

Combined fluctuating charge and polarizable dipole models: Application to a five-site water potential function

Harry A. Stern, F. Rittner, B. J. Berne,^{a)} and Richard A. Friesner
*Department of Chemistry and Center for Biomolecular Simulation, Columbia University,
New York, New York 10027*

(Received 12 February 2001; accepted 6 April 2001)

We present a general formalism for polarizable electrostatics based on fluctuating bond-charge increments and polarizable dipoles and its application to a five-site model for water. The parametrization is based largely on quantum-chemical calculations and should be easily transferable to other molecules. To examine basis-set effects we parametrized two models from two sets of quantum calculations, using the aug-cc-pVTZ and aug-cc-pVQZ basis sets. We computed several gas-phase and condensed-phase properties and compared with experiment or *ab initio* calculations as available. The models are quite similar and give condensed-phase properties at ambient conditions that are in reasonable accord with experiment, but evince errors consistent with a liquid-state dipole moment that is slightly too large. The model fit to the aug-cc-pVTZ basis set has a smaller liquid-phase dipole moment and thus gives a somewhat better description of liquid water at ambient conditions. This model also performs well away from room temperature, deviating less than 2% from the experimental density from 0 to 100 °C, and showing good agreement with experimental radial distribution functions, although the temperature of maximum density (~ 20 °C) is slightly too high and the model somewhat underpredicts the persistence of the hydrogen-bond network at elevated temperatures. © 2001 American Institute of Physics.
[DOI: 10.1063/1.1376165]

I. INTRODUCTION

Computer simulation at the atomic level of detail requires an expression for the potential energy (for Monte Carlo) and its gradient (for molecular dynamics or minimization) as a function of the position of the atoms. Though impressive progress has been made in *ab initio* molecular dynamics,^{1–5} in which these things are calculated “on the fly” through a quantum-mechanical electronic structure calculation, simulating a large system such as a protein solvated in water entirely by this method remains computationally unfeasible. Creating and improving simple potential energy functions for use in large-scale condensed-phase simulations therefore remains an important task.

Because of the biological importance and anomalous physical properties of water, more effort has been directed toward models for this molecule than perhaps all others combined. Water potentials may be roughly divided into two categories: empirical and *ab initio*. Empirical potentials are typically extremely simple in form and cheap to evaluate. The parameters of the model are fit to reproduce liquid properties, most often at room temperature, thus compensating for the inability of so simple a model to describe the true intermolecular interaction. Such potentials do quite well in accounting for liquid properties, in some cases even for thermodynamic states far away from those used for parametrization. The simplest empirical potentials have fixed charges and thus include many-body polarization effects implicitly.^{6–11} Later empirical potentials have included po-

larization explicitly in the form of one or more dipole polarizabilities^{12–25} or fluctuating charges.^{26–32} Including explicit polarizability has led to a better description of neat water and aqueous solutions. For instance, the TIP4P/FQ model²⁸ has recently been shown to have a density maximum at 4 °C,³² even though it was fit only to room-temperature properties.

In contrast to empirical models, *ab initio* models^{33–41} attempt to describe the true intermolecular potential energy surface. The functional form is often more complicated and is usually fit to quantum-chemical calculations on monomers, dimers, or small clusters. In principle such potentials should provide an accurate description of water in any thermodynamic state. In practice, while *ab initio* models provide a much better description of small clusters in the gas phase, their predictions of condensed-phase properties (when these have been computed) have not been significantly better than empirical models, especially given the greater computational expense. This is likely due to shortcomings in the quantum mechanical calculations used for parametrization, the finite number of points sampled on the potential energy surface, the fact that data are still fit to a relatively simple analytic function, and the inability to “parametrize away” errors in the simulation itself (finite system size, treatment of long-range electrostatics, quantum effects, use of a rigid model, etc.).

Our approach is a compromise between a strictly *ab initio* parametrization and a purely empirical model. We wish to retain a relatively simple functional form that is cheap enough for large-scale simulations. We are interested in a

^{a)}Electronic mail: berne@chem.columbia.edu

systematic method of parametrization that may be applied generally and easily to other molecules besides water. Indeed, we confess that the primary motivation for this work is not so much to add another water potential to the large collection in the literature as to obtain a measure of the reliability of a parametrization method by trial on a particularly challenging system. Finally, we are especially concerned with performance at physiological conditions (room temperature and pressure), although we view nonphysiological conditions as an important test of the robustness of the model. We have developed a new water potential with these goals in mind. The electrostatic model incorporates both polarizability and intramolecular charge transfer, and is parametrized to gas-phase quantum-chemical calculations, as is the short-range part of the pair potential. However, some details of the model such as the choice of geometry (a five-site rather than a three- or four-site model) have been influenced by performance in the condensed phase, and we have fit the long-range part of the pair potential directly to the liquid density at room temperature and pressure. We have tested the potential by computing a number of gas-phase and condensed-phase properties and comparing to experiment or high-level *ab initio* calculations.

This paper is organized as follows: Section II describes the electrostatic model in a general form as well its specific application to the water model, Sec. III describes the details of the methods used for parametrization and simulation, Sec. IV gives the results of the simulations, and Sec. V presents conclusions.

II. MODEL

A. Fluctuating bond-charge increment/polarizable dipole model

The current model represents the electrostatics of a molecule by a set of dipoles on sites and bond-charge increments between pairs of bonded sites. We use the term “site” to denote either an atom or an off-atom virtual site (for instance, a site intended to represent lone-pair electrons or the M-site in the TIP4P water geometry).

In the fluctuating bond-charge increment formalism,^{42–44} site charges result from the transfer of charge from one site to a bonded neighbor, such that no net charge is created or destroyed. We may represent transfer of (positive) charge from site i to site j by a bond-charge increment q_{ij} , which contributes a charge $-q_{ij}$ to site i and a charge $+q_{ij}$ to site j . The total charge on a site is then the sum of the contributions from all bond-charge increments containing that site.

In analogy to the fluctuating charge model,²⁸ we may expand the energy of transferring a charge q_{ij} between two sites to second order:

$$U(q_{ij}) = \chi_{ij}q_{ij} + \frac{1}{2}J_{ij}q_{ij}^2, \quad (1)$$

where χ_{ij} and J_{ij} are parameters depending only on the types of the sites. (It might be argued that χ_{ij} and J_{ij} should also depend on the distance between the sites, but the current water model does not include such a dependence, since it is rigid. Even for a flexible model, one might not want to include distance dependence, because it implies a force be-

tween the bonded atoms, and one might wish all such forces to be subsumed in the stretch/bend part of the model.)

One may express the parameters χ_{ij} and J_{ij} in terms of the electronegativities χ'_i , χ'_j and hardnesses J'_i , J'_j of the two bonded sites, as well as the coupling between them, J'_{ij} . According to the fluctuating charge model,²⁸ the energy of creating a bond-charge increment q_{ij} (that is, a charge $-q_{ij}$ on site i and $+q_{ij}$ on site j) is

$$U(q_{ij}) = -\chi'_i q_{ij} + \chi'_j q_{ij} + \frac{1}{2}J'_i q_{ij}^2 + \frac{1}{2}J'_j q_{ij}^2 - J'_{ij} q_{ij}^2. \quad (2)$$

Therefore

$$\chi_{ij} = \chi'_j - \chi'_i, \quad (3)$$

$$J_{ij} = J'_i + J'_j - 2J'_{ij}. \quad (4)$$

But there is no reason not to think of χ_{ij} and J_{ij} as fundamental, which has the virtue of replacing five parameters with two. The linear coefficient χ_{ij} is the electronegativity difference between the two sites; the quadratic coefficient J_{ij} is a measure of the work needed to transfer the charge. One might even go a step further and think of the bond-charge increment as a capacitor. Then $1/J_{ij}$ is the capacitance, since the work needed to “charge up” the capacitor is $J_{ij}q_{ij}^2/2$, and χ_{ij} is the “potential difference” of the “battery” hooked up to the capacitor, i.e., the potential difference imposed due to the difference in electronegativities of the two sites.

The expression for the energy of an induced dipole moment $\vec{\mu}_i$ on a site i is very similar:

$$U(\vec{\mu}_i) = \vec{\gamma}_i \cdot \vec{\mu}_i + \frac{1}{2}\vec{\mu}_i \cdot \vec{\alpha}_i^{-1} \cdot \vec{\mu}_i. \quad (5)$$

The quadratic term is the familiar self-energy of an induced dipole; $\vec{\alpha}_i$ is the polarizability of site i . The linear coefficient $\vec{\gamma}_i$ represents (the negative of) an “intrinsic” electric field at site i —that is, an electric field that exists even in the absence of any other sites or external fields. We would expect $\vec{\gamma}_i$ to be nonzero only if the site were part of an asymmetric molecule. The parameter $\vec{\gamma}_i$ is really just a way to introduce a “permanent,” nonzero dipole moment in an isolated molecule; we could have written Eq. (5) in the form $\frac{1}{2}(\vec{\mu}_i - \vec{\mu}_i^0) \cdot \vec{\alpha}_i^{-1} \cdot (\vec{\mu}_i - \vec{\mu}_i^0)$; however, Eq. (5) is more consistent with the notation used for fluctuating bond-charge increments and is somewhat more convenient in that one need keep track of only one dipole moment on a site (rather than both a permanent and induced dipole moment).

The electrostatics of a system of molecules is represented by a collection of interacting bond-charge increments and dipoles. We introduce a scalar coupling $J_{ij,kl}$ between bond-charge increments on sites i,j and k,l ; a vector coupling $\vec{S}_{ij,k}$ between a bond-charge increment on sites i,j and a dipole on site k ; and a rank-two tensor coupling $\mathbf{T}_{i,j}$ between dipoles on sites i and j . Then the total energy is

$$\begin{aligned}
 U(\{q_{ij}\}, \{\vec{\mu}_i\}) = & \sum_{ij} \left(\chi_{ij} q_{ij} + \frac{1}{2} J_{ij} q_{ij}^2 \right) \\
 & + \sum_i \left(\vec{\gamma}_i \cdot \vec{\mu}_i + \frac{1}{2} \vec{\mu}_i \cdot \alpha_i^{-1} \cdot \vec{\mu}_i \right) \\
 & + \frac{1}{2} \sum_{ij \neq kl} q_{ij} J_{ij,kl} q_{kl} + \sum_{ij,k} q_{ij} \vec{s}_{ij,k} \cdot \vec{\mu}_k \\
 & + \frac{1}{2} \sum_{i \neq j} \vec{\mu}_i \cdot \mathbf{T}_{i,j} \cdot \vec{\mu}_j. \quad (6)
 \end{aligned}$$

A natural choice for coupling of bond-charge increments and dipoles that are well-separated in space is the Coulomb interaction:

$$J_{ij,kl} = \frac{1}{r_{ik}} - \frac{1}{r_{il}} - \frac{1}{r_{jk}} + \frac{1}{r_{jl}}, \quad (7)$$

$$\vec{s}_{ij,k} = \frac{\vec{r}_{ik}}{r_{ik}^3} - \frac{\vec{r}_{jk}}{r_{jk}^3}, \quad (8)$$

$$\mathbf{T}_{i,j} = \frac{1}{r_{ij}^3} \left(1 - 3 \frac{\vec{r}_{ij} \vec{r}_{ij}}{r_{ij}^2} \right). \quad (9)$$

The Coulomb interaction diverges as the distance between bond-charge increments and dipoles goes to zero, so will not be appropriate if they are too close. Physically, this represents the fact that a point multipole description is only accurate from far enough away. This can be remedied by using a screening function rather than the bare Coulomb interaction at close distances, which is effectively the same thing as replacing the point multipoles with a “smeared-out” charge distribution; several variations on this approach are described in the literature.^{28,45–51} Alternately or in addition to screening, interactions may simply be omitted or scaled for atoms that are close neighbors within a molecule. Many molecular-mechanics force fields follow this approach.

For any set of spatial coordinates, the bond-charge increments and dipoles are determined by minimizing Eq. (6) with respect to these degrees of freedom; that is, requiring that

$$\frac{\partial U}{\partial q_{ij}} = 0, \quad (10)$$

$$\nabla_{\vec{\mu}_k} U = 0 \quad (11)$$

for all bond-charge increments on sites i, j and dipoles on sites k . It can be shown that for a system consisting only of dipoles, Eq. (11) is equivalent to the usual self-consistent field requirement on the induced dipoles. Likewise, for a system consisting only of fluctuating bond-charge increments, it can be shown that Eq. (10) is equivalent to the usual requirement of electronegativity equalization with a constraint of charge neutrality for every isolated collection of bond-charge increments.

Equations (10) and (11) may be solved by matrix diagonalization or by iterative methods. Alternately, the fluctuating bond-charge increments and dipole moments may be assigned fictitious masses and kinetic energies and integrated

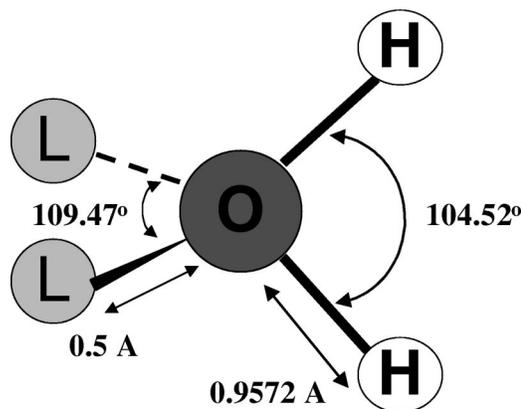


FIG. 1. POL5 model geometry.

along with the spatial coordinates in the extended Lagrangian scheme.^{1,16,28,46} The dynamics so generated is fictitious and functions only as a way to keep the electronic degrees of freedom close to the minimum-energy “Born–Oppenheimer” surface.

B. POL5 water potential

The current water potential, called POL5, has a tetrahedral geometry similar to the ST2 model of Stillinger and Rahman⁶ and the recent TIP5P model of Mahoney and Jorgensen.^{11,52} The OH bond length and HOH bond angles are set to the experimental gas-phase values, 0.9572 Å and 104.52°. Virtual sites denoted L_1 and L_2 are placed along lone-pair directions at a distance of 0.5 Å from the oxygen and making an angle of 109.47°. The geometry of the model is depicted in Fig. 1.

The electrostatics of a POL5 water monomer are represented by bond-charge increments q_{OH_1} and q_{OH_2} between the oxygens and the hydrogens and bond-charge increments q_{OL_1} and q_{OL_2} between the oxygens and the lone-pair sites, in addition to a dipole moment $\vec{\mu}_O$ on the oxygen only. The charge on each site is then

$$q_{H_1} = q_{OH_1}, \quad (12)$$

$$q_{H_2} = q_{OH_2}, \quad (13)$$

$$q_{L_1} = q_{OL_1}, \quad (14)$$

$$q_{L_2} = q_{OL_2}, \quad (15)$$

$$q_O = -q_{OH_1} - q_{OH_2} - q_{OL_1} - q_{OL_2}. \quad (16)$$

In the current model, the oxygen–hydrogen bond-charge increments are allowed to fluctuate, allowing charge transfer between these sites. However, the oxygen lone-pair bond-charge increments are fixed, since allowing these to fluctuate only negligibly improved the agreement of the model with gas-phase quantum-chemical calculations and decreased performance in the condensed phase. It appears that the description of out-of-plane polarization by a dipolar polarizability and charge transfer between lone-pair locations and oxygen is somewhat redundant. This does not appear to be the case with in-plane polarization: including charge transfer between

hydrogens and oxygens improved the model both in the gas phase and condensed phase. It should be noted that without any dipole moment at all, agreement with quantum-chemical gas-phase calculations was significantly poorer, in accord with previous findings,⁴³ so models consisting only of fluctuating charges were not pursued.

We omitted coupling between the two bond-charge increments, or between either bond-charge increment and the dipole moment within the same molecule. Including such coupling (in the form of a bare Coulomb interaction, a screened Coulomb interaction, or a constant coefficient) only marginally improved the agreement with quantum-chemical calculations used for parametrization, and the model is simpler without it. The electrostatic energy of a single POL5 water monomer is then

$$U_{\text{monomer}} = \chi_{\text{OH}}(q_{\text{OH}_1} + q_{\text{OH}_2}) + \frac{1}{2}J_{\text{OH}}(q_{\text{OH}_1}^2 + q_{\text{OH}_2}^2) + \vec{\gamma}_O \cdot \vec{\mu}_O + \frac{\mu_O^2}{2\alpha_O}. \quad (17)$$

Here we treat χ_{OH} , J_{OH} , $\vec{\gamma}_O$, and α_O as parameters fit to quantum-chemical calculations of the electrostatic potential, along with the fixed bond-charge increment q_{OL} , as described in Sec. III.

The coupling between bond-charge increments and dipoles on separate water molecules was taken to be Coulombic as specified in Eqs. (7)–(9), with the following modification. The Coulombic interaction with charges on the lone-pair sites was replaced with a cubic-spline screening function at distances closer than 2.0 Å, so as to represent the lone-pair charges as smeared out over a finite region of space rather than being concentrated at a point. The value of 2.0 Å was chosen because it is at this distance that probe interaction energies first deviate from those predicted by a Coulombic model; see Sec. IV. This distance is interpreted as the width of the lone-pair charge distribution, so that screening for interactions between two lone-pair sites begins at 4.0 Å, screening for interactions between a lone-pair site and any other site begins at 2.0 Å, and interactions between oxygen and hydrogen sites are unscreened. We initially tried placing a screening radius on the oxygen as well, but this did not significantly alter gas-phase results and in fact decreased performance in the condensed phase.

In addition to the electrostatic part, the POL5 potential has a pairwise interaction similar to a Buckingham potential⁵³ between the oxygens in order to take into account repulsion and dispersion:

$$U_{\text{pair}} = A e^{-r_{\text{OO}}/\sigma} - \frac{B}{r_{\text{OO}}^6}, \quad (18)$$

where A , B , and σ are parameters.

III. COMPUTATIONAL DETAILS

A. Parametrization

The electrostatic parameters of the model were fit in a manner similar to that described in Ref. 43. We applied a

TABLE I. Electrostatic parameters for the POL5 water model.

	POL5/TZ	POL5/QZ
$q_{\text{OL}}(e)$	−0.421 88	−0.428 66
$J_{\text{OH}}(\text{Å}^{-1})$	2.638 19	2.645 04
$\alpha_O(\text{Å}^3)$	1.059 99	1.084 16
$\chi_{\text{OH}}(e/\text{Å})$	−1.168 72	−1.165 42
$\vec{\gamma}_{\text{OH}}(e/\text{Å}^2)$	−0.153 59	−0.150 71
$\vec{\gamma}_{\text{OL}}(e/\text{Å}^2)$	0.144 88	0.142 16

series of electrostatic perturbations to a water molecule, in the form of dipolar probes consisting of two opposite charges of magnitude 0.78 e , 0.58 Å apart (for a dipole moment of 2.17 D—similar to that of nonpolarizable models for liquid water such as SPC/E⁹), placed at various locations. The outcome of the fitting procedure was relatively insensitive to the exact form of the perturbations, i.e., the magnitude or position of the probe charges. For each perturbation, the change in the electrostatic potential (ESP) at a set of grid points outside the van der Waals surface of the molecule was computed using density-functional theory (DFT) with the B3LYP method.^{54,55} All calculations were performed with the Jaguar electronic structure code.⁵⁶ The response parameters of the model—that is, J_{OH} and α_O —were chosen to minimize the mean-square deviation between the change in the ESP as given by model and by the DFT calculations. Next, χ_{OH} , $\vec{\gamma}_O$, and q_{OL} were fit so as to best reproduce DFT calculations of the ESP of an unperturbed water molecule. The vector $\vec{\gamma}_O$ was expressed as a sum of vector parameters pointing along the OH and OL bonds, $\vec{\gamma}_{\text{OH}}$ and $\vec{\gamma}_{\text{OL}}$. This avoids having to transform between the molecular frame and the lab frame in order to compute forces on the atoms. (For a rigid water model, this is not so complicated, but we wish our methodology to be readily extendible to larger molecules with internal degrees of freedom.)

To examine basis-set effects, the fitting procedure was performed twice: once with the aug-cc-pVTZ basis set and once with the aug-cc-pVQZ basis set, producing two models, denoted POL5/TZ and POL5/QZ, respectively. The parameters for these two versions of the POL5 potential are quite similar and are listed in Table I.

The cubic spline $f(x)$ used for close-range lone-pair interactions was chosen such that $f(1)=1$ and $f'(1)=-1$, that is, such that the function value and its first derivative match those of $1/x$ at $x=1$. All sites have a “screening radius” associated with them. As stated previously this radius was chosen to be 0 for oxygens and hydrogens and 2.0 Å for lone-pair sites. Screening for the interaction between a pair of sites is imposed at distances closer than s , the sum of the radii of the two sites. For $r < s$, the Coulombic $u(r)=1/r$ is replaced with

$$u(r) = \frac{1}{s} f(r/s). \quad (19)$$

Figure 2 shows a plot of $f(x)$ used for the POL5/TZ model; that used for the POL5/QZ model is almost identical. The parameters for the splines for the two models are given in Table II.

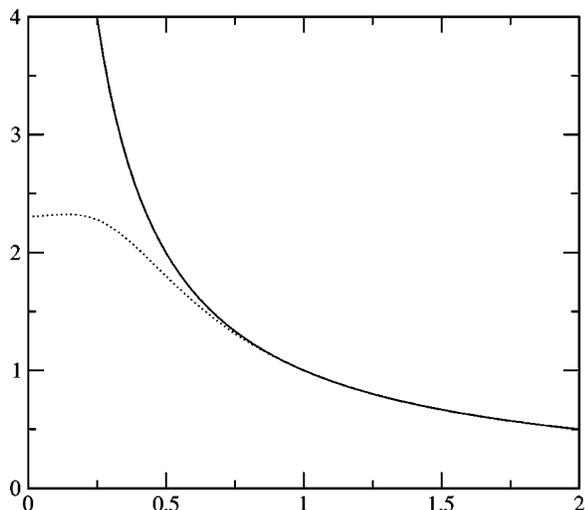


FIG. 2. Cubic-spline screening function (dotted line) used for close-range interactions with lone-pair sites, and $1/x$ (solid line).

The shape of the spline was fit to DFT calculations of the interaction energy of a water molecule and a dipolar probe placed at various distances to the oxygen and in an orientation such that the water mimics a hydrogen bond acceptor, as shown in Fig. 3.

The parameters A , σ , and B were chosen to reproduce the binding energy and O–O distance of the minimum-energy water dimer as well as the density of liquid water at 298.15 K. The target energy and distance were -4.96 kcal/mol and 2.8955 Å; these values were taken from high-level *ab initio* calculations of Halkier *et al.*^{57,58} The target density was 0.997 g/cm³. The resulting parameters for each of the two models are listed in Table III.

B. Simulation methods

All molecular dynamics runs were performed with $N = 256$ molecules and used a time step of 1 fs. Constant energy and volume (*NVE*) simulations were performed with the velocity Verlet integrator, constant temperature and volume (*NVT*) simulations with Nosé–Hoover chain (NHC) thermostats⁵⁹ on each molecule, and constant pressure and temperature (*NPT*) simulations with the Andersen–Hoover-type barostat of Martyna *et al.*⁶⁰ as well as NHC thermostats. All runs used cubic periodic boundary conditions and Ewald summation for the electrostatics.⁶¹ The real-space cutoff for the Ewald sum was $L/2$ where L is the box length, the reciprocal-space cutoff was $5/2\pi L$, and the screening parameter was set to $5.6/L$. The repulsion/dispersion pair potential

TABLE II. Cubic spline parameters. The splines pass through the points (x, y) , and have a derivative of 0 at $x=0$ and -1 at $x=1$.

POL5/TZ	POL5/QZ
(0.2,305)	(0.2,303)
(0.25,2.278)	(0.25,2.292)
(0.5,1.8)	(0.5,1.801)
(0.75,1.312)	(0.75,1.312)
(1,1)	(1,1)

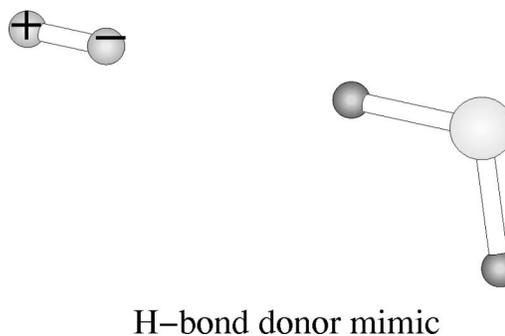
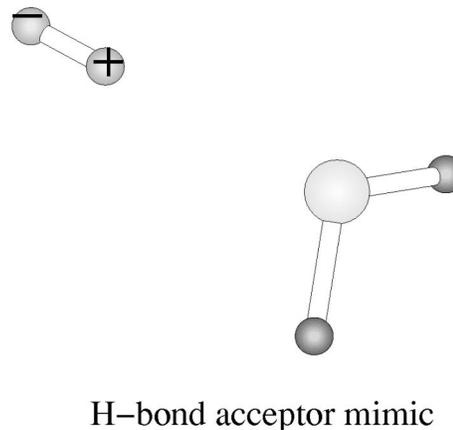


FIG. 3. Dipolar probe geometries. The hydrogen-bond acceptor mimic was used to fit the close-range screening function.

was also truncated at $L/2$. Long-range corrections to the energy and pressure (due to the B/r^6 portion only) were applied.⁶² For *NPT* simulations, the cutoff was scaled along with the box length in order for the long-range corrections to the energy and pressure to be thermodynamically consistent.⁶³

The “electronic” degrees of freedom (the fluctuating bond-charge increments and dipole moments) were propagated using the extended Lagrangian method^{1,16,28,46}—that is, assigned masses and integrated along with the spatial coordinates. The dynamics so generated is fictitious and functions only as a scheme to keep the electronic degrees of freedom close to the minimum-energy “Born–Oppenheimer” surface, without doing expensive iterative solves or matrix inversion. We used the following simple method for choosing the fictitious masses of the fluctuating bond-charge increments and dipole moments: given a single frequency ω , the mass of the bond-charge increment q_{OH} was set to J_{OH}/ω^2 and the fictitious mass of each component

TABLE III. Pair potential parameters.

	POL5/TZ	POL5/QZ
A (kcal/mol)	74 091	60 339
B (kcal/mol Å ⁶)	790	860
σ (Å)	0.298 37	0.306 16

TABLE IV. Computational expense for TIP4P, TIP4P/FQ, and POL5 relative to TIP5P. Results are for simulations using standard Ewald summation and 256 molecules.

Model	Relative expense
TIP4P	0.76
TIP4P/FQ	0.82
TIP5P	1.00
POL5	2.02

of the fluctuating dipole moment $\vec{\mu}_O$ was set to $1/\alpha^2\omega^2$. In this way, if the coupling between different fictitious degrees of freedom is weak, all the fictitious degrees of freedom will be in resonance. Arguably this is beneficial since any leaks of energy from the real system will be quickly distributed throughout the entire fictitious system rather than building up a ‘‘hotspot,’’ which could make the fictitious dynamics unstable. More importantly, if ω is chosen to be much larger than the frequencies of nuclear motion, then the fictitious degrees of freedom will be far from resonance with the nuclear degrees of freedom, little energy will be transferred from the ‘‘real’’ system to the fictitious system, and the electronic degrees of freedom will remain close to the minimum-energy surface as desired. In practice, the choice of $\omega = 50\,000\text{ cm}^{-1}$ worked extremely well: For all simulations, some of which were up to $\frac{1}{2}$ ns in duration with a (nuclear) temperature of $100\text{ }^\circ\text{C}$, the temperature of the fictitious system remained below 3.3 K without any thermostats needed on it.

Computational expense for a system of 256 molecules using standard Ewald summation and the extended Lagrangian method is compared for several water models in Table IV. The POL5 model is about twice as expensive as its counterpart without the dipole. It should be noted that Darden and co-workers have recently combined particle-mesh Ewald (PME) and the extended Lagrangian method for dipolar polarizability models and shown that the cost overhead is only 33% above PME simulations of fixed-charge models.⁶⁴

IV. RESULTS

In this section we present results from calculations of several properties in gas and condensed phase with the POL5 potential. We make comparisons to *ab initio* and experimen-

tal data where available. We also compare with three water potentials that have appeared recently in the literature: the TIP4P/FQ potential of Rick, Stuart, and Berne,²⁸ an empirical polarizable model with fluctuating charges; the TIP5P potential of Mahoney and Jorgensen,¹¹ an empirical fixed-charge model; and the MCDHO model of Saint-Martin *et al.*,⁴¹ an *ab initio* model including polarizability and flexibility.

A. Monomer properties

The first property we examined was simply the quality of the fit. For the POL5/TZ model, the rms deviation between the change in the ESP at the grid points due to the dipolar probe perturbations, as predicted by the model and as given by DFT/aug-cc-pVTZ calculations, was $0.37\text{ kcal/mol } e^{-1}$. For the isolated monomer, the rms deviation between the ESP of the model and the DFT calculations was $0.32\text{ kcal/mol } e^{-1}$. The quality of the fit for the POL5/QZ model was quite similar: the rms deviations between model and quantum mechanics for the change in the ESP due to perturbations, and the ESP for the isolated monomer, were 0.41 and $0.32\text{ kcal/mol } e^{-1}$, respectively.

Gas-phase electrostatic properties for an isolated monomer are listed in Table V. Both POL5 models reproduce the experimental dipole and quadrupole moments fairly well, as expected, since these quantities are also well reproduced by DFT calculations. The polarizability is slightly smaller than the experimental gas-phase value. This is due both to the fact that the quantum-chemical calculations used for parametrization underestimate the polarizability, and to the fact that the electrostatic model is not fit only to the long-range asymptotic response (i.e., the molecular dipolar polarizability), but to the response at close distances as well.

Interaction energies of a water molecule and a dipolar probe, for the geometries shown in Fig. 3 and at various probe distances, are shown in Figs. 4 and 5. Since the cubic-spline screening function is fit to the interaction energies in the acceptor geometry, the model is indistinguishable from the quantum mechanics. The model without any screening is also shown to illustrate the necessity of screening the lone-pair interactions. We note that the interaction energies in the hydrogen bond acceptor geometry first deviate from that given by the unscreened, Coulombic model at a lone-pair/probe distance of 2.0 \AA (i.e., an oxygen-probe distance of 2.5

TABLE V. Gas-phase dipole moment μ , quadrupole moment Q , and polarizability α of a water monomer. The y and z axes lie in the plane of the molecule with the z axis along the C_2 axis of symmetry.

	POL5/TZ	POL5/QZ	TIP4P/FQ ^a	TIP5P ^b	MCDHO ^c	Expt.
μ_z (D)	1.854	1.853	1.860	2.292	1.850	1.855 ^d
Q_{xx} (D \AA^2)	-2.335	-2.332	-1.785	-1.48	-2.44	-2.5 ^e
Q_{yy} (D \AA^2)	2.337	2.335	1.882	1.65	2.67	2.63
Q_{zz} (D \AA^2)	-0.002	-0.003	-0.098	-0.17	-0.24	-0.13
α_{xx} (\AA^3)	1.060	1.084	0.0	0.0	1.217	1.415 ^f
α_{yy} (\AA^3)	1.494	1.517	2.55	0.0	1.482	1.528
α_{zz} (\AA^3)	1.320	1.344	0.82	0.0	1.357	1.468

^aReference 28.^bReference 52.^cReference 41.^dReference 86.^eReference 87.^fReference 88.

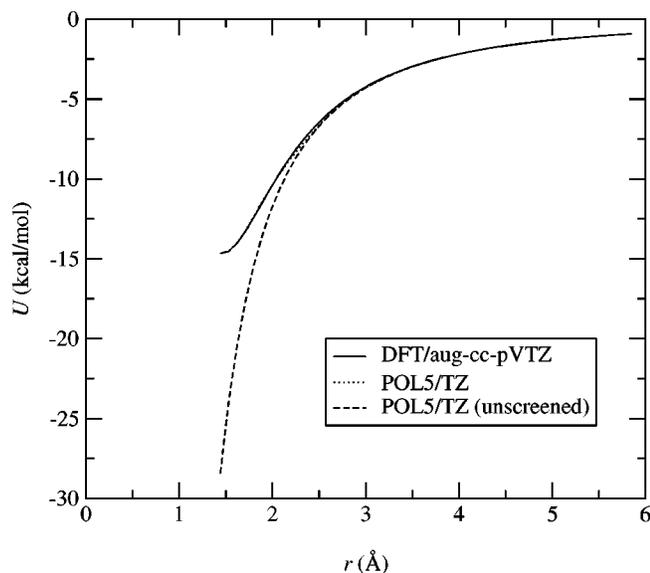


FIG. 4. Interaction energies between a water molecule and a probe where the water mimics a hydrogen-bond acceptor. The x axis is the distance r between the nearest probe charge and the oxygen. These energies were used to fit the close-range screening function, so the screened model is indistinguishable from the DFT calculations.

Å). It was this observation that prompted the choice of 2.0 Å for the “radius” of the lone-pair charge distribution, i.e., the distance at which the bare Coulomb interaction is replaced with the screened interaction. Furthermore, we observe that assigning such a screening width to the site on a hydrogen is unnecessary, and indeed, would be unphysical. Examining the probe interactions where the water simulates a hydrogen bond donor, we observe that the Coulombic behavior persists to close distances, and if anything, the model *underestimates* the DFT interaction energies. This is because the water hydrogen looks like a bare proton as the probe comes closer and penetrates the electron cloud.

B. Gas-phase properties

We determined the optimal dimer geometry as given by the POL5 model and compared with high-level *ab initio* calculations.^{57,58} These calculations constrained the mono-

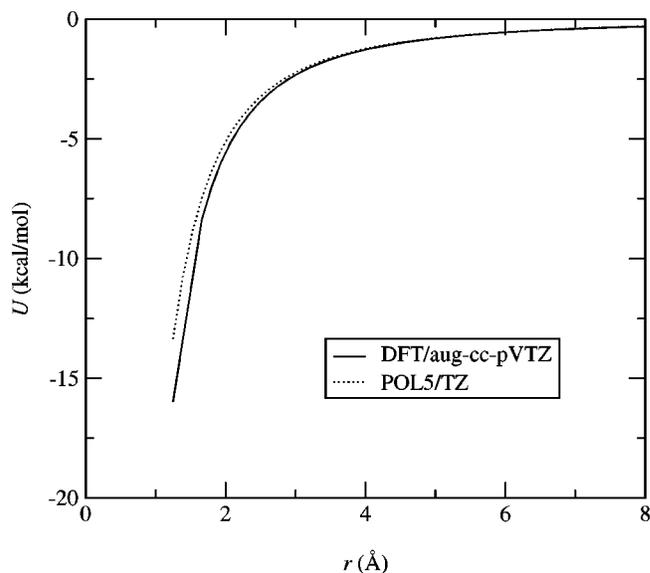


FIG. 5. Interaction energies between a water molecule and a probe where the water mimics a hydrogen-bond acceptor. The x axis is the distance r between the nearest probe charge and the oxygen. These energies were used to fit the close-range screening function, so the screened model is indistinguishable from the DFT calculations.

mers to be rigid (at the experimental geometry), optimized the dimer structure at the coupled-cluster single double triple level with the aug-cc-pVTZ geometry, and extrapolated from a series of calculations at the aug-cc-pVXZ level where X stands for D, T, Q, and 5 to obtain an estimate for the interaction energy in the complete-basis-set limit. Results are summarized in Table VI. The geometry is defined by the O–O distance and two angles as shown in Fig. 6. The minimum-energy dimer O–O distance and interaction energy are reproduced exactly because the short-range repulsion was fit to these quantities; the model also reproduces the orientation of the monomers quite well.

We computed the geometries of the cyclic trimer, tetramer, and pentamer, as well as four conformations of the hexamer: book, cage, cyclic, and prism, for the POL5 models as well as TIP4P/FQ and TIP5P. It should be noted that we did not perform a global conformational search, but started from *ab initio* geometries and performed a local minimiza-

TABLE VI. Water dimer optimum geometry, interaction energy U , net dipole moment μ , and average molecular dipole moment $\langle\mu\rangle$. The geometry is defined by the O–O distance r and the angles θ and ϕ as defined in Fig. 6.

	POL5/TZ	POL5/QZ	TIP4P/FQ ^a	TIP5P ^b	MCDHO ^c	<i>ab initio</i>	Expt.
U (kcal/mol)	-4.96	-4.96	-4.50	-6.78	-5.00	-4.96 ^d	-5.4 ± 0.7^e
r (Å)	2.896	2.896	2.924	2.676	2.916	2.896 ^f	2.98 ^e
θ (deg)	4.694	4.913	0.173	-1.610	3.8	4.754 ^f	0 ± 6^e
ϕ (deg)	62.638	62.385	27.170	50.222	56.1	57.281 ^f	58 ± 6^e
μ (D)	2.435	2.439	3.430	2.920	2.681	2.683 ^g	2.643 ^e
$\langle\mu\rangle$ (D)	2.063	2.065	2.055	2.292	2.086	2.1 ^g	

^aReference 28.

^bReference 11.

^cReference 41.

^dCCSD(T), extrapolation from aug-cc-pVXZ series from Ref. 58.

^eReferences 89,90.

^fCCSD(T)/aug-cc-pVTZ geometry with constrained monomers from Ref. 57.

^gReferences 91,92.

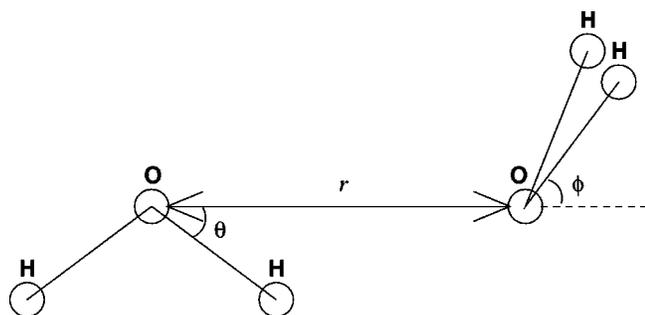


FIG. 6. Three parameters defining the minimum-energy water dimer geometry: r , θ , and ϕ .

tion with the model potential. All geometries were stationary points on the potential surface. Structures for the dimer, trimer, tetramer, and pentamer as given by the POL5/TZ potential are shown in Fig. 7; those given by the other potentials are very similar. Structures for the four hexamer conformations are shown in Fig. 8. From the model geometries, we computed the interaction energy of the cluster, the average distance between oxygens forming hydrogen bonds, the net dipole moment of the cluster, and the average dipole moment of the water molecules in the cluster. Results are summarized in Table VII. With the exception of the trimer, the TIP5P binding energies are consistently too large and hydrogen-bond distances are too short. This is to be expected as the potential implicitly incorporates many-body effects in the form of augmented charges. The water hexamer represents a crossover point, where noncyclic structures become more stable than cyclic ones.⁶⁵ Only MCDHO and the POL5 models appear to capture this behavior. For the hexamers, the two empirical models resemble each other closely in that they give nearly the same relative energies for the four hex-

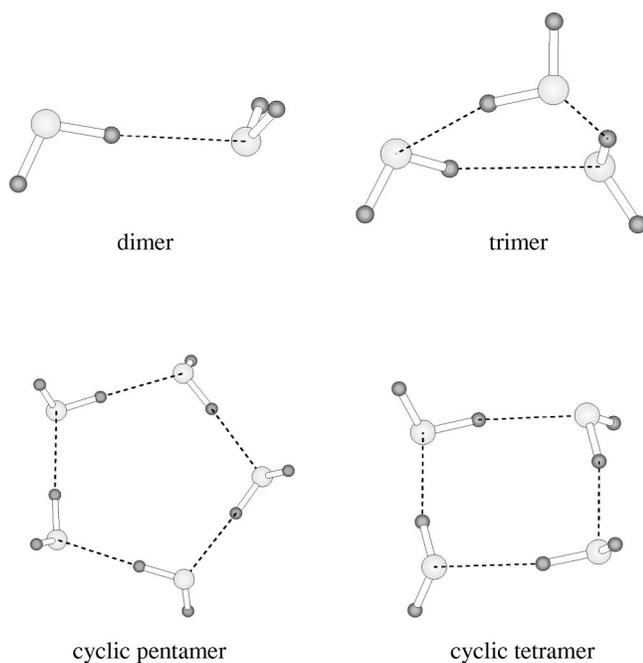


FIG. 7. POL5/TZ structures for small clusters.

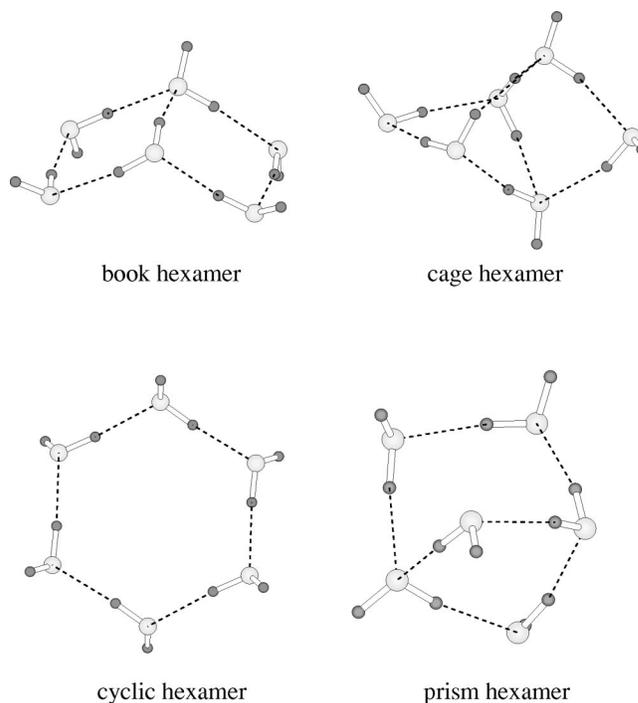


FIG. 8. POL5/TZ structures for four conformations of the water hexamer.

amer conformations, although the TIP5P binding energies are larger by 6 kcal/mol. Both predict the cyclic hexamer to be the most stable. While POL5 and MCDHO do somewhat better in predicting the crossover, they are still inaccurate. In particular, neither model ranks the cage or prism structure as the most stable. These two models give quite similar predictions for all binding energies; POL5 is slightly closer to the *ab initio* results for the smaller clusters, while MCDHO is slightly closer for the hexamers. In general, none of the potentials quantitatively predicts cluster binding energies or hydrogen-bond distances. It should be kept in mind that there is most likely significant uncertainty in the *ab initio* values used for comparison, and the *ab initio* calculations allowed the intramolecular coordinates to relax; calculations constraining the monomer geometries to be rigid might be more appropriate for comparison. All polarizable models reproduce the *ab initio* average molecular dipole moments in the clusters quite well, although the MCDHO model is in somewhat better agreement with the net dipole moments.

We computed the classical second virial coefficient $B_2(T)$ at several temperatures; these results are summarized in Table VIII. The second virial coefficient was computed for the MCDHO model only at temperatures higher than 373 K, so that model does not appear in the table. The TIP4P/FQ and POL5 models significantly underestimate the magnitude of the experimental virial; this is most likely due to the rigidity of these models. On the other hand the TIP5P model significantly overestimates the magnitude of experiment, consistent with its overestimation of cluster binding energies.

C. Liquid at room temperature and pressure

Table IX summarizes liquid-state properties computed at room temperature and pressure from molecular dynamics

TABLE VII. Cluster interaction energy U in kcal/mol, average distance between oxygens in hydrogen bonds (r_{O-O}) in Å, net dipole moment μ , and average molecular dipole moment ($\langle\mu\rangle$), both in D.

	POL5/TZ	POL5/QZ	TIP4P/FQ ^a	TIP5P ^b	MCDHO ^c	<i>ab initio</i>	Expt.
Trimer							
U	-13.416	-13.453	-12.576	-14.992	-13.982	-15.9 ^d	
$\langle r_{O-O} \rangle$	2.901	2.893	2.912	2.770	2.911	2.782 ^d	2.96 ^e
μ	1.205	1.205	0.417	1.074	1.114	1.071 ^f	
$\langle\mu\rangle$	2.218	2.228	2.216	2.292	2.270	2.1 ^f	
Tetramer							
U	-25.529	-25.665	-23.641	-28.431	-27.581	-23.8 ^g	
$\langle r_{O-O} \rangle$	2.769	2.759	2.809	2.673	2.806	2.743 ^g	2.79 ^h
μ	0.000	0.000	0.000	0.000	0.024	0.000 ^f	
$\langle\mu\rangle$	2.468	2.491	2.432	2.292	2.528	2.5 ^f	
Pentamer							
U	-34.111	-34.427	-32.954	-38.122	-35.229	-33.34 ^g	
$\langle r_{O-O} \rangle$	2.742	2.726	2.773	2.657	2.753	2.867 ^g	20.76 ⁱ
μ	1.190	1.191	0.401	1.219	0.922	0.927 ^f	
$\langle\mu\rangle$	2.570	2.607	2.546	2.292	2.689	2.6 ^f	
Hexamer (book)							
U	-42.464	-42.771	-40.152	-46.680	-43.977	-44.74 ^j	
$\langle r_{O-O} \rangle$	2.788	2.777	2.815	2.688	2.809	2.766 ^j	
μ	2.449	2.430	2.006	2.445			
$\langle\mu\rangle$	2.546	2.576	2.509	2.292			
Hexamer (cage)							
U	-41.783	-41.922	-39.297	-45.388	-43.690	-45.03 ^j	
$\langle r_{O-O} \rangle$	2.783	2.775	2.863	2.746	2.888	2.807 ^j	2.820 ^k
μ	2.442	2.454	1.788	2.178	2.034	2.05 ^f	1.904 ^k
$\langle\mu\rangle$	2.486	2.507	2.440	2.292	2.553	2.6 ^f	
Hexamer (cyclic)							
U	-41.785	-42.224	-41.368	-47.309	-44.264	-43.88 ^j	
$\langle r_{O-O} \rangle$	2.737	2.720	2.756	2.654	2.731	2.714 ^j	2.756 ^k
μ	0.017	0.003	0.000	0.000	0.134	0.000	
$\langle\mu\rangle$	2.621	2.662	2.607	2.292	2.791	2.7 ^f	
Hexamer (prism)							
U	-41.847	-42.135	-39.304	-45.805	-44.192	-45.12 ^j	
$\langle r_{O-O} \rangle$	2.792	2.782	2.819	2.773	2.892	2.840 ^j	
μ	2.905	2.931	3.254	2.692	2.627	2.701 ^f	
$\langle\mu\rangle$	2.516	2.544	2.482	2.292	2.558		

^aReference 28.^bReference 11.^cReference 41.^dReference 101.^eReferences 93,94.^fReferences 91,92.^gReference 98.^hReferences 102,103.ⁱReferences 95,96.^jReference 97.^kReferences 99,100.

simulations. The liquid-state energy, density, and average molecular dipole moment were computed in the NPT ensemble at 1 atm and 298.15 K. Runs were equilibrated for 100 ps and averaged over another 300 ps. Both models are parametrized to fit the experimental density, so they repro-

duce it exactly. POL5/TZ reproduces the experimental liquid energy exactly, while POL5/QZ is slightly overbound as is consistent with its larger polarizability and resulting average molecular dipole moment. The MCDHO model is also overbound and has an even larger average molecular dipole moment.

The static dielectric constant ϵ_0 was computed in the NVT ensemble at 0.997 g/cm³ and 298.15 K, from 20 separate runs of 100 ps each, by computing the mean-square fluctuation in the system dipole moment:

$$\epsilon_0 = \epsilon_\infty + \frac{4\pi\rho N}{3kT} (\langle \vec{M}^2 \rangle - \langle \vec{M} \rangle^2). \quad (20)$$

Here \vec{M} is net system dipole moment and ϵ_∞ is the infinite-frequency dielectric constant,

TABLE VIII. Second virial coefficient at various temperatures.

T (K)	$B_2(T)$ (L/mol)				Expt. ^c
	POL5/TZ	POL5/QZ	TIP4P/FQ ^a	TIP5P ^b	
298.15	-0.680	-0.671	-0.640	-2.935	-1.158
310.65	-0.569	-0.562	-0.539	-2.247	-0.966
323.15	-0.483	-0.478	-0.460	-1.765	-0.816
335.65	-0.416	-0.412	-0.398	-1.418	-0.696

^aReference 28.^bReference 11.^cReference 104.

TABLE IX. Liquid-state properties—energy U , density ρ , dipole moment μ , static dielectric constant ϵ_0 , infinite-frequency dielectric constant ϵ_∞ , translational diffusion constant D , and NMR relaxation time τ_{NMR} —at 1 atm and 298.15 K.

	POL5/TZ	POL5/QZ	TIP4P/FQ ^a	TIP5P ^b	MCDHO ^c	Expt.
U (kcal/mol)	-9.92 ± 0.01	-10.250 ± 0.007	-9.89 ± 0.02	-9.87 ± 0.01	-10.40 ± 0.01	-9.92^d
ρ (g/cm ³)	0.997 ± 0.001	0.998 ± 0.001	1.014 ± 0.002	0.999 ± 0.001	1.02 ± 0.01	0.997^d
μ (D)	2.712 ± 0.002	2.791 ± 0.002	2.6	2.29	3.01 ± 0.01	
ϵ_0	98±8	105±7	79±8	82±2		78.3 ^e
ϵ_∞	1.689 ± 0.001	1.708 ± 0.001	1.592 ± 0.003	1		1.79 ^f
D (10 ⁻⁹ m ² /s)	1.81 ± 0.06	1.25 ± 0.05	1.9 ± 0.1	2.62 ± 0.04^g		2.3 ^h
τ_{NMR} (ps)	2.6 ± 0.1	4.0 ± 0.2	2.1 ± 0.1	1.4 ± 0.1		2.1 ⁱ

^aReference 28.^bReference 11.^cReference 41.^dReference 83.^eReference 108.^fReference 105.^gReference 52.^hReference 106.ⁱReference 107.

$$\epsilon_\infty = 1 + \frac{4\pi}{3V} \text{Tr}\langle A \rangle, \quad (21)$$

where A is the polarizability tensor of the system.⁶⁶ While both POL5/TZ and POL5/QZ have infinite-frequency dielectric constants in good agreement with experiment, their static dielectric constants are slightly too high.

The translational diffusion constant was computed from the Einstein relation:

$$D = \lim_{t \rightarrow \infty} \frac{1}{6t} \langle |\vec{r}_i(t) - \vec{r}_i(0)|^2 \rangle, \quad (22)$$

where \vec{r}_i is the position of the center of mass of molecule i , and the average is taken over all molecules. Averages were taken from 20 separate trajectories, run in the NVE ensemble, starting from initial configurations sampled from the canonical distribution at 298.15 and 0.997 g/cm³. For each trajectory, a line was fit to the average mean-square displacement in the range from 3 to 10 ps, and the slope averaged over all trajectories. A plot of the mean-square displacement for one trajectory of the POL5/TZ model is shown in Fig. 9.

The nuclear magnetic resonance relaxation time was calculated from the zero-frequency component of the Fourier transform of the rotational correlation function

$$C_2^y(t) = \langle P_2[\hat{e}_i^y(0) \cdot \hat{e}_i^y(t)] \rangle, \quad (23)$$

where \hat{e}_i^y is the unit vector pointing along the axis connecting the hydrogens (the y axis) of molecule i , $P_2(x) = (3x^2 - 1)/2$ is the second Legendre polynomial, and the average as before is over all molecules. The zero-frequency part of the Fourier transform of $C_2^y(t)$ is given to a good approximation by $A_2^y \tau_2^y$, where $A_2^y \exp(-t/\tau_2^y)$ is an exponential fit to the long-range behavior of $C_2^y(t)$.²⁸ Values from twenty multiple trajectories were averaged in the same manner as described for the translational diffusion constant. A plot of the rotational correlation function for one trajectory of the POL5/TZ model is shown in Fig. 10. We also calculated τ_{NMR} for the TIP5P model of Mahoney and Jorgensen.¹¹

The dynamics of the POL5 models are in reasonable qualitative agreement with experiment, but are somewhat too slow, especially the POL5/QZ model. As with the overesti-

mation of the dielectric constant, the slow dynamics are most likely the result of a liquid-state dipole moment that is too large.

Radial distribution functions (RDFs) for the model and the most recent experiments of the Soper lab⁶⁷ at room temperature and pressure are shown in Figs. 11–13. Extracting RDFs from neutron or x-ray diffraction data is by no means a straightforward process and published results have varied somewhat over time.^{67–69} However, as shown in Fig. 14, Soper's most recent analysis of neutron diffraction data agrees quite well with the analysis of x-ray diffraction experiments from the Advanced Light Source (ALS) at Lawrence Berkeley National Laboratory from Glaeser and Head-Gordon,^{70,71} as well as the TIP5P model, which was parametrized completely independently, suggesting that the newest g_{OO} is quite reliable. The coordination number can be determined from the RDF by integrating $g_{\text{OO}}(r)$ over the first peak. Using the location of the first minimum in the experimental curve (3.36 Å) as the limit of integration, we

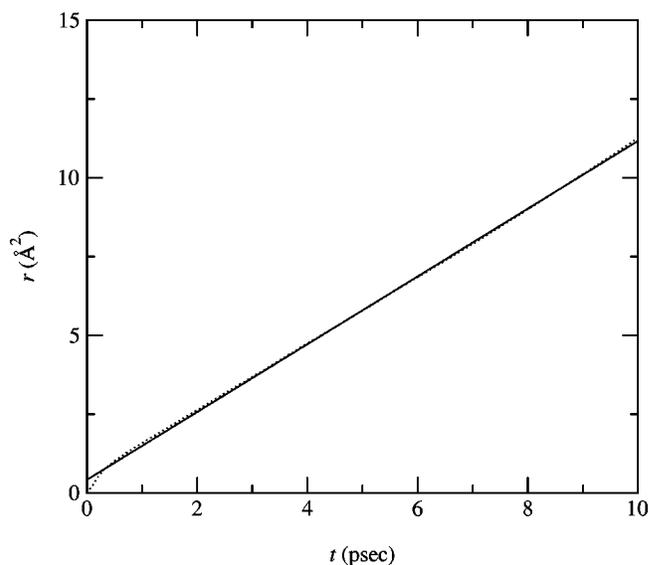


FIG. 9. Average root-mean-square displacement for a trajectory of the POL5/TZ model (dotted line) and the line of best fit in the range from 3 to 10 ps (solid line).

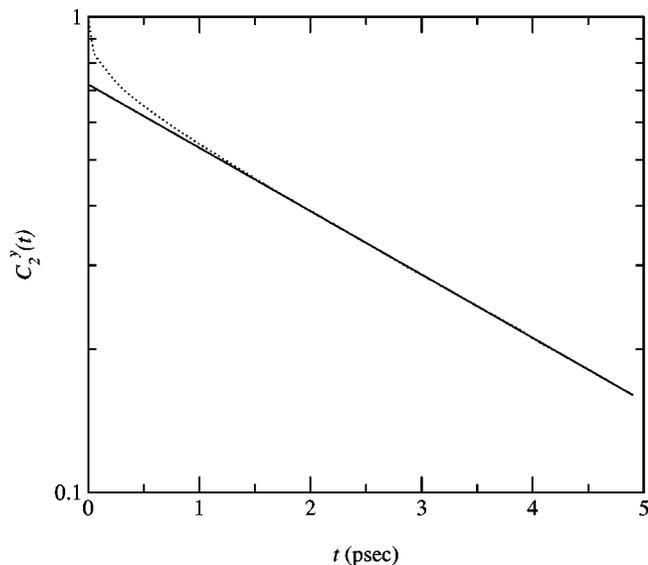


FIG. 10. Rotational correlation function for a trajectory of the POL5/TZ model (dotted line) and the exponential fit in the range of 2–5 ps (solid line).

obtain a coordination number of 4.5 for both versions of the model and for experiment. Agreement with the experimental RDFs is good but not perfect. The first peak for both POL5 models is slightly too short and too broad. POL5/TZ appears slightly understructured in that the first trough is not deep enough, nor is the second peak high enough; all features are also shifted very slightly out too far. POL5/QZ gives the correct depth for the first trough (although it, too, is shifted out slightly); however it appears slightly overstructured as the second peak, while at the correct location, is a bit too high.

There is at present some controversy in the literature about the “correct” value of the liquid-state dipole moment. Coulson and Eisenberg obtained a value of 2.6 D for ice I_h .⁷² However, Batista *et al.* repeated their analysis using

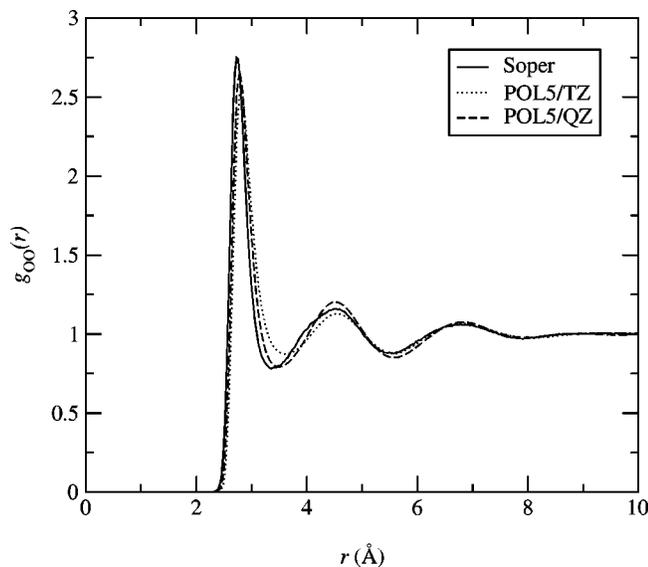


FIG. 11. Oxygen–oxygen radial distribution function $g_{OO}(r)$ at 298.15 K, 1 atm for the POL5 potential compared with the latest results of Soper.

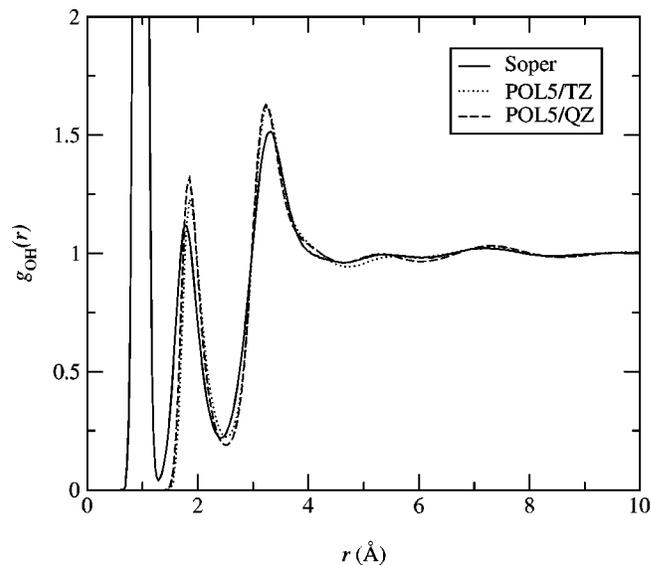


FIG. 12. Oxygen–hydrogen radial distribution function $g_{OH}(r)$ at 298.15 K, 1 atm for the POL5 potential compared with the latest results of Soper.

more recent input data and obtained a value of 3.09 D.⁷³ A recent analysis of x-ray diffraction data by the Soper group, the first experimental study of the average dipole moment in liquid water, inferred a value of 2.9 D at ambient conditions.⁷⁴ The dipole moment of most water models has tended to be somewhat lower. Sprik has observed that polarizable models need a liquid-state dipole moment close to 2.6 D in order to reproduce ϵ_0 ; the dielectric constant will be too large if this is not the case.¹⁴ For example, ϵ_0 for the NCC *ab initio* model, which has a liquid-state dipole moment of 2.7–2.8 D, was determined to be around 100.⁷⁵ Our data are clearly in agreement with this hypothesis. However, Silvestrelli and Parrinello have suggested that the correct value for the liquid-state dipole moment is somewhat larger, around 3.0 D, based on *ab initio* molecular dynamics simulations.^{4,76}

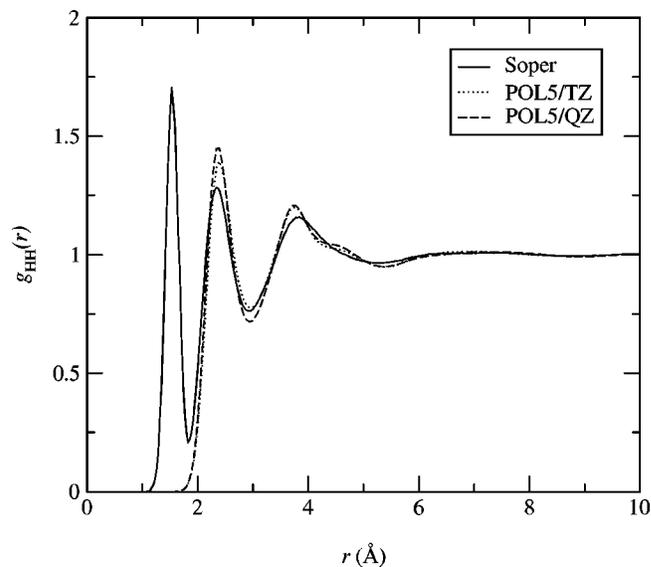


FIG. 13. Hydrogen–hydrogen radial distribution function $g_{HH}(r)$ at 298.15 K, 1 atm for the POL5 potential compared with the latest results of Soper.

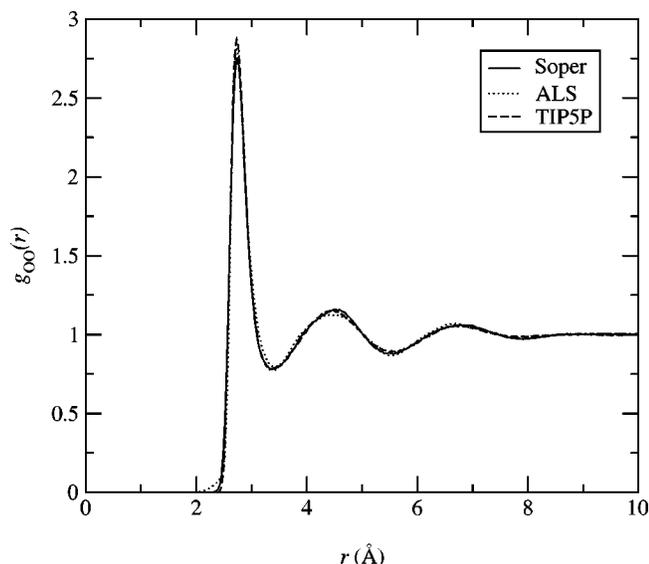


FIG. 14. Comparison of the oxygen–oxygen radial distribution functions $g_{OO}(r)$ from analysis of neutron data by Soper, analysis of x-ray data from the Advanced Light Source (ALS) at Lawrence Berkeley National Laboratory by Glaeser and Head-Gordon, and the TIP5P model.

The MCDHO *ab initio* model also predicts a moment of 3.0 D. However, dielectric constants have not been computed either for this model or for *ab initio* molecular dynamics simulations, although Silvestrelli and Parrinello have reported a rough estimate of 86.⁴ Furthermore, it is unclear whether a molecular dipole moment can even be unambiguously defined for an *ab initio* wave function, given that the calculated value depends strongly on the specific method used to partition the continuous charge distribution into molecules.⁷⁷

Given that *ab initio* models are unable to “parametrize away” quantum effects, it seems possible that large dipole moments and associated shortcomings such as too much structure, too high a heat of vaporization and dielectric constant, and too slow dynamics are artifacts of the classical treatment of the nuclear motion, and the “true” value for the dipole moment is somewhat closer to 2.6 D. While no quantum simulations of polarizable water appear to have been performed, path-integral simulations⁷⁸ of fixed-charge water models, both rigid⁷⁹ and flexible,^{80,81} indicate that for a given model, the quantum liquid is less structured, has a lower dielectric constant, and has faster dynamics (according to the approximate centroid molecular dynamics method). In particular, structural changes due to quantum effects were reported to be as large as those due to raising the temperature by 50 K,⁷⁹ dielectric constants were smaller for the quantum liquid by 15%–20%,⁸¹ and diffusion constants larger by 40%–70%.⁸¹ These changes would certainly bring the predictions of the POL5/QZ model better in line with experiment. Other explanations for a too-large dipole moment are also possible: For instance, that due to Pauli repulsion between molecules, it is harder to distort their electron clouds in the condensed phase, so that the polarizability for models in condensed phase should be somewhat smaller than the gas-phase value. This argument is supported by *ab initio* calculations of Morita and Kato,⁸² who found that the calcu-

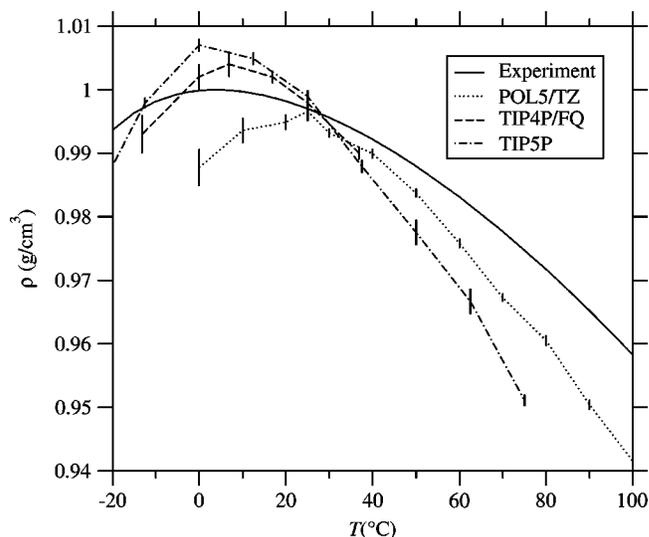


FIG. 15. Water density ρ at 1 atm as a function of temperature for POL5/TZ, TIP4P/FQ, TIP5P, and experiment.

lated polarizabilities of neutral species in water were decreased as much as 13%–18% from their values in vacuum. Whatever the cause, the slightly smaller basis set used to parametrize the POL5/TZ model produces a smaller polarizability and hence a smaller liquid-phase dipole moment, improving the dielectric constant and diffusion constant.

D. Liquid at other conditions

As POL5/TZ appears to be a slightly better model for the condensed phase than POL5/QZ (at least for classical simulations), we chose the former to perform additional water simulations at nonphysiological conditions. Water density at 1 atm as a function of temperature for POL5/TZ is shown in Fig. 15. Experimental data are from Ref. 83. As is well known, the density of cold water changes with temperature in a manner opposite to that of almost all other substances. As the open, tetrahedral structure breaks down, the density *increases* with temperature, leading to a maximum in the density at 4 °C. A density maximum is observed for POL5/TZ, but it is located at too high a temperature, 20–25 °C. Furthermore, the density decreases too sharply as the temperature is lowered. These discrepancies with experiment are similar to those observed with the ST2 model.⁶ The TIP5P model was parametrized to produce a density maximum at the correct location; however, the maximum is somewhat too sharp as well. In particular, the density decreases too rapidly as the temperature increases, yielding overestimates of the coefficient of thermal expansion. The density of the POL5/TZ model also decreases somewhat too rapidly, but less so than TIP5P, so that agreement with the experimental density from 0 to 100 °C is quite good, with maximum deviation less than 2%. The TIP4P/FQ model has a density maximum at the correct location; as with the other models, the maximum is too sharp.³²

The heat of vaporization ΔH_{vap} as a function of temperature is shown in Fig. 16. Experimental data are from Ref. 84. For all models, ΔH_{vap} decreases too quickly with increasing temperature. This is reflected in the constant-pressure heat

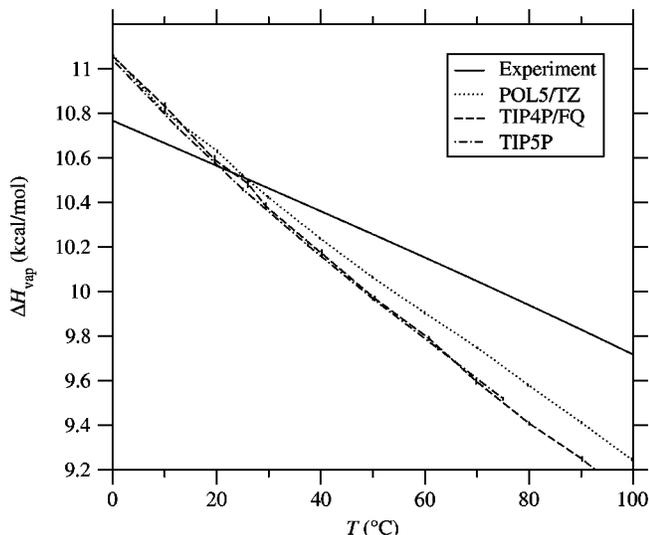


FIG. 16. Water heat of vaporization ΔH_{vap} at 1 atm as a function of temperature for POL5/TZ, TIP4P/FQ, TIP5P, and experiment.

capacity, $C_p = (\partial H / \partial T)_p$, which was computed using a centered-difference approximation and is shown in Fig. 17, along with experimental data from Refs. 84 and 85. Agreement with the experimental heat capacity is slightly better for the POL5/TZ model than for TIP4P/FQ or TIP5P. However, the experimental constancy of C_p from 0 to 100° is not well-reproduced by any of the models.

Radial distribution functions at several other thermodynamic state points were computed for the POL5/TZ model. These are plotted and compared with the latest results from Soper⁶⁷ in Figs. 18–20. In general the model agrees quite well with experiment. However, it is apparent that the model somewhat underpredicts the ability of the hydrogen-bond network to withstand increases in temperature and pressure. A well-defined first peak in the OH curve persists at all temperatures and pressures in the experimental curve. This peak is too small in the model at 423 K and is not well-defined at

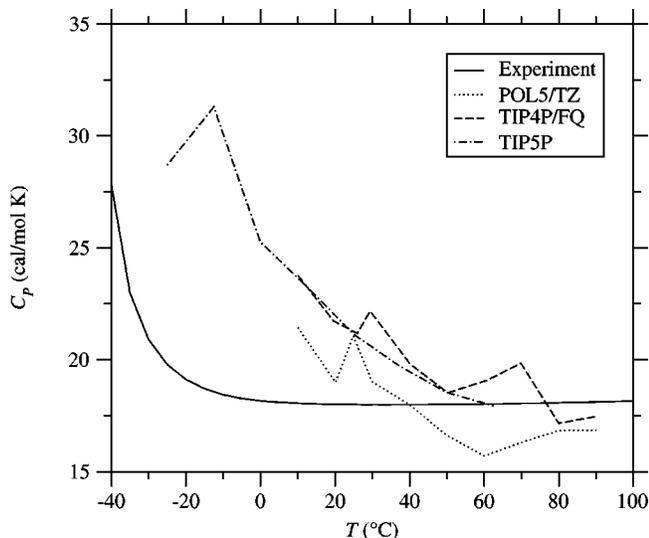


FIG. 17. Water constant-pressure heat capacity C_p at 1 atm as a function of temperature for POL5/TZ, TIP4P/FQ, TIP5P, and experiment.

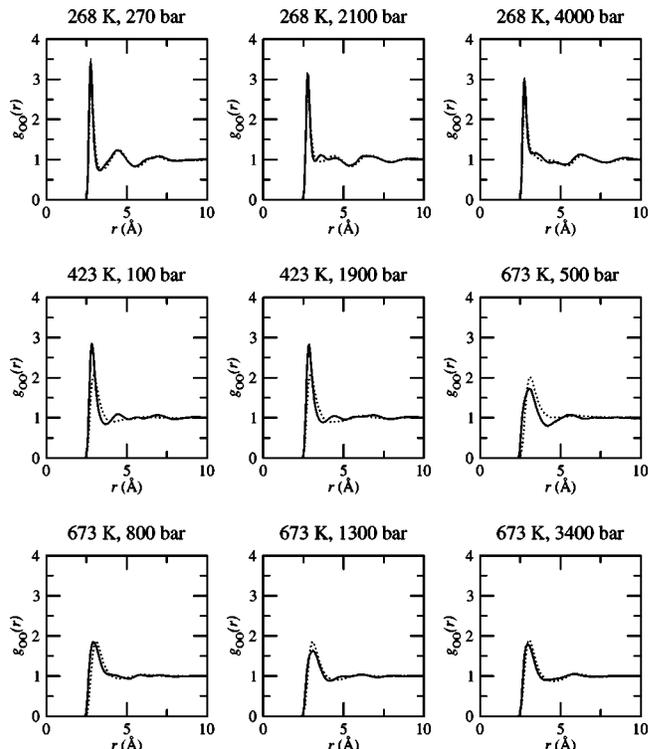


FIG. 18. Oxygen–oxygen radial distribution function $g_{\text{OO}}(r)$ at various thermodynamic states for POL5/TZ (dotted line) compared with the latest results of Soper (solid line).

673 K, becoming more of a shoulder. The lack of structure appears in the OO curve as well; at 423 K at both pressures and 673 K at the lowest pressure simulated (500 bar), the model RDF is almost featureless after the first peak, while

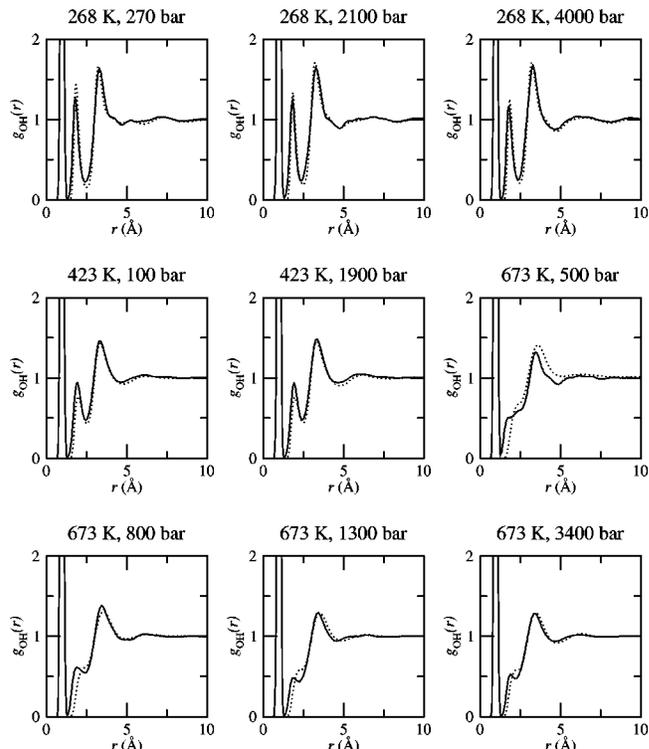


FIG. 19. Oxygen–hydrogen radial distribution function $g_{\text{OH}}(r)$ at various thermodynamic states for POL5/TZ (dotted line) compared with the latest results of Soper (solid line).

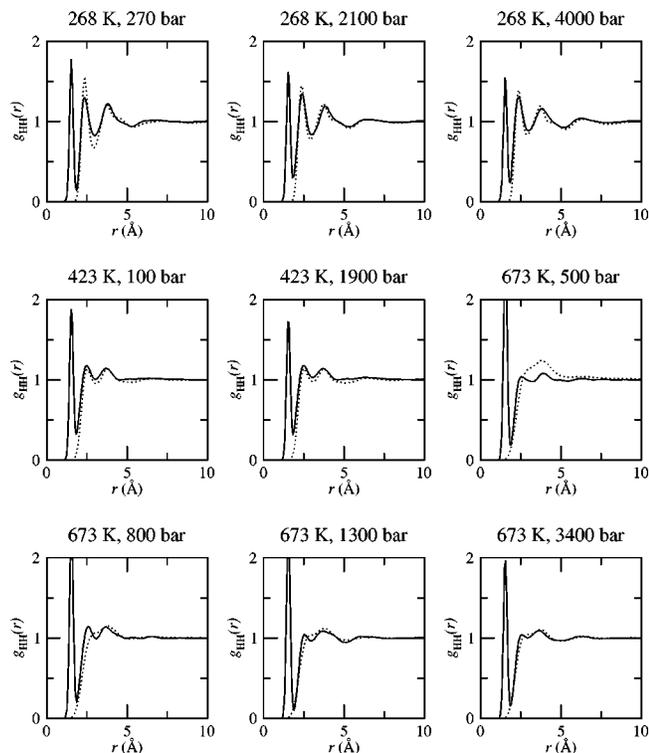


FIG. 20. Hydrogen–hydrogen radial distribution function $g_{\text{HH}}(r)$ at various thermodynamic states for POL5/TZ (dotted line) compared with the latest results of Soper (solid line).

the experimental curve shows remnants of tetrahedral order. Agreement is better at higher pressures, where both experiment and model show little or no ordering. Underprediction of the persistence of the hydrogen-bond network is consistent with a too-rapid decrease in density with increasing temperature. We also computed RDFs at these state points with the TIP4P/FQ and TIP5P models; the results are very similar, and are not shown here, but are available in electronic form.¹⁰⁹

V. CONCLUSIONS

We have presented a general formalism for polarizable electrostatics based on fluctuating bond-charge increments and polarizable dipoles, and applied it to a five-site model for water with a geometry similar to the ST2 and TIP5P models. The parametrization procedure, based largely on quantum-chemical calculations, is systematic, and should be easily transferable to other molecules. To examine basis-set effects we performed the fitting procedure with both the aug-cc-pVTZ and aug-cc-pVQZ basis sets. We computed several properties for both models, in both the gas phase and condensed phase, including the optimal dimer structure, binding energies of clusters, and the energy, density, dielectric constant, and diffusion constant of liquid water at ambient conditions. Only the models that were parametrized (at least in part) from *ab initio* data were able to capture the crossover point at which noncyclic conformations of clusters become more stable than cyclic ones. Predictions of condensed-phase properties from the model fit to DFT/aug-cc-pVQZ calculations are in reasonable accord with experiment, but show a

heat of vaporization and dielectric constant that are too large, a diffusion constant that is too small, and somewhat too much structure in the radial distribution functions. These errors are consistent with a liquid-state dipole moment that is too large, and are perhaps an artifact of running classical simulations on what is essentially an *ab initio* potential, or the result of the gas-phase polarizability being inappropriate for the condensed phase due to Pauli exclusion. Though the liquid-phase dipole moment is still slightly too high, the model fit to DFT/aug-cc-pVTZ calculations gives good results at ambient conditions and performs well away from room temperature, deviating less than 2% from the experimental density from 0 to 100°C, and showing good agreement with experimental radial distribution functions over a wide range of thermodynamic state points.

ACKNOWLEDGMENTS

This work was funded by the NIH under Grant Nos. GM52018 (R.A.F.) and GM43340 (B.J.B.) and by the NSF under Grant No. CHE-00-76279 (B.J.B.). The computational resources of the NIH Biomedical Technology Resource Center at Columbia University are gratefully acknowledged.

- ¹R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).
- ²D. Marx, M. Sprik, and M. Parrinello, Chem. Phys. Lett. **273**, 360 (1997).
- ³K. Laasonen, M. Sprik, M. Parrinello, and R. Car, J. Chem. Phys. **99**, 9080 (1993).
- ⁴P. L. Silvestrelli, M. Bernasconi, and M. Parrinello, Chem. Phys. Lett. **277**, 478 (1997).
- ⁵M. Diraison, G. J. Martyna, and M. E. Tuckerman, J. Chem. Phys. **111**, 1096 (1999).
- ⁶F. H. Stillinger and A. Rahman, J. Chem. Phys. **60**, 1545 (1974).
- ⁷H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans, in *Intermolecular Forces*, edited by B. Pullman (Reidel, Dordrecht, 1981), p. 331.
- ⁸W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
- ⁹H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, J. Phys. Chem. **91**, 6269 (1987).
- ¹⁰K. Watanabe and M. L. Klein, Chem. Phys. **131**, 157 (1989).
- ¹¹M. W. Mahoney and W. L. Jorgensen, J. Chem. Phys. **112**, 8910 (2000).
- ¹²J. A. C. Rullmann and P. T. van Duijnen, Mol. Phys. **63**, 451 (1988).
- ¹³J. Caldwell, L. X. Dang, and P. A. Kollman, J. Am. Chem. Soc. **112**, 9144 (1990).
- ¹⁴M. Sprik, J. Chem. Phys. **95**, 6762 (1991).
- ¹⁵L. X. Dang, J. E. Rice, J. Caldwell, and P. A. Kollman, J. Am. Chem. Soc. **113**, 2481 (1991).
- ¹⁶D. van Belle, M. Froeyen, G. Lippens, and S. J. Wodak, Mol. Phys. **77**, 239 (1992).
- ¹⁷A. Wallqvist and B. J. Berne, J. Phys. Chem. **97**, 13841 (1993).
- ¹⁸S.-B. Zhu and C. F. Wong, J. Phys. Chem. **98**, 4695 (1994).
- ¹⁹D. N. Bernardo, Y. Ding, K. Krogh-Jespersen, and R. M. Levy, J. Phys. Chem. **98**, 4180 (1994).
- ²⁰R. D. Mountain, J. Chem. Phys. **103**, 3084 (1995).
- ²¹P. Ahlström, A. Wallqvist, S. Engström, and B. Jönsson, Mol. Phys. **68**, 563 (1989).
- ²²Y. Ding, D. N. Bernardo, K. Krogh-Jespersen, and R. M. Levy, J. Phys. Chem. **99**, 11575 (1995).
- ²³L. X. Dang and T.-M. Chang, J. Chem. Phys. **106**, 8149 (1997).
- ²⁴L. X. Dang, J. Phys. Chem. B **102**, 620 (1998).
- ²⁵N. Yoshii, H. Yoshie, S. Miura, and S. Okazaki, J. Chem. Phys. **109**, 4873 (1998).
- ²⁶S.-B. Zhu, S. Yao, J.-B. Zhu, S. Singh, and G. W. Robinson, J. Phys. Chem. **95**, 6211 (1991).
- ²⁷S.-B. Zhu, S. Singh, and G. W. Robinson, J. Chem. Phys. **95**, 2791 (1991).
- ²⁸S. W. Rick, S. J. Stuart, and B. J. Berne, J. Chem. Phys. **101**, 6141 (1994).
- ²⁹I. M. Svishchev, P. G. Kusalik, J. Wang, and R. J. Boyd, J. Chem. Phys. **105**, 4742 (1996).

- ³⁰B. Chen, J. Xing, and J. I. Siepmann, *J. Phys. Chem. B* **104**, 2391 (2000).
- ³¹B. Chen, J. J. Potoff, and J. I. Siepmann, *J. Phys. Chem. B* **104**, 2378 (2000).
- ³²S. W. Rick, *J. Chem. Phys.* **114**, 2276 (2001).
- ³³G. C. Lie, E. Clementi, and M. Yoshimine, *J. Chem. Phys.* **64**, 2314 (1976).
- ³⁴O. Matsuoka, E. Clementi, and M. Yoshimine, *J. Chem. Phys.* **64**, 1351 (1976).
- ³⁵U. Niesar, G. Corongiu, E. Clementi, G. R. Kneller, and D. K. Bhattacharya, *J. Phys. Chem.* **94**, 7949 (1990).
- ³⁶C. Millot and A. J. Stone, *Mol. Phys.* **77**, 439 (1992).
- ³⁷P.-O. Åstrand, P. Linse, and G. Karlström, *Chem. Phys.* **191**, 195 (1995).
- ³⁸E. M. Mas, K. Szalewicz, R. Bukowski, and B. Jeziorski, *J. Chem. Phys.* **107**, 4207 (1997).
- ³⁹Y.-P. Liu, K. Kim, B. J. Berne, R. A. Friesner, and S. W. Rick, *J. Chem. Phys.* **108**, 4739 (1998).
- ⁴⁰C. Millot, J.-C. Soetens, M. T. C. M. Costa, M. P. Hodges, and A. J. Stone, *J. Phys. Chem. A* **102**, 754 (1998).
- ⁴¹H. Saint-Martin, J. Hernández-Cobos, M. I. Bernal-Uruchurtu, I. Ortega-Blake, and H. J. C. Berendsen, *J. Chem. Phys.* **113**, 10899 (2000).
- ⁴²J. L. Banks, G. A. Kaminski, R. Zhou, D. T. Mainz, B. J. Berne, and R. A. Friesner, *J. Chem. Phys.* **110**, 741 (1999).
- ⁴³H. A. Stern, G. A. Kaminski, J. L. Banks, R. Zhou, B. J. Berne, and R. A. Friesner, *J. Phys. Chem. B* **103**, 4730 (1999).
- ⁴⁴R. Chelli, P. Procacci, R. Righini, and S. Califano, *J. Chem. Phys.* **111**, 8569 (1999).
- ⁴⁵M. Sprik and M. L. Klein, *J. Chem. Phys.* **89**, 7556 (1988).
- ⁴⁶M. Sprik, *J. Phys. Chem.* **95**, 2283 (1991).
- ⁴⁷D. York and W. Yang, *J. Chem. Phys.* **104**, 159 (1996).
- ⁴⁸A. K. Rappé and W. A. Goddard, *J. Phys. Chem.* **95**, 3358 (1991).
- ⁴⁹B. T. Thole, *Chem. Phys.* **59**, 341 (1981).
- ⁵⁰P. T. V. Duijnen and M. Swart, *J. Phys. Chem. A* **102**, 2399 (1998).
- ⁵¹C. J. Burnham, J. Li, S. S. Xantheas, and M. Leslie, *J. Chem. Phys.* **110**, 4566 (1999).
- ⁵²M. W. Mahoney and W. L. Jorgensen, *J. Chem. Phys.* **114**, 363 (2001).
- ⁵³R. A. Buckingham and J. Corner, *Proc. R. Soc. London, Ser. A* **189**, 118 (1947).
- ⁵⁴A. D. Becke, *Phys. Rev. A* **38**, 3098 (1988).
- ⁵⁵C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988).
- ⁵⁶JAGUAR v3.0 Schrödinger, Inc., Portland, OR (1997).
- ⁵⁷A. Halkier, H. Koch, P. Jørgensen, O. Christiansen, I. M. B. Nielsen, and T. Helgaker, *Theor. Chem. Acc.* **97**, 150 (1997).
- ⁵⁸A. Halkier, W. Klopper, T. Helgaker, P. Jørgensen, and P. R. Taylor, *J. Chem. Phys.* **111**, 9157 (1999).
- ⁵⁹G. J. Martyna, M. L. Klein, and M. E. Tuckerman, *J. Chem. Phys.* **97**, 2635 (1992).
- ⁶⁰G. J. Martyna, D. J. Tobias, and M. L. Klein, *J. Chem. Phys.* **101**, 4177 (1994).
- ⁶¹W. Smith, *CCP5 Newsletter* **46**, 18 (1998).
- ⁶²M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Clarendon, Oxford, 1987).
- ⁶³G. J. Martyna, A. Hughes, and M. E. Tuckerman, *J. Chem. Phys.* **110**, 3275 (1999).
- ⁶⁴A. Toukmaji, C. Sagui, J. Board, and T. Darden, *J. Chem. Phys.* **113**, 10913 (2000).
- ⁶⁵J. K. Gregory, D. C. Clary, K. Liu, M. G. Brown, and R. J. Saykally, *Science* **275**, 814 (1997).
- ⁶⁶M. Neumann and O. Steinhauser, *Chem. Phys. Lett.* **106**, 563 (1984).
- ⁶⁷A. K. Soper, *Chem. Phys.* **258**, 121 (2000).
- ⁶⁸A. K. Soper and M. G. Phillipps, *Chem. Phys.* **107**, 47 (1986).
- ⁶⁹A. K. Soper, F. Bruni, and M. A. Ricci, *J. Chem. Phys.* **247**, 106 (1997).
- ⁷⁰G. Hura, J. M. Sorenson, R. M. Glaeser, and T. Head-Gordon, *J. Chem. Phys.* **113**, 9140 (2000).
- ⁷¹J. M. Sorenson, G. Hura, R. M. Glaeser, and T. Head-Gordon, *J. Chem. Phys.* **113**, 9149 (2000).
- ⁷²C. A. Coulson and D. Eisenberg, *Proc. R. Soc. London, Ser. A* **291**, 445 (1966).
- ⁷³E. R. Batista, S. S. Xantheas, and H. Jónsson, *J. Chem. Phys.* **109**, 4546 (1998).
- ⁷⁴Y. S. Badyal, M.-L. Saboungi, D. L. Price, S. D. Shastri, D. R. Haefner, and A. K. Soper, *J. Chem. Phys.* **112**, 9206 (2000).
- ⁷⁵J.-C. Soetens, M. T. C. M. Costa, and C. Millot, *Mol. Phys.* **94**, 577 (1998).
- ⁷⁶P. L. Silvestrelli and M. Parrinello, *Phys. Rev. Lett.* **82**, 3308 (1999).
- ⁷⁷E. R. Batista, S. S. Xantheas, and H. Jónsson, *J. Chem. Phys.* **111**, 6011 (1999).
- ⁷⁸R. P. Feynman, *Statistical Mechanics* (Addison-Wesley, Reading, MA, 1972).
- ⁷⁹R. A. Kuharski and P. J. Rossky, *J. Chem. Phys.* **82**, 5164 (1985).
- ⁸⁰A. Wallqvist and B. J. Berne, *Chem. Phys. Lett.* **117**, 214 (1985).
- ⁸¹J. Lobaugh and G. A. Voth, *J. Chem. Phys.* **106**, 2400 (1997).
- ⁸²A. Morita and S. Kato, *J. Chem. Phys.* **110**, 11987 (1999).
- ⁸³G. J. Kell, *J. Chem. Eng. Data* **20**, 97 (1975).
- ⁸⁴J. A. Riddick, W. B. Bunger, and T. K. Sakano, *Techniques of Chemistry, Vol. II: Organic Solvents, Physical Properties and Methods of Purification*, 4th ed. (Wiley, New York, 1986).
- ⁸⁵R. J. Speedy, *J. Phys. Chem.* **91**, 3354 (1987).
- ⁸⁶S. A. Clough, Y. Beers, G. P. Klein, and L. S. Rothman, *J. Chem. Phys.* **59**, 2254 (1973).
- ⁸⁷J. Verhoeven and A. Dymanus, *J. Chem. Phys.* **52**, 3222 (1970).
- ⁸⁸W. F. Murphy, *J. Phys. Chem.* **67**, 5877 (1977).
- ⁸⁹J. A. Odutola and T. R. Dyke, *J. Chem. Phys.* **72**, 5062 (1980).
- ⁹⁰J. A. Odutola, T. A. Hu, D. Prinslow, S. E. O'Dell, and T. R. Dyke, *J. Chem. Phys.* **88**, 5352 (1988).
- ⁹¹R. Krishnan, J. S. Binkley, R. Seeger, and J. A. Pople, *J. Chem. Phys.* **72**, 650 (1980).
- ⁹²T. Clark, J. Chandrasekhar, and P. v. R. Schleyer, *J. Comput. Chem.* **4**, 294 (1983).
- ⁹³N. Pugliano and R. J. Saykally, *Science* **257**, 1937 (1992).
- ⁹⁴M. R. Viant, J. D. Cruzan, D. D. Lucas, M. G. Brown, K. Liu, and R. J. Saykally, *J. Phys. Chem. A* **101**, 9032 (1997).
- ⁹⁵K. Liu, M. G. Brown, J. D. Cruzan, and R. J. Saykally, *Science* **271**, 62 (1996).
- ⁹⁶K. Liu, M. G. Brown, J. D. Cruzan, and R. J. Saykally, *J. Phys. Chem. A* **101**, 9011 (1997).
- ⁹⁷J. Kim and K. S. Kim, *J. Chem. Phys.* **109**, 5886 (1999).
- ⁹⁸S. S. Xantheas and J. T. H. Dunning, *J. Chem. Phys.* **98**, 8037 (1993).
- ⁹⁹K. Liu, M. G. Brown, C. Carter, R. J. Saykally, J. K. Gregory, and D. C. Clary, *Nature (London)* **381**, 501 (1996).
- ¹⁰⁰K. Liu, M. G. Brown, and R. J. Saykally, *J. Phys. Chem. A* **101**, 8995 (1997).
- ¹⁰¹I. M. B. Nielsen, E. T. Seidl, and C. L. Janssen, *J. Chem. Phys.* **110**, 9435 (1999).
- ¹⁰²J. D. Cruzan, L. B. Braly, K. Liu, M. G. Brown, J. G. Loeser, and R. J. Saykally, *Science* **271**, 59 (1996).
- ¹⁰³J. D. Cruzan, M. R. Viant, M. G. Brown, and R. J. Saykally, *J. Phys. Chem. A* **101**, 9022 (1997).
- ¹⁰⁴*CRC Handbook of Chemistry and Physics*, edited by D. R. Lide (Chemical Rubber Publishing Co., Boca Raton, FL, 1995).
- ¹⁰⁵A. D. Buckingham, *Proc. R. Soc. London, Ser. A* **238**, 235 (1956).
- ¹⁰⁶K. Krynicki, C. D. Green, and D. W. Sawyer, *Faraday Discuss. Chem. Soc.* **66**, 199 (1978).
- ¹⁰⁷J. Jonas, T. DeFries, and D. J. Wilber, *J. Chem. Phys.* **65**, 582 (1976).
- ¹⁰⁸D. Bertolini, M. Cassettari, and G. Salvetti, *J. Chem. Phys.* **76**, 3285 (1982).
- ¹⁰⁹See EPAPS Document No. E-JCPA6-115-503125 for radial distribution functions for the TIP4P/FQ and TIP5P water models (dotted line) versus experimental results from the Soper lab, at various thermodynamic states. This document may be retrieved via the EPAPS homepage (<http://www.aip.org/pubservs/epaps.html>) or from <ftp.aip.org> in the directory /epaps/. See the EPAPS homepage for more information.