Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew

Hans W. Horn,[a] William C. Swope, and Jed W. Pitera
IBM Almaden Research Center, San Jose, California 95120

Jeffry D. Madura and Thomas J. Dick
Department of Chemistry and Biochemistry, Center for Computational Sciences, Duquesne University, Pittsburgh, Pennsylvania 15282

Greg L. Hura
Graduate Group in Biophysics, University of California, Berkeley, California 94720

Teresa Head-Gordon
Graduate Group in Biophysics and Department of Bioengineering, University of California, Berkeley, California 94720

(Received 1 December 2003; accepted 22 January 2004)

A re-parameterization of the standard TIP4P water model for use with Ewald techniques is introduced, providing an overall global improvement in water properties relative to several popular nonpolarizable and polarizable water potentials. Using high precision simulations, and careful application of standard analytical corrections, we show that the new TIP4P-Ew potential has a density maximum at ~1 °C, and reproduces experimental bulk-densities and the enthalpy of vaporization, ΔH_{vap}, from −37.5 to 127 °C at 1 atm with an absolute average error of less than 1%. Structural properties are in very good agreement with x-ray scattering intensities at temperatures between 0 and 77 °C and dynamical properties such as self-diffusion coefficient are in excellent agreement with experiment. The parameterization approach used can be easily generalized to rehabilitate any water force field using available experimental data over a range of thermodynamic points. © *2004 American Institute of Physics.* [DOI: 10.1063/1.1683075]

I. INTRODUCTION

The goal of bio-molecular simulations is the accurate and predictive computer simulation of the physical properties of biological molecules in their aqueous environments. There are three main issues regarding the treatment of water in such simulations. The first is the accurate description of the protein–water interaction. While continuum models have shown some promise,[1] they cannot reproduce the microscopic details of the protein–water interface.[2] Consequently, most bio-molecular simulations are carried out with the solute surrounded by a droplet or periodic box of explicit water molecules. In a typical case, these water molecules will account for over 80% of the particles in the simulation. Water–water interactions dominate the computational cost of such simulations, so the model used to describe the water needs to be fast as well as accurate.

The second major issue is that proteins and nucleic acids are typically highly charged. This means that long-range Coulombic interactions need to be properly accounted for, whether by using Ewald summation techniques,[3] reaction field methods,[4] or other equivalents. Prior to the mid-1990s, most bio-molecular simulations used some sort of truncated Coulomb interaction to improve the speed of the simulation which greatly compromised accuracy. Algorithmic developments,[5,6] as well as continuing improvements in computer hardware, have made inclusion of long-range electrostatic effects ubiquitous in modern simulations. While the simulation methodology has improved, these simulations are typically carried out using water models that were originally parameterized using a truncated Coulomb interaction. Using these models with Ewald summation results in changes in both thermodynamic and kinetic properties: Densities are lower than in the original parameterization (Fig. 1); diffusion constants[7,8] are larger. Some attempt has been made to reparameterize the SPC and TIP4P models for use with a reaction field, but these variants are not widely used.[9]

The final issue is the area of comparison with experiment. In this case, it is critical that the model give good thermodynamic (structural observables, solvation free energies, etc.) and kinetic (diffusion, rotational correlation times, hydrogen bond dynamic, etc.) results for both neat liquids and solutions. In many cases, one is interested in a *temperature dependent* property of the bio-molecule, such as a melting curve[10] or structural fluctuations.[11] Water models in common use in bio-molecular simulation, however, have traditionally only been parameterized for a single temperature (~298 K: SPC,[12] SPC/E,[13] TIP3P/4P,[14] SPC/AL,[15] SPC/S[16]). More recent fixed charge (TIP5P[17]) and polarizable (SPC-Pol and TIP4P-Pol[18]) models have been developed using data over a range of temperatures.

Polarizable variants of various water models[19–24] have been introduced to overcome certain short-comings of their nonpolarizable ancestors. Although we anticipate that these

[a]Electronic mail: hans@almaden.ibm.com

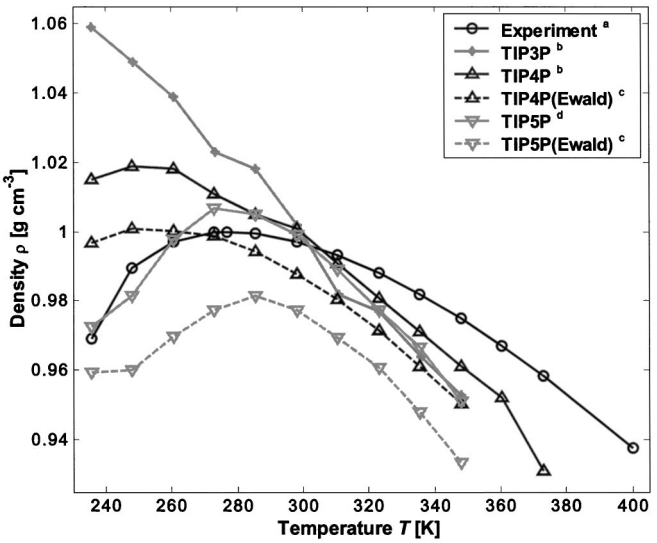

FIG. 1. Bulk-density $\rho(T)$ for conventional 3-, 4- and 5-site water models with and without Ewald. (a) From Refs. 40 (liquid) and 41 (supercooled and superheated liquid); (b) from Ref. 14; (c) from Ref. 8, no Lennard-Jones tail correction; $\rho(T)$ has a maximum at 256 ± 1 K [from seventh-order least-squares polynomial fit, Eq. (16)]; (d) from Ref. 17.

polarizable force fields will improve the accuracy of water potentials applicability for phase equilibria, mixtures, surface properties, dynamics, etc., they are much more expensive and thus implementations of polarizable models are entering main-stream bio-molecular simulation codes only slowly.

A good water model should not only reproduce bulk water properties over a range of thermodynamic states, but also work in concert with protein force fields to reproduce solution and interfacial thermodynamic and kinetic properties. Three-site water models are the most commonly used water potentials for protein–water simulations. Furthermore, common wisdom might suggest that the current generation of protein force fields are somewhat ''tuned'' for use with these three-site models. Therefore, it might be argued that the simple three-site models are the best choice for rehabilitation under an Ewald treatment. However, the protein force fields in common use were not, in fact, developed in a way that makes them particularly suited for the three-site water models.[25] Moreover, it is currently believed that any water model with demonstrable improvements in bulk water properties will also benefit solution properties. Taking these considerations together, the nonpolarizable TIP4P model was chosen as the model of choice for a reparameterization effort under an Ewald regime, since its simulated bulk-density with proper incorporation of long-ranged electrostatics is significantly better compared to other models (Fig. 1).

II. METHODS

A. The TIP4P-Ew model

Our purpose is to develop a re-parameterized TIP4P model (dubbed TIP4P-Ew) under inclusion of electrostatic and Lennard-Jones long-range interactions [see Eq. (6)]. Inclusion of the latter is important, as these interactions are always attractive (there is no partial cancellation as in electrostatic interactions) and have a noticeable influence on the density.[8]

Our goals are to produce a model appropriate for water in the *liquid phase* using experimental densities and enthalpies of vaporization at a number of different temperatures as input to a fitting procedure. A key issue for this approach is how one should extract information about the liquid phase of water from experimental enthalpies of vaporization, since this observable depends not only on the properties of liquid water, but also on those of water in the gas phase. One approach is to extract information about the intermolecular interaction energies of molecules in the liquid phase from these experimental enthalpies of vaporization by subtracting from them any effects due to the gas phase. One of these effects is due to the fact that gaseous water is a real gas.

Similarly, since our model uses fixed charges, we should account for the energetic effects of electrical polarization as a water molecule changes between the liquid and gas phases. Since our model is rigid, we should account in the experimental data for the omission of intramolecular vibrational degrees of freedom, including frequency shifts in these modes as water molecules transfer between phases. Finally, since we are developing a model that is to be used in a classical context, we should account for the fact that the experimental data include quantum effects. After considering all of these factors, we can produce from the experimental data the intermolecular interaction energies that a rigid, fixed charge, and classical water model should be able to reproduce. Alternatively, we could add all of these effects to our computed interaction energies to produce ''*computed*'' enthalpies of vaporization for direct comparison with the experimental enthalpies of vaporization. We have chosen the latter approach.

We should emphasize that since the production of the TIP4P-Ew model relies heavily on experimental densities and enthalpies of vaporization, the accuracy of the model should be assessed with respect to its ability to accurately predict *other* experimentally observable properties. However, it is still important to assess the degree to which our model is capable of describing the observed densities and enthalpies of vaporization. The goal is to achieve the best description of the relevant physics with a minimal set of parameters, and the ability to reproduce the data used in the fitting process is a measure of that.

In keeping with the tradition of its predecessor TIP4P, we have adopted the experimental gas-phase geometry of the water monomer[26] ($r_{OH}=0.9572$ Å and $\theta_{HOH}=104.52°$; see Fig. 2). Even though one might question the transferability of a molecular geometry from the gas phase into a highly associated liquid geometry such as in bulk water, we have kept these geometric parameters fixed. The other parameters $\{\xi\}$ of the model (see Fig. 2) are subjected to the parameter search procedure as outlined below.

B. Model energy expression

The total potential energy of the system is

$$U_{\text{total}}=U_{\text{Electrostatic}}+U_{\text{LJ}}, \tag{1}$$

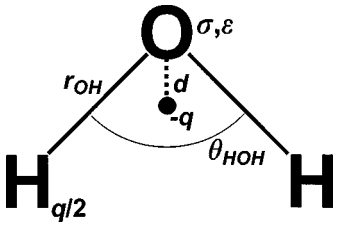

FIG. 2. Depiction of a TIP4P water molecule. The relevant model parameters are $\{\xi\}=\{q,d,\varepsilon,\sigma\}$, Eq. (7); parameters r_{OH} and θ_{HOH} are fixed. The fictitious center on the bisector of the θ_{HOH} angle is called the M-site.

where the electrostatic term is

$$U_{\mathrm{Electrostatic}}=\sum_{\substack{a\in I,b\in J\\ I<J}}\frac{q_a q_b e^2}{\varepsilon_r r_{ab}}, \tag{2}$$

and where the Lennard-Jones term is

$$U_{LJ}=\sum_{I<J}\left[u_{\mathrm{LJ}}(|\mathbf{r}_{O,I}-\mathbf{r}_{O,J}|)S(|\mathbf{r}_{O,I}-\mathbf{r}_{O,J}|)\right]+U_{\mathrm{LJ,tail}}. \tag{3}$$

In Eq. (2), a, b are charged sites on molecules I, J with charges q_a, q_b at separation r_{ab}, ε_r is the dielectric of the vacuum (ε_r is set to 1) and e is the charge of an electron.

In Eq. (3) $\mathbf{r}_{O,I}$ is the coordinate of an oxygen atom on molecule I and u_{LJ} is the usual Lennard-Jones functional form

$$u_{\mathrm{LJ}}(r)=4\varepsilon[(\sigma/r)^{12}-(\sigma/r)^6]. \tag{4}$$

The sums in Eqs. (2) and (3) go over all pairs of molecules I, J.

To ensure energy conservation and to avoid discontinuities due to truncation of the intermolecular potential, we employ a potential switching function S in Eq. (3) as described in Ref. 27. S is defined by a polynomial in $Z(r)=r^2-R_{\mathrm{lower}}^2$ that describes a function in the range from $Z=0$ ($r=R_{\mathrm{lower}}$) to $Z=R_{\mathrm{upper}}^2-R_{\mathrm{lower}}^2$

$$S(Z(r))=\begin{cases}1 & \text{if } r\leqslant R_{\mathrm{lower}}\\ 1+AZ^3+BZ^4+CZ^5 & \text{if } R_{\mathrm{lower}}<r\leqslant R_{\mathrm{upper}},\\ 0 & \text{if } r>R_{\mathrm{upper}}\end{cases} \tag{5}$$

with $A=-10/D^3$, $B=15/D^4$, $C=-6/D^5$, and $D=R_{\mathrm{upper}}^2-R_{\mathrm{lower}}^2$. This function is continuous and has continuous first and second derivatives at $r=R_{\mathrm{lower}}$ and $r=R_{\mathrm{upper}}$. The appropriate derivative of S is included in the intermolecular forces.

The long-range correction $U_{\mathrm{LJ,tail}}$ for the Lennard-Jones interaction energy[28] in Eq. (3) uses a mean-field approximation to account for neglected contributions to the Lennard-Jones energy U_{LJ}; its is obtained from integrals over the two intervals $[R_{\mathrm{lower}},R_{\mathrm{upper}}]$ and $[R_{\mathrm{upper}},\infty]$ as follows:

$$U_{\mathrm{LJ,tail}}=2\pi N\rho\left[\int_{R_{\mathrm{lower}}}^{R_{\mathrm{upper}}}(1-S(Z))r^2g(r)u_{\mathrm{LJ}}(r)dr+\int_{R_{\mathrm{upper}}}^{\infty}r^2g(r)u_{\mathrm{LJ}}(r)dr\right], \tag{6}$$

where we set $g(r)=1$ for $r>R_{\mathrm{lower}}$; N is the number of water molecules and $\rho=N/V$ is the number density. Under this approximation for $g(r)$, the integrals may be evaluated analytically.

The corresponding correction for the pressure[28] is obtained from a similar integral for the virial [$u_{\mathrm{LJ}}(r)$ is replaced by $r*du_{\mathrm{LJ}}(r)/dr$]. This is typically not done in current simulations. The long-range Lennard-Jones correction is always attractive, i.e., it lowers the potential energy, and causes a decrease of the internal pressure, or in a constant pressure simulation an increase of the bulk-density of about 0.5%–0.8%.[8]

C. Simulation protocol

A cubic box with edge length of 24.8 Å was filled with 512 water molecules. Molecular-dynamic (MD) simulations in an isothermal–isobaric (NpT) ensemble[29] at 1 atm and a range of temperatures were performed using an in-house simulation program. The equations of motion were integrated using the velocity Verlet algorithm[30] and with a time step size of 1 femtosecond. The velocity update was done using only forces on real sites after forces on fictitious sites (M-site, see Fig. 2) have been projected onto real sites.[31] The duration of equilibration runs was 100 picoseconds ($T>273$ K), 150 picoseconds ($T>248$ K), and 200 picoseconds ($T\leqq 248$ K). Typical production runs were longer than 5 nanoseconds. The intra-molecular geometry (r_{OH} and θ_{HOH}) was constrained by applying the M_SHAKE[32] and M_RATTLE[33] algorithms using an absolute geometric tolerance of 10^{-10} Å. Temperature and pressure were controlled using methods as described in Ref. 29, with velocity reassignment performed every 2000 MD time steps and a piston mass[34] of 0.000 40 amu Å^{-4} used in the context of isotropic expansion and contraction of the cubic simulation cell. Properties on the system (such as the instantaneous volume, potential energy, radial distribution functions, etc.) were sampled in 100 femtosecond intervals.

Coulomb interactions were computed using Ewald summation.[5] For the computation of the reciprocal space sum, 10 reciprocal space vectors in each direction were used, with a spherical cutoff for the reciprocal space sum of $n_x^2+n_y^2+n_z^2\leqslant 105$. The width of the screening Gaussian was 0.35 Å.

The values of R_{lower} and R_{upper} for the switching function S we use in the simulations are 9.0 and 9.5 Å, respectively.

The switch function in Eq. (5) (using the same settings for the switching parameters R_{lower} and R_{upper} as above) is also used as a molecule-based tapering function for the real-space Coulomb interaction energy in the Ewald summation.

The conversions factors and physical constants[35,36] used are listed in the supplemental material available from the publisher's website.[37] We note that the use of internally consistent and precise physical constants and conversion factors

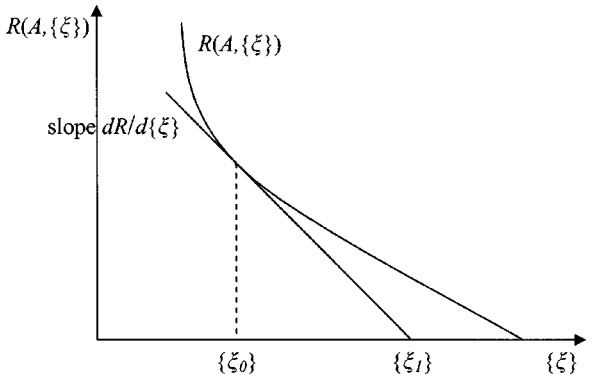

FIG. 3. Schematic representation of the residual function $R(A,\{\xi\})$, Eq. (8).

is crucial for comparison with our results as we strive for levels of precision (statistical uncertainties) that have traditionally not been common in the simulation community.

D. Parameter search

We want to modify the force field parameters $\{\xi\}$ for TIP4P water to get select simulated properties $A(T)$ [bulk-density $\rho(T)$, see Sec. II E 3; and enthalpy of vaporization $\Delta H_{\mathrm{vap}}(T)$, see Sec. II E 4] to agree closer with their experimental values over the range of temperatures of interest; $\{\xi\}$ designates the parameter set used in the force field (see Fig. 2)

$$\{\xi\}=\{q,d,\varepsilon,\sigma\}. \tag{7}$$

We define a residual function $R(A,\{\xi\})$ we wish to minimize with respect to $\{\xi\}$ in order to minimize the error in property A (see Fig. 3):

$$R(A,\{\xi\})=\sum_i \left[(A_{\mathrm{sim}}(T_i,\{\xi\})-A_{\mathrm{exp}}(T_i))/A_{\mathrm{exp}}(T_i)\right]^2$$

$$=\sum_i (\underline{A}_i-1)^2. \tag{8}$$

The sum is over temperatures. In the second form of Eq. (8) we are using the definition of the ''reduced property''

$$\underline{A}_i=A_{\mathrm{sim}}(T_i,\{\xi\})/A_{\mathrm{exp}}(T_i). \tag{9}$$

Note that $R(A,\{\xi\})$ is dimensionless by construction.

With the definition of the slope $dR/d\{\xi\}$ (see Fig. 3).

$$\left.\frac{dR}{d\{\xi\}}\right|_{\{\xi_0\}}=R(\{\xi_0\})/(\{\xi_0\}-\{\xi_1\}), \tag{10}$$

we can obtain an update $\{d\xi_1\}$ for parameter set $\{\xi_0\}$ by applying the Newton–Raphson rule

$$\{d\xi_1\}=\{\xi_1\}-\{\xi_0\}=-R(\{\xi_0\})\Big/\left.\frac{dR}{d\{\xi\}}\right|_{\{\xi_0\}}. \tag{11}$$

The dimension of $dR/d\{\xi\}$ is $[\{\xi\}]^{-1}$; to make $dR/d\{\xi\}$ dimensionless, we multiply it by $\{\xi\}$

$$dR/d\{\underline{\xi}\}\equiv\{\xi\}*dR/d\{\xi\};\ \text{ or }\ d\{\underline{\xi}\}=d\{\xi\}/\{\xi\}. \tag{12}$$

The definition of Eq. (12) enables us to use composite residuals (involving more than one property A). With Eq. (12), the parameter update Eq. (11) turns into an update relative to $\{\xi_0\}$

$$\{d\underline{\xi}_1\}=\{\xi_1/\xi_0\}-1=-R(\{\xi_0\})\Big/\left.\frac{dR}{d\{\underline{\xi}\}}\right|_{\{\xi_0\}}. \tag{13}$$

The slope $dR/d\{\underline{\xi}\}$ can be expressed in terms of derivatives of A with respect to $\{\xi\}$

$$dR/d\{\underline{\xi}\}=\frac{d}{d\{\underline{\xi}\}}\sum_i (\underline{A}_i-1)^2=2\sum_i (\underline{A}_i-1)d\underline{A}_i/d\{\underline{\xi}\}. \tag{14}$$

The derivatives of $\underline{A}$ with respect to $\{\underline{\xi}\}$ at different temperatures T_i in Eq. (14), $d\underline{A}_i/d\{\underline{\xi}\}$, can be obtained by numerical differentiation

$$d\underline{A}_i/d\{\underline{\xi}\}=(\underline{A}_i(\{\underline{\xi}^+\})-\underline{A}_i(\{\underline{\xi}^-\}))/(\{\underline{\xi}^+\}-\{\underline{\xi}^-\}). \tag{15}$$

An estimate for the reduced property $\underline{A}$ after applying the parameter update $\{d\underline{\xi}_1\}$ can be obtained as $\underline{A}_i(\{\xi_1\})=\underline{A}_i(\{\underline{\xi}_0\})+\{d\underline{\xi}_1\}*d\underline{A}_i/d\{\underline{\xi}_0\}$, which can be used to obtain an estimate for the new residual $R(A,\{\xi_1\})$ after the parameter update, a quantity useful when trying to judge the merit of updating one individual parameter over another. In order to obtain numerical derivatives [Eq. (15)], the perturbation $d\{\underline{\xi}^{\pm}\}=(\{\underline{\xi}^+\}-\{\underline{\xi}^-\})/2$ must be chosen carefully: It must be large enough to assure statistical significance of the computed finite difference $\underline{A}_i(\{\underline{\xi}^+\})-\underline{A}_i(\{\underline{\xi}^-\})$ for each of the properties $\underline{A}_i$ considered and it must be small enough to assure that $\underline{A}_i(\{\underline{\xi}\})$ is sufficiently linear in the interval $[\{\underline{\xi}^-\},\{\underline{\xi}^+\}]$. A recent chemometrics study[38] discusses sensitivity issues of the TIP4P model in this regard; here the authors used a perturbation of $\pm 5.0\%$ for each of the force field parameters, which we generally consider too high. Our choice of $d\{\underline{\xi}^{\pm}\}=\{0.50\%(q),4.0\%(d),1.50\%(\varepsilon),0.10\%(\sigma)\}$ is based on maximum allowed changes in properties $\underline{A}_i$ (± 0.010 g cm^{-3} for the density ρ and ± 0.30 kcal mol^{-1} for the enthalpy of vaporization ΔH_{vap}).

The numerical derivatives for the initial set of force field parameters, as listed in Table II, show some interesting characteristics. All four ΔH_{vap} derivatives are nearly constant across the entire temperature range. For the ρ derivatives three of them are either zero at $T\approx 273$ K ($\partial\rho/\partial q$ and $\partial\rho/\partial d$) or very small ($\partial\rho/\partial\varepsilon$). This means that changing parameter σ (i.e., the Lennard-Jones radius of oxygen) would lead to a mainly parallel shift in $\Delta H_{\mathrm{vap}}(T)$ and $\rho(T)$, while changing the other three parameters would permit changes to the shape of $\rho(T)$ (and perhaps influence the location of its maximum). These observations have led us to use the following parameter search strategy:

(1) increase parameter σ until $\rho(273\ \mathrm{K})\approx\rho_{\mathrm{exp}}(273\ \mathrm{K})$ using $\partial\rho/\partial\sigma$ information;
(2) vary parameters q, d, and ε (in this order) minimizing the density residual $R(\rho,\{q,d,\varepsilon\})$ alone;
(3) fine-tune all parameters minimizing the full combined residual $R(\rho+\Delta H_{\mathrm{vap}},\{q,d,\varepsilon,\sigma\})$.

TABLE I. Numeric values of the initial and the final set of force field parameters.

Force field parameter ξ^a	q [e$^-$]	d [Å]	ε [kcal mol^{-1}]	σ [Å]	Model dipole moment μ [D]
initial (TIP4P)	1.040	0.150	0.1550	3.15365	2.177
final (TIP4P-Ew)	1.04844	0.1250	0.162750	3.16435	2.321

[a]See Fig. 2.

The numerical derivatives for the final set of force field parameters (''TIP4P-Ew'') are listed in Table II; the corresponding residual R and its gradients are listed in Table III, while the numeric values of the force field parameters themselves are listed in Table I. It should be noted that this set of force field parameters cannot be changed without making either $\rho(T)$ or $\Delta H_{\text{vap}}(T)$ worse. It can be improved, however, if one is interested in creating a model that, for instance, best represents just $\rho(T)$ alone. It should also be noted that during the course of the re-parameterization the magnitude of the residual R was reduced by a factor of 20, while the magnitudes of its gradients were on average reduced by a factor of 10 (see Table III).

E. Property computation

1. Statistical uncertainties

Statistical uncertainties for averages of ''simple'' properties A (such as bulk-density, temperature, etc.) are estimated by fluctuation auto-correlation analysis (via the estimation of correlation times τ_A as outlined in Ref. 30). We note that the equilibrium average $\langle A\rangle$ (i.e., the quantity the uncertainty of which we wish to estimate) must relatively be well-known in order to estimate the correlation time τ_A reliably.

For properties that are determined by fluctuations (such as specific heat, isothermal compressibility, etc.) we do not attempt to estimate statistical uncertainties.

2. Polynomial fit

Various properties $A(T)$ discussed below are fit to an nth-order polynomial in T over the range of temperatures of interest in order to permit interpolation between temperatures or to obtain analytic derivatives. The general functional form used is

$$A(T)=\sum_{i=0}^{n} a(i)_A * T^i. \tag{16}$$

For simulation data, the coefficients $a(i)_A$ are determined from a weighted least-squares fitting procedure[39] where the weights are based on the statistical uncertainty of the data being fitted. For experimental data, the coefficients $a(i)_A$ are determined from an unweighed least-squares fit.

3. Bulk-density ρ

The average of the bulk-density $\langle\rho\rangle$ is computed from the average volume of the simulation box $\langle V\rangle$ using the molar mass M_{water} as listed in the supplemental material

$$\langle\rho\rangle=\frac{N_{\text{water}}M_{\text{water}}}{N_A\langle V\rangle};\quad \delta\langle\rho\rangle=\langle\rho\rangle\frac{\delta\langle V\rangle}{\langle V\rangle}. \tag{17}$$

Experimental density reference data have been taken from Ref. 40 (liquid phase) and from Ref. 41 (supercooled and superheated liquid).

4. Enthalpy of vaporization ΔH_{vap}

The enthalpy of vaporization ΔH_{vap} is the enthalpy change that occurs during the transition of one mol of substance from the liquid to the gas phase, where each of the phases is under the equilibrium pressure (i.e., the vapor pressure of the liquid).[42]

TABLE II. Numerical derivatives of $\underline{\rho}(T)$ and $\underline{\Delta H_{\text{vap}}}(T)$ for the initial and the final set of force field parameters.

	$\partial\underline{\rho}(T)/\partial\underline{\xi}^{a,b}$				$\partial\underline{\Delta H_{\text{vap}}}(T)/\partial\underline{\xi}^{a,b}$			
T [K]	$\underline{\xi}=\underline{q}$	$\underline{d}$	$\underline{\varepsilon}$	$\underline{\sigma}$	$\underline{q}$	$\underline{d}$	$\underline{\varepsilon}$	$\underline{\sigma}$
	initial							
235.5	−0.203 37	0.051 61	−0.059 18	−1.8809	2.9240	−0.528 25	−0.398 90	−7.0381
273	0.393 11	−0.072 14	−0.121 36	−3.3845	2.7807	−0.501 43	−0.380 32	−7.0452
323	0.790 43	−0.154 33	−0.135 94	−4.2806	2.7501	−0.516 99	−0.295 36	−6.9631
400	1.4589	−0.303 90	−0.218 24	−5.3538	2.8928	−0.560 00	−0.331 08	−6.4039
	final							
235.5	−0.685 64	0.077 94	0.071 92	−1.6333	3.1573	−0.432 55	−0.492 28	−7.8504
273	−0.023 01	−0.019 38	−0.041 34	−3.9111	2.9108	−0.388 93	−0.404 75	−7.1247
323	0.489 83	−0.105 86	−0.131 54	−4.5440	2.7093	−0.381 07	−0.362 47	−7.2635
400	1.1730	−0.192 43	−0.165 61	−5.7166	2.7158	−0.404 62	−0.307 96	−7.5861

[a]Reduced properties $\underline{\rho}$, $\underline{\Delta H_{\text{vap}}}$ as defined in Eq. (9) using experimental values from Table V and force field parameters from Table I.
[b]Reduced property derivatives as defined in Eq. (12).

TABLE III. Residual R and its derivatives for the initial and the final set of force field parameters.

	Residual R^a	$\partial R/\partial \underline{\xi}^b$			
		$\underline{\xi}=\underline{q}$	$\underline{\delta}$	$\underline{\varepsilon}$	$\underline{\sigma}$
Initial	0.015 866 7	−1.4188	0.268 19	0.171 10	3.4302
Final	0.000 893 35	0.143 33	−0.018 98	−0.021 60	−0.379 96

[a]Residual as defined by Eq. (8).
[b]Residual derivatives as defined by Eq. (14).

$$\Delta H_{\mathrm{vap}}(T)=H(p,T)_{\mathrm{gas}}-H(p,T)_{\mathrm{liquid}}=E_{\mathrm{gas}}-E_{\mathrm{liquid}}+p(V_{\mathrm{gas}}-V_{\mathrm{liquid}}), \tag{18}$$

with E being the total internal energy of the medium (consisting of a potential and a kinetic energy contribution: $E=U+K$). Under the assumption that the gas is ideal (potential energy $U_{\mathrm{gas}}=0$), and that the kinetic energies of a molecule in the gas and liquid phases are identical ($K_{\mathrm{gas}}=K_{\mathrm{liquid}}$) at a given temperature T, we can approximate Eq. (18) from quantities available from an NpT simulation as follows:

$$\begin{aligned}\Delta H_{\mathrm{vap}}(T)&\approx-\langle U(p,T)_{\mathrm{liquid}}\rangle/N+p(V_{\mathrm{gas}}-\langle V_{\mathrm{liquid}}\rangle)+C\\&\approx-\langle U(p,T)_{\mathrm{liquid}}\rangle/N+RT-p\langle V_{\mathrm{liquid}}\rangle+C,\end{aligned} \tag{19}$$

where $\langle U(p,T)_{\mathrm{liquid}}\rangle$ is the average intermolecular potential energy for N molecules at a given external pressure p and a bath temperature T, and $\langle V_{\mathrm{liquid}}\rangle$ is the average volume of the simulation box. The second term in Eq. (19) represents the work of expanding the gas against the external pressure p.

The correction term C in Eq. (19) corrects for the approximations made in the simulation and in the derivation of Eq. (19). That is, it accounts for vibrational, polarization, nonideal gas, and pressure effects:

$$C=C_{\mathrm{vib}}+C_{\mathrm{pol}}+C_{ni}+C_x. \tag{20}$$

In Eq. (20), the C_{vib} term accounts for intra- and intermolecular vibrational effects[43]

$$C_{\mathrm{vib}}=C_{\mathrm{vib,intra}}+C_{\mathrm{vib,inter}}. \tag{21}$$

For a classical harmonic oscillator, the energy for n_{vib} vibrational modes

$$E^{\mathrm{CM}}_{\mathrm{vib}}=n_{\mathrm{vib}}k_BT, \tag{22}$$

is independent of the modes' frequency.

For a quantum-mechanical harmonic oscillator, the energy is given by

$$E^{\mathrm{QM}}_{\mathrm{vib}}=\sum_{i=1}^{n_{\mathrm{vib}}}\left(h\nu_i/2+\frac{h\nu_i}{e^{h\nu_i/kT}-1}\right). \tag{23}$$

Quantum corrections to the vibrational energy are of two types. The first is due to the fact that our model is rigid, and so we must approximate the effect of the intra-molecular vibrations. Furthermore, the fundamental intra-molecular frequencies shift as a water molecule goes from the liquid into the gas phase. Therefore,

$$C_{\mathrm{vib,intra}}=E^{\mathrm{QM}}_{\mathrm{vib},g,\mathrm{intra}}-E^{\mathrm{QM}}_{\mathrm{vib},l,\mathrm{intra}}\quad(g=\mathrm{gas},\ \ l=\mathrm{liquid}), \tag{24}$$

where we use Eq. (23) and the intra-molecular vibrational frequencies of water in the gas and liquid phases[44] to compute both terms on the right.

The second type of quantum correction is due to the fact that several high frequency inter-molecular modes of the liquid are treated classically in the simulation, where they should have been treated quantum-mechanically; therefore,

$$C_{\mathrm{vib,inter}}=E^{\mathrm{QM}}_{\mathrm{vib},l,\mathrm{inter}}-E^{\mathrm{CM}}_{\mathrm{vib},l,\mathrm{inter}}\quad(l=\mathrm{liquid}), \tag{25}$$

where the first term on the right is evaluated using Eq. (23) and the high frequency inter-molecular modes of the liquid.[44] The second term on the right is from Eq. (22). Table IV lists the numeric values of C_{vib} for the temperatures of interest.

The second term in Eq. (20), C_{pol}, accounts for the depolarization energy that needs to be invested when a water molecule is transferred from the bulk to the gas phase[13] and was first proposed for the SPC/E model. A water molecule in the liquid described by a nonpolarizable effective pair potential like TIP4P has a higher dipole moment than a water molecule in the gas phase. The difference can be thought of as an induced dipole moment introduced by the bulk. C_{pol} can be approximated as[13]

TABLE IV. Corrections to ΔH_{vap} and c_p as computed by Eqs. (21), (27), and (31) for various temperatures.

T [K]	$C_{\mathrm{vib}}(T)^a$ [kcal mol^{-1}]	$C_{ni}(T)^b$ [kcal mol^{-1}]	$\partial E_{\mathrm{vib},l}(T)/\partial T^c$ [cal mol^{-1} K^{-1}]
235.5	−0.2247	−0.0001	−2.8960
248.0	−0.1894	−0.0003	−2.7499
260.5	−0.1559	−0.0007	−2.6115
273.0	−0.1241	−0.0014	−2.4806
285.5	−0.0938	−0.0027	−2.3571
298.0	−0.0651	−0.0048	−2.2408
310.5	−0.0378	−0.0079	−2.1313
323.0	−0.0118	−0.0125	−2.0284
335.5	0.0130	−0.0190	−1.9316
348.0	0.0365	−0.0276	−1.8407
360.5	0.0590	−0.0387	−1.7553
373.0	0.0804	−0.0527	−1.6750
400.0	0.1235	−0.0940	−1.5174

[a]Computed from Eq. (21); errors are ≤0.007 kcal mol^{-1} assuming an uncertainty in vibrational frequencies of 1 cm^{-1}.
[b]Computed from Eq. (27).
[c]Computed from Eq. (31); errors are ≤0.007 cal mol^{-1} K^{-1} assuming an uncertainty in vibrational frequencies of 1 cm^{-1}.

$$C_{\mathrm{pol}}=\frac{N}{2}(\mu_{\mathrm{gas}}-\mu_{\mathrm{liquid}})^2/\alpha_{\mathrm{gas}}, \tag{26}$$

where μ_{liquid} is the dipole moment of the effective pair model and μ_{gas} and α_{gas} are the dipole moment and the mean polarizability of a water molecule in the gas phase,[45] respectively.

The third term in Eq. (20), C_{ni}, accounts for the nonideality of the gas phase and is given by[46,47]

$$C_{ni}=\int_0^{p_{\mathrm{vap}}}\left(\frac{\partial H}{\partial p}\right)_T dp=\int_0^{p_{\mathrm{vap}}}\left[V(p)-\left(\frac{\partial V}{\partial T}\right)_p\right]dp$$
$$\approx p_{\mathrm{vap}}\left(B-T\frac{dB}{dT}\right). \tag{27}$$

In Eq. (27), the integration is done between the perfect-gas state ($p=0$) and the vapor pressure p_{vap} of the liquid and the virial equation of state is used. Experimental data for p_{vap} and the 2nd virial coefficient B have been taken from Ref. 48. Table IV lists the numeric values of C_{ni} for the temperatures of interest.

The last term in Eq. (20), C_x accounts for the fact that the simulation of the liquid is carried out at the external pressure $p=p_{\mathrm{ext}}$ rather than at the vapor pressure p_{vap}

$$C_x=\int_{p_{\mathrm{ext}}}^{p_{\mathrm{vap}}}\left(\frac{\partial H}{\partial p}\right)_T dp=\int_{p_{\mathrm{ext}}}^{p_{\mathrm{vap}}}\left[V(p)-T\left(\frac{\partial V(T)}{\partial T}\right)_p\right]dp$$
$$\approx\int_{p_{\mathrm{ext}}}^{p_{\mathrm{vap}}}[V(p_{\mathrm{ext}})[1-(p-p_{\mathrm{ext}})\kappa_T]-TV\alpha_p]dp.$$

Here, the second form uses the definition for the isothermal compressibility κ_T given in Eq. (32) [and a Taylor expansion for $V(p)$ around $p=p_{\mathrm{ext}}$] and the definition for the thermal expansion coefficient α_p given in Eq. (33). The numeric values of C_x for the range of temperatures studied ($p_{\mathrm{vap}}\lessapprox p_{\mathrm{ext}}$) are less than $4*10^{-4}$ kcal mol^{-1} and will, therefore, be neglected in this study.

Experimental ΔH_{vap} reference data have been taken from Ref. 40 ($T\geqq 253$ K). Values below 253 K have been obtained by extrapolation of a polynomial fit [Eq. (16)] of data from Ref. 40.

5. Isobaric heat capacity c_p

The isobaric heat capacity, c_p, is defined as

$$c_p=\left(\frac{\partial H}{\partial T}\right)_p. \tag{28}$$

We compute c_p in three different ways:

(a) by analytic differentiation of a polynomial fit of simulated enthalpies $\langle H(T)\rangle$ according to Eq. (16);

(b) by numeric differentiation of simulated enthalpies $\langle H(T)\rangle$ over the range of temperatures T of interest

$$c_p\approx\frac{\langle H_2\rangle-\langle H_1\rangle}{T_2-T_1}, \tag{29}$$

this allows to estimate statistical uncertainties of c_p

$$\delta c_p\approx\sqrt{\delta\langle H_1\rangle^2+\delta\langle H_2\rangle^2}/(T_2-T_1),$$

(a) by using the enthalpy fluctuation formula

$$c_p=\frac{\langle H^2\rangle-\langle H\rangle^2}{Nk_B\langle T\rangle^2}. \tag{30}$$

Due to the approximations made in the simulation (see discussion of ΔH_{vap} above), c_p values computed by Eqs. (16), (29), and (30) need to be corrected by a vibrational term

$$\left(\frac{\partial E_{\mathrm{vib},l}}{\partial T}\right)_p=\left(\frac{\partial E^{\mathrm{QM}}_{\mathrm{vib},l,\mathrm{intra}}}{\partial T}\right)_p+\left(\frac{\partial E^{\mathrm{QM}}_{\mathrm{vib},l,\mathrm{inter}}}{\partial T}\right)_p$$
$$-\left(\frac{\partial E^{\mathrm{CM}}_{\mathrm{vib},l,\mathrm{inter}}}{\partial T}\right)_p \quad (l=\mathrm{liquid}). \tag{31}$$

Table IV lists the numeric values of Eq. (31) for the temperatures of interest.

Experimental c_p reference data have been taken from Ref. 40 (liquid phase), from Ref. 49 (supercooled liquid), and from Ref. 41 (superheated liquid).

6. Isothermal compressibility κ_T

The isothermal compressibility κ_T is defined as

$$\kappa_T=-\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T=\frac{\langle V^2\rangle-\langle V\rangle^2}{k_B\langle T\rangle\langle V\rangle}. \tag{32}$$

We compute κ_T by using the volume fluctuation formula [second form of Eq. (32)]. Experimental κ_T reference data have been taken from Ref. 41.

7. Thermal expansion coefficient α_p

The thermal expansion coefficient, α_p, is defined as

$$\alpha_p=\frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_p. \tag{33}$$

We compute α_p in three different ways:

(a) By analytic differentiation of a polynomial fit of simulated bulk-densities $\langle\rho(T)\rangle$ according to Eq. (16)

$$\alpha_p=-\frac{d\ln\langle\rho(T)\rangle}{dT}, \tag{34}$$

(b) by numeric differentiation of simulated bulk-densities $\langle\rho(T)\rangle$ over the range of temperatures T of interest

$$\alpha_p\approx-\frac{\ln\langle\rho_2\rangle-\ln\langle\rho_1\rangle}{T_2-T_1}, \tag{35}$$

this allows to estimate statistical uncertainties of α_p

$$\delta\alpha_p\approx\sqrt{(\delta\langle\rho_1\rangle/\langle\rho_1\rangle)^2+(\delta\langle\rho_2\rangle/\langle\rho_2\rangle)^2}/(T_2-T_1),$$

(c) by using the enthalpy-volume fluctuation formula

$$\alpha_p=\frac{\langle VH\rangle-\langle V\rangle\langle H\rangle}{k_B\langle T\rangle^2\langle V\rangle}. \tag{36}$$

Experimental α_p reference data have been taken from Ref. 41.

TABLE V. Thermodynamic properties for the final TIP4P-Ew parameter set for a range of temperatures.[a]

T_{Bath} [K]	Duration [ns]	$\langle T_{\mathrm{internal}}\rangle$[b] [K]	$\langle U\rangle$[c] [kcal mol^{-1}]	$\langle\rho\rangle$ [g cm^{-3}]		$\langle\Delta H_{\mathrm{vap}}\rangle$ [kcal mol^{-1}]		c_p [cal mol^{-1} K^{-1}]			α_p [10^{-4} K^{-1}]		
				sim[d]	exp[e]	sim[f]	exp[g]	sim[h]	sim[i]	exp[j]	sim[k]	sim[l]	exp[m]
235.5	41	235.4	−6238.4	0.9845	0.9688	11.381	11.180	23.3		23.47	−9.2		
248.0	10	247.8	−6114.8	0.9935	0.98924	11.199	11.0372	21.9	21.9	19.34	−5.3	−5.7	−9.674
260.5	12	260.2	−6000.5	0.9986	0.99714	11.033	10.9029	20.9	20.9	18.38	−2.4	−2.5	−3.712
273.0	7	272.7	−5891.7	0.9996	0.99981	10.877	10.7732	20.1	20.2	18.170	−0.1	0.1	−0.705
285.5	5	284.9	−5790.4	0.9984	0.99953	10.731	10.6483	19.6	19.6	18.048	1.8	1.7	1.185
298.0	5	297.7	−5687.4	0.9954	0.99716	10.583	10.5176	19.2	19.2	18.004	3.4	3.1	2.558
310.5	4	309.4	−5596.8	0.9908	0.99362	10.452	10.3986	18.9	19.0	17.995	4.6	4.6	3.648
323.0	4	322.4	−5496.6	0.9843	0.98838	10.305	10.2640	18.7	18.7	18.004	5.5	5.4	4.567
335.5	5	335.3	−5401.2	0.9771	0.98207	10.163	10.1286	18.6	18.6	18.024	6.3	6.3	5.379
348.0	6	347.5	−5311.0	0.9688	0.97527	10.026	9.9993	18.5	18.6	18.054	7.0	7.4	6.121
360.5	6	360.1	−5219.3	0.9594	0.96737	9.883	9.8619	18.5	18.5	18.096	7.6	8.1	6.821
373.0	6	372.7	−5128.2	0.9492	0.95869	9.738	9.7206	18.5	18.4	18.151	8.3	9.2	7.498
400.0	7	399.3	−4939.7	0.9254	0.93803	9.426	9.4070	18.1		18.338	10.1		8.948

[a]The difference between the mean internal pressure and the control pressure (1 atm) is for all temperatures within one standard deviation (3 atm) of the mean internal pressure.
[b]Computed from the mean kinetic energy over the simulation along with the equipartition formula; note that this quantity is within two standard deviations (0.3 K) of the bath temperature T_{Bath}.
[c]Mean total potential energy including Lennard-Jones tail correction for system of 512 water molecules; statistical uncertainty $\delta\langle U\rangle \lessapprox 1.9$ kcal mol^{-1}.
[d]Statistical uncertainties: $\delta\langle\rho\rangle \lessapprox 0.0004$ g cm^{-3} ($T \leqq 248$ K); $\delta\langle\rho\rangle \lessapprox 0.0003$ g cm^{-3} ($T \geqq 260.5$ K).
[e]From Refs. 40 (liquid) and 41 (supercooled and superheated liquid); interpolated for actual simulation temperature $\langle T_{\mathrm{internal}}\rangle$.
[f]$\langle\Delta H_{\mathrm{vap}}\rangle$ values corrected according to Eq. (20), Sec. III E 4, statistical uncertainty $\delta\langle\Delta H_{\mathrm{vap}}\rangle \lessapprox 0.004$ kcal mol^{-1}.
[g]From Ref. 40, values interpolated for actual simulation temperature $\langle T_{\mathrm{internal}}\rangle$.
[h]From polynomial fit [Eq. (16)] of simulated enthalpies $\langle H\rangle$, Sec. II E 5 [part (a)]; with $n=6$ and $a_0 - 0.687\,867\,453\,03{*}10^{+2}$; $a_1 = 0.955\,591\,459\,04{*}10^{+0}$; $a_2 - 0.696\,621\,245\,80{*}10^{-2}$; $a_3 = 0.280\,812\,719\,77{*}10^{-4}$; $a_4 - 0.642\,690\,614\,43{*}10^{-7}$; $a_5 = 0.788\,654\,497\,06{*}10^{-10}$; $a_6 = -0.404\,309\,859\,83{*}10^{-13}$.
[i]From finite difference formula, Eq. (29), Sec. II E 5 [part (b)]; statistical uncertainty $\delta c_p \lessapprox 0.4$ cal mol^{-1} K^{-1}. $c_p(T)$ values corrected by $\partial E_{\mathrm{vib},I}(T)/\partial T$ terms [Eq. (31), 3rd column of Table IV]. This also applies to footnote[i].
[j]From Refs. 40 (liquid), 49 (supercooled liquid), and 41 (superheated liquid).
[k]From polynomial fit of simulated bulk-densities $\langle\rho\rangle$, Eq. (34), Sec. II E 7 [part (a)]; with $n=4$ and $a_0 = -0.138\,356\,556\,933{*}10^{+1}$; $a_1 = 0.267\,602\,057\,295{*}10^{-1}$; $a_2 = -0.110\,711\,242\,268{*}10^{-3}$; $a_3 = 0.202\,358\,222\,565{*}10^{-6}$; $a_4 = -0.141\,916\,114\,764{*}10^{-9}$.
[l]From finite difference formula, Eq. (35), Sec. II E 7 [part (b)]; statistical uncertainty $\delta\alpha_p \lessapprox 0.3{*}10^{-4}$ K^{-1}.
[m]From Ref. 41.

8. Self-diffusion coefficient D

Self-diffusion coefficients D were obtained for the final TIP4P-Ew model from the mean-square O–O displacement simulated under an NVE ensemble using the Einstein relation[28]

$$D = \lim_{t\to\infty} \frac{1}{6t}\langle|r(t)-r(0)|^2\rangle = \lim_{t\to\infty} \frac{1}{6t}\overline{R_{\mathrm{OO}}(t)}. \tag{37}$$

The starting configurations for the NVE simulations were sampled in 5 picosecond intervals from extended wellequili-

TABLE VI. Fluctuation properties for the final TIP4P-Ew parameter set for a range of temperatures

T_{Bath} [K]	Duration [ns]	$\langle T_{\mathrm{internal}}\rangle$[a] [K]	c_p[b] [cal mol^{-1} K^{-1}]	κ_T[10^{-6} atm^{-1}]		α_p[e] [10^{-4} K^{-1}]	Static dielectric $\varepsilon(0)$	
				sim[c]	exp[d]		sim[f]	exp[g]
235.5	41	235.4	24.2	54.3		−9.2	81.9±5.2	106.31±1.0 (238 K)
273.0	36	272.9	20.3	48.9	51.62	−0.1	70.8±1.4	87.96±0.04
298.0	36	297.9	19.2	48.1	45.86	3.2	63.9±0.9 62.9±1.0[h]	78.46
323.0	36	322.9	19.0	49.4	44.76	5.5	60.0±0.7	69.96±0.04
348.0	22	347.8	19.2	53.6	46.21	7.6	54.1±0.7	62.36
373.0	22	372.9	19.4	59.9	49.65	9.4	48.7±0.6	55.57±0.2
400.0	22	399.9	18.4	68.4	55.62	10.5	42.7±0.5	

[a]Computed from the mean kinetic energy over the simulation along with the equipartition formula; note that this quantity is within two standard deviations (0.1 K) within the bath temperature T_{Bath}.
[b]From enthalpy fluctuation formula, Eq. (30), Sec. II E 5 [part (c)]; $c_p(T)$ values corrected by $\partial E_{\mathrm{vib},I}(T)/\partial T$ terms [Eq. (31), 3rd column of Table IV].
[c]From volume fluctuation formula, Eq. (32), Sec. II E 6.
[d]From Ref. 41.
[e]From enthalpy-volume fluctuation formula, Eq. (36), Sec. II E 7 [part (c)].
[f]From dipole fluctuation formula, Eq. (38), Sec. II E 9.
[g]From Ref. 53.
[h]From weighted linear fit using external field formula, Eq. (40), Sec. II E 9.

brated NpT simulations. For each temperature studied, 80 NVE simulations were carried out for a duration of 80 picoseconds ($T=235.5$ K), 30 picoseconds ($T=273$ K), and 20 picoseconds ($T>273$ K), respectively. The average self-diffusion coefficient $\langle D(T)\rangle$ at each temperature (and its uncertainty) was obtained by averaging $D(T)$ values resulting from a weighted linear least-squares fit of $\overline{R_{OO}(t)}$ versus t for each of the NVE runs.

9. Static dielectric constant $\varepsilon(0)$

The static dielectric constant $\varepsilon(0)$ of a medium is determined by the magnitude and density of the molecular dipole moments and the extent to which the directions of the dipole moments are correlated. Within the Ewald approach using conducting boundary conditions, $\varepsilon(0)$ is given by[50]

$$\varepsilon(0)=1+4\pi\frac{\langle M^2\rangle}{3\langle V\rangle k_B\langle T\rangle}, \tag{38}$$

with $\langle M^2\rangle$ being the fluctuation of the total system dipole moment

$$\langle M^2\rangle=\langle \boldsymbol{M}\cdot\boldsymbol{M}\rangle-\langle \boldsymbol{M}\rangle\cdot\langle \boldsymbol{M}\rangle=\langle M_x^2+M_y^2+M_z^2\rangle-(\langle M_x\rangle^2+\langle M_y\rangle^2+\langle M_z\rangle^2). \tag{39}$$

A statistical uncertainty $\delta\varepsilon(0)$ for $\varepsilon(0)$ can be obtained from statistical uncertainties $\delta\langle V\rangle$, $\delta\langle T\rangle$, $\delta\langle(M_x^2+M_y^2+M_z^2)\rangle$, $\delta\langle M_x\rangle$, $\delta\langle M_y\rangle$, and $\delta\langle M_z\rangle$ using standard error propagation. Due to very long fluctuation auto-correlation times, $\langle M^2\rangle$ converges rather slowly, especially at low temperatures.

Alternatively, in the presence of an external electric field, E_0, $\varepsilon(0)$ is given by[50]

$$\varepsilon(0)=1+4\pi\frac{\langle P\rangle_{E_0}}{E_0},\quad \text{with } \langle P\rangle=\langle \boldsymbol{M}/V\rangle, \tag{40}$$

where the polarization $\langle P\rangle$ is the system dipole moment density.

For small "nonsaturating" fields[51] (i.e., $\mu E_0/3k_B\ll T$), $\langle P\rangle$ is linear in E_0; we note that the saturation field strength depends on the model's dipole moment μ, as well as on the temperature!

For polarizable molecules, a polarizability correction $\varepsilon_{\text{pol}}^{\text{corr}}$ is added to $\varepsilon(0)$ obtained from either Eqs. (38) or (40): $\varepsilon'(0)=\varepsilon(0)+\varepsilon_{\text{pol}}^{\text{corr}}$; $\varepsilon_{\text{pol}}^{\text{corr}}=4\pi N\alpha/\langle V\rangle$, $\varepsilon_{\text{pol}}^{\text{corr}}$ is related to the high-frequency dielectric, $\varepsilon(\infty)$: $\varepsilon_{\text{pol}}^{\text{corr}}=\varepsilon(\infty)-\varepsilon_{\text{vacuum}}=\varepsilon(\infty)-1$. The experimental value of $\varepsilon_{\text{pol}}^{\text{corr}}$ at room temperature is 0.79.[52] In this paper we compute $\varepsilon_{\text{pol}}^{\text{corr}}$ from the experimental mean polarizability α_{gas}[45] and the average volume of the simulation box, $\langle V\rangle$. Experimental $\varepsilon(0)$ reference data have been taken from Ref. 53.

10. Radial distribution functions $g(r)$ and scattering intensities $I(Q)$

The pair radial distribution function $g_{ab}(r)$ for atoms of type a and b is obtained from the simulation via[28]

$$g_{ab}(r)=\frac{V}{N^2}\left\langle\sum_{i,j\neq i}\delta(r_a^i)\,\delta(r_b^j-r)\right\rangle, \tag{41}$$

TABLE VII. Self-diffusion coefficients for a range of temperatures.

T_{NpT} [K] (1 atm)	$\langle T_{\text{NVE}}\rangle$ [K]	D [10^{-9} m^2 s^{-1}]	
		TIP4P-Ew	exp[a,b]
235.5	235.1	0.17±0.01	0.187±5% ($T=242.5$ K)[a]
273	272.2	1.2±0.02	1.05±5% ($T=273.5$ K)[a]
298	297.4	2.4±0.06	2.23±0.1[a] 2.299±0.2%[b]
323	321.6	3.9±0.06	3.575±0.2% ($T=318$ K)[b]
348	346.9	5.7±0.04	
373	371.6	7.8±0.1	
400	398.4	10.3±0.1	

[a]From Ref. 58.
[b]From Ref. 59.

$g(r)$ is related, through the spatial Fourier-type transform, to the structure factor $h_{ab}(Q)$

$$h_{ab}(Q)=4\pi\rho\int_0^\infty r^2(g_{ab}(r)-1)\frac{\sin(Qr)}{Qr}dr, \tag{42}$$

where $h_{ab}(Q)$ in turn is related to the scattering intensity $I(Q)$, an observable measured directly by experiment[54]

$$I(Q)\approx\underbrace{\sum_{ab}x_a x_b f_a(Q)f_b(Q)\frac{\sin(Qr_{ab})}{Qr_{ab}}}_{\text{intra-molecular term}}+\underbrace{\sum_{a\leqslant b}x_a x_b f_a(Q)f_b(Q)h_{ab}(Q)}_{\text{inter-molecular term}}. \tag{43}$$

Equation (43) makes the assumption (Debye approximation) that scattering effects can be separated into intra- and intermolecular contributions, and further, that scattering can be represented as arising from independent neutral atoms [described by atomic weights x and atomic form factors $f(Q)$]. In Ref. 54, the authors point out the difficulties and ambiguities of deriving radial distribution functions $g(r)$ from measured scattering intensities $I(Q)$, and devise a reverse procedure that produces $I(Q)$ curves from simulated $g(r)$, the results of which we adopt in this paper (see Fig. 10). We do realize that in doing so we violate common practice, which, traditionally, treated radial distribution functions $g(r)$ as the dedicated probe to assess whether a water model's structural properties agree with the experiment.

III. RESULTS AND DISCUSSION

The final TIP4P-Ew model (see Table I) has been characterized by computing thermodynamic properties $[\langle\rho\rangle,\langle\Delta H_{\text{vap}}\rangle,c_p,\kappa_T,\alpha_p,D,\varepsilon(0)]$ for a range temperatures in the interval [235.5 K, 400 K] listed in Tables V–VII together with experimental values. Results are also presented in Figs. 4–10, often with data from experiment and from selected other water models. In some cases we show comparisons between TIP4P-Ew and TIP4P-Pol2,[18] a potential we consider to be one of the better polarizable water models, especially with regard to structural properties.

The bulk-density $\rho(T)$ as a function of temperature for TIP4P-Ew, TIP4P-Pol2[18] and experiment is shown in Fig. 4.

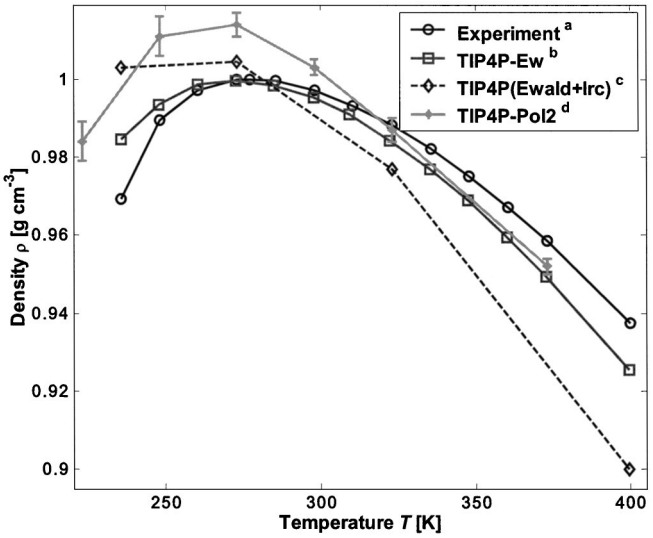

FIG. 4. Bulk-density $\rho(T)$ of water. (a) From Refs. 40 (liquid) and 41 (supercooled and superheated liquid); (b) this paper; final parameter set; statistical uncertainties $\lesssim 0.0004$ g cm^{-3}; $\rho(T)$ has a maximum at $\approx$274 K; (c) this paper; initial TIP4P parameter set; with Ewald and Lennard-Jones long-range correction; statistical uncertainties $\lesssim 0.0004$ g cm^{-3}; (d) from Ref. 18.

The temperature T_m where $\rho(T)$ has its maximum we find (from polynomial fit of $\rho(T)$ using Eq. (16), or from $\alpha_p(T)$ curve, see Fig. 6, bottom) $T_m \approx 274$ K for TIP4P-Ew (experiment: $T_m = 277$ K; TIP4P with Ewald (see Fig. 1): $T_m \approx 256$ K). The absolute average density error for our model is 0.58% or 0.0056 g cm^{-3} over the 165° temperature range (maximum deviation of +1.6% at −37.5 °C).

The enthalpy of vaporization $\Delta H_{\text{vap}}(T)$ as a function of temperature for TIP4P-Ew [corrected by Eq. (20)], TIP5P[17] and the experiment is shown in Fig. 5. The absolute average error of $\Delta H_{\text{vap}}(T)$ for our model is 0.69% or 0.074 kcal mol^{-1} (maximum deviation of +1.8% at −37.5 °C). The slope of the TIP4P-Ew curve is marginally steeper than the experimental curve, which manifests itself in heat capaci-

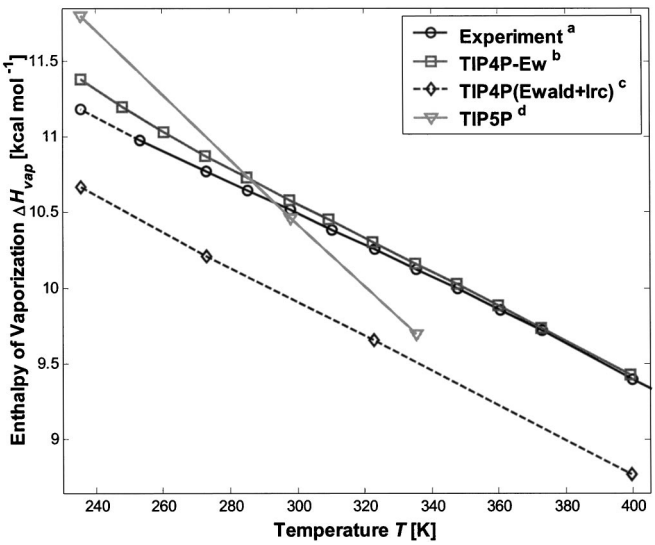

FIG. 5. Enthalpy of vaporization $\Delta H_{\text{vap}}(T)$ of water. (a) From Ref. 40; value at 235.5 K extrapolated (dashed line); (b) this paper; final parameter set; values corrected according to Eq. (20), Sec. II E 4; statistical uncertainties $\lesssim 0.004$ kcal mol^{-1}; (c) this paper; initial TIP4P parameter set; with Ewald and Lennard-Jones long-range correction; values corrected as in (b); statistical uncertainties $\lesssim 0.003$ kcal mol^{-1}; (d) from Ref. 17; errors 0.01 kcal mol^{-1}.

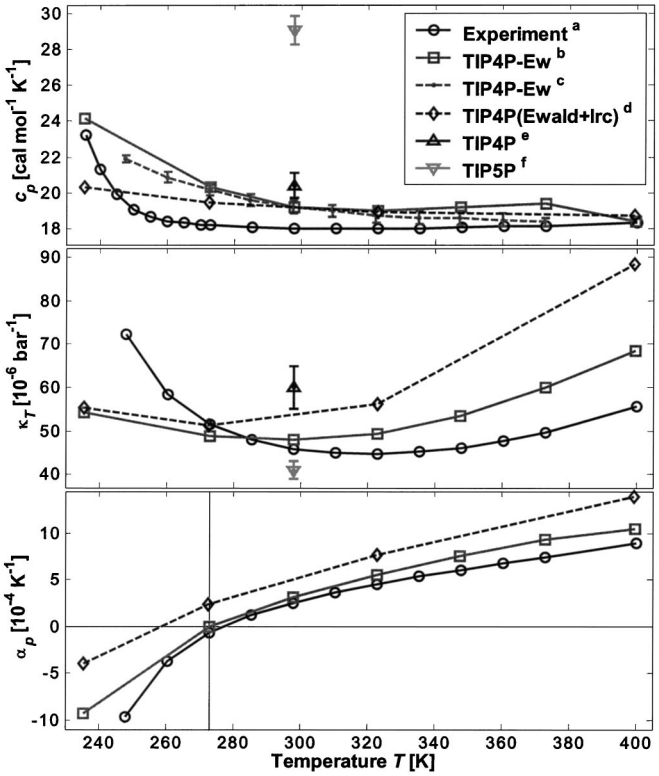

FIG. 6. Isobaric heat capacity $c_p(T)$, isothermal compressibility $\kappa_T(T)$ and thermal expansion coefficient $\alpha_p(T)$ of water. (a) $c_p(T)$ from Refs. 40 (liquid), 49 (supercooled liquid), and 41 (superheated liquid); $\kappa_T(T)$ and $\alpha_p(T)$ from Ref. 41.

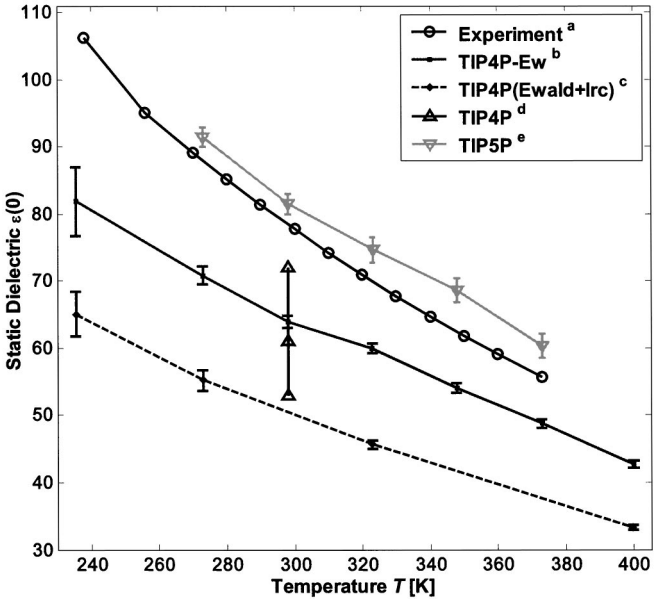

FIG. 7. Static dielectric constant $\varepsilon(0)$. (a) From Ref. 53; (b) this paper; final parameter set; from dipole fluctuation formula, Eq. (38); (c) this paper; initial TIP4P parameter set; with Ewald and Lennard-Jones long-range correction; from dipole fluctuation formula, Eq. (38); (d) from Ref. 60; (e) from Ref. 17.

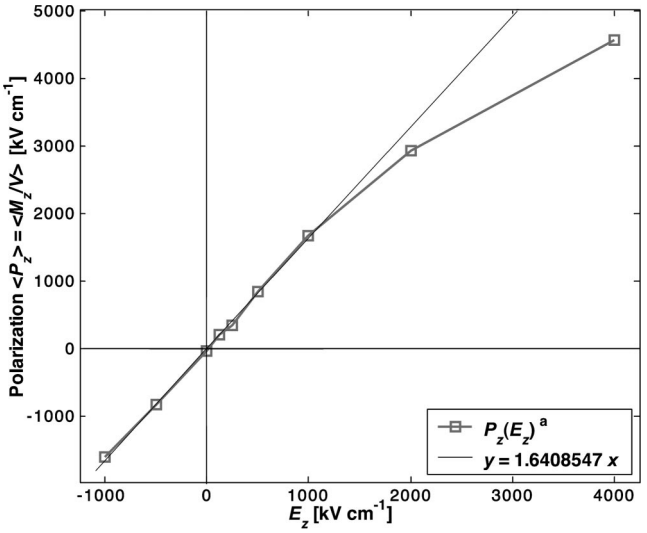

FIG. 8. Polarization $\langle P_z\rangle = \langle M_z/V\rangle$ as a function of an external field E_0 in z direction. (a) Statistical uncertainties $\delta\langle P_z\rangle \lessapprox 33$ kV cm^{-1}; the saturation field strength is found to be $\gtrsim 1000$ kV cm^{-1}.

ties $c_p(T)$ (see Fig. 6), that are slightly too high (≈ 1.5 cal mol^{-1} K^{-1} at 298 K). We note that the statistical uncertainties in $\Delta H_{\mathrm{vap}}(T)$ ($\lessapprox 0.004$ kcal mol^{-1}) are smaller than conservatively estimated errors in the corrections applied [$\delta(\partial C_{\mathrm{vib}}/\partial T)_p$, Table IV].

Plots of the isothermal compressibilities $\kappa_T(T)$ and thermal expansion coefficients $\alpha_p(T)$ as a function of temperature for TIP4P-Ew and the experiment are given in Fig. 6. For the model, $\kappa_T(T)$ is within about 8% of experiment, and $\alpha_p(T)$ is within $+10^{-4}$ K^{-1} of experiment, between the temperatures 273 and 310 K. The thermal expansion coefficient is zero, corresponding to the density maximum, at $T_m \approx 274$ K.

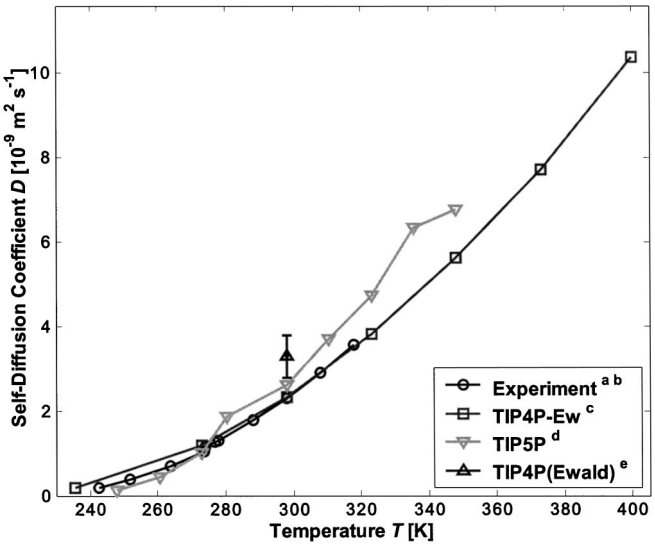

FIG. 9. Self-diffusion coefficient $D(T)$ of water. (a) 274–318 K: From Ref. 59; errors 0.2%; (b) 242.5–298 K: From Ref. 58; errors 5%, $0.1*10^{-9}$ m^2 s^{-1} at 298 K; (c) this paper; uncertainties $\delta D \lessapprox 0.1*10^{-9}$ m^2 s^{-1}; (d) from Ref. 61; errors $\lessapprox 0.10*10^{-9}$ m^2 s^{-1}; (e) from Ref. 62; error $0.5*10^{-9}$ m^2 s^{-1}.

Figure 7 compares the results for the static dielectric constant $\varepsilon(0)$ as a function of temperature for TIP4P-Ew, TIP4P with Ewald and Lennard-Jones long-range corrections and the experiment. Even after an extended simulation of >40 nanoseconds, the statistical uncertainty of $\varepsilon(0)$ at 235.5 K is significant (± 5). At one temperature ($T=298$ K), the dielectric was computed using Eq. (40) in the presence of an external field E_0. Figure 8 shows the polarization $\langle P\rangle$ as a function of E_0. We find the saturation field strength to be $\gtrsim 1000$ kV cm^{-1} for this temperature. From a linear weighted fit of $\langle P(E_0)\rangle$ we obtain a dielectric constant of $\varepsilon(0) = 62.9 \pm 1.0$, which agrees well with the result obtained from the dipole fluctuation formula Eq. (38) for this temperature [$\varepsilon(0) = 63.9 \pm 0.9$; see Table VI]. The overall agreement of TIP4P-Ew with the experiment is not great [$\delta\varepsilon(0) \lesssim -15\%$], but it halves the error of $\varepsilon(0)$ for TIP4P in its original parameterization (dashed curve). We admit that this is disappointing; the static dielectric constant $\varepsilon(0)$ may well be one of the properties that requires the use of a polarizable water model, the inclusion of molecular flexibility or the use of something other than conducting boundary conditions.[55]

Simulated self-diffusion coefficients D for the temperatures studied along with the experimental data are summarized in Table VII. Figure 9 plots data shown in Table VII; the agreement is remarkable where experimental data are available.

Figure 10 shows simulated O–O radial distribution functions $g_{\mathrm{OO}}(r)$ for TIP4P-Ew for a few select temperatures together with their corresponding scattering intensity curves $I(Q)$,[54] and comparisons with results for TIP4P and the polarizable model TIP4P-Pol2[18] [considered one of the best benchmark for $g(r)$[54]] against experiment. The TIP4P-Ew water structure shows considerable improvement over TIP4P, and comparable performance with TIP4P-Pol2, over the temperature range studied.

IV. CONCLUSION

We have presented a re-parameterization of the popular TIP4P water model (dubbed "TIP4P-Ew") with inclusion of both Coulomb and Lennard-Jones long-range interactions.

The new model has been tuned to reproduce both the bulk-density as well as the enthalpy of vaporization over the entire accessible liquid range (235.5–400 K) at ambient pressure. A careful characterization of the new model revealed that many thermodynamic and kinetic properties (bulk-density ρ, enthalpy of vaporization ΔH_{vap}, heat capacity c_p, compressibility κ_T, expansion coefficient α_p, self-diffusion coefficient D) are described very well over the entire temperature range.

We emphasize that even though the development of the new model used experimental densities and heats of vaporization as input, it is still important to assess the degree to which our model is capable of reproducing its input data as a measure of the model's ability to achieve the best description of the relevant physics with a minimal set of parameters.

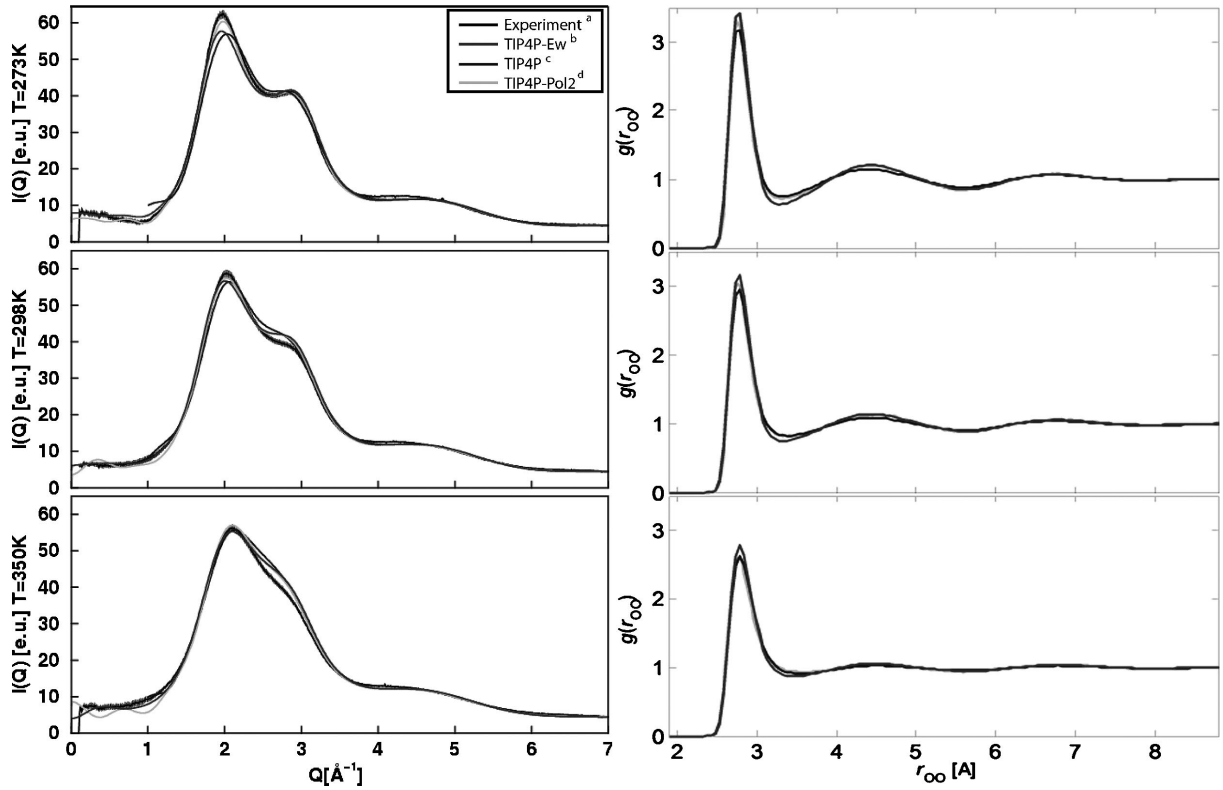

FIG. 10. Scattering intensities $I(Q)$ and O–O radial distribution functions $g(r)$ for water. (a) From Ref. 54 (x-ray scattering experiments at $T=275$, 298, and 350 K); (b) this paper ($T=273$, 298, and 348 K); (c) this paper; initial TIP4P parameter set; no Ewald and Lennard-Jones long-range correction ($T=273$, 298, and 348 K); (d) from Ref. 54 ($T=275$, 298, and 348 K); the authors wish to acknowledge J. Ilja Siepmann for providing TIP4P-Pol2 radial distribution function for our analyses.

The transferability of the new model, in particular its suitability to be used as solvent model in bio-molecular simulations is yet to be tested. Specifically, kinetic studies addressing hydrogen-bond life times as well as solvation studies (free energies of hydration for a selection of reference molecules in TIP4P-Ew water) need to be conducted.

Considering the success of the original TIP4P water model and the quality of the results obtained here, the new model should also be a good general-purpose water model. Further work is needed to confirm this statement. Simulations currently underway include the exploration of the phase diagram (ice–water properties, melting point; ice properties; vapor pressure, boiling point), interfacial properties, nonambient pressures, critical region, clusters, bilayers, molecule solvation, and bio-molecular systems.

APPENDIX: DERIVATION OF THE M_RATTLE EQUATIONS

For MD simulations of systems with bond constraints, one generally must resort to an iterative method for handling the constraints such as SHAKE[56] or RATTLE.[57] Generally, these methods are not rapidly converging, and considerable amounts of computer time can be spent addressing the constraints. Recently, the M_SHAKE algorithm has been described[32] for treating constraints in the context of Verlet or leap frog dynamical integration algorithms. Here, we show, for a fully rigid molecule, how we can use the M_SHAKE approach with the velocity Verlet algorithm.[30] This derivation closely follows the notation used in the description of the RATTLE algorithm.

In the velocity Verlet algorithm the fundamental equations are

$$\boldsymbol{r}(t+h)=\boldsymbol{r}(t)+h\boldsymbol{v}(t)+\frac{h^2}{2m}\boldsymbol{F}(t),$$

$$\text{compute } \boldsymbol{F}(t+h), \tag{A1}$$

$$\boldsymbol{v}(t+h)=\boldsymbol{v}(t)+\frac{h}{2m}(\boldsymbol{F}(t)+\boldsymbol{F}(t+h)),$$

where $\boldsymbol{r}$ is the position, $\boldsymbol{v}$ is the velocity Verlet velocity (an approximation to the true velocity), $\boldsymbol{F}$ is the force, m is the mass of the particle, and h is the time step size. When bond length constraints are introduced, they are of the form $\sigma_{ij}(|\boldsymbol{r}_i-\boldsymbol{r}_j|)=r_{ij}^2-d_{ij}^2=0$. These result in time-dependent forces of constraint, $\boldsymbol{F}_i^c(t)=\Sigma_{j\neq i}-\lambda_{ij}(t)\nabla_i\sigma_{ij}=\Sigma_{j\neq i}-2\lambda_{ij}(t)\boldsymbol{r}_{ij}$, where $\lambda_{ij}=\lambda_{ji}$ are scalar Lagrange multipliers whose values are determined so as to make the constraints satisfied at each time step. With velocity Verlet, there are actually constraints on the velocities as well as on the posi-

tions: $d\sigma_{ij}/dt = 2\boldsymbol{r}_{ij}\cdot\boldsymbol{v}_{ij} = 0$. This is a statement that there should be no component of the relative velocity parallel to a bond that is constrained and implies the use of both positional $[\lambda_{RRij}(t)]$ and velocity $[\lambda_{RVij}(t)]$ Lagrange multipliers. Following the RATTLE scheme, when constraints are introduced the velocity Verlet algorithm begins as follows:

$$\boldsymbol{r}_i(t+h) = \boldsymbol{r}_i(t) + h\boldsymbol{v}_i(t) + \frac{h^2}{2m_i} \times\left[\boldsymbol{F}_i(t) + \sum_{j\neq i} -2\lambda_{RRij}(t)\boldsymbol{r}_{ij}(t)\right].$$

The RATTLE algorithm provides a prescription for computing $\lambda_{RRij}(t)$ through an iterative process. However, the M_SHAKE algorithm can be applied directly for this step of the velocity Verlet algorithm. The M_SHAKE algorithm is a matrix-based iterative procedure that is more rapidly converging than SHAKE and RATTLE for obtaining positional Lagrange multipliers for system of coupled constraints. Once the $\lambda_{RRij}(t)$ are determined, the positions at time $t+h$ are known and the forces at that time can be computed. The last thing to be done in a velocity Verlet algorithm is to update the velocities with the following expression:

$$\boldsymbol{v}_i(t+h) = \boldsymbol{v}_i(t) + \frac{h}{2m_i}\left[\boldsymbol{F}_i(t) + \sum_{j\neq i} -2\lambda_{RRij}(t)\boldsymbol{r}_{ij}(t) + \boldsymbol{F}_i(t+h) + \sum_{j\neq i} -2\lambda_{RVij}(t+h)\boldsymbol{r}_{ij}(t+h)\right]. \tag{A2}$$

In Eq. (A2) everything on the right is known except the $\lambda_{RVij}(t+h)$, the velocity Lagrange multiplier in the RATTLE scheme, which are chosen to make $\boldsymbol{r}_{ij}(t+h)^*\boldsymbol{v}_{ij}(t+h) = 0$. For a fully rigid molecule these may be found in closed form on a molecule-by-molecule basis. Consider a rigid three site molecule, which could be our TIP4P-Ew model (the M-site position can be constructed from the positions of the oxygen and the two hydrogens). Label the three sites 0, 1, and 2. For our fully rigid system we are concerned with coordinates $\boldsymbol{r}_0$, $\boldsymbol{r}_1$, $\boldsymbol{r}_2$, the corresponding velocities and the three constraints σ_{01}, σ_{02}, σ_{12}. Define

$$g_{ij} \equiv -h\lambda_{RVij}(t+h) \tag{A3}$$

and

$$\boldsymbol{q}_i \equiv \boldsymbol{v}_i(t) + \frac{h}{2m_i}\left[\boldsymbol{F}_i(t) + \sum_{j\neq i} -2\lambda_{RRij}(t)\boldsymbol{r}_{ij}(t) + \boldsymbol{F}_i(t+h)\right],$$

so that Eq. (A2) becomes

$$\boldsymbol{v}_i(t+h) = \boldsymbol{q}_i + \frac{1}{m_i}\sum_{j\neq i} g_{ij}\boldsymbol{r}_{ij}(t+h). \tag{A4}$$

For each molecule, we need to determine the three values g_{01}, g_{02}, and g_{12}, which result in $\boldsymbol{r}_{ij}(t+h)^*\boldsymbol{v}_{ij}(t+h) = 0$. Dropping the $t+h$ argument for simplicity one obtains for the three relative velocities

$$\boldsymbol{v}_{01} = \boldsymbol{q}_{01} + \frac{1}{m_0}g_{01}\boldsymbol{r}_{01} + \frac{1}{m_0}g_{02}\boldsymbol{r}_{02} + \frac{1}{m_1}g_{01}\boldsymbol{r}_{01} - \frac{1}{m_1}g_{12}\boldsymbol{r}_{12},$$

$$\boldsymbol{v}_{02} = \boldsymbol{q}_{02} + \frac{1}{m_0}g_{01}\boldsymbol{r}_{01} + \frac{1}{m_0}g_{02}\boldsymbol{r}_{02} + \frac{1}{m_2}g_{02}\boldsymbol{r}_{02} + \frac{1}{m_2}g_{12}\boldsymbol{r}_{12},$$

$$\boldsymbol{v}_{12} = \boldsymbol{q}_{12} - \frac{1}{m_1}g_{01}\boldsymbol{r}_{01} + \frac{1}{m_1}g_{12}\boldsymbol{r}_{12} + \frac{1}{m_2}g_{02}\boldsymbol{r}_{02} + \frac{1}{m_2}g_{12}\boldsymbol{r}_{12}.$$

In writing this we have taken advantage of the fact that $\boldsymbol{r}_{ij} = -\boldsymbol{r}_{ji}$. If we take appropriate dot products we have that

$$\boldsymbol{v}_{01}\cdot\boldsymbol{r}_{01} = \boldsymbol{q}_{01}\cdot\boldsymbol{r}_{01} + \left(\frac{1}{m_0}+\frac{1}{m_1}\right)g_{01}\boldsymbol{r}_{01}\cdot\boldsymbol{r}_{01} + \frac{1}{m_0}g_{02}\boldsymbol{r}_{01}\cdot\boldsymbol{r}_{02} - \frac{1}{m_1}g_{12}\boldsymbol{r}_{01}\cdot\boldsymbol{r}_{12},$$

$$\boldsymbol{v}_{02}\cdot\boldsymbol{r}_{02} = \boldsymbol{q}_{02}\cdot\boldsymbol{r}_{02} + \frac{1}{m_0}g_{01}\boldsymbol{r}_{01}\cdot\boldsymbol{r}_{02} + \left(\frac{1}{m_0}+\frac{1}{m_2}\right)g_{02}\boldsymbol{r}_{02}\cdot\boldsymbol{r}_{02} + \frac{1}{m_2}g_{12}\boldsymbol{r}_{02}\cdot\boldsymbol{r}_{12},$$

$$\boldsymbol{v}_{12}\cdot\boldsymbol{r}_{12} = \boldsymbol{q}_{12}\cdot\boldsymbol{r}_{12} - \frac{1}{m_1}g_{01}\boldsymbol{r}_{01}\cdot\boldsymbol{r}_{12} + \frac{1}{m_2}g_{02}\boldsymbol{r}_{02}\cdot\boldsymbol{r}_{12} + \left(\frac{1}{m_2}+\frac{1}{m_1}\right)g_{12}\boldsymbol{r}_{12}\cdot\boldsymbol{r}_{12},$$

to satisfy the three constraints, each of these equations must evaluate to zero. However, note that all the dot products of the form $\boldsymbol{r}_{ij}\cdot\boldsymbol{r}_{kl}$ are simple functions of the geometry of the rigid molecule. We may write the equations in the form of a matrix equation for a vector $\boldsymbol{g} = (g_{01}, g_{02}, g_{12})$ as follows:

$$\begin{pmatrix} \left(\frac{1}{m_0}+\frac{1}{m_1}\right)\boldsymbol{r}_{01}\cdot\boldsymbol{r}_{01} & \frac{1}{m_0}\boldsymbol{r}_{01}\cdot\boldsymbol{r}_{02} & -\frac{1}{m_1}\boldsymbol{r}_{01}\cdot\boldsymbol{r}_{12} \\ \frac{1}{m_0}\boldsymbol{r}_{01}\cdot\boldsymbol{r}_{02} & \left(\frac{1}{m_0}+\frac{1}{m_2}\right)\boldsymbol{r}_{02}\cdot\boldsymbol{r}_{02} & \frac{1}{m_2}\boldsymbol{r}_{02}\cdot\boldsymbol{r}_{12} \\ -\frac{1}{m_1}\boldsymbol{r}_{01}\cdot\boldsymbol{r}_{12} & \frac{1}{m_2}\boldsymbol{r}_{02}\cdot\boldsymbol{r}_{12} & \left(\frac{1}{m_2}+\frac{1}{m_1}\right)\boldsymbol{r}_{12}\cdot\boldsymbol{r}_{12} \end{pmatrix} \begin{pmatrix} g_{01} \\ g_{02} \\ g_{12} \end{pmatrix} = \begin{pmatrix} -\boldsymbol{q}_{01}\cdot\boldsymbol{r}_{01} \\ -\boldsymbol{q}_{02}\cdot\boldsymbol{r}_{02} \\ -\boldsymbol{q}_{12}\cdot\boldsymbol{r}_{12} \end{pmatrix}. \tag{A5}$$

The nine elements of the matrix on the left are constant, since the molecule is rigid. This matrix is nonsingular and may be constructed once and its inverse computed at the beginning of the simulation. The three-vector on the right changes with every dynamical time step, but it is known at the time we need to solve for g_{01}, g_{02}, and g_{12}, which we do by simply multiplying the right side of Eq. (A5) by the inverse of the matrix on the left. Once the $\boldsymbol{g}$ values are known we can very easily use them with Eqs. (A3) and (A4) to update the velocities so that the constraints are satisfied.

This new noniterative approach for addressing velocity constraints for fully rigid molecules (''M_RATTLE'') is much faster than iterative schemes.

[1] J. W. Pitera and W. C. Swope, Proc. Natl. Acad. Sci. U.S.A. **100**, 7587 (2003).
[2] P. J. Rossky and Y. K. Cheng, Nature (London) **392**, 696 (1998).
[3] P. P. Ewald, Ann. Phys. (Leipzig) **64**, 253 (1921).
[4] L. Onsager, J. Am. Chem. Soc. **58**, 1486 (1936).
[5] S. W. De Leeuw, Proc. R. Soc. London, Ser. A **373**, 27 (1980).
[6] T. Darden, D. York, and L. Pedersen, J. Chem. Phys. **98**, 10089 (1993).
[7] M. Lisal, J. Kolafa, and I. Nezbeda, J. Chem. Phys. **117**, 8892 (2002).
[8] J. W. Pitera, K. Armstrong, and W. C. Swope (in preparation).
[9] D. van der Spoel, P. J. van Maaren, and H. J. C. Berendsen, J. Chem. Phys. **108**, 10220 (1998).
[10] R. H. Zhou, B. J. Berne, and R. Germain, Proc. Natl. Acad. Sci. U.S.A. **96**, 14931 (2001).
[11] P. J. Steinbach and B. R. Brooks, Chem. Phys. Lett. **226**, 447 (1994).
[12] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans, in *Interaction Models for Water in Relation to Protein Hydration. In Intermolecular Forces*, edited by B. Pullmann (Reidel, Dordrecht, 1981), p. 331.
[13] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, J. Phys. Chem. **91**, 6269 (1987).
[14] W. L. Jorgensen, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[15] A. Glaettli, X. Daura, and W. F. van Gunsteren, J. Chem. Phys. **116**, 9811 (2002).
[16] A. Glaettli, X. Daura, and W. F. van Gunsteren, J. Comput. Chem. **24**, 1087 (2003).
[17] M. W. Mahoney and W. L. Jorgenson, J. Chem. Phys. **112**, 8910 (2000).
[18] B. Chen, J. Xing, and J. I. Siepmann, J. Phys. Chem. B **104**, 2391 (2000).
[19] S. W. Rick, S. J. Stuart, and B. J. Berne, J. Chem. Phys. **101**, 6141 (1994).
[20] C. J. Burnham and S. S. Xantheas, J. Chem. Phys. **116**, 1479 (2002).
[21] H. Yu, T. Hansson, and W. F. van Gunsteren, J. Chem. Phys. **118**, 221 (2003).
[22] J. Jeon, A. E. Lefohn, and G. A. Voth, J. Chem. Phys. **118**, 7504 (2003).
[23] P. Ren and J. W. Ponder, J. Phys. Chem. B **107**, 5933 (2003).
[24] G. Lamoureux, A. D. MacKerell, Jr., and Benoit Roux, J. Chem. Phys. **119**, 5185 (2003).
[25] Alex MacKerell, David Spellmeyer, Wendy Cornell, James Caldwell, Chris Bailey (personal communication).
[26] W. S. Benedict, N. Gailar, and E. K. Plyer, J. Chem. Phys. **24**, 1139 (1956).
[27] T. A. Andrea, W. C. Swope, and H. C. Andersen, J. Chem. Phys. **79**, 4576 (1983).
[28] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, 1st ed. (Clarendon, Oxford, 1987)
[29] H. C. Andersen, J. Chem. Phys. **72**, 2384 (1980).
[30] W. C. Swope, H. C. Andersen, H. P. Berens, and K. R. Wilson, J. Chem. Phys. **76**, 637 (1982).
[31] K. A. Feenstra, B. Hess, and H. J. C. Berendsen, J. Comput. Chem. **20**, 786 (1999).
[32] V. Kraeutler, W. F. van Gunsteren, and P. H. Huenenberger, J. Comput. Chem. **22**, 501–508 (2001).
[33] See Appendix.
[34] amu = atomic mass unit.
[35] 1999 CODATA formulation (see http://physics.nist.gov/cuu/Constants/index.html)
[36] IUPAC Pure Appl. Chem. **66**, 2423 (1994).
[37] See EPAPS Document No. E-JCPSA6-512415 for TIP4P-Ew supplemental material. A direct link to this document may be found in the online article's HTML reference section. The document may also be reached 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.
[38] M. Z. Hernandes, J. B. P. da Silva, and R. L. Longo, J. Comput. Chem. **24**, 1 (2003).
[39] J. R. Taylor, *Introduction to Error Analysis*, 2nd ed. (University Science Books, Sausalito, 1997), Vol. 1, Chap. 8.
[40] W. Wagner and A. Pruss, J. Phys. Chem. Ref. Data **31**, 387 (2003).
[41] G. S. Kell, J. Chem. Eng. Data **20**, 97 (1975).
[42] The reference state for the gas phase at any given temperature T is the **real gas**.
[43] J. P. M. Postma, A molecular dynamics study of water. PhD thesis, University of Groningen. the Netherlands (1985).
[44] Fundamental vibrational frequencies ν (cm^{-1}); from infrared (IR) spectra of water gas and liquid from Ref. 43: intra-molecular(gas): {3755.7, 3656.6, 1594.6}; intra-molecular(liquid): {3490, 3280, 1645}; inter-molecular(liquid): {800, 500, 200, 50}.
[45] The experimental values used for μ_{gas} and α_{gas} are: $\mu_{\mathrm{gas}} = 1.854\,989$ D and $\alpha_{\mathrm{gas}} = 1.6633\ 10^{-40}$ Fm^2 as recommended by IAPWS; Guideline on the Use of Fundamental Physical Constants and Basic Constants of Water (2001).
[46] W. L. Jorgensen and J. D. Madura, Mol. Phys. **56**, 1381 (1985).
[47] G. N. Lewis and M. Randall, *Thermodynamics*, 2nd ed. (McGraw-Hill, New York, 1961), Chap. 16.
[48] *CRC Handbook of Chemistry and Physics*, 75th ed., edited by D. R. Lide (CRC, Boca Raton, 1994), Vol. 1, Chap. 6.
[49] D. G. Archer, J. Phys. Chem. B **104**, 8563 (2000).
[50] M. Neumann, Mol. Phys. **50**, 841 (1983).
[51] W. A. B. Evans and J. G. Powles, Mol. Phys. **45**, 695 (1982).
[52] A. D. Buckingham, Proc. R. Soc. London, Ser. A **238**, 235 (1956).
[53] D. P. Fernandez, A. R. H. Goodwin, E. W. Lemmon, J. M. Levelt Singers, and R. C. Williams, J. Phys. Chem. Ref. Data **26**, 1125 (1997).
[54] G. Hura, D. Russo, R. M. Glaeser, and T. Head-Gordon, Phys. Chem. Chem. Phys. **5**, 1981 (2003).
[55] P. E. Smith and W. F. van Gunsteren, J. Chem. Phys. **100**, 3169 (1994).
[56] J. P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen, J. Comput. Phys. **23**, 327 (1977).
[57] H. C. Andersen, J. Comput. Phys. **52**, 24 (1983).
[58] K. T. Gillen, D. C. Douglas, and M. J. R. Hoch, J. Chem. Phys. **57**, 5117 (1972).
[59] R. Mills, J. Phys. Chem. **77**, 685 (1970).
[60] W. L. Jorgensen and C. Jenson, J. Comput. Chem. **19**, 1179 (1998).
[61] M. W. Mahoney and W. L. Jorgenson, J. Chem. Phys. **114**, 363 (2001).
[62] K. Watanabe and M. L. Klein, Chem. Phys. **131**, 157 (1989).