Supplementary Materials

The Quantum Mechanics-Based Polarizable Force Field for Water Simulations

Saber Naserifar and William A. Goddard III*

Materials and Process Simulation Center, California Institute of Technology, Pasadena, California 91125, USA

This document includes:

- The Comparison between PQEq and QM Charge Distributions in Water Dimer
- The Crystal Structures of Solid Hydrogen and Solid Oxygen
- The Two-body van der Waals Energy Curves
- The Energy Change with the Volume of the Hydrogen Crystal
- The Bond Energy to Bond Order Relationship
- The Structures of 10 Low-lying Stationary Points of the Water Dimer
- The Global Minimized Structures of 19 Water Clusters
- The Comparison between RexPoN and QM for the 4 Scan Cases
- The Comparison between RexPoN and QM Energies for Water Dimers and Clusters
- The Liquid Density Evaluation Using MD-*NPT* Simulations
- The Computational Cost of RexPoN
- References

Figures. S1 to S9

Tables S1 and S2

The Comparison between PQEq and QM Charge Distributions in Water Dimer

We optimized the PQEq parameters of water to ensure that the charge distribution on oxygen and hydrogen atoms are in reasonable ranges when two water molecules get very close to each other. This is important in order to describe the equation of state of water for high pressures.

Thus, we computed the QM (ESP and Mulliken) charge distributions on water dimer over a range of short distances and adjusted the PQEq parameters such that they produce reasonable charges, as shown in Fig. S1.

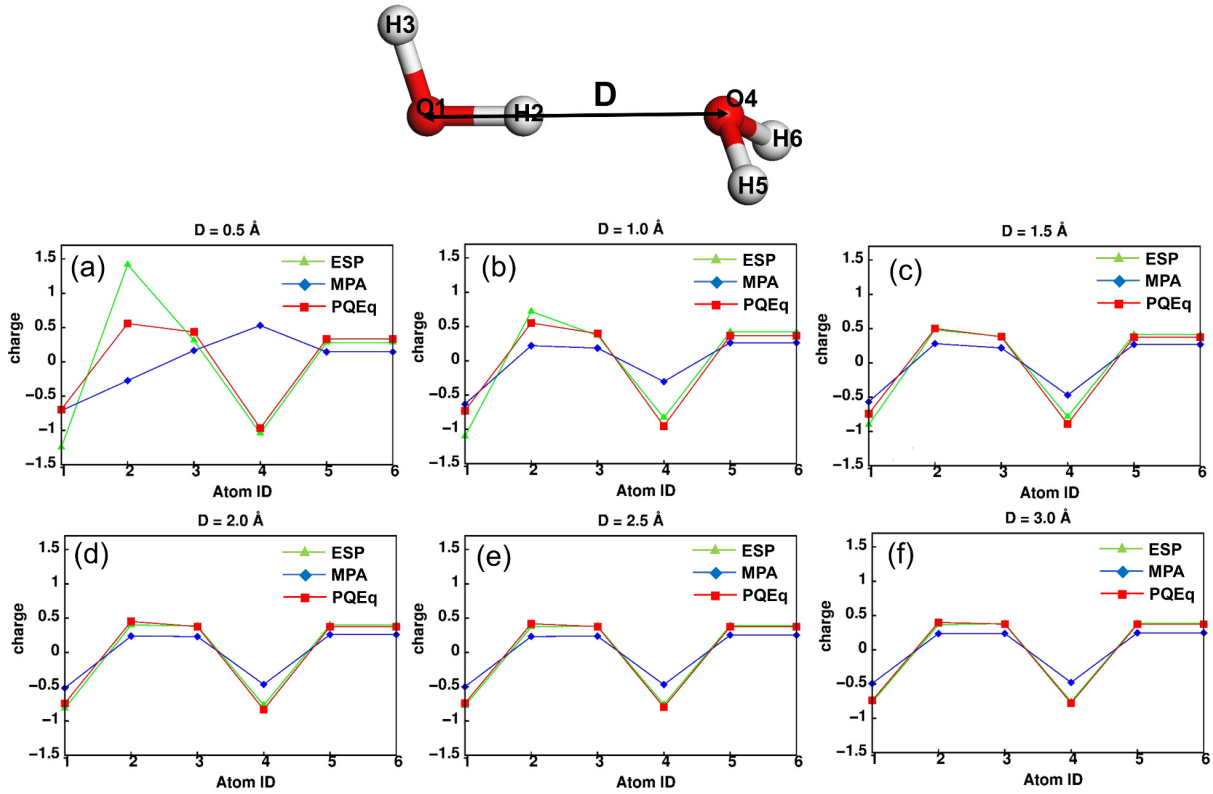


Figure S1. Comparison of atomic charges computed by PQEq, QM Mulliken population analysis (MPA), and QM electrostatic potential (ESP) for water dimer at distances of (a) 0.5 Å, (b) 1.0 Å, (c) 1.5 Å, (d) 2.0 Å, (e) 2.5 Å, and (f) 3.0 Å. The computed charges by PQEq are in a reasonable range compared to MPA and ESP charges. The *x*-axis shows the ID of atoms that are depicted in the water dimer structure at the top.

Crystal Structures of Solid Hydrogen and Solid Oxygen

Based on theoretical studies, the most stable crystal structure of hydrogen has *Pca2₁* space group containing a hexagonal close packed structure with four molecules per unit cell. For solid oxygen the most stable structure is *α*-phase, which according to Neutron-diffraction and X-ray measurements has monoclinic base-center structure with *C2/m* space group[1, 2]. See the text of manuscript for more information. The crystal structures of hydrogen and oxygen are shown in Figure S2.

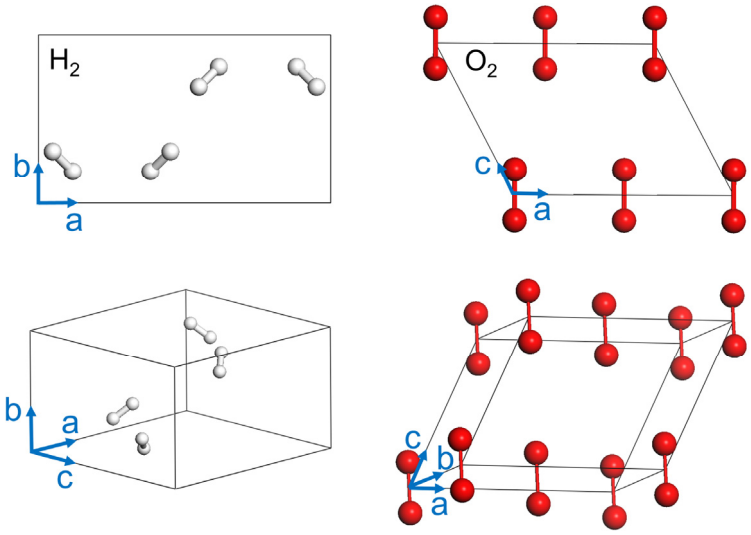


Figure S2. Crystal structures of hydrogen (left) and oxygen (right) from two views. The space groups of hydrogen and oxygen crystals are *Pca2₁* and *C2/m*, respectively.

Two-body van der Waals Energy Curves

The two-body van der Waals (vdW) energy curves of H and O were determined from the equation of states (EOS) of solid hydrogen and solid oxygen based on the density functional theory (DFT) including the empirical Grimme D3 dispersion correction (see manuscript).

To describe HO vdW interaction, we used the geometric mean average of the two-body vdW potential energy curves of H and O (see Equations 8 to 10 of the manuscript). Figure S3 shows the two-body vdW energy curves of HH, OO, and OH.

Energy Change with the Volume of the Hydrogen Crystal

The total PBE-D3 energy of hydrogen crystals at 0 and 300 K for different volumes is shown in Figure S4. The total energy curve at 300 K was fitted to third-order Birch-Murnaghan isothermal equation of state[3] (E_{BM}). The computed energy by E_{BM} at different volumes (blue) is also shown in Figure S4 which matches well with the PBE-D3 curve at 300 K (red). Then, the E_{BM} curve was used to compute pressure-volume (*P-V*) EOS and the results were compared with experimental *P-V*) EOS. See Equation 11 and Figure 3 of the manuscript.

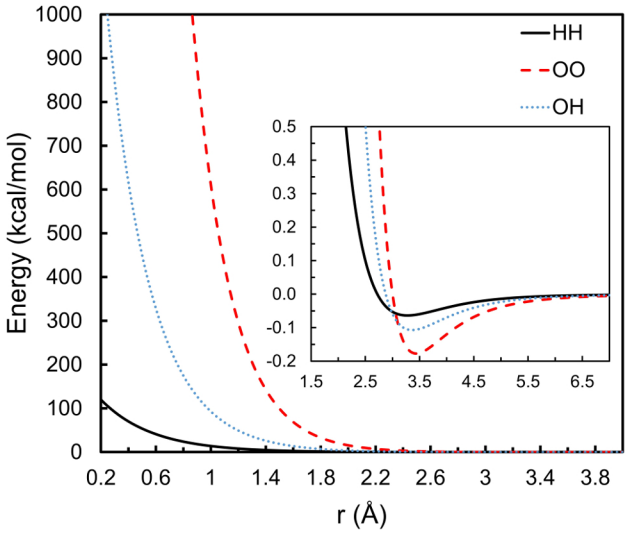


Figure S3. Two-body vdW energy curves as a function of distance for HH, OO, and OH. To describe HO vdW interaction, the geometric mean average of the two-body vdW potential energy curves of HH and OO were used.

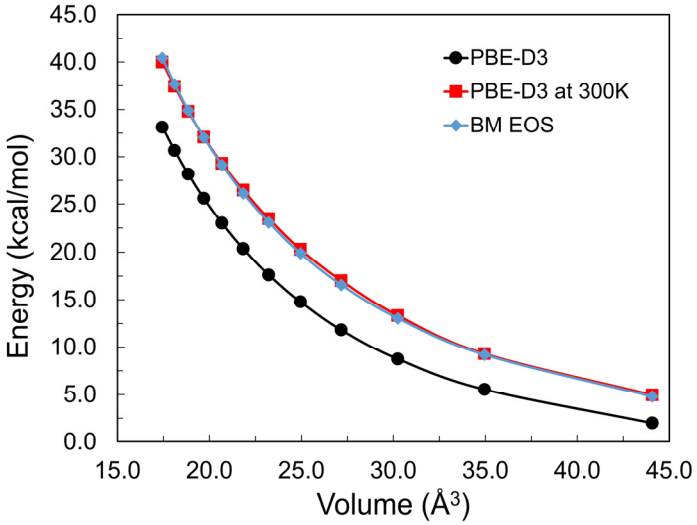


Figure S4. The energy change with the volume of the hydrogen crystal. The PBE-D3 curve (black) has been corrected for zero-point energy and thermal effects at 300 K (red). A third-order Birch-Murnaghan isothermal equation of state has been fitted (blue) to the corrected PBE-D3 (red) curve.

The Bond Energy to Bond Order Relationship

Figure S5A shows the bond order to bond distance relationships for the HH, OO, and HO-H bond dissociations. Figures S4B-D show the total normalized bond energy (NBE=E_{Q2}/E_{Re}) versus bond order (BO) for HH, OO, and HO-H bond dissociations. Note that for RexPoN the NBE is not monotonically attractive as it was in ReaxFF. Rather the NBE has a maximum near R_e (BO=1) and decreases for small distances (larger BO). This is expected from valence bond theory where we know that the bonding contribution is dominated by the decrease in kinetic energy resulting from opposite gradients in the two valence bond orbitals (contragradience) for the regions between the atoms[4, 5]. For small distances this goes to zero.

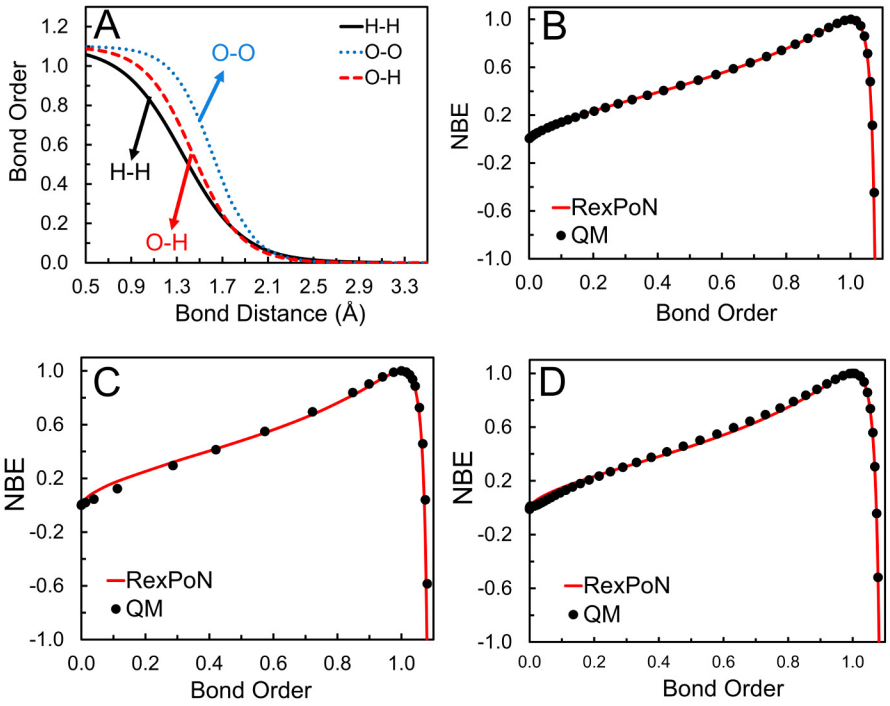


Figure S5. (A) Bond order versus bond distance and normalized bond energy (NBE) versus bond order for (B) HH, (C) OO, and (D) HO-H bond dissociations.

The Structures of 10 Low-lying Stationary Points of the Water Dimer

The structures of 10 stationary points of water dimer from accurate water potential energy surfaces (Bowman PES[6]) are shown in Figure S6.

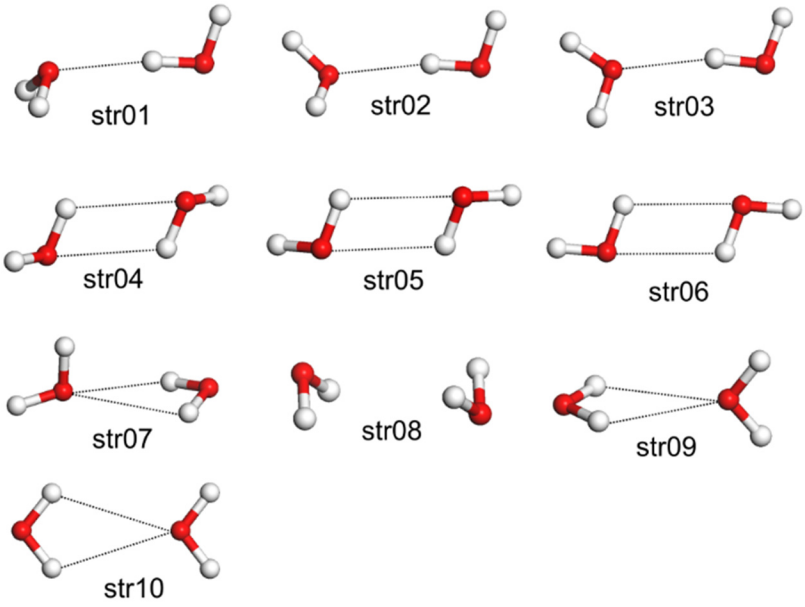
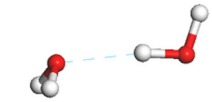


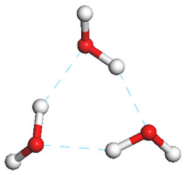
Figure S6. The structures of 10 low-lying stationary point of the water dimer. Str01 is the most stable (global minimum) and str08 is the least stable (no hydrogen bonding) structures.

The Global Minimized Structures of 19 Water Clusters

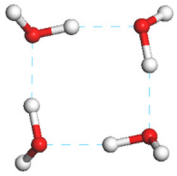
To account for non-additive hydrogen bond effects that exist in a bulk system, the training set was expanded to include the global minima of water clusters, $(H_2O)_n$, containing up to n=19 waters. The reference energies and global minimized structures were taken from X3LYP[7] hybrid DFT calculations which leads to excellent structures for large water clusters. The cluster structures are provided in Figure S7.



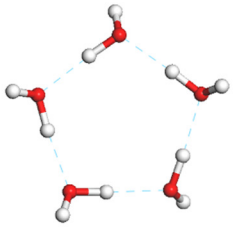
2-Linear



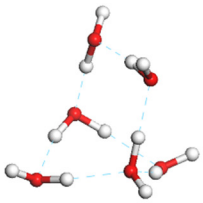
3-Cyclic



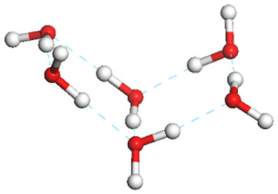
4-Cyclic



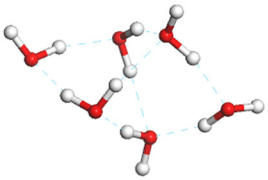
5-Cyclic



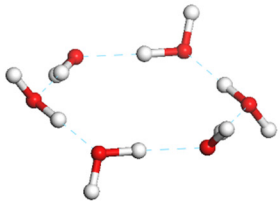
6-Bag



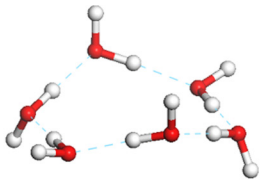
6-Book



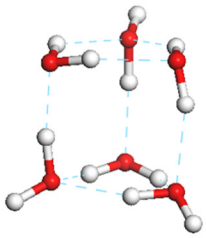
6-Cage



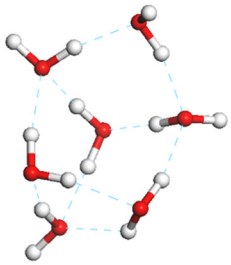
6-Cyclic



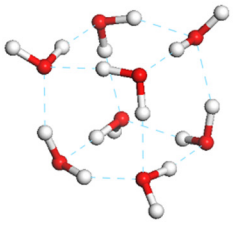
6-Cyclic'



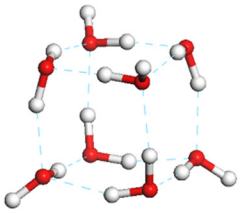
6-Prism



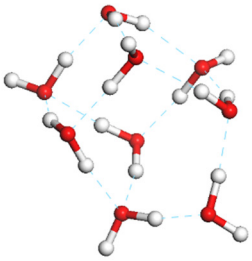
7-Prism



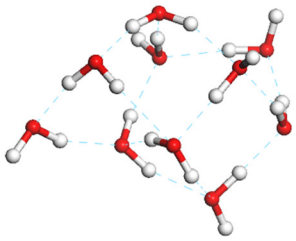
8-D2d



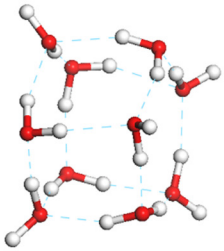
8-S4



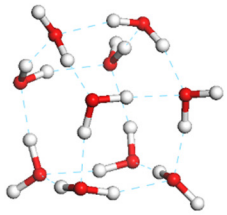
9-D2dDD



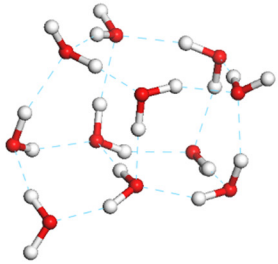
10-Butterfly



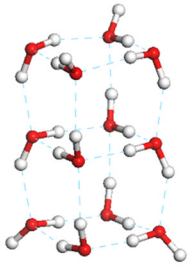
10-Prism



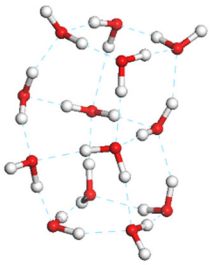
10-Prism'



11-Pr443



12-Pr444



13-Pr454

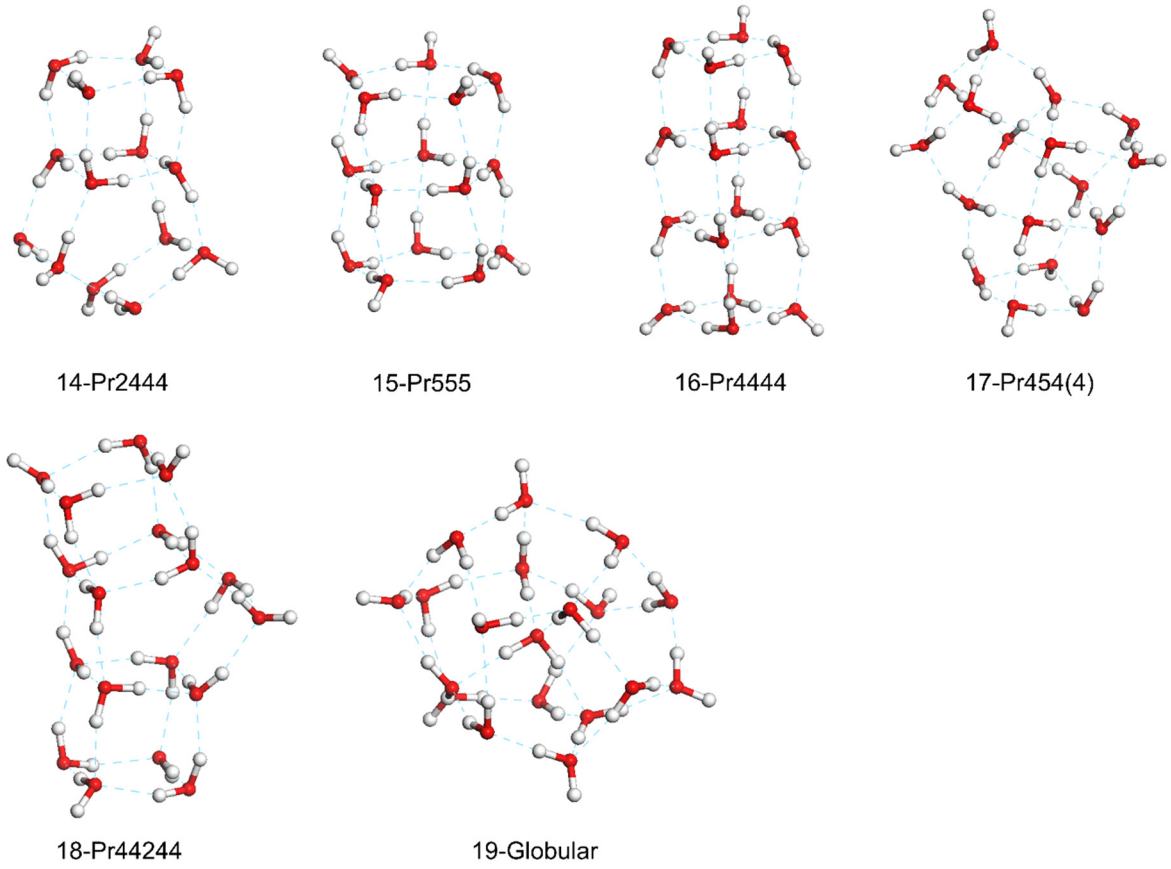


Figure S7. Water cluster, $(H_2O)_n$, structures for n=2 to19. The global minimized structures were taken from X3LYP[7] hybrid DFT calculations which leads to excellent structures for large water clusters.

The Comparison between RexPoN and QM for the 4 Scan Cases

Figure S8 provides the comparison between RexPoN (using FF2) and QM for the 4 scan cases. See Figure 7 and the text of manuscript for the descriptions of 4 cases.

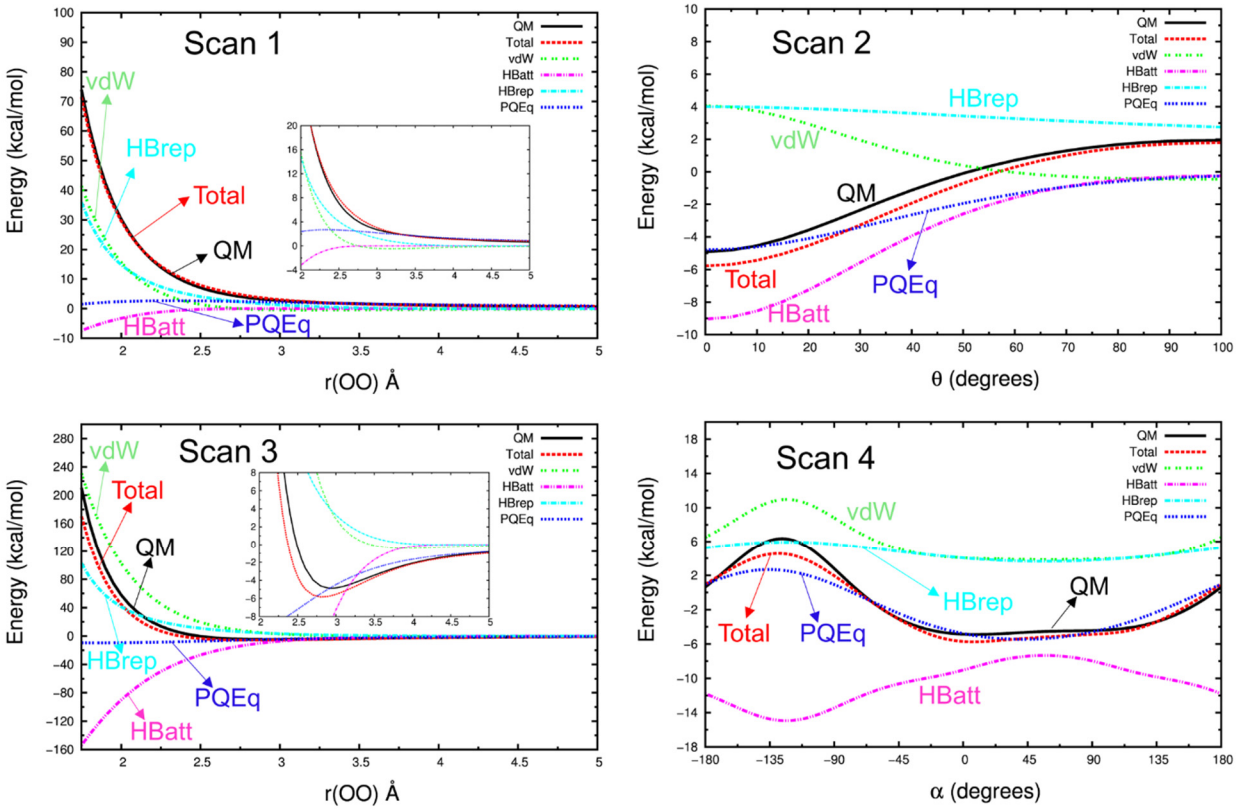


Figure S8. Energy comparison between RexPoN and QM using the FF2 optimized parameter set for the 4 water dimer scan cases. The energy components of the RexPoN are also shown for comparison.

The Comparison between RexPoN and QM Energies for Water Dimers and Clusters

The FF2 optimized force field was used to compute the RexPoN energies for the 10 water dimers and 19 clusters. The results are shown in Table S1. See the text of manuscript for more explanations.

Table S1. Comparison of computed energies in kcal/mol/Nmolec of QM and RexPoN (FF2) for 10 low-lying stationary points of water dimers (see Figure S6) and 19 water clusters (see Figure S7).

No. of waters	Geometry	QM Energy	RexPoN Energy
2	str01	-2.49	-2.86
2	str02	-2.23	-2.44
2	str03	-2.20	-2.50
2	str04	-2.12	-1.74
2	str05	-2.00	-1.36
2	str06	-1.97	-1.27
2	str07	-1.61	-1.54
2	str08	-0.74	-0.40
2	str09	-1.63	-1.88
2	str10	-1.17	-1.60
2	Linear	-2.50	-2.92
3	Cyclic	-5.17	-4.39
4	Cyclic	-6.93	-5.90
5	Cyclic	-7.31	-6.12
6	Bag	-7.40	-6.49
6	Book	-7.53	-6.42
6	Cage	-7.46	-6.56
6	Cyclic	-7.51	-6.04
6	Cyclic'	-7.35	-5.87
6	Prism	-7.45	-7.10
7	Prism	-7.90	-7.50
8	D2d	-8.92	-8.56
8	S4	-8.92	-8.61
9	D2dDD	-8.93	-8.38
10	Butterfly	-8.41	-7.44
10	Prism	-9.12	-8.74
10	Prism'	-9.18	-8.93
11	Pr443	-8.89	-8.35
12	Pr444	-9.35	-9.22
13	Pr454	-9.42	-8.74
14	Pr2444	-9.56	-9.39
15	Pr555	-9.49	-8.93
16	Pr4444	-9.57	-9.19
17	Pr454(4)	-9.60	-9.66
18	Pr44244	-9.76	-9.52
19	Globular	-9.69	-9.41

The Liquid Density Evaluation Using MD-*NPT* Simulations

Figure 12 showed that calculating the density of water at 298K and 1 atm, by doing the equation of state with multiple NVT MD simulations leads to a predicted density of 0.9965 in exact agreement with the experimental value.

We report here the result of MD-*NPT* simulations were utilized for 1 nanosecond (ns) at 298 K to evaluate the liquid density of water for a system of 216 water molecules. See the text of the manuscript for the details of simulations. The change of pressure and density with time for the last 0.5 ns of simulations are shown in Figure S9. The data were averaged over 1 picosecond (ps) time intervals. The average of density gives a density of 0.9960 gr/cm^3 in excellent agreement with the experimental density of 0.9965 gr/cm^3.

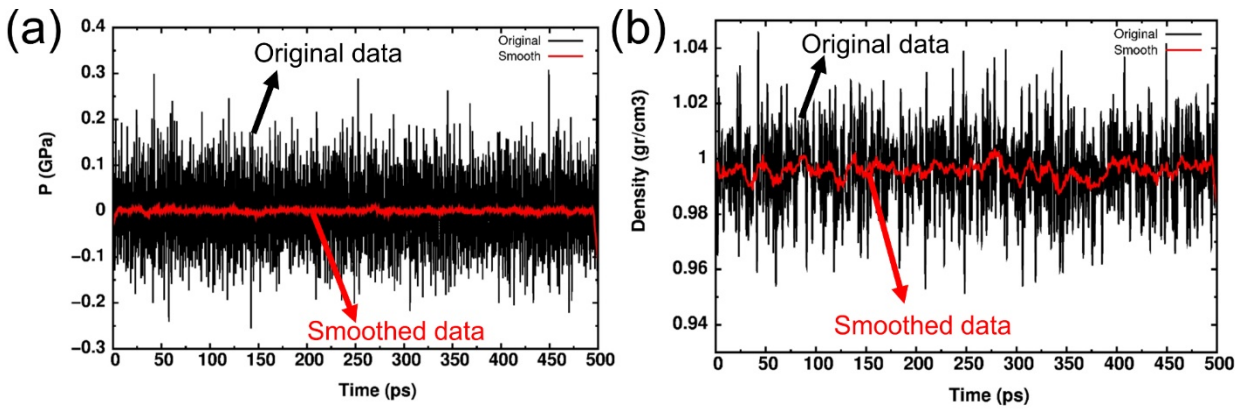


Figure S9. The change of (a) pressure and (b) density with time during 0.5 ns MD-*NPT* simulations of liquid water at 298 K. The simulation cell contains 216 water molecules. The original data (black) were averaged over 1 ps time intervals (red curve). The average of density curve gives a density of 0.9960 gr/cm^3 in excellent agreement with the experimental density of 0.9965 gr/cm^3.

The Computational Cost of RexPoN compared to the TIP3P standard water force field

The computational cost of RexPoN in terms of valence (bond and angle), hydrogen-bond, and non-bond calculations is very similar to other water models. However, RexPoN has additional costs from allowing the charge and polarization to change dynamically (PQEq) and from our use of a 12A cutoff for nonbond (van der Waals) interactions instead of the 8A cutoff commonly used in water FF.

To provide a measure of these extra costs, we ran 10 ps NVT-MD simulations of 216 water molecules at 10 K using RexPoN and TIP3P models. We used a single processor on the same machine for all calculations. The computational costs in terms of central processing unit (CPU seconds) are given in Table S2. We used several scenarios (Cases 1 to 6) that include different cutoff values, variable and fixed charge and shell position in order to provide a better comparison.

When charges and shells are not updated and the same cutoff of 8.0 Å is used (Case 1), the computational cost of RexPoN and TIP3P (also fixed charges) are reasonably close (TIP3P=108 vs. RexPoN=175 CPU seconds). The higher computational cost of RexPoN is due to the use of a Gaussian Charge distribution instead of a point charge.

Using a smaller cutoff value of 8.0 Å but updating the charges and polarization every time step increases the RexPoN time to 1567 CPU seconds (Case 6). This is the most likely level to be used for large protein simulations. We expect the accuracy of the RexPoN with the smaller cutoff to still be far better than any other water force field. With a standard protein MD simulation using a standard forcefield (Amber or Charmm) the increase in costs for the RexPoN would be relatively less.

Larger values of cutoff (12.0 Å in Case 2) but fixed charges and polarization results in an increase of computational costs by a factor of ~3.4 in both models (TIP3P=367 vs. RexPoN=592 CPU seconds).

Only letting charges to be updated without shell polarization (Case 3) the computational cost increases to 2797 CPU seconds. Allowing only shells to polarize (Case 4) the computational cost is 1037 CPU seconds.

Doing full PQEq calculations where charges are variable and shell positions are dynamic (Case 5) the computational cost increases to 4361 CPU seconds.

Table S2. Comparisons of computational processing units (CPUs) of RexPoN and TIP3P model. The *NVT*-MD simulations were performed for 10000 steps (10 ps) at the temperature of 10 K. Computational costs of PQEq components are studied using different scenarios (Cases 1 to 6).

	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
Variable Charge	No	No	Yes	No	Yes	Yes
Shell Polarization	No	No	No	Yes	Yes	Yes
Cutoff (Å)	8.0	12.0	12.0	12.0	12.0	8.0
TIP3P (cpu)	**108**	**367**	**-**	**-**	**-**	**-**
RexPoN (cpu)	**175**	**592**	**2797**	**1037**	**4361**	**1567**

References

1. C. Barrett, L. Meyer and J. Wasserman, J. Chem. Phys. **47** (2), 592-597 (1967).
2. Y. A. Freiman and H.-J. Jodl, Physics Reports **401** (1-4), 1-228 (2004).
3. F. Birch, Phys. Rev. **71** (11), 809 (1947).
4. W. A. Goddard and C. W. Wilson, Theor. Chim. Acta **26** (3), 211-230 (1972).
5. C. W. Wilson and W. A. Goddard, Theor. Chim. Acta **26** (3), 195-210 (1972).
6. A. Shank, Y. Wang, A. Kaledin, B. J. Braams and J. M. Bowman, J. Chem. Phys. **130** (14), 144314 (2009).
7. J. T. Su, X. Xu and W. A. Goddard, J. Phys. Chem. A **108** (47), 10518-10526 (2004).