Many-body force field models based solely on pairwise Coulomb screening do not simultaneously reproduce correct gas-phase and condensed-phase polarizability limits

Timothy J. Giese and Darrin M. York

Articles you may be interested in

Many-body effects are essential in a physically motivated CO2 force field
J. Chem. Phys. 136, 034503 (2012); 10.1063/1.3672810

Approximating quantum many-body intermolecular interactions in molecular clusters using classical polarizable force fields
J. Chem. Phys. 130, 164115 (2009); 10.1063/1.3121323

Comparing polarizable force fields to ab initio calculations reveals nonclassical effects in condensed phases
J. Chem. Phys. 122, 234107 (2005); 10.1063/1.1931653

Elucidating the role of many-body forces in liquid water. I. Simulations of water clusters on the VRT(ASP-W) potential surfaces
J. Chem. Phys. 120, 4777 (2004); 10.1063/1.1645777

A transferable polarizable electrostatic force field for molecular mechanics based on the chemical potential equalization principle
J. Chem. Phys. 117, 9175 (2002); 10.1063/1.1515773

Many-body force field models based solely on pairwise Coulomb screening do not simultaneously reproduce correct gas-phase and condensed-phase polarizability limits

Timothy J. Giese and Darrin M. York[a]
Department of Chemistry, University of Minnesota, Minneapolis, Minnesota 55415

It is demonstrated that many-body force field models based solely on pairwise Coulomb screening cannot simultaneously reproduce both gas-phase and condensed-phase polarizability limits. Several many-body force field model forms are tested and compared with basis set-corrected *ab initio* results for a series of bifurcated water chains. Models are parameterized to reproduce the *ab initio* polarizability of an isolated water molecule, and pairwise damping functions are set to reproduce the polarizability of a water dimer as a function of dimer separation. When these models are applied to extended water chains, the polarization is over-predicted, and this over-polarization increased as a function of the overlap of molecular orbitals as the chains are compressed. This suggests that polarizable models based solely on pairwise Coulomb screening have some limitations, and that coupling with non-classical many-body effects, in particular exchange terms, may be important. © *2004 American Institute of Physics.* [DOI: 10.1063/1.1756583]

The development of new-generation many-body force fields for biomolecular simulations[1] is an area of intense effort,[2–6] and offers considerable promise toward the improved modeling of biological systems.[7–9] Although these force fields represent a significant step forward, preliminary tests,[7,8,10] along with recent high-level *ab initio* quantum results,[11] suggest that models parameterized to obtain accurate gas-phase polarizabilities for individual molecules tend to be over-polarized in the condensed phase (e.g., in solution). The results presented here suggest that the *functional forms* used in these force fields are not able to simultaneously reproduce the gas-phase and condensed-phase polarizabilities due to the neglect of explicit coupling between polarization and *many-body* exchange effects. These effects might prove important for the design of transferable biomolecular force fields where molecules are exposed to drastically different environments ranging from hydrophobic pockets in proteins to direct coordination with divalent metal ions to polarized hydrogen bonding with charged amino or nucleic acids.

To overcome problems related to the polarizability at close separation, the most noteworthy being the so-called "polarization catastrophe,"[12] the Coulomb interactions between polarizable sites may be screened to model the diffuseness of an effective response density, as first suggested by Thole.[13] Modern-day polarizable force fields have adopted various screening methods, including the use of cutoff screening lengths,[14] cubic spline functions,[6] and other strategies.[7,13,15–17] All of these methods lead to improvement. However, none of them explicitly considers non-classical components of the many-body response,[3] such as polarization-exchange coupling. It is therefore instructive to explore the limits as to how well the forms of these Coulomb screening functions can reproduce the polarization response in different environments.

In this work, the polarizability of a series of water chains are examined with basis set-corrected *ab initio* methods and compared with values predicted by polarizable force field models with pairwise Coulomb screening. Models are developed such that the *ab initio* polarizability of an isolated water is reproduced, and one model is further developed, through adjustment of the Coulomb screening function, to reproduce the polarizability of a bifurcated water dimer as a function of internuclear separation. These models are demonstrated to be over-polarized when applied to extended chains, and this effect arises mainly from neglect of coupling between many-body exchange and polarization. The use of pairwise Coulomb screening functions leads to improvement, but is still unable to simultaneously reproduce the polarizability of both isolated water and the extended chain limit, especially for strongly interacting chains.

Polarizability of bifurcated water chains. The *ab initio* characterization of the polarizability of water as a function of number of waters in a cluster requires extensive statistical sampling, and tedious procedures to systematically remove basis set effects for different cluster sizes. For the purposes of this work, the complications of the cluster approach[11,18] are avoided by consideration of a series of linear bifurcated water chains (Fig. 1) that allow a consistent basis to be used, and afford more facile examination of trends in the change in polarization as a function of chain length and compression.

Density-functional (DFT) calculations of water chain polarizabilities were performed using the PBE0 functional[19] and the d-aug-cc-pVDZ basis.[20,21] The polarizability of water predicted at the PBE0/d-aug-cc-pVDZ level (Table I) is in close agreement with experiment (around 2% error), is well converged in terms of basis set cardinal index and augmen-

---
[a]Author to whom correspondence should be addressed. Electronic mail: york@chem.umn.edu
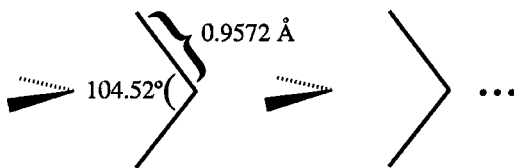
FIG. 1. Schematic of a bifurcated water chain. Bifurcated chains of length 2, 4, 6, 8, 10, and 12 were constructed from uniformly spaced monomers of fixed internal geometry ($r_{O-H} = 0.9572$ Å and $\theta_{H-O-H} = 104.52°$), with C_2 symmetry axis in the z direction. Chains were constructed at intermolecular O–O distances of 2.970 and 2.481 Å, corresponding to the equilibrium (R_e) and zero energy (σ) distances, respectively, on the binding curve obtained from one-dimensional minimization at the B3LYP/6-311++G($3df,2p$) level.

tation level (around 0.1% variation with respect to values calculated with aug-cc-pV5Z and q-aug-cc-pVDZ basis sets). This level of theory has been demonstrated to give reliable results in applications to the polarizability of water in clusters.[11] All calculations were performed with GAUSSIAN 03.[22]

The focus of the current work is to quantify the degree to which polarizability changes in going from an isolated system to an interacting limit (that of an infinitely long bifurcated chain). Special attention must be paid to the removal of basis set effects that could systematically alter the polarizability of the chain relative to that of a single water. Toward this end, polarizability calculations on bifurcated waters chains of length $N = 2, 4, 6, 8, 10,$ and 12 were performed in the basis of the 12 water system and corrected to be size consistent with the polarizability calculation of a single water. Only the component of the polarizability along the chain (z-direction) is considered (α_{zz}). The equation for the basis set-corrected chain polarizability is

$$\alpha_{zz}^{BSC}(N) = \alpha_{zz}(N;12) + \sum_{i=1}^{N} (\alpha_{zz}(1;1) - \alpha_{zz}(i;12)), \quad (1)$$

where $\alpha_{zz}^{BSC}(N)$ is the basis set-corrected zz polarizability component of the N-water chain, $\alpha_{zz}(1;1)$ is the zz component of a single water in the 1-water basis, and $\alpha_{zz}(i;12)$ is the zz component of the ith water in the N-water chain in the basis of the 12-water system. The N-water systems were placed at the center of the 12-water basis.

Polarizable force field models and pairwise Coulomb screening. The polarizable model for water used in this work

TABLE I. Polarizability of water (zz component).

Basis	α_{zz} (a.u.)
cc-pVDZ	5.345
aug-cc-pVDZ	9.055
d-aug-cc-pVDZ	**9.689**
t-aug-cc-pVDZ	9.677
q-aug-cc-pVDZ	9.678
aug-cc-pVTZ	9.485
d-aug-cc-pVTZ	9.723
aug-cc-pVQZ	9.629
aug-cc-pV5Z	9.683
Experiment[a]	*9.907*

[a]Experimental value taken from Ref. 24.

is a variation of the AMOEBA water model, described in detail by Ren and Ponder[7] and implemented in the TINKER molecular modeling package.[23] The focus of this work is polarization, therefore we limit our attention to the terms in the AMOEBA model that relate to the polarizability. From this perspective, the AMOEBA water model involves a set of three polarizable point dipoles located at the nuclei with screened Coulomb interactions modeled by a form of damping function first suggested by Thole.[13] The screened Coulomb interactions, described in detail elsewhere,[7] model the effect of a smeared charge density of the form

$$\rho = \frac{3a}{4\pi} \exp(-au^3), \quad (2)$$

where $u = R_{ij}/(\alpha_i \alpha_j)^{1/6}$, α_i and α_j are the polarizabilities of the dipoles in the interaction of sites i and j and a is a parameter (a constant).

It is not the purpose of the work to critique the AMOEBA water model for the purposes for which it was designed and performs very well[7]—but rather to investigate the limits of the functional form of the model itself in its ability to reproduce simultaneously the polarizability of noninteracting and interacting systems. Toward this end, a model is developed specifically to reproduce the PBE0/d-aug-cc-pVDZ polarizability component α_{zz} for a single water molecule through adjustment of the α_i and a parameters. This model, which screens only intramolecular interactions ($a \rightarrow \infty$ for intermolecular interactions), is designated POL-W1. A second model, designated POL-W2, was developed to further reproduce the polarizability values for the bifurcated water dimer as a function of internuclear separation through introduction of a distance-dependent form of the parameter $a(R_{ij})$ used for intermolecular interactions. The constraint that the bifurcated dimer polarizability be reproduced at each internuclear separation uniquely determines the parameter $a(R_{ij})$ for all R_{ij}. The resulting $a(R_{ij})$ could have been represented on a spline, however an excellent fit was obtained with the form

$$a(R_{ij}) = A \exp(-BR_{ij}^2) + C/R_{ij}. \quad (3)$$

A comparison of the original AMOEBA water model, POL-W1 and POL-W2 models with the basis set-corrected DFT results is shown in Fig. 2. The POL-W2 model agrees with the DFT results over the entire range of separations, whereas the POL-W1 model converges to the DFT results beyond 3.75 Å, but is observed to be considerably too polarized at shorter distances.

Comparison of predicted results for the average polarizability of water chains. Figure 3 compares the chain axis polarizability component per water (α_{zz}/N) calculated from the AMOEBA, POL-W1 and POL-W2 models with the basis set-corrected DFT values. The POL-W1 model, which reproduces the polarizability of a single water and does not screen intermolecular interactions, significantly over-predicts the polarizability per water in the infinite chain limit. The POL-W2 model, which additionally reproduces the α_{zz} polarizability of the dimer at each internuclear separation, closely reproduces the average polarizability of the water chains at the R_e separations. However, when the chains are
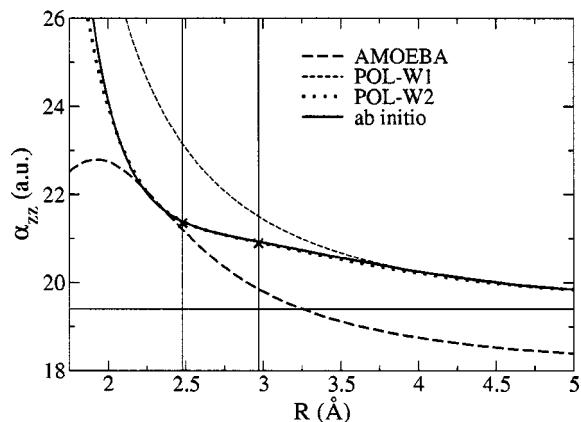
FIG. 2. Comparison of the *ab initio* and model polarizability component (α_{zz}) of a bifurcated water dimer as a function of separation. Shown also are the *ab initio* α_{zz} asymptotic limit (horizontal line), values calculated in the basis of the 12-water bifurcated chains (X marks), and R_e and σ distances (vertical lines). The parameters of POL-W2 (in a.u.) are $\alpha(H) = 3.333$, $\alpha(O) = 4.320$, $a = 0.079\,24$. The intermolecular screening parameters are: $A = 20.6438$, $B = 0.4372$, and $C = 0.4095$.

slightly compressed to their σ separation, the model considerably over-predicts the polarizability of the asymptotic infinite chain limit. The conclusion is that use of a pairwise Coulomb screening alone—here designed to reproduce the polarizability of the bifurcated water dimer over all distances—leads to significant improvement, but is not suf-
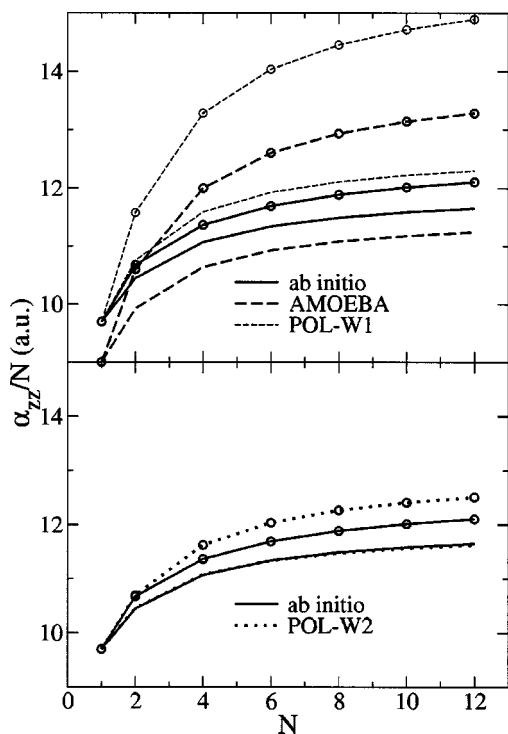

FIG. 3. Comparison of the average polarizability component per water (α_{zz}/N) as a function of chain length N. Unmarked lines and lines marked with circles correspond to chains at R_e and σ separations, respectively. The limiting α_{zz}/N values (a.u.) for (R_e, σ) chains are: *ab initio* (11.80,12.27), AMOEBA (11.42,13.62), POL-W1 (12.51,15.29), POL-W2 (11.77,12.72). σ and R_e are the distances defined by the potential energy surface of the bifurcated dimer which correspond to zero and minimum interaction energies, respectively.

ficient to obtain the polarizability in the infinite chain limit in the more extreme case of strongly interacting molecules.

These results are consistent with the observations of others[8] that suggest that polarizable models based on accurate quantum calculations with large basis sets (including diffuse functions) lead to over-polarization in the condensed phase. This heuristic reasoning has lead to parameterization strategies whereby quantum calculations are performed in the absence of diffuse augmentation functions that lead to isolated molecule polarizabilities that may be systematically too low, but that perform well in condensed phase simulations. The results presented here lend credence to the speculation[8,10] that the overlap of molecular orbitals and Pauli exclusion principle leads to condensed-phase molecular polarizability that are effectively reduced relative to that of isolated molecules. Moreover, the present work highlights a limitation of current polarizable force fields for which future improvement might be made: the inclusion of explicit coupling of polarization and many-body exchange. These effects are likely important for the design of highly accurate transferable models for biomolecular simulation where a wide range of vastly different chemical environments exist. The continued development and application of new-generation polarizable models[1] will ultimately bring these matters to bear, and serve to inspire novel advances in the design of improved quantum models for biomolecular simulations.

[1] T. A. Halgren and W. Damm, Curr. Opin. Struct. Biol. **11**, 236 (2001).
[2] S. W. Rick, S. J. Stuart, and B. J. Berne, J. Chem. Phys. **101**, 6141 (1994).
[3] D. York and W. Yang, J. Chem. Phys. **104**, 159 (1996).
[4] R. Chelli, P. Procacci, R. Righini, and S. Califano, J. Chem. Phys. **111**, 8569 (1999).
[5] C. Bret, M. J. Field, and L. Hemmingsen, Mol. Phys. **98**, 751 (2000).
[6] H. A. Stern, F. Rittner, B. J. Berne, and R. A. Friesner, J. Chem. Phys. **115**, 2237 (2001).
[7] P. Ren and J. W. Ponder, J. Phys. Chem. B **107**, 5933 (2003).
[8] G. A. Kaminski, H. A. Stern, B. J. Berne, and R. A. Friesner, J. Phys. Chem. A **108**, 621 (2004).
[9] N. Gresh, J. E. Šponer, N. Špačková, J. Leszczynski, and J. Šponer, J. Phys. Chem. B **107**, 8669 (2003).
[10] G. Lamoureux, A. D. MacKerell, Jr., and B. Roux, J. Chem. Phys. **119**, 5185 (2003).
[11] A. Morita, J. Comput. Chem. **23**, 1466 (2002).
[12] A. Stone, *The Theory of Intermolecular Forces* (Clarendon Press, Oxford, 1996).
[13] B. T. Thole, Chem. Phys. **59**, 341 (1981).
[14] H. A. Stern, G. A. Kaminski, J. L. Banks, R. Zhou, B. J. Berne, and R. A. Friesner, J. Phys. Chem. B **103**, 4730 (1999).
[15] P. van Duijnen and M. Swart, J. Phys. Chem. A **102**, 2399 (1998).
[16] M. Sprik and M. L. Klein, J. Chem. Phys. **89**, 7556 (1988).
[17] A. K. Rappe and W. A. Goddard III, J. Phys. Chem. **95**, 3358 (1991).
[18] A. Morita and S. Kato, J. Chem. Phys. **110**, 11987 (1999).
[19] C. Adamo and V. Barone, J. Chem. Phys. **110**, 6158 (1999).
[20] Basis sets were obtained from the Extensible Computational Chemistry

Environment Basis Set Database, Version 6/19/03, as developed and distributed by the Molecular Science Computing Facility, Environmental and Molecular Sciences Laboratory which is part of the Pacific Northwest Laboratory, P.O. Box 999, Richland, Washington 99352.

[21] D. E. Woon and T. H. Dunning, Jr., J. Chem. Phys. **100**, 2975 (1994).

[22] M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, Gaussian 03, Revision B.01, Gaussian, Inc., Pittsburgh, PA, 2003.

[23] J. W. Ponder, TINKER: Software Tools for Molecular Design, 4.1 ed., Washington University School of Medicine, Saint Louis, MO, 2004.

[24] W. F. Murphy, J. Phys. Chem. **67**, 5877 (1977).