A single-site multipole model for liquid water
Kelly N. Tran, Ming-Liang Tan, and Toshiko Ichiye

Citation: The Journal of Chemical Physics 145, 034501 (2016); doi: 10.1063/1.4958621
View online: http://dx.doi.org/10.1063/1.4958621
View Table of Contents: http://aip.scitation.org/toc/jcp/145/3
Published by the American Institute of Physics

Articles you may be interested in
Dynamics of water-alcohol mixtures: Insights from nuclear magnetic resonance, broadband dielectric spectroscopy, and triplet solvation dynamics
The Journal of Chemical Physics 140, 114503114503 (2014); 10.1063/1.4868003

The phase diagram of water at negative pressures: Virtual ices
The Journal of Chemical Physics 131, 034510034510 (2009); 10.1063/1.3182727

The random phase approximation applied to ice
The Journal of Chemical Physics 140, 084502084502 (2014); 10.1063/1.4865748

Effects of surface interactions on heterogeneous ice nucleation for a monatomic water model
The Journal of Chemical Physics 141, 084501084501 (2014); 10.1063/1.4892804
A single-site multipole model for liquid water

Kelly N. Tran, Ming-Liang Tan, and Toshiko Ichiye

Department of Chemistry, Georgetown University, Washington, DC 20057, USA

(Received 6 May 2016; accepted 28 June 2016; published online 18 July 2016)

Accurate and efficient empirical potential energy models that describe the atomistic interactions between water molecules in the liquid phase are essential for computer simulations of many problems in physics, chemistry, and biology, especially when long length or time scales are important. However, while models with non-polarizable partial charges at four or five sites in a water molecule give remarkably good values for certain properties, deficiencies have been noted in other properties and increasing the number of sites decreases computational efficiency. An alternate approach is to utilize a multipole expansion of the electrostatic potential due to the molecular charge distribution, which is exact outside the charge distribution in the limits of infinite distances or infinite orders of multipoles while partial charges are a qualitative representation of electron density as point charges. Here, a single-site multipole model of water is presented, which is as fast computationally as three-site models but is also more accurate than four- and five-site models. The dipole, quadrupole, and octupole moments are from quantum mechanical–molecular mechanical calculations so that they account for the average polarization in the liquid phase, and represent both the in-plane and out-of-plane electrostatic potentials of a water molecule in the liquid phase. This model gives accurate thermodynamic, dynamic, and dielectric properties at 298 K and 1 atm, as well as good temperature and pressure dependence of these properties. Published by AIP Publishing.

INTRODUCTION

Water has many unique and anomalous properties, which make it essential for life as we know it. These properties are generally attributed to the tetrahedral network of hydrogen bonds between water molecules; however, even though water is the most ubiquitous and the most studied liquid on this planet, the molecular features that give rise to this network are still not completely clear. Therefore, it is not surprising that water is difficult to model using empirical potential energy functions in computer simulations. Furthermore, although ab initio molecular dynamics simulations hold promise, they have neither reached the accuracy nor the computational efficiency for typical problems involving water today.

The molecular properties that ultimately give rise to the hydrogen-bonded network of the liquid come from the charge distribution due to the nuclei and electron density comprising the molecule. A water molecule in the gas phase has long been known from experiment to have a large dipole moment and a large quadrupole moment. It is also generally accepted that a water molecule in the liquid phase has an even larger average dipole moment due to the average electronic polarization by the environment; empirical water models that include the average effects of polarization by increasing the moments are referred to as nonpolarizable. The simplest are “3-site” models in which “partial charges” are placed only on the nuclei; widely used 3-site models include SPC/E and TIP3P. Since 3-site models have quadrupole moments even smaller than the gas phase quadrupole, “4-site” models move the negative partial charge from the oxygen to a dummy site slightly away from the oxygen center along the dipole vector towards the hydrogens to increase the quadrupole. The 4-site models include Bernal-Fowler, TIP4P, and many current models such as TIP4P-Ew and TIP4P/2005. In another strategy, the “5-site” models split the negative partial charge into two dummy sites on each side of the molecular plane in tetrahedral directions with respect to the hydrogens, based on the idea of “lone pairs” due to sp³ hybridization. The 5-site models include BNS, ST2, and more current models such as TIP5P and TIP5P-E. However, these models have even smaller quadrupoles than 3-site models.

Quantum mechanical (QM) studies of water molecules have also provided information about the charge distribution. Large basis sets using methods that account for electron correlation were found necessary to reproduce the experimental dipole and quadrupole moments in the gas phase. In addition, sp³ hybridization has long been discounted and the three highest occupied molecular orbitals are best characterized as p-orbitals on the oxygen. Quantum studies of water molecules in the liquid phase have been more difficult because of the computational power needed. Ab initio molecular dynamics studies using plane wave basis sets and density functional theory have shown the increase in the dipole moment in the liquid phase, although the exact values depend on how the electron density is partitioned between molecules. Quantum mechanical–molecular mechanical (QM/MM) studies in which one water is treated quantum mechanically using large basis sets and electron correlation methods while the surrounding water is treated as partial charges also show the increase in dipole moment over the gas phase. The dipole moments obtained in both types of QM...
studies agree well with those from a mean field approach to extracting the dipole from the refractive index (2.95 ± 0.2 D) and from an analysis of structure factors from x-ray and neutron diffraction (2.9 D). However, they are generally somewhat larger than in empirical models for liquid water. Also, the QM/MM studies show a somewhat smaller increase in the quadrupole and octupole moments. In addition, a study showed that the multipoles through the octupole could only be reproduced via partial charges if partial charges were placed on six sites: the three nuclei, two out-of-plane dummy sites almost co-linear with the oxygen to reflect the $p$-orbital character of the highest occupied molecular orbital, and a dummy site in the molecular plane similar to the 4-site models.

Given the complexity of the electron density distribution, many nonpolarizable empirical potential energy functions are able to reproduce many properties of liquid water in simulations surprisingly well with both 4- and 5-site models although both types have notable deficiencies. Although the deficiencies have generally been attributed to the lack of electronic polarizability, an alternative (or additional) explanation is that both the large quadrupole and out-of-plane character are important. In support of the latter, a recent comparison of thirteen water models showed that the degree of symmetry governed by the out-of-plane charge, which is measured by the octupole, and the magnitude of the quadrupole appear correlated with reproducing different liquid properties. In particular, it was shown that these two molecular properties govern the nature of the tetrahedral order of the first hydration shell, which is described by order parameters for the magnitude $S^2$ and symmetry $\Delta S^2$, so that a direct correlation between the multipole moments and the tetrahedral order parameters was made. Since $S^2$ and $\Delta S^2$ were also shown to correlate with liquid properties, these liquid properties were shown to depend on the multipoles. Furthermore, it was shown that the diffusion coefficient and the dielectric permittivity depend on the local and long-range structures, respectively, of water and therefore assess different features of water structure. However, partial charges on at least six sites are required to reproduce both a large quadrupole and out-of-plane character, which leads to greater computational time. Given the reasonable properties of the more recent models, and the increased computational time of additional partial charges, the obvious question is whether it is worthwhile to seek improved empirical models, or if they are "good enough," particularly when water is being modeled as a solvent and not the main focus of interest.

The balance of accuracy and efficiency of a water model is important for one of the major uses of water in simulations today, which is as the solvent for biological macromolecules. Simple 3-site water models may be good enough for many problems, such as when the focus of interest is an active site deeply buried in a large protein. However, the macromolecule–water interactions are critical in other problems involving differences in solvation, such as protein folding, or large surface-to-volume ratios, such as peptides and small proteins, disordered regions of proteins, membranes, and nucleic acids. One problem in current nonpolarizable empirical site models is that they predict the wrong hydration number of small cations such as sodium and potassium, indicating that the aqueous solvation of charged species is incorrect. Another problem is that they are not able to predict the deep minimum in the partial molar volume of ethanol in water at low concentrations of ethanol, which was one of the initial experiments leading to current notions about the nature of the hydrophobic effect. Disturbingly, this indicates that the hydrophobic effect may not be modeled properly in common water models. Yet another problem is found in TIP3P, which is widely used for solvating biomolecules; namely, it has a self-diffusion coefficient that is about 2½ times too large (see below). This indicates that coupling of allosteric or collective motions of proteins to solvent may be misrepresented since the dynamics of solutes are determined by the dynamics of the solvent.

Considering the accuracy of describing the electrostatics, although “partial charges” are often used to describe charge distributions of a molecule at a conceptual level, they are not an accurate description of electron density. In fact, the electrostatic potential due to a charge distribution can be described by a multipole expansion outside of the charge distribution, which becomes exact in the limits of infinite distances or infinite order of multipoles. For a water molecule, it has been shown as necessary to include multipoles up to the octupole because the octupole is the first moment that breaks the tetrahedral symmetry of the positive charge from the hydrogens and the negative charge from the $p$-orbital density. In addition, the success of the soft-sticky dipole-quadrupole-octupole (SSDQO) model of water in reproducing the electrostatic potential, as well as liquid state properties in computer simulations, of common water models using multipoles from the models indicates that it is sufficient. Furthermore, when the multipoles were optimized to reproduce properties of liquid water at 25 °C and 1 atm, the resulting SSDQO1 parameters were able to predict accurate properties of water as a function of temperature and pressure as well as the minimum in the partial molar volume of ethanol in ethanol-water mixtures.

Another consideration is efficiency since ultimately a quantum mechanical description is necessary to accurately describe water. A partial charge model has the advantage of conceptual simplicity over a multipole expansion since Coulomb’s law is used to describe the interactions between partial charges while the multipole expansion of the interaction energy between molecules involves higher order tensor multiplications. However, once the equations are programmed into a computer program, a multipole expansion of the interaction energy can gain in efficiency for a complex charge distribution because only one distance needs to be calculated between the molecules while $n^2$ distances need to be calculated for a model with $n$ partial charges per molecule, but multipole expansions can also lose in efficiency due to the tensor multiplications. This was partially solved in the SSDQO formalism by truncating at the $1/r^5$ term, which includes the dipole–octupole and quadrupole–quadrupole interactions, and further approximating the $1/r^5$ term by including only lower order tensor multiplications.

More recently, the implementation of a computationally fast module for multipole interactions, referred to as MPOLE,
in the CHARMM molecular mechanics/dynamics program\textsuperscript{37} affords the opportunity for exploring accurate and efficient multipole models of water for biological simulations. MPOLE gains its efficiency by solving the multipole expansions via an efficient spherical harmonic approach within a particle mesh Ewald (PME) framework. However, the MPOLE module truncates in order of multipole rather than in power of distance. The importance of truncating in power of distance has recently been demonstrated for water since the higher order multipoles are large\textsuperscript{38} and appropriate modifications have been made in the CHARMM program. These studies also demonstrated that the exact expansion to \(1/r^5\) can mimic the electrostatics of a 3-site model in simulations of liquid water with approximately the same computational efficiency.

As a final note, although nonpolarizable models can account for the average polarization in the homogeneous liquid phase since the average environment of a given water molecule is the same, empirical models with electronic polarization should perform better in inhomogeneous phases and there are current reviews of polarizable models in the literature.\textsuperscript{39,40} However, a good polarizable model should also exceed the performance of the best nonpolarizable model for the homogeneous liquid but this goal has been relatively elusive.\textsuperscript{41} Since the effects of explicit electronic polarization added to a site model may include correcting for deficiencies in the gas phase model, such as the lack of gas phase out-of-plane charge density in a 4-site model, this type of unphysical polarization may lead to other errors. In particular, only a few polarizable models have both large quadrupoles and out-of-plane character built into the unpolarized partial charges\textsuperscript{42,43} even though they are also present in the gas phase. Thus, a good nonpolarizable water model serves as a standard for polarizable models. In particular, it can serve as a standard for inhomogeneous environments in biological problems, where the balance between the accuracy possible in good models for polarization must be balanced by the computational efficiency of nonpolarizable models.

Here, a single-site multipole (SSMP) model for liquid water is developed using MPOLE,\textsuperscript{26} in the CHARMM molecular mechanics/dynamics program.\textsuperscript{37} Although the SSDQO1 parameters have been shown to be very accurate for reproducing properties of liquid water, the SSDQO model utilizes the above-mentioned approximation for the \(1/r^5\) term while here the exact \(1/r^5\) term is used for SSMP. In addition, the multipoles can now be obtained from the QM/MM calculations of the higher multipole moments of water\textsuperscript{24} rather than by fitting them as parameters as in the original SSDQO1 parameters. Therefore, here the multipoles up to the octupole from QM/MM calculations are used in simulations of liquid water in the CHARMM program. Since the QM/MM calculations do not account for charge exchange and overlap between the QM water and the neighboring water molecules, an \textit{ad hoc} method for considering these effects in the multipoles is also presented. Finally, a standard 12-6 Lennard-Jones potential is used for SSMP to simplify combining rules with typical force fields for biomolecules, while SSDQO1 utilized a \(1/r^9\) repulsion. The Lennard-Jones parameters are the only adjustable parameters, which are fit to give good radial distribution functions compared to experiment as well as the liquid state properties of density, heat of vaporization, self-diffusion coefficient, and dielectric permittivity at 298 K and 1 atm. The temperature and pressure dependence of these properties is then explored. Results are compared to different types of site models; specifically, SPC/E, TIP3P, TIP4P-Ew, and TIP5P-E were chosen because they have been used with Ewald-type methods generally used in simulations of biological macromolecules in explicit water. In addition, while TIP4P-Ew and TIP4P/2005 were both optimized to give good properties of water over a range of temperatures and give similar results, TIP4P-Ew was developed for use as a solvent model in biomolecular simulations. Thus, TIP4P-Ew tends to be better near 298 K especially for the dielectric permittivity while TIP4P/2005 is better over a wider range of temperatures.

\textbf{METHODS}

The geometry and multipole moments (\(\mu_i^{(QM)}\)), the dipole; \(\Theta_i^{(QM)}\), the linear and planar quadrupoles; and \(\Omega_i^{(QM)}\) and \(\Omega_j^{(QM)}\), the linear and cubic octupoles) are from the QM/MM calculations of a water-like cluster.\textsuperscript{24} The geometry (OH bond length of 0.965 Å and HOH bond angle of 105.9\textdegree) gives rise to the moments of inertia, but does not affect the electrostatics. Additional calculations were performed to test the QM methods; however, methods used for the original calculations were judged superior so these are summarized in the supplementary material in Table S1.\textsuperscript{44} Since the multipoles were calculated for a central QM water molecule surrounded by four hydrogen bonded MM water molecules, partitioning the electron density of the central water molecule was not necessary. Here, charge transfer and overlap between this molecule and its neighbors was accounted for by assuming a reduction in the spread of charge. Thus, the dipole was reduced from \(\mu_0^{(QM)}\) to a new value \(\mu_0\), and the quadrupole and octupole moments were scaled according to

\[
\Theta_i/\mu_0^2 = \Theta_i^{(QM)}/[\mu_0^{(QM)}]^2
\]

and

\[
\Omega_i/\mu_0^3 = \Omega_i^{(QM)}/[\mu_0^{(QM)}]^3,
\]

where \(i = 1, 2\). The value of \(\mu_0\) (i.e., the final value used for SSMP) was constrained so that the magnitude of the tetrahedral order parameter of the hydration shell \(S^2\),

\[
S^2 = \langle \Sigma_j P_2(\mathbf{u}_j \cdot \mathbf{d}_j) \rangle,
\]

where \(P_2\) is the second order Legendre polynomial, \(\mathbf{d}\) is the unit vector between the central water and the \(j\)th nearest neighbor, \(\mathbf{u}\) is the unit vector between the central water and nearest of four corners of a perfect tetrahedron, and \(j\) is summed over the four nearest neighbors,\textsuperscript{20} was predicted to be the same as predicted from the moments of SSDQO1. In other words, subject to the constraint

\[
S^2(\text{predicted}) = 0.52 + 0.038 \Theta_2^2 - 0.057 \mu_0 (\Omega_2 + 5/4 \Omega_0),
\]

(4)
where \( S^2 \) (predicted) is the predicted order parameter based on the moments, which is equal to 0.76 for SSDQO1 based on its moments, and Eq. (4) is from Ref. 26. The multipoles used here along with those of a few selected models appear in Table I.

The Lennard-Jones 12-6 parameters, \( \sigma \) and \( \epsilon \), were allowed to vary. Initial values were chosen such that the \( 1/\sigma^6 \) (i.e., dispersion) coefficient was near the quantum mechanical value of 625 kcal-Å\(^6\)/mol for the water dimer. However, since the Lennard-Jones parameters account for a variety of differences between the empirical potential and the real potential between two water molecules, the final value of the dispersion was somewhat lower.

Molecular dynamics (MD) simulations and analysis were performed using the molecular mechanics package CHARMM version c40a2. The multipole expansion for the interaction energy truncated at the quadrupole is evaluated using the MPOLE module and a separate dipole-octupole \( 1/r^3 \) term is added explicitly. MPOLE uses the particle mesh Ewald (PME) method for the long-range interactions, with a \( \beta \)-spline coefficient equal to 6, a \( k \) value of 0.55 Å\(^{-1}\), and a real space truncation at 7 Å for all possible interactions between dipoles and quadrupoles; in addition, the dipole-octupole term is calculated in real space only and is turned off at long-distances using the CHARMM switching function from 5 to 7 Å since it is short-ranged. The Lennard-Jones potential was turned off at long-range using the CHARMM switching function from 8 to 10 Å. The cutoff for the pair and image lists was 12 Å, which were updated heuristically. Since orientations are needed for the multipole operations, sites at the hydrogens were generated according to the QM/MM geometry to define a molecular coordinate system centered on the oxygen of each molecule; the SHAKE algorithm was then used to maintain this geometry.

The MD simulations employed the leapfrog Verlet algorithm with a time step of 1 fs and cubic periodic boundary conditions. The simulations were prepared using coordinates from a pre-equilibrated periodic cubic box of 512 water molecules. Initially, velocities were assigned according to a Gaussian distribution at a given temperature every 200 fs for 50 ps and then scaled every 1 ps if the temperature is outside a 5 K range for 250 ps. The simulations then ran unperturbed in the NPT ensemble using the Nosé-Hoover thermostat-barostat. Timings for these simulations are reported in Table II in the supplementary material. After a period of equilibration as noted below, the average properties were calculated using coordinates at 1 ps interval.

The properties used in assessing the Lennard-Jones parameters were the radial distribution function \( g(r) \), density \( \rho \), heat of vaporization \( \Delta H_{\text{vap}} \), diffusion coefficient \( D \), and dielectric permittivity \( \epsilon \). \( \Delta H_{\text{vap}} \) was calculated using the average intermolecular potential energy and volume for the liquid to obtain the enthalpy of the liquid, the ideal gas enthalpy for the gas, and a correction for the rigid, nonpolarizable model,

\[
C = C_{vib} + C_{pol} + C_{ni} + C_x,
\]

where \( C_{vib}, C_{pol}, C_{ni}, \) and \( C_x \) are the vibrational, polarization, nonideal gas, and pressure effects, similar to Ref. 9. According to Ref. 9, \( C_x \) was found to be small and therefore neglected and \( C_{pol} \) was given by

\[
C_{pol} = -\frac{1}{2}(\mu_{\text{gas}} - \mu_{\text{liquid}})^2/\alpha_{\text{gas}},
\]

where \( \mu_{\text{gas}} \) and \( \mu_{\text{liquid}} \) are the dipole moments of water in the gas and liquid phases respectively, and \( \alpha_{\text{gas}} \) is the mean polarizability of a water molecule in the gas phase. Here, \( \mu_{\text{gas}} \) and \( \alpha_{\text{gas}} \) were the same as in Ref. 9 while \( \mu_{\text{liquid}} \) was from the model being simulated; \( C_{vib} \) and \( C_{ni} \) were obtained by interpolation of the values in Ref. 9. \( D \) was calculated from the mean-square O–O displacement as a function of time. A correction for the simulation size dependence of \( D \) was calculated by

\[
D = D_{PBC} + 2.837 29k_BT/(6\pi\eta L),
\]

where \( D_{PBC} \) is diffusion coefficient calculated directly from the simulation. \( L \) is the simulation box length, and \( \eta \) is the viscosity. However, while Ref. 51 used the viscosity calculated for each simulation, here the experimental viscosity of water at each temperature was used; values for \( T = 260 \) K and below were obtained by extrapolation of a polynomial fit of the experimental data. \( \epsilon \) was calculated using conducting boundary conditions.

### Table I. The multipole moments and Lennard-Jones parameters for several water models.

<table>
<thead>
<tr>
<th>Source</th>
<th>( \mu_0 )</th>
<th>( \Theta_0 )</th>
<th>( \Theta_2 )</th>
<th>( \Omega_0 )</th>
<th>( \Omega_2 )</th>
<th>( \Theta_2/\mu_0^2 )</th>
<th>( \Omega_2/\mu_0^2\Theta_2 )</th>
<th>( \sigma ) (Å)</th>
<th>( \epsilon ) (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expt. gas phase</td>
<td>1.86 0.11</td>
<td>2.57 NA</td>
<td>NA</td>
<td>0.74</td>
<td>NA</td>
<td>0.74</td>
<td>0.74</td>
<td>NA</td>
<td>0.74</td>
</tr>
<tr>
<td>MP2 gas phase</td>
<td>1.86 0.11</td>
<td>2.54 −1.35</td>
<td>1.91</td>
<td>0.73</td>
<td>0.40</td>
<td>0.73</td>
<td>0.40</td>
<td>NA</td>
<td>0.73</td>
</tr>
<tr>
<td>MP2/4/MM</td>
<td>2.49 0.13</td>
<td>2.93 −1.73</td>
<td>2.09</td>
<td>0.47</td>
<td>0.29</td>
<td>0.47</td>
<td>0.29</td>
<td>NA</td>
<td>0.47</td>
</tr>
<tr>
<td>SPC/E</td>
<td>2.35 0.00</td>
<td>2.04 −1.57</td>
<td>1.96</td>
<td>0.37</td>
<td>0.41</td>
<td>0.37</td>
<td>0.41</td>
<td>3.169</td>
<td>0.155</td>
</tr>
<tr>
<td>TIP3P</td>
<td>2.35 0.23</td>
<td>1.72 −1.21</td>
<td>1.68</td>
<td>0.31</td>
<td>0.42</td>
<td>0.31</td>
<td>0.42</td>
<td>3.151</td>
<td>0.152</td>
</tr>
<tr>
<td>TIP4P-Ew</td>
<td>2.32 0.21</td>
<td>2.16 −1.53</td>
<td>2.11</td>
<td>0.40</td>
<td>0.42</td>
<td>0.40</td>
<td>0.42</td>
<td>3.164</td>
<td>0.163</td>
</tr>
<tr>
<td>TIP4P-E</td>
<td>2.29 0.13</td>
<td>1.56 −1.01</td>
<td>0.59</td>
<td>0.30</td>
<td>0.16</td>
<td>0.30</td>
<td>0.16</td>
<td>3.097</td>
<td>0.178</td>
</tr>
<tr>
<td>SSDQO1</td>
<td>2.12 0.00</td>
<td>2.13 −1.34</td>
<td>1.15</td>
<td>0.47</td>
<td>0.25</td>
<td>0.47</td>
<td>0.25</td>
<td>3.433</td>
<td>0.089</td>
</tr>
<tr>
<td>SSMP</td>
<td>2.28 0.11</td>
<td>2.46 −1.33</td>
<td>1.61</td>
<td>0.47</td>
<td>0.29</td>
<td>0.47</td>
<td>0.29</td>
<td>3.275</td>
<td>0.106</td>
</tr>
</tbody>
</table>

\(^a\)All values have been rounded for ease of comparison so it is suggested to refer to the original papers for the exact values of each model.

\(^b\)Dipole from Ref. 2 quadrupole from Ref. 3.

\(^c\)Uses approximate multipole expansion and a \( 1/r^6 \) repulsion.
To optimize the Lennard-Jones parameters, different values were simulated at $T = 298$ K and $P = 1$ atm for 1.2 ns; the last 1 ns of the trajectory was used for a preliminary analysis based on the $g(r)$, $\rho$, and $\Delta H_{vap}$. For selected sets of parameters with good $g(r)$, $\rho$, and $\Delta H_{vap}$, an additional 4 ns was simulated and used to calculate $D$ and $\varepsilon$ (i.e., a total of 1.2 ns equilibration, 4 ns production). For selected sets of these parameters with good $\rho$, $\Delta H_{vap}$, $D$, and $\varepsilon$, additional simulations at temperatures from 238 to 318 K at 20 K intervals with an additional 4-ns segments. These simulations were 4 ns. A final set of parameters was chosen from these simulations based on the temperature dependence of these properties.

The final results reported here are at temperatures from 238 to 338 K at 20 K intervals with an additional temperature of 268 K with $P = 1$ atm and at pressures of 500, 1000, 5000, and 10000 atm with $T = 298$ K. These results were from more extensive simulations than used during the optimization of the Lennard-Jones parameters, and generally consisted of 1.2 ns of equilibration and 12 ns of production run. To estimate error bars, standard deviations were calculated by dividing the 12 ns trajectories into three 4-ns segments. For the 238 and 258 K at 1 atm simulations and for the 1000 and 5000 atm at 298 K simulations, two additional 12 ns trajectories from different seeds were generated and error bars were estimated by averaging the three 12-ns segments. While further optimizations of the Lennard-Jones parameters or the multipole moments could be done to improve agreement with temperature dependence, testing of the current choice of parameters with solutes was deemed more important than more exact fitting of temperature dependence. In addition, since no pressure dependent data for TIP4P-Ew appear in Ref. 9, TIP4P-Ew was simulated at pressures of 500, 1000, 5000, and 10000 atm with $T = 298$ K using conditions similar to Ref. 9, i.e., 512 molecules with 0.2 ns of equilibration and very short 6 ns production runs.

**RESULTS**

The multipole moments for SSMP and several other models (SPC/E, TIP3P, TIP4P-Ew, TIP5P-E) are given in Table I. Also, the moments are given for the original QM/MM calculation of a water-like cluster at the MP2/aug-cc-pVQZ level/basis set.\textsuperscript{24} referred to as MP2 + 4MM, $\Theta_2/\mu_0^2$, which is a measure of the relative magnitude of the quadrupole, and $\Omega_2/\mu_0-\Theta_2$, which is a measure of the out-of-plane character,\textsuperscript{26} for the different models are also given in Table I. While SSMP has a dipole similar to common site models, comparing $\Theta_2/\mu_0^2$ shows that the quadrupole is larger, reflective of quantum mechanical studies of the liquid phase. In addition, comparing $\Omega_2/\mu_0-\Theta_2$, which is zero if the negative charges are anti-symmetric to the positive charges in charge and distance to the oxygen, is intermediate between TIP5P-E and TIP4P-Ew, indicating significant out-of-plane character. The final set of Lennard-Jones parameters for SSMP are also given in Table I, which are used in the rest of the results presented.

Simulations of SSMP at 298 K and 1 atm give properties that agree well with experiment. The radial distribution functions $g(r)$ of SSMP water agree well with experimental functions\textsuperscript{54,55} as do the other models studied here except TIP3P (Fig. 1). The spatial distribution functions\textsuperscript{56} of the nearest neighbors (Fig. 2) show that SSMP is similar to both TIP4P-Ew and TIP5P-E, although perhaps closer to TIP4P-Ew, in the asymmetric order around the hydrogens versus the ”lone pair” directions, with a defined separation of the neighbors near the oxygen into two lobes in the lone pair directions. In addition, the density $\rho$, heat of vaporization $\Delta H_{vap}$, diffusion coefficient $D$, and dielectric permittivity $\varepsilon$ at 298 K and 1 atm of SSMP water were calculated with the corrections described in Methods, which agree well with experimental values (Table II). The uncorrected values of $D$ and $\Delta H_{vap}$ for SSMP are given in Table SIII in the supplementary material.\textsuperscript{44} Values for these quantities for the above-mentioned models are also given; these come from Ref. 26 but the corrections for $\Delta H_{vap}$ and $D$ have also been included. With the correction for $D$, it is apparent that the site models all tend to be too fast, indicating that they are a little understructured while SSMP is very slightly overstructured.

![FIG. 1. Radial distribution functions $g(r)$ at 298 K and 1 atm for water: (a) the OO function and (b) the OH function and the HH function, shifted upwards by 1.5. SSMP (red), TIP3P (magenta), TIP4P-Ew (green), TIP5P-E (blue), experiment\textsuperscript{54} (dashed black), and experiment\textsuperscript{55} (dotted black). Larger versions of this figure appear in the supplementary material as FIG. S1.\textsuperscript{44}](image-url)
FIG. 2. Spatial distribution functions of water around a water molecule (red and white stick) in the pure liquid, with the red surface indicating a probability density greater than 3 for (a) SSMP, (b) TIP3P, (c) TIP4P-Ew, and (d) TIP5P-E at 298 K and 1 atm.

based on the correlation of $D$ with the degree of order of the first hydration shell. In addition, $\varepsilon$ of SSMP is much better than either TIP4P-Ew, which is too low, or TIP5P-E, which is too high, characteristic of 4-site and 5-site models, respectively. This indicates that the degree of symmetry between neighbors in the hydrogen and “lone pair” directions is correct based on the correlation of $D$ with the asymmetry of the first hydration shell. Since the 3-site models give the worst agreement with experiment and more information on them appear in many other sources, no further comparisons were made here.

The temperature dependence of SSMP was also examined. The data for TIP4P-Ew and TIP5P-E are from Refs. 9 and 14, respectively, except that the corrections for $\Delta H_{\text{vap}}$ and $D$ have also been included. For $\rho$, SSMP generally falls between TIP4P-Ew and TIP5P-E in reproducing the experimental results except at the highest temperatures (Fig. 3). However, only the Lennard-Jones parameters were varied for SSMP, whereas the partial charges, positions of dummy sites, and Lennard-Jones parameters were optimized for good $\rho$ and $\Delta H_{\text{vap}}$ temperature dependence in TIP4P-Ew and good $\rho$ temperature dependence in TIP5P-E. All three models underestimate the density at very high temperatures while they bracket experiment at low temperatures. In particular, the temperature of maximum density $T_{\text{MD}}$ is 270 K for SSMP, which is similar to the TIP4P-Ew value of 273 K and in reasonable agreement with the experimental value of 277 K. As in prior studies, a $D$ in good agreement with experiment at 298 K is correlated with a $T_{\text{MD}}$ that is in good agreement with experiment since both are determined by the order of the nearest neighbors. Also, the $\Delta H_{\text{vap}}$ temperature dependence is quite good for SSMP and comparable to TIP4P-Ew, while TIP5P-E is worse (Fig. 4), although TIP5P-E was parameterized without considering corrections for being a rigid, nonpolarizable model. In addition, the temperature dependence of $D$ for SSMP is excellent and in much better agreement with experiment than either TIP4P-Ew or TIP5P-E (Fig. 5). Finally, the temperature dependence of $\varepsilon$ for SSMP is in good agreement with experiment near 298 K compared to TABLE II. Selected properties at $\sim$300 K and 1 atm.

<table>
<thead>
<tr>
<th>Source</th>
<th>$\rho$ (g/cm$^3$)</th>
<th>$\Delta H_{\text{vap}}$ (kcal/mol)</th>
<th>$D$ ($10^{-5}$ cm$^2$ s$^{-1}$)</th>
<th>$\varepsilon$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expt.</td>
<td>0.997</td>
<td>10.52</td>
<td>2.30</td>
<td>78</td>
</tr>
<tr>
<td>SPC/E</td>
<td>0.995</td>
<td>10.55</td>
<td>2.87</td>
<td>69</td>
</tr>
<tr>
<td>TIP3P</td>
<td>0.979</td>
<td>9.27</td>
<td>5.76</td>
<td>94</td>
</tr>
<tr>
<td>TIP4P-Ew</td>
<td>0.991</td>
<td>10.50</td>
<td>2.61</td>
<td>66</td>
</tr>
<tr>
<td>TIP5P-E</td>
<td>1.003</td>
<td>9.71</td>
<td>3.03</td>
<td>101</td>
</tr>
<tr>
<td>SSDQO1</td>
<td>0.993</td>
<td>9.31</td>
<td>2.67</td>
<td>73</td>
</tr>
<tr>
<td>SSMP</td>
<td>0.995</td>
<td>10.54</td>
<td>2.24</td>
<td>72</td>
</tr>
</tbody>
</table>

$^a$Experimental values for $\rho$ and $\varepsilon$ come from Ref. 60, for $\Delta H_{\text{vap}}$ come from Ref. 61 and for $D$ come from Ref. 62. Values for models other than SSMP come from the supplementary material for Ref. 26 except for $\Delta H_{\text{vap}}$ which were from the same simulations as in the supplementary material but reported here for the first time. Corrections for $\Delta H_{\text{vap}}$ and $D$ for all models have been added as noted in Methods.

FIG. 3. The density $\rho$ as a function of temperature at 1 atm. SSMP (red), TIP4P-Ew (green), TIP5P-E (blue), and experiment (black).

FIG. 4. The heat of vaporization $\Delta H_{\text{vap}}$ as a function of temperature at 1 atm. SSMP (red), TIP4P-Ew (green), TIP5P-E (blue), and experiment (black).
either TIP4P-Ew or TIP5P-E; however, the slope is somewhat too steep (Fig. 6). Since the average dipole moment decreases with temperature in simulations of the QDO polarizable model for water, one of the few or perhaps only polarizable water model that reproduces the temperature dependence of $\rho$, $\Delta H_{\text{vap}}$, and $\varepsilon$, the overestimate appears to be due to using average polarization appropriate to 298 K. In addition, the error bars at 238 K indicate that even with the extra sampling noted in the methods, the sampling is insufficient to obtain a reliable value of $\varepsilon$, as noted for similar sampling for TIP4P-Ew; however, the sampling appears sufficient for $\rho$, $\Delta H_{\text{vap}}$, and $D$ based on the error bars. This is because $\varepsilon$ is dependent on the long-range structure of the liquid while $\rho$, $\Delta H_{\text{vap}}$, and $D$ are much more dependent on the local structure.

In addition, the pressure dependence of SSMP was examined. The data for TIP5P-E are from Ref. 14, respectively, except that the corrections for $\Delta H_{\text{vap}}$ have also been included, and the data for TIP4P-Ew were generated here. The pressure dependence of $\rho$ for SSMP is good at lower pressures (Fig. 7). This is because it has the best $\rho$ at 1 atm, although it appears to be somewhat more compressible than experiment at high pressure. No experimental data for the pressure dependence of $\Delta H_{\text{vap}}$ were available but SSMP appears reasonable since it is close to TIP4P-Ew (Fig. 8). Also, $D$ for SSMP is close to experiment while TIP4P-Ew is somewhat too high compared to experiment (Fig. 9), as might be predicted from the 1 atm values, and both show the evidence of a slight decrease at high pressures as in experiment. Finally, the pressure dependence of $\varepsilon$ for both SSMP and TIP4P-Ew appears poor compared to experiment (Fig. 10). However, the statistics are insufficient. The TIP4P-Ew simulations were very short (6 ns), which is insufficient time to obtain a converged value of $\varepsilon$. The SSMP simulations were longer, and the additional sampling at 1000 and 5000 atm appears to give converged values $\varepsilon$; however, the dependence of $\varepsilon$ on the long-range structure may mean insufficient sampling of long-range structure, which has a much longer time scale especially at high pressures where fluctuations are smaller. (Note, given the unreliability of $\varepsilon$ from the 1000 and 5000 atm simulations, the 10000 atm results are not shown in Fig. 10.) Alternatively, the apparent discrepancies between SSMP and experiment could be due to
changes in electronic polarization at higher pressures. More sampling at higher pressures with different starting conditions is needed to obtain reliable values of $\varepsilon$ for each model. However, based on the low temperature results, the results for $\rho$, $\Delta H_{\text{vap}}$, and $D$ are adequately sampled since they depend on the short-range order.

**DISCUSSION AND CONCLUSIONS**

A nonpolarizable model of water referred to as SSMP (single-site multipole) that uses only a single site on the oxygen of a water molecule for use in computer simulations is presented here. The single-site is the center of a 12-6 Lennard-Jones potential and a multipole moment expansion is presented here. The single-site is the center of a 12-6 oxygen of a water molecule for use in computer simulations (single-site multipole) that uses only a single site on the oxygen.

The authors are grateful for support from the National Science Foundation through Grant No. CHE-1464766 and from the McGowan Foundation. This work used time on the LoBoS cluster at the Laboratory for Computational Biology, National Heart, Lung, and Blood Institute, National Institutes of Health, which was generously provided by Dr. Bernard R. Brooks, the Extreme Science and Engineering Discovery Environment (XSEDE) granted via No. MCB990010, which is supported by National Science Foundation Grant No. OCI-1053575, and the Medusa cluster maintained by University Information Services at Georgetown University. We also thank Phil E. Mason and John W. Brady for the program used to generate the spatial distributions in Figure 2.
