Development of a Many-Body Force Field for Aqueous Alkali Metal and Halogen Ions: An ALMO-EDA Guided Approach

Akshaya Kumar Das[§,†] Meili Liu[§,†,‡] and Teresa Head-Gordon[*,†,¶]

†*Pitzer Center for Theoretical Chemistry, Department of Chemistry, University of California, Berkeley 94720*

‡*Department of Chemistry, Beijing Normal University, Beijing 100875, China*

¶*Departments of Bioengineering and Chemical and Biomolecular Engineering, University of California, Berkeley, California 94720*

§ *These authors contributed equally*

E-mail: thg@berkeley.edu

Abstract

Aqueous solutions of alkyl/alkaline metal and halide ions play a crucial functional role in biological systems such as proteins, membranes, and nucleic acids, and for interfacial chemistry in geomedia and in the atmosphere. We present the MB-UCB many-body force field for monovalent and divalent ions that includes polarization, charge penetration to describe the short-range permanent electrostatics accurately, as well as a model for charge transfer to better describe the quantum mechanical potential energy surface and its components obtained from the absolutely localized molecular orbitals energy decomposition (ALMO-EDA) analysis. We find that the MB-UCB force field is in very good agreement with a validation suite of ion-ion and ion-water cluster data, exhibiting overall better cancellation of errors among energy

components, unlike the case for other many-body potentials that do not utilize an EDA scheme. However limitations in the functional form for the many-body energy components limit achievable accuracy through complete cancellation of error.

1 Introduction

Molecular dynamics (MD) simulations are widely used to study physical and chemical properties of biological systems like metalloenzymes as well as for non-biological systems such as metal organic frameworks and zeolites, and various interfacial chemistries such as clays and clouds. The predictive power of MD simulations depends on how accurately it describes the underlying potential energy surface associated with the system, especially for condensed-phase properties for which an accurate description of intermolecular interactions is crucial. The most commonly used force fields (FFs) such as AMBER,[1,2] CHARMM,[3] GROMOS,[4,5] and OPLS[6,7] use fixed point charges and pairwise additive terms to describe the intermolecular interactions. However, due to explicit electronic response effects such as many-body polarization, charge penetration, charge transfer, and higher order dispersion and Pauli repulsion, the pairwise additivity assumption can breakdown in asymmetric environments.[8–11] Therefore, accurate descriptions of intermolecular potentials require both better functional forms and correct partitioning of the energy among the molecular interactions (permanent electrostatics, polarization, charge transfer, dispersion, etc.).

There has been increasing efforts among research groups to improve the underlying physics of force fields, leading with many-body polarization introduced by means of differing functional forms such as the Drude model, charge fluctuation, or point induced dipoles.[10,12] Several complete polarizable force fields have been developed such as the CHARMM Drude model[13,14] and AMOEBA point dipole models,[15–17] whereas other advanced force fields include additional terms beyond polarization, such as a charge transfer model used in SIBFA (Sum of Interactions Between Fragments Ab initio computed)[18] and Gaussian smeared charges such as GEM (Gaussian Electrostatic Model).[19,20] Although these many-body mod-

els often perform well with much greater transferability compared to fixed charge force fields, they can still exhibit surprising failures in some contexts, such as that reported in the SAMPL community evaluations for guest-host interactions using advanced models.[21,22] This we hypothesize arises from the fact that their force field parameters are developed with respect to the total *ab initio* energy and forces and/or experimental properties, and thus must rely on the cancellation of errors of the individual intermolecular terms that contribute unequally in the overlapping region of molecule-molecule interactions. But with guidance from only total energies and forces or averaged properties, and without the breakdown into the energy components of electrostatics, polarization, charge transfer, exchange-repulsion and dispersion, these terms may be out of balance.

What has transpired recently is the recognition that energy decomposition analysis (EDAs) techniques such as Symmetry-Adapted Perturbation Theory (SAPT) and variational absolutely localized-molecular-orbitals (ALMO-EDA), can offer a strategy for separating intermolecular energy terms into chemically sensible energy components.[23–25] Although the EDAs too are model chemistries, i.e. the QM energy is not uniquely decomposable, they can aid in the development of physically motivated many-body force fields which will include a balanced breakdown among short-ranged charge transfer, intermediate range polarization, and long-ranged permanant electrostatics. In fact our MB-UCB force field for water was developed based on the principles of the many-body expansion for permanent electrostatics and polarization and with guidance from the ALMO-EDA method for charge transfer, and yielding water properties in excellent agreement with experiment.[26,27]

Here we extend that success to the development of the MB-UCB force field for alkyl/alkaline metals (M^{n+} = Li^+, Na^+, K^+, Cs^+, Mg^{2+}, and Ca^{2+}) and halide (X^-= Cl^-, Br^- and I^-) ions that again benefits from the many-body expansion and direct guidance from the ALMO-EDA approach. We have used 2-body and 3-body ion-water systems evaluated with ALMO-EDA to help parameterize all piecewise terms of MB-UCB without changing the original MB-UCB water model. We show that better control in the cancellation of smaller errors of an EDA

guided approach is superior to unguided force field developments that rely on cancellation of large magnitude errors or missing molecular interactions. Even so, the limitations in the many-body functional form for charge transfer and the damping functions for polarization, and the inflexibility of the 14-7 van der Waals function, do limit the goal of perfect error cancellation. But in the end the validation suite of ion-ion, solvent separated ion pairs, and a companion paper that shows that MB-UCB performs well for all condensed phase ion properties, indicates that the MB-UCB model is a significant improvement over existing polarizable force fields.

2 THEORY

Permanent Electrostatics. The permanent electrostatics for the MB-UCB force field is modeled using atom centered point multipoles, and the electrostatic interactions between the atoms is expressed as

$$E_{\text{elec}} = \sum_{i<j} \mathbf{M}_i^T \mathbf{T}_{ij} \mathbf{M}_j \tag{1}$$

where $\boldsymbol{M}_i^T = [q_i, \mu_{ix}, \mu_{iy}, \mu_{iz}, Q_{ixx}, Q_{ixy}, Q_{ixz}, Q_{iyy}, Q_{iyz}]$ is the multipole vector and T_{ij} is the multipole interaction tensor that consists of appropriate derivatives of $1/r_{ij}$ according to the multipole expansion. Multipole expansions of the permanent electrostatics is intended to describe the anisotropic electrostatic potential more accurately. However, in the very short-range, when two atomic electron clouds overlap, the electrostatic multipole expansion breaks down due to missing charge penetration effects. Charge penetration in force fields is modeled based on a strategy of separating the atomic charge into a core nuclear charge and smeared electron cloud charge. In fact many advanced force fields have begun to adopt some variation of the CP models.[8,18,20,28–35] proposed by either Slipchenko and co-workers[28,31] or Piquemal et al.[18]

The main difference between these two type of CP models is the use of different damping

functions to approximate the value of the overlap integral. While the recent AMOEBA+ model[36] adopts the functional form of Slipchenko et al.[28], we have instead used the Piquemal model for the MB-UCB force fields to account for the charge penetration effect, but only applying it to monopole-monopole ($q-q$) interactions[37]. Therefore, the modified charge-charge electrostatic interactions between two atoms A and B with atomic charges q_A and q_B are expressed as

$$\begin{aligned} E_{\text{elec}}^{q-q} =& \frac{Z_{\text{A}} Z_{\text{B}}}{r} - \frac{Z_{\text{A}}\left(Z_{\text{B}} - q_{\text{B}}\right)}{r} f_{\text{damp}} - \frac{Z_{\text{B}}\left(Z_{\text{A}} - q_{\text{A}}\right)}{r} f_{\text{damp}} \\ &+ \frac{\left(Z_{\text{A}} - q_{\text{A}}\right)\left(Z_{\text{B}} - q_{\text{B}}\right)}{r} f_{\text{damp}}^{\text{overlap}} \end{aligned} \tag{2}$$

The first term in Equation 2 is for the core–core interaction, where Z is the effective core charge (and equal to the number of valence electrons). The second and third terms describe the interaction between the core monopole charge q and electron clouds of the other atoms, and the fourth term accounts for the electron–electron interactions, where Z–q describes the magnitude of the negatively charged electron cloud. The two damping functions, $f_{damp} = (1-\exp(-\alpha r))$ and $f_{damp}^{overlap} = (1\text{–}exp(-\beta_A r))(1\text{–}exp(-\beta_B r))$, require two parameters, α and β, to control the damping of core–electron and electron–electron interactions, respectively, in order for the charge penetration effects to vanish rapidly and to recover the classical Coulombic multipolar interactions at longer distances.

Polarization Energy. Many-body polarization is explicitly incorporated in the MB-UCB model by induced dipoles at each polarizable site located on all atomic centers.[15] The induced dipoles (μ_{ind}) at a polarizable site i with atomic polarizability α_i is expressed as

$$\boldsymbol{\mu}_i^{\text{ind}} = \alpha_i \left[\sum_j \mathbf{T}_{ij} \mathbf{M}_j - \sum_{j \neq i} \mathbf{T}_{ij}^{\text{d-d}} \boldsymbol{\mu}_j^{\text{ind}} \right] \tag{3}$$

where, M_j and T_{ij} are polytensor permanent multipoles and the multipole-multipole interaction matrix, respectively, and T_{ij}^{d-d} are the dipole-dipole interaction tensor. The first term in parentheses on the right hand side of the Equation 3 is the electric field E_i from the

the permanent multipoles and the second term is the electric field from the mutual induced dipoles from all other sites. Reorganizing Equation 3 we arrive at the following expression

$$\begin{aligned} \boldsymbol{\mu}_i^{\text{ind}} &= \alpha_i \left[\mathbf{E}_i - \sum_{j \neq i} \mathbf{T}_{ij}^{\text{d-d}} \boldsymbol{\mu}_j^{\text{ind}} \right] \\ \Rightarrow \alpha_i^{-1} \boldsymbol{\mu}_i^{\text{ind}} &+ \sum_{j \neq i} \mathbf{T}_{ij}^{\text{d-d}} \boldsymbol{\mu}_j^{\text{ind}} = \mathbf{E}_i \end{aligned} \tag{4}$$

or more generally Equation 4 can be represented in matrix form

$$\begin{pmatrix} \alpha_1^{-1} & T_{12}^{\text{d-d}} & \cdots & T_{1N}^{\text{d-d}} \\ T_{12}^{\text{d-d}} & \alpha_2^{-1} & \cdots & T_{2N}^{\text{d-d}} \\ \vdots & \vdots & \ddots & \vdots \\ T_{1N}^{\text{d-d}} & T_{2N}^{\text{d-d}} & \cdots & \alpha_N^{-1} \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_N \end{pmatrix} = \begin{pmatrix} E_1 \\ E_2 \\ \vdots \\ E_N \end{pmatrix} \tag{5}$$

In Equation 5, the diagonal blocks are the inverse of the atomic polarizability, while the off-diagonal blocks T_{d-d} are the Thole damped[38] Cartesian interaction tensors between induced dipoles of two polarizable sites i and j. The first order Thole damped T_{d-d} can be expressed as

$$T_{ij,\gamma} = -\left[1 - \exp\left(-au^3\right)\right] \frac{r_{ij,\gamma}}{r_{ij}^3}, \quad \gamma = x, y, z \tag{6}$$

where $u = r_{ij} / \left(\alpha_i \alpha_j\right)^{1/6}$. Unlike other polarizable force fields such as AMOEBA and AMOEBA+ that use rotationally invariant isotropic atomic polarizabilities[36,39,40], MB-UCB uses a rank two tensor anisotropic atomic polarizablity.

$$\alpha_i^{-1} = \begin{pmatrix} \alpha_{i,xx} & \alpha_{i,xy} & \alpha_{i,xz} \\ \alpha_{i,yx} & \alpha_{i,yy} & \alpha_{i,yz} \\ \alpha_{i,zx} & \alpha_{i,zy} & \alpha_{i,zz} \end{pmatrix}^{-1} \tag{7}$$

Finally, the polarization energy can expressed in terms of induced dipoles as

$$E_{\text{pol}} = -\frac{1}{2}\sum_i \boldsymbol{\mu}_i^{\text{ind}} \mathbf{E}_i \tag{8}$$

and the induced dipoles at each multipole site are obtained by solving Equation 5 self-consistently by using conjugate gradient method[41]. More advanced algorithms exist such as the iEL/0-SCF approach for efficient computation of induced dipoles,[42,43] but were not used here.

Charge Transfer Energy. MB-UCB introduced a new charge transfer model to fill the gap between ALMO-EDA energy components and force fields. MB-UCB uses an empirical many-body function similar to the polarization energy induced multipoles to incorporate the many-body charge transfer energy.[44]

$$\begin{aligned} E_{\text{CT-ind}} &= -\frac{1}{2}\sum_i \mu_i^{\text{ct-ind}} \mathbf{E}_i^{\text{ct}} \\ \boldsymbol{\mu}_i^{\text{ct-ind}} &= \alpha_i^{\text{ct}} \left[\sum_j \mathbf{T}_{ij}^{\text{ct}} \mathbf{M}_j - \sum_{j \neq i} \mathbf{T}_{ij}^{\text{ct[d-d]}} \boldsymbol{\mu}_j^{\text{ct-ind}} \right] \end{aligned} \tag{9}$$

where α_i^{ct} controls the charge transfer energy between two multipole sites through a response to the permanent electrostatics, and the multipole interaction matrix (T^{ct}) elements are damped with an exponential damping function.

$$\mathbf{T}_\zeta^{\text{ct}} = -\left[1 - d\exp\left(-bu^3\right)\right] \frac{r_\zeta}{r_{ij}^3}, \quad \zeta = x, y, z \qquad u = \frac{r_{ij}}{\left(\alpha_i^{\text{ct}} \alpha_j^{\text{ct}}\right)^{1/6}} \tag{10}$$

The three parameters α_{ct}, b and d are responsible for the fast exponential decay of the charge transfer energy, which should be more short-ranged than polarization. However, unlike the Deng et al. model[44], which only considered the direct charge transfer between the atomic sites. We consider both direct and mutual CT terms that recovers a larger amount of the many-body character of the charge transfer, albeit with the understanding that no explicit charge flow is operative. For AMOEBA+ the calculated induced dipole polarization energy is subtracted from the SAPT induction energy to obtain the CT energy, since the

version of SAPT used for AMOEBA+ development does not separate charge transfer from polarization.[36] This may result in over or under subtraction of charge transfer depending on the amount of polarization.

van der Waals Interactions. The remaining energy terms are Pauli repulsion and dispersion, and are modeled in MB-UCB as a van der Waals interaction using a buffered 14-7 pairwise-additive function proposed by Halgren[45]

$$E_{\mathrm{vdW}} = \sum_{i<j} \epsilon_{ij} \left(\frac{1+\delta}{\sigma_{ij}+\delta} \right)^{7} \left(\frac{1+\gamma}{\sigma_{ij}^{7}+\gamma} - 2 \right) \tag{11}$$

where, ϵ defines the energy scale, and $\sigma = R_0/r$ is the distance between two atoms, where R_0 is the distance corresponding to the minimum energy. Like AMOEBA,[16] we set the two constants δ and γ to 0.12 and 0.07, respectively.

2.1 METHODS

Using a many-body ansatz and without changing the original MB-UCB water parameters,[27] we follow an ALMO-EDA-guided strategy for developing the MB-UCB force fields for ions. In the first step a systematic set of ion-water dimer and trimer complexes were generated along intermolecular distances and angles, d and θ, and ALMO-EDA calculations were performed for water-ion dimers and trimers containing monovalent and divalent ions complexed with one or two water molecules. The equilibrium optimized geometries (always defined as $d = 0$ on the distance axis in figures below) is chosen as the reference geometry. All of the ALMO-EDA calculations were performed at the level of ωB97X-V/def2-QZVPPD.

For the two-body systems, the directions of distance scans are the M-O separation in the H_2OM^{n+} systems (Figure 1 a) and $X-H$ for the $X^- - H_2O$ systems (Figure 1 b). The angle scan of the water-ion dimers are performed along the water bisector axis (Figure 1 c). For the trimer systems, $(H_2O)_2M^{n+}$, distance scans are performed along the M-O distances by displacing two water molecules simultaneously from the metal center (Figure 1 d). For the

distance scans for the $H_2O(M^{n+})_2$ systems, the starting structure placed the two metal ions at the equilibrium distance of the corresponding H_2OM^{n+} dimer systems. For the distance scan in the $(H_2O)_2X^-$ systems, one of the anions is displaced from the centroid of the triangle formed by the three heavy atoms (Figure 1 g). For the $H_2O(X^-)_2$ system, the distance scan is along the $X-H$ and two ions X^{-1} are displaced simultaneously. Furthermore, all of the three-body distance scans are performed at four different water-ion-water or ion-water-ion angles depending on the system (see Figure 1 d, e, f, g, i).

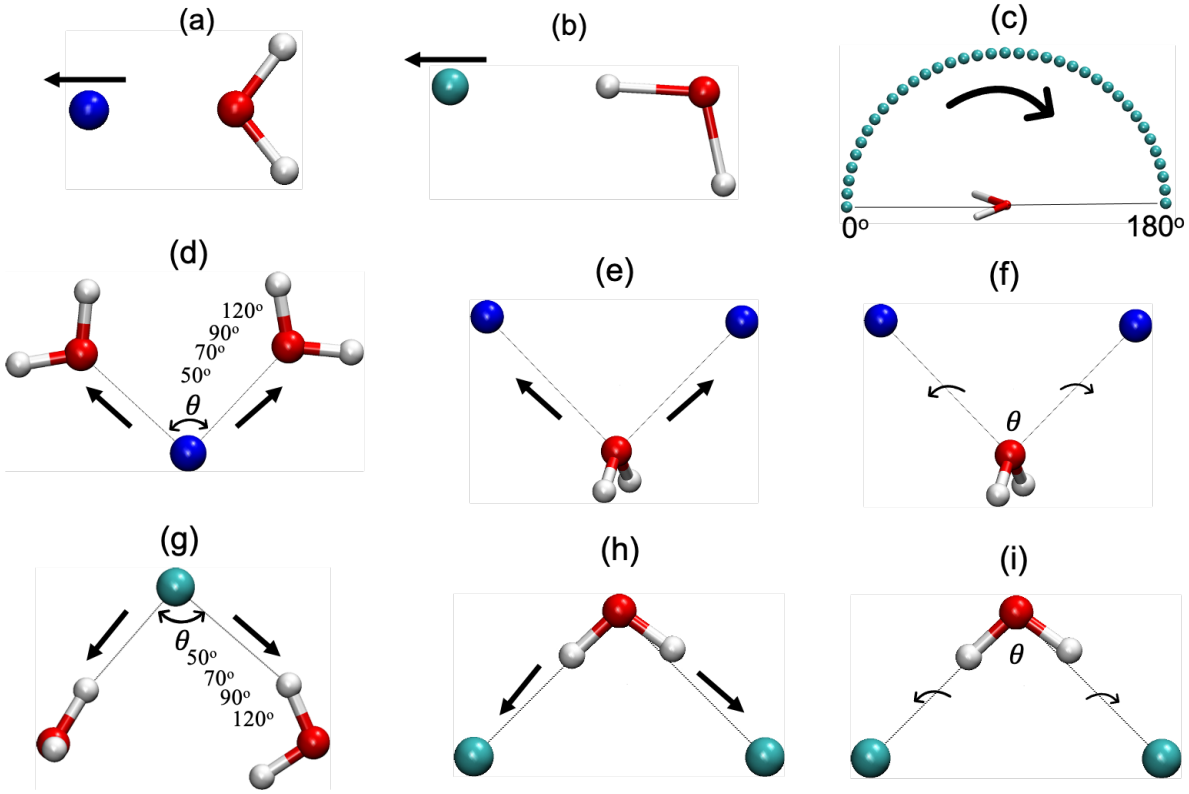

Figure 1: Illustrations of the distance scans and angle scans performed for two-body and three-body systems. (a, c)H_2OM^{n+} distance and angle scan, (b, c)H_2OX^- distance and angle scan, (d, g)$(H_2O)_2M^{n+}/X^-$ distance scan at four different fixed angles, (e, f)$H_2O(M^{n+})_2$ distance and angle scan,(h, i)$H_2O(X^-)_2$ distance and angle scan.

For the ion electrostatics we only use monopoles, with integral charge of +1 for monovalent cations, +2 for divalent cations, and -1 for halide ions. Charge penetration parameters for the metal ions are obtained by fitting them to the two body and three body ion-water ALMO-EDA electrostatics. However, for anions, the CP damping parameters for α and β

are used directly from the literature for haloalkanes[18] and by simply changing the valence Z for halide ions, and was found to be sufficient to describe the water-ion electrostatics. For polarization, atomic polarizabilities for ions are directly taken from the AMOEBA force field,[15] but the Thole damping parameters are optimized to reproduce the two-body and three-body ion-water ALMO-EDA polarization energy. The charge transfer parameters are obtained by fitting with respect to the ALMO-EDA charge transfer energy for the configurations consisting of dimers and trimers. Finally, vdW parameters for ions are optimized to reproduce the total *ab initio* ion-water intermolecular energy.

3 RESULTS

3.1 Parameterization strategy using 2-body and 3-body energies

We directly use the intermolecular energy breakdown of the two-body and three-body water-ion systems obtained from ALMO-EDA to guide parameterization of individual energy components of MB-UCB, and compare the total MB-UCB energy to the ωB97X-V/def2-QZVPPD DFT benchmark. Transferability of the MB-UCB force field for cations and anions in more complicated environments is assessed later in the validation of water-ion systems not in the training set.

The leading order term in the many-body expansion is the 2-body energies. Figure 2 shows the intermolecular 2-body total energy and energy components for the Na^+ and Cl^- ions interacting with a single water molecule. For both Na^+ and Cl^- ions, errors in the total interaction energy are less than 1 kcal/mol compared to the *ab initio* benchmark near equilibrium (Figure 2 a,b). The ALMO-EDA energy decompositions give some insight into the leading 2-body result for the MB-UCB force field. For both cations and anions, the maximum contribution to the total interaction energy is from the permanent electrostatics, and MB-UCB reproduces the same ALMO-EDA component throughout the distance scan except in the highly compressed region (Figure 2 c,d). The inclusion of charge penetration

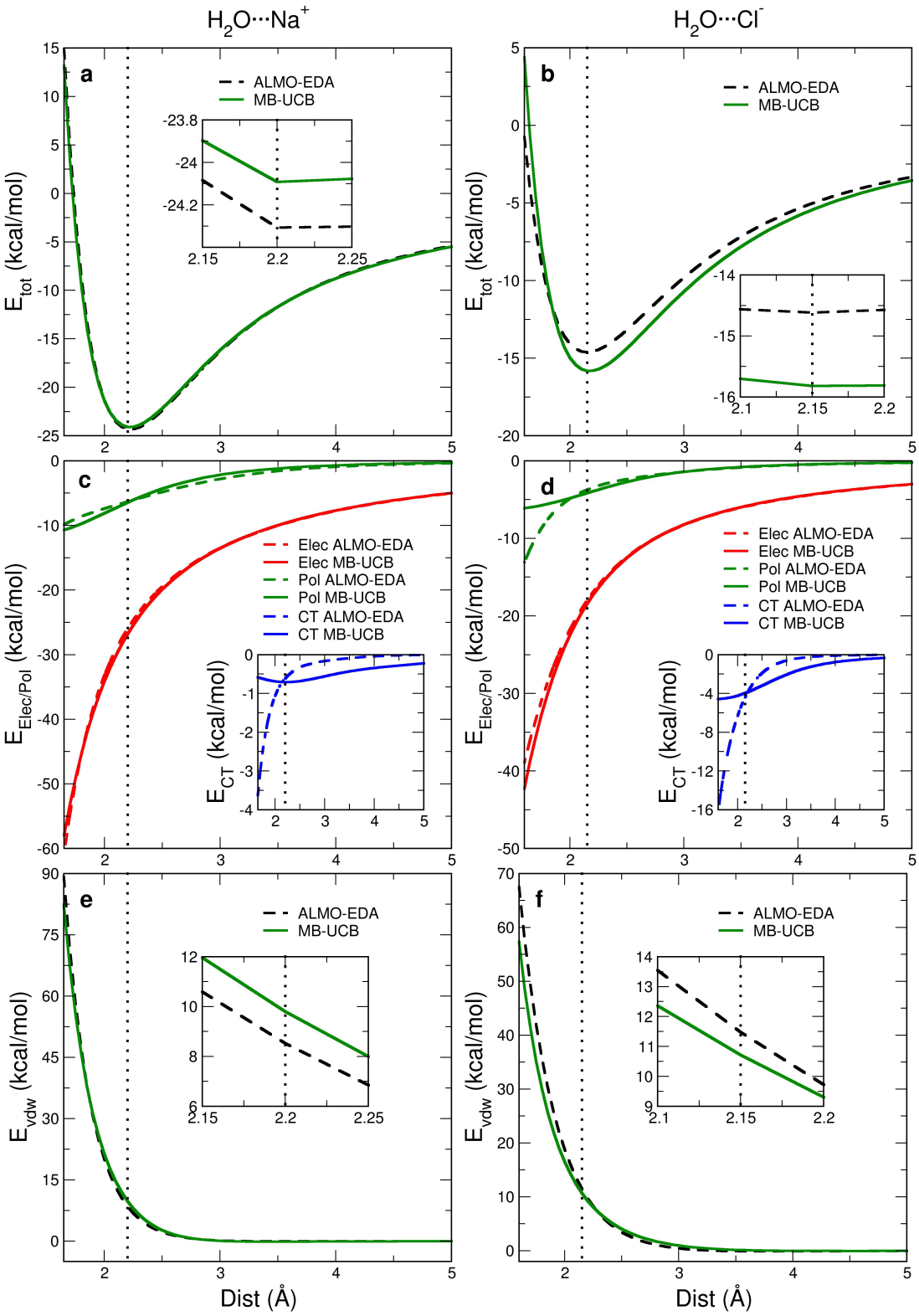

Figure 2: *Total energy and energy decomposition at the 2-body level for* $(H_2O)_2Na^+$ *(left) and* $(H_2O)_2Cl^-$ *(right).* Distance scan for ALMO-EDA (dash) and MB-UCB (solid): (a,b) total interaction energy; (c,d) permanent electrostatics (red), polarization (green), and charge transfer (inset, blue); (e,f) van der Waals energy. Equilibrium distances are represented by dotted lines.

in the MB-UCB model helps to reproduce the short-ranged electrostatics for MB-UCB when it is compared with the electrostatics from the AMOEBA force field, although more recent incarnations such as AMOEBA+ have also shown the benefits of charge penetration.[8] Similar to permanent electrostatics, the error in the polarization energy near equilibrium is also very small compared to the ALMO-EDA for both Na^+ and Cl^- ions (Figure 2 c,d), although it is overdamped for Cl^- in the compressed region. When comparing the ALMO-EDA and MB-UCB charge transfer models, it is evident that the differences around the equilibrium geometries are very small for the Cl^- ion, and within 0.5 kcal/mol for the Na^+ ion with some overdamping observed at short-ranged (Figure 2 c,d inset). Even so, the contribution of the CT energy towards the total interaction energy is smaller compared to the other energy components. We therefore rely on cancellation of error through the van der Waals term (Figure 2 e,f), which appears to be an adequate strategy to compensate for $Na^+ - H_2O$, but the underdamping errors in the polarization that occur mostly in the compressed region for Cl^- are too large to benefit completely with this strategy.

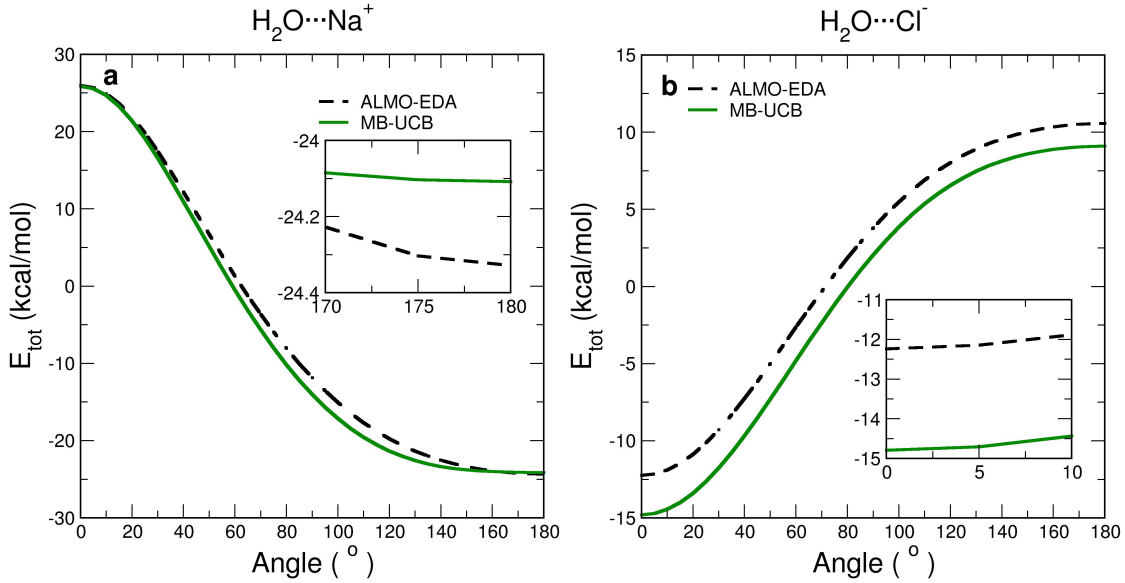

Figure 3: *Total interaction energy of the 2-body level angular scans for (a)* $(H_2O)Na^+$ *and (b)* $(H_2O)Cl^-$ *(right).* Scan of the angle formed by Na-O and O-Cl bond and water bisector axis. See Figure 2 for further details.

The van der Waals compensation can be further evaluated with a stress test of angular

scans as given in Figure 3. We wish to emphasize that the angular scans corresponding to regions that are away from equilibrium geometries (i.e. corresponding to angle values $< 90^\circ$ for cations and $> 130^\circ$ for anions) are high in energy and hence will almost never be sampled in a simulation near room temperature. Even so all energy components reasonably track the trends of the ALMO-EDA results for all energy components, and the resulting total energy error is less than 1 kcal/mol for Na^+ and Cl^- throughout the angular scan with this strategy. The energy decompositions for the angular scans in terms of permanent electrostatics, polarization, charge transfer, and van der Waals interactions are given in the Supplementary Information Figure S1.

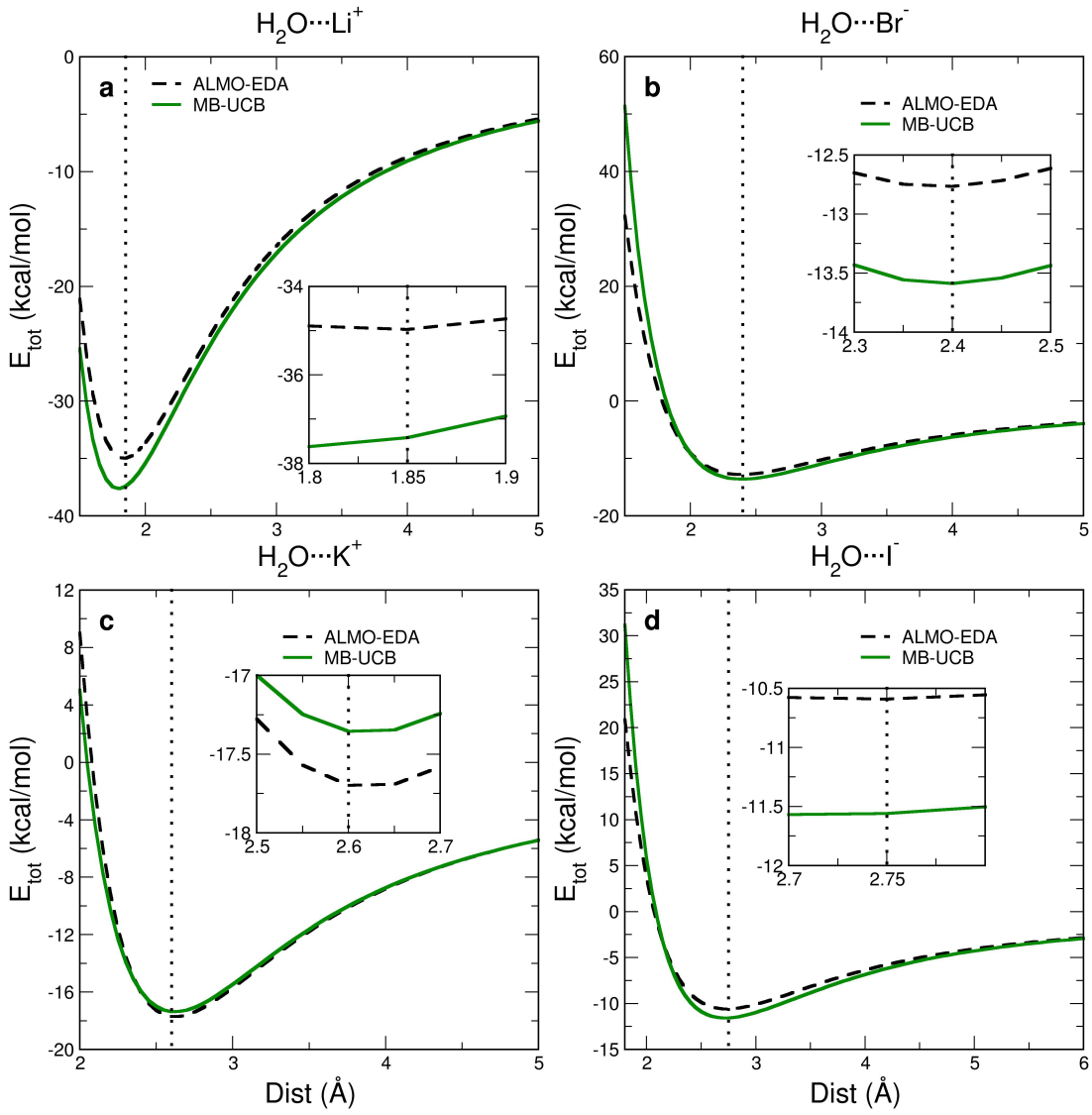

Figure 4: *Total energy at the 2-body level for* $(H_2O)Li^+$ *and* $(H_2O)K^+$ *(left) and* $(H_2O)Br^-$ *and* $(H_2O)I^-$ *(right).* See Figure 2 for further details.

Identical conclusions are reached in regards the MB-UCB force field performance on all the monovalent ions and their interactions with water when compared to the DFT benchmark. Figure 4 shows the distance scans for the total energy of interaction between water and the Li^+, K^+, Br^- and I^- ions. The energy decompositions in terms of permanent electrostatics, polarization, charge transfer, and van der Waals interactions for the distance and angle scans are given in the Supplementary Information Figures S2-S5, and show again the same trends with the ALMO-EDA decompositions as found for Na^+ and Cl^-. We see

that the errors in the CT model are largest for cations while polarization errors are largest for anions due to the damping function, and that angular scans using MB-UCB track the ALMO-EDA for both monovalent anions and cations. Figure S6 also reports the energy and energy decompositions for distance and angle scans of Cs^+ with no loss of generality in conclusions in regards the 2-body monovalent ion results.

The next term in the many-body expansion is the 3-body energies. First we consider the three-body $(H_2O)_2Na^+$ and $(H_2O)_2Cl^-$ at an equilibrium angle of 90° for the cation and 70° for the anion that are thoroughly analyzed in Figure 5. The total interaction energy at the 3-body level for both ions reproduce the DFT benchmark with very small errors of less than 1 kcal/mol near equilibrium and in the asymptotic region, with most errors arising in the compressed region (Figure 5 a,b). Comparing individual energy components, the permanent electrostatic and polarization interactions are in near perfect agreement between MB-UCB and ALMO-EDA for both the Na^+ and Cl^- ions (Figure 5 c,d), but with expected deviations in the polarization energy at short-range due to the damping function as already seen in the 2-body systems. The CT for Na^+ show noticeable errors at both the compressed and asymptotic regions (Figure 5 c,d insets), but these are compensated for through the vdW energy component (Figure 5e). Again the overdamped polarization for Cl^- ion is not perfectly compensated for as seen in Figure 5f, although the total interaction errors near the basin remain small, i.e. < 1 kcal/mol. Similar trends, energy accuracies, and conclusions are reached for $(H_2O)_2Na^+$ and $(H_2O)_2Cl^-$ as the scan angle decreases (and the energies are highly repulsive) or increases (when the energies are stabilizing) as shown in Figures S7-S9.

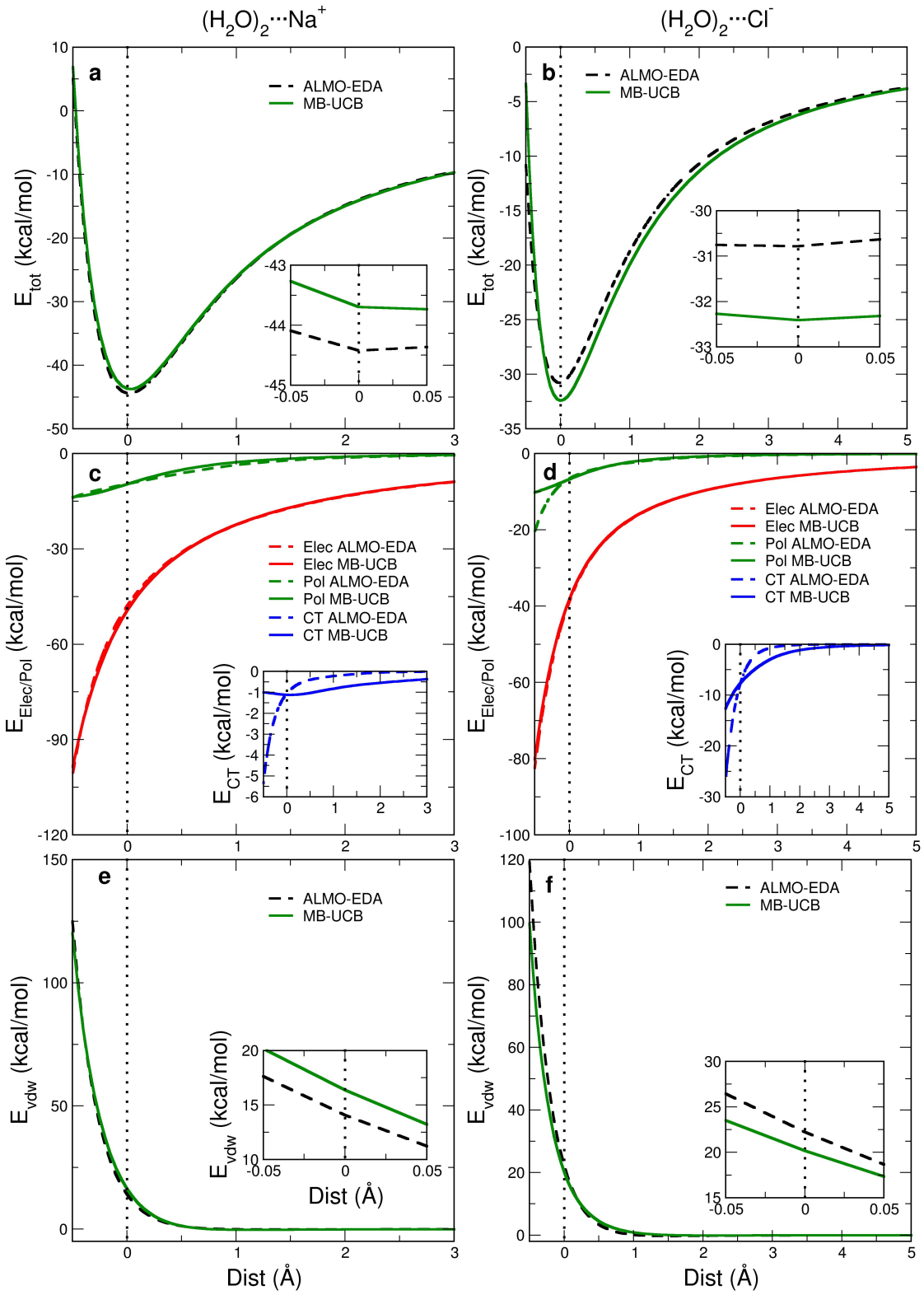

Figure 5: *Total interaction energy and energy decomposition at the 3-body level for* $(H_2O)_2Na^+$ *(left) and* $(H_2O)_2Cl^-$ *(right).* Scan is along the O-Na-O distance and H-Cl-H distances at a 3-body angle of 90° for Na^+ and 70° for Cl^-. See Figure 2 for further details.

Again, similar conclusions are reached in regards the other monovalent ions and their 3-body interactions when compared to the DFT benchmark. Figure 6 shows the distance scans for the total energy of interaction between water and the Li^+, K^+, Br^- and I^- ions with two waters at their equilibrium angles. The energy decompositions in terms of permanent electrostatics, polarization, charge transfer, and van der Waals interactions for the distance.

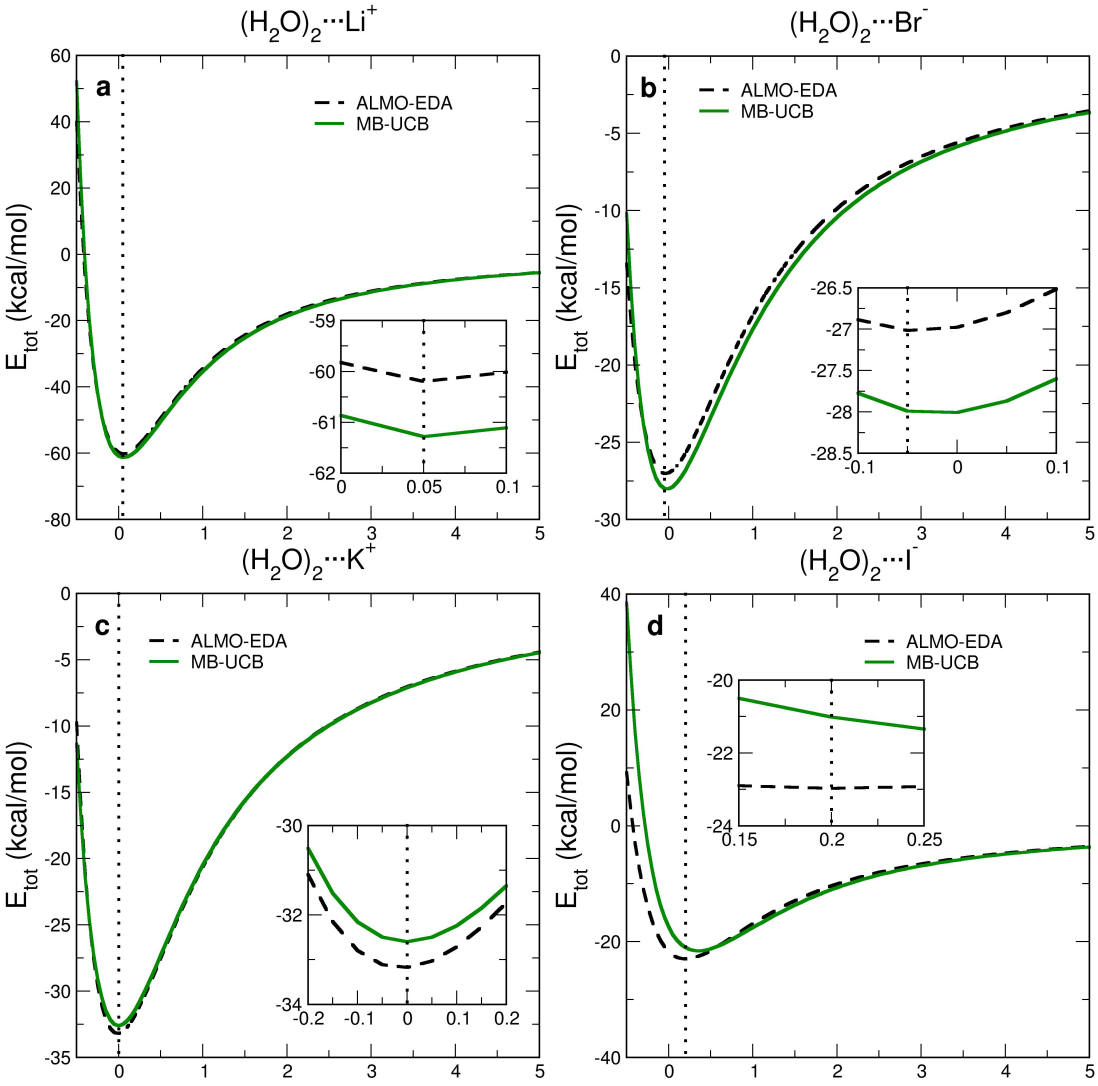

Figure 6: *Total energy at the 3-body level for* $(H_2O)_2Li^+$ *and* $(H_2O)_2K^+$ *(left) and* $(H_2O)_2Br^-$ *and* $(H_2O_2)I^-$ *(right).* Scan is along the O-cation distance and H-anion distances at a 3-body angle of 90 degree for cations and 70 degree for anions. See Figure 2 for further details

.

and angle scans for 3-body interactions for these ions as well as Cs^+, are given in the Supplementary Information Figure S10-S19. They show the same trends with the ALMO-

EDA decompositions as found for $(H_2O)_2Na^+$ and $(H_2O)_2Cl^-$.

The final results for this section concern the ability of the MB-UCB model to reproduce the ALMO-EDA results for the more difficult case of the divalent cations, and illustrated here for Ca^{2+} and Mg^{2+}. Figure 7 shows that the 2-body total interaction energy of the divalent cations with water as a function of distance or angle (Figure S20) exhibit larger errors relative to the monovalent ions by ∼2.5-3.0 kcal/mol. Indeed the divalent ions have proven much more difficult at short-range, with interaction strengths approaching that of covalent bonding and the classical description breaks down. Although the CT model remains a problem, overall error cancellation has been reasonably successful in the case of divalent ions, and this is further supported in the 3-body interactions for $(H_2O)_2Ca^{2+}$ and $(H_2O)_2Mg^{2+}$ that are in excellent agreement with the DFT benchmark as shown in Figure S21-S24. We return to these systems when we consider validation involving ion clusters for both monovalent and divalent systems in the next section.

3.2 Validation of MB-UCB for more complex clusters

As a first validation of the MB-UCB ion force field, we consider larger water-ion clusters. For each monovalent and divalent ion we selected 100 random structures of the ions with 6-8 first shell neighbouring water molecules, extracted from the MD simulations of 1 ns using the the AMOEBA force field and the geometries are extracted every 10 ps. Based on the experimental first shell coordination number for each ion, the closest 6-8 water molecules to the ions are chosen to obtain the final configurations for validation. For each configuration we calculated the total interaction energy and contributions of the different energy components for both the MB-UCB and the AMOEBA polarizable force fields and compared to ALMO-EDA.

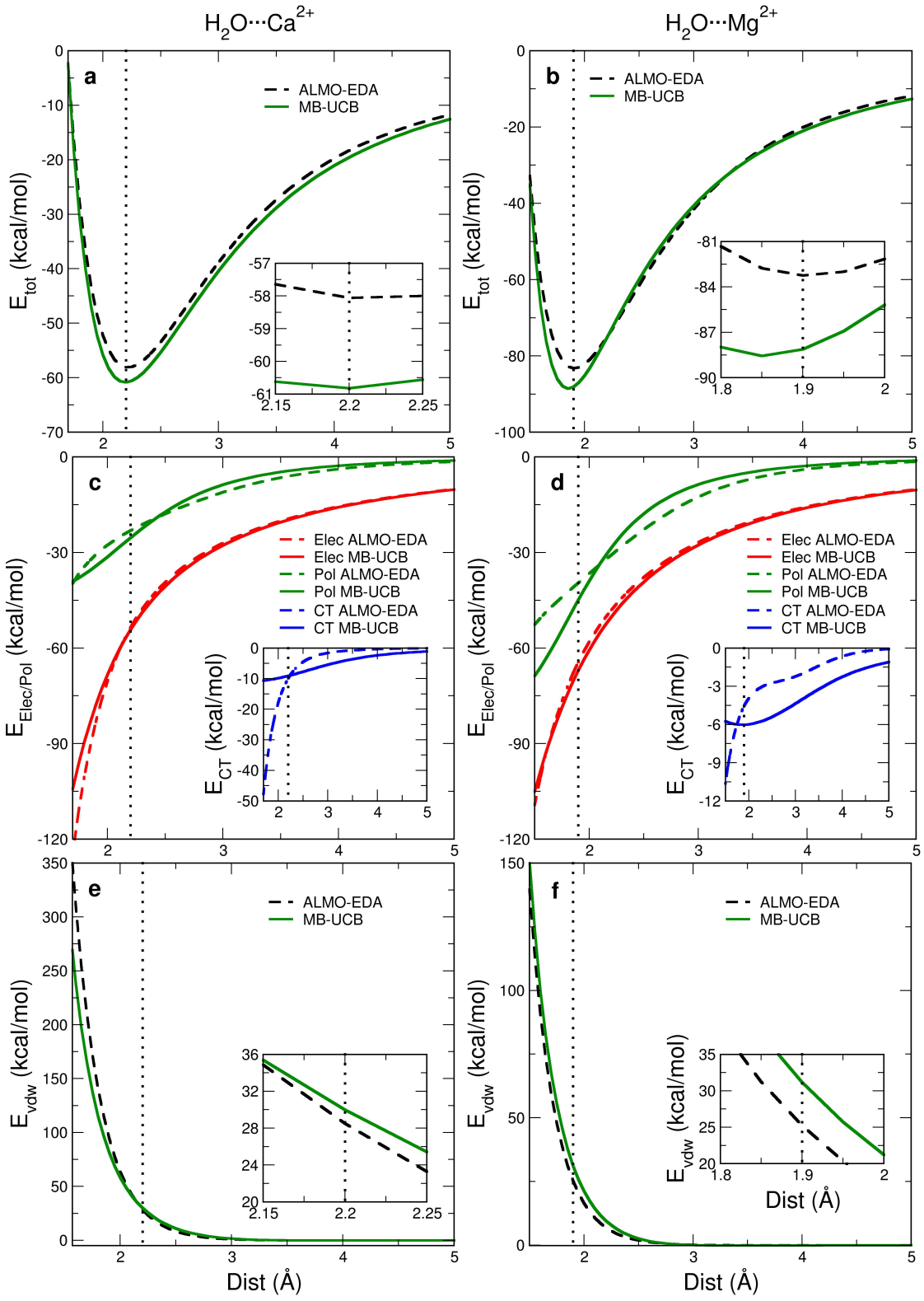

Figure 7: *Total interaction energy and energy decomposition at the 2-body level for* $(H_2O)Ca^{2+}$ *(left) and for* $(H_2O)Mg^{2+}$ *(right).* Distance scan (top) and angle scan (bottom). See Figure 2 for further details.

Table 1 shows the mean signed error (MSE) to test for overbinding or underbinding, and root mean square error (RMSE) to assess larger energy differences, for MB-UCB and AMOEBA against the DFT benchmark. Overall MB-UCB performs significantly better than AMOEBA for all ions, and with fewer egregious outliers, with the exception of the Cl^- water-ion clusters although error remains reasonably small ($\sim$ 1 kcal/mol) on an absolute scale. From Table 1 it is seen that for all the ions AMOEBA systematically underbinds in the total interaction energy.

Table 1: Mean signed errors and root mean squared errors in the total energy of MB-UCB force field and AMOEBA force field with DFT for systems containing 6-8 first shell water molecules along with different ions.

Ions	Mean signed error (MSE) (kcal/mol)		Root mean square error (RMSE) (kcal/mol)	
	MB-UCB	AMOEBA	MB-UCB	AMOEBA
Li^+	-2.26	-4.94	3.44	5.17
Na^+	-0.08	-1.92	1.52	2.19
K^+	-0.27	-0.42	1.29	1.01
Cs^+	-0.02	-6.45	2.54	6.58
Mg^{2+}	-0.12	-12.11	2.83	12.26
Ca^{2+}	-0.53	-3.42	2.27	3.73
Cl^-	-0.73	-0.08	1.64	1.25
Br^-	-1.04	-1.69	1.75	2.22
I^-	-4.79	-2.30	6.32	2.52

Figure 8 uses the Na^+ and Cl^- water-ion clusters to show the origin of errors through visual inspection of the correlation between the DFT benchmark for the total interaction energy and ALMO-EDA decompositions compared to MB-UCB and the AMOEBA polarizable force fields. While MB-UCB correlates very well with the ALMO-EDA electrostatic components due to good reproduction of the leading order 2- and 3-body terms, AMOEBA shows substantial deviation from the ALMO-EDA result due to lack of charge penetration (Figure 8 c,d), resulting in errors that are as large as 20 kcal/mol in the electrostatic interactions. It is important to note that AMOEBA+ has corrected that problem by introducing charge penetration for the permanent electrostatics,[8] and is now becoming a largely solved problem. For the polarization energy component of $(H_2O)_6Na^+$ interactions (Figure 8e), both MB-UCB and AMOEBA agree very well with ALMO-EDA, however for the Cl^- ion AMOEBA substantially overestimates the polarization energy whereas MB-UCB not only

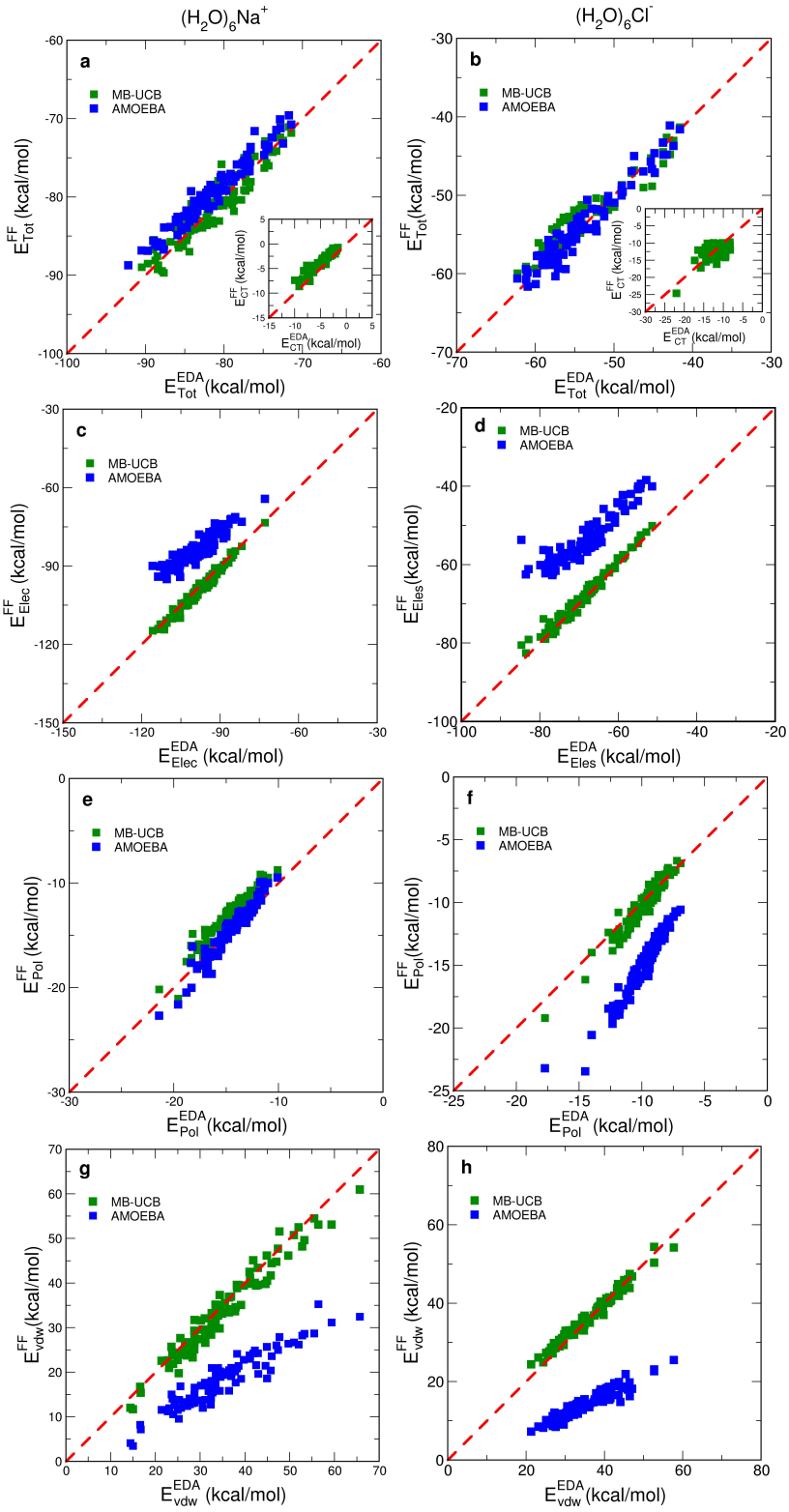

Figure 8: Correlation with DFT total energy and ALMO-EDA energy components obtained for $(H_2O)_6Na^+$ (left) and $(H_2O)_6Cl^-$ (right) from AMOEBA (blue) and MB-UCB (green). (a,b) total interaction energy correlation with DFT; CT for MB-UCB is provided in the inset. Correlation with ALMO-EDA for (c,d) permanent electrostatics; (e,f) polarization energy; (g,h) vdW energy.

correlates well but also is in good quantitative agreement with the ALMO-EDA polarization. Finally, AMOEBA compensate for very large error through the van der Waals interaction, but the error compensation is seen to be much smaller for MB-UCB. Nearly identical outcomes are reached for the other monovalent alkali cations and halogens as seen in Figures S25-S27.

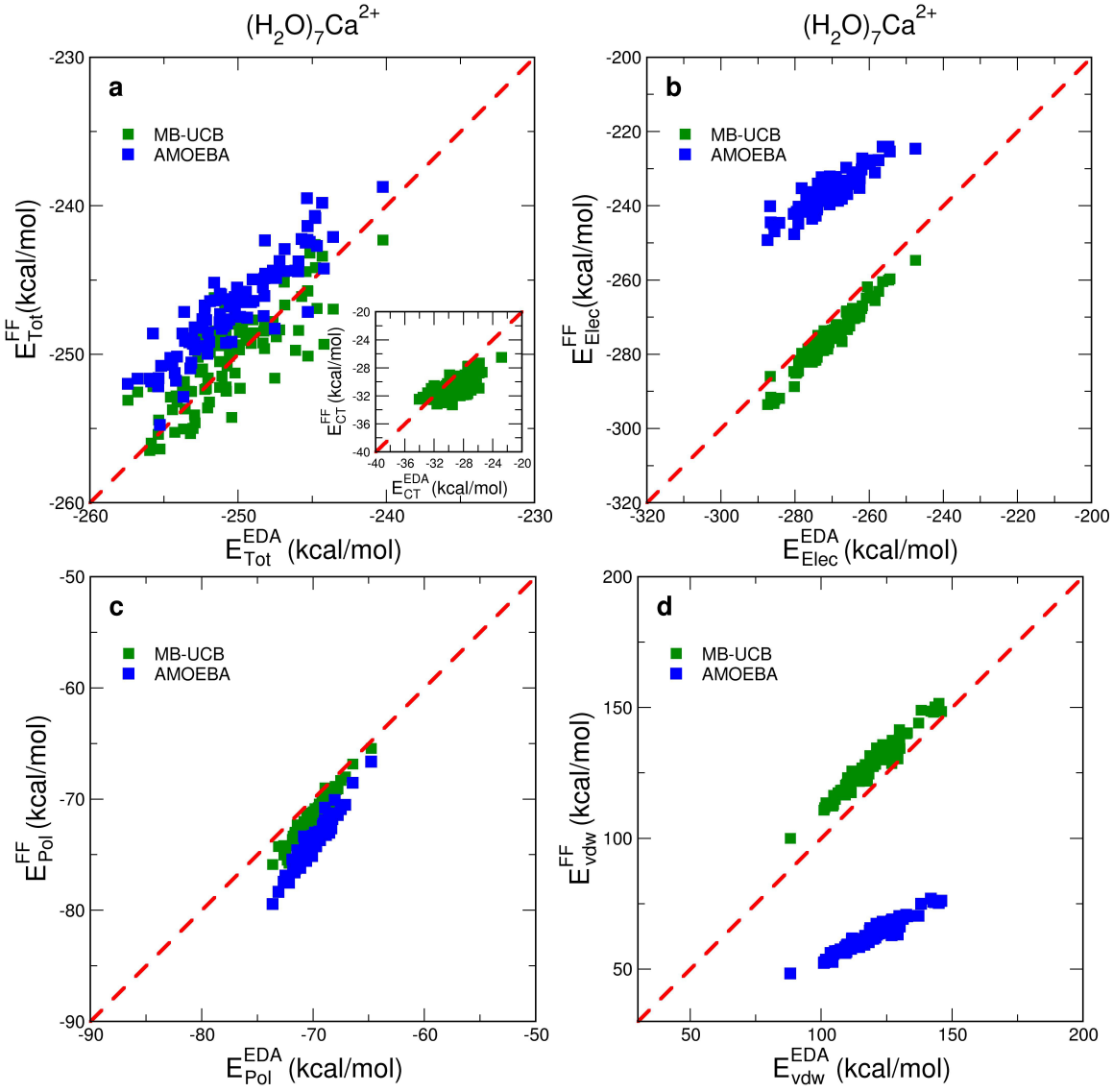

Figure 9: Correlation with DFT and ALMO-EDA for total energy and energy components obtained from AMOEBA and MB-UCB at the 8-body level for 100 MD extracted $(H_2O)_7Ca^{2+}$ configurations. (a) total interaction energy correlation with DFT for AMOEBA (blue) and MB-UCB (green); CT for MB-UCB is provided in the inset. (b) permanent electrostatic energy correlation with ALMO-EDA; (c) polarization energy correlation with ALMO-EDA; (d) van der Waals energy correlation with ALMO-EDA.

Figure 9 shows the correlation of total interaction energy along with the different energy components obtained from the AMOEBA and MB-UCB force fields with DFT/ALMO-EDA for the first hydration layer near the divalent Ca^{2+} ion. Considering the individual energy components of the $(H_2O)_7Ca^{2+}$ cluster, AMOEBA substantially underestimates the permanent electrostatics by as much as 50 kcal/mol (while also overestimating the polarization energy), and requiring massive underestimation of Pauli+dispersion energies by nearly 50 kcal/mol to compensate for these errors and the missing CT energy. MB-UCB slightly overbinds for both the permanent electrostatics and polarization energy with respect to the ALMO-EDA result, and while the charge transfer component has excellent correlation with the corresponding ALMO-EDA energy, the spread of the energy distribution is due to the associated limitations of the MB-UCB CT functional form. Hence the van der Waals energy shows a relatively small compensation effect for MB-UCB as seen in Figure 9 to yield total energies that are an overall improvement with respect to AMOEBA. Even so, the error compensation is not as precise for the $(H_2O)_7Mg^{2+}$ cluster as seen in Figure S28, although it greatly improves the MSE and RMSE with respect to AMOEBA as seen in Table 1.

An additional set of validations were performed for ion-ion and ion-water-ion interactions, as they were never considered in the MB-UCB force field parameterization. Figure 10 shows the interaction between two ions, Na^+ and Cl^-, as a function of their separation and it is clear that not only the total interaction energy agrees well with the DFT benchmark, but also show good correlations with the various ALMO-EDA decompositions such as electrostatics, polarization and vdW. Although the CT component is not as well reproduced by the MB-UCB force field, the CT contribution is very small and of relative insignificance when compared to the magnitude of the total interaction energy.

Finally we consider the solvent separated ion pair, in particular the Na^+-H_2O-Cl^- system, where both Na^+ and Cl^- are placed along the water bisector axis and the Na^+ ion is directed towards the water O atom and the Cl^- towards the two water H atoms. Distance scans are performed along the water bisector axis and where both Na^+ and Cl^- ions are moved from

Figure 10: Correlation of MB-UCB (green) and AMOEBA (blue) for Na^+Cl^- along the Na-Cl distance compared with DFT for total energy and ALMO-EDA energy components.(a) total interaction energy; (b) permanent electrostatic and polarization energy correlation with ALMO-EDA; (c) charge transfer energy correlation with ALMO-EDA for MB-UCB; (d) van der Waals energy correlation with ALMO-EDA.

the water oxygen simultaneously. Figure 11 shows the comparison of the total interaction energy and energy contributions obtained from ALMO-EDA, MB-UCB and AMOEBA. We find that MB-UCB better reproduces the total interaction energy at the equilibrium distance, with better agreement with the EDA energy components. In this most extreme example, AMOEBA exhibits poor cancellation of errors among energy components, resulting in an overstabilizing error of ∼5 kcal/mol and shift in the equilibrium position.

Figure 11: Correlation of MB-UCB (green) and AMOEBA (blue) for Na^+-H_2O-Cl^- compared with DFT for total energy and ALMO-EDA energy components. Scan is along the ratio of the Na-O and O-Cl distances: $r = \frac{d_{\mathrm{NaO}}}{d_{\mathrm{ClO}}}$. (a) total interaction energy; (b) permanent electrostatic and polarization energy; (c) charge transfer energy; (d) van der Waals energy correlation.

Table 2 reports the final parameters for the MB-UCB model for all ions investigated in this work, and water parameters are reported in our previous work.[27]

Table 2: Many-body MB-UCB force field parameters for various ions.

Ions	**Pol term**		**CP term**			**CT term**			**vdw term**	
	α (Å^3)	a	Z	α (Å^{-1})	β (Å^{-1})	α^{ct} (Å^3)	b	d	r^0 (Å)	ϵ (kcal/mol)
Li^+	0.2800	0.04553	2	3.95703	3.73790	0.99555	0.91364	1.00000	0.5059	0.0083
Na^+	0.1200	0.07963	8	3.17000	3.23700	0.96417	0.01064	1.00000	2.7354	0.1407
K^+	0.7800	0.39000	8	3.62583	3.01059	0.98635	0.00505	1.00000	3.7050	0.6005
Cs^+	2.2600	0.39000	8	5.20000	2.68500	1.02035	0.93844	1.00000	4.4621	0.6240
Ca^{2+}	0.5500	0.15850	8	3.93953	3.11579	0.02570	0.00846	1.00000	4.1826	0.1177
Mg^{2+}	0.0800	0.09520	8	2.79570	2.85620	0.00120	0.00120	1.00000	2.1513	0.1618
Cl^-	4.0000	0.22700	8	3.00000	2.59400	0.01284	0.01168	1.00000	5.5070	0.0065
Br^-	5.6500	0.21074	8	2.70000	2.33600	0.00586	0.00702	1.00000	5.7154	0.0084
I^-	7.2500	0.21647	8	2.40000	2.05900	0.01048	0.12704	1.00000	5.7181	0.0357

4 CONCLUSIONS

In summary, the many-body MB-UCB force field for various simple cations and anions has been developed using the 2-body and 3-body terms of the many-body expansion combined with ALMO-EDA as a guide for parameterization to minimize the errors in the individual physically relevant energy components. At the 2-body level the total interaction energy is chemically accurate for equilibrium geometries and in the asymptotic regions, with the largest errors of ∼1-2 kcal/mol observed in the overlapping region where configurations are more improbable due to the dominant effect of Pauli-repulsion at short-range. We also showed that the MB-UCB ion model is transferable and exhibits good accuracy for the ensemble of configurations for the primary hydration shells of all ions that includes 6-8 water molecules, and for ion pairing and ion-solvent separated states that will be important for concentrated salt solutions.

It is clear that force fields fitted to total energies and forces, and/or condensed phase data, can suffer from poor cancellation of errors as illustrated here for AMOEBA. New generation force fields such as MB-UCB and AMOEBA+ are now better grounded in physically motivated energy decomposition approaches such as SAPT[36] and variational EDAs,[27] as well as reliance on the many-body expansion to ensure better transferability.[12] We attribute the success of the MB-UCB model for ion-water systems in particular to a highly accurate water model,[26,27] and better electrostatic descriptions beyond polarization, such as charge penetration as also shown by other groups,[8,18,28] and introduction of a many-body charge transfer component to model all the ALMO-EDA energy components independently.

Even so, there are limitations to the current chosen functional forms used by advanced force fields. The empirical CT model used in MB-UCB is found to be only sufficient for the halogen anions when reproducing the ALMO-EDA CT energy, degrading in performance for monovalent and divalent cations. The damping function used for polarization is also problematic in the short-ranged that limits force field accuracy. We conclude that cancellation of errors using the limited form of the van der Waals energy function, while successful in the

equilibrium and asymptotic regions, is less successful so in the highly compressed regions and alternative formulations for short-ranged repulsion with more flexible functional forms would be desirable.[46] This is also a problem due to the arbitrariness in the "mixing rules" which are only qualitatively correct at best as showed by us previously.[47]

These areas will be directions we will return to in future work for MB-UCB, that in turn will better reduce the need to rely on cancellation of errors through the van der Waals component to reproduce the total interaction energy. But regardless of ions and ion-water cluster size, MB-UCB is found to more accurately reproduce the total interaction energy of the DFT benchmark, and we expect it will perform well for the condensed phase properties of aqueous ions, results that we will report in the near future.

5 ACKNOWLEDGMENTS

We thank the National Science Foundation under grant CHE-1955643 for support of this work. M. Liu thanks the China Scholarship Council for a visiting scholar fellowship. This research used computational resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. ED-AC02-05CH11231.

References

(1) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **2006**, *65*, 712–25.

(2) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **2010**, *78*, 1950–8.

(3) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmuller, H.; MacKerell, J., A. D. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods* **2017**, *14*, 71–73.

(4) Horta, B. A. C.; Fuchs, P. F. J.; van Gunsteren, W. F.; Hünenberger, P. H. New Interaction Parameters for Oxygen Compounds in the GROMOS Force Field: Improved Pure-Liquid and Solvation Properties for Alcohols, Ethers, Aldehydes, Ketones, Carboxylic Acids, and Esters. *Journal of Chemical Theory and Computation* **2011**, *7*, 1016–1031.

(5) Reif, M. M.; Hünenberger, P. H.; Oostenbrink, C. New Interaction Parameters for Charged Amino Acid Side Chains in the GROMOS Force Field. *Journal of Chemical Theory and Computation* **2012**, *8*, 3705–3723.

(6) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **1983**, *79*, 926–935.

(7) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *Journal of the American Chemical Society* **1996**, *118*, 11225–11236.

(8) Rackers, J. A.; Wang, Q.; Liu, C.; Piquemal, J.-P.; Ren, P.; Ponder, J. W. An optimized charge penetration model for use with the AMOEBA force field. *Physical Chemistry Chemical Physics* **2017**, *19*, 276–291.

(9) Van Vleet, M.; Misquitta, A.; Schmidt, J. New angles on standard force fields: towards a general approach for treating atomic-level. *Journal of Chemical Theory and Computation* **2017**, *14*, 739–758.

(10) Demerdash, O.; Yap, E. H.; Head-Gordon, T. Advanced potential energy surfaces for condensed phase simulation. *Annu Rev Phys Chem* **2014**, *65*, 149–74.

(11) Price, S. L. *Reviews in Computational Chemistry*; John Wiley Sons, 2000; pp 225–289.

(12) Demerdash, O.; Wang, L.-P.; Head-Gordon, T. Advanced models for water simulations. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2018**, *8*, e1355.

(13) Lopes, P. E. M.; Huang, J.; Shim, J.; Luo, Y.; Li, H.; Roux, B.; MacKerell, J. A. D. Polarizable Force Field for Peptides and Proteins Based on the Classical Drude Oscillator. *Journal of chemical theory and computation* **2013**, *9*, 5430–5449.

(14) Huang, J.; Lopes, P. E. M.; Roux, B.; MacKerell, A. D. Recent Advances in Polarizable Force Fields for Macromolecules: Microsecond Simulations of Proteins Using the Classical Drude Oscillator Model. *The Journal of Physical Chemistry Letters* **2014**, *5*, 3144–3150.

(15) Ren, P.; Ponder, J. W. Polarizable atomic multipole water model for molecular mechanics simulation. *The Journal of Physical Chemistry B* **2003**, *107*, 5933–5947.

(16) Ponder, J. W.; Case, D. A. *Protein Simulations*; Advances in Protein Chemistry; Academic Press, 2003; Vol. 66; pp 27 – 85.

(17) Ren, P.; Ponder, J. W. Polarizable Atomic Multipole Water Model for Molecular Mechanics Simulation. *The Journal of Physical Chemistry B* **2003**, *107*, 5933–5947.

(18) Piquemal, J.-P.; Gresh, N.; Giessner-Prettre, C. Improved formulas for the calculation of the electrostatic contribution to the intermolecular interaction energy from multipolar expansion of the electronic distribution. *The Journal of Physical Chemistry A* **2003**, *107*, 10353–10359.

(19) Cisneros, G. A.; Piquemal, J.-P.; Darden, T. A. Generalization of the Gaussian electrostatic model: Extension to arbitrary angular momentum, distributed multipoles, and speedup with reciprocal space methods. *The Journal of chemical physics* **2006**, *125*, 184101.

(20) Cisneros, G.; Tholander, S. N.-I.; Parisel, O.; Darden, T.; Elking, D.; Perera, L.; Piquemal, J.-P. Simple formulas for improved point-charge electrostatics in classical force fields and hybrid quantum mechanical/molecular mechanical embedding. *International journal of quantum chemistry* **2008**, *108*, 1905–1912.

(21) Amezcua, M.; Khoury, L.; Mobley, D. SAMPL7 Host–Guest Challenge Overview: assessing the reliability of polarizable and non-polarizable methods for binding free energy calculations. *Journal of Computer-Aided Molecular Design* **2021**, *35*, 1–35.

(22) Ding, Y.; Xu, Y.; Qian, C.; Chen, J.; Zhu, J.; Huang, H.; Shi, Y.; Huang, J. Predicting partition coefficients of drug-like molecules in the SAMPL6 challenge with Drude polarizable force fields. *Journal of Computer-Aided Molecular Design* **2020**, *34*, 421–435.

(23) McDaniel, J.; Schmidt, J. Physically-Motivated Force Fields from Symmetry-Adapted Perturbation Theory. *The journal of physical chemistry. A* **2013**, *117*, 2053–2066.

(24) McDaniel, J. G.; Schmidt, J. Next-Generation Force Fields from Symmetry-Adapted Perturbation Theory. *Annual Review of Physical Chemistry* **2016**, *67*, 467–488.

(25) Naseem-Khan, S.; Gresh, N.; Misquitta, A.; Piquemal, J.-P. Assessment of SAPT and Supermolecular EDA Approaches for the Development of Separable and Polarizable Force Fields. *Journal of Chemical Theory and Computation* **2021**, *17*, 2759–2774.

(26) Das, A. K.; Demerdash, O. N.; Head-Gordon, T. Improvements to the AMOEBA force field by introducing anisotropic atomic polarizability of the water molecule. *Journal of chemical theory and computation* **2018**, *14*, 6722–6733.

(27) Das, A. K.; Urban, L.; Leven, I.; Loipersberger, M.; Aldossary, A.; Head-Gordon, M.; Head-Gordon, T. Development of an Advanced Force Field for Water Using Variational Energy Decomposition Analysis. *J Chem Theory Comput* **2019**, *15*, 5001–5013.

(28) Slipchenko, L. V.; Gordon, M. S. Damping functions in the effective fragment potential method. *Molecular Physics* **2009**, *107*, 999–1016.

(29) Wang, B.; Truhlar, D. G. Including charge penetration effects in molecular modeling. *Journal of chemical theory and computation* **2010**, *6*, 3330–3342.

(30) Stone, A. J. Electrostatic damping functions and the penetration energy. *The Journal of Physical Chemistry A* **2011**, *115*, 7017–7027.

(31) Freitag, M. A.; Gordon, M. S.; Jensen, J. H.; Stevens, W. J. Evaluation of charge penetration between distributed multipolar expansions. *The Journal of Chemical Physics* **2000**, *112*, 7300–7306.

(32) Piquemal, J.-P.; Cisneros, G. A.; Reinhardt, P.; Gresh, N.; Darden, T. A. Towards a force field based on density fitting. *The Journal of chemical physics* **2006**, *124*, 104101.

(33) Spackman, M. A. The use of the promolecular charge density to approximate the penetration contribution to intermolecular electrostatic energies. *Chemical physics letters* **2006**, *418*, 158–162.

(34) Tafipolsky, M.; Engels, B. Accurate Intermolecular Potentials with Physically Grounded Electrostatics. *Journal of Chemical Theory and Computation* **2011**, *7*, 1791–1803, PMID: 26596442.

(35) Wang, B.; Truhlar, D. G. Partial atomic charges and screened charge models of the electrostatic potential. *Journal of chemical theory and computation* **2012**, *8*, 1989–1998.

(36) Liu, C.; Piquemal, J.-P.; Ren, P. AMOEBA+ classical potential for modeling molecular interactions. *Journal of chemical theory and computation* **2019**, *15*, 4122–4139.

(37) Wang, Q.; Rackers, J. A.; He, C.; Qi, R.; Narth, C.; Lagardere, L.; Gresh, N.; Ponder, J. W.; Piquemal, J.-P.; Ren, P. General model for treating short-range electrostatic

penetration in a molecular mechanics force field. *Journal of chemical theory and computation* **2015**, *11*, 2609–2618.

(38) Thole, B. T. Molecular polarizabilities calculated with a modified dipole interaction. *Chemical Physics* **1981**, *59*, 341–350.

(39) Wang, L.-P.; Head-Gordon, T.; Ponder, J. W.; Ren, P.; Chodera, J. D.; Eastman, P. K.; Martinez, T. J.; Pande, V. S. Systematic improvement of a classical molecular model of water. *The Journal of Physical Chemistry B* **2013**, *117*, 9956–9972.

(40) Laury, M. L.; Wang, L.-P.; Pande, V. S.; Head-Gordon, T.; Ponder, J. W. Revised parameters for the AMOEBA polarizable atomic multipole water model. *The Journal of Physical Chemistry B* **2015**, *119*, 9423–9437.

(41) Wang, W.; Skeel, R. D. Fast evaluation of polarizable forces. *The Journal of chemical physics* **2005**, *123*, 164107.

(42) Albaugh, A.; Demerdash, O.; Head-Gordon, T. An efficient and stable hybrid extended Lagrangian/self-consistent field scheme for solving classical mutual induction. *The Journal of Chemical Physics* **2015**, *143*, 174104.

(43) Albaugh, A.; Niklasson, A. M. N.; Head-Gordon, T. Accurate classical polarization solution with no self-consistent field iterations. *J. Phys. Chem. Lett.* **2017**, *8*, 1714–1723.

(44) Deng, S.; Wang, Q.; Ren, P. Estimating and modeling charge transfer from the SAPT induction energy. *Journal of computational chemistry* **2017**, *38*, 2222–2231.

(45) Halgren, T. A. The representation of van der Waals (vdW) interactions in molecular mechanics force fields: potential form, combination rules, and vdW parameters. *Journal of the American Chemical Society* **1992**, *114*, 7827–7843.

(46) Van Vleet, M. J.; Misquitta, A. J.; Stone, A. J.; Schmidt, J. R. Beyond Born–Mayer: Improved Models for Short-Range Repulsion in ab Initio Force Fields. *Journal of Chemical Theory and Computation* **2016**, *12*, 3851–3870, PMID: 27337546.

(47) Nerenberg, P. S.; Head-Gordon, T. Optimizing Protein-Solvent Force Fields to Reproduce Intrinsic Conformational Preferences of Model Peptides. *Journal of Chemical Theory and Computation* **2011**, *7*, 1220–1230, PMID: 26606367.

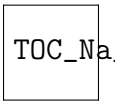

Figure 12: Graphics Abstract