- <sup>28</sup>J. W. Ponder and D. A. Case, *Protein Simul.* **66**, 27–85 (2003).
- <sup>29</sup>L. P. Wang, T. Head-Gordon, J. W. Ponder, P. Ren, J. D. Chodera, P. K. Eastman, T. J. Martinez, and V. S. Pande, *J. Phys. Chem. B* **117**(34), 9956–9972 (2013).
- <sup>30</sup>D. Bratko, L. Blum, and A. Luzar, *J. Chem. Phys.* **83**(12), 6367–6370 (1985).
- <sup>31</sup>Y. Liu and T. Ichiye, *J. Phys. Chem.* **100**(7), 2723–2730 (1996).
- <sup>32</sup>T. Ichiye and M. L. Tan, *J. Chem. Phys.* **124**(13), 134504 (2006).
- <sup>33</sup>J. A. Te and T. Ichiye, *J. Chem. Phys.* **132**(11), 114511 (2010).
- <sup>34</sup>S. L. Carnie and G. N. Patey, *Mol. Phys.* **47**(5), 1129–1151 (1982).
- <sup>35</sup>P. G. Kusalik and G. N. Patey, *J. Chem. Phys.* **89**(9), 5843–5851 (1988).
- <sup>36</sup>K. T. Wikfeldt, E. R. Batista, F. D. Vila, and H. Jonsson, *Phys. Chem. Chem. Phys.* **15**(39), 16542–16556 (2013).
- <sup>37</sup>P. A. Golubkov and P. Y. Ren, *J. Chem. Phys.* **125**(6), 64103 (2006).
- <sup>38</sup>P. A. Golubkov, J. C. Wu, and P. Y. Ren, *Phys. Chem. Chem. Phys.* **10**(15), 2050–2057 (2008).
- <sup>39</sup>V. Molinero and E. B. Moore, *J. Phys. Chem. B* **113**(13), 4008–4016 (2009).
- <sup>40</sup>U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, *J. Chem. Phys.* **103**(19), 8577–8593 (1995).
- <sup>41</sup>P. Y. Ren, C. J. Wu, and J. W. Ponder, *J. Chem. Theory Comput.* **7**(10), 3143–3161 (2011).
- <sup>42</sup>P. Ren, T. Liu, J. Q. Qin, and C. T. Chen, *Spectrochim. Acta, Part A* **59**(5), 1095–1101 (2003).
- <sup>43</sup>T. A. Halgren, *J. Am. Chem. Soc.* **114**(20), 7827–7843 (1992).
- <sup>44</sup>C. Sagui, L. G. Pedersen, and T. A. Darden, *J. Chem. Phys.* **120**(1), 73–87 (2004).
- <sup>45</sup>B. T. Thole, *Chem. Phys.* **59**(3), 341–350 (1981).
- <sup>46</sup>B. J. Smith, D. J. Swanton, J. A. Pople, H. F. Schaefer, and L. Radom, *J. Chem. Phys.* **92**, 1240–1247 (1990).
- <sup>47</sup>D. M. Bates and G. S. Tschumper, *J. Phys. Chem. A* **113**(15), 3555–3559 (2009).
- <sup>48</sup>S. S. Xantheas and E. Apra, *J. Chem. Phys.* **120**, 823–828 (2004).
- <sup>49</sup>S. Bulusu, S. Yoo, E. Apra, S. Xantheas, and X. C. Zeng, *J. Phys. Chem. A* **110**(42), 11781–11784 (2006).
- <sup>50</sup>G. S. Fanourgakis, E. Apra, and S. S. Xantheas, *J. Chem. Phys.* **121**, 2655–2663 (2004).
- <sup>51</sup>G. S. Fanourgakis, E. Apra, W. A. de Jong, and S. S. Xantheas, *J. Chem. Phys.* **122**, 134304 (2005).
- <sup>52</sup>R. P. Steele, R. A. DiStasio, Jr., Y. Shao, J. Kong, and M. Head-Gordon, *J. Chem. Phys.* **125**, 074108 (2006).
- <sup>53</sup>R. P. Steele, R. A. DiStasio, and M. Head-Gordon, *J. Chem. Theory Comput.* **5**, 1560–1572 (2009).
- <sup>54</sup>T. H. Dunning, *J. Chem. Phys.* **90**, 1007–1023 (1989).
- <sup>55</sup>Y. Shao, L. F. Molnar, Y. Jung, J. Kusmann, C. Ochsenfeld, S. T. Brown, A. T. Gilbert, L. V. Slipchenko, S. V. Levchenko, D. P. O'Neill, R. A. DiStasio, Jr., R. C. Lochan, T. Wang, G. J. Beran, N. A. Besley, J. M. Herbert, C. Y. Lin, T. Van Voorhis, S. H. Chien, A. Sodt, R. P. Steele, V. A. Rassolov, P. E. Maslen, P. P. Korambath, R. D. Adamson, B. Austin, J. Baker, E. F. Byrd, H. Dachsel, R. J. Doerksen, A. Dreuw, B. D. Dunietz, A. D. Dutoi, T. R. Furlani, S. R. Gwaltney, A. Heyden, S. Hirata, C. P. Hsu, G. Kedziora, R. Z. Khallullin, P. Klunzinger, A. M. Lee, M. S. Lee, W. Liang, I. Lotan, N. Nair, B. Peters, E. I. Proynov, P. A. Pieniazek, Y. M. Rhee, J. Ritchie, E. Rosta, C. D. Sherrill, A. C. Simmonett, J. E. Subotnik, H. L. Woodcock III, W. Zhang, A. T. Bell, A. K. Chakraborty, D. M. Chipman, F. J. Keil, A. Warshel, W. J. Hehre, H. F. Schaefer III, J. Kong, A. I. Krylov, P. M. Gill, and M. Head-Gordon, *Phys. Chem. Chem. Phys.* **8**, 3172–3191 (2006).
- <sup>56</sup>K. Levenberg, *Q. Appl. Math.* **2**, 164–168 (1944).
- <sup>57</sup>D. W. Marquardt, *J. Soc. Ind. Appl. Math.* **11**, 431–441 (1963).
- <sup>58</sup>J. E. Dennis, D. M. Gay, and R. E. Welsch, *ACM Trans. Math. Software* **7**, 348–368 (1981).
- <sup>59</sup>J. J. More and D. C. Sorensen, *SIAM J. Sci. Stat. Comput.* **4**, 553–572 (1983).
- <sup>60</sup>L.-P. Wang, J. Chen, and T. van Voorhis, *J. Chem. Theory Comput.* **9**, 452–460 (2013).
- <sup>61</sup>M. Di Pierro and R. Elber, *J. Chem. Theory Comput.* **9**, 3311–3320 (2013).
- <sup>62</sup>P. Eastman, M. S. Friedrichs, J. D. Chodera, R. J. Radmer, C. M. Bruns, J. P. Ku, K. A. Beauchamp, T. J. Lane, L.-P. Wang, D. Shukla, T. Tye, M. Houston, T. Stich, C. Klein, M. R. Shirts, and V. S. Pande, *J. Chem. Theory Comput.* **9**(1), 461–469 (2013).
- <sup>63</sup>P. Eastman and V. S. Pande, *Comput. Sci. Eng.* **12**(4), 34–39 (2010).
- <sup>64</sup>V. S. Pande and P. Eastman, 2013, <http://www.Simtk.org> for OpenMM.
- <sup>65</sup>J. W. Ponder, “Electrostatics in the AMOEBA force field: The long and the short of it,” presented at 246th National Meeting of the American-Chemical-Society (ACS), Indianapolis, IN, 246, 110-COMP, September 2013.
- <sup>66</sup>B. Abdul-Wahid, L. Yu, D. Rajan, H. Feng, E. Darve, D. Thain, and J. A. Izaguirre, presented at the IEEE International Conference on E-Science, Chicago, IL, 8-12 October 2012.
- <sup>67</sup>P. Bui, D. Rajan, B. Abdul-Wahid, J. A. Izaguirre, and D. Thain, presented at the Workshop on Python for High Performance and Scientific Computing (PyHPC), 2011.
- <sup>68</sup>M. Albrecht, D. Rajan, and D. Thain, presented at the IEEE International Conference on CLUSTER, Indianapolis, IN, 1–8 September 2013.
- <sup>69</sup>M. R. Shirts and J. D. Chodera, *J. Chem. Phys.* **129**(12), 124105 (2008).
- <sup>70</sup>M. R. Shirts and D. L. Mobley, *Methods Mol. Biol.* **924**, 271–311 (2013).
- <sup>71</sup>L.-P. Wang, 2013, <http://www.Simtk.org> for ForceBalance.
- <sup>72</sup>G. J. Martyna, M. E. Tuckerman, D. J. Tobias, and M. L. Klein, *Mol. Phys.* **87**(5), 1117–1157 (1996).
- <sup>73</sup>I. C. Yeh and G. Hummer, *J. Phys. Chem. B* **108**(40), 15873–15879 (2004).
- <sup>74</sup>G. D. J. Phillips, *Macromolecules* **35**(19), 7414–7418 (2002).
- <sup>75</sup>H. Morhenn, S. Busch, H. Meyer, D. Richter, W. Petry, and T. Unruh, *Phys. Rev. Lett.* **111**(17), 173003 (2013).
- <sup>76</sup>J. C. Wu, G. Chatterjee, and P. Y. Ren, *Theor. Chem. Acc.* **131**(3), 1138 (2012).
- <sup>77</sup>X. Zeng, J. Li, H. Xie, and L. Liu, *Chemosphere* **93**(7), 1288–1294 (2013).
- <sup>78</sup>F. Bemani and R. Sadighi-Bonabi, *Phys. Rev. E* **87**(1), 013004 (2013).
- <sup>79</sup>N. Gailar and E. K. Plyler, *J. Chem. Phys.* **24**(6), 1139–1165 (1956).
- <sup>80</sup>P. Hocht, S. Boresch, W. Bitomsky, and O. Steinhauser, *J. Chem. Phys.* **109**(12), 4927–4937 (1998).
- <sup>81</sup>L. X. Dang and B. M. Pettitt, *J. Phys. Chem.* **91**(12), 3349–3354 (1987).
- <sup>82</sup>C. Vega and J. L. F. Abascal, *Phys. Chem. Chem. Phys.* **13**(44), 19663–19688 (2011).
- <sup>83</sup>See supplementary material at <http://dx.doi.org/10.1063/1.4923338> for tables of uAMOEBA and *ab initio* (respectively, experimental) reference values for water cluster interaction energies (respectively, liquid thermodynamic properties), figures showing change of thermodynamic properties during the optimization process, and system size dependence of diffusion coefficient.
- <sup>84</sup>M. L. Laury, L. P. Wang, V. S. Pande, T. Head-Gordon, and J. W. Ponder, “Revised parameters for the AMOEBA polarizable atomic multipole water model,” *J. Phys. Chem. B* (published online).
- <sup>85</sup>S. A. Clough, Y. Beers, G. P. Klein, and L. S. Rothman, *J. Chem. Phys.* **59**(5), 2254–2259 (1973).
- <sup>86</sup>J. Verhoeve and A. Dymanus, *J. Chem. Phys.* **52**(6), 3222 (1970).
- <sup>87</sup>W. F. Murphy, *J. Chem. Phys.* **67**(12), 5877–5882 (1977).
- <sup>88</sup>D. S. Eisenberg and W. Kauzmann, *The Structure and Properties of Water* (Clarendon Press, Oxford, 1969).
- <sup>89</sup>G. Maroulis, *Chem. Phys. Lett.* **289**, 403–411 (1998).
- <sup>90</sup>A. Grossfield, P. Y. Ren, and J. W. Ponder, *J. Am. Chem. Soc.* **125**(50), 15671–15682 (2003).
- <sup>91</sup>R. D. Mountain and D. Thirumalai, *Comput. Phys. Commun.* **62**(2-3), 352–359 (1991).
- <sup>92</sup>H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura, and T. Head-Gordon, *J. Chem. Phys.* **120**(20), 9665–9678 (2004).
- <sup>93</sup>F. N. Keutsch and R. J. Saykally, *Proc. Natl. Acad. Sci. U. S. A.* **98**(19), 10533–10540 (2001).
- <sup>94</sup>I. Tunon, M. T. C. MartinsCosta, C. Millot, and M. F. Ruiz-Lopez, *J. Mol. Model.* **1**(4), 196–201 (1995).
- <sup>95</sup>G. Hura, J. M. Sorenson, R. M. Glaeser, and T. Head-Gordon, *J. Chem. Phys.* **113**(20), 9140–9148 (2000).
- <sup>96</sup>L. B. Skinner, C. C. Huang, D. Schlesinger, L. G. M. Pettersson, A. Nilsson, and C. J. Benmore, *J. Chem. Phys.* **138**(7), 074506 (2013).
- <sup>97</sup>G. A. Jeffrey, *An Introduction to Hydrogen Bonding* (Oxford University Press, New York, 1997).
- <sup>98</sup>A. K. Soper and M. G. Phillips, *Chem. Phys.* **107**(1), 47–60 (1986).
- <sup>99</sup>A. K. Soper, *Chem. Phys.* **258**(2-3), 121–137 (2000).

- <sup>100</sup>J. Rezac, K. E. Riley, and P. Hobza, *J. Chem. Theory Comput.* **7**(8), 2427–2438 (2011).
- <sup>101</sup>J. M. Turney, A. C. Simmonett, R. M. Parrish, E. G. Hohenstein, F. A. Evangelista, J. T. Fermann, B. J. Mintz, L. A. Burns, J. J. Wilke, M. L. Abrams, N. J. Russ, M. L. Leininger, C. L. Janssen, E. T. Seidl, W. D. Allen, H. F. Schaefer, R. A. King, E. F. Valeev, C. D. Sherrill, and T. D. Crawford, *WIREs: Comput. Mol. Sci.* **2**(4), 556–565 (2012).
- <sup>102</sup>W. Wagner and A. Pruss, *J. Phys. Chem. Ref. Data* **31**(2), 387–535 (2002).
- <sup>103</sup>M. A. Gonzalez and J. L. F. Abascal, *J. Chem. Phys.* **132**(9), 096101 (2010).
- <sup>104</sup>W. B. Yu, P. E. M. Lopes, B. Roux, and A. D. MacKerell, *J. Chem. Phys.* **138**(3), 034508 (2013).
- <sup>105</sup>S. Izadi, R. Anandkrishnan, and A. V. Onufriev, *J. Phys. Chem. Lett.* **5**(21), 3863–3871 (2014).
- <sup>106</sup>Y. J. Mao and Y. W. Zhang, *Chem. Phys. Lett.* **542**, 37–41 (2012).
- <sup>107</sup>S. Imoto, S. S. Xantheas, and S. Saito, *J. Chem. Phys.* **138**(5), 054506 (2013).
- <sup>108</sup>S. Yoo and S. S. Xantheas, *J. Chem. Phys.* **134**(12), 121105 (2011).
- <sup>109</sup>P. T. Kiss and A. Baranyai, *J. Chem. Phys.* **138**(20), 204507 (2013).
- <sup>110</sup>P. Paricaud, M. Predota, A. A. Chialvo, and P. T. Cummings, *J. Chem. Phys.* **122**(24), 244511 (2005).
- <sup>111</sup>Q. T. Wang, J. A. Rackers, C. F. He, R. Qi, C. Narth, L. Lagardère, N. Gresh, J. W. Ponder, J. P. Piquemal, and A. P. Ren, *J. Chem. Theory Comput.* **11**(6), 2609–2618 (2015).
- <sup>112</sup>M. K. Gilson and H. X. Zhou, *Annu. Rev. Biophys. Biomol. Struct.* **36**, 21–42 (2007).
- <sup>113</sup>Y. Shi, Z. Xia, J. J. Zhang, R. Best, C. J. Wu, J. W. Ponder, and P. Y. Ren, *J. Chem. Theory Comput.* **9**(9), 4046–4063 (2013).