# Toward a Universal Water Model: First Principles Simulations from the Dimer to the Liquid Phase

3 **AUTHORS**, INCLUDING:

Greg Medders
University of California, San Diego
**11** PUBLICATIONS **165** CITATIONS

SEE PROFILE

Francesco Paesani
University of California, San Diego
**92** PUBLICATIONS **2,354** CITATIONS

SEE PROFILE

# Toward a universal water model:
# First principles simulations from the dimer to the liquid phase

V. Babin,[1, a)] G.R. Medders,[1, a)] and F. Paesani[1, b)]
*Department of Chemistry and Biochemistry, University of California, San Diego*
*La Jolla, California 92093*

(Dated: 29 October 2012)

A full-dimensional molecular model of water, HBB2-pol, derived entirely from "first principles", is introduced and employed in computer simulations ranging from the dimer to the liquid. HBB2-pol provides excellent agreement with the measured second and third virial coefficients and, by construction, reproduces the dimer vibration-rotation tunneling spectrum. The model also predicts the relative energy differences between isomers of small water clusters within the accuracy of highly correlated electronic structure methods. Importantly, when combined with simulation methods that explicitly include zero-point energy and quantum thermal motion, HBB2-pol accurately describes both structural and dynamical properties of the liquid phase. The predictive power of the HBB2-pol quantum simulations opens the door to the long-sought molecular-level understanding of water under different conditions and in different environments.

Given the central role that water plays in nature, it is not surprising that many studies have attempted to develop a microscopic picture of its unique properties. In particular, since the advent of computer simulations, myriad molecular models based on both force fields (including different degrees of empiricism) and ab initio approaches have been proposed. Force field-based models range from coarse-grained representations with no atomistic details (e.g, see Ref. 1), to classical parameterizations in terms of point charges and rigid bonds (see Ref. 2) for a recent review), to yet more sophisticated models that account for molecular flexibility (e.g., Ref. 3), electronic polarization (e.g., Ref. 4), and charge transfer[5]. On the other hand, due to the associated computational cost, purely ab initio models so far have been limited to the use of density functional theory (DFT) (e.g., Ref. 6). However, despite much recent progress, none of the existing models is capable of correctly describing the properties of water from isolated molecules and small clusters in the gas phase up to the liquid and solid phases. Due to the broad spectrum of (often conflicting) predictions derived from these models, the microscopic behavior of water under different conditions and in different environments remains the subject of continuing debate[7–12].

Independently of the number of molecules, the "first principles" modeling of water within the Born-Oppenheimer approximation builds upon two components: a faithful representation of the electronic potential energy surface (PES) and the proper treatment of the nuclear motion at a quantum-mechanical level. In principle, the multidimensional PES can be accurately approximated at the coupled cluster level of theory including single, double and perturbative triple excitations, CCSD(T)[13], which currently represents the gold standard in quantum chemistry. For the nuclear degrees
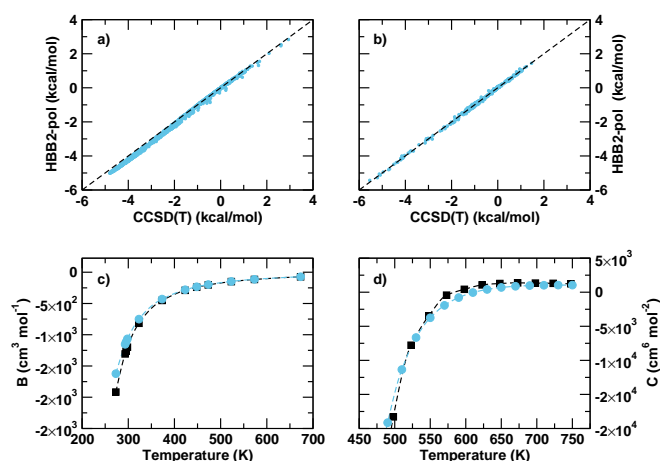


FIG. 1. 2-body (a) and 3-body (b) interactions along with the second (c) and third (d) virial coefficients. Experimental data for second[33] and third[34] virial coefficients are shown as black squares and the HBB2-pol values as light blue circles.

of freedom, quantum dynamics methods based on basis set expansions of the nuclear wavefunction are well suited to the study of small complexes[14], while simulation approaches based on the path-integral formalism allow the fully quantum-mechanical modeling of water in condensed phases[15]. In the end, these two components must be combined in an efficient computational scheme that enables the calculation of statistically converged quantities.

Unfortunately, the computational cost associated with CCSD(T) makes these calculations prohibitively expensive even for small water clusters. To overcome this computational barrier while still providing an ab initio representation of the molecular interactions, less accurate but more efficient DFT approaches have been widely applied to water simulations. However, the choice of the most appropriate DFT model for water remains the sub-

a)Contributed equally to this work
b)Electronic mail: fpaesani@ucsd.edu

ject of ongoing research[16,17]. Alternatively, the PES for a system of N water molecules can be expressed through the many-body expansion of interaction energies, consisting of 1-body, 2-body, ..., N-body terms[18]. It has been shown that this expansion converges rapidly for water, such that it is sufficient to take into account only the first few[19]. Since the low-order interactions can be accurately calculated using CCSD(T), the many-body expansion effectively enables the representation of the multidimensional PES at the CCSD(T) level of theory[20].

This strategy has been rigorously followed in the development of the CC-pol[21] and WHBB[29] models. In CC-pol, which describes the water molecules as rigid, the 2-body term was derived from CCSD(T) data while Hartree-Fock calculations were used to fit the 3-body term. WHBB, which allows for molecular flexibility, was parameterized using CCSD(T) and MP2 reference data for the 2-body and 3-body interaction terms, respectively. Both models include higher-body interactions through point polarizable dipoles. CC-pol accurately reproduce the vibration-rotation tunneling (VRT) spectrum of the water dimer and has been used in classical molecular dynamics (MD) simulations of liquid water[21]. However, these simulations explicitly neglect nuclear quantum effects such as zero-point energy and quantum thermal motion, which have been shown to be important[15]. A new version of CC-pol with flexible monomers has recently been applied to study the water dimer[22]. WHBB also reproduces the VRT spectrum of the dimer with high accuracy[23]. Due to the computational cost associated with the high dimensionality of the polynomials used to represent the 3-body interactions, quantum simulations with the WHBB model so far have been limited to small clusters[24].

Here, we report on structural and dynamical properties of water calculated with the newly developed full-dimensional model, HBB2-pol, derived entirely from first principles. Similarly to CC-pol and WHBB, HBB2-pol is built upon the many-body expansion of the molecular interactions. The 1-body term associated with intramolecular distortion is described by the spectroscopically accurate PES developed by Partridge and Schwenke[25]. The 2-body interaction at short range is represented by the HBB2 potential[26], which smoothly transitions between 5.5 /AA and 7.5 /AA into the sum of electrostatic and dispersion interactions, reproducing the correct asymptotic behavior. The induction contributions to non-pairwise additive interactions are taken into account using Thole-type point polarizable dipoles on all atomic sites. In addition, an explicit 3-body component is introduced to account for short-range exchange-repulsion and charge transfer, which have been shown to make a significant contribution to the 3-body interaction[27,28].

The 3-body interaction of HBB2-pol is unique in that it is currently the first that has been fitted to CCSD(T) data. Importantly, the inclusion of induction interactions in the 3-body term enables the use of lower-degree polynomials than previously reported[29], resulting in a
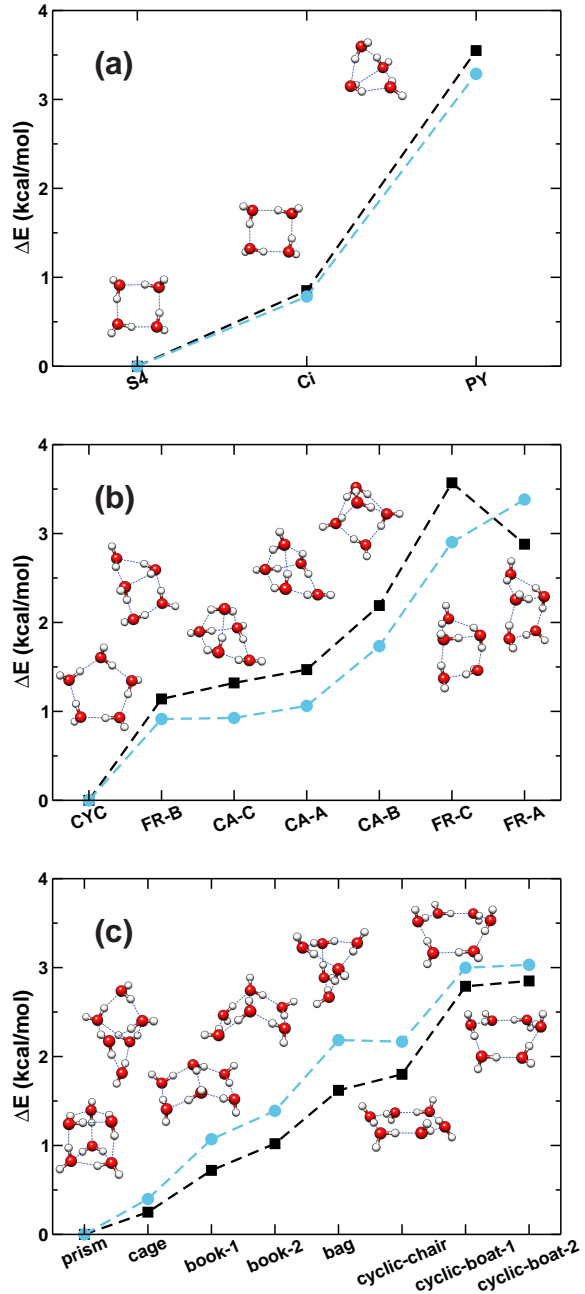


FIG. 2. Relative energies of the low-lying isomers of the water tetramer (a), pentamer (b), and hexamer (c) computed with the HBB2-pol model (light blue circles) and state-of-the-art quantum chemistry methods (black squares)[35,36].

sizeable decrease in the computational cost associated with the many-body expansion. Specifically, the explicit 3-body component takes the form of a third-degree polynomial in the exponential of the interatomic distances, which is invariant with respect to the permutations of equivalent atoms. This short-range correction was found

to be indispensable to overcome the limitations of current polarizable force fields (e.g., Ref. 30), which attempt to describe all non-pairwise additive interactions through point polarizable dipoles. These findings suggest that the neglect of accurate short-range 3-body interactions invariably leads to an incorrect description of the liquid structure.

The ability of HBB2-pol to accurately reproduce the molecular interactions of water is established by comparison to high-level electronic structure calculations and experimental measurements reported in the literature. Figures 1a and 1b show the correlation plots of the HBB2-pol and CCSD(T) 2-body and 3-body interaction energies, respectively. For this analysis approximately 1400 dimer and 500 trimer configurations were extracted from classical MD simulations of the hexamer, ice Ih, and liquid water. CCSD(T) energies were calculated with the aug-cc-pVTZ basis set[31] and corrected for the basis set superposition error through the counterpoise method[32]. Nearly perfect agreement is found over the entire range of energies, indicating that HBB2-pol accurately reproduces the energetics of both favorably interacting and distorted dimers and trimers.

An additional measure of the overall accuracy of the 2-body and 3-body interactions is provided by the second and third virial coefficients, which directly probe dimer and trimer interactions, respectively. The HBB2-pol results are shown in Figures 1c and 1d along with the corresponding experimental data[33,34]. In the calculation of the second virial coefficient, nuclear quantum effects were explicitly included through the path-integral formalism (see Supporting Information for details). Since experimental data for the third virial coefficient are only available at relatively high temperatures where quantum effects are less important, these calculations were carried out at the classical level with each water monomer held fixed at the ground-state vibrationally-averaged geometry. For both virial coefficients, the calculated values are in close agreement with the corresponding experimental data over the entire ranges of temperature, providing further support of the accuracy of the 2-body and 3-body interactions of HBB2-pol.

The energetics of small water clusters with more than three molecules allows for a quantitative assessment of the ability of HBB2-pol to correctly describe N-body interactions with N > 3 and represents a stringent test for the overall many-body expansion. For this purpose, the relative energies of the low-lying isomers of the water tetramer, pentamer, and hexamer are compared in Figure 2 with the corresponding ab initio values reported in the literature[35,36]. Water clusters in this size range are among the largest ones for which highly-correlated electronic structure calculations are still feasible. In all cases, HBB2-pol predicts the correct energy ordering of the isomers, with the energy differences being within the intrinsic accuracy of the ab initio methods[13].

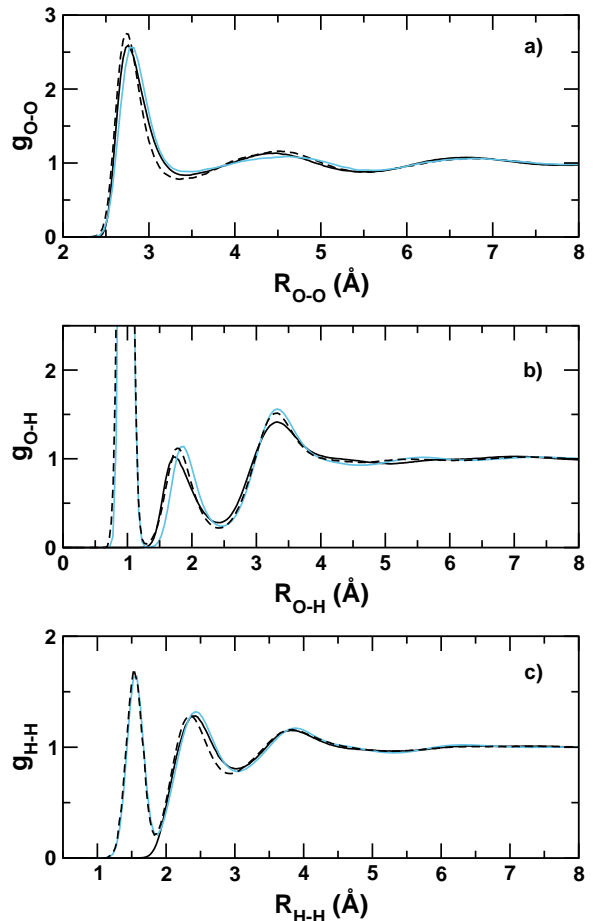The comparisons shown in Figures 1 and 2 demonstrate that the degree of accuracy of HBB2-pol is com-



FIG. 3. Oxygen-oxygen, O-O (a), oxygen-hydrogen, O-H (b), hydrogen-hydrogen, H-H (c), radial distribution functions. Experimental data from Refs. 37 and 38 are shown as dashed and solid black lines, respectively. The quantum RDFs calculated with the HBB2-pol model are shown as light blue lines.

parable to that of state-of-the-art quantum chemistry methods. While these highly-correlated calculations are restricted to small clusters, quantum simulations with HBB2-pol enable a microscopically detailed characterization of water in condensed phases. To this end, HBB2-pol was used in fully quantum molecular dynamics simulations of liquid water at ambient conditions (T = 298.15 K and $\rho = 0.997$ g cm$^{-3}$). Specifically, path-integral molecular dynamics (PIMD) and centroid molecular dynamics (CMD) simulations were performed to determine structural and dynamical properties of the liquid phase, respectively[15]. Both PIMD and CMD are based on Feynman's formulation of statistical mechanics in terms of path-integrals and have been shown to accurately describe nuclear quantum effects in condensed phases of water[15]. All simulations were carried out with a system consisting of 256 molecules in a periodic cubic box.

The oxygen-oxygen (O-O), oxygen-hydrogen (O-H),

TABLE I. Comparison between the experimental and calculated diffusion coefficient (D) and orientational relaxation time ($\tau_2$) of liquid water at ambient conditions.

|  | Experiment | Simulation |
|---|---|---|
| D ($\text{Å}^2\text{ps}^{-1}$) | 0.23[a] | 0.23 ± 0.5 |
| $\tau_2$ (ps) | 2.5[b] | 2.5 ± 0.2 |

[a] Ref. 40
[b] Ref. 41

and hydrogen-hydrogen (H-H) radial distribution functions (RDFs) calculated from an 80 ps PIMD simulation are compared in Figure 3 with two sets of experimental data[37,38]. The PIMD simulations correctly predict a lower first peak in the O-O RDF, as determined by the most recent experiments. This feature has been proven difficult to reproduce by current force field-based and ab initio models. Some differences between the PIMD results and the experimental data exist for the second peak of the O-H RDF. However, both position and shape of this peak describing the spatial correlation between O and H atoms directly involved in hydrogen bonds are difficult to be experimentally determined as demonstrated by the appreciable differences between the two sets of experimental data. The diffusion coefficient (D) and orientational relaxation time ($\tau_2$) calculated at the quantum-mechanical level by averaging over six CMD trajectories of 10 ps each are listed in Table I. For both quantities, the CMD results are in quantitative agreement with the corresponding experimental values, providing further evidence of the accuracy of the HBB2-pol model. The average structure of the first hydration shell was determined by labeling the water molecules according to the number of donating hydrogen bonds as non-donor, single-donor, and double-donor. Using the geometric criterion proposed in Ref. 7, the PIMD simulations predict that ∼57% of the molecules are double donors and ∼38% are single donors. Each molecule is involved, on average, in ∼3 hydrogen bonds.

In summary, the full-dimensional HBB2-pol model, based entirely on first principles, has been introduced and employed in calculations of the water properties from the dimer to the liquid phase. Being derived from state-of-the-art quantum chemistry calculations, HBB2-pol represents a major step toward the long-sought "universal model" capable of describing the behavior of water under different conditions and in different environments[39]. Future simulation studies with the HBB2-pol model will allow the resolution of current controversies regarding structural, thermodynamic, and dynamical properties of bulk, interfacial, and supercooled water.

[1] V. Molinero, E. B. Moore, *J. Phys. Chem. B* **113**, 4008 (2009).
[2] C. Vega, J. Abascal, *Phys. Chem. Chem. Phys.* **13**, 19663 (2011).
[3] L. X. Dang, B. M. Pettitt, *J. Phys. Chem.* **91**, 3349 (1987).
[4] G. Fanourgakis, S. Xantheas, *J. Chem. Phys.* **128**, 074506 (2008).
[5] A. J. Lee, S. W. Rick, *J. Chem. Phys.* **134**, 184507 (2011).
[6] K. Laasonen, M. Sprik, M. Parrinello, R. Car, *J. Chem. Phys.* **99**, 9080 (1993).
[7] P. Wernet, *et al.*, *Science* **304**, 995 (2004).
[8] G. N. Clark, C. D. Cappa, J. D. Smith, R. J. Saykally, T. Head-Gordon, *Mol. Phys.* **108**, 1415 (2010).
[9] P. A. Pieniazek, C. J. Tainter, J. L. Skinner, *J. Am. Chem. Soc.* **133**, 10360 (2011).
[10] S. Nihonyanagi, *et al.*, *J. Am. Chem. Soc.* **133**, 16875 (2011).
[11] P. Kumar, G. Franzese, H. E. Stanley, *J. Phys. Condens. Matter* **20**, 244114 (2008).
[12] D. T. Limmer, D. Chandler, *J. Chem. Phys.* **135**, 134503 (2011).
[13] K. Raghavachari, G. Truck, J. Pople, M. Head-Gordon, *Chem. Phys. Lett.* **157**, 479 (1989).
[14] H.-D. Meyer, U. Manthe, L. Cederbaum, *Chem. Phys. Lett.* **165**, 73 (1990).
[15] F. Paesani, G. A. Voth, *J. Phys. Chem. B* **113**, 5702 (2009).
[16] C. Zhang, J. Wu, G. Galli, F. Gygi, *J. Chem. Theory Comput.* **7**, 3054 (2011).
[17] I.-C. Lin, A. P. Seitsonen, I. Tavernelli, U. Rothlisberger, *J. Chem. Theory Comput.* **8**, 3902 (2012).
[18] D. Hankins, J. Moskowitz, F. Stillinger, *J. Chem. Phys.* **53**, 4544 (1970).
[19] S. Xantheas, *J. Chem. Phys.* **100**, 7523 (1994).
[20] U. Góra, R. Podeszwa, W. Cencek, K. Szalewicz, *J. Chem. Phys.* **135**, 224102 (2011).
[21] R. Bukowski, K. Szalewicz, G. Groenenboom, A. van der Avoird, *Science* **315**, 1249 (2007).
[22] C. Leforestier, K. Szalewicz, A. van der Avoird, *J. Chem. Phys.* **137**, 014305 (2012).
[23] C. Leforestier, *Phil. Trans. R. Soc. A* **370**, 2675 (2012).
[24] Y. Wang, V. Babin, J. Bowman, F. Paesani, *J. Am. Chem. Soc.* **134**, 11116 (2012).
[25] H. Partridge, D. W. Schwenke, *J. Chem. Phys.* **106**, 4618 (1997).
[26] A. Shank, Y. Wang, A. Kaledin, B. Braams, J. Bowman, *J. Chem. Phys.* **130**, 144314 (2009).
[27] W. Chen, M. Gordon, *J. Phys. Chem.* **100**, 14316 (1996).
[28] E. Mas, R. Bukowski, K. Szalewicz, *J. Chem. Phys.* **118**, 4386 (2003).
[29] Y. Wang, X. Huang, B. Shepler, B. Braams, J. Bowman, *J. Chem. Phys.* **134**, 094509 (2011).
[30] C. Burnham, D. Anick, P. Mankoo, G. Reiter, *J. Chem. Phys.* **128**, 154519 (2008).
[31] T. Dunning, *J. Chem. Phys.* **90**, 1007 (1989).
[32] S. Boys, F. Bernardi, *Mol. Phys.* **19**, 553 (1970).
[33] A. Harvey, E. Lemmon, *J. Phys. Chem. Ref. Data.* **33**, 369 (2004).
[34] G. Kell, G. McLaurin, E. Whalley, *Proc. R. Soc. Lond. A* **425**, 49 (1989).
[35] B. Temelso, K. A. Archer, G. C. Shields, *J. Phys. Chem. A* **115**, 12034 (2011).
[36] D. Bates, G. Tschumper, *J. Phys. Chem. A* **113**, 3555 (2009).
[37] A. Soper, *Chem. Phys.* **258**, 121 (2000).
[38] A. Soper, C. Benmore, *Phys. Rev. Lett.* **101**, 065502 (2008).
[39] F. N. Keutsch, R. J. Saykally, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10533 (2001).
[40] K. Krynicki, C. D. Green, D. W. Sawyer, *Faraday Discuss.* **66**, 199 (1978).
[41] Y. L. A. Rezus, H. J. Bakker, *J. Chem. Phys.* **123**, 114502 (2005).
[42] J. Caldwell, L. Dang, P. Kollman, *J. Am. Chem. Soc* **112**, 9144 (1990).
[43] C. Tainter, P. Pieniazek, Y. Lin, J. Skinner, *J. Chem. Phys.* **134**, 184501 (2011).