Supporting Information for:

Building Water Models, A Different Approach

Saeed Izadi, Ramu Anandakrishnan, and Alexey V. Onufriev*

Department of Biomedical Engineering and Mechanics, Department of Computer Science, Departments of Computer Science and Physics, Virginia Tech, Blacksburg, VA 24060

E-mail: alexey@cs.vt.edu

*To whom correspondence should be addressed

Analytical solution for optimal point charges

Here we present the analytical equations to find three point charges that optimally reproduce the dipole, the quadrupole and the octupole moments of the water molecule. In the coordinate system shown in Fig. 2 (main text), the elements of the traceless dipole $\mathbf{p}_i$, quadrupole $\mathbf{Q}_{ij}$ and octupole $\mathbf{O}_{ijk}$ tensors[1] are

$$\mathbf{p}_i = (0, 0, \mu) \tag{1}$$

$$\mathbf{Q}_{ij} = \begin{pmatrix} -Q_T - Q_0/2 & 0 & 0 \\ 0 & Q_T - Q_0/2 & 0 \\ 0 & 0 & Q_0 \end{pmatrix} \tag{2}$$

$$\mathbf{O}_{ijk} = \begin{pmatrix} -\Omega_T - \Omega_0/2 & 0 & 0 \\ 0 & \Omega_T - \Omega_0/2 & 0 \\ 0 & 0 & \Omega_0 \end{pmatrix} \tag{3}$$

where $i, j = x, y$ and $k = z$, and μ, Q_0, Q_T, Ω_0 and Ω_T are the dipole, the linear component of the quadrupole, the square component of the quadrupole, the linear component of the octupole, the square component of the octupole, respectively.[2,3] The other elements of the octupole tensor ($k = x, y$) can be found by symmetry. The optimal charge values and positions are calculated so that these three moments are sequentially reproduced, starting with the lowest order moments.[1] The first two lowest order moments of the water molecule, the dipole and the quadrupole, are fully reproduced by requiring

$$\mu = 2q(z_2 - z_1) \tag{4}$$

$$Q_0 = -2q(\frac{y^2}{2} - z_2^2 + z_1^2) \tag{5}$$

$$Q_T = \frac{3qy^2}{2} \tag{6}$$

where z_2, z_1, y and q are independent unknown parameters that characterize the three point charge model (see Fig. 2). The above three equations are solved to find three geometrical parameters (z_2, z_1 and y), as follows

$$z_{1,2} = \frac{2Q_T + 3Q_0}{6\mu} \mp \frac{\mu}{4q} \tag{7}$$

$$y = \sqrt{\frac{2Q_T}{3q}} \tag{8}$$

For a given value of q, the values of z_2, z_1 and y found as above exactly reproduce the dipole (μ) and the quadrupole (Q_0 and Q_T) moments of interest. The only remaining unknown parameter, q, is found to optimally reproduce the next order moment, the octupole, which is described by two independent parameters (Ω_0 and Ω_T). The components of the octupole moment are related to the charge distribution parameters through

$$\Omega_0 = -2q(\frac{3}{2}y^2 z_2 - z_2^3 + z_1^3) \tag{9}$$

$$\Omega_T = \frac{5qy^2 z_2}{2} \tag{10}$$

The octupole tensor (Eq. 3) can be optimally approximated if the largest absolute principal value of the octupole tensor (i.e. $(\Omega_T - \Omega_0/2)$ for the water molecule) is reproduced.[1] Therefore, we set $(\Omega_T - \Omega_0/2)$ from Eqs. 9 and 10 and solve for q as

$$q = -3\frac{\sqrt{\mu^4(256Q_T^2 + \xi)} + 16Q_T\mu^2}{2\xi} \tag{11}$$

where

$$\xi = 52Q_T^2 + 60Q_TQ_0 - 9(3Q_0^2 + 8(\Omega_T - \Omega_0/2)\mu)$$

The above solution is valid only when $\xi < 0$. For $\xi \geq 0$, the point charge positions

converge to a singular point and the charge values go to infinity. The corresponding region in $\mu - Q_T$ map (Fig. 3 in the main text) leading to this condition is displayed in deepest red (zero score).

van der Waals Parameters

The usual 12-6 Lennard-Jones (LJ) potential is employed to model the van der Waals interaction among the oxygens. The Lennard-Jones function, E_{LJ}, can be written as

$$E_{LJ}(r_{oo}) = 4\epsilon_{LJ}[(\frac{\sigma_{LJ}}{r_{oo}})^{12} - (\frac{\sigma_{LJ}}{r_{oo}})^{6}] = \frac{A_{LJ}}{r_{oo}^{12}} - \frac{B_{LJ}}{r_{oo}^{6}} \tag{12}$$

The values of A_{LJ} and B_{LJ}, unlike σ_{LJ} and ϵ_{LJ},[4] are nearly independent.[5] The value of A_{LJ}, which is mainly responsible for characterizing the short-ranged repulsive interactions, is selected so that the location of the first peak of RDF $g_{oo}(r)$ is in agreement with the experiment.[6] Next, the value of B_{LJ}, which does not affect the structure significantly, is varied so that the experimental value for density is achieved.

Solvation free energy calculations

Standard thermodynamics integration (TI) protocol was adopted from Ref.[7] The Merck-Frosst implementation of AM1-BCC[8,9] was used to assign the partial charges. The topology and coordinates for the molecules were obtained from Ref.[7] Molecules were solvated in triclinic box with at least 12 Å from the solute to the nearest box edge. After minimization and equilibration, we performed standard free energy perturbation calculations using 20 λ values. Real space electrostatic cutoff was 10 Å. All bonds were restrained using the LINCS algorithm. Production NPT simulations were performed for 5ns. Identical simulations were performed for TIP3P, TIP4PEw, and OPC.

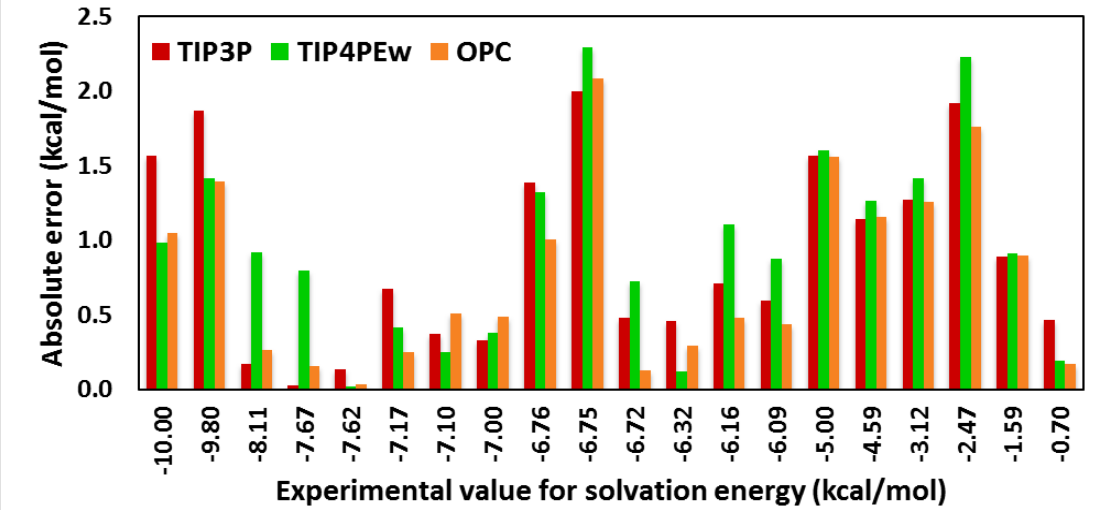

Figure 1: Absolute error relative to experiment in solvation free energies of a set of 20 small molecules calculated using TIP3P, TIP4PEw and the proposed OPC models.

Calculating the bulk properties

The calculation of bulk properties were done based on standard equations in the literature.[10–13] Unless stated otherwise, values of OPC at ambient temperature (Table 3) are given as averages over six independent simulations of 65 ns each, except for those quantities that are derived from temperature dependent results. The temperature dependent results are calculated from one simulation of 65 ns for each temperature point, i.e. 12.5K intervals in a temperature range [248K, 373K]. Details of the calculations of studied quantities are described below.

Static dielectric constant

The static dielectric constant ϵ_0 is determined through[10,12,13]

$$\epsilon_0 = 1 + \frac{4\pi}{3k_BTV}(< \mathbf{M}^2 > - < \mathbf{M} >^2) \tag{13}$$

where $\mathbf{M} = \Sigma_i q_i \mathbf{r}_i$, $\mathbf{r}_i$ is the position of atom i, k_B is the Boltzmann constant, T is the absolute temperature and V is the simulation box average volume.

Self diffusion coefficient

The self-diffusion coefficient D is obtained using the Einstein relation[10,11,13]

$$D = \lim_{t \to \infty} \frac{1}{6t} < |r(t) - r(0)|^2 > \tag{14}$$

The simulation protocol to compute the self-diffusion coefficient is similar to the protocol described in Ref.;[10] the well equilibrated NPT simulations were followed up with 80 successive intervals of NVE (20 ps) and NPT (5 ps) ensembles. The self diffusion was obtained by averaging D values over all the NVE runs.

Heat of vaporization

The heat of vaporization ΔH_{vap} is obtained following the method described in Ref.,[10] as

$$\Delta H_{vap} \approx -U_{liq}/N + RT - pV - E_{pol} + C \tag{15}$$

where U_{liq} is the potential energy of the liquid with N molecules at a given external pressure p and a temperature T, and V is the average volume of the simulation box. R is the ideal gas constant. E_{pol} accounts for the energetic cost of the effective polarization energy, and can be approximated as

$$E_{pol} = \frac{(\mu - \mu_{gas})^2}{2\alpha_{gas}} \tag{16}$$

where μ is the dipole moment of the corresponding rigid model and μ_{gas} and α_{gas} are the dipole moment and the mean polarizability of a water molecule in the gas phase,[14] respectively. The OPC's dipole is close to experiment and larger than that of common rigid models which yields a relatively larger value of E_{pol} for OPC compared to common rigid models. The last term in Eq. 15, C, is a correction to account for the change in the intramolecular vibrational modes and for nonideal gas behavior, which for various temperatures is calculated

and reported by Horn et al.[10]

Isobaric heat capacity

The isobaric heat capacity c_p is determined through numeric differentiation of simulated enthalpies $H(T)$ over the range of temperatures T of interest[10,11]

$$C_p \approx \frac{< H(T_2) > - < H(T_1) >}{T_2 - T_1} + \Delta C_{QM} \tag{17}$$

where ΔC_{QM} (≈ -2.2408 at $T = 298.0K$) is a quantum correction term accounting for the quantized character of the neglected intramolecular vibrations. The values of ΔC_{QM} for different temperatures are taken from Ref.[10] The numeric differentiation is calculated from simulations in the temperature range [248K, 373K] in 12.5K increments.

Thermal expansion coefficient

The thermal expansion coefficient α_p can be approximated through numeric differentiation of simulated bulk-densities $\rho(T)$ over a range of temperatures T of interest[10,11,13]

$$\alpha_p \approx -(\frac{\ln < \rho(T_2) > - \ln < \rho(T_1) >}{T_2 - T_1})_P \tag{18}$$

The reported value at ambient conditions is calculated from a numeric differentiation of bulk-densities at T_1=296K and T_2=300K, averaged over 4 independent simulations.

Isothermal compressibility

The isothermal compressibility κ_T is calculated from volume fluctuations in NPT simulation using a Langevin thermostat with coupling constant 2.0 ps^{-1} and a Monte Carlo barostat with coupling constant of 3.0 ps^{-1}, via the following formula[10,12,13]

$$\kappa_T = \frac{< V^2 > - < V >^2}{k_B T < V >} \tag{19}$$

Simulations of 65ns and 15ns time length were performed to obtain the temperature dependent results for ($T \leq 298K$) and ($T > 298K$), respectively.

Propensity for Charge Hydration Asymmetry

Propensity of a water model to cause Charge Hydration Asymmetry (CHA) for a similar size cation/anion pair (B^+/A^-) such as K^+/F^- is defined in Ref.[15] as

$$\eta^*(B^+/A^-) = \frac{\Delta G(B^+) - \Delta G(A^-)}{1/2|\Delta G(B^+) + \Delta G(A^-)|} \approx 2\frac{\frac{\widetilde{Q}_{zz}}{\mu}}{R_{iw}} \tag{20}$$

where the term on the right is an approximation of propensity for CHA for point charge water models ,[15] R_{iw} is the ion-water distance, ΔG is the free energy of hydration, and μ and $\widetilde{Q}_{zz}$ are the dipole and the nontraceless quadrupole moment of the model, respectively.[2]

Additional bulk properties, comparison with most recent models

O-O and O-H radial distribution functions

Each test OPC model is parametrized to exactly reproduce the position of the first peak. The positions and the heights of the remaining peaks are very accurately reproduced with these parameters. The height of the first peak is however slightly high, which leads to an average O-O coordination number (n_{oo}) larger than experiment. This may be because of the r^{-12} repulsion in the LJ potential that is known to create an over structured liquid.[16,17] It is argued that using a softer potential (e.g. a simple exponential in the form of Ae^{Br}) can correct the height of the first peak.[16] We employ a 12-6 potential to achieve compatibility

with standard biomolecular force fields. While TIP3P is the only model that accurately reproduces the height of the first peak, it lacks structure beyond the first coordination shell (Fig. 2).

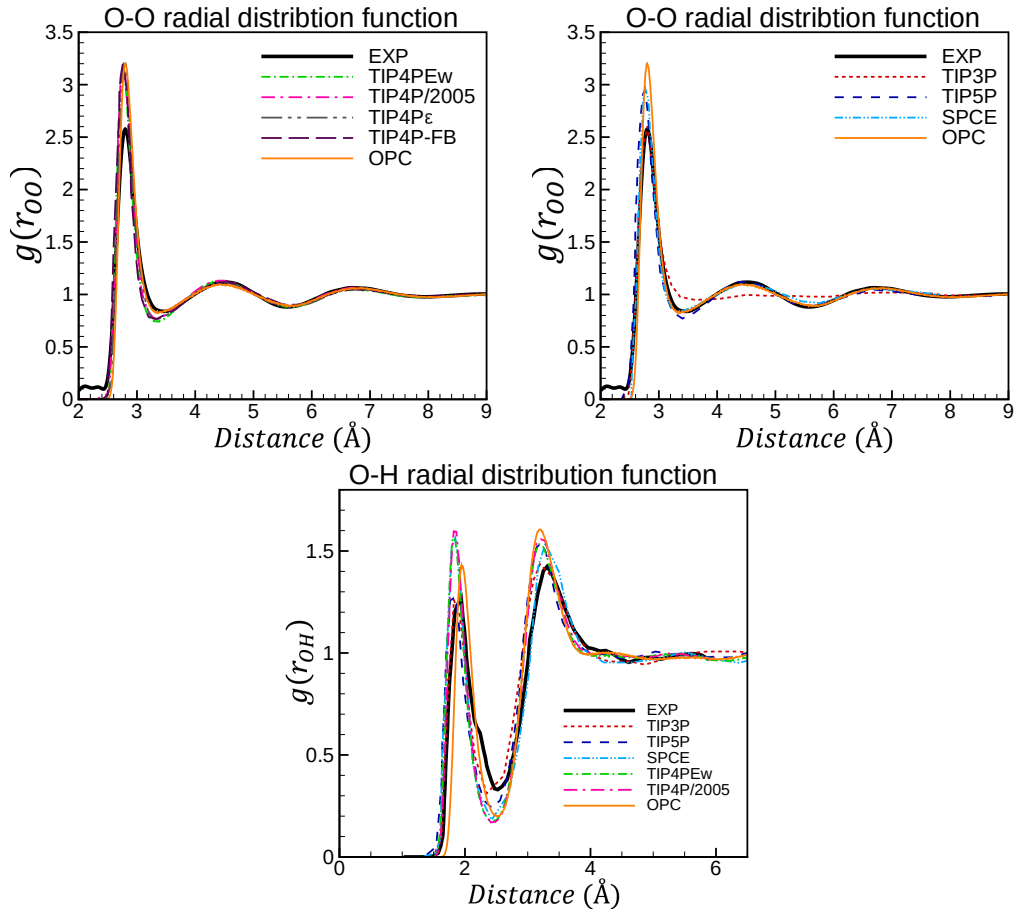

Figure 2: O-O and O-H radial distribution functions of liquid water at 298.16 K, 1 bar. The OPC model is compared to the commonly used rigid models as well as some recent rigid models (TIP4P-FB,[18] TIP4Pϵ[13] and TIP4P/2005[12]). The experimental data is taken from.[6] TIP4PEw result is from,[10] TIP4P-FB from,[18] TIP4Pϵ from,[13] SCPE from,[19] TIP3P from,[20] TIP5P from[21] and TIP4P/2005 from.[12] For simplicity, we approximated locations of the protons in OPC water by locations of the positive point charges.

Isobaric heat capacity, isothermal compressibility, recent models

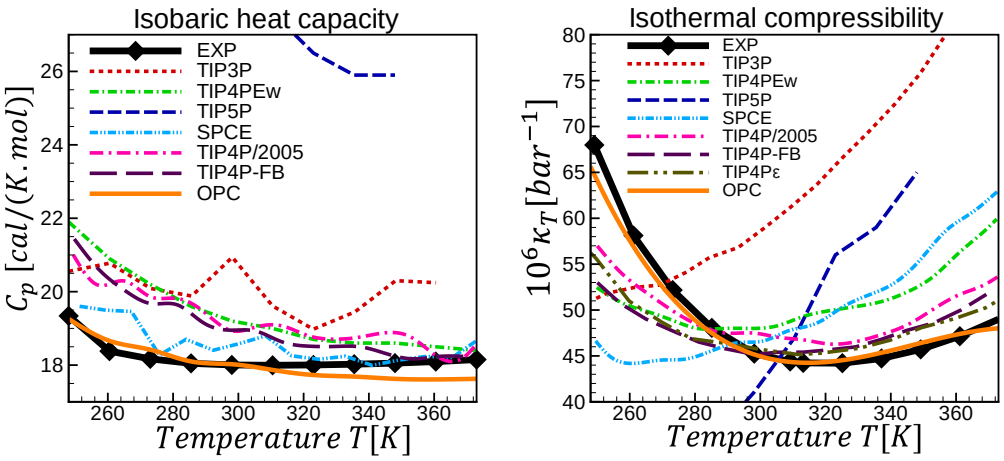

Figure 3: Variation of isobaric heat capacity and isothermal compressibility of liquid phase water with temperature. OPC model (this work) is compared to several common rigid models, some recent rigid models (TIP4P-FB,[18] TIP4Pϵ[13] and TIP4P/2005[12]) and experiment. TIP4PEw results are from,[10] TIP5P from,[21] TIP3P from,[18,22] SPCE and TIP4P-FB from,[18] and TIP4Pϵ from.[13]

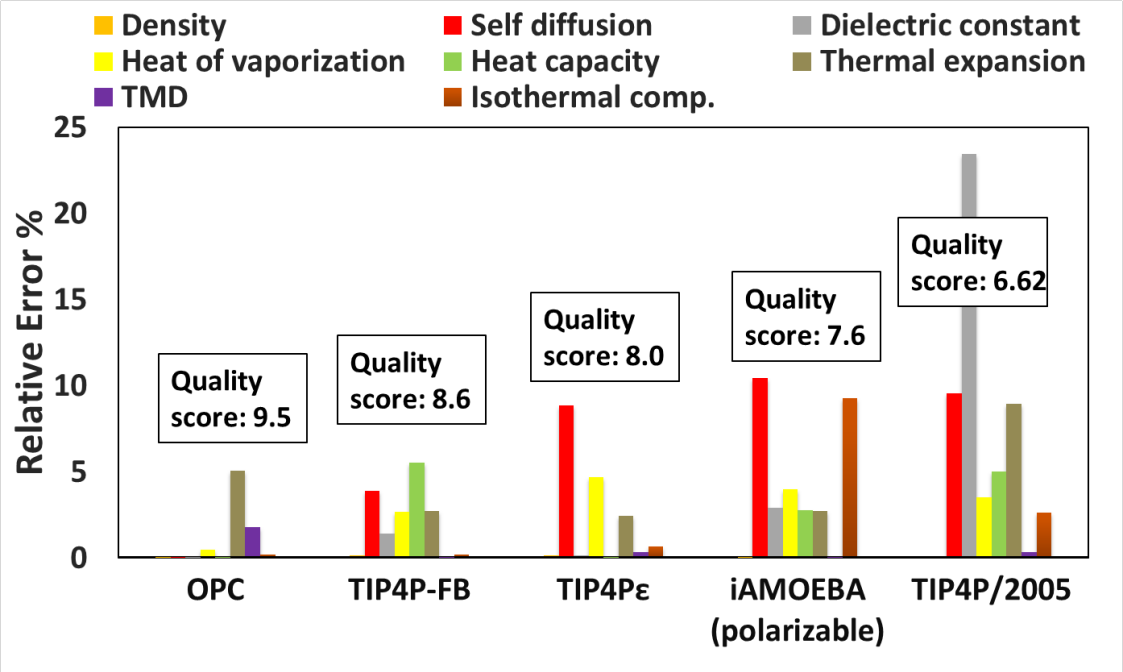

Figure 4: Comparing the accuracy of OPC to some recent rigid water models (TIP4P-FB,[18] TIP4Pϵ,[13] and TIP4P/2005[12]), including a polarizable one (iAMOEBA[23]). The quality scores (see *Methods*) represent the overall performance of each model in reproducing eight key properties, i.e. density ρ, self diffusion coefficient D, static dielectric constant ϵ_0, heat of vaporization ΔH_{vap}, isobaric heat capacity C_p, isothermal compressibility κ_T and thermal expansion coefficient α_p, at ambient conditions, as well as the temperature of maximum density (TMD). The heat capacity value for TIP4Pϵ is not reported in the original reference,[13] and therefore was excluded from the quality score calculated for this model.

References

(1) Anandakrishnan, R.; Baker, C.; Izadi, S.; Onufriev, A. V. Point Charges Optimally Placed to Represent the Multipole Expansion of Charge Distributions. *PLoS ONE* **2013**, *8*, e67715.

(2) Stone, A. *The Theory of Intermolecular Forces*; International Series of Monographs on Chemistry; Clarendon Press, 1997.

(3) Niu, S.; Tan, M. L.; Ichiye, T. The large quadrupole of water molecules. *J Chem Phys* **2011**, *134*, 134501+.

(4) Stöbener, K.; Klein, P.; Reiser, S.; Horsch, M.; Küfer, K.-H.; Hasse, H. Multicriteria optimization of molecular force fields by Pareto approach. *Fluid Phase Equilibria* **2014**, *373*, 100 – 108.

(5) Rick, S. W. A reoptimization of the five-site water potential (TIP5P) for use with Ewald sums. *J Chem Phys* **2004**, *120*, 6085–6093.

(6) Skinner, L. B.; Huang, C.; Schlesinger, D.; Pettersson, L. G. M.; Nilsson, A.; Benmore, C. J. Benchmark oxygen-oxygen pair-distribution function of ambient water from x-ray diffraction measurements with a wide Q-range. *J Chem Phys* **2013**, *138*, 074506.

(7) Mobley, D. L.; Bayly, C. I.; Cooper, M. D.; Shirts, M. R.; Dill, K. A. Small Molecule Hydration Free Energies in Explicit Solvent: An Extensive Test of Fixed-Charge Atomistic Simulations. *J Chem Theor Comp* **2009**, *5*, 350–358.

(8) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method. *J Comp Chem* **2000**, *21*, 132–146.

(9) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J Comp Chem* **2002**, *23*, 1623–1641.

(10) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J Chem Phys* **2004**, *120*, 9665–9678.

(11) Wu, Y.; Tepper, H. L.; Voth, G. A. Flexible simple point-charge water model with improved liquid-state properties. *J Chem Phys* **2006**, *124*, 024503+.

(12) Abascal, J. L. F.; Vega, C. A general purpose model for the condensed phases of water: TIP4P/2005. *J Chem Phys* **2005**, *123*, 234505+.

(13) Fuentes-Azcatl, R.; Alejandre, J. Non-Polarizable Force Field of Water Based on the Dielectric Constant: TIP4P/ϵ. *J Phys Chem B* **2014**, *118*, 1263–1272.

(14) *Guideline on the Use of Fundamental Physical Constants and Basic Constants of Water*; The International Association for the Properties of Water and Steam: Gaithersburg, Maryland, 2001.

(15) Mukhopadhyay, A.; Fenley, A. T.; Tolokh, I. S.; Onufriev, A. V. Charge hydration asymmetry: the basic principle and how to use it to test and improve water models. *J Phys Chem B* **2012**, *116*, 9776–9783.

(16) Kiss, P. T.; Baranyai, A. A systematic development of a polarizable potential of water. *J Chem Phys* **2013**, *138*, 204507+.

(17) Guillot, B. A reappraisal of what we have learnt during three decades of computer simulations on water. *J Mol Liq* **2002**, *101*, 219–260.

(18) Wang, L. P.; Martinez, T. J.; Pande, V. S. Building Force Fields: An Automatic, Systematic, and Reproducible Approach. *J Phys Chem Lett* **2014**, *5*, 1885–1891.

(19) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The missing term in effective pair potentials. *J Phys Chem* **1987**, *91*, 6269–6271.

(20) Jorgensen, W. L.; Jenson, C. Temperature dependence of TIP3P, SPC, and TIP4P water from NPT Monte Carlo simulations: Seeking temperatures of maximum density. *J Comp Chem* **1998**, *19*, 1179–1186.

(21) Mahoney, M. W.; Jorgensen, W. L. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J Chem Phys* **2000**, *112*, 8910–8922.

(22) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **1983**, *79*, 926–935.

(23) Wang, L.-P.; Head-Gordon, T.; Ponder, J. W.; Ren, P.; Chodera, J. D.; Eastman, P. K.; Martinez, T. J.; Pande, V. S. Systematic Improvement of a Classical Molecular Model of Water. *J Phys Chem B* **2013**, *117*, 9956–9972.