Quantum mechanical force field for water with explicit electronic polarization

Jaebeom Han,[a] Michael J. M. Mazack,[a),b)] Peng Zhang, Donald G. Truhlar,[b)] and Jiali Gao[b)]

Department of Chemistry and Supercomputing Institute, University of Minnesota, 207 Pleasant Street, SE, Minneapolis, Minnesota 55455-0431, USA

(Received 26 April 2013; accepted 8 July 2013; published online 1 August 2013)

A quantum mechanical force field (QMFF) for water is described. Unlike traditional approaches that use quantum mechanical results and experimental data to parameterize empirical potential energy functions, the present QMFF uses a quantum mechanical framework to represent intramolecular and intermolecular interactions in an entire condensed-phase system. In particular, the internal energy terms used in molecular mechanics are replaced by a quantum mechanical formalism that naturally includes electronic polarization due to intermolecular interactions and its effects on the force constants of the intramolecular force field. As a quantum mechanical force field, both intermolecular interactions and the Hamiltonian describing the individual molecular fragments can be parameterized to strive for accuracy and computational efficiency. In this work, we introduce a polarizable molecular orbital model Hamiltonian for water and for oxygen- and hydrogen-containing compounds, whereas the electrostatic potential responsible for intermolecular interactions in the liquid and in solution is modeled by a three-point charge representation that realistically reproduces the total molecular dipole moment and the local hybridization contributions. The present QMFF for water, which is called the XP3P (explicit polarization with three-point-charge potential) model, is suitable for modeling both gas-phase clusters and liquid water. The paper demonstrates the performance of the XP3P model for water and proton clusters and the properties of the pure liquid from about 900 × 10^6 self-consistent-field calculations on a periodic system consisting of 267 water molecules. The unusual dipole derivative behavior of water, which is incorrectly modeled in molecular mechanics, is naturally reproduced as a result of an electronic structural treatment of chemical bonding by XP3P. We anticipate that the XP3P model will be useful for studying proton transport in solution and solid phases as well as across biological ion channels through membranes. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4816280]

I. INTRODUCTION

Critical to the success of dynamical simulations of chemical and biological systems is the potential energy function used to describe intermolecular interactions.[1,2] Because of the importance of aqueous solution and its unique roles in biomolecular interactions, water has been a subject of extensive and continuous investigation (a review in 2002 included a partial list of 46 water models,[3] while at least two dozen new models have appeared since that time).[4,5] An accurate and efficient model for liquid water also serves as an anchor for developing force fields for proteins, nucleic acids, and carbohydrates. Traditionally, the Lifson-type of effective, pairwise potentials have been used,[1,2,6] and much effort has also been devoted to incorporating many-body polarization effects into such force fields.[7] However, unlike the development of pairwise potentials, there is a great deal of uncertainty in the treatment of polarization effects, both in the choice of functional form and in the associated parameters. This is reflected in the fact that simple point charge models such as SPC,[8] TIP3P and TIP4P[9] quickly emerged as the standards in the 1980s for biomolecular force fields, but no standard polarizable force fields have emerged although dozens of polarizable potentials for water have been proposed.[3,4,10] We have developed a quantum mechanical framework in which each individual molecular fragment is treated by electronic structure theory.[11–14] Since polarization effects are naturally included in the self-consistent field (SCF) optimization of molecular wave functions, we call this method the explicit polarization (X-Pol) theory.[11,14,15] Recent studies demonstrated the feasibility of X-Pol as a next generation force field for biomolecular simulations,[13] and encouraging results have been obtained using standard semiempirical Hamiltonians.[12,16] In the present paper we report a novel model for water, called XP3P, based on X-Pol theory and a three-point charge representation of the electrostatic potential, as a first step in our effort to develop a full quantum mechanical X-Pol force field for biomolecular and materials simulations.

The present quantum mechanical force field (QMFF) may be compared with phenomenological representations of electronic polarization in three commonly used methods in molecular mechanics, namely induced-dipole, Drude-oscillator, and fluctuating-charge models. In the

[a)] J. Han and M. J. M. Mazack contributed equally to this work.
[b)] Electronic addresses: mazack@mazack.org; truhlar@umn.edu; and gao@jialigao.org. Telephone: 1-612-625-0769. Fax: 1-612-624-0741.

induced-dipole approach,[17–20] atomic polarizabilities are assigned to the interaction sites, typically located on, but not limited to, atomic centers, from which induced point dipoles, representing the total electric field of the system, are obtained.[21] A commonly used method to assign atomic polarizabilities is the dipole interaction model (DIM) popularized by Applequist *et al.*[22] and extended by Thole[23] to incorporate short-range damping functions. Remarkably, the values optimized in DIM are quite transferrable,[24] requiring typically one parameter per element. The Drude-oscillator model may be considered as a point-charge equivalent of the induced-dipole method.[25,26] Here, one or a set of point charges are harmonically linked to a polarizable site, in which the directions and distances of the Drude oscillators give rise to the corresponding induced dipole moments. The fluctuating-charge[27–30] approach employs a chemical potential equalization scheme, in which the instantaneous partial charges minimize the energy of the system. The fundamental parameters used in the fluctuating-charge model correspond to the atomic electronegativity and hardness that are rigorously defined in density functional theory.[31]

Each of these classical methods has its advantages and shortcomings in practice. In the fluctuating-charge model, unphysical charge transfer effects between distant monomers can occur. Thus, charge constraints are required. On the other hand, the induced-dipole and the Drude-oscillator model are difficult to use for representing molecular polarization involving a significant charge delocalization such as that across a conjugated polyene chain and the polarization of push-pull compounds (e.g., the crystal of *p*-nitroaniline). The Drude-oscillator model has the advantage of simplicity in practice since any dynamics simulation code can be conveniently adapted to treat polarization effects by that method.

The X-Pol method relies on the partition of a large, condensed-phase system into molecular or submolecular fragments (or blocks),[11,12,15] which can be single solvent molecules like water, amino acid residues or nucleotide bases, small ions or enzyme cofactors, or a collection of these small units. The wave function of each molecular fragment is described by a Slater determinant of block-localized molecular orbitals that are expanded over basis functions located on atoms of the fragment. The total molecular wave function is approximated as a Hartree product of these fragmental, determinant functions. Consequently, Coulombic interactions between different fragments are naturally incorporated into the Hamiltonian, but short-range exchange repulsion, charge delocalization (also called charge transfer) and long-range dispersion interactions are not explicitly treated in the quantum chemical formalism.[32–34] These effects are included and optimized empirically to strive for accuracy (and efficiency) in X-Pol in the same spirit as that in force field development. The determinantal wave function for each monomer fragment can be approximated by wave function theory (WFT) at either an *ab initio* or a semiempirical level,[12,35] the density may be approximated by density functional theory (DFT),[35,36] or one can combine levels of theory,[37] but in this paper we use only semiempirical wave function theory. Although the present work involves only water, we note that the X-Pol theory can be used to model electronic polarization involving conjugated systems and significant charge delocalization contributions,[38] and the X-Pol model is also a reactive force field for modeling systems involving bond-forming and bond-breaking processes.

Semiempirical methods employing neglect diatomic differential overlap (NDDO)[39] are especially suited for QMFF development because of their computational efficiency. However, most such semiempirical models were not optimized to describe intermolecular interactions that are essential for modeling condensed-phase systems.[40–43] Part of the problems has been remedied through the incorporation of empirically damped dispersion functions.[44–49] Another important deficiency of many semiempirical models for treating nonbonded interactions is that molecular polarization is systematically underestimated. Recently, we have introduced a polarized molecular orbital (PMO) alternative,[49–51] in which a set of *p*-orbitals are added to each hydrogen atom.[52] It was found that the computed molecular polarizabilities for a range of compounds containing hydrogen, carbon, and oxygen are significantly improved.[49,51] Employing this strategy, we report here a parametrization of the PMO model for water (PMOw), which can be used in X-Pol for liquid simulations.

In the following, Sec. II summarizes the PMO parameterization for water and the development of the XP3P model liquid water. Computational details are given in Sec. III. In Sec. IV, we present results and discussion. Section V concludes the paper with a summary of major findings.

II. METHOD

The X-Pol quantum mechanical force field is designed to model condensed phase systems with or without bond-forming and bond-breaking processes. Thus, the X-Pol method can be used as a general-purpose force field in dynamics simulations of solvated proteins or as a reactive force field to model chemical reactions in solutions and in enzymes. In this section, we first describe the quantum chemical model designated as PMOw for water and compounds containing oxygen and hydrogen atoms. The acronym PMO is used to describe the general semiempirical model in which, in addition to a minimal basis set, a set of *p*-orbitals is added to hydrogen atoms.[49,51] Then, we highlight its incorporation in X-Pol, called the XP3P model, for simulation of liquid water.

II.A. Polarized molecular orbital model for water

The PMOw model is a new parameterization of the PMO method,[49] which is based on the MNDO formalism[53] with three key enhancements. First, a set of diffuse *p*-type basis functions is added on the hydrogen atoms.[50] This greatly improves the quality of the computed molecular polarizabilities and hence the treatment of hydrogen bonding interactions. Second, a damped dispersion function, following the work of Tang and Toennies in wave function theory[54] and Grimme in density functional theory,[55] is included as a post-SCF correction to the electronic energy. In the present implementation, we have adopted the method and parameters proposed by Hillier and co-workers in the PM3-D method.[44–46]

TABLE I. Semiempirical parameters for H and O Atoms in the PMOw model.

	H	O
U_{ss} (eV)	− 11.15043	− 111.86028
U_{pp} (eV)	− 7.35459	− 78.64105
β_s (eV)	− 6.88125	− 25.57063
β_p (eV)	− 3.52628	− 31.90404
ζ_s (bohr^{-1})	1.17236	3.05303
ζ_p (bohr^{-1})	1.05333	3.12265
α (Å^{-1})	3.05440	3.76880
g_{ss} (eV)	12.73667	17.36659
g_{sp} (eV)	8.04688	13.37288
g_{pp} (eV)	6.98401	14.78196
$g_{pp'}$ (eV)	10.65161	13.49319
h_{sp}[a] (eV)	1.92149	4.42643

[a]The derived parameter, h_{pp}, is determined from g_{pp} and $g_{pp'}$ and has been set to a minimum value of 0.1 eV as implemented in the MOPAC program, $h_{pp} = \max\{0.1\,eV, (g_{pp} - g_{pp'})/2\}$.

TABLE II. Additional semiempirical parameters for oxygen and hydrogen in the polarized molecular orbital model and the Lennard-Jones parameters in explicit polarization model for liquid water.

Parameter	Value
A_{sp}	0.03000
A_{pp}	0.15000
B	1.00000
κ_{sp} (Å^{-1})	0.47069
κ_{pp} (Å^{-1})	0.47069
λ (Å^{-2})	1.10000
$\hat{\alpha}^{H}$ (Å^{-1})	2.52552
$\hat{\alpha}^{O}$ (Å^{-1})	3.03253
ζ^{HH} (bohr^{-1})	1.28000
ζ^{OO} (bohr^{-1})	2.76400
σ_H (Å)	0.800
σ_O (Å)	3.225
ε_H (kcal/mol)	0.05
ε_O (kcal/mol)	0.15

The inclusion of the damped dispersion terms further improves the description of intermolecular interactions and the performance of PMO on small molecular clusters.[44–47,49,51] Third, the PMOw model is parameterized for general applications to a specific class of compounds (see Sec. IV.A for the set of parametrization data), and the optimization targets include molecular polarizabilities and non-bonded interactions as well as other properties used in the traditional semiempirical parameterization.[49] The parameters presented here are optimized for compounds containing oxygen and hydrogen atoms, especially for studying liquids, aqueous solutions, and proton transport. We note here that, in the same way that atoms are assigned types in molecular mechanics, the parameters for oxygen and hydrogen atoms in functional groups other than water (e.g., peptide bonds) need not be restricted to the same as used for such atoms in water. This departs from the philosophy that has usually been used in semiempirical methods,[56,57] in which general atomic parameters are used for all functionalities.

In the MNDO formalism,[53,58] there are 12 atomic parameters for each element, and the PMOw values for water and other compounds containing oxygen and hydrogen are listed in Table I. These values are similar in many respects to the PMOv1 model introduced previously,[49] but they result from a new parametrization presented below. Three exceptions were made to the MNDO functional forms because of the addition of diffuse p basis functions on hydrogen atoms,[49] and they are listed as follows:

1. For the resonance integral involving p orbitals on hydrogen, the following conventions are used:

$$\beta_{lp}^{HH} = 0, \tag{1a}$$

$$\beta_{lp}^{OH} = \frac{\beta_l^O + \beta_p^H}{2} S_{lp} A_{lp} e^{\kappa_{lp} R_{OH}}, \tag{1b}$$

where l is the angular momentum quantum number, having the values of 0 (s orbital) and 1 (p orbital), and the subscript p denotes a p-orbital on hydrogen. Notice that Eq. (1b) is slightly different from the expression used in Ref. 49, in which the exponential function is absent. In Eq. (1b), β_l^O and β_p^H are standard MNDO-type parameters, A_{lp} and κ_{lp} are additional parameters introduced in PMO, and R_{OH} is the distance between oxygen and hydrogen atoms. S_{lp} in Eq. (1b) is an overlap integral ($\langle O_l | H_p \rangle$) between oxygen and hydrogen Slater-type orbitals using the parameters listed in Table I, but specific exponents, ζ^{OO} and ζ^{HH}, are used for H–H and O–O pairs, respectively, in PMOw.
2. In standard MNDO,[53,58] the nucleus-electron attraction integral, $H_{\mu\nu}^A$, between electronic charge density on atom A and nucleus B is evaluated on the basis of the two-electron repulsion integral, $\langle \mu_A \nu_A | s_B s_B \rangle$.[59] In PMOw, if both A and B are hydrogen atoms, for a distribution of p orbitals (pp'), this is screened as follows:

$$H_{pp'}^{H} = \left[1 - Be^{-\lambda R_{HH'}^2}\right] \left(H_{pp'}^{H}\right)_{MNDO}. \tag{2}$$

3. For the homonuclear core-core repulsion integrals,[49,53,58] the standard values for α^O and α^H are replaced by $\hat{\alpha}^O$ and $\hat{\alpha}^H$. Note that α^O and α^H are used as in standard MNDO for core-core repulsion integrals between oxygen and hydrogen atoms.

The parameters in the standard MNDO formalism[53] (Table I) and the additional parameters (Table II) described above were adjusted by iterative optimization using a genetic algorithm for some of the systems and properties listed in Table S1 in the supplementary material.[149] In comparison with the results in Ref. 49, the present parameter set further improves the calculated molecular polarizability and dipole moment of water in the gas phase as well as the binding energy and dipole moment of water dimer (Table III).

II.B. Explicit polarization theory

In X-Pol,[11,12,15] the system is partitioned into molecular or submolecular fragments, in which the total wave function of the system is assumed to be a Hartree product of the determinant wave functions of the individual fragments. In the

TABLE III. Computed equilibrium properties for water monomer and dimer from different polarizable water models and *ab initio* MP2/(CBS) with CCSD(T) corrections along with experimental data.

		PMOw	XP3P	AMOEBA	POL5.TZ	*Ab initio*	Expt.[99]
H_2O	AE (kcal/mol)	233.0	233.0			229.3[100]	232.2
	IP (eV)	13.20	13.20			12.42	12.68
	r (Å)	0.955	0.957	0.957	0.957	0.9589[103]	0.9572
	θ (deg)	104.6	104.5	108.5°	104.5	104.16[103]	104.52
	α (Å^3)	1.27	1.27	1.41	1.29	1.45[102]	1.45
	q^{H} (e)	0.16	0.34	0.26		0.35	N/A
	q^{O} (e)	−0.31	−0.67	−0.52		−0.70	N/A
	μ (debye)	1.88	1.88	1.77	1.85	1.84[102]	1.86[76]
$(H_2O)_2$	ΔE_b	−5.1	−5.2	−4.96	−4.96	−5.0[77]	−5.44
	R_{OO}	2.89	2.90	2.89	2.90	2.92	2.98
	α	6.2	1.3	4.2	4.7	4.8	-1 ± 10
	ϕ	115	165	123	117	125	123 ± 10
	$\langle\mu_{\mathrm{mol}}\rangle$	2.10	2.16				
	μ	2.39	3.85	2.54	2.44	2.65	2.64

present case, each fragment is simply a single water molecule, and the overall wave function is

$$\Phi = \prod_{a=1}^{N} \Psi_a, \tag{3}$$

where N is the number of fragments in the system, and Ψ_a is a Slater determinant of doubly-occupied molecular orbitals (MOs) block-localized on molecule (fragment) a. The approximation of Eq. (3) implies neglect of the short-range exchange repulsion[33] and long-range dispersion interactions[60] between different fragments, which are corrected empirically below.[11,12,15] Use of Eq. (3) reduces the computational costs, allowing molecular dynamics and Monte Carlo simulations to be carried out for large systems efficiently with sufficient sampling.[12,13]

The effective Hamiltonian of the system is given by

$$H = \sum_{a=1}^{N} H_a^o + \frac{1}{2}\sum_{a=1}^{N}\sum_{b\neq a}^{N} H_{ab}, \tag{4}$$

where H_a^o is the electronic Hamiltonian of fragment a in the gas phase and H_{ab} represents the effective interactions between molecules a and b:

$$H_{ab}(\rho_b) = -\sum_{i=1}^{M} V_i(\rho_b) + \sum_{A=1}^{Q} Z_A^a V_A(\rho_b) + E_{ab}^{\mathrm{XD}}, \tag{5}$$

where M is the number of electrons and Q is the number of atoms in fragment a, Z_A^a would be the nuclear charge of atom A of fragment a if all electrons were treated explicitly but here it is the core charge since $1s$ electrons of oxygen atoms are in the core, and E_{ab}^{XD} is the exchange-dispersion correlation energy. The electrostatic potential $V_x(\rho_b)$, either at the electronic ($x = i$) or at the nuclear ($x = A$) position, due to the instantaneous charge density of fragment b is given by

$$V_x(\rho_b) = -\int \frac{\rho_b(\mathbf{r})d\mathbf{r}}{|\mathbf{r}_x - \mathbf{r}|} + \sum_{B=1}^{Q} \frac{Z_B^b}{|\mathbf{r}_x - \mathbf{R}_B^b|}. \tag{6}$$

Here, $\rho_b(\mathbf{r})$ is the electron density of fragment b derived from the corresponding wave function Ψ_b (or Kohn–Sham Slater determinant),[11,12] and $\mathbf{R}_B^b$ denotes the nuclear coordinates.

We define the total interaction energy of a condensed phase system by

$$E_{\mathrm{tot}} = \langle\Phi|H|\Phi\rangle - \sum_{a=1}^{N}\langle\Psi_a^o|H_a^o|\Psi_a^o\rangle. \tag{7}$$

The energy defined in Eq. (7) corresponds to the total energy of the condensed-phase system relative to that of infinitively separated fragments. Since all molecules are identical in pure liquid water in the present study, the last summation term in Eq. (7) is simply NE_a^o with $E_a^o = \langle\Psi_a^o|H_a^o|\Psi_a^o\rangle$ being the energy of an isolated monomer. It is often useful for interpretive purposes to consider the dimeric interaction energies between two fragments even for a potential that includes many-body polarization effects as in the present X-Pol potential. To this end, we define the interaction energy between fragments a and b by[12]

$$E_{ab} = \frac{1}{2}(\langle\Psi_a|H_{ab}|\Psi_a\rangle + \langle\Psi_b|H_{ba}|\Psi_b\rangle). \tag{8}$$

The two terms in Eq. (8) corresponds to a embedding in b and b embedding in a, respectively, both in the presence of the rest of the system, and they are not always numerically equivalent in practice[11] even though they describe the same intermolecular interactions. The definition of Eq. (8) ensures that $E_{ab} = E_{ba}$.

The exchange-dispersion correlation energy can be incorporated with an explicit density dependent term and added to the Fock operator as described in the work of York and coworkers.[32,61] Alternatively, the damped dispersion term that is an intrinsic part of the PMOw model can be used with the addition of a repulsive potential. Here, in the spirit of simplicity for a force field, we adopt a Lennard-Jones potential to approximate the remaining energy contributions[11,12,15] not included in the PMOw electronic structure method.[49] (Thus there are two R^{-6} terms, one in PMOw for

intrafragment interactions and one in the Lennard-Jones term associated with interfragment interactions.) The Lennard-Jones term introduces two empirical parameters per atom type:

$$E_{ab}^{\mathrm{XD}} = \sum_{A}^{Q}\sum_{B}^{Q} 4\varepsilon_{AB}\left[\left(\frac{\sigma_{AB}}{R_{AB}}\right)^{12} - \left(\frac{\sigma_{AB}}{R_{AB}}\right)^{6}\right], \quad (9)$$

where ε_{AB} and σ_{AB} are obtained from the geometric mean of atomic parameters such that $\varepsilon_{AB} = (\varepsilon_A\varepsilon_B)^{1/2}$ and $\sigma_{AB} = (\sigma_A\sigma_B)^{1/2}$. These parameters are also listed in Table II.

II.C. The XP3P model for liquid water

The electrostatic potential (ESP) in Eq. (6) can be determined explicitly by evaluating the associated one and two-electron integrals in SCF calculations. However, this would have not saved much computational time, and would have missed the point of developing a fragment-based technique in electronic structure calculations. As we have proposed previously,[11,15,35] it is desirable to employ a more computationally efficient method to approximate the external potential $V_x(\Psi_b)$. In the present application to liquid water, we use a simple, three-point-charge approximation to $V_x(\Psi_b)$. Consequently, we call this X-Pol potential with three-point charges for water the XP3P model.

Several methods based on atomic partial charges for approximating the quantum external potential were described originally for the X-Pol potential,[11,12] and some of them were adopted later in other fragment-based molecular orbital models.[62] Although the use of atomic charges obtained from fitting the quantum mechanical $V_x(\Psi_b)$ has been successfully used in several molecular mechanics force fields,[63,64] it is known that the ESP-fitting method sometimes yields unreasonably large partial charges on structurally buried atoms.[65] In addition, large variations could occur as a result of structural fluctuations to expose buried atoms during a dynamics simulation. A general approach is the multi-center multipole expansion of the quantum mechanical ESP,[66] and this method has been used in the effective fragment potential model;[67] multi-center multipolar representations could also be used with X-Pol.[68] A conceptually simple alternative is to use atomic charges derived from a population analysis such as the Mulliken or Löwdin population method.[69] When used with small, well balanced basis sets, the Mulliken or Löwdin charges can provide a good representation of the relative atomic electronegativity and they are computationally efficient. Scaled Mulliken population charges have been used and shown to be effective in statistical mechanical Monte Carlo simulations of liquid water using an explicit QMFF.[12]

Another way of approximating the external potential for intermolecular interactions is to employ partial atomic charges that are mapped from the density matrix to reproduce experimental dipole moments (in contrast to ESP fitting). This has been called a class IV charge model, and it can be parametrized to show good consistency for a variety of electronic structure methods and basis sets.[70,71] Alternatively, partial atomic charges can be derived to rigorously reproduce the molecular moments to any order of accuracy from a Lagrangian multiplier procedure. Following the method proposed by Thole and van Duijnen[72] and extended by Swart and van Duijnen,[73] we applied the Lagrangian multiplier approach to semiempirical methods,[74] which are known to yield excellent molecular dipole moments in comparison with experiments. In this approach, both the total molecular dipole moment and the local atomic hybridization contributions of the approximate NDDO wave function are reproduced exactly. In the present implementation, we preserve the total and local molecular dipole moments. In addition, we included in the procedure the capability to reproduce experimental molecular polarizability and its atomic decomposition according to the dipole interaction model.[74] We called this method the dipole preserving and polarization consistent (DPPC) charge model.[74]

Specifically, the DPPC charge has two contributions, the Mulliken population charge q_A^{MP} and the residual charges Δq_A^B due to preservation of atomic s and p hybridization dipole moments:[74]

$$q_A^{\mathrm{DPPC}} = q_A^{\mathrm{MP}} + \sum_{B=1}^{Q} \Delta q_A^B, \quad (10)$$

where the residual charge Δq_A^B on atom A due to the constraint that the residual moment is identical to the atomic hybridization contribution from atom B:

$$\boldsymbol{\mu}_B^{hyb} = -(\mathbf{P}_{sp})_B \cdot \mathbf{D}_B = \sum_{A=1}^{Q} \Delta q_A^B \mathbf{R}_A, \quad (11)$$

where $(\mathbf{P}_{sp})_B$ is a diagonal matrix with the densities $P_{sp_x}^B$, $P_{sp_y}^B$, and $P_{sp_z}^B$, on atom B, $\mathbf{D}_B$ is the corresponding dipole integral, and $\mathbf{R}_A$ denotes the coordinates of atom A. The residual charges Δq_A^B that reproduce the hybridization component of molecular dipole moment, $\boldsymbol{\mu}_B^{hyb}$, are predominantly localized on atoms closest to atom B. Since the molecular dipole moment is determined from

$$\boldsymbol{\mu}_{QM} = \sum_{A=1}^{Q} q_A^{\mathrm{MP}} \mathbf{R}_A + \sum_{B=1}^{Q} \boldsymbol{\mu}_B^{hyb} \quad (12)$$

in semiempirical methods employing the NDDO approximation,[75] it is clear that the atomic charges given in Eq. (10) reproduce exactly the full quantum mechanical dipole moment and the local, atomic hybridization contributions:

$$\sum_{A=1}^{Q} q_A^{\mathrm{DPPC}} \mathbf{R}_A = \boldsymbol{\mu}_{QM}. \quad (13)$$

The residual charges depend on geometry and atomic electronegativity, and an expression for them was given in Ref. 74. The advantage of using the DPPC charges over the ESP-fitted ones is that local properties of the dipole integrals are explicitly accounted for and fully utilized to generate the partial atomic charges. The method to generate DPPC charges

is applicable both to neutral and ionic molecules, independent of the origin of coordinates.[74]

III. COMPUTATIONAL DETAILS

The parameterization of the PMOw model was carried out by iterative optimization using a genetic algorithm that has been detailed in Ref. 49. The PMOv1 set of parameters overestimated the dipole moment of water (2.19 D) and underestimated the interaction energy for the water dimer (4.7 kcal/mol) in comparison with the target values of 1.85 D from experiment[76] and 5.0 kcal/mol from CCSD(T) and MP2/(CBS) calculations.[77] The PMOw parametrization improves these quantities for application to water and its ions.

Statistical mechanical Monte Carlo simulations were performed on a system consisting of 267 water molecules in a cubic box, employing the PMOw Hamiltonian. Based on procedures described previously,[12,16] periodic boundary conditions were used along with the isothermal-isobaric ensemble (NPT) at 1 atm and temperature ranging from −40 to 100 °C. As in the development of other empirical potentials including the successful SPC,[8] TIP3P, and TIP4P models[9] and the polarizable AMEOBA,[10] SWM4-NDP[26] and POL5/TZ[78] potentials for water (and many other water models not explicitly compared in this paper), the parameterization was performed only at 25 °C. The XP3P model based on the PMOw Hamiltonian has four Lennard-Jones parameters, ε_O, ε_H, σ_O, and σ_H. We have kept the ε_H and σ_H values used in a previous X-Pol simulation of liquid water with the AM1 Hamiltonian (called the MODEL potential for water), and we made small adjustments of the other two values (3.24 Å and 0.16 kcal/mol)[12] to reproduce the liquid density and heat of vaporization within 1% of the experimental values at 25 °C. In the parameterization stage, spherical cutoff with a switching function between 8.5 Å and 9.0 Å based on oxygen-oxygen separations was employed, and a long-range correction to the Lennard-Jones potential was included. (The SPC and TIP3P/TIP4P models[8,9] and later the TIP5P model[79] were also developed using cutoff distances, which were as small as 7.5 Å with a box of 125 or 216 water molecules.) Although it is possible to use Ewald sums to treat long-range electrostatic interactions,[80] we have not used the particle-mesh Ewald implementation in the present Monte Carlo calculation. In Monte Carlo simulations, new configurations were generated by randomly translating and rotating a randomly selected water molecule within ranges of ±0.13 Å and ±13°. In addition, the volume of the system was changed randomly within the limit of ±150 Å^3 on every 550th attempted move, and the coordinates of oxygen atoms were scaled accordingly. (Note that in the Monte Carlo calculations, the waters are rigid, so the hydrogen positions also adjust when the oxygen positions are adjusted.) These options were slightly adjusted to maintain an acceptance rate of about 45% at each temperature in the Metropolis sampling. In each simulation, at least 5×10^6 configurations were discarded for equilibration, which was followed by an additional 1×10^7 to 1.1×10^8 configurations for averaging. About 6×10^6 configurations can be executed per day on a 6-core Intel Xeon X7542 Westmere processor at 2.66 GHz.

The XP3P model was further examined in molecular dynamics simulations for 500 ps in the NVT ensemble, using the Lowe-Andersen thermostat[81,82] and a volume fixed at the average value from the Monte Carlo simulation; the number of water molecules in the dynamics simulations was also 267. The monomer geometries were enforced by the SHAKE/RATTLE procedure.[83] Although long-range electrostatic interactions can be computed using the particle-mesh Ewald summation that has been extended for the X-Pol potential,[80,84] we have used 9.0 Å cutoff in the present study. The velocity Verlet integration algorithm was used with a 1fs time step.

The total energy of the system was obtained from fully converged wave functions for each water molecule for each microscopic configuration, although different procedures were utilized in the Monte Carlo sampling and in molecular dynamics simulations. In Monte Carlo, an initial set of DPPC charges, derived from an initial guess of the X-Pol wave function, e.g., that from the previous configuration (with random perturbation to some randomly selected elements in the density matrix), are incorporated into the Fock matrix in terms of one-electron integrals (as in combined QM/MM schemes) in the subsequent iteration step during the self-consistent field (SCF) optimization. Then, a new set of orbital coefficients is obtained to generate updated DPPC charges for the next iteration until the electronic energy is converged to 5×10^{-5} eV for each monomer and to 10^{-5} for the partial atomic charges (in atomic units) between consecutive iterations. In Monte Carlo simulations, the Fock operator is constructed analogously to a combined QM/MM scheme,[85] which is not fully variational with respect to the change of the charge density; the external potential does incorporate the complete electrostatic effects in a self-consistent manner.[11,12] The procedure is efficient in Monte Carlo simulations since the electronic integrals are not required from all other molecular fragments, and it does not pose problems because gradients are not needed. This is the method proposed in the original development of the method for Monte Carlo calculations,[11,12] and it was used a few years later in the fragment molecular orbital model of Kitaura and co-workers.[86] For molecular dynamics simulations, a fully variational Fock operator for each monomer was used in which the external potential consists of contributions both from the DPPC charges and the explicit electron densities of all other fragments.[14,35] Here, analytic gradients can be directly obtained from the optimized X-Pol wave function. In molecular dynamics simulations, the criteria for energy and density conversion were set as 10^{-9} eV for energy and 10^{-6} for density matrix elements. The average energy difference from the two approaches in Monte Carlo and molecular dynamics is less than 1.5% in the computed heat of vaporization.

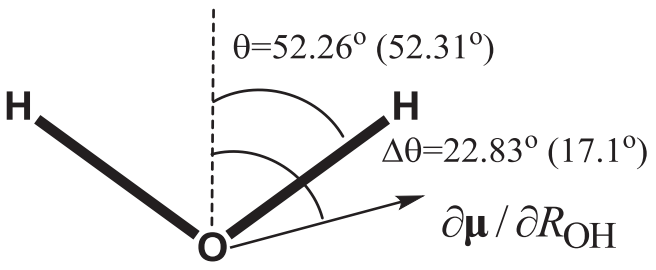
The Monte Carlo simulations were performed using the MCSOL program for X-Pol simulations,[87] while molecular dynamics simulations were carried out using a newly developed X-Pol program[88] written in C++ which has been interfaced both with CHARMM[89] and NAMD.[90] All *ab initio* electronic structure calculations were performed using GAUSSIAN 09.[91] All calculations were run on a constellation of clusters at the Minnesota Supercomputing Institute.

IV. RESULTS AND DISCUSSION

IV.A. Gas-phase properties

Properties for the optimized water monomer and dimer using the PMOw and XP3P models are listed in Table III along with experimental data and the results from two empirical polarizable potentials that have been examined by Ren and Ponder.[10] The PMOw parameters were optimized against experimental or high-level *ab initio* data for a series of small molecules containing hydrogen and oxygen atoms (supplementary material[149]), including the properties listed in Table III. In particular, the computed atomization energy (233.0 kcal/mol) and dipole moment (1.88 D) for water from PMOw agree with the corresponding experimental data that have been summarized in Ref. 49 (232.6 kcal/mol and 1.85 D, respectively). The Mulliken population charges from the PMOw wave function and the DPPC charges used in the XP3P potential are also listed in Table III; the latter yields exactly the same molecular dipole moment as that from the QM calculation. An important quantity critical to describing hydrogen-bonding interactions is the molecular polarizability, which also shows good agreement with experiment (a deviation of 14%). This represents a major improvement over all previous NDDO-based models, which typically have errors more than 60% for water. Nevertheless, a question arises on whether or not the somewhat smaller polarizability would affect liquid properties. To address this issue, it is interesting to consider polarizable potential functions for water, in which the experimental gas-phase electrostatic properties are not always enforced. This is illustrated by the use of smaller molecular polarizabilities in these empirical force fields, and this was justified as to reflect the relatively larger electric field than the mean field of the bulk due to the highly inhomogeneous environment in the first solvation shell;[92] for example, polarizabilities are set to 1.41, 1.29, and 0.98 Å^3 in the AMOEBA,[10] POL5/TZ[10,78] and SWM4-NDP[26] models, respectively, all of which yield similar heats of vaporization and similar densities of liquid water at ambient conditions.

The optimized bond length and bond angle for water are 0.9552 Å and 104.61° using PMOw; these values are in excellent agreement with the experimental values of 0.9572 Å and 104.54°.[93] Thus, either the optimized or the experimental monomer geometry can be used in the XP3P potential for liquid simulations discussed below. The change of the molecular dipole moment with geometry variation for the water monomer has an intriguing nonlinear dependence, which is not correctly reproduced in nearly all polarizable and nonpolarizable potentials for water, except the TTM2-F model[94] that was specifically fitted with a function to reproduce an accurate *ab initio* dipole moment surface.[95] This is illustrated in Scheme 1, which shows that the dipole derivative with respect to an O–H stretch, $\partial\boldsymbol{\mu}/\partial R_{\mathrm{OH}}$, lies significantly outside of the two O–H bonds of water. An angle of $\Delta\theta = 22.8°$ was obtained based on the vibrational absorption intensities.[94,96,97] For comparison, the present PMOw model yields a value of $\Delta\theta = 17.1°$, in reasonable agreement with experiment. This is encouraging since this information was not included in the PMOw parametrization process; it is purely a result of the qualitatively correct treatment of chemical bonding interactions in the present quantum mechanical model.



SCHEME 1. Illustration of the angle between the molecular dipole moment derivative and the O–H bond vector in water monomer. Experimental values are given first, followed by the PMOw results in parentheses.

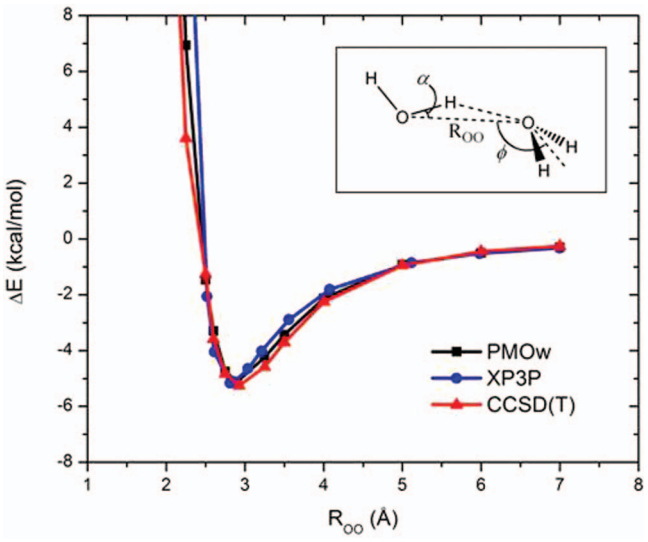


FIG. 1. Potential energy profiles for a water dimer at the hydrogen bonding configuration from the PMOw (black) and the XP3P (blue) models for water along with CCSD(T) results (red). Definition of the geometrical parameters listed in Table III are given in the structure shown as inset in the upper right-hand corner. The CCSD(T) results are obtained with the aug-cc-pVDZ basis set on fully optimized geometries at various fixed O–O distances. Studies have shown that extrapolation to the complete basis set limit from the current size does not affect the computed energies by more than 0.2 kcal/mol.[101] All other geometric parameters are optimized.

The potential energy profile for the water dimer along the O–O separation is illustrated in Figure 1, and the computed binding energies from PMOw and the XP3P potential are −5.1 and −5.2 kcal/mol, respectively, slightly greater than the best estimate of −5.0 kcal/mol from *ab initio* calculations using MP2/(CBS)+Δ CCSD(T) with the 6-311++G(d,p) basis set,[77] but somewhat smaller than an estimated value (−5.4 kcal/mol) based on measured molecular vibrations.[98] For comparison, both the POL5/TZ[78] model and the AMOEBA model yield a binding energy of −5.0 kcal/mol.[10] The equilibrium structures optimized using the full PMOw Hamiltonian and the fragmental XP3P potential are listed in Table III.[99–103] The O–O distances from the PMOw and XP3P models agree well with those from POL5/TZ and AMOEBA, which yield 2.89 Å and with the *ab initio* value of 2.91 Å.[77] Ren and Ponder found that the flap angle θ (the flap angle is defined as the angle between the C_2 axis of the hydrogen bond acceptor monomer and the O–O distance vector,

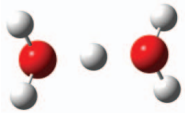
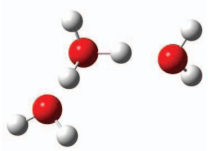
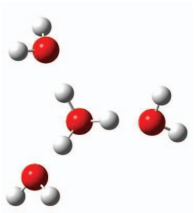
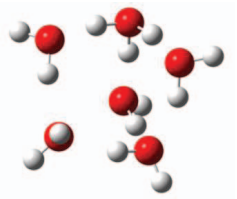
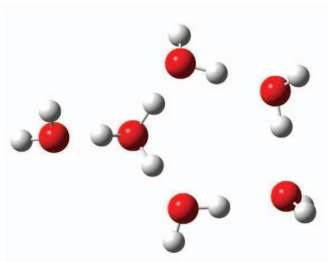
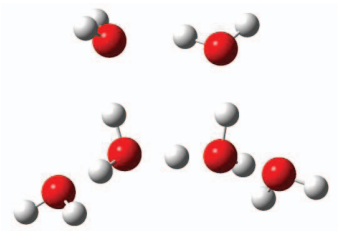
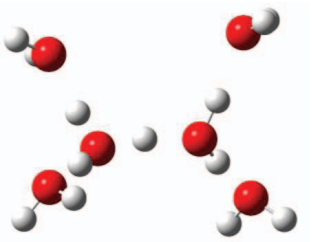
TABLE IV. Computed and experimental properties for water clusters. The angle $\langle\phi\rangle$ is the average O…H–O angle of the hydrogen bonds in a given cluster.

		PMOw	XP3P	POL5/TZ[10]	AMOEBA[10]	*Ab initio*[77,104]	Expt.
Trimer cyclic	ΔE_b	− 14.8	− 15.7	− 13.4	− 15.3	− 15.8	
	$\langle R_{OO}\rangle$	2.87	2.77	2.90	2.81	2.81	2.845
	$\langle\phi\rangle$	105.1	125.6		151.5	110.4	152
	$\langle\mu_{mol}\rangle$	2.14	2.46	2.22	2.29	2.3	
	μ	1.19	0.01	1.21	1.09	1.07	
Tetramer cyclic	ΔE_b	− 27.5	− 28.9	25.5	27.7	− 27.4	
	$\langle R_{OO}\rangle$	2.74	2.68	2.769	2.76	2.75	2.79
	$\langle\phi\rangle$	116.5	145.9		168.0	121.6	
	$\langle\mu_{mol}\rangle$	2.22	2.71	2.47	2.55	2.6	
	μ	0.0	0.0	0.0	0.0	0.0	
Pentamer cyclic	ΔE_b	− 35.7	− 39.7	34.1	36.5	− 35.9	
	$\langle R_{OO}\rangle$	2.73	2.66	2.74	2.76	2.73	2.76
	$\langle\phi\rangle$	126	159		176	132	
	$\langle\mu_{mol}\rangle$	2.26	2.82	2.57	2.64	2.7	
	μ	1.17	0.02	1.19	0.92	0.93	
Hexamer cyclic	ΔE_b	− 43.3	− 49.0	41.8	44.8	− 44.3	
	$\langle R_{OO}\rangle$	2.72	2.65	2.74	2.75	2.72	2.76
	$\langle\phi\rangle$	130	167		179	139	
	$\langle\mu_{mol}\rangle$	2.28	2.86	2.62	2.70	2.7	
	μ	0.0	0.0	0.02	0.0		
Hexamer prism	ΔE_b	− 47.8	− 44.4	41.9	45.9	− 45.3	
	$\langle R_{OO}\rangle$	2.84	2.76	2.79	2.80	2.86	
	$\langle\phi\rangle$	121.0	128.7			123.1	
	$\langle\mu_{mol}\rangle$	2.24	2.72	2.52	2.60		
	μ	2.40	3.29	2.91	2.57	2.70	
Hexamer cage	ΔE_b	− 47.8	− 45.2	41.8	45.9	− 46.0	
	$\langle R_{OO}\rangle$	2.80	2.76	2.78	2.80	2.83	2.82
	$\langle\phi\rangle$	118	126			121	
	$\langle\mu_{mol}\rangle$	2.22	2.72	2.49	2.58	2.6	
	μ	2.05	2.01	2.44	2.16	1.90	1.82–2.07[104]
Hexamer book	ΔE_b	− 46.2	− 48.3	42.5	45.8	− 45.8	
	$\langle R_{OO}\rangle$	2.75	2.70	2.79	2.78	2.78	
	$\langle\phi\rangle$	121	144			127	
	$\langle\mu_{mol}\rangle$	2.24	2.79	2.55	2.63		
	μ	2.40	2.22	2.45	2.29		
Octamer	ΔE_b	− 77.7	− 69.5			− 72.6[a]	
	$\langle R_{OO}\rangle$	2.74	2.72			2.81	
	$\langle\phi\rangle$	163	164			163	
	$\langle\mu_{mol}\rangle$	2.20	2.86				
	μ	0.0	0.0			0.0	

[a]MP2/(CBS) limit, see Ref. 101.

depicted in the inset of Figure 1) is dependent on the monomer quadrupole moment, and that it was necessary to use explicit quadrupole terms in the AMOEBA model to yield a flap angle in agreement with the *ab initio* results. The results on the flap angle in the water dimer from the PMOw and XP3P models are also good, and the small tilt angle, α, from the hydrogen bond donor is also predicted. However, the large flap angle is not preserved in the XP3P model. The structures and energies on other stationary points of water dimer are given in the supplementary material.[149]

Small water clusters (Figure S1), including the cyclic configurations of the trimer, tetramer, and pentamer, four configurations of the hexamer, and the cubic D_{2h} arrangement of the octamer have been examined (Table IV). All clusters were fully optimized with PMOw using the conjugated gradient method with NAMD.[88,90] A configuration was considered optimized when its gradient norm fell below 0.5 kcal mol^{-1} Å^{-1}. The best theoretical estimates for these systems are from the work of Bryantsev *et al.*, who performed single-point MP2/(CBS) along with a CCSD(T) correction (simply called CCSD(T) results in this discussion) at the B3LYP/6-311++G(2d,2p) optimized structures.[77] As in the work of Ren and Ponder,[10] we list in Table IV the total binding energies, the average O–O distances (R_{OO}), average O…H–O hydrogen bond angles ($\langle\phi\rangle$), and the total (μ) and average monomer ($\langle\mu_{mol}\rangle$) dipole moments. Of all water clusters, the average monomer dipole moments from the POL5/TZ and AMOEBA models[10] fall between the values computed

	
$H^+(H_2O)_2$	$H^+(H_2O)_3$
	
$H^+(H_2O)_4$	$H^+(H_2O)_6$ (I)
	
H+$(H_2O)_6$ (II)	$H^+(H_2O)_6$ (III)
	
$H^+(H_2O)_6$ (IV)	

SCHEME 2. Optimized geometries of $H^+(H_2O)_n$ clusters from PMOw.

using the PMOw and the XP3P method, and the trends are in accord with that estimated by Gregory *et al.*[104] using a portioning scheme for the electron density. Overall, the computed binding energies rom PMOw and XP3P methods are in good agreement with the CCSD(T) results, with root-mean-square (RMS) deviations of 1.2 and 2.4 kcal/mol, respectively. The performance of the AMOEBA force field is excellent, whereas the POL5/TZ model slightly underestimates the binding energies.[10,78] For the hexamers, the ordering of relative stability is cage > book > prism > cyclic from CCSD(T), and cage = prism > book > cyclic from PMOw. For comparison, the ordering from the MP2/CBS+ΔCCSD(T) calculations with 6-311++G(d,p) basis[77] and AMOEBA optimizations is prism > cage > book > cyclic.[10] In any event, the three non-cyclic structures of the water hexamer are energetically similar in binding, whereas the cyclic configuration is noticeably less stable than the other three.

We have also examined several configurations of microsolvated proton $H^+(H_2O)_n$, where n = 2, 3, 4, and 6 (Scheme 2). Depicted in Figure 2 are the potential energy profile for a proton migration between two water molecules at fixed O–O distances of the global minimum R_{min}(OO), R_{min}(OO)+ 0.2 Å, and R_{min}(OO)+ 0.4 Å from PMOw, MP2/aug-cc-pVDZ, B3LYP/aug-cc-pVTZ, and M06-2X/aug-cc-pVTZ optimizations. The equilibrium structure has an R_{min}(OO) separation of 2.46, 2.40, 2.41, and 2.39 Å, respectively, from these theoretical models. With a basis set comparable to aug-cc-pVDZ, the MP2 results on these proton clusters are very close to CCSD(T)-F12 results with jun-cc-pVTZ basis.[105] The PMOw O–O distance is about 0.05 Å longer than the MP2 result, while DFT values are in

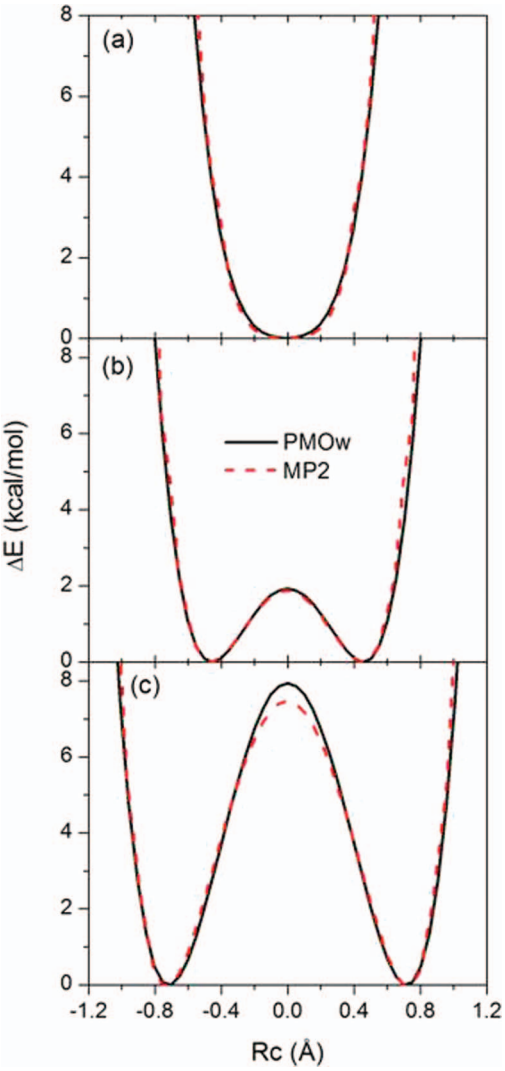


FIG. 2. Potential energy profile for $H_5O_2^+$ in the gas phase as a function of the proton transfer coordinate , defined as the distance from the mid-point between the two oxygen atoms, (a) at the minimum geometry , (b) at a fixed O–O separation of Å, and (c) at a fixed O–O distance of Å from PMOw (black), and CCSD(T)/aug-cc-pVDZ (red) calculations. Geometries were optimized with fixed O–O distances.

close agreement with MP2. In all cases, the proton is essentially symmetrically located between the two water molecules (Figure 2(a)). A small barrier appears when the O–O distance is stretched by 0.2 Å. The PMOw model yields a barrier of 1.9 kcal/mol, compared to 1.9, 1.4, and 1.3 kcal/mol from MP2, B3LYP, and M06-2X. Further stretching the O–O distance to $R_{\min}$(OO)+ 0.4 Å increases the barrier heights to 7.9, 7.5, 6.7, and 6.9 kcal/mol, respectively. There are numerous studies of proton-water clusters and proton transfer barriers with a variety of computational methods;[105–108] a thorough comparison with earlier studies is beyond the scope of the present work.

The binding energies between additional water molecules and $H_5O_2^+$ are listed in Table V, along with the MP2/aug-cc-pVDZ results. Overall, the agreement is good, with a mean-signed deviation of 1.6 kcal/mol. Note that unconstrained optimization of the structure $H^+(H_2O)_6$ (IV) using PMOw collapses to isomer (III). Thus, the value in Table V was obtained by fixing the relative torsion angles of the hydrogen atoms of the central $H_5O_2^+$ unit to the MP2 values. Overall, the results from the PMOw model are in good accord with MP2 calculations and other theoretical models.

TABLE V. Computed interaction energies for $H^+(H_2O)_n$ complexes from the PMOw and MP2 methods.[a]

Complex	PMOw	MP2
$H_5O_2^+ \ldots H_2O$	−21.4	−23.8
$H_5O_2^+ \ldots (H_2O)_2$	−39.8	−43.8
$H_5O_2^+ \ldots (H_2O)_4$ (Isomer I)	−68.9	−71.8
$H_5O_2^+ \ldots (H_2O)_4$ (Isomer II)	−67.3	−71.8
$H_5O_2^+ \ldots (H_2O)_4$ (Isomer III)	−66.6	−71.0
$H_5O_2^+ \ldots (H_2O)_4$ (Isomer IV)	−60.5	−69.7

[a]Interaction energies are calculated by $\Delta E = E(\text{cluster}) - \{E(H_5O_2^+) + nE(H_2O)\}$, where n is the number of water molecules.

IV.B. Liquid properties

IV.B.1. Properties at 25 °C

The computed and experimental thermodynamic and dynamic properties of liquid water at 25 °C and 1 atm are listed in Table VI, along with the results from TIP3P,[9,79,109] AMOEBA,[10] and SWM4-NDP.[26] The standard errors ($\pm 1\sigma$) were obtained from fluctuations of separate averages over blocks of (2–4) × 10^5 configurations. A correction, by integrating the Lennard-Jones potential beyond the cutoff distance, for the Lennard-Jones potential neglected by the cutoff has been included, and this contributes to the total computed heat of vaporization by about 1%. Long-range electrostatic interactions were not corrected in the Monte Carlo simulations. Previous studies using empirical force fields indicate that there is little size dependency of the computed properties for liquid water, and these effects will be investigated in a future study. (The TIP3P and TIP4P potential functions were developed with 125 water molecules with a cutoff of 7.5 Å without long-range corrections.[9,79,109,110])

The average density of XP3P is 0.996 ± 0.001 g/cm^3 at 25 °C, which is within 1% of the experimental value and is similar to results obtained with other polarizable and nonpolarizable force fields (Table VI).[10,26,79,109] The total energy per monomer of liquid water, $E_i(l)$, is related to the heat of vaporization by

$$\Delta H_v = -E_i(l) + P(V_{gas} - V_{liq}) + \Delta Q - (H^o - H), \tag{14}$$

where V_{gas} and V_{liq} are the molar volumes of water in the gas phase (ideal) and in the liquid, P is the pressure, ΔQ is the quantum corrections to inter and intramolecular degrees of freedom between the gas and liquid, and the last term, $(H^o - H)$, is the enthalpy departure function.[111] Although ΔQ and $(H^o - H)$ has been tabulated and can be explicitly included[9,12,110] and this amount to a total correction of −0.06 kcal/mol at 25 °C, they have typically been neglected.[10,26,79] In this case, ΔH_v is simply approximated by $-E_i(l) + RT$, which is also adopted in the present study (Table VI). The

TABLE VI. Liquid properties of the XP3P model for water along with those from experiments, and the TIP3P, AMOEBA, and SWM4-NDP models.

	XP3P	TIP3P[79,109]	AMOEBA[10]	SWM4-NDP[26]	Expt.[a]
$E(l)$ (kcal/mol)	-9.83 ± 0.01[b]	-9.82	-9.89	-9.92	-9.98
ΔH_v (kcal/mol)	10.42 ± 0.01[b]	10.41	10.48	10.51	10.51
d (g/cm^3)	0.996 ± 0.001	1.002	1.000	1.000	0.997
C_p (cal mol^{-1} K^{-1})	21.8 ± 1.0	20.0	20.9		18.0
$10^6\ \kappa$ (atm^{-1})	25 ± 2	60			46
$10^5\ \alpha$ (K^{-1})	37 ± 3	75			26
μ_{gas} (D)	1.88	2.31	1.77	1.85	1.85
μ_{liq} (D)	2.524 ± 0.002	2.31	2.78	2.33	2.3–2.6
$10^5\ D$ (cm^2/s)	2.7	5.1	2.02	2.3	2.3
ε	97 ± 8	92	82	79 ± 3	78
τ_D (ps)	8.8			11 ± 2	8.3
τ_{NMR} (ps)	2.6			1.87 ± 0.03	2.1

[a]See text for details.
[b]The average $E(l)$ from molecular dynamics simulations employing the variational Fock operator is -9.99 kcal/mol over 400 ps. This gives a heat of vaporization of 10.52 kcal/mol.

calculated heat of vaporization from the XP3P model is 10.42 ± 0.01 kcal/mol using the non-variational approximation in Monte Carlo simulations,[11,12] and the value is increased to 10.58 kcal/mol using the variational Fock operator in molecular dynamics.[14,35,88] The variational X-Pol approach used in molecular dynamics simulations lowers the interaction energy of the liquid by about 1.5% relative to the non-variational approach used in Monte Carlo. Overall, the agreement with experiment[112,113] is good, although there is greater deviation in the non-variational approach. The quality of the XP3P quantum mechanical potential for these two critical thermodynamic properties is comparable to that of the widely used SPC, TIP3P and TIP4P models for water[8,9] and to that of the recent polarizable models.[10,26,78,114]

The distribution of the magnitudes of monomer dipole moments from polarized wave functions in the liquid is shown in Figure 3; these dipole moments span a range from 2.1 to 2.9 D, and they yield an average $\langle \mu_{\text{liq}} \rangle$ of 2.524 ± 0.002 D. The width at half maximum in the dipole distribution is 0.30 D (a half-width of 0.8 D was reported for the AMOEBA model,[10] which seems to be unrealistically large). Clearly, there is a major enhancement of the molecular dipole moment in the liquid, amounting to an increase over 35% relative to the gas phase value. For comparison, the AMOEBA model produced a much greater average, 2.78 D, or 50% greater than its gas phase value. The SWM4-NDP model yielded an average of 2.46 D,[26] similar to the present XP3P quantum mechanical model. Our previous investigation, employing the AM1 Hamiltonian to represent water monomers in X-Pol, resulted in an average dipole moment of 2.29 D;[12] however, the smaller value is partly due to the much smaller molecular polarizability from AM1, and the weak polarization effect was corrected by scaling Mulliken population charges in that study. There is no experimental value for the dipole moment of liquid water (and in fact this quantity is not well defined), but values ranging from 2.3 to 2.6 D have been cited based on an estimate for ice Ih.[115,116] Finally, we note that *ab initio* molecular dynamics simulations yielded dipole moments ranging from 2.3 D to 3.8 D, depending on the method and functional used in DFT.[117] *Ab initio* molecular dynamics simulations seem to produce greater average dipole moments than polarizable force fields and the present XP3P model.

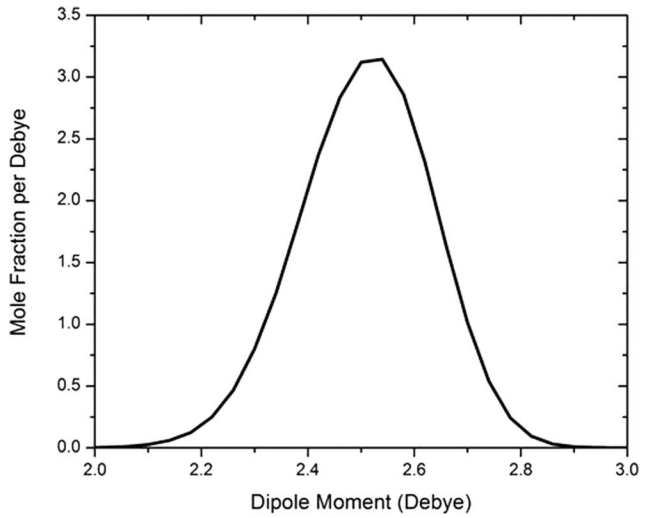


FIG. 3. Distribution of the scalar molecular dipole moment in liquid water from Monte Carlo simulations with the XP3P potential at 25 °C and 1 atm. The units for the ordinate are mole percent per debye.

The dielectric constant of the liquid is related to the fluctuations of the total dipole moment of the simulation box and it is dependent on the boundary conditions used to treat long-range electrostatics.[118,119] We employed the reaction field approximation in the NVT ensemble at 25 °C and experimental density, where intermolecular interactions are truncated at $R_{\text{cut}} = 9.0$ Å. Under these conditions, a reaction field contribution is added to the electrostatic potential in Eq. (6):[118,120]

$$V_x^{RF}(\rho_b) = V_x(\rho_b)\left\{1+\frac{2(\varepsilon_{RF}-1)}{2\varepsilon_{RF}+1}\left(\frac{|\mathbf{r}_x-\mathbf{R}_B^b|}{R_{\text{cut}}}\right)^3\right\}, \quad (15)$$

where ε_{RF} is the dielectric constant of the continuum. The static dielectric constant ε is determined from Eq. (16).[120–122]

$$\frac{(\varepsilon-1)(2\varepsilon_{RF}+1)}{2\varepsilon_{RF}+\varepsilon}=\frac{4\pi}{3k_BT}\frac{\langle \mathbf{M}^2\rangle}{\langle V\rangle}, \tag{16}$$

where **M** is the total dipole moment of the simulation box and $\langle V\rangle$ is the average volume per monomer. Ideally the reaction field dielectric ε_{RF} should be the same as that of the liquid in the cutoff sphere, although previous studies suggest that a choice of ε_{RF} in the range of $\varepsilon \le \varepsilon_{RF} \le \infty$ typically yields consistent results,[123] and a value of 160 has been used in the present study. The liquid dipole fluctuation converges slowly, and we have carried out 16 separate simulations, each lasting about 15×10^6 configurations at 25 °C. An average value of 97 ± 8 was obtained by removing the two highest and two lowest values; the present average is greater than the experimental value of 78. Interestingly, Sprik argued that an average dipole moment of 2.5–2.6 D in liquid water would lead to the correct dielectric constant at room temperature,[124] and a similar observation was used in the parameter optimization process by Lamoureux *et al.*[92] In view of the average dipole moment from the XP3P liquid, which falls in the middle of this range, it is likely that a better agreement with experiment could be obtained if the simulations were further converged by extending the simulation to 100×10^6 configurations or more in each simulation. It is interesting to note that Ren and Ponder obtained a static dielectric constant of 82, in spite of a significantly larger dipole moment of 2.78 D of the liquid from the AMEOBA potential.[10] In that work, the authors argued that the correct average H–O–H angle was responsible for the good agreement between experimental and calculated liquid dielectric constant.[10,125,126]

Displayed in Figure 4 are the distributions of the binding energies per monomer in liquid water at a temperature range of −40 °C to 100 °C. The binding energies in Figure 4 correspond to the interaction energy of one monomer with the rest of the system. In a polarizable model, the total energy of the liquid also includes the energy cost needed to polarize the electronic wave function (also called self-energy, see below). Thus, in contrast to the use of a pairwise potential, the average energy, $E_i(l)$, per monomer in Table VI is not exactly equal to half of the binding energy at 25 °C from Figure 4. but it is smaller by the amount of the self-energy. This is a reflection of the non-additive nature of a polarizable force field.[127] Note that such a self-energy term has been used to develop the SPC/E model.[128]

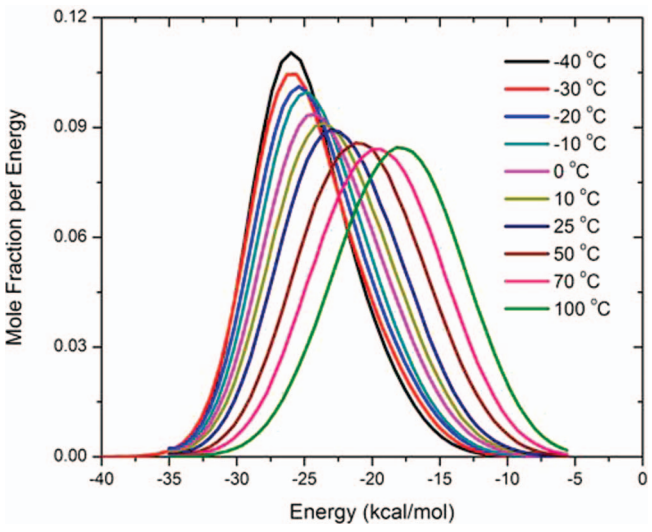


FIG. 4. Distribution of the binding energies of water in the liquid at temperatures ranging from −40 °C to 100 °C. The binding energy corresponds to the total interaction energy of one water with the rest of the bulk solvent.

We have estimated several thermodynamic properties involving molecular fluctuations. The intermolecular contribution to the isobaric heat capacity C_p of water is defined below and can also be computed from the enthalpy fluctuations by

$$C_P=\left(\frac{\partial\langle H_i(l)\rangle}{\partial T}\right)_P+3R=\frac{\langle H_i(l)^2\rangle-\langle H_i(l)\rangle^2}{RT^2}+3R, \tag{17}$$

where $H_i(l) = E_i(l) + PV_{liq}$ is the average enthalpy of the system per monomer. The total heat capacity of the liquid C_P for a rigid monomer model is determined by adding the classical kinetic energy contributions from translation and rotation of a water molecule ($3R$).[79] The average from the fluctuation formula in Eq. (17) is 22 ± 1 cal mol^{-1} K^{-1}, which is greater than the experimental value at 25 °C.[129,130] Path integral simulations by Vega *et al.* showed that inclusion of nuclear quantum effects lowers the computed heat capacity by up to 6 cal/mol.[131] Quantities based on the fluctuation formula, including C_P (isobaric heat capacity), α (coefficient of thermal expansion), and κ (isothermal compressibility) are difficult to converge; they can also be estimated from the numerical derivatives of their definitions. The derivative estimate from liquid enthalpies vs. T yields a C_p of 19 cal mol^{-1} K^{-1} at 25 °C. The coefficient of thermal expansion (α) and the isothermal compressibility (κ) are determined from fluctuations of volume and enthalpy, with a computed value of 37×10^{-5} K^{-1} for α and 25×10^{-6} atm^{-1} for κ, respectively. These quantities show relatively large deviations from experiment ($\alpha = 25.6 \times 10^{-5}$ K^{-1} and $\kappa = 45.8 \times 10^{-6}$ atm^{-1})[129] due to their convergence.

The self-diffusion coefficient of liquid water was determined using the Einstein formula[122] from molecular dynamics simulations with constant volume and temperature:

$$D=\lim_{t\to\infty}\frac{1}{6t}\langle|\mathbf{r}(t)-\mathbf{r}(0)|^2\rangle, \tag{18}$$

where $\mathbf{r}(t)$ is the position of the oxygen atom of water at time t. The diffusion coefficient was obtained as the slope from a linear fit of $\langle|\mathbf{r}(t) - \mathbf{r}(0)|^2\rangle/6$ as a function of t, and we obtained a value of 2.7×10^5 cm^2 s^{-1}, which agrees with experiment.[132] It is known that non-polarizable potentials for water, such as SPC, TIP3P, and TIP4P, tend to overestimate the self-diffusion coefficient, while most polarizable force fields, including the present XP3P model, show significant improvement.[10,26,78,114] The computed diffusion coefficient is also affected by finite size of the simulation box, and extrapolation to infinity will further increase the value of the diffusion coefficient.[133]

The rotational correlation times, τ_2^α, of water with respect to the H–H and O–H axes are obtained from least-square fits of the orientational time-correlation function to a single exponential function, $C_2^\alpha(t) = Ae^{-t/\tau_2^\alpha}$, where α specifies the

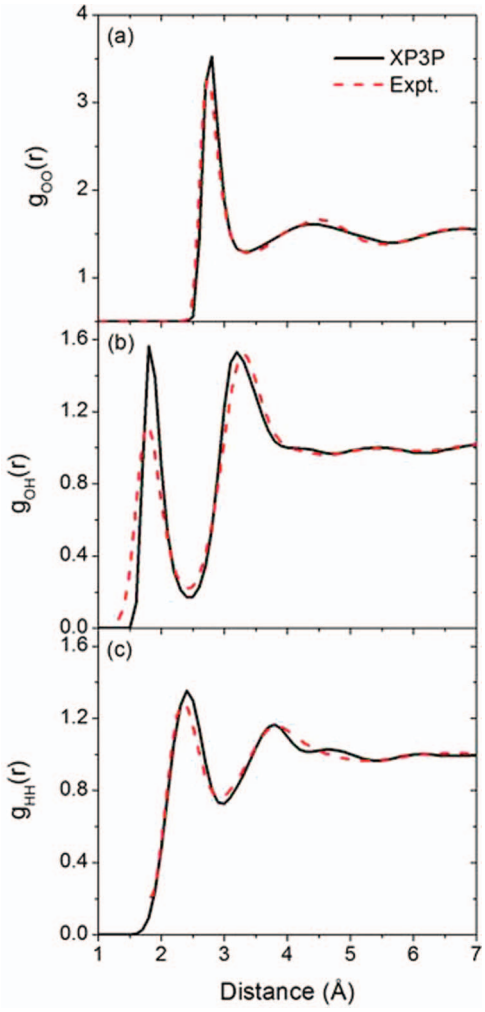


FIG. 5. Computed (black) and experimental (red, dashed) oxygen-oxygen (a), oxygen-hydrogen (b), and hydrogen-hydrogen (c) radial distribution functions of liquid water at 25 °C and 1 atm.

rotation axis. The orientation time-correlation function is defined as follows:[122]

$$C_2^{\alpha}(t) = \langle P_2[\mathbf{u}_i^{\alpha}(t)\mathbf{u}_i^{\alpha}(0)]\rangle, \tag{19}$$

where P_2 is the second-order Legendre polynomial, and $\mathbf{u}_i^{\alpha}(t)$ is the unit vector along the α rotation axis of molecule i at time t. The time-integral of Eq. (19), $A\tau_2^{\mathrm{HH}}$, corresponds to the NMR rotational relaxation time of H_2O, τ_{NMR};[134] the present XP3P model yields a value of 2.6 ps, which may be compared with the experimental value (2.1 ps).[135] For comparison, the SWM4-NDP model predicts a τ_{NMR} value of 1.9 ps.[26] Similarly, the Debye dielectric relaxation time was determined from an exponential fit to the normalized autocorrelation function of the total dipole moment $\mathbf{M}$ of the system:[122]

$$C_{\mathrm{D}}(t) = \frac{\langle \mathbf{M}(t)\mathbf{M}(0)\rangle}{\langle \mathbf{M}^2(0)\rangle}. \tag{20}$$

The Debye relaxation time characterizes the relaxation time of the hydrogen bonding network in the liquid. The XP3P model shows that the Debye relaxation time is about 6% faster than the observed values (8.3 ps).[136] In comparison with other models, the present XP3P model performs well for these dynamic properties.[10,26,114]

The structure of liquid water is characterized by radial distribution functions (RDFs), $g_{xy}(r)$, which gives the probability of finding an atom of type y at a distance r from an atom of type x relative to the bulk. The RDFs computed at 25 °C from Monte Carlo simulations are shown in Figure 5 along with the neutron diffraction data. Overall, the agreement with experimental results is excellent. For the XP3P potential, the location of the maximum of the first peak of the O–O RDF is 2.78 ± 0.05 Å with a peak height of 3.0 (Figure 5(a)). For comparison, the corresponding experimental values are 2.73 Å and 2.8 from neutron diffraction.[137,138] Integration of the O–O RDF to the first minimum at 3.30 Å yields an estimated coordination number of 4.5, which is in good agreement with the neutron diffraction result of 4.51 (integrated to 3.36 Å), but somewhat smaller than the X-ray diffraction result (4.7). The oxygen-hydrogen and hydrogen-hydrogen radial distribution functions are also in accord with experiments.

IV.B.2. Temperature-dependent liquid properties

The computed liquid properties for ΔH_v, C_P, ρ, α, and κ, at different temperatures ranging from −40 to 100 °C are listed in Table S5, and some of these are compared with experimental data in Figures 6–8. The formulas involving fluctuations of enthalpy and volume for C_P, α, and κ are known to have slow convergence even when Monte Carlo simulations were extended to over hundreds of millions of configurations. In the present simulations, C_P and α can also be determined directly from the enthalpy and volume derivatives with respect to temperature. For the isothermal compressibility, the fluctuation formula was used since the pressure was not changed in the present study.

The heats vaporization from −40 to 100 °C were obtained from the average energies plus RT for the PV term of an ideal gas; here, we have ignored the small corrections for the quantum vibrational energy difference and enthalpy departure function. For comparison, we have included the computed heats of vaporization in Figure 6 from the TIP5P model. The XP3P model agrees with the results from TIP5P

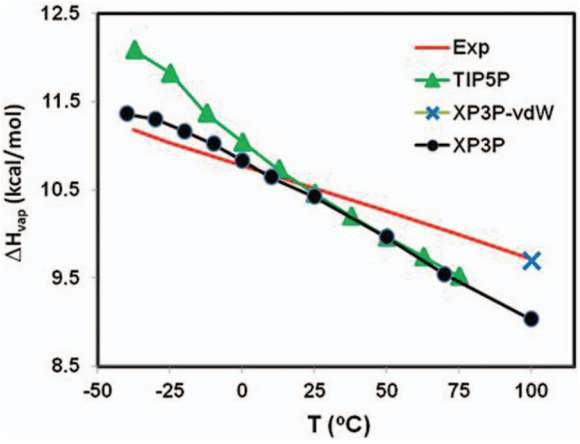


FIG. 6. Computed (black) and experimental (red) heats of vaporization for liquid water. The results from the TIP5P model are illustrated in green.

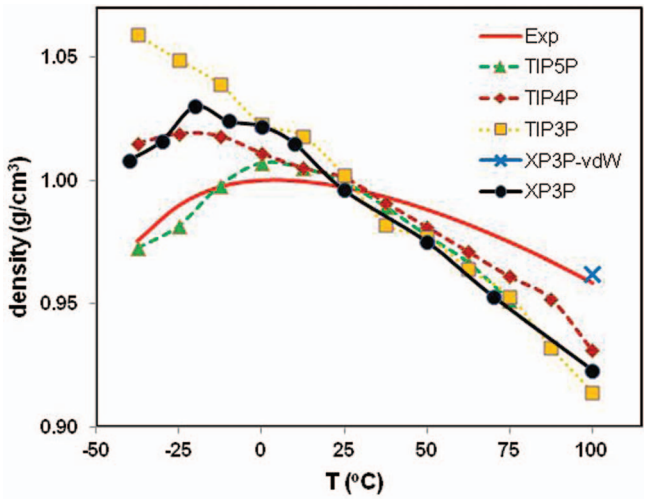


FIG. 7. Computed (black) and experimental (red) densities for liquid water, along with those from the TIP3P (brown), the TIP4P (maroon), and the TIP5P (green) models.

quantitatively at temperature above 25 °C. Both XP3P and TIP5P overestimate ΔH_v at temperature lower than 25 °C, but the TIP5P model yielded a greater deviation on supercooled water. Figure 6 shows that the change in ΔH_v is nearly linear over the entire temperature range considered. This agrees with the experimental results on heat capacity, which is nearly constant at about 20 cal mol^{-1} K^{-1}.[129] The changes of heat capacity with temperature are given in Figure S5. The trends are in reasonable agreement with experiment at temperatures above 0 °C, although the sharp increase of C_P below 20 °C is not reproduced by the present simulations.

The liquid density as a function of temperature is presented in Figure 7 along with the experimental density of liquid water. The XP3P model, which is optimized to reproduce the heat of vaporization and density at 25 °C, yields a maximum density at about −20 °C. Although the density maximum is significantly lower than the experimental value at 4 °C,[129] it is in fact remarkable in that there is a density maximum at all from the present model because other three-point-charge models do not possess this property with a reasonable temperature (except the SPC/E with much enhanced electrostatics). The computed density at temperature greater than 25 °C shows more rapid decline with increasing temperatures than experimental results.[129] This trend is similar to that found in the TIP*x*P series of models.[79] The densities for supercooled water are overestimated by 2%–5% compared with the experimental data.[129] For comparison, among the non-polarizable models that do possess a density maximum, SPC/E[128] has a density maximum at −38 °C,[139] TIP4P at −15 °C,[109] and TIP5P at about 0 °C; the TIP5P model was optimized to reproduce the temperature dependence of liquid density of water.[79] The AMOEBA model has a density maximum at 17 °C.[126]

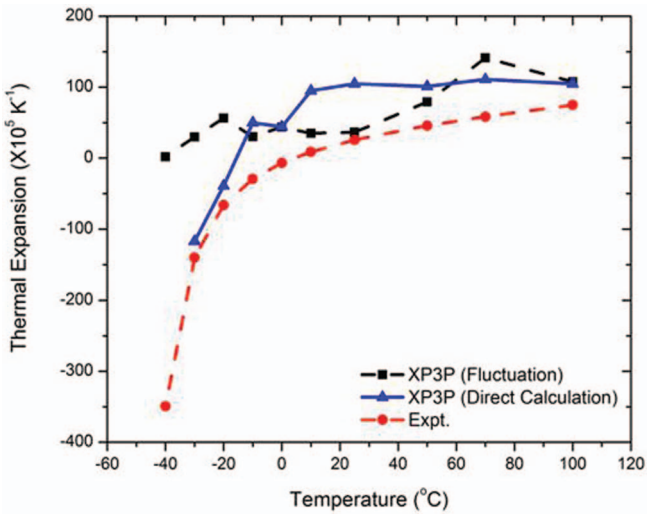


FIG. 8. Computed and experimental coefficients of thermal expansion (α) for liquid water. The α values are determined from numerical derivatives of liquid volume variations with temperature.

The temperature dependences of the computed density and ΔH_v from the non-polarizable TIP*x*P series of models[79,109] and the polarizable AMEOBA potential[126] indicate that it is difficult, with fixed empirical parameters, to obtain good agreement (within 1%) with experiment for the entire temperature range from the supercooled liquid to the boiling point. This difficulty has been pointed out by Siepmann and co-workers, who used a charge-dependent van der Waals radius for oxygen in a fluctuating charge model for water.[140] Giese and York[32] developed a density-dependent van der Waals potential that can be directly incorporated into QM/MM style simulations. We have further optimized σ_O at 100 °C to yield a better agreement with the experimental liquid density ρ. We found that a small change in σ_O from 3.225 to 3.205 Å is sufficient to produce a liquid density (0.962 g/cm^3) in good agreement with experiment (0.958 g/cm^3). This is shown by the blue cross point in Figure 7. Interestingly, the computed ΔH_v (9.70 kcal/mol) was also found to be in excellent agreement with experiment (9.72 kcal/mol)[112] after this small adjustment (blue cross point in Figure 6). With this change, the average dipole moment is computed to be 2.470 ± 0.001 D, representing an increase of 0.042 D from 2.428 D computed with the original Lennard-Jones parameters in Table II.

In view of the small change in the σ_O value, we suggest a simple temperature-dependent relationship for σ_O,

$$\sigma_O(T) = 3.225 - 2.667 \times 10^{-4}(T - 298.15), \qquad (21)$$

in Å^3 where T is the absolute temperature. Alternatively, Eq. (21) may be rewritten in terms molecular dipole moment, which translates the expression to an aesthetically appealing, density-dependent one. In any event, it is straightforward to use Eq. (21) in Monte Carlo simulations, while it can be conveniently incorporated into a thermostat algorithm in molecular dynamics simulations.[82,141,142] However, a thorough examination of the performance of temperature-dependent van der Waals parameters is beyond the scope of the present work.

The computed coefficient of thermal expansion, α, follows the experimental trends nicely in Figure 8, and the negative values for supercooled water are consistent with the experimental values as a result of the existence of a density maximum vs. temperature.

The average dipole moment from the XP3P model decreases monotonically with increasing temperature (inset of

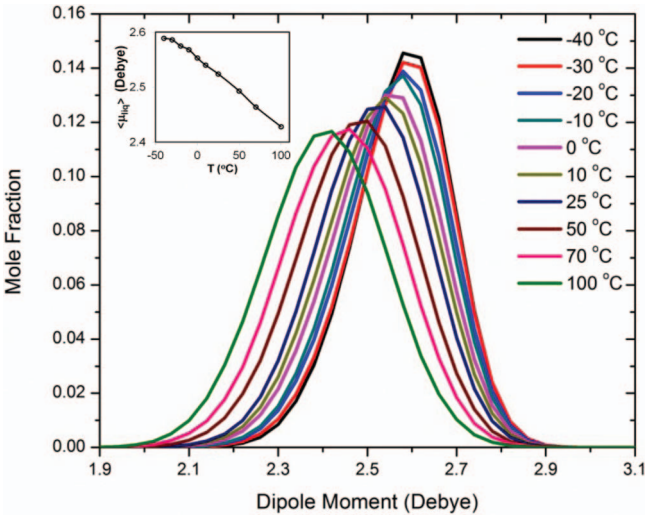


FIG. 9. Computed average molecular dipole moments for liquid water at different temperatures.

Figure 9). The distributions of scalar dipole moment in the liquid at different temperatures are given in Figure 9. Consistent with Figure 4, the maximum positions are shifted towards smaller values as temperature increases, and this shift is accompanied by an increase in half width from about 0.26 D to about 0.32 D. The broader distribution of molecular dipole moment in liquid water at higher temperature reflects greater variations in the local hydrogen bonding networks and reduced average binding energies (Figure 4) and heats of vaporization (Figure 6). It is interesting to notice that the maximum dipole values in the distributions are not shifted at different temperatures (Figure 9); it is the population of the molecular dipole moment in the liquid that is broadened. This results in a shift of the maximum position towards smaller average values as the temperature increases. In a recent study, Raabe and Sadus suggested that the introduction of bond and angle flexibility in a water model is responsible for the decrease in the dipole moment with increased temperature and for the good performance on computed dielectric constant and pressure-temperature-density behavior using a flexible water model.[143] However, the water geometry was severely distorted from the gas-phase structure and the average bond lengths and angles in the liquid states are both significantly larger than commonly accepted values of liquid water.[143,144] The results displayed in Figure 9 show that the change in electronic polarization at different thermodynamic state points also makes critical contributions to the variation of the molecular dipole moment.

Computed radial distribution functions, which exhibit the expected trends as functions of temperature, are given in the supplementary material.[149] The loss of the liquid structure is observed with increasing temperature, and the height of the first peak in $g_{OO}(r)$ declines with broadening of the peak as the first minimum disappears at high temperature (Figure S7). On the contrary, $g_{OO}(r)$ at low temperatures exhibits more structured RDFs. Similar trends are observed in both $g_{OH}(r)$ and $g_{HH}(r)$ as functions of temperature (Figures S8 and S9).

IV.B.3. Energy decomposition analysis of liquid water

The total binding energy, $E_i(l)$, from the XP3P water can be decomposed into specific contributing factors,[12,85,127] including vertical interaction energy and polarization energy. This analysis is useful for understanding the energy terms that are implicitly fitted in the development of polarizable or non-polarizable empirical potentials.

The vertical interaction energy represents the total energy of the liquid in which the wave function of each water molecule is not polarized, corresponding to that in the gas phase,

$$\Delta E_{\text{vert}} = \frac{1}{2}\sum_{a=1}^{N}\sum_{b\neq a}^{N}\left\langle \Psi_a^o \middle| H_{ab}^o(\rho_b^o) \middle| \Psi_a^o \right\rangle + E^{\text{XD}}, \tag{22}$$

where $H_{ab}^o(\rho_b^o)$ is the interaction Hamiltonian between molecules a and b, in which the electrostatic potential defined in Eqs. (5) and (6) is obtained using the density of molecule b in the gas phase, ρ_b^o, and $E^{\text{XD}} = \sum_{a>b} E_{ab}^{\text{XD}}$ is the total van der Waals (i.e., the exchange-correlation term approximated by the Lennard-Jones potential in Eq. (9)).

We emphasize that the term "vertical interaction energy" in energy decomposition analysis (EDA) is used to describe the interaction energy of the solvent molecules with their gas-phase, non-polarized electronic wave function relative to that of non-interacting molecules (Eq. (22)).[85,127,145] This differs from the meaning of "vertical" that is associated with processes such as ionization and electronic excitation, where the geometries of the solute and the surrounding solvent are hypothetically kept in the un-ionized or the ground-state equilibrium configuration. In both cases the electronic wave function of the solute does change. In condensed-phase simulations, however, the energy accompanying the change of the electronic wave function is called "polarization energy". Therefore, the term vertical is used to specify the interaction energy from an electronic state that is kept to remain in its gas-phase (electronic) configuration, prior to polarization.

The wave functions of the solvent molecules are polarized in the liquid, and the energy change induced by the mutual interactions with the rest of the system corresponds to the polarization interaction energy, which is defined by Eq. (23).[12,85,127]

$$\Delta E_{\text{pol}} = \left(\langle\Phi|H|\Phi\rangle - NE_a^o\right) - \Delta E_{\text{vert}} = E_{\text{tot}} - \Delta E_{\text{vert}}. \tag{23}$$

The polarization energy can be further separated into two physically significant terms, corresponding to the so-called self-energy, ΔE_{self}, which is an energy cost (also called energy penalty) needed to pay for distorting the molecular wave function, and a net stabilizing contribution, ΔE_{stab}, which is responsible for polarizing the electronic wave function to lower the total energy of the system. These energy terms are given below,[12,38,85,127]

$$\Delta E_{\text{self}} = \sum_{a=1}^{N}\left[\langle\Psi_a|H_a^o|\Psi_a\rangle - \left\langle\Psi_a^o\middle|H_a^o\middle|\Psi_a^o\right\rangle\right] = \sum_{a=1}^{N}\Delta E_a, \tag{24}$$

TABLE VII. Temperature-dependent energy components (unit in kcal/mol).

T (°C)	$E_i(l)$	E_{vert}	E_{pol}	ΔE_{stab}	ΔE_{self}	E^{XD}	E_{ele}
−40	−10.89	−6.62	−4.27	−7.98	3.71	3.17	−14.06
−30	−10.81	−6.57	−4.24	−7.92	3.68	3.12	−13.93
−20	−10.66	−6.52	−4.14	−7.71	3.57	2.97	−13.63
−10	−10.50	−6.44	−4.06	−7.56	3.50	2.90	−13.40
0	−10.29	−6.35	−3.94	−7.30	3.36	2.74	−13.03
10	−10.08	−6.26	−3.82	−7.07	3.25	2.62	−12.70
25	−9.83	−6.17	−3.66	−6.76	3.10	2.49	−12.32
50	−9.32	−5.90	−3.42	−6.26	2.84	2.20	−11.52
70	−8.86	−5.69	−3.17	−5.78	2.61	1.97	−10.83
100	−8.28	−5.38	−2.90	−5.23	2.33	1.69	−9.97

$$\Delta E_{stab} = \frac{1}{2}\sum_{a=1}^{N}\sum_{b\neq a}^{N}\left[\langle\Psi_a|\hat{H}_{ab}(\rho_b)|\Psi_a\rangle - \left\langle\Psi_a^o\middle|\hat{H}_{ab}^o\left(\rho_b^o\right)\middle|\Psi_a^o\right\rangle\right]$$

$$= \frac{1}{2}\sum_{a=1}^{N}\sum_{b\neq a}^{N}\Delta\Delta E_{ab}. \quad (25)$$

Shown in Table VII and Figure 10 are the XP3P energy components at different temperatures. The vertical interaction energy contributes an almost constant percentage of the total binding energy, ranging from 60.8% at −40 °C to 65.0% at 100 °C. The increase of the percentage with increasing temperature can be attributed to the increased volume of the system and reduced polarization effects at higher temperatures. At all temperatures used in the simulations, polarization effects are significant, contributing 35.0%–39.2% of the total binding energies. At 25 °C, the average polarization energy is −3.66 kcal/mol (37.2% of $E_i(l)$). The van der Waals (or exchange-dispersion) term E^{XD} is dominated by the repulsive potential. The total electrostatic (non-van der Waals) component of the binding energy, $E_i(l)$, is the sum of the vertical and polarization interaction energies less the E^{XD} term, and it is about 20%–30% greater than the total binding energy in the 140 °C temperature range.

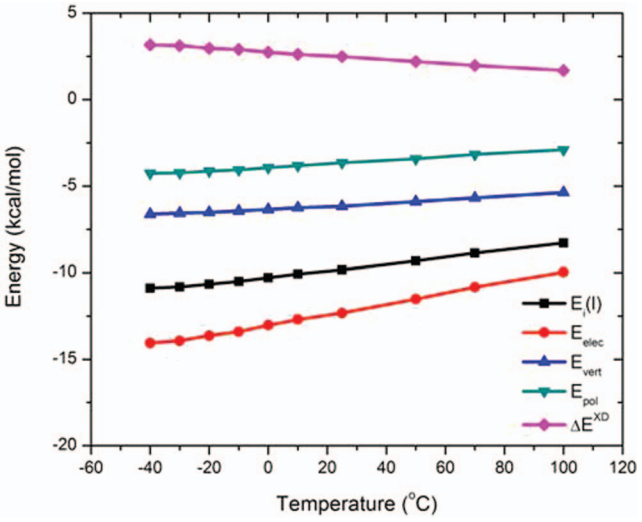


FIG. 10. Average total interaction energies (black) per water in the liquid and their contributing components, including vertical interaction energies (blue), polarization energies (green), total electrostatic interaction energies (red), and exchange-dispersion correlation energies (magenta).

Table VII shows that the average energy cost, i.e., self-energy (Eq. (24)), needed to polarize the molecular wave function, is 3.10 ± 0.01 kcal/mol from the XP3P mode at 25 °C. This value is somewhat greater than the value estimated using the AM1 Hamiltonian (3.03 ± 0.01 kcal/mol).[12] If the classical expression for the self-energy,[128]

$$\Delta E_{self}^{cl} = \Delta\mu_{ind}^2/2\alpha \quad (26)$$

is used, where $\Delta\mu_{ind}$ is the induced dipole moment in the liquid, which is 0.64 D at 25 °C, and α is the molecular polarizability (1.27 Å^3) from the XP3P model, we obtain a self-energy of 2.35 kcal/mol, somewhat smaller than the quantum mechanical result (Eq. (23)). The self-energy was used to correct the total energy of liquid water in the SPC/E model,[128] which has an effective dipole of 2.35 D ($\Delta\mu_{ind} = 0.50$ D). In that work, an estimate of $\Delta E_{self}^{cl} = 1.25$ kcal/mol was used as an energy correction based on experimental polarizability of water. Table VII shows that over the temperature range of −40 to 100 °C, ΔE_{self} varies from 3.69 kcal/mol to 2.33 kcal/mol, and the corresponding total polarization energies change from −4.25 to −2.90 kcal/mol.

V. CONCLUSIONS

A quantum mechanical force field (QMFF) for water with the explicit treatment of electronic polarization (X-Pol) has been described. Moving beyond the current Lifson-type, molecular mechanics force fields (MMFF) that have been under continuous development in the past half century,[146–148] the present QMFF represents the condensed-phase system explicitly by an electronic structure method. Consequently, the internal energy terms in the traditional MMFF are replaced by a quantum mechanical formalism that naturally includes electronic polarization. An important aspect of the present procedure is the partition of a solution into molecular fragments such that the total wave function of the system is approximated as a Hartree product of antisymmetric, fragment wave functions. This approximation requires an empirical treatment of short-range intermolecular exchange repulsion and long-range dispersion interactions between different molecular fragments; however, one can model these effects using customary empirical formalisms. To this end, we have introduced a polarizable molecular orbital (PMO) model in the framework of the neglect diatomic differential overlap approximation. The present study represents a first step towards the goal of developing a full QMFF for the dynamic simulations of macromolecular systems as traditionally carried out with MMFF.

In this work, we introduce the first generation of a QMFF for water, making use of the PMO model specifically parameterized for compounds composed of hydrogen and oxygen, i.e., PMOw. The electrostatic potential responsible for the interactions among different fragments is model by a three-point charge representation that reproduces the total molecular dipole moment and the local hybridization contributions exactly. Consequently, the present QMFF for water, suitable for modeling gas-phase clusters, pure liquids, solid

isomorphs, aqueous solutions, and the self-dissociation along with proton and anion transport, is called the XP3P model. The paper highlights the performance of the PMOw model for small water and proton clusters and simple proton transfer reactions, and the properties of liquid water using XP3P from a conglomeration of about 900×10^6 self-consistent-field calculations on a periodic system consisting of 267 water molecules. It is no exaggeration to say that this is the longest quantum mechanical simulation performed to date. More significantly, the unusual dipole derivative behavior of water, which is incorrectly modeled in molecular mechanics, but is critical for a flexible water model, is naturally reproduced as a result of an electronic structural treatment of chemical bonding by XP3P. Much remains to be tested and investigated in future studies with the combined use of large clusters treated by PMOw embedded the XP3P liquid water. We anticipate that the present model is useful for studying proton transport in solution and solid phases as well as across biological membranes through ion channels.

ACKNOWLEDGMENTS

This work has been partially supported by the National Institutes of Health, Grant Nos. GM46376 and RC1-GM091445, and by the National Science Foundation Grant Nos. CHE09-56776 and CHE09-57162. Part of the computations has been performed on an SGI Altix cluster acquired through an NIH grant (Grant No. S10-RR029467). We thank an anonymous referee for several insightful comments.

[1] A. D. MacKerell, Jr., J. Comput. Chem. **25**, 1584 (2004).
[2] J. W. Ponder and D. A. Case, Adv. Protein Chem. **66**, 27 (2003).
[3] B. Guillot, J. Mol. Liq. **101**, 219 (2002).
[4] B. Schropp and P. Tavan, J. Phys. Chem. B **112**, 6233 (2008).
[5] C. Vega, J. L. F. Abascal, and P. G. Debenedetti, Phys. Chem. Chem. Phys. **13**, 19660 (2011).
[6] J. D. Bernal and R. H. Fowler, J. Chem. Phys. **1**, 515 (1933).
[7] W. L. Jorgensen, J. Chem. Theory Comput. **3**, 1877 (2007).
[8] H. J. C. Berendsen, J. P. M. Postama, W. F. van Gunsteren, and J. Hermans, in *Intermolecular Forces*, edited by B. Pullmann (D. Reidel Publishing Company, Dordrecht, 1981), p. 331.
[9] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[10] P. Ren and J. W. Ponder, J. Phys. Chem. B **107**, 5933 (2003).
[11] J. Gao, J. Phys. Chem. B **101**, 657 (1997).
[12] J. Gao, J. Chem. Phys. **109**, 2346 (1998).
[13] W. Xie, M. Orozco, D. G. Truhlar, and J. Gao, J. Chem. Theory Comput. **5**, 459 (2009).
[14] W. Xie, L. Song, D. G. Truhlar, and J. Gao, J. Chem. Phys. **128**, 234108 (2008).
[15] W. Xie and J. Gao, J. Chem. Theory Comput. **3**, 1890 (2007).
[16] S. J. Wierzchowski, D. A. Kofke, and J. Gao, J. Chem. Phys. **119**, 7365 (2003).
[17] F. J. Vesely, J. Comput. Phys. **24**, 361 (1977).
[18] A. E. Howard, U. C. Singh, M. Billeter, and P. A. Kollman, J. Am. Chem. Soc. **110**, 6984 (1988).
[19] D. N. Bernardo, Y. Ding, K. Krogh-Jespersen, and R. M. Levy, J. Phys. Chem. **98**, 4180 (1994).
[20] J. Gao, D. Habibollazadeh, and L. Shao, J. Phys. Chem. **99**, 16460 (1995).
[21] J. M. Stout and C. E. Dykstra, J. Phys. Chem. A **102**, 1576 (1998).
[22] J. Applequist, J. R. Carl, and K.-K. Fung, J. Am. Chem. Soc. **94**, 2952 (1972).
[23] B. T. Thole, Chem. Phys. **59**, 341 (1981).
[24] P. T. van Duijnen and M. Swart, J. Phys. Chem. A **102**, 2399 (1998).
[25] H. B. Yu, T. Hansson, and W. F. van Gunsteren, J. Chem. Phys. **118**, 221 (2003).
[26] G. Lamoureux, E. Harder, I. V. Vorobyov, B. Roux, and A. D. MacKerell, Chem. Phys. Lett. **418**, 245 (2006).
[27] A. K. Rappe and W. A. Goddard, J. Phys. Chem. **95**, 3358 (1991).
[28] S. W. Rick, S. J. Stuart, and B. J. Berne, J. Chem. Phys. **101**, 6141 (1994).
[29] G. A. Kaminski, H. A. Stern, B. J. Berne, and R. A. Friesner, J. Phys. Chem. A **108**, 621 (2004).
[30] S. M. Valone, J. Chem. Theory Comput. **7**, 2253 (2011).
[31] W. Kohn, A. D. Becke, and R. G. Parr, J. Phys. Chem. **100**, 12974 (1996).
[32] T. J. Giese and D. M. York, J. Chem. Phys. **127**, 194101 (2007).
[33] A. Cembran, P. Bao, Y. Wang, L. Song, D. G. Truhlar, and J. Gao, J. Chem. Theory Comput. **6**, 2469 (2010).
[34] J. Gao, A. Cembran, and Y. Mo, J. Chem. Theory Comput. **6**, 2402 (2010).
[35] L. Song, J. Han, Y. L. Lin, W. Xie, and J. Gao, J. Phys. Chem. A **113**, 11656 (2009).
[36] J. B. Han, D. G. Truhlar, and J. L. Gao, Theor. Chem. Acc. **131**, 1161 (2012).
[37] Y. J. Wang, C. P. Sosa, A. Cembran, D. G. Truhlar, and J. L. Gao, J. Phys. Chem. B **116**, 6781 (2012).
[38] Y. R. Mo, P. Bao, and J. L. Gao, Phys. Chem. Chem. Phys. **13**, 6760 (2011).
[39] J. A. Pople, D. P. Santry, and G. A. Segal, J. Chem. Phys. **43**, S129 (1965).
[40] M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart, J. Am. Chem. Soc. **107**, 3902 (1985).
[41] J. J. P. Stewart, J. Comput.-Aided Mol. Des. **4**, 1 (1990).
[42] G. B. Rocha, R. O. Freire, A. M. Simas, and J. J. P. Stewart, J. Comput. Chem. **27**, 1101 (2006).
[43] J. J. P. Stewart, J. Mol. Model **13**, 1173 (2007).
[44] C. A. Morgado, J. P. McNamara, I. H. Hillier, and N. A. Burton, J. Chem. Theory Comput. **3**, 1656 (2007).
[45] J. P. McNamara and I. H. Hillier, Phys. Chem. Chem. Phys. **9**, 2362 (2007).
[46] J. P. McNamara, R. Sharma, M. A. Vincent, I. H. Hillier, and C. A. Morgado, Phys. Chem. Chem. Phys. **10**, 128 (2008).
[47] T. Tuttle and W. Thiel, Phys. Chem. Chem. Phys. **10**, 2159 (2008).
[48] M. Korth and W. Thiel, J. Chem. Theory Comput. **7**, 2929 (2011).
[49] P. Zhang, L. Fiedler, H. R. Leverentz, D. G. Truhlar, and J. L. Gao, J. Chem. Theory Comput. **7**, 857 (2011).
[50] L. Fiedler, J. L. Gao, and D. G. Truhlar, J. Chem. Theory Comput. **7**, 852 (2011).
[51] M. Isegawa, L. Fiedler, H. R. Leverentz, Y. J. Wang, S. Nachimuthu, J. L. Gao, and D. G. Truhlar, J. Chem. Theory Comput. **9**, 33 (2013).
[52] K. Jug and G. Geudtner, J. Comput. Chem. **14**, 639 (1993).
[53] M. J. S. Dewar and W. Thiel, J. Am. Chem. Soc. **99**, 4899 (1977).
[54] K. T. Tang and J. P. Toennies, J. Chem. Phys. **80**, 3726 (1984).
[55] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, J. Chem. Phys. **132**, 154104 (2010).
[56] J. J. P. Stewart, Rev. Comput. Chem. **1**, 45 (1990).
[57] M. C. Zerner, Rev. Comput. Chem. **2**, 313 (1991).
[58] M. J. S. Dewar and W. Thiel, J. Am. Chem. Soc. **99**, 4907 (1977).
[59] M. J. S. Dewar and W. Thiel, Theoret. Chim. Acta **46**, 89 (1977).
[60] J. Gao and Y. Wang, J. Chem. Phys. **136**, 071101 (2012).
[61] T. J. Giese, H. Y. Chen, T. Dissanayake, G. M. Giambasu, H. Heldenbrand, M. Huang, E. R. Kuechler, T. S. Lee, M. T. Panteva, B. K. Radak, and D. M. York, J. Chem. Theory Comput. **9**, 1417 (2013).
[62] T. Nakano, T. Kaminuma, T. Sato, K. Fukuzawa, Y. Akiyama, M. Uebayasi, and K. Kitaura, Chem. Phys. Lett. **351**, 475 (2002).
[63] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, J. Am. Chem. Soc. **117**, 5179 (1995).
[64] P. Cieplak, J. Caldwell, and P. Kollman, J. Comput. Chem. **22**, 1048 (2001).
[65] J. Wang, P. Cieplak, and P. A. Kollman, J. Comput. Chem. **21**, 1049 (2000).
[66] A. J. Stone, *The Theory of Intermolecular Forces* (Oxford University Press, Oxford, 1996).
[67] M. S. Gordon, L. Slipchenko, H. Li, and J. H. Jensen, Annu. Rep. Comp. Chem. **3**, 177 (2007).
[68] H. R. Leverentz, J. L. Gao, and D. G. Truhlar, Theor. Chem. Acc. **129**, 3 (2011).
[69] W. J. Hehre, L. Radom, P. v. R. Schleyer, and J. A. Pople, *Ab Initio Molecular Orbital Theory* (John Wiley & Sons, New York, 1986).
[70] J. Li, T. Zhu, C. J. Cramer, and D. G. Truhlar, J. Phys. Chem. A **102**, 1820 (1998).

[71] A. V. Marenich, S. V. Jerome, C. J. Cramer, and D. G. Truhlar, J. Chem. Theory Comput. **8**, 527 (2012).
[72] B. T. Thole and P. T. van Duijnen, Theor. Chim. Acta **63**, 209 (1983).
[73] M. Swart, P. T. Van Duijnen, and J. G. Snijders, J. Comput. Chem. **22**, 79 (2001).
[74] P. Zhang, P. Bao, and J. L. Gao, J. Comput. Chem. **32**, 2127 (2011).
[75] J. A. Pople and G. A. Segal, J. Chem. Phys. **43**, S136 (1965).
[76] S. A. Clough, Y. Beers, G. P. Klein, and L. S. Rothman, J. Chem. Phys. **59**, 2254 (1973).
[77] V. S. Bryantsev, M. S. Diallo, A. C. T. van Duin, and W. A. I. Goddard, J. Chem. Theory Comput. **5**, 1016 (2009).
[78] H. A. Stern, F. Rittner, B. J. Berne, and R. A. Friesner, J. Chem. Phys. **115**, 2237 (2001).
[79] M. W. Mohoney and W. L. Jorgensen, J. Chem. Phys. **112**, 8910 (2000).
[80] P. Zhang, D. G. Truhlar, and J. Gao, Phys. Chem. Chem. Phys. **14**, 7821 (2012).
[81] H. C. Andersen, J. Chem. Phys. **72**, 2384 (1980).
[82] E. A. Koopman and C. P. Lowe, J. Chem. Phys. **124**, 204103 (2006).
[83] S. Miyamoto and P. A. Kollman, J. Comput. Chem. **13**, 952 (1992).
[84] K. Nam, J. Gao, and D. M. York, J. Chem. Theory Comput. **1**, 2 (2005).
[85] J. Gao and X. Xia, Science **258**, 631 (1992).
[86] K. Kitaura, E. Ikeo, T. Asada, T. Nakano, and M. Uebayasi, Chem. Phys. Lett. **313**, 701 (1999).
[87] J. Gao, J. Han, and P. Zhang, MCSOL, version 2012xp, Minneapolis, 2012.
[88] M. Mazack and J. Gao, X-Pol, version 2013a1, University of Minnesota, 2013.
[89] B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus, J. Comput. Chem. **30**, 1545 (2009).
[90] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten, J. Comput. Chem. **26**, 1781 (2005).
[91] M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN 09, Rev. A.02, Gaussian, Inc., Wallingford, CT, 2009).
[92] G. Lamoureux, A. D. MacKerell, Jr., and B. Roux, J. Chem. Phys. **119**, 5185 (2003).
[93] W. S. Benedict, N. Gailar, and E. K. Plyler, J. Chem. Phys. **24**, 1139 (1956).
[94] C. J. Burnham and S. S. Xantheas, J. Chem. Phys. **116**, 5115 (2002).
[95] H. Partridge and D. W. Schwenke, J. Chem. Phys. **106**, 4618 (1997).
[96] E. Whalley and D. D. Klug, J. Chem. Phys. **84**, 78 (1986).
[97] L. S. Rothman, C. P. Rinsland, A. Goldman, S. T. Massie, D. P. Edwards, J. M. Flaud, A. Perrin, C. Camy-Peyret, V. Dana, J. Y. Mandin, J. Schroeder, A. McCann, R. R. Gamache, R. B. Wattson, K. Yoshino, K. V. Chance, K. W. Jucks, L. R. Brown, V. Nemtchinov, and P. Varanasi, J. Quant. Spectrosc. Radiat. Transf. **111**, 1568 (2010).
[98] L. A. Curtiss, D. J. Frurip, and M. Blander, J. Chem. Phys. **71**, 2703 (1979).
[99] K. Liu, M. G. Brown, and R. J. Saykally, J. Phys. Chem. A **101**, 8995 (1997).
[100] M. Piris, J. M. Matxain, X. Lopez, and J. M. Ugalde, J. Chem. Phys. **132**, 031103 (2010).
[101] S. S. Xantheas, Struct. Bond **116**, 119 (2006).
[102] G. Maroulis, J. Chem. Phys. **113**, 1813 (2000).
[103] G. S. Tschumper, M. L. Leininger, B. C. Hoffman, E. F. Valeev, H. F. Schaefer, and M. Quack, J. Chem. Phys. **116**, 690 (2002).
[104] J. K. Gregory, D. C. Clary, K. Liu, M. G. Brown, and R. J. Saykally, Science **275**, 814 (1997).
[105] S. Nachimuthu, J. L. Gao, and D. G. Truhlar, Chem. Phys. **400**, 8 (2012).
[106] S. Sadhukhan, D. Munoz, C. Adamo, and G. E. Scuseria, Chem. Phys. Lett. **306**, 83 (1999).
[107] R. Kumar, R. A. Christie, and K. D. Jordan, J. Phys. Chem. B **113**, 4111 (2009).
[108] P. Goyal, M. Elstner, and Q. Cui, J. Phys. Chem. B **115**, 6790 (2011).
[109] W. L. Jorgensen and C. Jenson, J. Comput. Chem. **19**, 1179 (1998).
[110] W. L. Jorgensen and J. D. Madura, Mol. Phys. **56**, 1381 (1985).
[111] B. G. Kyle, *Chemical and Process Thermodynamics* (Prentice Hall PTR, 1999).
[112] W. Wagner and A. Pruss, J. Phys. Chem. Ref. Data **31**, 387 (2002).
[113] L. Haar, E. Gallagher, and G. Kell, *NBS/NRC Steam Tables: Thermodynamic and Transport Properties and Computer Programs for Vapor and Liquid States of Water in SI Units* (Hemisphere Publishing Corporation, Washington, 1984).
[114] J. Wang, P. Cieplak, Q. Cai, M. J. Hsieh, J. M. Wang, Y. Duan, and R. Luo, J. Phys. Chem. B **116**, 7999 (2012).
[115] C. A. Coulson and D. Eisenberg, Proc. R. Soc. London, Ser. A **291**, 445 (1966).
[116] J. W. Caldwell and P. A. Kollman, J. Phys. Chem. **99**, 6208 (1995).
[117] P. L. Silvestrelli and M. Parrinello, Phys. Rev. Lett. **82**, 3308 (1999).
[118] M. Neumann, J. Chem. Phys. **82**, 5663 (1985).
[119] J. L. Aragones, L. G. MacDowell, and C. Vega, J. Phys. Chem. A **115**, 5745 (2011).
[120] H. E. Alper and R. M. Levy, J. Chem. Phys. **91**, 1242 (1989).
[121] J. A. Barker and R. O. Watts, Mol. Phys. **26**, 789 (1973).
[122] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, Oxford, 1987),
[123] S. B. Zhu and C. F. Wong, J. Chem. Phys. **98**, 8892 (1993).
[124] M. Sprik, J. Chem. Phys. **95**, 6762 (1991).
[125] P. Hochtl, S. Boresch, W. Bitomsky, and O. Steinhauser, J. Chem. Phys. **109**, 4927 (1998).
[126] P. Ren and J. W. Ponder, J. Phys. Chem. B **108**, 13427 (2004).
[127] J. Gao, J. Comput. Chem. **18**, 1062 (1997).
[128] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, J. Phys. Chem. **91**, 6269 (1987).
[129] G. S. Kell, J. Chem. Eng. Data **20**, 97 (1975).
[130] C. A. Angell, M. Oguni, and W. J. Sichina, J. Phys. Chem. **86**, 998 (1982).
[131] C. Vega, M. M. Conde, C. McBride, J. L. F. Abascal, E. G. Noya, R. Ramirez, and L. M. Sese, J. Chem. Phys. **132**, 046101 (2010).
[132] K. Krynicki, C. D. Green, and D. W. Sawyer, Faraday Discuss. **66**, 199 (1978).
[133] S. Tazi, A. Botan, M. Salanne, V. Marry, P. Turq, and B. Rotenberg, J. Phys. Cond. Matter **24**, 284117 (2012).
[134] A. Abragam, *The Principles of Nuclear Magnetism* (Clarendon Press, Oxford, England, 1961),
[135] D. J. Wilbur, T. Defries, and J. Jonas, J. Chem. Phys. **65**, 1783 (1976).
[136] J. Barthel, K. Bachhuber, R. Buchner, and H. Hetzenauer, Chem. Phys. Lett. **165**, 369 (1990).
[137] A. K. Soper, Chem. Phys. **258**, 121 (2000).
[138] T. Head-Gordon and M. E. Johnson, Proc. Nat. Acad. Sci. **103**, 7973 (2006).
[139] L. A. Baez and P. Clancy, J. Chem. Phys. **101**, 9837 (1994).
[140] B. Chen, J. Xing, and J. I. Siepmann, J. Phys. Chem. B **104**, 2391 (2000).
[141] S. Nose, J. Chem. Phys. **81**, 511 (1984).
[142] W. G. Hoover, Phys. Rev. A **31**, 1695 (1985).
[143] G. Raabe and R. J. Sadus, J. Chem. Phys. **134**, 234501 (2011).
[144] K. Ichikawa, Y. Kameda, T. Yamaguchi, H. Wakita, and M. Misawa, Mol. Phys. **73**, 79 (1991).
[145] Y. Mo and J. Gao, J. Phys. Chem. B **110**, 2976 (2006).
[146] S. Lifson, J. Chim. Phys. Physicochim. Biol. **65**, 40 (1968).
[147] M. Levitt and S. Lifson, J. Mol. Biol. **46**, 269 (1969).
[148] M. Levitt, Nat. Struct. Biol. **8**, 392 (2001).
[149] See supplementary material at http://dx.doi.org/10.1063/1.4816280 for optimized geometries and computed properties for water clusters and proton-water clusters using the PMOw and XP3P method and various *ab initio* molecular orbital and density functional theory approaches mentioned in the text, and average thermodynamic properties for liquid water at temperature ranging from −40 to 100 °C. In addition, figures depicting optimized structures for water clusters, computed reorientation and molecular dipole time-correlation functions, root-of-mean square displacement, heat capacities, isothermal compressibilities, and radial distributions functions for liquid water are provided.