Assessing ion-water interactions in the AMOEBA force field using energy decomposition analysis of electronic structure calculations

Omar Demerdash,[†,¶] Yuezhi Mao,[†,¶] Martin Head-Gordon,[*,†] and Teresa Head-Gordon[*,‡]

†*Kenneth S. Pitzer Center for Theoretical Chemistry, Department of Chemistry, University of California at Berkeley Berkeley, CA 94720*

‡*Kenneth S. Pitzer Center for Theoretical Chemistry, Department of Chemistry, Department of Bioengineering, Department of Chemical and Biomolecular Engineering, University of California at Berkeley Berkeley, CA 94720*

¶*These two authors contributed equally*

E-mail: mhg@cchem.berkeley.edu; thg@berkeley.edu

Abstract

AMOEBA is a molecular mechanics force field that addresses some of the shortcomings of a fixed partial charge model, by including permanent atomic point multipoles through quadrupoles, as well as many-body polarization through the use of point inducible dipoles. In this work we investigate how well AMOEBA formulates its non-bonded interactions, and how it implicitly incorporates quantum mechanical effects

such as charge penetration (CP) and charge transfer (CT), for water-water and water-ion interactions. We find that AMOEBA's total interaction energies as a function of distance and over angular scans for the water dimer and for a range of water-monovalent cations agree well with an advanced density functional theory (DFT) model, whereas the water-halides and water-divalent cations show significant disagreement with the DFT result, especially in the compressed region when the two fragments overlap. We use a second generation energy decomposition analysis (EDA) scheme based on absolutely localized molecular orbitals (ALMOs) to show that in the best cases AMOEBA relies on cancellation of errors by softening of the van der Waals (vdW) wall to balance permanent electrostatics that are too unfavorable, thereby compensating for the missing CP effect. CT, as another important stabilizing effect not explicitly accounted for in AMOEBA, is also found to be incorporated by the softened vdW interaction. For the water-halides and water-divalent cations, this compensatory approach is not as well executed by AMOEBA over all distances and angles, wherein permanent electrostatics remains too unfavorable and polarization is over-damped in the former while overestimated in the latter. We conclude that the DFT-based EDA approach can help refine a next generation AMOEBA model that either realizes a better cancellation of errors for problematic cases like those illustrated here, or to guide the parameterization of explicit functional forms for short-range contributions from CP and/or CT.

1 Introduction

Condensed-phase simulations with molecular dynamics or Monte Carlo methods afford the ability to probe physical properties not easily accessible by experimental means, but only when there is a reliable model of the potential energy surface. However for simulations on scales representing tens of thousands to millions of atoms, quantum mechanical (QM) methods become extremely intractable, if not impossible, to serve that purpose. Therefore,

there exists a keen interest in reducing the computational cost by representing such systems classically, and considering moreover only the nuclear degrees of freedom explicitly under the Born-Oppenheimer separation of the timescales of the electronic and nuclear degrees of freedom.

In molecular mechanics (MM) formulations, a potential energy function of the positions of the atomic nuclear positions is used to describe covalent interactions (bond stretching, angle bending, and torsional rotation, and sometimes other cross-terms) as well as non-covalent interactions such as van der Waals (vdW) and point-charge permanent electrostatics. Fundamentally the MM description involves idealized functional forms for these terms such as

$$\begin{aligned} U = & \sum_{i}^{N} k_b (b - b_0)^2 + \sum_{i}^{N} k_\theta (\theta - \theta_0)^2 + \sum_{i}^{N} k_\phi \cos(n\phi + \delta) \\ & + \sum_{i<j} 4\epsilon_{ij}[(\sigma_{ij}/r_{ij})^{12} - (\sigma_{ij}/r_{ij})^{6}] + \sum_{i<j} q_i q_j / r_{ij} \end{aligned} \tag{1}$$

For bonded atoms, Eq. (1) permits only small fluctuations around the equilibrium bond-lengths and bond angles. It employs nuclear-centered point charges for the permanent electrostatics, and uses simple pairwise additivity and mixing rules for the vdW interactions. It has also become more common to replace point charges with an extended set of point multipoles, and to include non-pairwise additive interactions such as polarization which is discussed in detail below.

While the long-range asymptotic behavior of these functional forms are correctly modeled for electrostatics and dispersion,[1] the short-ranged asymptotics in the compressed region are more approximate, for example in the steep region of the van der Waals function which assumes a power law for the Pauli exclusion term as opposed to the exponential form from QM. Other short-ranged QM features such as a non-local electron density extent from the nucleus, which is related to the charge penetration (CP) effect, or interactions such as charge

transfer (CT), are ignored altogether in Eq. (1), although very recently approximate classical formulations to these terms are just starting to appear in the literature.[2–18]

Although the classical functional forms for the Born-Oppenheimer potential energy surface (BO-PES) are largely correct, and are starting to approach a level of completeness for all relevant QM interactions, the methods by which the parameters are obtained for these functions can be more arbitrary. This is especially true for the soft, non-covalent degrees of freedom that fluctuate much more in distance than that experienced due to the covalent interactions at typical ambient conditions. For example, deriving point charge or point multipole parameters by fitting to the QM-derived electrostatic potential at a series of distances outside of the vdW surface is often a poorly defined problem, with charges or multipoles of buried atoms far from the surface being able to vary unphysically and still reproduce the potential,[19] and such problems are compounded when including the additional polarization terms. The vdW parameters are especially difficult to fit since the attractive part of the vdW interaction, the dispersion interaction, cannot be fitted easily to QM data owing to the difficulty of finding accurate yet tractable levels of theory that adequately capture electron correlation.[20–22] Together with the repulsive wall, the overall vdW function is typically tuned through a reliance on experimental data, by adjusting the parameters to reproduce densities and heats of vaporization of neat liquids.[23–27] In typical force field development, parameters for the individual terms in Eq. (1) are determined independently and then refined together in an iterative fashion to capture missing features like charge penetration, polarization if it is absent, or even directly targeting better properties for a particular phase (such as liquid water).[28]

Electronic structure calculations have been routinely used to benchmark the quality of force field parameterization, since they facilitate the direct comparison of resulting interaction energies and other properties within a given model system. Nevertheless, considering the piecemeal parameterization of a force field based on several minimally correlated ap-

proximating assumptions and the prevalence of incomplete error cancellation between terms in the final form, it would be helpful to benchmark against a method that is able to ascertain the quality of individual terms of the force field. Energy decomposition analysis (EDA) affords a way to determine the relative contributions of several physically meaningful terms out of the QM interaction energy, e.g., permanent electrostatics, Pauli repulsion, polarization, dispersion, etc.[29–31] The asymptotic components of an EDA method are uniquely defined for a given electronic structure method. Furthermore, despite some arbitrariness that exists in the definitions of these terms (since none of them are true experimental observables), well-defined EDAs can yield a reasonable and chemically sensible separation of energy components in the overlapping regime. Therefore, by comparing the corresponding terms between an EDA scheme and a force field, one can in principle obtain insight into the strengths and weaknesses of MM formulations.

In this work, we analyze the energy decomposition of the popular polarizable MM force field AMOEBA.[26,27,32,33] The AMOEBA model goes beyond typical fixed-charge force field by including both higher-order permanent multipoles and inducible dipoles, in an attempt to more faithfully reproduce the BO-PES from more accurate QM methods. As such, AMOEBA results can be directly compared to and parameterized against a corresponding BO-PES derived from *ab initio* electronic structure calculations, while remaining compatible with the goal of increased computational efficiency and ease of parameterization for a variety of systems including ions,[34,35] small organic molecules,[27] transition metal complexes,[36,37] proteins,[38] and nucleic acids either in the gas or condensed phases. The subject of this paper is to assess how well the individual energy terms of the AMOEBA model are recapitulated when compared to QM on the simplest of systems — namely the water dimer and a series of water-ion dimers — over a range of distances and angular orientations.

Of the available energy decomposition schemes applied to force field development, two of the more well known examples include symmetry-adapted perturbation theory (SAPT)[7,8,39–46]

and variational based EDA approaches.[35,39,47–52] SAPT evaluates intermolecular interactions via a perturbative approach,[20,22,53] and the resulting interaction energies are decomposed into the contributions of electrostatics, exchange, induction, and dispersion. The development of SAPT(DFT)[54–57] or DFT-SAPT,[58–60] which tackles intramolecular correlation by means of Kohn-Sham (KS)-DFT,[61] offers a reasonable balance between accuracy and efficiency ($O(N^5)$ scaling with density-fitting[62,63]) so that it can be used for generating the training data for force field parametrization.[46] While SAPT-based methods facilitate the development of many advanced force fields (the readers are referred to the review in Ref. 46), it still faces several challenges. Most importantly, the accuracy of the total interaction energy relies on the convergence of the perturbative expansion, which is often difficult to assure when methods like DFT-SAPT are used for systems with strong induction effects.[64,65] Also, there is no clean separation between polarization and charge transfer in the conventional formulation of SAPT (they both belong to the induction term), although several approaches have been proposed to extract the CT energy.[66–68]

We shall evaluate AMOEBA using the variational absolutely localized molecular orbital (ALMO)-EDA scheme,[49,52,69] which partitions the total intermolecular interaction energy into contributions of frozen orbital interaction (which contains permanent electrostatics, Pauli repulsion, and dispersion), polarization and CT. New advances made in the ALMO-EDA scheme include (1) the ability to reach a meaningful complete basis set (CBS) limit for polarization and CT using the fragment electric-field response function (FERF) model[70] and (2) the ability to disentangle the contributions from the aforementioned three components of the frozen term,[71] which is not further separated in the original scheme. These advances define a second generation ALMO-EDA[52] method which is employed in the present paper.

Apart from the robustness of this EDA scheme, another important motivation for employing a DFT-based EDA approach is the significantly reduced errors of state-of-the-art density functionals for non-covalent interactions.[72–76] It is noteworthy that the functionals recently

developed by Mardirossian and Head-Gordon[77–79] demonstrate very good accuracy for non-covalent interactions when large basis sets are employed. Based on recent benchmark by Lao et al.,[65] the mean absolute error (MAE) of ωB97X-V[77] (using the aug-cc-pVTZ (aTZ)[80,81] basis set) for a composite dataset comprising neutral-anion, neutral-cation, and cation-anion interactions (43 data points in total) is 0.55 kcal/mol, which is superior to that of the popular DFT-SAPT method (1.43 kcal/mol) and comparable to the most accurate SAPT result available (0.43 kcal/mol, as computed at the SAPT2+3-δMP2/aTZ level of theory). We note that such high-level SAPT methods are computationally costly ($O(N^7)$ scaling) and offer a less unambiguous energy decomposition due to the coupling between terms. Beyond equilibrium binding energies, accurate PESs generated by ωB97X-V for water-water[76] and water-anion (F^-, Cl^-) dimers[82] have also been recently reported. Therefore, we use ωB97X-V for all the ALMO-EDA calculations in the present paper.

The remainder of this paper is outlined as follows. Pertinent details of the AMOEBA force field are summarized in Sec. 2.1, followed in Sec. 2.2 by a concise summary of the version of the ALMO-EDA that will be applied in this work and the mapping between its terms and those in AMOEBA. The resulting data and analysis for four categories of interactions are demonstrated and discussed in Sec. 3. For each category, we first compare total energies evaluated with ωB97X-V and AMOEBA, and assess the agreement near the equilibrium configuration, as well as at short-range (the so-called compressed region) and long-range (the asymptotic region). We then compare the relative contributions of the energetic components, including permanent electrostatics, polarization, and vdW interactions, in various regions, via which we elucidate why the resultant total interaction energy profile given by AMOEBA for each system has satisfactory or poor agreement with the ωB97X-V result. Moreover, in cases where total interaction energies are in reasonable agreement between DFT and AMOEBA, we investigate how the effects of CP and CT might be accounted for implicitly by AMOEBA, as it lacks explicit functional forms for these effects. The insights gained from

these benchmark calculations are discussed in Sec. 4.

2 Computational methods

2.1 Non-covalent terms in the AMOEBA force field

The non-covalent (intermolecular) terms in the AMOEBA force field comprises permanent electrostatics (U_{ele}^{perm}), induced electrostatics (U_{ele}^{ind}), and van der Waals (vdW) interactions (U_{vdw}). An atom-centered point multipole model is adopted for permanent electrostatics: on each atomic site i, the (vector) of permanent multipoles $\mathbf{M}_i$ includes monopole (q_i), dipole ($\boldsymbol{\mu}_i$), and quadrupole ($\mathbf{Q}_i$) moments:

$$\mathbf{M}_i^T = [q_i, \mu_{ix}, \mu_{iy}, \mu_{iz}, Q_{ixx}, Q_{ixy}, Q_{ixz}, Q_{iyy}, Q_{iyz}] \tag{2}$$

The total permanent electrostatics contribution is then evaluated as the pairwise sum of interactions between different atomic sites:

$$U_{ele}^{perm} = \sum_{i<j} \mathbf{M}_i^T \mathbf{T}_{ij} \mathbf{M}_j \tag{3}$$

where $\mathbf{T}_{ij}$ is the "composite" multipole interaction tensor between sites i and j, that contain appropriate powers of $1/r_{ij}$ according to the permanent multipole expansion for the AMOEBA potential.[26,32] For water and simple ions, the interaction sites i and j are constrained to be on the different molecular or atomic fragments; and for the mono-atomic ions studied in this work, the RHS of Eq. (2) contains monopole (q_i) only. The set of multipoles are derived from a QM electronic density using the distributed multipole analysis (DMA),[83] whose values are further refined by a fit to an electrostatic potential generated by a higher level of QM theory.[27]

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

$$\boldsymbol{\mu}_i^{ind} = \alpha_i(\mathbf{E}_i + \mathbf{E}_i') \tag{4}$$

where $\mathbf{E}_i$ is the electric field owing to the permanent multipoles on other fragments, and $\mathbf{E}_i'$ is the field generated by the induced dipoles on all the other atomic sites:

$$\mathbf{E}_i = \sum_j \mathbf{T}_{ij}\mathbf{M}_j^{(d)} \tag{5}$$

$$\mathbf{E}_i' = \sum_{j\neq i} \mathbf{T}_{ij}'\boldsymbol{\mu}_j^{ind} \tag{6}$$

where $\mathbf{T}_{ij}'$ now refers to appropriate powers of $1/r_{ij}$ according to the dipole induction and the superscript (d) refers to special scaling factors used for electrostatic interactions in AMOEBA[27] (not involved in this work). Since the RHS of Eq. (4) relies on the induced dipoles, $\{\boldsymbol{\mu}_i^{ind}\}$ are solved self-consistently in order to complete the calculation of the many-body polarization effects. With converged $\{\boldsymbol{\mu}_i^{ind}\}$, the associated energy lowering (polarization energy) is determined by

$$U_{ele}^{ind} = -\frac{1}{2}\sum_i \boldsymbol{\mu}_i^{ind} \cdot \mathbf{E}_i \tag{7}$$

The atomic polarizability parameters, $\{\alpha_i\}$, are derived by a fit to available experimental molecular polarizabilities.[26,32]

One artifact of the distributed interactive induced electrostatics model is the so-called "polarization catastrophe", i.e., the electric field generated by point multipoles can severely overpolarize at short range and even lead to divergence. To ensure the finite nature of the

intermolecular induction effect, a Thole-style damping scheme is employed by AMOEBA,[84,85] which is equivalent to replacing a point multipole with a smeared charge distribution. For example, the damping function for monopoles has the following form:

$$\rho = \frac{3a}{4\pi}\exp(-au^3), \quad u = \frac{r_{ij}}{(\alpha_i\alpha_j)^{1/6}} \tag{8}$$

where r_{ij} is the distance between sites i and j, α_i, α_j are their polarizabilities, and a is a dimensionless width parameter. The damping functional forms for higher multipoles are reported in Ref. 26. In practice, the damping functions are built in the formation of multipole interaction tensors in Eq. (5) and (6).

For the vdW interaction, AMOEBA adopts a pairwise additive buffered 14-7 potential originally proposed by Halgren:[86]

$$U_{vdw} = \sum_{i<j}\epsilon_{ij}\left(\frac{1+\delta}{\rho_{ij}+\delta}\right)^7\left(\frac{1+\gamma}{\rho_{ij}^7+\gamma}-2\right) \tag{9}$$

where ϵ_{ij} is the depth of the potential well, ρ_{ij} is the dimensionless distance between sites i and j: $\rho_{ij} = r_{ij}/R^0_{ij}$, where R^0_{ij} is the minimum energy separation, and γ and δ are two constants whose values are set to 0.12 and 0.07, respectively. If we further expand the RHS of Eq. (9), the repulsive "14" term (U^{rep}_{vdw}) mostly accounts for Pauli repulsion, while the attractive "7" term (U^{attr}_{vdw}) in principle yields a more accurate series expansion for dispersion. In the current parameterization regime of AMOEBA, the homonuclear vdW parameters ϵ_{ii} and R^0_{ii} are obtained by a fit to dimer energies calculated at the MP2/aTZ level of theory or higher, followed by a refinement stage where experimental liquid densities and heats of vaporization are reproduced;[27] and the heteronuclear parameters (ϵ_{ij} and R^0_{ij}) are obtained by using the combination rules. For more details we refer the reader to the original AMOEBA references.[26,27,32,33]

2.2 Energy decomposition analysis

The initial partitioning of the total interaction energy in the ALMO-EDA method[49,52,69] is as follows:

$$E_{int} = E_{gd} + E_{frz} + E_{pol} + E_{ct}. \tag{10}$$

The first term (geometry distortion) describes the energy change due to the geometric distortion of monomers to the complex structure from their fully relaxed geometries, which is *not* considered in this work (in AMOEBA it is captured by the bonded terms). The frozen orbital interaction, E_{frz}, is defined as the energy difference between the frozen orbital wavefunction (which corresponds to a frozen one-particle density matrix $\mathbf{P}_{frz}$) and the sum of monomer energies that are computed individually:

$$E_{frz} = E(\mathbf{P}_{frz}) - \sum_A E_A(\mathbf{P}_A), \tag{11}$$

It represents the energy change when fragments approach each other without any variational relaxation of their orbitals or density, apart from ensuring that they obey the Pauli Principle.

The frozen interaction can be further decomposed into contributions from permanent electrostatics, Pauli repulsion and dispersion interactions. Our original approach[71] is based on the partitioning of $\mathbf{P}_{frz}$ into a sum of fragment contributions: $\mathbf{P}_{frz} = \sum_A \tilde{\mathbf{P}}_A$, where interfragment orthogonality is enforced between $\tilde{\mathbf{P}}_A$'s. This is achieved via a constrained minimization of the "kinetic energy pressure" (KEP) objective function, as described in Ref. 71.

$\tilde{\mathbf{P}}_A$ can be regarded as the deformed (but not yet relaxed) density of each individual fragment upon the formation of the initial complex. The dispersion energy is then defined as the remainder of interfragment exchange-correlation (XC) energy after subtracting the part

that is "dispersion-free" (DF) in nature:

$$E_{disp} = \left(E_{xc}[\mathbf{P}_{frz}] - \sum_A E_{xc}[\tilde{\mathbf{P}}_A] \right) - \left(E_{xc}^{DF}[\mathbf{P}_{frz}] - \sum_A E_{xc}^{DF}[\tilde{\mathbf{P}}_A] \right), \tag{12}$$

For this purpose, an auxiliary DFXC functional is required, and our previous work[52,71] suggests that Hartree-Fock (HF) theory is an appropriate choice for dispersion-corrected range-separated hybrid (RSH) functionals like ωB97X-V.

The approach presented in Ref. 71 utilizes $\{\tilde{\mathbf{P}}_A\}$ to define permanent electrostatics and Pauli repulsion as well, i.e., all three energy components are computed making use of the properly antisymmetrized wavefunction. That approach properly describes the deformation of monomer densities due to Pauli repulsion (antisymmetrization of the supersystem wavefunction).[71] However such deformations cannot be captured by AMOEBA or any other force field whose permanent electrostatics is described by multipole moments that are invariant with intermolecular separations. Therefore, in order to make ALMO-EDA's permanent electrostatics physically compatible with AMOEBA, we *step back* to adopt the "classical electrostatics" definition,[29,87–89] which describes the Coulomb interaction between charge distributions of isolated fragments:

$$E_{elec}^{cls} = \sum_{A<B} \int_{r_1} \int_{r_2} \rho_A^{tot}(\mathbf{r}_1) \frac{1}{\mathbf{r}_{12}} \rho_B^{tot}(\mathbf{r}_2) \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 \tag{13}$$

where $\rho_A^{tot}(\mathbf{r}) = \rho_A^{el}(\mathbf{r}) + \rho_A^{nuc}(\mathbf{r})$. The modified Pauli term is then simply defined as the remainder of the frozen interaction, which still comprises the repulsive interaction stemming from Pauli exclusion principle and interfragment exchange effects. Taken together, the

decomposition of the frozen term adopted in this work can be expressed as

$$E_{frz} = E_{elec}^{cls} + E_{Pauli}^{mod} + E_{disp} \tag{14}$$

The contribution of polarization is determined by variationally optimizing the density matrix associated with a fragment-block-diagonal (absolutely localized) MO coefficient matrix, using locally projected self-consistent field (SCF) techniques[90–92] or a gradient-based minimization.[70] This procedure is called "SCF for molecular interactions" (SCF-MI). The resulting ALMO density matrix, $\mathbf{P}_{ALMO}$, is employed to evaluate the polarization energy:

$$E_{pol} = E[\mathbf{P}_{ALMO}] - E[\mathbf{P}_{frz}] \tag{15}$$

The use of ALMOs ensures that the net charge on each fragment is conserved under the Mulliken population definition, i.e., CT between fragments is not allowed. The contribution of CT is then defined as the energy difference between this "CT-forbidden" SCF-MI wavefunction and the fully relaxed one:

$$E_{ct} = E[\mathbf{P}_{SCF}] - E[\mathbf{P}_{ALMO}] \tag{16}$$

In practice, the AO-based fragment partitioning used in SCF-MI breaks down when very large AO basis sets (especially those with diffuse functions) are used, which results in an overestimated polarization energy (contaminated by CT).[68,70,93] In order to judiciously choose the degrees of freedom that are relevant to polarization, Horn and Head-Gordon introduced the fragment electric-field response function (FERF) model[70] which defines the fragment subspaces based on the response of MOs to an external electric field (and its spatial derivatives). The FERFs are able to capture the relaxation of fragment occupied orbitals under a weak external field, which is deemed as the physical essence of polarization (see

Ref. 70 for more details on how FERFs are constructed). In this work, we use the "FERF-nDQ" model (non-orthogonal FERFs accounting for dipole and quadrupole responses, which require 3 and 5 FERFs for each occupied orbital, respectively) to compute the polarization energy instead of the original AO-based approach. The FERF-nDQ model appears to give a satisfactory description[52,70] of polarization effects with a well-behaved basis set limit. The equations utilized to determine the contributions of polarization and CT are identical to Eq. (15) and (16) in form.

Here we also briefly compare the intermolecular interaction components generated by ALMO-EDA with those from a standard DFT-SAPT calculation, due to the popularity of the latter approach in developing physically-motivated force fields. The DFT-SAPT interaction energy can be written as (following the notation of Ref. 64):

$$\begin{aligned} E_{int}^{\text{DFT-SAPT}} = & [E_{elst}^{(1)}]_{elst} + [E_{exch}^{(1)}]_{exch} \\ & + [E_{ind}^{(2)} + E_{exch-ind}^{(2)} + \delta E_{HF}^{(2)}]_{ind} \\ & + [E_{disp}^{(2)} + E_{exch-disp}^{(2)}]_{disp} \end{aligned} \tag{17}$$

Subscripts outside the brackets indicate the grouping of terms into four energy components: electrostatics, exchange, induction and dispersion. While the first-order electrostatic term is identical to ALMO-EDA's (classical) electrostatics (Eq. (13)) in form, the other components are computed differently in DFT-SAPT due to its perturbative approach. Nonetheless, based on the physical meaning of each term, there exists a clear correspondence between the terms in DFT-SAPT (left) and ALMO-EDA (right):

- Exchange → Pauli repulsion
- Induction → Polarization + CT
- Dispersion → Dispersion

Numerically their resulting energy components should be comparable at least to some extent. As an example, we compare the results of DFT-SAPT (provided in Ref. 94) and ALMO-EDA for CCSD(T)[95]-optimized structures of the "linear" and "bifurcated" water dimer (taken from Ref. 96), which are presented in Table S1 of the Supplementary Material. In general, there is *no* qualitative difference between the results of these two decomposition schemes (in contrast, the difference between the energy components of AMOEBA and either EDA can be much larger), while the separation of polarization and CT in ALMO-EDA seems to further facilitate the comparison with AMOEBA.

Finally, the correspondence between terms in AMOEBA and the ALMO-EDA scheme used in this work is summarized in Table. 1. Note that in the following discussion, we use "total electrostatics" to represent the sum of permanent electrostatic interactions and polarization (induced electrostatics); and "vdW interaction" refers to the entire 14-7 potential for AMOEBA, while for ALMO-EDA it refers to the sum of modified Pauli repulsion and dispersion.

Table 1: Correspondence of terms in AMOEBA and ALMO-EDA

Components	AMOEBA	ALMO-EDA
Permanent Electrostatics	U_{ele}^{perm}	E_{elec}^{cls}
Induced Electrostatics	U_{ele}^{ind}	E_{pol}
Pauli Repulsion	U_{vdw}^{rep}	E_{Pauli}^{mod}
Dispersion	U_{vdw}^{attr}	E_{disp}
Charge Transfer	no explicit	E_{ct}

2.3 Computational details

Energy calculations using the AMOEBA force field were performed in the Tinker7 molecular modeling package.[97] The most recently released parameters were used for all species: "amoebapro13" was used for water-water, water-cation, and water-Cl^- interactions, while

for water-F^- and water-Br^-, parameters were obtained from "amoeba09", which corresponds to the latest parameterization for these two halides.[26,27,34,38,98,99] Neither periodic boundary conditions nor distance cutoffs were adopted for any of these calculations; therefore, permanent electrostatics and polarization were performed in a standard, no-cutoff direct space interaction. Induced dipoles were converged to 10^{-12} Debye. For the calculation of the buffered 14-7 vdW potential, a trivial modification to the source code was made to allow the repulsive and attractive terms to be reported separately.

All the ALMO-EDA calculations in this work were performed with a standard version of the Q-Chem 4.4 software package.[100] The ωB97X-V functional,[77] which is a range-separated hybrid GGA that incorporates the VV10[101] non-local correlation functional for the description of dispersion, is used for modeling the intermolecular interactions. The large def2-QZVPPD[102] basis set (augmented quadruple-ζ) is employed *without* counterpoise corrections[103] for basis set superposition errors (BSSE), since for dimer interactions, the BSSE associated with the use of this basis set should be almost negligible compared to the magnitude of the investigated interactions.[104] The numerical integration of the XC functional is performed on a $(99, 590)$ grid (99 radial shells with 590 Lebedev points on each), while the SG-1 grid[105] is used to integrate the non-local correlation functional.

The frozen energy decomposition is based on the modified scheme defined by Eq. (14), using Hartree-Fock as the DF functional for the separation of dispersion. The polarization contribution is determined through the "FERF-nDQ" model, as introduced in Sec. 2.2. All the variational energy minimizations involved in ALMO-EDA (SCF on isolated fragments, SCF-MI, and supersystem SCF) are converged to 10^{-8} a.u..

For the PES scans, we start from the equilibrium geometry optimized at the ωB97X-V/def2-QZVPPD level of theory. Then, we stretch/compress the complex along one single chosen coordinate, *without* relaxing other degrees of freedom (a rigid dissociation/compression). Unless otherwise specified, for all the distance scans, the separation between the oxygen in

water and the ion ($O\cdots O$ distance for the water dimer case) is selected as the coordinate being modified, and the interval between neighboring data points is 0.05 Å. AMOEBA and ALMO-EDA results are then generated on the same set of configurations.

In order to further validate the QM model chemistry employed in this work, we compare the interaction energies of the water dimer and five water-ion dimers (at compressed, equilibrium and stretched geometries) evalutated by ωB97X-V/def2-QZVPPD and AMOEBA against the ΔCCSD(T)/CBS reference values (Table 2). More computational details for this benchmark are provided in the caption of Table 2. ωB97X-V provides sub-kcal/mol accuracy for almost all the investigated systems except for the stretched configuration of $H_2O\cdots Mg^{2+}$. ωB97X-V tends to slightly overbind the water-cation complexes in the long range due to self-interaction error, which is most pronounced in $H_2O\cdots Mg^{2+}$. The development of ALMO-EDA for wavefunction-based correlation methods[106,107] could provide useful alternatives in such scenarios. The error of ωB97X-V is typically over 5 times smaller than AMOEBA's so that we can use ALMO-EDA to trace the source of error among AMOEBA's energy components. In cases where the ratio is errors is smaller (e.g., for the equilibrium water dimer, the error of AMOEBA is even smaller than that of ωB97X-V), the interaction energy given by AMOEBA is usually fairly accurate, and ALMO-EDA can be utilized to uncover the origins of such well-behaved cases.

3 Results

The equilibrium intermolecular distances and interaction energies for all the studied dimer complexes are summarized in Table 3. Note that the AMOEBA interaction energies in Table 3 are evaluated at MM-relaxed geometries so they do *not* correspond to any points on the potential energy curves in the figures presented below. We note that the intermolecular interaction energies of $H_2O\cdots H_2O$ and $H_2O\cdots Cl^-$ given by AMOEBA turn out to be *less*

Table 2: Total interaction energies (in kJ/mol) of the water dimer and five water-ion dimers evaluated at equilibrium, compressed (10%) and stretched (10%) configurations, where the compressions and stretches are applied to the same coordinates as in the rigid PES scans. For the ΔCCSD(T)/CBS reference, core-valence correlated Dunning basis sets aug-cc-pCVTZ and aug-cc-pCVQZ[108,109] are employed for a two-point extrapolation of the MP2 correlation energies,[110] and the correction for higher-order correlation (E[CCSD(T)] - E[MP2]) is computed at the aug-cc-pCVTZ level. All the correlation energies are computed without the frozen-core approximation.

		ωB97X-V	AMOEBA	Ref.	Error(ωB97X-V)	Error(AMOEBA)
Compressed	$H_2O\cdots H_2O$	-14.31	-13.70	-14.86	0.55	1.16
	$H_2O\cdots Li^+$	-133.37	-127.34	-133.72	0.35	6.38
	$H_2O\cdots Na^+$	-89.49	-84.97	-89.81	0.32	4.84
	$H_2O\cdots Mg^{2+}$	-317.40	-298.55	-317.38	-0.02	18.83
	$H_2O\cdots F^-$	-106.43	99.68	-108.35	1.92	208.03
	$H_2O\cdots Cl^-$	-52.32	-43.93	-52.77	0.45	8.84
Equilibrium	$H_2O\cdots H_2O$	-21.07	-21.35	-21.25	0.18	-0.10
	$H_2O\cdots Li^+$	-146.27	-139.26	-145.25	-1.02	5.99
	$H_2O\cdots Na^+$	-101.69	-97.83	-100.70	-0.99	2.87
	$H_2O\cdots Mg^{2+}$	-347.79	-326.19	-344.86	-2.93	18.67
	$H_2O\cdots F^-$	-132.79	-97.18	-133.12	0.33	35.94
	$H_2O\cdots Cl^-$	-64.79	-68.91	-64.72	-0.07	-4.19
Stretched	$H_2O\cdots H_2O$	-18.53	-17.86	-18.49	-0.04	0.63
	$H_2O\cdots Li^+$	-139.40	-131.87	-137.69	-1.71	5.82
	$H_2O\cdots Na^+$	-95.85	-92.53	-94.36	-1.49	1.83
	$H_2O\cdots Mg^{2+}$	-331.64	-302.60	-326.85	-4.79	24.25
	$H_2O\cdots F^-$	-122.08	-128.31	-121.87	-0.21	-6.44
	$H_2O\cdots Cl^-$	-59.55	-63.05	-59.19	-0.36	-3.86

favorable at their AMOEBA-optimized geometries than at the QM minima (the interaction energies for the latter are given in Table 2), which, nevertheless, are compensated by the intramolecular relaxation of the involved water monomers.

Table 3: Equilibrium intermolecular separations (Å) and total interaction energies (kJ/mol) for the water dimer and eight water-ion dimers, based on geometries optimized at the ωB97X-V/def2-QZVPPD level of theory (left) and with the AMOEBA force field (right). The distance from the oxygen atom of the water molecule to the ion in each complex ($O\cdots O$ distance for the water dimer case) is reported.

	ωB97X-V		AMOEBA	
	distance	E_{int}	distance	E_{int}
$H_2O\cdots H_2O$	2.92	-21.07	2.89	-20.86
$H_2O\cdots Li^+$	1.84	-146.27	1.82	-140.67
$H_2O\cdots Na^+$	2.22	-101.69	2.23	-98.04
$H_2O\cdots K^+$	2.62	-74.10	2.60	-73.16
$H_2O\cdots Mg^{2+}$	1.91	-347.79	1.88	-332.87
$H_2O\cdots Ca^{2+}$	2.22	-242.55	2.22	-228.65
$H_2O\cdots F^-$	2.45	-132.79	2.64	-119.81
$H_2O\cdots Cl^-$	3.12	-64.79	3.15	-66.44
$H_2O\cdots Br^-$	3.31	-55.95	3.36	-55.07

3.1 The water dimer

We first assess the performance of AMOEBA for the water dimer interaction against the QM results. Since AMOEBA was initially designed as an advanced polarizable water model, we expect it to give a high-quality depiction of the PES for the water dimer, the prototypical system for water-water interactions. The top left panel of Figure 1 shows an overall good match between the potential energy curve evaluated by AMOEBA and ωB97X-V. For a rigid dissociation curve, AMOEBA and ωB97X-V predict the same $O\cdots O$ distance for the energy minimum at 2.90 Å, with a minimal discrepancy in energy: AMOEBA is slightly more bound by -0.35 kJ/mol, a difference that is close to the intrinsic error of the functional for this system. The agreement between the two potential energy curves near the equilibrium separation is more clearly demonstrated by the inset plot. In the highly compressed region

(< 2.70 Å), AMOEBA gives a distinctly harder repulsive wall. Its long-range interaction is also less attractive than the ωB97X-V profile, although the discrepancy is fairly small (the maximum deviation is about 1 kJ/mol around $R_{O\cdots O} = 3.6$ Å).

The component breakdowns of QM and AMOEBA interaction energies in the same range are plotted in the middle and bottom left panels. The permanent electrostatic component of AMOEBA is less attractive compared to its "classical electrostatics" counterpart in ALMO-EDA at short range, although they agree in the asymptotic region. By contrast close agreement for the polarization energy is achieved over all intermolecular separations except for the highly compressed region ($R_{O\cdots O} < 2.6$ Å), where the polarization energy of AMOEBA becomes less favorable than that given by ALMO-EDA due to the onset of Thole damping in the former. The profiles for total electrostatics mostly reflect the above mentioned discrepancy in permanent electrostatics.

From the bottom left plot, we see a very large discrepancy between AMOEBA's total vdW interaction and the physically pertinent terms in ALMO-EDA (modified Pauli + dispersion): AMOEBA's total vdW interaction is more favorable by about 17 kJ/mol at equilibrium, and the difference becomes much more pronounced at shorter separations. Since the attractive component of AMOEBA's vdW agrees with the dispersion term given by ALMO-EDA for most distances (except for the highly compressed region), the large difference in their total vdW interaction must reside in the repulsive part of AMOEBA's 14-7 potential which turns out to be excessively soft. This may seem surprising given the overly repulsive wall of AMOEBA's total interaction energy profile. However, the comparison to EDA shows that the softened repulsive vdW potential of AMOEBA is accounting for two attractive contributions that are important in the short range but which are not explicitly included in AMOEBA: (i) the effect of CP, which renders the QM permanent electrostatics more favorable, and (ii) the CT from the proton acceptor to the proton donor, as an extra stabilizing effect. If we combine the CT term with the total vdW interaction given by ALMO-EDA, it almost halves

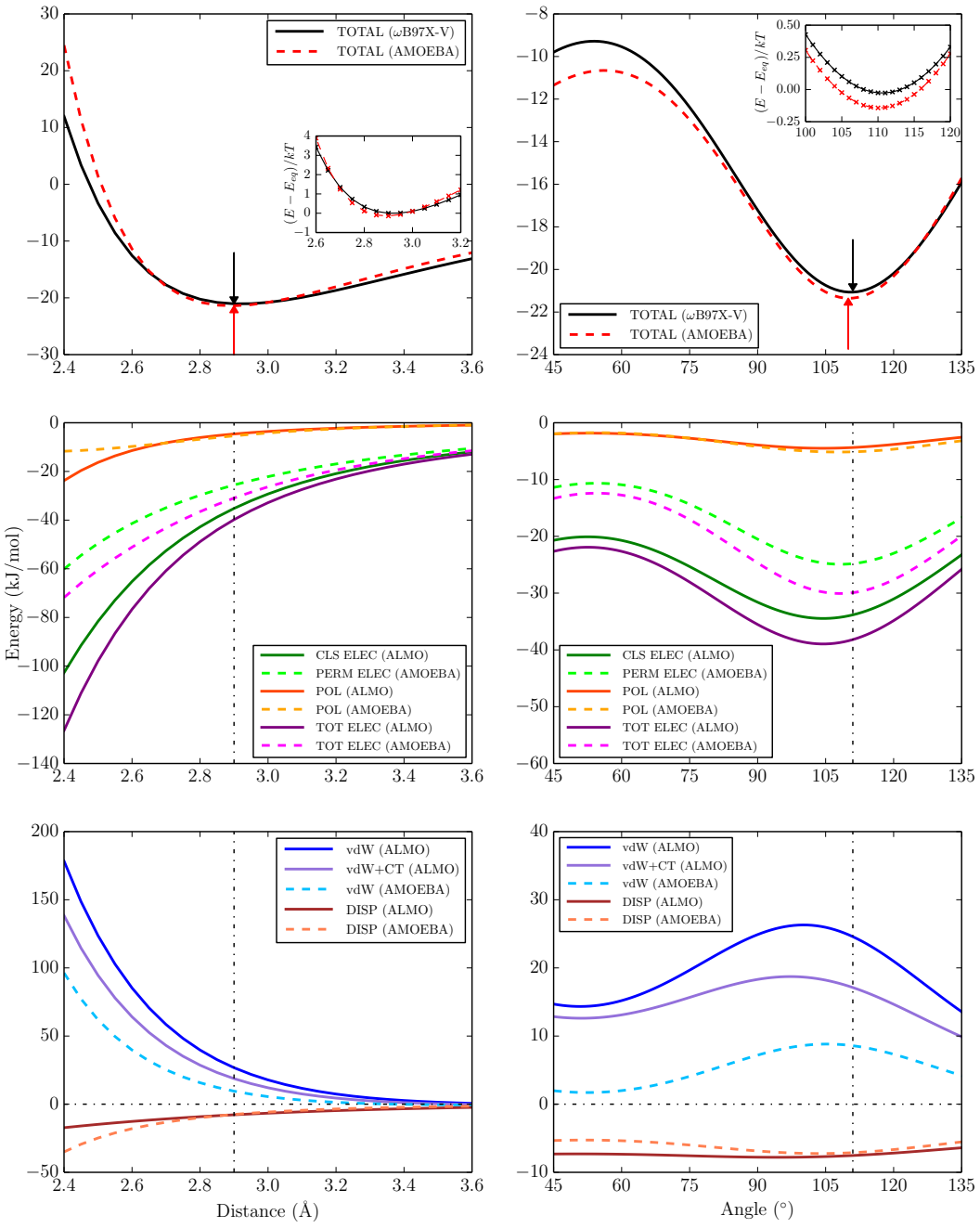

Figure 1: Distance (left) and angular (right) dependence of the total interaction energy and its breakdowns (in kJ/mol) for the water dimer. Top: total interaction energy; middle: permanent and induced electrostatics; bottom: vdW interaction. The inset plots in the two top panels show the zoomed-in near-equilibrium region in the units of kT, and the arrows indicate the location of energy minima for QM and AMOEBA interactions, while the dash-dotted lines in the lower four panels indicate the position of QM minimum.

the gap between AMOEBA and ALMO-EDA's vdW profiles, as the difference at equilibrium reduces to 9.5 kJ/mol. Thus, the energetically favorable CP and CT contributions seem to be implicitly accounted for via the softened vdW potential in AMOEBA, which results in the cancellation of errors that yields good agreement in total interaction energies. However, in the short range, the cancellation of errors turns out to be imperfect, and the extra hardness of AMOEBA's repulsive wall actually resides in its too unfavorable permanent electrostatics and polarization.

It is often deemed to be an important and challenging task for a force field to correctly reproduce the directionality of hydrogen bonds.[43,50,94] Therefore, an assessment of the angular dependence of the water dimer interaction will be instructive. The angular scan is performed by modifying the θ angle illustrated in Figure 2 at the equilibrium O$\cdots$O distance (2.92 Å) with all the other degrees of freedom fixed. The results for $\theta = 45°$–$135°$ are plotted in the right three panels. According to the ALMO-EDA results, the directionality of the hydrogen bonding interaction in the water dimer is a consequence of the interplay of permanent electrostatics, Pauli repulsion, and CT, while the angular dependence of polarization and dispersion is less appreciable. AMOEBA reproduces the angular dependence predicted by ωB97X-V fairly well in the favorable region: the energy minimum appears at 110°, which is only minimally different from the QM result (111°), and the energy discrepancy is less than 0.3 kJ/mol in the entire low-energy region of the potential well (100-120°, see the inset plot in the top right panel). In the "more exotic" higher energy region ($< 70°$), the AMOEBA curve is slightly too favorable by 1.0–1.5 kJ/mol.

Nevertheless, when turning to the energy breakdowns, we see a sharp difference between QM and AMOEBA's permanent electrostatics, while their polarization profiles exhibit very good agreement in general, which are consistent with the trend observed in the distance scan. As a result, AMOEBA's total electrostatics is too unfavorable by a considerable amount compared to the ALMO-EDA results (the maximum discrepancy is about 11 kJ/mol

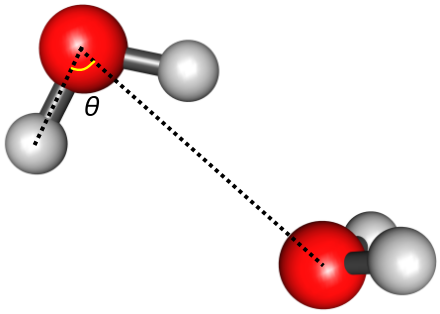

Figure 2: Illustration of the angular scan performed for the water dimer: θ is the angle being modified, while the O$\cdots$O distance remains unchanged.

at $\theta \approx 80°$), and the angle that corresponds to most favorable total electrostatics is shifted from 104° (the QM value) to 108°. We expect that such a discrepancy is offset by AMOEBA's vdW interaction in order to restore the correct angular dependence. This is confirmed by the bottom right plot: the 14-7 potential of AMOEBA is less repulsive by roughly the same amount when measured against the vdW+CT contribution determined by ALMO-EDA, and the maximum of this net repulsive energetic contribution is also shifted from 97° to 105°.

3.2 Water-monovalent ions

The same analysis is then performed on water interacting with Na^+ and Cl^-, as two representative monovalent ions. Compared to the water dimer case where both fragments are neutral, we expect to see much stronger permanent electrostatic interactions (led by charge-dipole interaction) and polarization effects in these systems. The total interaction energies calculated by ωB97X-V and AMOEBA upon rigid dissociation of the $H_2O\cdots Na^+$ complex (conserving C_{2v} symmetry) are shown in the top left panel of Figure 3. We see very good agreement between them (even in the repulsive region), although we should bear in mind that the energy scale here (increment of the y axis) is much larger than that in the water dimer case. The O$\cdots$Na$^+$ distance corresponding to AMOEBA's energy minimum (2.25 Å) matches the QM value (2.20 Å) closely, while AMOEBA slightly underbinds the complex in

the whole plotted range (by about 4 kJ/mol in the vicinity of equilibrium).

The magnitude of permanent electrostatics for this system is over 100 kJ/mol at the equilibrium separation. Due to the lack of CP in AMOEBA, its permanent electrostatics is consistently less favorable than its ALMO-EDA counterpart. The discrepancy at equilibrium is roughly 8.5 kJ/mol. They start to match to within 1 kJ/mol not far beyond that (> 2.75 Å), which validates the accuracy of the distributed-multipole scheme in describing permanent electrostatics in the long range.

The discrepancy between polarization energies evaluated by QM and AMOEBA varies with intermolecular separation. While AMOEBA only slightly overestimates the polarization energy (by about 3.5 kJ/mol) at equilibrium, it significantly overpolarizes in the short range (almost twice as favorable as ALMO-EDA's polarization in the more compressed region). This suggests inadequate damping from the Thole model. Nevertheless, at short range, AMOEBA's overestimated polarization appears to balance its too unfavorable permanent electrostatics, which results in reasonable agreement with QM for the total electrostatic contributions.

The almost superimposed curves for AMOEBA and ALMO-EDA's vdW interactions (see the bottom left panel) indicates their close agreement, despite the sharp difference seen for their attractive component (dispersion). Also, the contribution from CT is almost negligible for this system compared to the magnitude of the total interaction energy.

The performance of AMOEBA is also assessed on the $H_2O\cdots Cl^-$ complex, and the results are shown in the right three panels of Figure 3. While comparison of the two top panels suggests a more considerable difference between QM and AMOEBA's total interaction energy profiles, it should be kept in mind that the energy range plotted for this system is much smaller. In fact, the performance of AMOEBA (size of errors) in the vicinity of equilibrium is close to that in the $H_2O\cdots Na^+$ case: it shifts the energy minimum to longer distance by 0.05Å, and overestimates the equilibrium interaction energy by 4-5 kJ/mol. However,

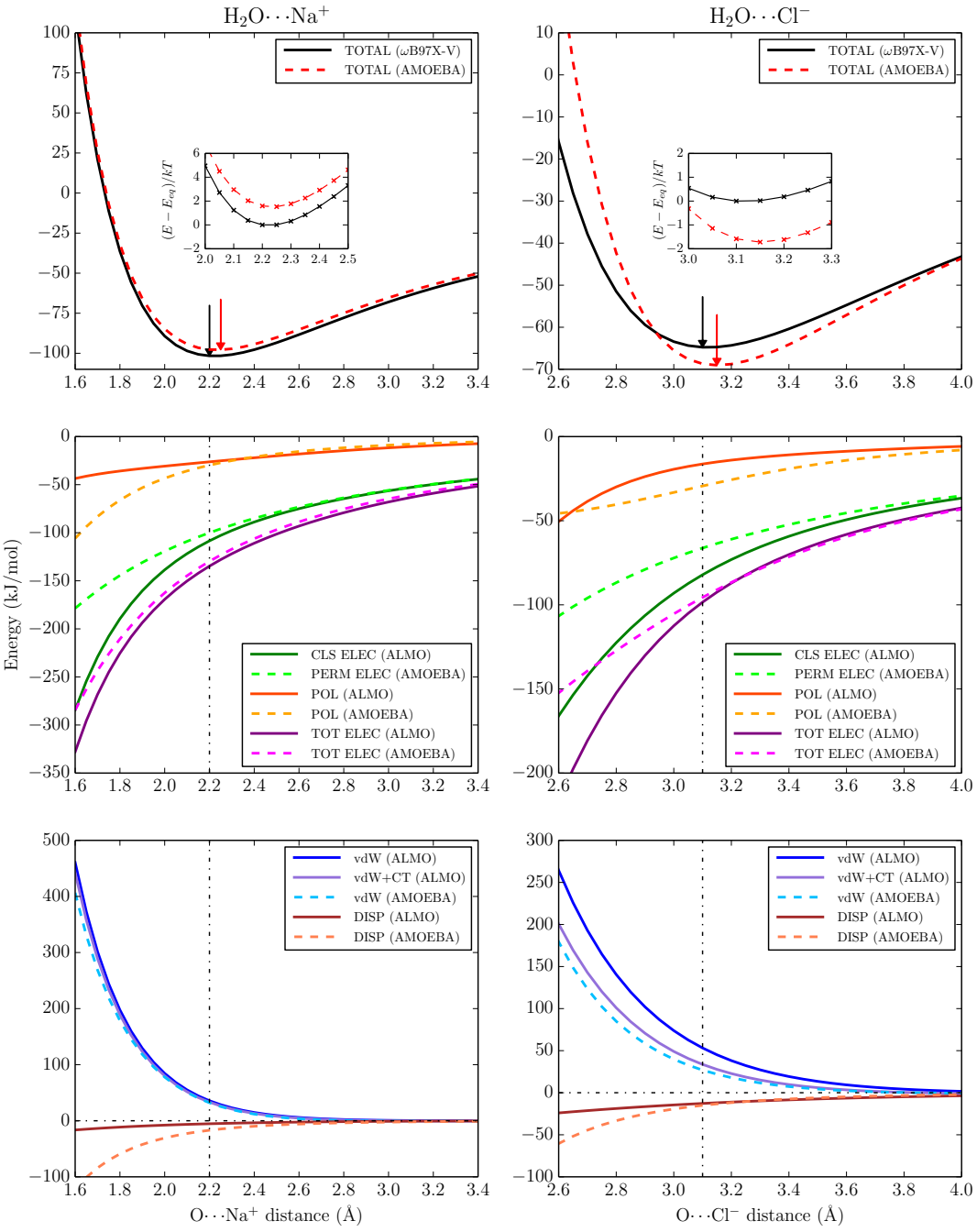

Figure 3: The total interaction energy and its breakdowns (in kJ/mol) upon a rigid dissociation of $H_2O\cdots Na^+$ and $H_2O\cdots Cl^-$ complexes. The plotting details (arrangement of figures and symbols used) are the same as in Figure 1.

the total interaction energies given by QM and AMOEBA start to differ significantly when entering the compressed region. AMOEBA becomes repulsive more rapidly for $R_{O\cdots Cl^-} < 2.9$ Å, although it is overbound adjacent to and beyond equilibrium. Also, for $R_{O\cdots Cl^-} > 4.2$ Å (not shown on the plot), AMOEBA's total interaction energy again becomes less favorable than the QM result.

One might ascribe AMOEBA's much harder repulsive wall to an improperly trained 14-7 potential. However, according to the bottom right panel of Figure 3, the 14-7 potential of AMOEBA is less repulsive than ALMO-EDA's vdW contribution across the entire plotted range. Its fairly reasonable agreement (differing by ∼7 kJ/mol at equilibrium) with ALMO-EDA's vdW+CT term indicates again that AMOEBA implicitly incorporates CT through its 14-7 potential.

Therefore, similar to the water dimer case, the deficiency of AMOEBA for the short-range interaction between H_2O and Cl^- mostly arises from the failure to fully compensate for the missing effect of CP via the 14-7 potential. The middle right plot of Figure 3 shows that AMOEBA's permanent electrostatic interaction is significantly underestimated (by about 15 kJ/mol) in the vicinity of the energy minimum. This energetic discrepancy, nonetheless, is largely canceled by its overestimated polarization energy in the same range, which leads to fairly close agreement in total electrostatics between QM and AMOEBA at and beyond equilibrium. The error at the bottom of AMOEBA's total interaction potential well is mostly due to the aforementioned slight difference (whose size is only a few kJ/mol) between the 14-7 potential of AMOEBA and ALMO-EDA's vdW+CT contribution. However, at shorter intermolecular separations, this small discrepancy in vdW interactions is overwhelmed by the error due to AMOEBA's far too unfavorable permanent electrostatics. At the same time, AMOEBA's overestimation of polarization also diminishes gradually with reduced intermolecular distance, due to the onset of Thole damping, which leaves permanent electrostatics as the culprit for the excessively hard repulsive wall in the total interaction

potential of $H_2O\cdots Cl^-$.

Comparing the results of permanent electrostatics in $H_2O\cdots Na^+$ and $H_2O\cdots Cl^-$ at their individual equilibrium separations, AMOEBA underestimates this energetic component by 8% and 19%, respectively, due to imperfect compensation for the missing effect of CP. It is not surprising that the latter system is more prone to CP, since the charge distribution of Cl^- is much more diffuse than that of Na^+.

The difference between AMOEBA's polarization profiles for these two systems at short range is also intriguing, and must arise from the effect of Thole damping through the effectively smeared point multipoles. Based on Eq. (8), atomic sites with higher polarizability are more smeared in the damping procedure. The polarizability value for Cl^- is 4.00 $Å^{-3}$, which is one of the largest amongst all atom types available in the AMOEBA force field. For $H_2O\cdots Cl^-$, although AMOEBA's polarization energy is more favorable than ALMO-EDA's over almost the entire plotted range, the onset of Thole damping is clearly manifested in the curvature in the AMOEBA polarization profile. Indeed there is an inflection point near the equilibrium distance, and therefore AMOEBA polarization crosses with ALMO-EDA's polarization curve at $R_{O\cdots Cl^-} = 2.65$ Å. In contrast, the onset of damping is not apparent in $H_2O\cdots Na^+$, for which AMOEBA polarization due to the cation is significantly overestimated in the short range. This is due to the considerably smaller atomic polarizability of Na^+ (0.12 $Å^{-3}$), rendering the damping effect through Eq. (8) negligible unless Na^+ and H_2O are in extremely close contact.

3.3 Water-divalent cations

It is well-known that the description of water-divalent cation interactions is challenging for classical force fields.[111,112] We next assess the agreement between AMOEBA and ωB97X-V results upon rigid dissociation of $H_2O\cdots Mg^{2+}$ and $H_2O\cdots Ca^{2+}$ complexes (with C_{2v} symmetry). The results are collected in Figure 4. For both systems, AMOEBA correctly reproduces

the position of energy minima, while it underestimates the magnitude of the binding energies across the board. At the equilibrium distance, AMOEBA underbinds $H_2O\cdots Mg^{2+}$ and $H_2O\cdots Ca^{2+}$ by 21 kJ/mol and 17 kJ/mol, respectively, which is 6–7% of the total interaction energies evaluated by ωB97X-V.

To understand the much larger deviations compared to that in the $H_2O\cdots Na^+$ case, we again measure the individual terms of AMOEBA against the energy components given by ALMO-EDA. Due to the +2 charge, permanent and induced electrostatics dominate the strongly favorable total interactions. Some qualitative similarities emerge for these two systems, as seen in the middle two panels of Figure 4. First is the consistently less favorable short-range permanent electrostatics of AMOEBA, which is a common issue for all the systems assessed so far. Second is the crossing of the AMOEBA and ALMO-EDA polarization energy curves slightly beyond the equilibrium distance.

The permanent electrostatic interactions in these two systems are *identically* described by the AMOEBA model (+2 point monopole for both Mg^{2+} and Ca^{2+}). However, the short-range discrepancy from ALMO-EDA's classical electrostatics is larger in $H_2O\cdots Ca^{2+}$, indicating a more pronounced CP effect. The CP effect (regarded as the difference between these two models) is determined by (i) the extent of charge distribution on each fragment and (ii) the intermolecular distance. The former should be the dominant factor here, since otherwise we would expect the more significant CP effect in $H_2O\cdots Mg^{2+}$ whose equilibrium distance is 0.3 Å shorter. This trend (and interpretation) also applies to the interactions between water and alkali metal cations (results for $H_2O\cdots Li^+$ and $H_2O\cdots K^+$ are shown in Figure S1 of the Supplementary Material), where the effect of CP increases as $Li^+ < Na^+ < K^+$.

The more diffuse charge distribution in $H_2O\cdots Ca^{2+}$ is also be reflected in the (modified) Pauli term in ALMO-EDA. At equilibrium, the Pauli term for $H_2O\cdots Ca^{2+}$ is 16 kJ/mol more repulsive than that for $H_2O\cdots Mg^{2+}$ despite the 0.3 Å longer intermolecular separation

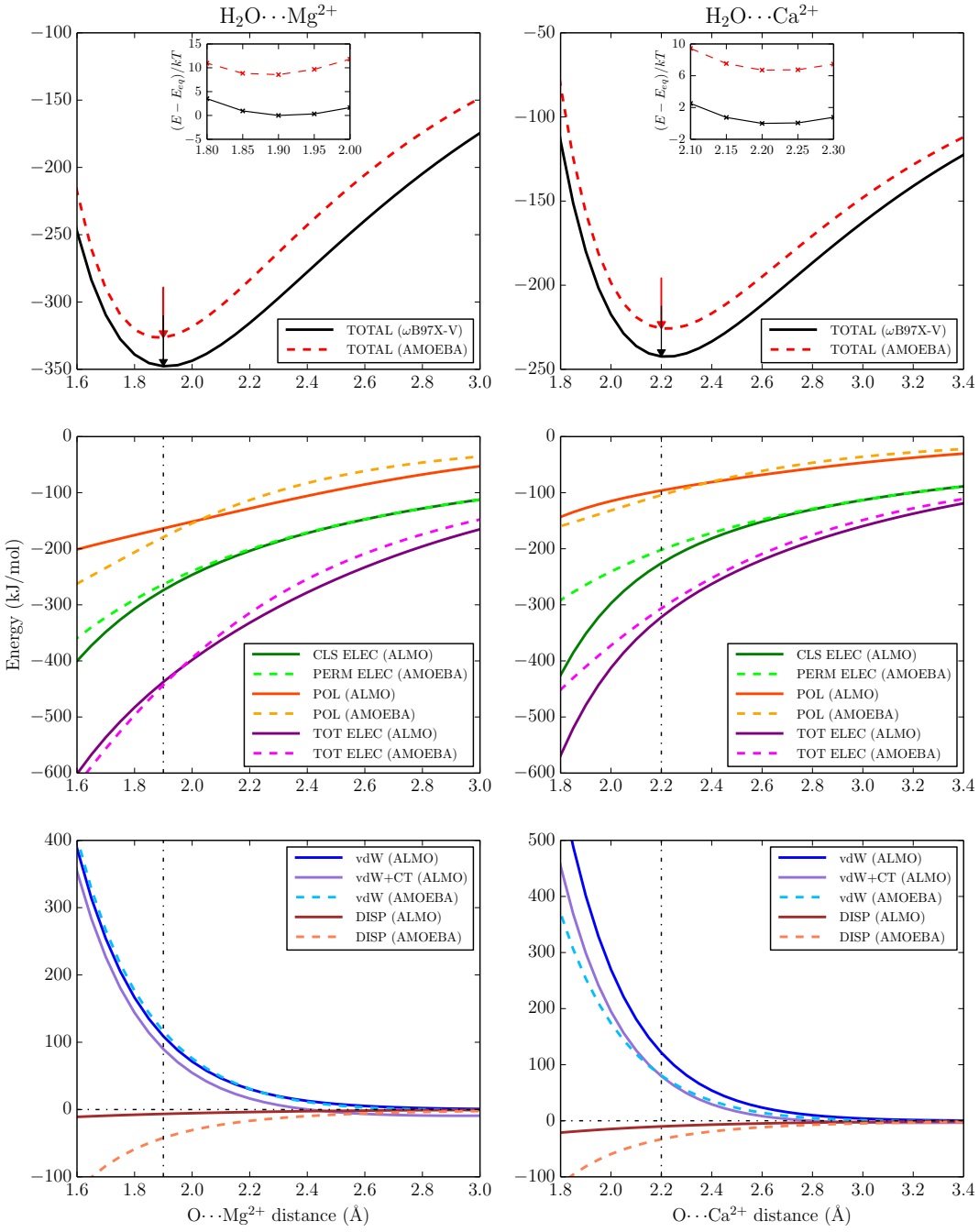

Figure 4: The total interaction energy and its breakdowns (in kJ/mol) upon a rigid dissociation of $H_2O\cdots Mg^{2+}$ and $H_2O\cdots Ca^{2+}$ complexes. The plotting details (arrangement of figures and symbols used) are the same as in Figure 1.

of the former, indicating stronger overlap between fragment charge densities resulting from the much more diffuse charge distribution of Ca^{2+}.

The polarization energies given by AMOEBA for these two systems, at their equilibrium intermolecular separations, are slightly overestimated (by roughly 9%). This order is reversed (AMOEBA polarization becomes too weak) in the long range. In the compressed region plotted in Figure 4, AMOEBA overpolarizes in both cases. However, one nuanced distinction exists: in $H_2O\cdots Mg^{2+}$, the AMOEBA over-polarization increases monotonically with reducing $O\cdots Mg^{2+}$ distance, while in the $H_2O\cdots Ca^{2+}$ case, the difference reaches a maximum at 1.90 Å (18.5 kJ/mol), and a crossing point emerges at more compressed distance (1.65 Å, not shown in the figure). This behavior is similar to what was discussed previously for $H_2O\cdots Cl^-$, which was regarded as a signature of the onset of Thole damping. Here it can also be explained by the same polarizability-based argument. The AMOEBA atomic polarizability for Mg^{2+} (0.08 $Å^{-3}$) is so small that the effect of damping can hardly be seen in the entire plotted range. By contrast, the value for Ca^{2+} is significantly larger (0.55 $Å^{-3}$) so that the damping effect is manifested in the slightly compressed region. The same trend is observed with the polarization energy curves of water-alkali metal cation series (the results for Li^+ and K^+ are shown in Figure S1 of the Supplementary Material), where AMOEBA immensely overestimates the polarization energy of $H_2O\cdots Li^+$ in the short range, while the agreement with ALMO-EDA's polarization is much better for $H_2O\cdots K^+$.

In terms of total electrostatics, AMOEBA agrees closely with ALMO-EDA near the equilibrium distance in the $H_2O\cdots Mg^{2+}$ case (the former is more favorable by 5 kJ/mol at equilibrium), thanks to error cancellation between permanent electrostatics and polarization. The deviations in the short range and long range have opposite signs, which, in contrast to many other systems, are both dominated by the discrepancy in polarization. For $H_2O\cdots Ca^{2+}$, AMOEBA's total electrostatics is less favorable across the entire range, and the discrepancy at equilibrium is about 3 times larger than that in the $H_2O\cdots Mg^{2+}$

case, as AMOEBA underestimates the permanent electrostatic interaction more severely in $H_2O\cdots Ca^{2+}$ due to the increased importance of CP.

Turning to the vdW terms, a remarkable common feature of the bottom two panels of Figure 4 is the much more attractive $1/R^7$ potential in AMOEBA relative to the dispersion component determined by ALMO-EDA. Specifically, at the equilibrium distance of each system, AMOEBA's dispersion (attractive vdW potential) is 6.5 times more favorable in $H_2O\cdots Mg^{2+}$, and 3 times more favorable in $H_2O\cdots Ca^{2+}$. It is inevitable for the $1/R^7$ potential of AMOEBA to be excessively attractive in the strongly overlapping regime since it is not appropriately damped, however, we think that it should be able to approximately match the dispersion energy given by ALMO-EDA near the equilibrium region. Here (and in $H_2O\cdots Li^+$ and $H_2O\cdots Na^+$ as well) the large difference against ALMO-EDA's dispersion term could be accounted for by the incorporation of other stabilizing effects like CP or CT in AMOEBA's $1/R^7$ potential, as we discussed before. On the other hand, it is possible that such differences are related to the fact that AMOEBA's 14-7 potential is parameterized simultaneously so that its repulsive and attractive components may not correspond to their presumed physical meanings when scrutinized individually.

For the total vdW interaction between H_2O and Mg^{2+}, AMOEBA exhibits a slightly more repulsive potential than its ALMO-EDA counterpart when $R_{O\cdots Mg^{2+}} < 2.2$ Å. The inclusion of CT further enlarges the discrepancy, which is the opposite of the general trend observed in other systems. The plot for $H_2O\cdots Ca^{2+}$, on the other hand, demonstrates the common trend, where AMOEBA's total 14-7 potential is considerably softened and matches ALMO-EDA's vdW+CT curve fairly well except in the strongly unfavorable region. It should be noted that the contribution of CT is appreciable when water interacts with these divalent cations, especially for $H_2O\cdots Ca^{2+}$ where the CT energy at equilibrium is -42 kJ/mol (roughly 1/6 of the total interaction energy). According to our results, this significant stabilizing energy component is implicitly incorporated in the softened 14-7 potential of

AMOEBA.

So is there an overall reason for why AMOEBA underbinds these two systems? We suggest, based on the assessment against the ALMO-EDA results, that for $H_2O\cdots Mg^{2+}$ the error mostly resides in AMOEBA's overly repulsive vdW potential. On the other hand, AMOEBA's too unfavorable permanent electrostatics (due to the lack of CP) is the main origin of the underestimated interaction energy between H_2O and Ca^{2+}. Therefore, although these two divalent cations belong to the same main group on the periodic table, and AMOEBA manifests rather similar systematic errors on them, we think that they actually arise for different reasons.

3.4 Other water-halide interactions

Although AMOEBA underestimates the total $H_2O\cdots Mg^{2+}$ and $H_2O\cdots Ca^{2+}$ interaction energies, it gives a reasonable description of the shape of their PESs, which is rather important for accurately computing the intermolecular forces. In contrast, the shape of AMOEBA's total interaction energy profile for $H_2O\cdots Cl^-$ agrees less satisfactorily with that generated by QM. Therefore, we complete the full series of water-halide interactions by assessing the performance of AMOEBA on $H_2O\cdots F^-$ and $H_2O\cdots Br^-$ against QM, and the results are shown in Figure 5.

As for $H_2O\cdots Cl^-$, AMOEBA gives overly repulsive walls for water-halide interactions in the short range, and also exhibits overly long intermolecular distances. The discrepancy between QM and AMOEBA is most exaggerated in the $H_2O\cdots F^-$ case, where the AMOEBA equilibrium $O\cdots F^-$ distance is 0.25 Å too long. At the distance corresponding to the QM minimum ($R_{O\cdots F^-} = 2.45$ Å), AMOEBA is underbound by about 35 kJ/mol. On the other hand, the total interaction energy profile of AMOEBA for $H_2O\cdots Br^-$ exhibits features that are rather similar to the $H_2O\cdots Cl^-$ case: the equilibrium distance is slightly overestimated by 0.05 Å, and the energy discrepancy at $R_{O\cdots Br^-} = 3.3$ Å (the QM minimum) is rather

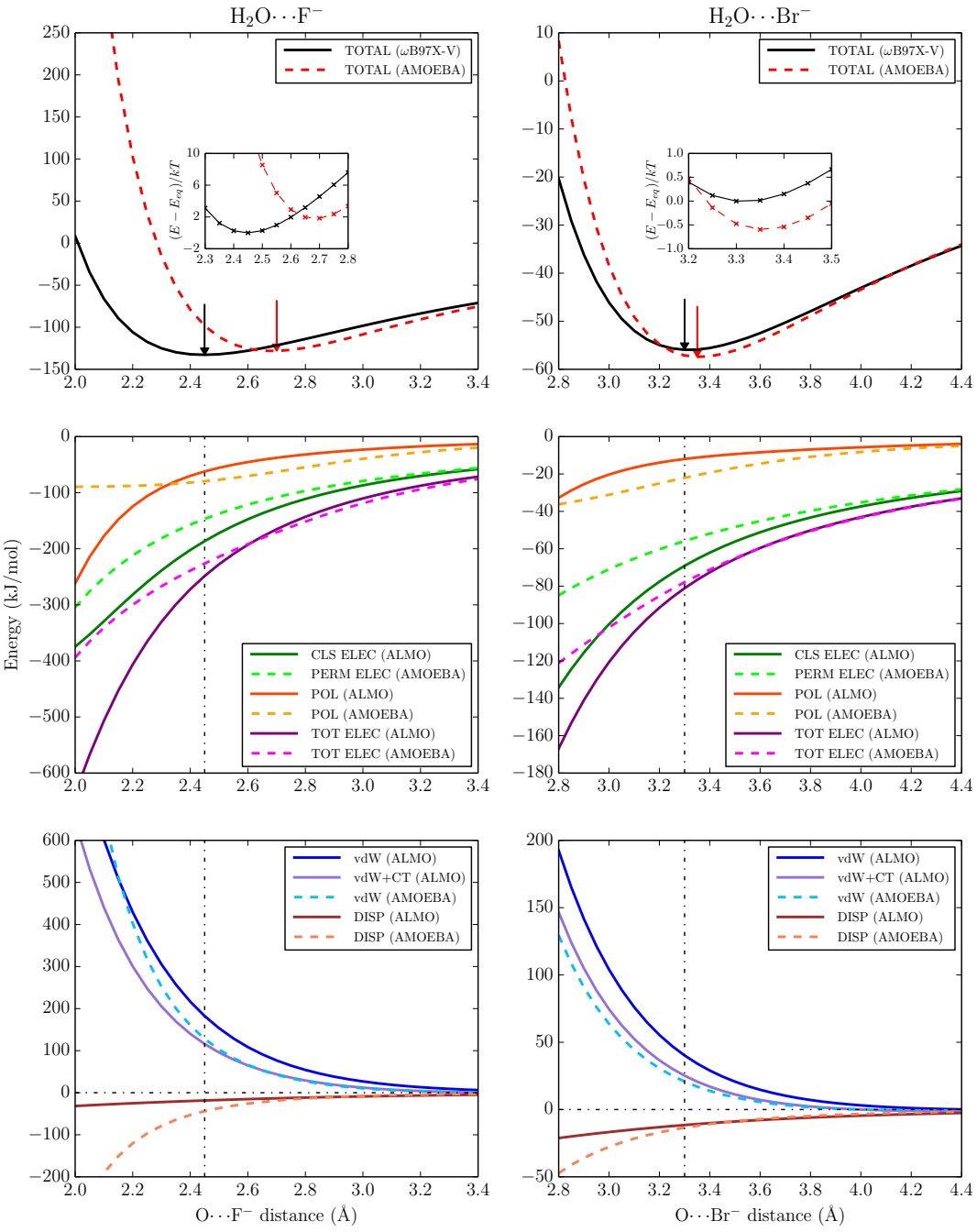

Figure 5: The total interaction energy and its breakdowns (in kJ/mol) upon a rigid dissociation of $H_2O\cdots F^-$ and $H_2O\cdots Br^-$ complexes. The plotting details (arrangement of figures and symbols) are the same as in Figure 1.

small (about 1 kJ/mol).

Based on the ALMO-EDA results, we investigate the reason for such a tremendous discrepancy between QM and AMOEBA's total interaction energy profiles for $H_2O\cdots F^-$. According to the middle left panel of Figure 5, at the QM energy minimum, the permanent electrostatic interaction given by AMOEBA is less favorable than ALMO-EDA's classical electrostatics by over 20% (40 kJ/mol). Due to the diffuse charge distribution of F^- and the relatively small equilibrium distance, the effect of CP should be the main origin of such a difference. Nevertheless, AMOEBA's too unfavorable permanent electrostatics here is not qualitatively different from the situations in other water-halide interactions: e.g., for $H_2O\cdots Cl^-$, AMOEBA underestimates the permanent electrostatic interaction by roughly 20% as well at equilibrium and by over 1/3 in the most compressed region. Therefore, the neglect of CP should *not* be the only culprit for AMOEBA's poor performance on $H_2O\cdots F^-$ *overall*.

As in $H_2O\cdots Cl^-$, AMOEBA's polarization energy is more favorable than that given by ALMO-EDA at and beyond the QM minimum. However, it seems to be damped too quickly after entering the compressed region. Specifically, the AMOEBA polarization curve is almost flat when $R_{O\cdots F^-} < 2.45$ Å, and it becomes less favorable than ALMO-EDA's polarization energy when $R_{O\cdots F^-} \leq 2.3$ Å. The overdamping of AMOEBA's polarization at short range is essentially similar to what we observed in other water-halide interactions, but it is far more pronounced in this system, most likely due to the much shorter interfragment distance. Therefore, the permanent and induced electrostatics of AMOEBA are *both* too unfavorable in the short range, contributing to the striking difference between AMOEBA and ALMO-EDA's total electrostatic contributions together. It is also noteworthy that AMOEBA's total electrostatics turns out to be more favorable than its ALMO-EDA counterpart when $R_{O\cdots F^-} > 2.6$ Å (the largest difference beyond that point is about 6 kJ/mol), where the less attractive permanent electrostatics of AMOEBA is outweighed by its more favorable

polarization energy, which is also observed in $H_2O\cdots Cl^-$ and $H_2O\cdots Br^-$ systems (though to a lesser extent).

In the previous discussion of the $H_2O\cdots Cl^-$ complex, we ascribed the overly repulsive AMOEBA potential wall to its too unfavorable permanent electrostatics. While AMOEBA's vdW interaction appears to be softer than ALMO-EDA's vdW+CT contribution, it partially compensates for this deficiency in electrostatics. According to the two bottom panels of Figure 5, a similar trend emerges for $H_2O\cdots Br^-$, and it again leads to the reasonable agreement between QM and AMOEBA's total interaction energies, at least near the bottom of the potential well.

However, the situation is rather different for $H_2O\cdots F^-$. The AMOEBA vdW potential matches ALMO-EDA's vdW+CT contribution closely at and beyond the QM equilibrium. However, the repulsion rises too rapidly at compressed $O\cdots F^-$ distances. For instance, at 2.2 Å (0.25 Å shorter than the (QM) equilibrium distance), the vdW interaction of AMOEBA is already 70 kJ/mol more unfavorable than the ALMO-EDA's vdW+CT contribution. Therefore, based on the analysis above, AMOEBA's total electrostatic and vdW interactions are both too unfavorable in the short range. These two errors accumulate, rather than cancel each other, which result in AMOEBA's very poor description of the total PES for the $H_2O\cdots F^-$ complex.

4 Discussion

In comparing the breakdown of AMOEBA's classical molecular interactions against the energy decomposed QM results for the water-water and the water-monovalent cation interactions, it is apparent that the observed good agreement in total interaction energy rests on a very delicately balanced cancellation of errors, as would be expected for most if not all empirical force fields. In the compressed region, where AMOEBA exhibits polarization that is

typically too favorable until the onset of Thole damping is realized, it is overwhelmed by the excessively unfavorable permanent electrostatics that originates in the failure of the point multipole model to account for charge penetration, an important stabilizing effect which becomes more pronounced at short range. The fact that the buffered 14-7 potential contains short-range softening effects to implicitly compensate for the electrostatics is perhaps not surprising given the parameterization procedure of AMOEBA wherein the permanent multipoles and polarization parameters are determined first and all the remaining non-covalent energetic effects are folded into the 14-7 potential.[26,27]

For the water-halide cases where agreement in total interaction energies between QM and AMOEBA are poorer, the cancellation of errors is insufficient for two reasons. First, too much is asked of the limited functional form of the 14-7 vdW potential to account for favorable CP and CT effects at short range, which are each more important in anions than in cations and to account for permanent electrostatics that are excessively unfavorable at short range. Second, for halides, the AMOEBA polarization suffers from overdamping in roughly the same region, which is a consequence of the exclusive functional dependency of the Thole damping on the atomic polarizabilities (Eq. (8)), whose values are very large for the halides.

For the water-divalent cation systems, the total interaction energies in AMOEBA are significantly underestimated compared to QM throughout the distance range. For $H_2O\cdots Ca^{2+}$, the breakdown of the interaction energy into individual components exhibit the same problems observed for the water-halide systems. However, it may also be exacerbated by the fact that the original fit of the AMOEBA divalent cation parameters was to a somewhat limited QM benchmark, where the counterpoise corrections for BSSE were performed with insufficiently large basis sets,[35,98] which may undershoot the correct interaction energies of these systems. Interestingly, $H_2O\cdots Mg^{2+}$ shows another difference in its 14-7 potential, which is even more unfavorable than the sum of ALMO-EDA's Pauli repulsion and dispersion so it

evidently cannot implicitly incorporate any CT. In addition, its permanent electrostatics at short range is less unfavorable versus the $H_2O\cdots Ca^{2+}$ case, mostly due to the comparatively more compact Mg^{2+} charge distribution.

Although errors in AMOEBA's polarization contribute less significantly to errors in total interaction energy in general, it clearly deviates from the ALMO-EDA polarization profiles in all cases. Taken together, these point to either a shortcoming in the functional form given by Eq. (8), or the way in which the damping parameters are determined. In the original Thole model, atomic polarizabilities and the damping parameter were determined by a fit to molecular polarizabilities.[84,85] However, molecular polarizability is *not* very sensitive to either the value of the damping parameter or even the functional form of the damping, since a linear model performs as well as the currently used exponential model.[84] On the other hand, it has been found that the same "dimensionless width" parameter, a, in Eq. (8) cannot simultaneously reproduce the gas-phase molecular polarizability of water and QM energies of small clusters of water molecules, which indicates the higher sensitivity of polarization energetics to the choice of this value.[27,32]

It has been suggested that the need to modify the damping parameter, is a manifestation of exchange-polarization coupling between molecules that occurs in the regime of overlapping charge distribution.[48,113,114] Such an effect is intrinsically accounted for in ALMO-EDA, since its polarization term is computed using a properly antisymmetrized wavefunction. Therefore, the discrepancies in polarization energies between ALMO-EDA and AMOEBA observed in this work may also be related to this missing effect in AMOEBA. It is noteworthy that only a single value of the damping parameter, a, is employed for almost every type of interaction in AMOEBA (except for aromatic carbon atoms and divalent cations[27]), and therefore the damping may benefit from greater chemical specificity in choosing this parameter in order to capture differences in diffuseness of atomic charge distributions, as well as the exchange-polarization coupling effect at short range. Moreover, a similar argument to incorporate

greater chemical specificity may also apply to atomic polarizabilities, since by simply reproducing the isolated molecular polarizabilities within the current parameterization, the resulting polarizability values may not yield correct polarization energetics under more complicated interacting environment.

An additional potential limitation of the AMOEBA functional form may reside in its reliance on isotropic polarizabilities, which can be seen in the angular scans for water interacting with alkali metal cations that show larger discrepancies between the AMOEBA and ALMO-EDA polarization energies in the strained region (see Figures S3 and S4 in the Supplementary Material), due to the presence of stronger electrostatic fields. This points to two aspects of the anisotropy of polarization. The first is the issue of whether additional polarization sites other than the atom centers are needed to faithfully reproduce properties determined by QM. The polarization profiles of AMOEBA seem to overemphasize the angular dependence, which might be related to the lack of extra inducible sites around the water oxygen. Indeed, it has been shown that MM models that incorporate polarization centers at the lone pair sites of the water oxygen in addition to the atom centers more faithfully reproduce QM-derived energies and dipole moments for several sets of water oligomers than models that use atom-centered polarization only.[115]

The second aspect of the anisotropy is the replacement of an isotropic scalar polarizability with a polarizability tensor. Harder et al. examined both of the major aspects of polarization anisotropy in the context of the Drude oscillator model.[116] They observed that the use of both lone pair polarizability sites and anisotropic polarizabilities is essential to faithfully reproduce the QM electrostatic potentials along a curvilinear coordinate, compared with using isotropic polarizabilities either with or without lone-pair polarizability sites. In terms of the energetics of oligomeric systems of small organic molecules in water clusters, the use of lone pair polarization sites turns out to be the more important aspect, and the additional use of an anisotropic polarizability tensor further reduces the error with respect to the QM

results.

5 Conclusion

In this work, we have compared the energetic profiles of water-water and water-ion dimers generated by an advanced MM force field, AMOEBA, and those obtained with the ωB97X-V density functional whose accuracy was verified by a comparison to the highly accurate ΔCCSD(T)/CBS reference values. More importantly, we have appraised the force field by comparing the relative contributions of its non-covalent terms with corresponding terms generated by a decomposition of the DFT total interaction energy using the second generation of the ALMO-EDA.

Taken together, a number of trends have emerged. It is clear that the physical effects that are missing in AMOEBA, specifically charge penetration (CP) and charge transfer (CT), have been captured implicitly in the 14-7 vdW potential. This is a consequence of both the short-range buffering that perhaps renders it amenable to capturing such short-range softening (stabilizing) effects, and the fact that the vdW parameters are determined at the end of the parameterization of the non-covalent terms in AMOEBA to match the binding energies given by QM reference. However, it is also clear that this implicit accounting of short-ranged softening effects by the 14-7 potential is imperfect, as revealed by the investigation on water-divalent cation and water-halide interactions. In addition, the Thole-damping of AMOEBA's polarization was sometimes found to yield unphysical results at short range (underdamped for water-cations while overdamped for water-anions, in general), which is related to the exclusive dependence of the damping effect on atomic polarizabilities. Perhaps the ALMO-EDA (or related methods) can help refine a next generation AMOEBA model that realizes a better cancellation of errors for the problematic ion-water cases.

Alternatively, the ALMO-EDA could also be used to guide the development of explicit

functional forms and the associated parameterization for short-range effects like CP and CT. Indeed, efforts are underway by others to explicitly account for the effect of CP under the framework of the AMOEBA force field,[8,9] and simply adding corrections for CP to the monopole-monopole term in AMOEBA's permanent electrostatics has been shown to yield considerably improved agreement with the permanent electrostatics given by the SAPT2+ level of theory.[8] Nonetheless, to yield a balanced force field, AMOEBA's 14-7 potential would need to be reparameterized in the context of explicit inclusion of CP to avoid overcounting this effect. Similarly, any future effort to incorporate CT explicitly would require reparameterization of the 14-7 potential as well. The separation of these short-range softening effects from the vdW potential has benefits. In particular, its repulsive and attractive components could be parameterized individually according to their accepted physical meaning (i.e., Pauli repulsion and dispersion, respectively) which have precise definitions within QM methods such as the ALMO-EDA.

6 Associated Content

- **Supporting Information** Numerical comparison between the ALMO-EDA and DFT-SAPT results for the linear and bifurcated water dimers, plots of the distance scan PESs of water-Li^+, K^+ and the angular scan PESs for all the water-ion dimers studied in this work, and optimized structures for all the investigated water-water and water-ion dimers are provided. These materials are available free of charge via the Internet at http://pubs.acs.org.

7 Author information

O.D. and Y.M contributed equally to this work.

8 Acknowledgments

This work was supported by the National Science Foundation. THG acknowledges CHE-1363320 and MHG acknowledges CHE-1363342.

References

(1) Stone, A. J. *The Theory of Intermolecular Forces*, 2nd ed.; Oxford University Press: Oxford, 2013; p 285.

(2) Day, P. N.; Jensen, J. H.; Gordon, M. S.; Webb, S. P.; Stevens, W. J.; Krauss, M.; Garmer, D.; Basch, H.; Cohen, D. *J. Chem. Phys.* **1996**, *105*, 1968–1986.

(3) Freitag, M. A.; Gordon, M. S.; Jensen, J. H.; Stevens, W. J. *J. Chem. Phys.* **2000**, *112*, 7300–7306.

(4) Slipchenko, L.; Gordon, M. S. *J. Comput. Chem.* **2007**, *28*, 276–291.

(5) Piquemal, J.-P.; Gresh, N.; Giessner-Prettre, C. *J. Phys. Chem. A* **2003**, *107*, 10353–10359.

(6) Stone, A. J. *J. Phys. Chem. A* **2011**, *115*, 7017–7027.

(7) Wang, B.; Truhlar, D. G. *J. Chem. Theory Comput.* **2014**, *10*, 4480–4487.

(8) Wang, Q.; Rackers, J. A.; He, C.; Qi, R.; Narth, C.; Lagardere, L.; Gresh, N.; Ponder, J. W.; Piquemal, J.-P.; Ren, P. *J. Chem. Theory Comput.* **2015**, *11*, 2609–2618.

(9) Narth, C.; Lagardere, L.; Polack, E.; Gresh, N.; Wang, Q.; Bell, D. R.; Rackers, J. A.; Ponder, J. W.; Ren, P. Y.; Piquemal, J.-P. *J. Comput. Chem.* **2016**, *37*, 494–506.

(10) Lee, A. J.; Rick, S. W. *J. Chem. Phys.* **2011**, *134*, 184507.

(11) Soniat, M.; Rick, S. W. *J. Chem. Phys.* **2012**, *137*, 044511.

(12) Soniat, M.; Rick, S. W. *J. Chem. Phys.* **2014**, *140*, 184703.

(13) Soniat, M.; Rick, S. W. *J. Chem. Theory Comput.* **2015**, *11*, 1658–1667.

(14) Soniat, M.; Kumar, R.; Rick, S. W. *J. Chem. Phys.* **2015**, *143*, 044702.

(15) Soniat, M.; Pool, G.; Franklin, L.; Rick, S. W. *Fluid Phase Equilibria* **2016**, *407*, 31–38.

(16) Chelli, R.; Pagliai, M.; Procacci, P.; Cardini, G. *J. Chem. Phys.* **2005**, *122*, 074504.

(17) Piquemal, J.-P.; Chevreau, H.; Gresh, N. *J. Chem. Theory Comput.* **2007**, *3*, 824–837.

(18) Millot, C.; Soetens, J.-C.; Costa, M. T. C. M.; Hodges, M. P.; Stone, A. J. *J. Phys. Chem. A* **1998**, *102*, 754–770.

(19) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269–10280.

(20) Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chem. Rev.* **1994**, *94*, 1887–1930.

(21) Chalasinski, G.; Szczesniak, M. M. *Chem. Rev.* **1994**, *94*, 1723–1765.

(22) Hohenstein, E. G.; Sherrill, C. D. *WIREs Comput. Mol. Sci.* **2012**, *2*, 304–326.

(23) Jorgensen, W. L. *J. Phys. Chem.* **1983**, *87*, 5304–5314.

(24) Kaminski, G.; Duffy, E. M.; Matsui, T.; Jorgensen, W. L. *J. Phys. Chem.* **1994**, *98*, 13077–13082.

(25) Yin, D.; MacKerell Jr., A. D. *J. Comput. Chem.* **1998**, *19*, 334–348.

(26) Ren, P.; Ponder, J. W. *J. Phys. Chem. B* **2003**, *107*, 5933–5947.

(27) Ren, P.; Wu, C.; Ponder, J. W. *J. Chem. Theory Comput.* **2011**, *7*, 3143–3161.

(28) MacKerell Jr., A. D. In *Computational Biochemistry and Biophysics*; Becker, O. M., MacKerell Jr., A. D., Roux, B., Watanabe, M., Eds.; Marcel Dekker, Inc.: New York, New York, U.S.A., 2001; Chapter 2, pp 7–38.

(29) Morokuma, K. *Acc. Chem. Res.* **1977**, *10*, 294–300.

(30) von Hopffgarten, M.; Frenking, G. *WIREs Comput. Mol. Sci.* **2012**, *2*, 43–62.

(31) Phipps, M. J. S.; Fox, T.; Tautermann, C. S.; Skylaris, C.-K. *Chem. Soc. Rev.* **2015**, *44*, 3177–3211.

(32) Ren, P.; Ponder, J. W. *J. Comput. Chem.* **2002**, *23*, 1497–1506.

(33) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. *J. Phys. Chem. B* **2010**, *114*, 2549–2564.

(34) Grossfield, A.; Ren, P.; Ponder, J. W. *J. Am. Chem. Soc.* **2003**, *125*, 15671–15682.

(35) Piquemal, J.-P.; Perera, L.; Cisneros, G. A.; Ren, P.; Pedersen, L. G.; Darden, T. A. *J. Chem. Phys.* **2006**, *125*, 054511.

(36) Xiang, J. Y.; Ponder, J. W. *J. Comput. Chem.* **2013**, *34*, 739–749.

(37) Xiang, J. Y.; Ponder, J. W. *J. Chem. Theory Comput.* **2013**, *10*, 298–311.

(38) Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Ponder, J. W.; Ren, P. *J. Chem. Theory Comput.* **2013**, *9*, 4046–4063.

(39) Kumar, R.; Wang, F.-F.; Jenness, G. R.; Jordan, K. D. *J. Chem. Phys.* **2010**, *132*, 014309.

(40) Tafipolsky, M.; Engels, B. *J. Chem. Theory Comput.* **2011**, *7*, 1791–1803.

(41) Ansorg, K.; Tafipolsky, M.; Engels, B. *J. Phys. Chem. B* **2013**, *117*, 10093–10102.

(42) McDaniel, J. G.; Schmidt, J. *J. Phys. Chem. A* **2013**, *117*, 2053–2066.

(43) Tafipolsky, M.; Ansorg, K. *J. Chem. Theory Comput.* **2016**, *12*, 1267–1279.

(44) Van Vleet, M. J.; Misquitta, A. J.; Stone, A. J.; Schmidt, J. R. *J. Chem. Theory Comput.* **2016**, *12*, 3851–3870.

(45) Misquitta, A. J.; Stone, A. J. *J. Chem. Theory Comput.* **2016**,

(46) McDaniel, J. G.; Schmidt, J. *Annu. Rev. Phys. Chem.* **2016**, *67*, 467–488.

(47) Piquemal, J.-P.; Cisneros, G. A.; Reinhardt, P.; Gresh, N.; Darden, T. A. *J. Chem. Phys.* **2006**, *124*, 104101.

(48) Gresh, N.; Cisneros, G. A.; Darden, T. A.; Piquemal, J.-P. *J. Chem. Theory Comput.* **2007**, *3*, 1960–1986.

(49) Khaliullin, R. Z.; Cobar, E. A.; Lochan, R. C.; Bell, A. T.; Head-Gordon, M. *J. Phys. Chem. A* **2007**, *111*, 8753–8765.

(50) Lu, Z.; Zhou, N.; Wu, Q.; Zhang, Y. *J. Chem. Theory Comput.* **2011**, *7*, 4038–4049.

(51) Zhou, N.; Lu, Z.; Wu, Q.; Zhang, Y. *J. Chem. Phys.* **2014**, *140*, 214117.

(52) Horn, P. R.; Mao, Y.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2016**, *18*, 23067–23079.

(53) Szalewicz, K. *WIREs Comput. Mol. Sci.* **2012**, *2*, 254–272.

(54) Heßelmann, A.; Jansen, G. *Chem. Phys. Lett.* **2002**, *357*, 464–470.

(55) Heßelmann, A.; Jansen, G. *Chem. Phys. Lett.* **2002**, *362*, 319–325.

(56) Heßelmann, A.; Jansen, G. *Chem. Phys. Lett.* **2003**, *367*, 778–784.

(57) Jansen, G. *Wiley Interdiscip Rev: Comput. Mol. Sci.* **2014**, *4*, 127–144.

(58) Misquitta, A. J.; Jeziorski, B.; Szalewicz, K. *Phys. Rev. Lett.* **2003**, *91*, 033201.

(59) Misquitta, A. J.; Szalewicz, K. *J. Chem. Phys.* **2005**, *122*, 214109.

(60) Misquitta, A. J.; Szalewicz, K. *J. Chem. Phys.* **2005**, *122*, 214109.

(61) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133.

(62) Heßelmann, A.; Jansen, G.; Schütz, M. *J. Chem. Phys.* **2005**, *122*, 014103.

(63) Podeszwa, R.; Bukowski, R.; Szalewicz, K. *J. Chem. Theory Comput.* **2006**, *2*, 400–412.

(64) Parker, T. M.; Burns, L. A.; Parrish, R. M.; Ryno, A. G.; Sherrill, C. D. *J. Chem. Phys.* **2014**, *140*, 094106.

(65) Lao, K. U.; Schaffer, R.; Jansen, G.; Herbert, J. M. *J. Chem. Theory Comput.* **2015**, *11*, 2473–2486.

(66) Stone, A. J.; Misquitta, A. J. *Chem. Phys. Lett.* **2009**, *473*, 201–205.

(67) Misquitta, A. J. *J. Chem. Theory Comput.* **2013**, *9*, 5313–5326.

(68) Lao, K. U.; Herbert, J. M. *J. Chem. Theory Comput.* **2016**, *12*, 2569–2582.

(69) Horn, P. R.; Sundstrom, E. J.; Baker, T. A.; Head-Gordon, M. *J. Chem. Phys.* **2013**, *138*, 134119.

(70) Horn, P. R.; Head-Gordon, M. *J. Chem. Phys.* **2015**, *143*, 114111.

(71) Horn, P. R.; Mao, Y.; Head-Gordon, M. *J. Chem. Phys.* **2016**, *144*, 114107.

(72) Klimeš, J.; Michaelides, A. *J. Chem. Phys.* **2012**, *137*, 120901.

(73) Corminboeuf, C. *Acc. Chem. Res.* **2014**, *47*, 3217–3224.

(74) DiStasio Jr, R. A.; Gobre, V. V.; Tkatchenko, A. *J. Phys.: Condens. Matter* **2014**, *26*, 213202.

(75) Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. *Chemical reviews* **2016**, *116*, 5105–5154.

(76) Brauer, B.; Kesharwani, M. K.; Kozuch, S.; Martin, J. M. *Phys. Chem. Chem. Phys.* **2016**, *18*, 20905.

(77) Mardirossian, N.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9904–9924.

(78) Mardirossian, N.; Head-Gordon, M. *J. Chem. Phys.* **2015**, *142*, 074111.

(79) Mardirossian, N.; Head-Gordon, M. *J. Chem. Phys.* **2016**, *144*, 214110.

(80) Dunning Jr, T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(81) Woon, D. E.; Dunning Jr, T. H. *J. Chem. Phys.* **1993**, *98*, 1358–1371.

(82) Lao, K. U.; Herbert, J. M. *J. Phys. Chem. A* **2015**, *119*, 235–252.

(83) Stone, A. J. *Chem. Phys. Lett.* **1981**, *83*, 233–239.

(84) Thole, B. T. *Chem. Phys.* **1981**, *59*, 341–350.

(85) van Duijnen, P. T.; Swart, M. *J. Phys. Chem. A* **1998**, *102*, 2399–2407.

(86) Halgren, T. A. *J. Am. Chem. Soc.* **1992**, *114*, 7827–7843.

(87) Kitaura, K.; Morokuma, K. *Int. J. Quantum Chem.* **1976**, *10*, 325–340.

(88) Ziegler, T.; Rauk, A. *Theor. Chem. Acc.* **1977**, *46*, 1–10.

(89) Ziegler, T.; Rauk, A. *Inorg. Chem.* **1979**, *18*, 1558–1565.

(90) Stoll, H.; Wagenblast, G.; Preuß, H. *Theor. Chem. Acc.* **1980**, *57*, 169–178.

(91) Gianinetti, E.; Raimondi, M.; Tornaghi, E. *Int. J. Quantum Chem.* **1996**, *60*, 157–166.

(92) Khaliullin, R. Z.; Head-Gordon, M.; Bell, A. T. *J. Chem. Phys.* **2006**, *124*, 204105.

(93) Azar, R. J.; Horn, P. R.; Sundstrom, E. J.; Head-Gordon, M. *J. Chem. Phys.* **2013**, *138*, 084102.

(94) Tafipolsky, M. *J. Phys. Chem. A* **2016**, *120*, 4550–4559.

(95) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479–483.

(96) Tschumper, G. S.; Leininger, M. L.; Hoffman, B. C.; Valeev, E. F.; Schaefer III, H. F.; Quack, M. *J. Chem. Phys.* **2002**, *116*, 690–701.

(97) Ponder, J. W. *Tinker–Software Tools for Molecular Design*, tinker 7.1 ed.; Washington University: St. Louis, MO, 2015.

(98) Jiao, D.; King, C.; Grossfield, A.; Darden, T. A.; Ren, P. *J. Phys. Chem. B* **2006**, *110*, 18553–18559.

(99) Wu, J. C.; Piquemal, J.-P.; Chaudret, R.; Reinhardt, P.; Ren, P. *J. Chem. Theory Comput.* **2010**, *6*, 2059–2070.

(100) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; Ghosh, D.; Goldey, M.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Khaliullin, R. Z.; Kuś, T.; Landau, A.; Liu, J.;

Proynov, E. I.; Rhee, Y. M.; Richard, R. M.; Rohrdanz, M. A.; Steele, R. P.; Sundstrom, E. J.; Woodcock, H. L.; Zimmerman, P. M.; Zuev, D.; Albrecht, B.; Alguire, E.; Austin, B.; Beran, G. J. O.; Bernard, Y. A.; Berquist, E.; Brandhorst, K.; Bravaya, K. B.; Brown, S. T.; Casanova, D.; Chang, C.-M.; Chen, Y.; Chien, S. H.; Closser, K. D.; Crittenden, D. L.; Diedenhofen, M.; DiStasio, R. A.; Do, H.; Dutoi, A. D.; Edgar, R. G.; Fatehi, S.; Fusti-Molnar, L.; Ghysels, A.; Golubeva-Zadorozhnaya, A.; Gomes, J.; Hanson-Heine, M. W.; Harbach, P. H.; Hauser, A. W.; Hohenstein, E. G.; Holden, Z. C.; Jagau, T.-C.; Ji, H.; Kaduk, B.; Khistyaev, K.; Kim, J.; Kim, J.; King, R. A.; Klunzinger, P.; Kosenkov, D.; Kowalczyk, T.; Krauter, C. M.; Lao, K. U.; Laurent, A.; Lawler, K. V.; Levchenko, S. V.; Lin, C. Y.; Liu, F.; Livshits, E.; Lochan, R. C.; Luenser, A.; Manohar, P.; Manzer, S. F.; Mao, S.-P.; Mardirossian, N.; Marenich, A. V.; Maurer, S. A.; Mayhall, N. J.; Neuscamman, E.; Oana, C. M.; Olivares-Amaya, R.; ONeill, D. P.; Parkhill, J. A.; Perrine, T. M.; Peverati, R.; Prociuk, A.; Rehn, D. R.; Rosta, E.; Russ, N. J.; Sharada, S. M.; Sharma, S.; Small, D. W.; Sodt, A.; Stein, T.; Stück, D.; Su, Y.-C.; Thom, A. J.; Tsuchimochi, T.; Vanovschi, V.; Vogt, L.; Vydrov, O.; Wang, T.; Watson, M. A.; Wenzel, J.; White, A.; Williams, C. F.; Yang, J.; Yeganeh, S.; Yost, S. R.; You, Z.-Q.; Zhang, I. Y.; Zhang, X.; Zhao, Y.; Brooks, B. R.; Chan, G. K.; Chipman, D. M.; Cramer, C. J.; Goddard, W. A.; Gordon, M. S.; Hehre, W. J.; Klamt, A.; Schaefer, H. F.; Schmidt, M. W.; Sherrill, C. D.; Truhlar, D. G.; Warshel, A.; Xu, X.; Aspuru-Guzik, A.; Baer, R.; Bell, A. T.; Besley, N. A.; Chai, J.-D.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; Gwaltney, S. R.; Hsu, C.-P.; Jung, Y.; Kong, J.; Lambrecht, D. S.; Liang, W.; Ochsenfeld, C.; Rassolov, V. A.; Slipchenko, L. V.; Subotnik, J. E.; Van Voorhis, T.; Herbert, J. M.; Krylov, A. I.; Gill, P. M.; Head-Gordon, M. *Mol. Phys.* **2015**, *113*, 184–215.

(101) Vydrov, O. A.; Van Voorhis, T. *J. Chem. Phys.* **2010**, *133*, 244103.

(102) Rappoport, D.; Furche, F. *J. Chem. Phys.* **2010**, *133*, 134105.

(103) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553–566.

(104) Witte, J.; Neaton, J. B.; Head-Gordon, M. *J. Chem. Phys.* **2016**, *144*, 194306.

(105) Gill, P. M.; Johnson, B. G.; Pople, J. A. *Chem. Phys. Lett.* **1993**, *209*, 506–512.

(106) Azar, R. J.; Head-Gordon, M. *J. Chem. Phys.* **2012**, *136*, 024103.

(107) Thirman, J.; Head-Gordon, M. *J. Chem. Phys.* **2015**, *143*, 084124.

(108) Woon, D. E.; Dunning Jr, T. H. *J. Chem. Phys.* **1995**, *103*, 4572–4585.

(109) Peterson, K. A.; Dunning Jr, T. H. *J. Chem. Phys.* **2002**, *117*, 10548–10560.

(110) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, *286*, 243–252.

(111) Gresh, N.; Garmer, D. R. *J. Comput. Chem.* **1996**, *17*, 1481–1495.

(112) Ponder, J. W.; Case, D. A. *Adv. Prot. Chem.* **2003**, *66*, 27–85.

(113) Gresh, N. *J. Phys. Chem. A* **1997**, *101*, 8680–8694.

(114) Giese, T. J.; York, D. M. *J. Chem. Phys.* **2004**, *120*, 9903–9906.

(115) Piquemal, J.-P.; Chelli, R.; Procacci, P.; Gresh, N. *J. Phys. Chem. A* **2007**, *111*, 8170–8176.

(116) Harder, E.; Anisimov, V. M.; Vorobyov, I. V.; Lopes, P. E. M.; Noskov, S. Y.; MacKerrel Jr., A. D.; Roux, B. *J. Chem. Theory Comput.* **2006**, *2*, 1587–1597.

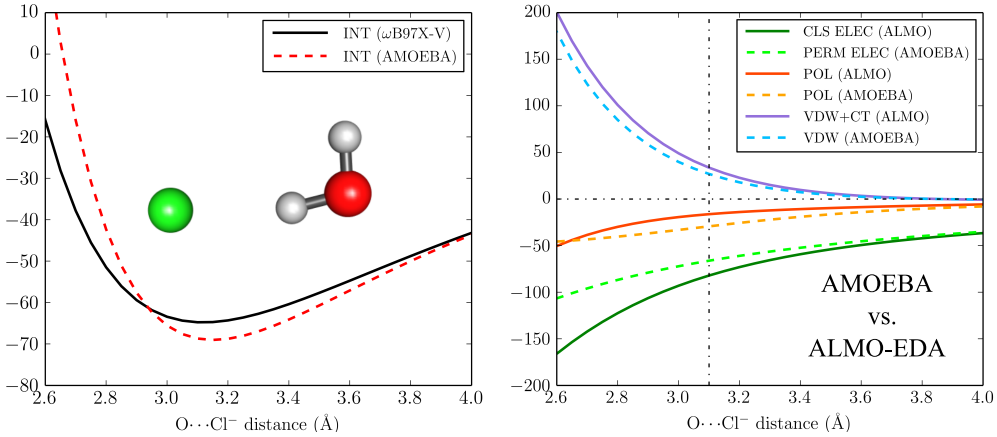

Figure 6: For Table of Contents (TOC) only