UC Berkeley
UC Berkeley Previously Published Works

Title
Assessing many-body contributions to intermolecular interactions of the AMOEBA force field using energy decomposition analysis of electronic structure calculations

Permalink
https://escholarship.org/uc/item/2cd9w4dr

Journal
The Journal of Chemical Physics, 147(16)

ISSN
0021-9606

Authors
Demerdash, Omar
Mao, Yuezhi
Liu, Tianyi
et al.

Publication Date
2017-10-28

DOI
10.1063/1.4999905

Peer reviewed

Assessing many-body contributions to intermolecular interactions of the AMOEBA force field using energy decomposition analysis of electronic structure calculations

Omar Demerdash[1†3], Yuezhi Mao[1†3], Tianyi Liu[1], Martin Head-Gordon[1†], Teresa Head-Gordon[1,2,3,†*]

[1]*Departments of Chemistry,* [2]*Bioengineering,* [3]*Chemical and Biomolecular Engineering*

[†]*Kenneth S. Pitzer Center for Theoretical Chemistry*

University of California, Berkeley, CA 94720 USA

Abstract

In this work we evaluate the accuracy of the classical AMOEBA model for representing many-body interactions such as polarization, charge transfer, and Pauli repulsion and dispersion, through comparison against an energy decomposition method based on absolutely localized molecular orbitals (ALMO-EDA) for the water trimer and a variety of ion-water systems. When the 2- and 3-body contributions according to the many-body expansion are analyzed for the ion-water trimer systems examined here, the 3-body contributions to Pauli repulsion and dispersion are found to be negligible under ALMO-EDA, thereby supporting the validity of the pairwise-additive approximation in AMOEBA's 14-7 van der Waals term. However AMOEBA shows imperfect cancellation of errors for the missing effects of charge transfer and incorrectness in the distance dependence for polarization when compared with the corresponding ALMO-EDA terms. We trace the larger 2-body followed by 3-body polarization errors to the Thole damping scheme used in AMOEBA, and although the width parameter in Thole damping can be changed to improve agreement with the ALMO-EDA polarization for points about equilibrium, the correct profile of polarization as a function of intermolecular distance cannot be reproduced. The results suggest there is a need for re-examining the damping and polarization model used in the AMOEBA force field, and provide further insights into the formulations of polarizable force fields in general.

[3]Authors equally contributed
*corresponding author

1. INTRODUCTION

Classical molecular mechanics (MM) force fields are a computationally tractable approximation to the true quantum mechanical (QM) potential energy surface such that they afford the ability to calculate condensed-phase properties for large systems. Typically, force fields rely on a potential energy functional form of the atomic nuclear positions involving a sum of terms: a harmonic potential for covalent bonds and angles, a truncated Fourier series for describing torsional motions, and a model for non-covalent interactions that is often represented by pairwise-additive point-charge electrostatics and a Lennard-Jones potential (or a close variation thereof) for Pauli repulsion and dispersion. Recently, there has been significant growth in the development of force fields that take into account more advanced molecular interactions such as charge penetration[1-6], charge transfer[7-15], as well as the leading many-body effect of polarization explicitly[9, 11, 16-25]. In principle, these additions would improve the transferability of the force field by enhancing the ability to better describe heterogeneous molecular systems as well as to be able to model behavior across the phase diagram from gas to liquid to solid[22, 26-32]. Moreover, improvements in software implementations[33-38] have rendered advanced force fields more amenable to the types of computationally intensive applications that heretofore were only accessible by the simpler pairwise-additive force fields.

Nevertheless, while advanced force fields that include many-body interactions in principle afford better physics, their inherent accuracy is also inextricably linked to the choice of the underlying functional form of the relevant molecular interactions and the correctness of their parameterization. To illustrate, polarization of well-separated molecular species can be described rigorously as a power series expansion in terms of polarizabilities and higher-order hyperpolarizabilities[39, 40]. The formulation of polarization in the polarizable AMOEBA force field[19, 22, 41, 42] is based on this description, truncating at first order and formulating its polarization as atom-centered point inducible dipoles that respond linearly to the electrostatic field according to an isotropic polarizability parameter, α. While this model for polarizability is formulated to be accurate at long range, the simple description of distorted electron distributions as point inducible dipoles becomes inaccurate for small intermolecular separations, and therefore accounting for the finite nature of the distorted electron distribution becomes necessary for an accurate model. This short-ranged polarization description is typically achieved by a damping function, often exponential or Gaussian in form, which effectively "smears out" the point multipoles[43-45]. In addition to providing an improved model of the physics, the damping function has the practical mathematical effect of obviating the so-called polarization catastrophe, wherein atomic sites at short distance separation may polarize each other to infinity. In AMOEBA, damping is implemented using a

damping function originally developed by Thole[43] that assumes the following smeared charge distribution:

$$\rho = \frac{3a}{4\pi}\exp\left(-a\frac{r_{ij}^3}{\left(a_i a_j\right)^{\frac{1}{2}}}\right)$$ (1)

where r_{ij} is the distance between atomic sites i and j, α_i and α_j are their polarizabilities, and a is a dimensionless width parameter that effectively controls the strength of the damping[43]. It should be noted that this is strictly an empirical formula, and is one of many that Thole formulated in his original work.

Regarding the issue of parameterization, the overall accuracy of classical force fields rests on a delicate cancellation of errors among the non-bonded energetic terms that are accounted for explicitly – such as permanent electrostatics and van der Waals (vdW) interactions – and their ability to implicitly account for the missing terms such as charge penetration (formally two-body) and the many-body effects in terms such as charge transfer, Pauli repulsion, dispersion, and polarization if any of these are absent from the force field. Another assumption is that while the individual non-covalent terms in the potential are parameterized separately, they are finely tuned together in a final stage to reproduce properties that are a function of all the terms together, such as total QM energies of molecular clusters or condensed-phase properties such as densities and heats of vaporization[46]. For example, AMOEBA assumes isotropic, strictly atom-centered polarizabilities, and furthermore assumes that atomic polarizabilities for use in condensed-phase simulations may be adequately parameterized by a fit to gas-phase molecular polarizabilities. It is realized that such polarizabilities must be subsequently more finely tuned, as in the case of AMOEBA's water model, whose polarizability parameters, α (for O and H) and a, were further adjusted to better reproduce the QM energies of water clusters[22].

Quantum mechanical calculations are used in part or in total [9, 11, 23, 47-54] to parameterize force fields, as well as to appraise their quality. While the current practice of comparing total MM energies and forces to total QM energies and forces is helpful, it is desirable to be able to appraise the quality of individual non-covalent terms in the force field given its piece-wise nature. Energy decomposition analysis[55-57] (EDA) is a means to decompose the QM interaction energy into chemically sensible terms that in some instances have direct correspondence to the non-covalent terms of force fields, or can evaluate how physical effects not evaluated explicitly in dedicated force field terms are captured implicitly in other terms. Broadly speaking, EDA methods may be categorized into perturbative[58-60] or variational approaches[55, 61-68]. While the asymptotic behavior of the components is usually (but not

always) uniquely defined for a given level of wavefunction or DFT theory for either type of EDA, the decomposition of the total QM energy into chemically intuitive components becomes more difficult to uniquely define in the regime where the electron densities of the different chemical species overlap. Nonetheless, many well-designed EDA schemes are able to yield chemically sensible decompositions in the overlapping regime.

In a previous article[69], we assessed the accuracy of the non-covalent terms in AMOEBA for the water dimer and a series of water-ion dimers using a variational EDA approach based on absolutely localized molecular orbitals (ALMO) [67, 70] at the highly accurate ωB97X-V/def2-QZVPPD QM level of theory[71, 72]. The properties of the second generation ALMO-EDA approach that we find desirable include: (1) the energy decomposition relies on a series of intermediate wavefunctions that each obeys anti-symmetry; (2) the better handling of strongly interacting systems where the accuracy of perturbative approaches might be undermined by convergence issues; (3) the use of a high-quality density functional that adequately captures electron correlation and other effects for weakly interacting systems as benchmarked against the references obtained with high-level correlated methods (see Ref. [61] for the protocols of the latter); and (4) the ability to separate the induction term into polarization and charge transfer contributions, which is not naturally enabled in a perturbative approach such as symmetry-adapted perturbation theory (SAPT)[73, 74].

In our study, we found that the parameterization of the AMOEBA model indeed rested on a delicate cancellation of error among the permanent electrostatics, buffered 14-7 vdW potential[75], and many-body polarization to yield any agreement in the intermolecular energy of water and ion-water dimers as a function of distance and orientation with the ALMO-EDA result. We found universally that AMOEBA's multipolar permanent electrostatics were systematically too unfavorable, becoming most pronounced in the overlapping regime, and this was clearly attributable to a lack of charge penetration, an effect completely absent from the point multipole formalism of AMOEBA's permanent electrostatics. This missing effect, however, was partially but imperfectly recovered through AMOEBA's vdW potential in an implicit manner, entailing that there is some degree of softening of the repulsive wall. In addition, this softened vdW potential also accounts for charge transfer, a formally many-body non-covalent term, that has no explicit correspondence among the energy components of AMOEBA. Naively, one might expect the charge transfer to be incorporated implicitly into AMOEBA's polarization, as charge transfer and polarization may be viewed collectively under the more general term "induction", with charge transfer being considered an extreme case of polarization wherein the distorted electron density on one molecule is delocalized into another to which it is not covalently bound. However,

combining ALMO-EDA's charge transfer with its polarization term did not systematically improve agreement with AMOEBA's polarization throughout the entire range of intermolecular separations, especially in the compressed region where charge transfer is most pronounced. On the other hand, for almost all the investigated systems, adding the charge transfer term to the counterparts of the vdW term in ALMO-EDA (Pauli repulsion and dispersion) systematically improved the agreement with AMOEBA's 14-7 vdW potential throughout the range of intermolecular separations.

The agreement of AMOEBA's explicit accounting of polarization with that of ALMO-EDA was found to be system-dependent and of variable quality depending on the region of the distance and orientation scans examined. Near equilibrium, AMOEBA's polarization of water-halide complexes was too favorable compared with that of ALMO-EDA. However, this trend was reversed in the compressed region, where the Thole damping per Eq. (1) was too strong due to the larger polarizabilities of the halides. In the cases of the water-metal complexes, for the compressed and even for the equilibrium region in some cases, AMOEBA was found to be over-polarized due to the smaller polarizabilities of the metals and therefore weaker damping. Interestingly, while the asymptotic behavior of the polarization is reasonable for the water-monovalent metal complexes, the polarization is underestimated in the asymptotic region of the water-divalent ion complexes, which is noteworthy since divalent cations should provide the best "stress test" as the most strongly polarizing environment.

Clearly, the formulation of polarization in AMOEBA deserves further attention. In this work, we extend our analysis of the many-body, non-covalent interactions, represented solely by polarization in AMOEBA, and by polarization, charge transfer, Pauli repulsion, and dispersion under ALMO-EDA, for the water trimer and a series of water-halides and water-metal ionic trimer systems. To this end, we analyze the total energies and energetic components, and isolate separate 2- and 3-body contributions to the energetic components by means of the many-body expansion (MBE) [76-78]:

$$U_N \sim U_1 + U_2 + U_3 + \cdots \tag{2a}$$

where

$$U_1 = \sum_{i=1}^{N} U(i) \qquad U_2 = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} U(i,j) - U(i) - U(j) \tag{2b}$$

$$U_3 = \sum_{i=1}^{N-2} \sum_{j=i+1}^{N-1} \sum_{k=j+1}^{N} U(i,j,k) - U(i,j) - U(i,k) - U(j,k) + U(i) + U(j) + U(k) \tag{2c}$$

For the trimer systems examined in this work, Eq. (2) expanded to 3rd order (*N*=3) is exact; furthermore, for the AMOEBA model the non-covalent energy for the 1-body term is zero, so only 2- and 3-body contributions will be analyzed. In addition to our appraisal of polarization in AMOEBA, we assess the validity of the pairwise-additive approximation of van der Waals interactions as represented by the

buffered 14-7 potential for Pauli repulsion and dispersion, and how the 3-body contributions from charge-transfer are accounted for in the MM model. Building on our previous finding concerning the strong disagreement of polarization with ALMO-EDA in the compressed regime, we also assess whether we can tune the Thole damping parameters that gives maximal agreement with the corresponding ALMO-EDA term.

This article is organized as follows. In Section 2, we present the relevant details of the AMOEBA MM force field and the ALMO-EDA. In Section 3, we present the total energies and energy term breakdowns in the asymptotic, equilibrium, and compressed regions of the distance scans for the water trimer, as well as a series of water-ion trimer systems comprised of either two water molecules and one ion, or one water molecule and two ions. The implications of our findings are discussed in Section 4.

2. THEORY AND METHODS

AMOEBA force field. The details of the AMOEBA model have been described elsewhere[22, 41], although we briefly describe its salient features for non-covalent interactions here. The non-covalent terms of the AMOEBA force field include van der Waals (vdW) interaction, as well as permanent and induced electrostatics. For the vdW interaction, AMOEBA adopts a pairwise additive (strictly 2-body) buffered 14-7 originally proposed by Halgren[75], where the repulsive "14" term mimics Pauli repulsion, while the attractive "7" term in principle corresponds dispersion.

$$U_{\text{vdW}} = \sum_{i<j} \varepsilon_{ij} \left(\frac{1+\delta}{\sigma_{ij}+\delta} \right)^7 \left(\frac{1+\gamma}{\sigma_{ij}^7+\gamma} - 2 \right) \tag{3}$$

where ε_{ij} is the depth of the potential well, $\sigma_{ij} = r_{ij}/R_{ij}^0$ is the dimensionless distance between sites i and j (R_{ij}^0 refers to the minimum energy separation), and γ and δ are two constants whose values are set to 0.12 and 0.07, respectively.

The permanent electrostatics in AMOEBA is modeled by interactions between atom-centered point multipoles that generally includes monopoles q, dipoles μ, and quadrupoles Q:

$$\boldsymbol{M}_i^T = \left[q_i, \mu_{ix}, \mu_{iy}, \mu_{iz}, Q_{ixx}, Q_{ixy}, Q_{ixz}, Q_{iyy}, Q_{iyz} \right] \tag{4}$$

For the atomic ions studied in this work, the RHS of Eq. (4) contains monopoles only. The total permanent electrostatics contribution is then evaluated as a pairwise sum

$$U_{\text{elec}}^{\text{perm}} = \sum_{i<j} \boldsymbol{M}_i^T \boldsymbol{T}_{ij} \boldsymbol{M}_j \tag{5}$$

where $\boldsymbol{T}_{ij}$ is the multipole interaction tensor whose elements contain appropriate derivatives of $1/r_{ij}$ according to the multipole expansion. The set of permanent multipoles are derived from a QM electronic density using the distributed multipole analysis (DMA) due to Stone[79], whose values are further refined by a fit to an electrostatic potential generated by a higher level of QM theory[41]. The electrostatic energy in AMOEBA for covalently bound atoms is computed according to exclusion rules similar to those of other force fields, with 1-2 and 1-3 interactions set to zero; 1-4 interactions and those further separated are evaluated without any attenuation [41, 42].

As we have mentioned, permanent electrostatic interaction described by distributed point multipoles starts to break down in the overlapping regime due to the charge penetration effect. Two models previously explored by Rackers et $al.$[80] that are able to incorporate this effect in AMOEBA are considered in this work: one by Gordon and co-workers.[1, 3] (originally used in the EFP model) and another by Piquemal et $al.$[2] (used in the SIBFA model). In both models, the monopole charge is divided into a core nuclear charge and an electron portion that is smeared in a spherically symmetric fashion about the nucleus. Thus, the charge-charge interaction between two sites can be split into three contributions as follows:

$$U_{\text{elec}}^{\text{chg-chg}} = U^{\text{core-core}} + U^{\text{core-smeared}} + U^{\text{smeared-smeared}} \tag{6}$$

The two models share the common functional forms for the first two terms in the above equation:

$$U^{\text{core-core}}(r_{ij}) = \frac{Z_i Z_j}{r_{ij}} \tag{7a}$$

$$U^{\text{core-smeared}}(r_{ij}) = \frac{Z_i q_j}{r_{ij}} f_j^{\text{damp}}(r_{ij}) + \frac{Z_j q_i}{r_{ij}} f_i^{\text{damp}}(r_{ij}) \tag{7b}$$

where Z_i and q_i are the core and smeared charges at site i, respectively. $f_i^{\text{damp}}(r_{ij})$ is an exponential damping function that has a common form in these two models:

$$f_i^{\text{damp}}(r_{ij}) = 1 - \exp(-\beta_i r_{ij}) \tag{8}$$

In the model by Gordon and co-workers, the atomic number of each site is taken as the core charge, while in the Piquemal model the number of valence electrons is used instead. Besides that, the most pronounced difference between these two models rests in the description of interaction between two smeared charges: the former model uses the so-called "one-center" approach[81] to approximate this interaction:

$$U^{\text{smeared-smeared}} = \frac{q_i q_j}{r_{ij}} \left[1 - \frac{\beta_j^2}{(\beta_j^2 - \beta_i^2)} \exp(-\beta_i r_{ij}) - \frac{\beta_i^2}{(\beta_i^2 - \beta_j^2)} \exp(-\beta_j r_{ij}) \right] \tag{9}$$

whereas the latter adopts a different formula:

$$U^{\text{smeared}-\text{smeared}} = \frac{q_i q_j}{r_{ij}}\left[1 - \exp(-k\beta_i r_{ij})\right]\left[1 - \exp(-k\beta_j r_{ij})\right] \tag{10}$$

Here β_i is the same parameter as in Eq. (8), and k is a universal scaling parameter for all atom pairs. Note that when $\beta_i = \beta_j$, Eq. (9) becomes singular so a slightly modified form

$$U^{\text{smeared}-\text{smeared}} = \frac{q_i q_j}{r_{ij}}\left[1 - \exp(-\beta_i r_{ij})\left(1 + \frac{1}{2}\beta_i r_{ij}\right)\right] \tag{11}$$

is employed. Taken together, the Coulomb interaction between two smeared charges can be described by the following general form:

$$U^{\text{smeared}-\text{smeared}} = \frac{q_i q_j}{r_{ij}} f_{ij}^{\text{damp}}(r_{ij}) \tag{12}$$

where $f_{ij}^{\text{damp}}(r_{ij})$ is a two-site damping function that relies on both β_i and β_j (different than the one-site damping function $f_i^{\text{damp}}(r_{ij})$).

For permanent electrostatics, we only implement these charge penetration models for interactions between permanent monopoles (the "charge-charge damping" model in Ref. [71]. This is mostly for the sake of simplicity, as the accuracy of the strictly 2-body permanent electrostatics is *not* the focus of our present paper. The equations required for extending these models to higher-order multipoles can be derived by appropriately differentiating the one- and two-site damping functions, which were reported by Rackers *et al.*[80]

The polarization effect in AMOEBA is modeled by induced dipoles, $\boldsymbol{\mu}_i^{\text{ind}}$, placed on each atomic site i, whose magnitude is determined by a site-specific isotropic polarizability α_i and the total external electric field exerted:

$$\boldsymbol{\mu}_i^{\text{ind}} = \alpha_i(\boldsymbol{E}_i + \boldsymbol{E'}_i) \tag{13}$$

where $\boldsymbol{E}_i$ is the electric field due to permanent multipoles:

$$\boldsymbol{E}_i = \sum_j \boldsymbol{T}_{ij}^{\text{d}} \boldsymbol{M}_j \tag{14}$$

and $\boldsymbol{E}_i'$ is the field arising from induced dipoles on other sites:

$$\boldsymbol{E}_i' = \sum_{j \neq i} \boldsymbol{T}_{ij}^{\text{d}-\text{d}} \boldsymbol{\mu}_j^{\text{ind}} \tag{15}$$

Here T_{ij}^{d} refers to the blocks of the multipole interaction tensor (T_{ij}) that correspond to interaction between dipole and all other multipoles, and $T_{ij}^{\mathrm{d-d}}$ is specifically for dipole-dipole interactions. Once all μ_i^{ind}'s are self-consistently solved, the polarization energy is determined by

$$U_{\mathrm{elec}}^{\mathrm{ind}} = -\frac{1}{2}\sum_i \mu_i^{\mathrm{ind}} E_i \tag{16}$$

As already mentioned, a Thole-type damping scheme is employed in the original AMOEBA model assuming a smeared charge distribution described by Eq. (1). It is consistently applied to the evaluation of electric field due to permanent multipoles and induced dipoles (Eqs. (14) and (15)). Denoting $u = r_{ij}/(\alpha_i \alpha_j)^{1/6}$, the block of T_{ij} corresponding to the electric field arising from monopoles can be expressed as

$$T_{ij,\gamma}^{\mathrm{d-c}} = -[1 - \exp(-au^3)]\frac{r_{ij,\gamma}}{r_{ij}^3}, \qquad \gamma = x, y, z \tag{17}$$

and the Thole-damped electric field due to dipoles and quadrupoles can be derived by differentiating the RHS of Eq. (17) with respect to $\mathbf{r}_i$. The reader is referred to Ref. [22] where many of the specifics of AMOEBA's formulation of polarization were introduced.

Similar to other non-covalent terms such as permanent electrostatics, the calculation of the polarization energy for covalently bound atoms is also governed by exclusion rules that affect the calculation of the direct and mutual electrostatic fields. The direct polarization is defined by so-called polarization groups, whereby direct polarization can occur among atoms from different groups, but not among atoms within a group [41, 42]. Here, the only species with any covalent structure is the water molecule, and all its atoms are designated as a single polarization group. Mutual polarization, on the other hand, occurs among all atoms in the system, regardless of whether or not they are covalently bound [41, 42].

In this work, we also investigate the effect of charge penetration on AMOEBA's polarization. Similar to how the Thole-damped higher-order elements of T_{ij} can be derived by taking the appropriate derivatives of Eq. (17) according to the multipole expansion, so can the same procedure be followed using the charge penetration model for the damping. A key yet subtle distinction between the two damping models is that the lowest-order interaction in the Thole damping corresponds to the induced dipole-permanent monopole interaction, whereas the lowest-order interaction in the charge penetration model is the charge-charge interaction whose damped **T** matrix element is given by:

$$T_{ij,\gamma}^{c-c} = [1 - \exp(-\beta_j r_{ij})]\frac{1}{r_{ij}} \tag{18}$$

The analogue of Eq. (17) corresponding to the induced dipole-monopole interaction under charge penetration can be obtained by taking the first derivative of Eq. (18):

$$T_{ij,\gamma}^{d-c} = -\left\{1 - (1 + \beta_j r_{ij})\exp(-\beta_j r_{ij})\right\}\frac{r_{ij,\gamma}}{r_{ij}^3}, \qquad \gamma = x, y, z \tag{19}$$

In order to investigate the full effect of charge penetration, the smearing is also applied to permanent dipoles and quadrupoles, which is analogous to the application of Thole damping to higher-order multipoles[22, 80]. The damped $\boldsymbol{T}_{ij}$ elements corresponding to induced dipole-permanent dipole and induced dipole-permanent quadrupole are given, respectively, by Eq. (20a) and Eq. (20b) below:

$$T_{ij,\gamma\varepsilon}^{d-d} = -\left\{1 - (1 + \beta_j r_{ij})\exp(-\beta_j r_{ij})\right\}\frac{\delta_{\gamma\varepsilon}}{r_{ij}^3}$$

$$+ \left\{1 - \left(1 + \beta_j r_{ij} + \tfrac{1}{3}(\beta_j r_{ij})^2\right)\exp(-\beta_j r_{ij})\right\}\frac{3 r_{ij,\gamma} r_{ij,\varepsilon}}{r_{ij}^5} \qquad \gamma, \varepsilon = x, y, z \tag{20a}$$

$$T_{ij,\gamma\varepsilon\zeta}^{d-q} = \left\{1 - \left(1 + \beta_j r_{ij} + \tfrac{1}{3}(\beta_j r_{ij})^2\right)\exp(-\beta_j r_{ij})\right\}\frac{3(r_{ij,\gamma}\delta_{\varepsilon\zeta} + r_{ij,\varepsilon}\delta_{\zeta\gamma} + r_{ij,\zeta}\delta_{\gamma\varepsilon})}{r_{ij}^5}$$

$$- \left\{1 - \left(1 + \beta_j r_{ij} + \tfrac{2}{5}(\beta_j r_{ij})^2 + \tfrac{1}{15}(\beta_j r_{ij})^3\right)\exp(-\beta_j r_{ij})\right\}\frac{15 r_{ij,\gamma} r_{ij,\varepsilon} r_{ij,\zeta}}{r_{ij}^7} \quad \gamma, \varepsilon, \zeta = x, y, z \tag{20b}$$

Following the convention adopted in Ref. [22], we display for the sake of clarity in Eq. (21) below the damping coefficients $\lambda_1, \lambda_3, \lambda_5, \lambda_7$, corresponding to terms $\sim 1/r$, $\sim 1/r^3$, $\sim 1/r^5$, and $\sim 1/r^7$, respectively, from Eq. (18-20):

$$\lambda_1 = 1 - \exp(-\beta_j r_{ij}) \tag{21a}$$

$$\lambda_3 = 1 - (1 + \beta_j r_{ij})\exp(-\beta_j r_{ij}) \tag{21b}$$

$$\lambda_5 = 1 - \left(1 + \beta_j r_{ij} + \tfrac{1}{3}(\beta_j r_{ij})^2\right)\exp(-\beta_j r_{ij}) \tag{21c}$$

$$\lambda_7 = 1 - \left(1 + \beta_j r_{ij} + \tfrac{2}{5}(\beta_j r_{ij})^2 + \tfrac{1}{15}(\beta_j r_{ij})^3\right)\exp(-\beta_j r_{ij}) \tag{21d}$$

Note that once a charge penetration model is applied, the original Thole damping scheme is no longer applied for the electric field due to those smeared multipole moments (and their paired core charges), while it remains for the mutual induction between inducible dipoles (the evaluation of $\boldsymbol{E}_i'$).

Energy decomposition analysis. As in our prior work [69], a slightly modified version of the second-generation ALMO-EDA [67, 70] is utilized to separate the intermolecular interaction energy into

contributions from permanent electrostatics, Pauli repulsion, dispersion, polarization and charge transfer:

$$E_{int} = E_{elec} + E_{pauli} + E_{disp} + E_{pol} + E_{ct} \tag{22}$$

Here we briefly recapitulate the feature of this EDA approach. The first three terms on the RHS of Eq. (22) constitute the frozen interaction energy (E_{frz}), which corresponds to the energy change associated with bringing infinitely separated fragments into the complex geometry without relaxing their molecular orbitals (MOs):

$$E_{frz} = E[\mathbf{P}_{frz}] - \sum_A E_A[\mathbf{P}_A] \tag{23}$$

where $\mathbf{P}_{frz}$ is the one-particle density matrix (1PDM) associated with the antisymmetric wavefunction constructed from unrelaxed fragment MOs, and E_A's are converged SCF energies of fragments computed in isolation. In order to further decompose the frozen term, we first define the permanent electrostatic contribution as the classical Coulomb interaction between charge distributions of isolated fragments (including their nuclei and electrons):

$$E_{elec} = \sum_{A<B} \iint \rho_A^{tot}(\mathbf{r}_1) \frac{1}{\mathbf{r}_{12}} \rho_B^{tot}(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \tag{24}$$

The choice of this "classical" definition of permanent electrostatics (rather than its original definition in Ref. [73]) is because of its better physical correspondence to AMOEBA's permanent electrostatics based on distributed multipoles. The definition of the dispersion term, however, remains unchanged as in Ref. [73], which is defined as the difference between the exchange-correlation contribution to the frozen energy and the same quantity evaluated with a "dispersion-free" (DF) functional:

$$E_{disp} = \left(E_{xc}[\mathbf{P}_{frz}] - \sum_A E_{xc}[\widetilde{\mathbf{P}}_A] \right) - \left(E_{xc}^{DF}[\mathbf{P}_{frz}] - \sum_A E_{xc}^{DF}[\widetilde{\mathbf{P}}_A] \right) \tag{25}$$

where $\widetilde{\mathbf{P}}_A$'s are distorted fragment densities obtained from an orthogonal decomposition of $\mathbf{P}_{frz}$ (the reader is referred to Horn *et al.*[70] for more details on this procedure). Finally, the contribution from Pauli repulsion is then simply defined as the remainder of the frozen term:

$$E_{pauli} = E_{frz} - E_{elec} - E_{disp} \tag{26}$$

The separation of the polarization contribution utilizes the so-called "SCF for molecular interaction" (SCF-MI) approach [82-84], which variationally lowers the energy of the complex under the constraint that the MOs of each fragment are expanded only by the basis functions on its own fragment

(absolutely localized). Once the SCF-MI calculation is converged, the energy lowering due to the on-fragment relaxation of MOs is defined as the polarization term:

$$E_{\text{pol}} = E[\mathbf{P}_{\text{ALMO}}] - E[\mathbf{P}_{\text{frz}}] \qquad (27)$$

where $\mathbf{P}_{\text{ALMO}}$ is the 1PDM constructed from the polarized ALMOs. The last contribution, charge transfer, is then given by performing a fully relaxed SCF calculation:

$$E_{\text{ct}} = E[\mathbf{P}_{\text{SCF}}] - E[\mathbf{P}_{\text{ALMO}}] \qquad (28)$$

In the original ALMO-EDA scheme[67], fragment AO basis functions were utilized to enforce the absolute locality of the MOs. This approach lacks a well-defined basis set limit, as when large basis sets with diffuse functions are used (desired by modern density functionals for intermolecular binding energies), the boundary between intra- and intermolecular relaxations becomes obscure[74, 85]. In this work, the basis of each fragment is prepared by means of the recently developed approach based on fragment electrical response functions [85] (FERFs), which is then utilized to compute the polarization energy via a generalized SCF-MI scheme. The FERFs provide only the variational degrees of freedom that are pertinent to the response of fragment MOs to a uniform external electric field (dipolar response) and its field gradient (quadrupolar response). Unlike the original AO-based approach, the resulting polarization energy manifests a well-defined basis set limit, as established by numerical tests.

Computational details. Energy calculations using the AMOEBA force field are performed in the Tinker7 molecular modeling package[86]. The most recently released parameters are used for all species: "amoebapro13" is used for water-water, water-cation, and water-Cl⁻ interactions, while for water-F⁻ and water-Br⁻, parameters are obtained from "amoeba09", which corresponds to the latest parameterization for these two halides. Neither periodic boundary conditions nor distance cutoffs are adopted for any of these calculations; therefore, permanent electrostatics and polarization are performed in a standard, no-cutoff direct space interaction. Induced dipoles are self-consistently determined and the convergence criterion is set to 10^{-12} Debye.

Unless otherwise noted, the corrections for charge penetration are *not* employed in the AMOEBA calculations. The implementation details of the charge penetration models for permanent electrostatics and polarization have been introduced in Sec. 2, and the parameters required are taken from the work by Rackers *et al.*[80], which were derived by a fit to SAPT2+ electrostatics [87]. The SAPT2+ electrostatics corresponds to the interaction between charge distributions of isolated monomers without any effects of antisymmetry and, as such, are equivalent to the classical definition of permanent electrostatics in ALMO-EDA (Eq. (24)). The only difference between the electrostatics of these two energy decomposition methods rests in the treatment of intramolecular electron correlation (perturbation

theory vs. KS-DFT), while similar results should be produced by these two approaches. Note that in Ref. [71], the parameters for halogens were parameterized for covalently bonded environment but not for the halide anions. Nonetheless, these parameters are employed for F^-, Cl^- and Br^- in this work for a proof-of-concept purpose. Also, no charge penetration parameters were developed for metal cations in Ref. [71], so in the present work we only investigate the effects of charge penetration on pure water and water-halide clusters.

All the ALMO-EDA calculations are performed with the Q-Chem 4.4 software package [88]. The ωB97X-V functional[71], which is a range-separated hybrid GGA that incorporates the VV10 nonlocal correlation functional [89] for dispersion correction, is used together with the def2-QZVPPD basis set [72] to compute the intermolecular interaction energies (without counterpoise corrections since they are negligible in this basis). The XC functional is evaluated on a (99, 590) grid, while the SG-1 grid[90] is employed for the integration of the nonlocal correlation functional (VV10). In ALMO-EDA calculations, the "FERF-DQ" model [85] that incorporates dipolar and quadrupolar responses is employed for the evaluation of E_{pol}, which, in principle, provides eight polarization functions per occupied orbital. Also, in order to separate out the dispersion term, the Hartree-Fock theory is chosen to be the auxiliary "dispersion-free" functional paired with ωB97X-V, which is used to compute the latter two terms in Eq. (25).
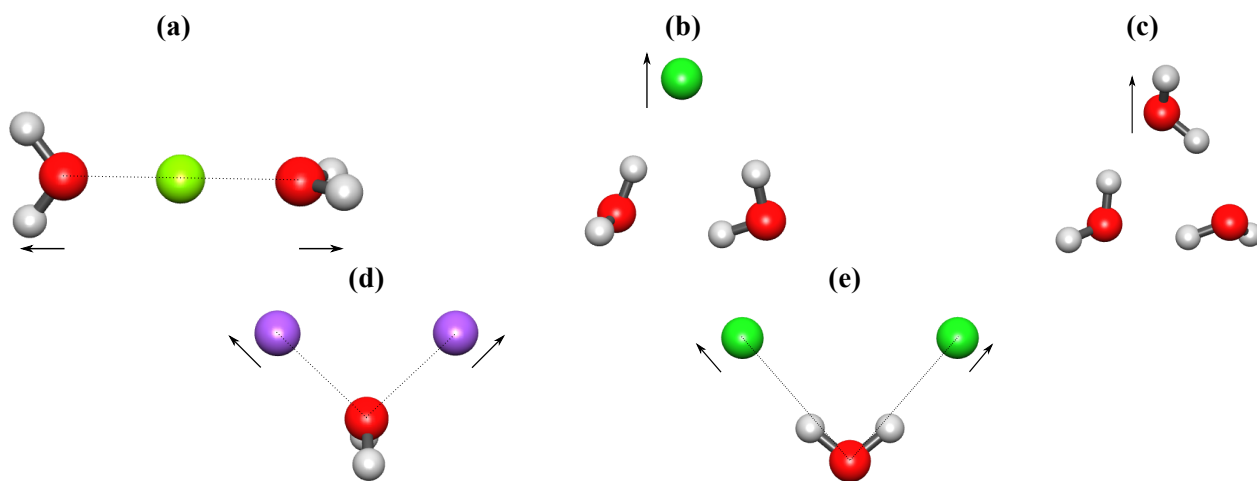

Figure 1. *Illustrations of the scans performed for the water trimer and ion-water trimers.* (a) $(H_2O)_2M^{n+}$, (b) $(H_2O)_2X^-$, (c) water trimer, (d) $H_2O(M^{n+})_2$, (e) $H_2O(X^-)_2$.

Definition of distance scans. The distance scans performed for the study of 2- and 3-body interactions are illustrated in Figure 1, in which the monomer geometries all remain unrelaxed. For the water trimer and systems containing two water molecules and one ion, we choose their equilibrium geometries optimized at the ωB97X-V/def2-QZVPPD level of theory as the starting points of the scans

(which are marked as the origins in Figure 1). The distance coordinate corresponds to displacement from equilibrium from the reference geometries shown here, in the directions indicated by arrows.

The $M^{n+}(H_2O)_2$ systems are of C_{2v} point group symmetry, and two O-M^{n+} distances are altered simultaneously along the 2-fold axis. For the $X^-(H_2O)_2$ systems (or the water trimer), we shift X^- (or one of the water molecules) along the distance from the centroid (not mass-weighted) of the triangle formed by the three heavy atoms at the equilibrium structure to X^- (or to the oxygen atom of the displaced water). The structures for the $H_2O(M^{n+})_2$ and $H_2O(X^-)_2$ systems are prepared differently, since no meaningful equilibrium geometries can be obtained due to the strong repulsion between two like ions. For systems containing two cations, the angles between the O-M^{n+} distances and the bisector of the H_2O molecule are constrained to be $45°$, and the initial O-M^{n+} distances are chosen to be the same as that in the optimized structure of the corresponding $M^{n+}(H_2O)$ dimer. For systems containing two anions, the initial structure for the scan is simply prepared by a reflection of the O-H--X^- angle under the equilibrium geometry of the $X^-(H_2O)$ dimer so that the supersystem has C_{2v} symmetry, and two O--X^- distances are then modified simultaneously.

3. RESULTS

The general positive features and documented problems of the AMOEBA potential revealed in our prior study based on the water dimer and an extensive series of ion-water dimers persist in all of the trimer systems examined here, which is not surprising given the dominant effect of 2-body compared to 3-body contributions to the MBE. Figure 2a shows that for the water trimer the total intermolecular energy curve is underbound throughout the entire range of distances, which is primarily due to the imperfect cancellation of errors between the dominant 2-body terms: the insufficiently favorable permanent electrostatics that arise from use of point multipoles, and the 14-7 van der Waals, whose excessively soft repulsive wall only partially compensates. Similar to the case of the water trimer, the water-cation $(H_2O)_2M^{n+}$ and $H_2O(M^{n+})_2$ systems (M^{n+}=Li^+, Na^+, K^+, Ca^{2+} and Mg^{2+}; Figure S1) are also underbound throughout the distance scan (shown specifically for $(H_2O)_2Mg^{2+}$ in Figure 2b). This again is due to a lack of charge penetration as observed for the water trimer, but the onset of charge penetration error occurs at shorter range for the cation-water systems due to the more compact electron density distributions of the cations than that of the water molecule. As was observed in our study of the dimer, all of the water-halide trimer systems $(H_2O)_2X^-$ and $H_2O(X^-)_2$ systems (X^-= F^-, Cl^-, and Br^-) exhibit a distorted shape and shifted minima in their total intermolecular AMOEBA energy profiles compared

15

with the ωB97X-V benchmark; this is shown specifically for $(H_2O)_2F^-$ in Figure 2c, which serves to illustrate the same general trends we observed for all halide-water systems (Figure S1).
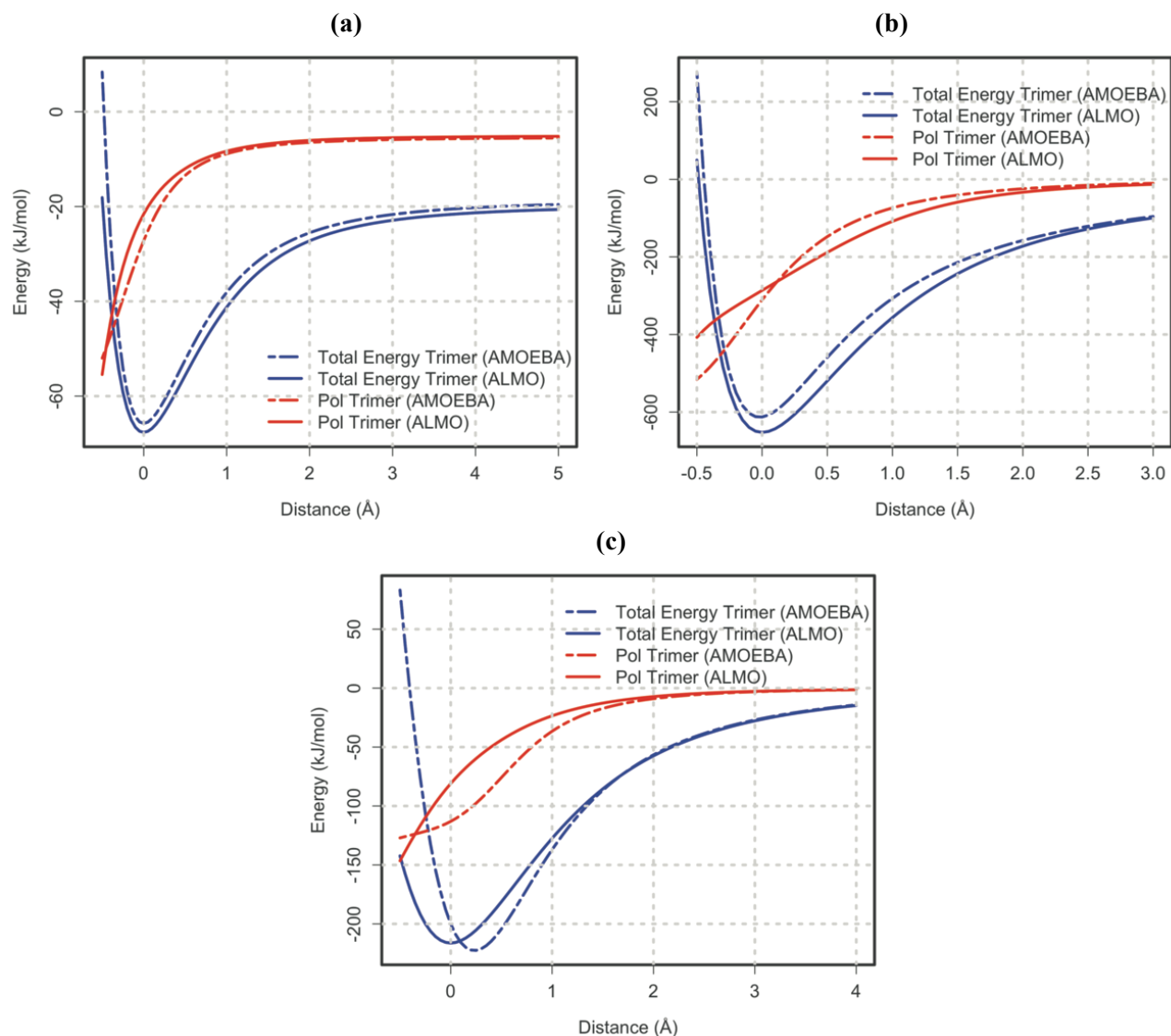
(a)

(b)

(c)

Figure 2. *Comparison of the total intermolecular energy profile for AMOEBA and the ωB97X-V DFT benchmark and AMOEBA polarization against the ALMO-EDA polarization. (a) water trimer, (b)* $(H_2O)_2Mg^{2+}$, *and (c)* $(H_2O)_2F^-$. These three representative examples illustrate the trends observed for the water trimer and the replacement of one or two waters for the alkali, alkali earth metals, or halide ions. In this and all subsequent plots, the distance coordinate corresponds to displacement from equilibrium from the reference geometries.

The primary new feature of this study is how AMOEBA reproduces the 2- as well as 3-body energies in the distance scans for the genuinely many-body terms of ALMO, which are the Pauli repulsion, dispersion, polarization, and charge transfer contributions. Since AMOEBA's only many-body term arises from Thole-damped polarization, the question we address here is how successfully it renders agreement with the corresponding ALMO-EDA polarization. While Figure 2a-2c and S1 show

that the total polarization for AMOEBA ranges from quite good agreement for the water trimer, to quite poor agreement for the ion-water systems, we examine whether the 3-body sum of ALMO's Pauli repulsion, dispersion and charge transfer terms are thus being captured by 3-body polarization or whether it is spread "incoherently" across, for example, the 2- and 3-body polarization contributions that might explain the discrepancies in the total polarization profile.

At the start we can state the 3-body sum of ALMO's Pauli repulsion and dispersion terms is negligible at equilibrium, and only becomes significant in the compressed region (illustrated in Figure 3 for the water trimer and the $(H_2O)_2Cl^-$ trimer). The trends of the 3-body contributions to the Pauli and dispersion terms are opposite to those of their 2-body counterparts, with an attractive 3-body Pauli term and a repulsive 3-body dispersion term that is smaller in magnitude, which together combine to form an attractive (cooperative) term. However, even in the compressed region, the sum of the 3-body Pauli and dispersion amounts to no more than 0.5-1.0% of the total Pauli repulsion and dispersion contribution across all trimer systems examined here, and therefore we do not consider these negligible 3-body terms further. Therefore, the trimer systems examined in this study afford the opportunity to ascertain the success of the AMOEBA potential when it is broken down into 2- and 3-body polarization contributions that can be compared against the same 2- and 3-body polarization terms for the ALMO-EDA profiles, and to assess the quite sizeable 3-body charge transfer, both of which can show cooperative and anti-cooperative behavior.
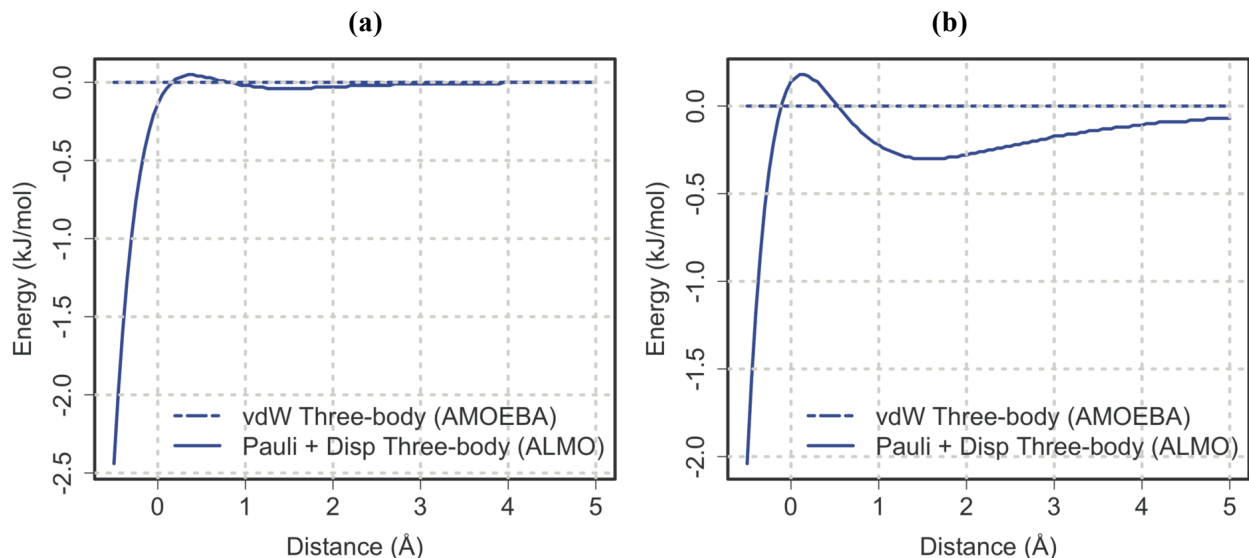

Figure 3. *The sum of the ALMO's 3-body contribution to the vdW interaction (Pauli repulsion + dispersion)* for the (a) water trimer and (b) $(H_2O)_2Cl^-$ trimer. AMOEBA's buffered 14-7 van der Waals potential is strictly pairwise additive and contributes no 3-body contribution.

In order to proceed with the many-body analysis of agreement between AMOEBA and ALMO-EDA, we return to one of the dominant errors in the 2-body interactions, namely the lack of charge penetration for the permanent electrostatics. This is because the many-body polarization term (Eq. (16)) is formulated from the dipole response to the electric field arising from point permanent multipoles (Eq. (14)) – which is known to be in error– as well as the field arising from other inducible dipoles (Eq. (15)). Therefore, in this work we have explored two charge penetration models (introduced in Section 2) that better reproduce ALMO-EDA's permanent electrostatics, thereby better isolating remaining differences (if any) to just the polarization model. Compared to the previous study by Rackers *et al.*[80], what differs here is that we evaluate the improvement in the AMOEBA electrostatics against ALMO-EDA as opposed to their comparison to SAPT2+[87], and we also explicitly consider the effect of these charge penetration models on AMOEBA's polarization as well, as the electric field due to permanent multipoles is altered once they are smeared.

Figures 4 and S2 show unambiguously that either monopole charge penetration model alters AMOEBA's permanent electrostatics to dramatically improve agreement with the ALMO-EDA result, with the improvement most noticeable for atoms with larger electron clouds. Even so, using the supplied parameters for the monopole-only charge penetration models derived by other groups, the Piquemal model is quantitatively better across the halide systems that have been examined (Figure 4 and S2). Since the optimization of those parameters was based on fits to SAPT2+, it indirectly implies that the permanent electrostatic component given by ALMO-EDA and SAPT2+ decompositions are in very good accord. What is even more striking is the dramatic improvement achieved in the $H_2O(Cl^-)_2$ system, even though the parameters for covalently bound (neutral) Cl atom is utilized, suggesting that the charge penetration parameters are relatively robust to changes in atomic radii or chemical microenvironment at least in this specific case. Nonetheless, we do not make strong distinction between the Piquemal and the Slipchenko-Gordon charge penetration models for their performances on these ion-water systems, since (1) the overall functional forms are very similar, (2) the possibility that the charge penetration model extended to dipoles and quadrupoles may change the quantitative conclusion, (3) parameters for covalently bonded halogen atoms are used here, and (4) the latter model could rather easily be reparameterized to improve agreement with the ALMO-EDA result if desired.

For the assessment of permanent electrostatic energies, the charge penetration corrections are only applied to the permanent monopoles. However, as mentioned in Section 2, in our investigation of the effect of charge penetration on polarization, smeared permanent dipoles and quadrupoles are also employed to compute their contributions to the electric field. Note that in the calculation of the

permanent field contribution to the induced dipoles (Eq. (14)) and the calculation of the polarization energy (Eq. (16)), the original Thole-damped terms are replaced by Eqs. (19-20), as the charge penetration and Thole damping models assume different smeared charge distributions, and it would be unreasonable to include both damping models for the permanent field terms. However, we find that the Thole model must be retained for the damping of the mutual polarization between induced dipoles (Eq. (15)), as use of the damping function from charge penetration models instead results in a failure for the induced dipoles to converge properly.
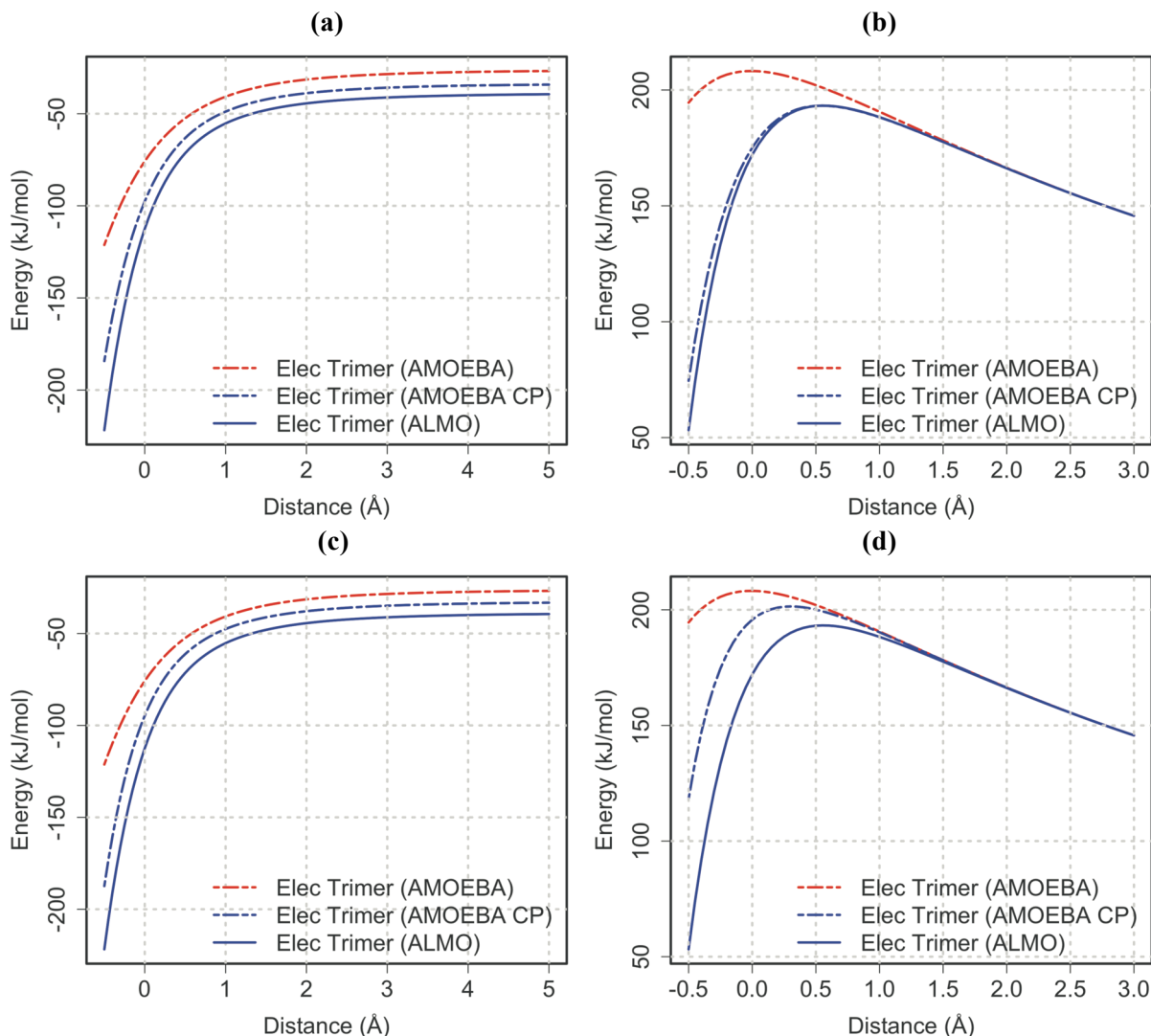
Figure 4. *Comparison of no charge penetration and with charge penetration (monopole only) for AMOEBA permanent electrostatics against the ALMO-EDA classical electrostatics using $(H_2O)_3$ (a and c) and $H_2O(Cl^-)_2$ (b and d) as examples.* a and b: charge penetration model from Piquemal[2] and (c and d) charge penetration model due to Gordon and Slipchenko[1, 3, 45] with parameters derived for both models by Rackers *et al.*[80]

Figures 5 and S3 show that the incorporation of charge penetration for the higher order multipoles results in reasonable behavior in the long range, but overpolarization in the compressed region that can be quite drastic (as shown in Figure 5a and Figure S3(a-c)). This is probably for the trivial reason that the AMOEBA force field is now imbalanced with the introduction of the new physics that essentially assumes a wholly different charge distribution than the one assumed by Thole damping, and a total reparameterization would have to be undertaken to account for the different charge distribution. Furthermore, it is known that separating the nuclear and electron charges and smearing only the latter makes the issue of the missing exchange effect in polarization more pronounced,[91] which could otherwise curb the overpolarization observed in the short range. The main conclusion from the charge penetration analysis is that we can proceed with the analysis of the AMOEBA polarization against the ALMO-EDA polarization using the original Thole prescription for charge distribution, since the error in the permanent electrostatic field arising from point multipoles does not dominate the polarization errors.
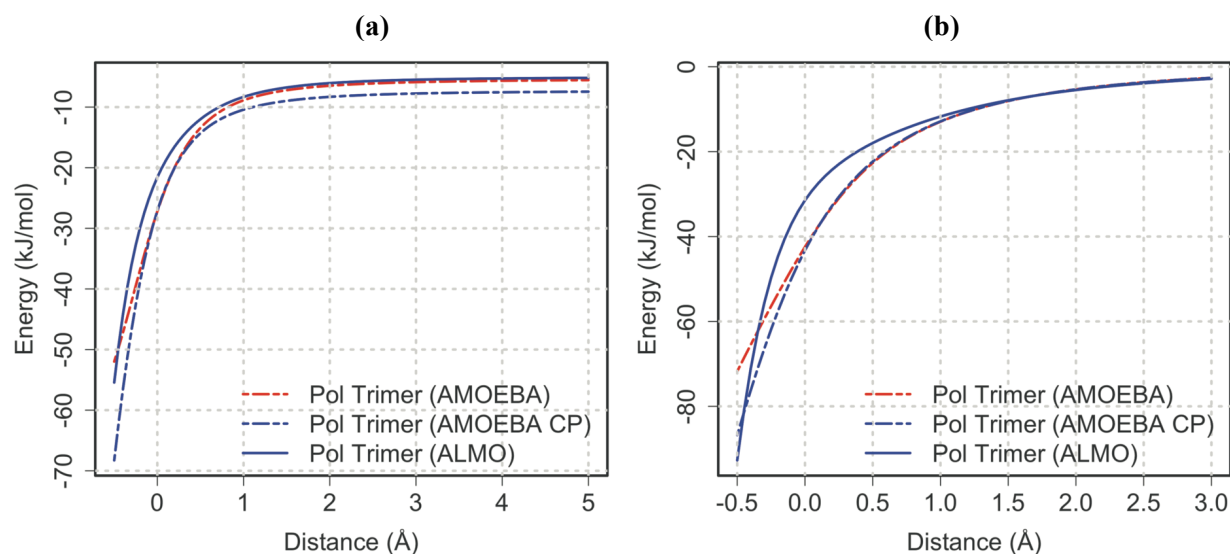

Figure 5. *Comparison of AMOEBA's polarization energies evaluated with permanent field arising from point multipoles and multipoles with charge penetration against ALMO-EDA's polarization term for (a) the water trimer and (b) $H_2O(Cl^-)_2$ as an example.* Results are shown here for the Piquemal model and parameters from Rackers *et al.* [80].

The sole many-body effect in AMOEBA, the polarization, shows small quantitative differences to large qualitative disagreement with the ALMO-EDA profiles in the equilibrium and compressed regimes for all trimer systems studied. Figure 6a shows that the 2-body polarization for the water trimer is well reproduced by the AMOEBA potential in the asymptotic region and near equilibrium, but then shows an inflection point that makes it unfavorable in the compressed region – a short-range effect arising from Thole damping. In contrast, AMOEBA's 3-body polarization energy is seen to be too

negative throughout most of the intermolecular distances examined, and especially in the compressed region and near equilibrium distances as seen in Figure 6b. Interestingly, the agreement between 3-body effects for the water trimer improves significantly if we instead compare AMOEBA's 3-body polarization against the sum of ALMO's 3-body polarization and 3-body charge transfer, which mostly (but not completely) corrects for the 3-body differences.
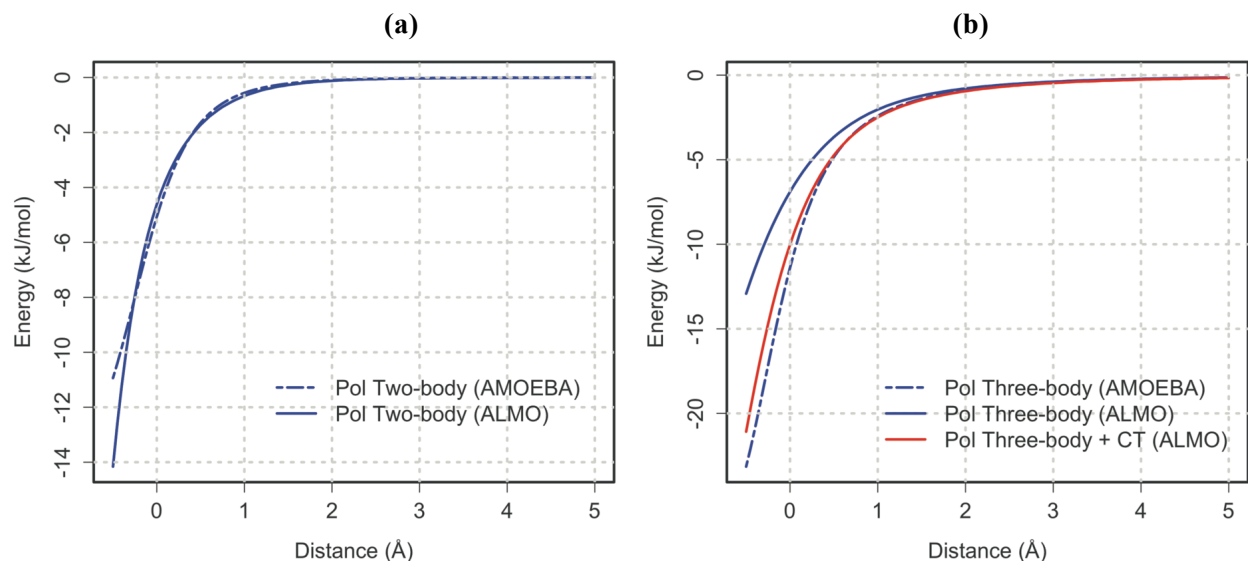

Figure 6. *Comparison of AMOEBA 2-body and 3-body polarization against the ALMO-`EDA polarization for the water trimer.* (a) The 2-body polarization energy for one of the three pairs in the trimer. (b) The 3-body polarization as well as the sum of ALMO's 3-body polarization and charge transfer terms.

Unlike the water trimer system, where both 2- and 3-body polarization energies are favorable, we find in the $(H_2O)_2M^{n+}$ systems that the 2- and 3-body polarization terms are favorable (Figure 7a) and unfavorable (Figure 7b), respectively, as shown for the representative case of $(H_2O)_2Mg^{2+}$. In general, both the 2- and 3-body polarization energies of $(H_2O)_2M^{n+}$ are underdamped at short range; this is partly attributable to the functional dependency of Thole damping on the polarizability, which is smaller for both cations relative to water. Although the damping parameter has been adjusted to much smaller values for Mg^{2+} (a =0.0952)[28] relative to that of water and the monovalent cations and anions (a=0.39), corresponding to a stronger damping effect according to Eq. (17), it is not enough to correct for the underdamping in the short range. While the effect of charge transfer is negligible for the $(H_2O)_2Mg^{2+}$ system, it is more significant for the $(H_2O)_2Na^+$ system (Figure S4c and S4d), and we see that AMOEBA's 3-body polarization captures 3-body charge transfer at best in the compressed region but not at longer distances, unlike the case of the water trimer.
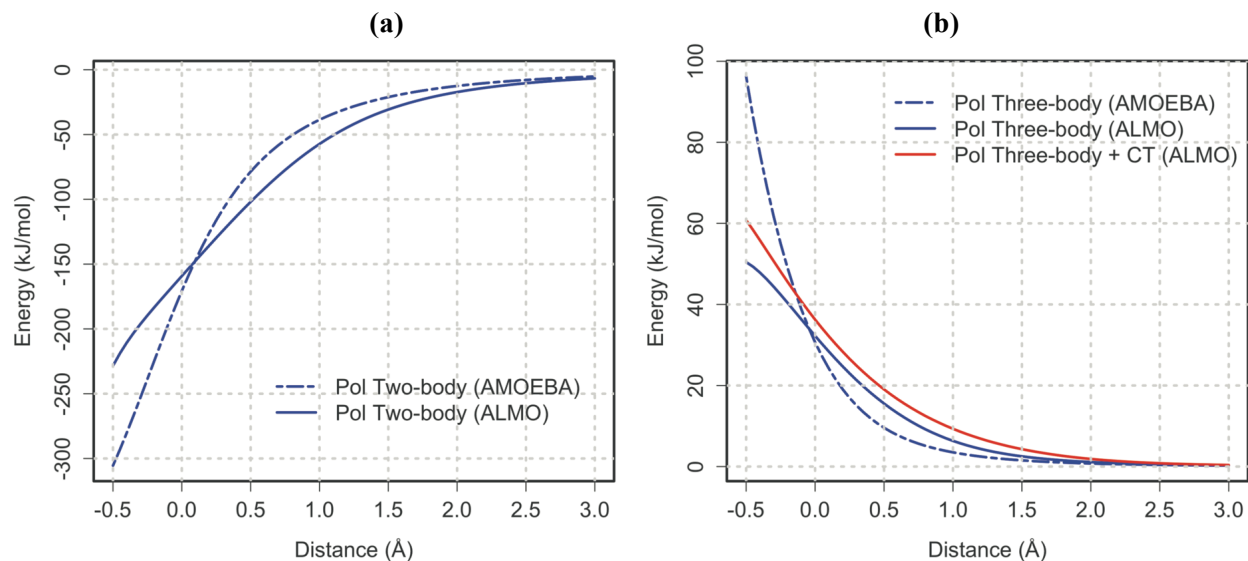
Figure 7. *Comparison of AMOEBA 2-body and 3-body polarization against the ALMO-EDA polarization for the $(H_2O)_2Mg^{2+}$ system.* (a) The 2-body polarization energy for one of the ion-water pairs in the trimer. (b) The 3-body polarization as well as the sum of ALMO's 3-body polarization and charge transfer terms.

The water dimer-halide $(H_2O)_2X^-$ (X=F, Cl, and Br) systems also deserve to be considered together due to overall similarities in their total intermolecular energies and energy breakdowns upon the distance scan, since all exhibit a distorted shape and shifted minima in their total intermolecular AMOEBA energy profiles compared with the ωB97X-V benchmark (Figure 2 and S1). While this poor overall agreement arises in part from inadequate cancellation of error in the strictly 2-body terms as discussed previously, it also arises from errors in the net polarization, with defects in the anti-cooperative 3-body polarization as large as the leading order cooperative 2-body term as illustrated for $(H_2O)_2F^-$ (Figure 8). Both 2- and 3-body polarization exhibit inflection points in the compressed region due to the onset of Thole damping. As the halides are of higher polarizabilities, these systems are subject to stronger damping effect, but again, polarization is overestimated relative to the ALMO-EDA result at almost all investigated ranges. AMOEBA's 3-body polarization energy is in better agreement with the sum of ALMO-EDA's 3-body polarization plus charge transfer than with polarization solely, and this is notable particularly for $(H_2O)_2F^-$. Nonetheless, this correspondence between AMOEBA's 3-body polarization and ALMO's 3-body polarization and charge transfer is not as clear-cut as for the water trimer, as AMOEBA's 3-body polarization for the halide systems is still too anti-cooperative (Figure 8 and S4).
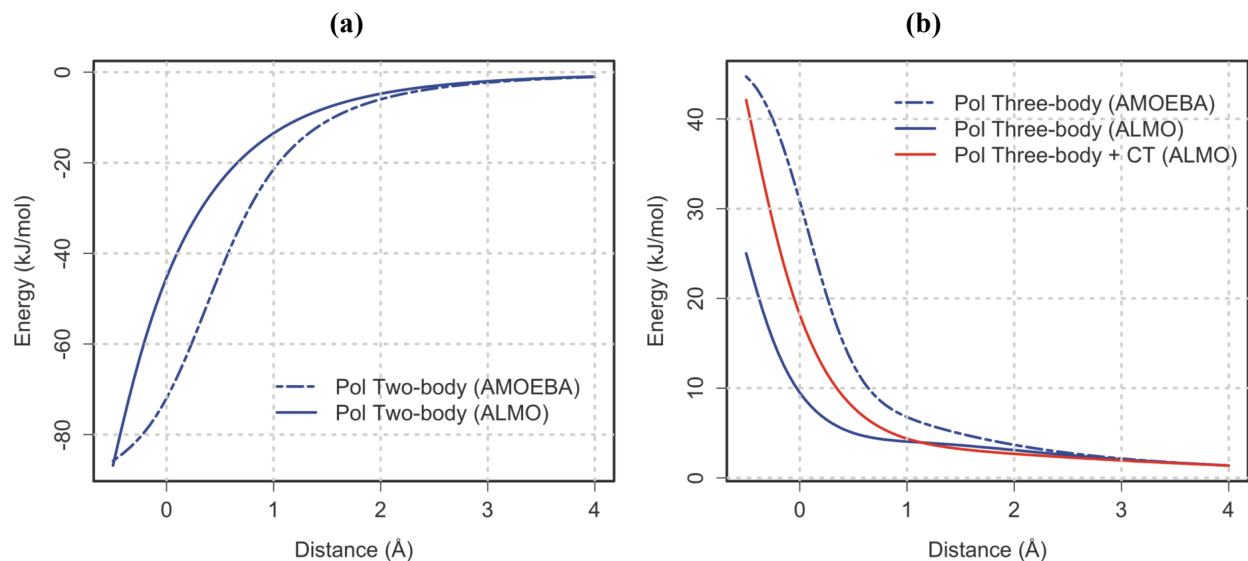
Figure 8. *Comparison of AMOEBA 2-body and 3-body polarization against the ALMO-EDA polarization for the $(H_2O)_2F^-$ system.* (a) The 2-body polarization energy for one of the ion-water pairs in the trimer. (b) The 3-body polarization as well as the sum of ALMO's 3-body polarization and charge transfer terms for the system.

To assess the robustness of AMOEBA under a more strongly polarizing environment, we also performed scans consisting of two ions and a single water molecule (Figure 9, S5, and S6); to a large degree, the profiles in these scans only serve to amplify any defects observed in the case consisting of a single ion with two waters, but there are notable differences among some of the ion-water profiles that are worth mentioning. On a positive note, we see that the 2-body ion-ion polarization energies are in much better agreement with those of ALMO compared with the 2-body ion-water terms (Figure 9(a)-(c), Figure S5, and Figure S6). However, the overall trend of AMOEBA's 3-body polarization for $H_2O(Na^+)_2$ is incorrect just beyond equilibrium, lacking the anti-cooperative behavior of that given by ALMO-EDA (Figure 9d). Similar to $H_2O(Na^+)_2$, AMOEBA's 3-body polarization for $H_2O(Mg^{2+})_2$ (Figure S5i) also displays an incorrect cooperative trend just beyond equilibrium compared with the anti-cooperative result from ALMO-EDA. $H_2O(Li^+)_2$ (Figure S5c) and $H_2O(Ca^{2+})_2$ (Figure 9e) display even more striking disagreement in their 3-body polarization, as they completely fail to show the correct anti-cooperative behavior of ALMO's corresponding 3-body polarization, which for these two cations is anti-cooperative throughout the distance range. Surprisingly, this qualitatively incorrect behavior for the 3-body polarization is not borne out for $H_2O(K^+)_2$, where the correct anti-cooperative behavior is observed. The 3-body polarization terms of the $H_2O(X^-)_2$ systems also display the correct anti-cooperative behavior (Figure 9f, S5l, and S5o).
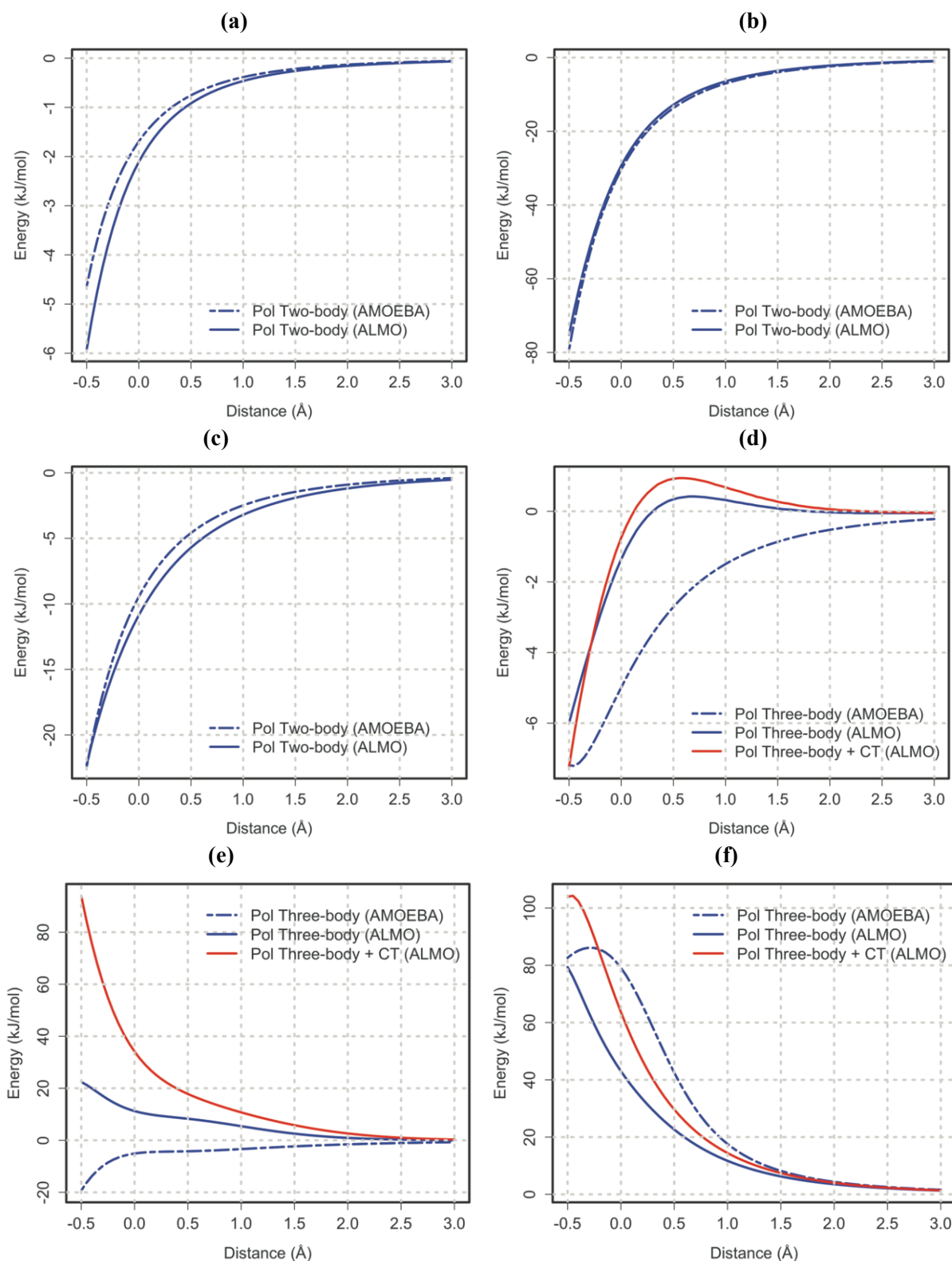
Figure 9. *Comparison of AMOEBA 2-body and 3-body polarization against the ALMO-EDA polarization.* (a)-(c): The 2-body polarization energy for the ion-ion pair in the trimer and (d)-(f): The 3-body polarization as well as the sum of ALMO's 3-body polarization and charge transfer terms for the systems, for the $H_2O(Na^+)_2$, $H_2O(Ca^{2+})_2$, and $H_2O(F^-)_2$ systems.

Given the functional form for Thole damping in Eq. (1), we sought to determine whether we could tune the atomic polarizabilities (α_i and α_j) and/or the Thole damping parameter a to improve the agreement with ALMO-EDA's polarization. AMOEBA assumes isotropic, strictly atom-centered polarization, and furthermore assumes that atomic polarizabilities for use in condensed-phase simulations may be adequately parameterized by a fit to gas-phase molecular polarizabilities. The appropriateness of this is controversial, as there are strong reasons to believe that the effective polarizabilities of atoms will be different in the condensed phase as compared with their values in the gas phase. In fact, the observed AMOEBA polarization profiles compared with those of ALMO-EDA suggest a general trend that halide polarizabilities are too large and those of metal cations are too small.
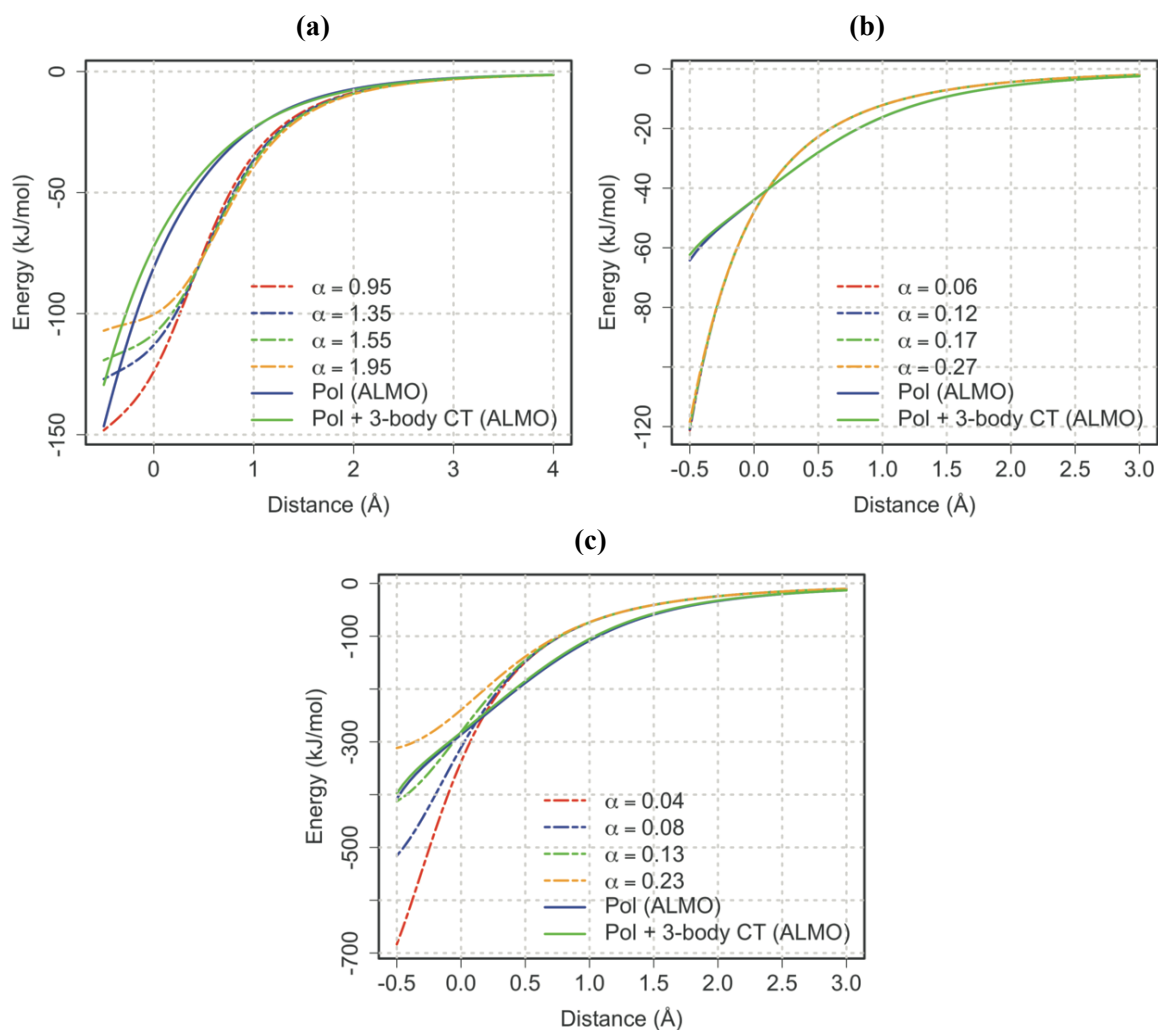

Figure 10. The effect of modifying the ion polarizability for water-ion interactions in systems representative of the types of interacting species: a.) $(H_2O)_2F^-$. b.) $(H_2O)_2Na^+$. c.) $(H_2O)_2Mg^{2+}$. In all cases, AMOEBA's default ion polarizability is represented by the blue dashed curve curve.

We first examined the water trimer and confirmed that AMOEBA's default polarizabilities yielded the best agreement between AMOEBA's total polarization energy and the sum of ALMO-EDA's total

polarization and 3-body charge transfer energy (Figure S7a). Keeping the polarizabilities of the water's atoms fixed at their default values, we then attempted to tune the ion polarizabilities. Interestingly, the improvements in polarization near equilibrium afforded by tuning the ion polarizabilities are negligible for all ion-water trimers, with any apparent effect only in the compressed region due to the dependency of the damping on the polarizability (Figure 10 and S7).

We then turn to the dimensionless width parameter a. We first assessed the Thole damping parameter of the water trimer and found that the default value of 0.39 in AMOEBA did in fact provide the best agreement with the sum of ALMO's total polarization and 3-body charge transfer (Figure S8a).
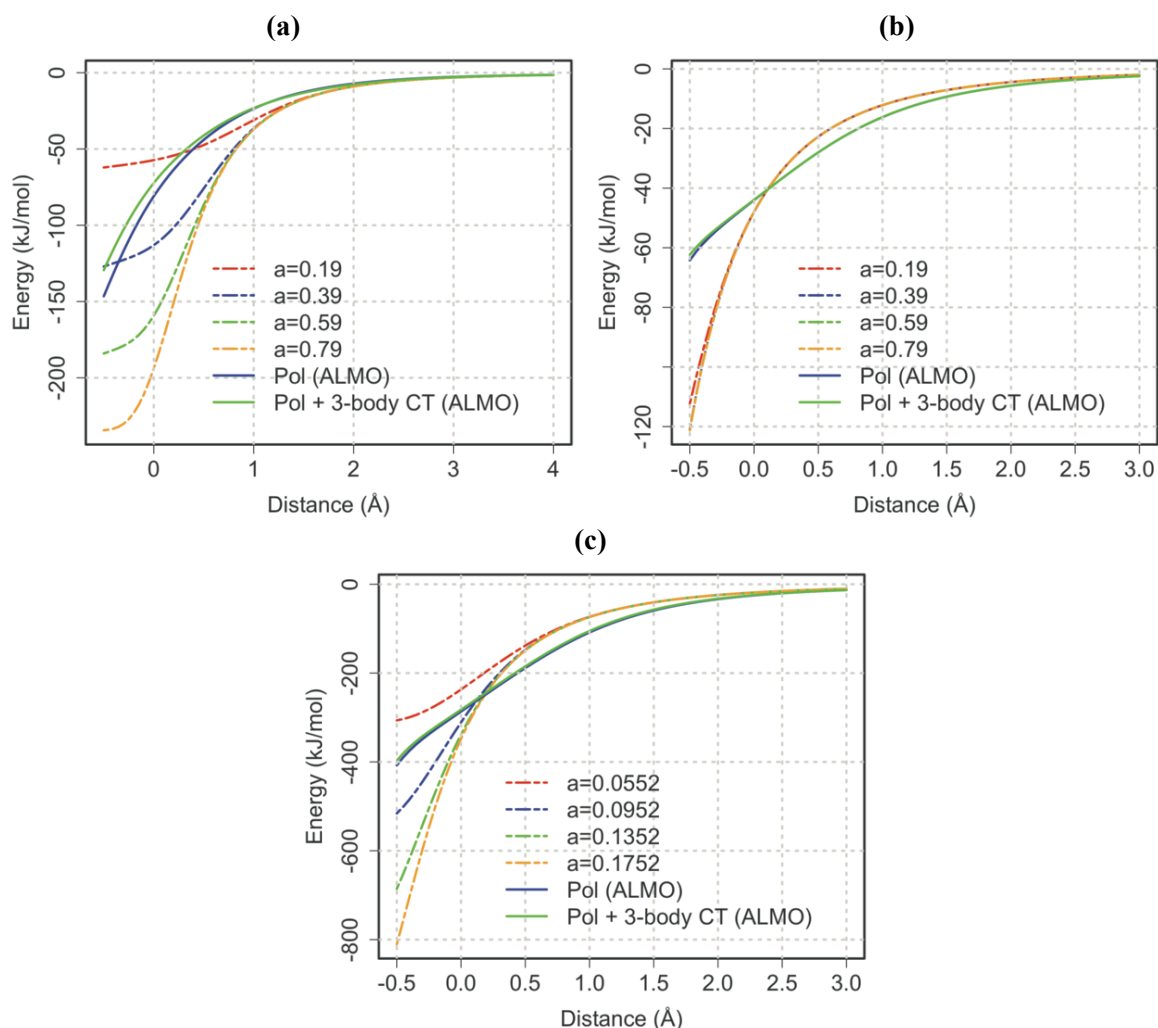

Figure 11. The effect of modifying the Thole damping parameter a for water-ion interactions in systems representative of the types of interacting species: a.) $(H_2O)_2F^-$. b.) $(H_2O)_2Na^+$. c.) $(H_2O)_2Mg^{2+}$. For all species, the default values for a in AMOEBA is 0.39, except for Mg^{2+}, whose a value is 0.0952; in all cases, the curve corresponding to the default value is dashed blue.

Following this, we then fixed the damping parameter between water O and H at 0.39, and varied the damping parameter of the water-ion interactions. In all the systems studied, we find that no single value

of a can yield a polarization profile that agreed with the ALMO curve throughout the distance range. If we focus on just points near equilibrium, in all the systems studied, a smaller a that corresponds to stronger damping could improve the agreement with ALMO's polarization (Table S2). Significantly, the same value that leads to improvement in the 2-body polarization can be as or as near optimal for the 3-body polarization and implicit incorporation of 3-body charge transfer (Figure 11 and S8). The real test is to see whether ion solvation free energies would benefit from these parameter changes. Nonetheless, we note that since the shape of the energy curve is wrong for a range of values, energy gradients that are critical for accurate MD would likely be in error.

Recently, the polarizabilities, damping parameter, and two different models of the damping were investigated and optimized by Liu and co-workers for a range of interactions, including ion/water and pure water systems[92], and therefore differ from the publically available versions from the TINKER designated download that we use here. Both models that were developed had slightly modified polarizabilities for Na^+, Mg^{2+}, and Ca^{2+} and were derived from RI-MP2 data. Figure S9 show that the 3-body energies of ALMO-EDA generated with the ωB97X-V functional are in good agreement with the 3-body energies of the RI-MP2 benchmark used in [92]. We then tested whether these models provide improved agreement for the systems and coordinate scans presented here against ALMO-EDA (Figure S10-S14). Model 1 used a slightly stronger damping parameter a (0.34) for all interactions, except for that of the direct field for Mg^{2+} and Ca^{2+}, for which a was set to 0.060 and 0.280, respectively. Overall Model 1 yields marginal improvement over standard AMOEBA, primarily near equilibrium geometries of the trimer systems examined here, compared to ALMO-EDA. Model 2 adopts an exponent of 3/2 for u instead of the original cubic power in the damping of the direct field (Eq. (7) in [92]) along with different damping parameters a, while the functional form and damping parameter for the mutual field are unchanged from standard AMOEBA. Model 2 is similar to AMOEBA and Model 1 with the exception of the $(H_2O)_2Cl^-$ and $(H_2O)_2Br^-$ trimers, where Model 2 is a significant improvement near equilibrium geometries over AMOEBA when benchmarked against ALMO-EDA. However, Model 2 shows large discrepancies with ALMO-EDA in the asymptotic region (especially for the water trimer) that are not problematic for standard AMOEBA and Model 1.

4. DISCUSSION

In this work, we assessed how well the many-body energy is captured by the AMOEBA force field by comparing against a variational, single-determinant EDA of a DFT energy calculated using a range-separated, dispersion-corrected functional with exact Hartree-Fock exchange. In our previous work on

water-ion and pure water dimers, we demonstrated good agreement of the ωB97M-V functional with the CCSD(T)/CBS "gold standard", indicating that the dispersion correction provides a reasonable description of true electron correlation. While we did not directly compare total DFT energies against CCSD(T) for the trimer systems in this work, we refer to work by Beran, Hobza, and co-workers that showed that for complexes dominated by polarization, such as those represented here, the most important component of a density functional in capturing many-body energy was the amount of true Hartree-Fock exchange that was present[93], justifying the use of the functional used here. Moreover, we showed that the major contributions to ALMO's 3-body energy are in near perfect accord with those obtained in a recent study using a variant of MP2 (RI-MP2)[92].

Overall, the AMOEBA force field benefits from a finely tuned cancellation of errors, yielding total energies that are in reasonable agreement with QM energies near equilibrium and for symmetric structures containing monovalent ions[54, 92]. However, further from equilibrium, agreement with QM breaks down, as demonstrated by angular and distance scans[69, 94]. Here we use the ALMO-EDA to reveal which components of AMOEBA are contributing to the 2- and 3-body errors and to what degree. We note that an EDA is not uniquely defined in the compressed regime, and the use of other EDA schemes might give rise to different conclusions in terms of the comparison against force field terms. For example, the density-based EDA (DEDA)[66] suggested by Wu et al. intuitively would be ideal to guide force field[95, 96], as its frozen density energy is variationally optimized while constraining the total system density to reproduce the sum of the densities of isolated monomers[96], yielding a more favorable frozen energy and smaller polarization and CT terms than those of the ALMO-EDA. Nevertheless, it was recently revealed that the energy lowering caused by the variational relaxation of the frozen wavefunction is largely due to the so-called "constant-density CT" effect[68], indicating that a strongly many-body effect is effectively absorbed into the Pauli term. Since the repulsive vdW term of AMOEBA is purely pairwise additive, the results of the many-body analyses of AMOEBA and QM might be even more discrepant if DEDA is employed.

Taken together, these findings point to a deficiency in how AMOEBA models the polarization of ion-water interactions, and the imperfect transferability of AMOEBA's water parameters and model to other systems. Although the validity of the assumption of linear dependency of the point inducible dipole on the field may be questioned, and quadratic and higher-order hyperpolarizabilities terms may be important[39, 40], these higher-order terms are deemed to be non-negligible only in the presence of strong electrostatic fields. In addition, the polarization may also involve dependencies on field gradients in addition to the field itself. Dependencies on the first and second derivatives of the field are associated

with induced quadrupoles and octupoles, respectively, and in fact, the FERF-DQ model used in the ALMO-EDA calculations takes the effect of quadrupolar polarization into account explicitly. However, according to the investigation by Soderhjelm *et al.*[97], these higher-order polarization effects are found to be of much less importance and urgency compared to the dominant effect of the need to account for Pauli repulsion through damping.

Therefore, the leading order error that lies in the form of the damping function within the point dipole model deserves further scrutiny. In his original work, Thole introduced a host of empirical damping functions, but only the functional form of Eq. (1) is in use in the AMOEBA force field. Masia *et al.* [44] found that for cation-water interactions, a linear form was superior in reproducing *ab initio* results to the form of Eq. (1). Slipchenko and Gordon tested damping functions for a range of interactions in the context of EFP calculations, including charge penetration, dispersion, and polarization[45]. For polarization, a Gaussian form and a form where the argument of the exponential is linear in interatomic distance were investigated, and it was found that the Gaussian damping function is better for benzene dimers, while the exponential form performs better for hydrogen-bonded complexes. Clearly, more attention needs to be paid to the functional form and system-specific predilections of the damping. This was recognized by developers of the AMOEBA potential, and modification of both the damping parameter and the functional form were proposed[92]. These modifications were tested on the trimer systems presented here, and showed that no single modified damping could consistently improve the agreement with the QM reference across the board for all systems, consistent with the observations of Slipchenko and Gordon [45] and Masia *et al.*[44] that different damping schemes were optimal for different systems. In this work, we have explored an additional damping model for the permanent field by incorporating the effect of charge penetration, which, like polarization damping, is a manifestation of the finite nature of the electron density. The caveat here is that we had to adopt a certain prescription for partitioning the monopole into core and smeared charge, and that the parameters are optimized to fit SAPT2+ permanent electrostatics, not the polarization. Indeed, the modeling of exchange-polarization coupling, manifested in the polarization damping function, used in conjunction with a charge penetration model for the permanent electrostatics may require special considerations in the parameterization of polarizable force fields, even though they are both models for the finite extent of molecular charge distributions.

It is also remarkable how much poorer AMOEBA's polarization model captures 3-body charge transfer in the ion-water trimer systems than in the water trimer. To shed some light on this discrepancy, we turn to the parameterization process of AMOEBA, specifically, the fine-tuning of AMOEBA's sole

many-body polarization term. In the course of parameterizing the AMOEBA water model, it was found that the Thole damping parameter had to be reduced from the original value of 0.59, which had been found ideal for reproducing gas-phase molecular polarizabilities, to its current value of 0.39, which results in a total molecular mechanical energy that better matches QM[22]. In other words, the polarization parameters for pure water in AMOEBA have been tuned to be more appropriate for calculations in the condensed phase, and an indirect result is that the many-body charge transfer is better captured implicitly through AMOEBA's sole many-body term. In contrast, no such procedure was used to optimize the polarizability parameters for ion-water clusters. While the Thole damping parameters were optimized for Ca^{2+} and Mg^{2+}, this was performed only for water-ion dimers[28], and therefore we cannot expect that the 3-body charge transfer would be captured in a systematic fashion through the polarization for these systems.

Lastly, we turn to the issue of anisotropy, which can be manifested in two ways. The first of these is the polarizability, which in principle is a rank-two tensor. In AMOEBA and other point inducible dipole models, the anisotropic polarizability is approximated as a single scalar isotropic polarizability. The other aspect of anisotropy is the addition of extra polarizable sites in areas of significant electron density distal from the atom center. A clear example of this in the systems studied here is the electron density corresponding to the lone-pair electrons on the water oxygen. Harder et al.[98] found that both anisotropic polarizability and inclusion of polarization centers on lone-pair oxygen sites improved geometries and energies for water dimers in the context of their Drude oscillator-based model. Interestingly, they found that polarization centers on lone-pair sites was more important than the use of anisotropic polarizabilities, and this conclusion was also found to be true in the development of the SIBFA potential (personal communication). Our results suggest the importance of anisotropy in the polarization for lone-pair species in two respects. First, we note the poorer performance of AMOEBA in modeling its 2-body ion-water polarizations compared with ion-ion polarizations in the $H_2O(M^{n+}/X^-)_2$ systems. Secondly, the poorer performance of AMOEBA in modeling the 3-body polarization for $H_2O(M^{n+})_2$ than for $H_2O(X^-)_2$ is likely related to the proximity of the cations to the lone-pair sites on the water oxygen in the distance scans.

5. CONCLUSION

In this article, we have extended the prior findings from our work on the dimer by focusing on the contrast of specifically many-body energetic components of the polarizable, multipole-based force field AMOEBA with their *ab initio* counterparts obtained by means of the ALMO-EDA method. By

performing distance scans of $(H_2O)_2M^{n+}$ and $(H_2O)_2X^-$ for a "dilute" ion regime as well as trimer systems consisting of $H_2O(M^{n+}/X^-)_2$ to mimic a concentrated ion environment, we have assessed how well the pairwise-additive approximation in AMOEBA holds for describing many-body dispersion and Pauli repulsion, how well AMOEBA's distributed inducible point dipoles with empirical Thole damping models a QM-derived polarization term, and how well charge transfer is implicitly captured through other terms in the AMOEBA force field.

First, the validity of the pairwise-additive approximation for the Pauli repulsion and dispersion terms, often collectively referred to as the vdW term modeled by the buffered 14-7 potential in AMOEBA, is supported by ALMO-EDA's Pauli and dispersion terms whose 3-body effects only become significant at very close intermolecular separation and amount to less than 1% of the total energy for all trimer systems evaluated. We note that dispersion is not the dominant binding force for the systems investigated here, although there is evidence for the importance of many-body dispersion in the condensed phase.[86] Thus it is worthwhile to extend the analysis to systems consisting of organic or hydrophobic moieties for which dispersion plays a more significant role so that its many-body effect may be non-negligible.

Therefore, in trimers, the remaining many-body effects that are appraised rest in polarization and charge transfer, the former of which is explicitly accounted for in AMOEBA whereas the latter is not. Given the fact that charge transfer and polarization may be grouped together as induction, we would have perhaps expected the 2-body polarization in AMOEBA to implicitly capture charge transfer, as does AMOEBA's 3-body polarization. Interestingly, and as a general rule across all dimer and trimer systems we have analyzed, the 2-body ALMO-EDA charge transfer is largely captured through AMOEBA's pairwise additive 14-7 van der Waals term, whereas AMOEBA's 3-body polarization agrees better with the sum of ALMO-EDA's 3-body charge transfer and polarization, particularly for the water trimer system. This may arise from the parameterization protocol. In the parameterization of AMOEBA03 water model, the parameters for the 14-7 potential were tuned to simultaneously reproduce the water dimer binding energy and geometry as well as the density and internal energy of liquid water [22], offering a likely explanation for the absorption of effects like charge transfer that are not accounted for explicitly at the 2-body level into the vdW potential. Furthermore, the polarization of water is also tuned (through its Thole width parameter) to match the reference total QM energies of water clusters, thereby providing a possibility for the absorption of the 3-body charge transfer. Such a tuning procedure was, however, not performed with the water-ion clusters, and therefore the 3-body polarization in water-ion trimer systems does not account for charge transfer as well as in the water trimer.

While AMOEBA's total polarization for the water trimer matches reasonably well the sum of ALMO-EDA's total polarization plus its 3-body charge transfer term, it is often not the case for ion-water systems. For the cation-water systems, AMOEBA's total polarization profile is excessively unfavorable (underbound) at and beyond equilibrium but excessively favorable in the compressed region, while the opposite trend is usually observed for the halide-water systems. In the halide-water systems, the effect of Thole damping to reduce the polarization in the compressed region is pronounced owing to its functional dependency on polarizabilities, which are larger for the halides. When turning to the many-body effects, the 3-body polarization in AMOEBA is excessively large in magnitude for the $(H_2O)_2M^{n+}$, $(H_2O)_2X^-$, and $H_2O(X^-)_2$ systems, even when measured against the sum of ALMO's 3-body polarization and charge transfer. The anti-cooperative behavior, on the other hand, is correctly displayed for these halide-water systems. In contrast, for all the $H_2O(M^{n+})_2$ systems with the exception of $H_2O(K^+)_2$, a qualitatively incorrect cooperative behavior is observed in all or part of the distance range for the 3-body polarization in AMOEBA.

While the incorporation of charge penetration improves agreement of AMOEBA's permanent electrostatic energy with its counterpart in ALMO-EDA dramatically, the same improvement is not observed for the polarization term when corrections for charge penetration are applied to the evaluation of the permanent field. Rather, the total polarization energy tended to be excessively favorable in the compressed region, reflecting the dominant 2-body trend, and the 3-body polarization is grossly inflated in magnitude with charge penetration in place. This is due in part to the documented effect that the improper modeling of the coupling of Pauli exchange and polarization becomes a more pronounced issue when permanent multipoles are separated into core and smeared valence components[91]. It could be argued that the Thole damping of the multipoles and the effect of charge penetration on the permanent electrostatic field are both manifestations of the finite nature of the charge distribution. However, once the Thole damping for the permanent electric field in Eq. (13) is replaced entirely with the damping terms that are prescribed by the charge penetration model, the resulting polarization energies are often significantly overestimated in the short range, suggesting that such a direct substitution might not be the proper way to incorporate the charge penetration model consistently for both permanent electrostatics and polarization. Further scrutiny on this is essential for the development of the next-generation AMOEBA force field that accounts for the finite extent of charge distributions.

SUPPORTING INFORMATION. The complete set of numerical comparison between AMOEBA and ALMO-EDA over distance scan PESs of any water trimer and water-ion trimers not represented in the main text.

REFERENCES

1.   M. A. Freitag, M. S. Gordon, J. H. Jensen and W. J. Stevens, J Chem Phys **112** (17), 7300-7306 (2000).
2.   J. P. Piquemal, N. Gresh and C. Giessner-Prettre, J Phys Chem A **107** (48), 10353-10359 (2003).
3.   L. V. Slipchenko and M. S. Gordon, J Comput Chem **28** (1), 276-291 (2007).
4.   B. Wang and D. G. Truhlar, J Chem Theory Comput **6** (11), 3330-3342 (2010).
5.   A. J. Stone, J Phys Chem A **115** (25), 7017-7027 (2011).
6.   Q. T. Wang, J. A. Rackers, C. He, R. Qi, C. Narth, L. Lagardere, N. Gresh, J. W. Ponder, J. P. Piquemal and P. Y. Ren, J Chem Theory Comput **11** (6), 2609-2618 (2015).
7.   N. Gresh, P. Claverie and A. Pullman, Int J Quantum Chem **29** (1), 101-118 (1986).
8.   N. Gresh, J Comput Chem **16** (7), 856-882 (1995).
9.   N. Gresh, J Phys Chem A **101** (46), 8680-8694 (1997).
10.   H. Li, M. S. Gordon and J. H. Jensen, J. Chem. Phys. **124** (21), 214108 (2006).
11.   N. Gresh, G. A. Cisneros, T. A. Darden and J. P. Piquemal, J Chem Theory Comput **3** (6), 1960-1986 (2007).
12.   A. J. Lee and S. W. Rick, J Chem Phys **134** (18) (2011).
13.   M. Soniat and S. W. Rick, J Chem Phys **140** (18) (2014).
14.   M. Soniat, L. Hartman and S. W. Rick, J Chem Theory Comput **11** (4), 1658-1667 (2015).
15.   M. Soniat, R. Kumar and S. W. Rick, J Chem Phys **143** (4) (2015).
16.   M. Sprik and M. L. Klein, J Chem Phys **89** (12), 7556-7560 (1988).
17.   P. O. Astrand and G. Karlstrom, Mol Phys **77** (1), 143-155 (1992).
18.   S. W. Rick, S. J. Stuart and B. J. Berne, J Chem Phys **101** (7), 6141-6156 (1994).
19.   P. Y. Ren and J. W. Ponder, J Comput Chem **23** (16), 1497-1506 (2002).
20.   G. Lamoureux, A. D. MacKerell and B. Roux, J Chem Phys **119** (10), 5185-5197 (2003).
21.   G. Lamoureux and B. Roux, J Chem Phys **119** (6), 3025-3039 (2003).
22.   P. Y. Ren and J. W. Ponder, J Phys Chem B **107** (24), 5933-5947 (2003).
23.   A. G. Donchev, V. D. Ozrin, M. V. Subbotin, O. V. Tarasov and V. I. Tarasov, P Natl Acad Sci USA **102** (22), 7829-7834 (2005).
24.   K. Palmo, B. Mannfors, N. G. Mirkin and S. Krimm, Chem Phys Lett **429** (4-6), 628-632 (2006).
25.   M. S. Gordon, L. Slipchenko, H. Li and J. H. Jensen, Ann Rep Comp Chem **3**, 177-193 (2007).
26.   A. Grossfield, P. Y. Ren and J. W. Ponder, J Am Chem Soc **125** (50), 15671-15682 (2003).
27.   A. Grossfield, P. Y. Ren and J. W. Ponder, Biophys J **84** (2), 94a-94a (2003).
28.   D. Jiao, C. King, A. Grossfield, T. A. Darden and P. Y. Ren, J Phys Chem B **110** (37), 18553-18559 (2006).
29.   B. A. Bauer and S. Patel, J Chem Phys **131** (8) (2009).

30.   M. E. Johnson, C. Malardier-Jugroot, R. K. Murarka and T. Head-Gordon, J Phys Chem B **113** (13), 4082-4092 (2009).

31.   H. B. Yu, T. W. Whitfield, E. Harder, G. Lamoureux, I. Vorobyov, V. M. Anisimov, A. D. MacKerell and B. Roux, J Chem Theory Comput **6** (3), 774-786 (2010).

32.   Y. Shi, C. J. Wu, J. W. Ponder and P. Y. Ren, J Comput Chem **32** (5), 967-977 (2011).

33.   F. Lipparini, L. Lagardere, B. Stamm, E. Cances, M. Schnieders, P. Y. Ren, Y. Maday and J. P. Piquemal, J Chem Theory Comput **10** (4), 1638-1651 (2014).

34.   A. Albaugh, O. Demerdash and T. Head-Gordon, J Chem Phys **143** (17) (2015).

35.   L. Lagardere, F. Lipparini, E. Polack, B. Stamm, E. Cances, M. Schnieders, P. Y. Ren, Y. Maday and J. P. Piquemal, J Chem Theory Comput **11** (6), 2589-2599 (2015).

36.   A. C. Simmonett, F. C. Pickard, Y. H. Shao, T. E. Cheatham and B. R. Brooks, J Chem Phys **143** (7) (2015).

37.   A. Albaugh, H. A. Boateng, R. T. Bradshaw, O. N. Demerdash, J. Dziedzic, Y. Z. Mao, D. T. Margul, J. Swails, Q. Zeng, D. A. Case, P. Eastman, L. P. Wang, J. W. Essex, M. Head-Gordon, V. S. Pande, J. W. Ponder, Y. H. Shao, C. K. Skylaris, I. T. Todorov, M. E. Tuckerman and T. Head-Gordon, J Phys Chem B **120** (37), 9811-9832 (2016).

38.   A. Albaugh, A. M. N. Niklasson and T. Head-Gordon, Journal of Physical Chemistry Letters **8** (8), 1714-1723 (2017).

39.   A. J. Stone, *The theory of intermolecular forces*. (Oxford University Press Inc., New York, 1996).

40.   C. E. Dykstra, Chemical Reviews **93** (7), 2339-2353 (1993).

41.   P. Y. Ren, C. J. Wu and J. W. Ponder, J Chem Theory Comput **7** (10), 3143-3161 (2011).

42.   Y. Shi, Z. Xia, J. J. Zhang, R. Best, C. J. Wu, J. W. Ponder and P. Y. Ren, J Chem Theory Comput **9** (9), 4046-4063 (2013).

43.   B. T. Thole, Chem Phys **59** (3), 341-350 (1981).

44.   M. Masia, M. Probst and R. Rey, J Chem Phys **123** (16) (2005).

45.   L. V. Slipchenko and M. S. Gordon, Mol Phys **107** (8-12), 999-1016 (2009).

46.   A. D. MacKerell, Jr., *Computational Biochemistry and Biophysics*. (Marcel Dekker, New York, 2001).

47.   A. Wallqvist, P. Ahlstrom and G. Karlstrom, Journal of Physical Chemistry **94**, 1649-1656 (1990).

48.   W. Chen and M. S. Gordon, J. Phys. Chem. **100** (34), 14316--14328 (1996).

49.   J. L. Gao, J Phys Chem B **101** (4), 657-663 (1997).

50.   G. A. Cisneros, J.-P. Piquemal and T. A. Darden, The Journal of chemical physics **123** (4), 044109-044109 (2005).

51.   V. Babin, C. Leforestier and F. Paesani, J Chem Theory Comput **9** (12), 5395-5403 (2013).

52.   V. Babin, G. R. Medders and F. Paesani, J Chem Theory Comput **10** (4), 1599-1607 (2014).

53.   G. R. Medders, V. Babin and F. Paesani, J Chem Theory Comput **10** (8), 2906-2910 (2014).

54.   D. J. Arismendi-Arrieta, M. Riera, P. Bajaj, R. Prosmiti and F. Paesani, J. Phys. Chem. B **120** (8), 1822-1832 (2016).

55.   K. Kitaura and K. Morokuma, Int. J. Quantum Chem. **10** (2), 325--340 (1976).

56.   M. von Hopffgarten and G. Frenking, WIREs Comput. Mol. Sci. **2** (1), 43--62 (2012).

57.   M. J. S. Phipps, T. Fox, C. S. Tautermann and C.-K. Skylaris, *Energy decomposition analysis approaches and their evaluation on prototypical protein-drug interaction patterns*. (Royal Society of Chemistry, 2015).

58.   B. Jeziorski, R. Moszynski and K. Szalewicz, Chem. Rev. **94** (7), 1887--1930 (1994).

59.   E. G. Hohenstein and C. D. Sherrill, WIREs Comput. Mol. Sci. **2** (2), 304--326 (2012).

60.   K. Szalewicz, WIREs Comput. Mol. Sci. **2** (2), 254--272 (2012).

61.   P. S. Bagus, K. Hermann and C. W. Bauschlicher, Jr., J. Chem. Phys. **80** (9), 4378--4386 (1984).

62.   W. J. Stevens and W. H. Fink, Chem. Phys. Lett. **139** (1), 15--22 (1987).
63.   Y. Mo, J. Gao and S. D. Peyerimhoff, J. Chem. Phys. **112** (13), 5530--5538 (2000).
64.   Y. Mo, P. Bao and J. Gao, Phys. Chem. Chem. Phys. **13** (15), 6760--6775 (2011).
65.   F. M. Bickelhaupt and E. J. Baerends, **15**, 1--86 (2007).
66.   Q. Wu, P. W. Ayers and Y. Zhang, J. Chem. Phys. **131** (16), 164112 (2009).
67.   R. Z. Khaliullin, E. A. Cobar, R. C. Lochan, A. T. Bell and M. Head-Gordon, J. Phys. Chem. A **111** (36), 8753--8765 (2007).
68.   P. R. Horn, Y. Mao and M. Head-Gordon, J. Chem. Phys. **144**, 114107 (2016).
69.   Y. Z. Mao, O. Demerdash, M. Head-Gordon and T. Head-Gordon, J Chem Theory Comput **12** (11), 5422-5437 (2016).
70.   P. R. Horn, Y. Z. Mao and M. Head-Gordon, Phys Chem Chem Phys **18** (33), 23067-23079 (2016).
71.   N. Mardirossian and M. Head-Gordon, Phys. Chem. Chem. Phys. **16** (21), 9904--9924 (2014).
72.   D. Rappoport and F. Furche, J. Chem. Phys. **133** (13), 134105 (2010).
73.   A. J. Stone and A. J. Misquitta, Chem. Phys. Lett. **473** (1-3), 201--205 (2009).
74.   K. U. Lao and J. M. Herbert, J Chem Theory Comput **12** (6), 2569-2582 (2016).
75.   T. A. Halgren, J Am Chem Soc **114** (20), 7827-7843 (1992).
76.   D. Hankins, J. W. Moskowitz and F. H. Stillinger, J Chem Phys **53** (12), 4544-+ (1970).
77.   H. Stoll and H. Preuss, Theor Chim Acta **46** (1), 11-21 (1977).
78.   R. M. Richard and J. M. Herbert, J Chem Phys **137** (6) (2012).
79.   A. J. Stone, Chem Phys Lett **83** (2), 233-239 (1981).
80.   J. A. Rackers, Q. Wang, C. Liu, J. P. Piquemal, P. Ren and J. W. Ponder, Phys Chem Chem Phys **19**, 276-291 (2017).
81.   C. A. Coulson, P Camb Philos Soc **38**, 210-223 (1942).
82.   H. Stoll, G. Wagenblast and Preu, Theor. Chem. Acc. **57** (2), 169--178 (1980).
83.   E. Gianinetti, M. Raimondi and E. Tornaghi, Int. J. Quantum Chem. **60** (1), 157--166 (1996).
84.   R. Z. Khaliullin, M. Head-Gordon and A. T. Bell, J. Chem. Phys. **124** (20), 204105 (2006).
85.   P. R. Horn and M. Head-Gordon, J. Chem. Phys. **143** (11), 114111 (2015).
86.   J. W. Ponder, (Washington University, St. Louis, MO, 2015), Vol. 7.1.
87.   T. M. Parker, L. A. Burns, R. M. Parrish, A. G. Ryno and C. D. Sherrill, J Chem Phys **140** (9) (2014).
88.   Y. H. Shao, Z. T. Gan, E. Epifanovsky, A. T. B. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. T. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Kus, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M. Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. Woodcock, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. O. Beran, Y. A. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C. M. Chang, Y. Q. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, R. A. DiStasio, H. Do, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. D. Hanson-Heine, P. H. P. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T. C. Jagau, H. J. Ji, B. Kaduk, K. Khistyaev, J. Kim, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. D. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S. P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, E. Neuscamman, C. M. Oana, R. Olivares-Amaya, D. P. O'Neill, J. A. Parkhill, T. M. Perrine, R. Peverati, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, S. M. Sharada, S. Sharma, D. W. Small, A. Sodt, T. Stein, D. Stuck, Y. C. Su, A. J. W. Thom, T. Tsuchimochi, V. Vanovschi, L. Vogt, O. Vydrov, T. Wang, M. A. Watson, J. Wenzel, A. White, C. F. Williams, J. Yang, S. Yeganeh, S. R. Yost, Z. Q. You, I. Y. Zhang, X. Zhang, Y. Zhao, B. R. Brooks, G. K. L. Chan, D. M. Chipman, C. J. Cramer, W. A. Goddard, M. S. Gordon, W. J. Hehre, A. Klamt, H. F.

Schaefer, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J. D. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C. P. Hsu, Y. S. Jung, J. Kong, D. S. Lambrecht, W. Z. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. Van Voorhis, J. M. Herbert, A. I. Krylov, P. M. W. Gill and M. Head-Gordon, Mol Phys **113** (2), 184-215 (2015).
89.   O. A. a. Vydrov, J. Chem. Phys. **133** (24), 244103 (2010).
90.   P. M. W. Gill, B. G. Johnson and J. A. Pople, Chem Phys Lett **209** (5-6), 506-512 (1993).
91.   Y. Z. Mao, Y. Shao, J. Dziedzic, C. K. Skylaris, T. Head-Gordon and M. Head-Gordon, J Chem Theory Comput **Just Accepted** (2017).
92.   C. Liu, R. Qi, Q. Wang, J. P. Piquemal and P. Ren, J Chem Theory Comput **13** (6), 2751-2761 (2017).
93.   J. Řezáč, Y. Huang, P. Hobza and G. J. O. Beran, J Chem Theory Comput **11** (7), 3065-3079 (2015).
94.   M. Riera, A. W. Gotz and F. Paesani, Phys Chem Chem Phys **18** (44), 30334-30343 (2016).
95.   Z. Lu, N. Zhou, Q. Wu and Y. Zhang, J Chem Theory Comput **7** (12), 4038-4049 (2011).
96.   N. Zhou, Z. Lu, Q. Wu and Y. Zhang, The Journal of Chemical Physics **140** (21), 214117 (2014).
97.   P. Soderhjelm, A. Ohrn, U. Ryde and G. Karlstrom, J Chem Phys **128** (1) (2008).
98.   E. Harder, V. M. Anisimov, I. V. Vorobyov, P. E. M. Lopes, S. Y. Noskov, A. D. MacKerell and B. Roux, J Chem Theory Comput **2** (6), 1587-1597 (2006).