Study of interactions between metal ions and protein model compounds by energy decomposition analyses and the AMOEBA force field

Zhifeng Jing, Rui Qi, Chengwen Liu, and Pengyu Ren^{a)}

Department of Biomedical Engineering, The University of Texas at Austin, Austin, Texas 78712, USA

(Received 1 June 2017; accepted 14 September 2017; published online 17 October 2017)

The interactions between metal ions and proteins are ubiquitous in biology. The selective binding of metal ions has a variety of regulatory functions. Therefore, there is a need to understand the mechanism of protein-ion binding. The interactions involving metal ions are complicated in nature, where short-range charge-penetration, charge transfer, polarization, and many-body effects all contribute significantly, and a quantitative description of all these interactions is lacking. In addition, it is unclear how well current polarizable force fields can capture these energy terms and whether these polarization models are good enough to describe the many-body effects. In this work, two energy decomposition methods, absolutely localized molecular orbitals and symmetry-adapted perturbation theory, were utilized to study the interactions between $\text{Mg}^{2+}/\text{Ca}^{2+}$ and model compounds for amino acids. Comparison of individual interaction components revealed that while there are significant charge-penetration and charge-transfer effects in Ca complexes, these effects can be captured by the van der Waals (vdW) term in the AMOEBA force field. The electrostatic interaction in Mg complexes is well described by AMOEBA since the charge penetration is small, but the distance-dependent polarization energy is problematic. Many-body effects were shown to be important for protein-ion binding. In the absence of many-body effects, highly charged binding pockets will be over-stabilized, and the pockets will always favor Mg and thus lose selectivity. Therefore, many-body effects must be incorporated in the force field in order to predict the structure and energetics of metalloproteins. Also, the many-body effects of charge transfer in Ca complexes were found to be non-negligible. The absorption of charge-transfer energy into the additive vdW term was a main source of error for the AMOEBA many-body interaction energies. *Published by AIP Publishing.* <https://doi.org/10.1063/1.4985921>

I. INTRODUCTION

Metal ions play a major role in many biological processes, such as cell signaling, maintenance of osmotic pressure, and catalysis. Dysfunction of metalloproteins is associated with a wide variety of diseases^{1–4} including cancers, infectious diseases, and neurodegenerative disorders.^{5,6} Therefore, understanding the interactions between metal ions and proteins and how these interactions help define specificity in ion-protein binding have broad impact. Specifically, Mg and Ca are among the most abundant ions in natural metalloproteins. Ca and Mg are similar as they are both hard ions that preferably bind to oxygen atoms, although Mg binding tends to be stronger.^{7,8} Ca-binding proteins can selectively bind Ca against a background of much higher concentrations of Mg, which is crucial for their functions. The selective binding between Mg and Ca has been a topic of numerous experimental and theoretical studies.^{9–14}

Molecular dynamics simulations based on classical force fields are a promising tool for probing protein-ion interactions. However, due to the complexity of the interactions between highly charged species, these systems remain

challenging for classical computational models.¹⁵ The short-range electrostatic interactions, such as charge penetration, charge transfer (CT), and many-body polarization effects, are critical in shaping ion-protein binding potential energy surfaces and thus determining the specificity and selectivity.^{16,17} Unfortunately, these effects are missing in the fixed-point-charge models such as AMBER,¹⁸ CHARMM,¹⁹ OPLS-AA,²⁰ and GROMOS.²¹ The ion parameters of these models are usually fitted to the experimental solvation free energies,²² which are not transferable to the protein environment.

Significant efforts have been made to develop polarizable force fields for biomolecular simulations.^{23–29} For example, three main approaches have been designed to explicitly treat the electronic polarization effect: fluctuating-charge,^{30–33} Drude oscillator,^{34–38} and induced dipole models.^{39–43} The AMOEBA force field utilizes atomic induced dipoles to represent the response to the environment.^{42,43} The AMOEBA model well captures many-body interactions of water and the structural stability of metalloproteins.^{44–46} However, a polarizable force field may fail to represent systems in which a significant charge transfer occurs.¹³ Among others,^{47,48} the SIBFA (sum of interactions between fragments *ab initio* computed)⁴⁹ force field is an example that has an explicit charge-transfer term. For transition metals, there are electronic effects such as the Jahn–Teller distortion and the trans influence

^{a)}Author to whom correspondence should be addressed: pren@mail.utexas.edu.

that may need special treatment. The Jahn–Teller effect can be modeled effectively by the angular overlap model or the valence bond theory (e.g., LFMM,⁵⁰ SHAPES,⁵¹ and VALBOND⁵²). LFMM⁵³ and VALBOND⁵⁴ force fields have also been extended to include the trans influence. While these models were originally developed for non-polarizable force fields, they can also be incorporated into polarizable force fields.^{55,56}

Quantum mechanical (QM) calculations can provide an accurate description of ion-protein interactions,^{57,58} although QM calculations on large systems are impractical due to their computational high cost. Among them, energy decomposition analysis (EDA) methods decompose the interaction energy into several well-defined physical terms. For example, symmetry-adapted perturbation theory (SAPT) partitions the interaction energy into electrostatic, induction, London dispersion, and exchange-repulsion.⁵⁹ The sign and magnitude of these individual terms are useful both for a quantitative understanding of protein-ligand binding and for the evaluation of empirical force field.

In this work, the EDA is performed to understand the effect of polarization, charge-transfer, and many-body interactions on protein-ion binding. The quality of the additive and non-additive terms of the AMOEBA model is evaluated based on the EDA. In particular, whether the polarization model in AMOEBA is good enough to capture the many-body effect of highly charged systems is discussed.

II. METHODS

A. AMOEBA force field

A detailed description of the function forms of the AMOEBA force field can be found elsewhere.⁴³ Here we briefly summarize the aspects of AMOEBA that are relevant to the current work. The total energy is expressed as the sum of valence terms and non-bonded terms. The non-bonded terms include permanent electrostatics, polarization, and van der Waals (vdW) interaction, which are responsible for the intermolecular interaction energy. The permanent electrostatics is represented by atomic multipoles up to quadrupole. The polarization energy arises from induced dipoles. The electric field produced by permanent electrostatics and induced dipoles induces a dipole at each polarizable atomic site. The magnitude of the induced dipole is related to an isotropic polarizability defined for each atom type. In addition, the polarization is damped at the short range according to the charge distribution

$$\rho = \frac{3a}{4\pi} \exp(-au^3), \quad (1)$$

where a is a dimensionless width parameter of the charge distribution and controls the strength of damping, and $u = R_{ij}(\alpha_i\alpha_j)^{-1/6}$ is an effective distance as a function of the distance between and the atomic polarizabilities of sites i, j . All the atoms except metal ions share the same damping parameter $a = 0.39$, which is based on a fit to water cluster energies. For metal ions, it is necessary to optimize a to account for the different charge distributions. The vdW interactions are described by Halgren's buffered 14–7 function, which is slightly softer in the repulsive region than the Lennard-Jones 12–6 function.

B. Energy decomposition analysis (EDA) methods

An energy decomposition analysis interprets the interaction energy in terms of physical meaningful contributions. ALMO (Absolutely Localized Molecular Orbital) decomposes the total interaction energy as $E_{\text{int}} = E_{\text{froz}} + E_{\text{pol}} + E_{\text{CT}}$. The frozen orbital interaction E_{froz} represents the energy change when fragments approach each other without relaxation of the fragment orbitals. It contains contributions from dispersion, Pauli repulsion, and permanent electrostatics, $E_{\text{froz}} = E_{\text{disp}} + E_{\text{exch}} + E_{\text{elst}} = E_{\text{vdW}} + E_{\text{elst}}$. The polarization energy E_{pol} is the energy change when fragments are allowed to respond to weak electric fields. And the remainder charge-transfer (CT) term E_{CT} accounts for the donor-acceptor interactions. The sum of polarization and charge transfer is the total induction energy, $E_{\text{ind}} = E_{\text{pol}} + E_{\text{CT}}$. For simplicity, the text in the subscript will be used to refer to the energy components.

SAPT (Symmetry-Adapted Perturbation Theory) directly calculates the intermolecular interaction using perturbation. The interaction energy is divided into terms from the Rayleigh-Schrödinger (RS) perturbation theory expansion and terms from symmetry forcing. The first-order energy corresponds to the electrostatics interaction between charge distributions of isolated monomers. The second-order energy can be decomposed into induction and dispersion energies. Although the first- and second-order energy corrections are the major contributions to the asymptotic intermolecular interactions, higher order terms can be important for polar systems.⁵⁹ In practice, several levels of SAPT can be used. SAPT2 denotes the use of second-order corrections in the monomer fluctuation potential, which yields asymptotic equivalence to the MP2 interaction energy. SAPT2+ further incorporates the influence of intramolecular correlation on the dispersion energy.⁵⁹ SAPT is free from basis set superposition error. However, it does not provide a unique separation of charge-transfer energy since it is unclear how to decompose the higher order terms.^{60,61}

C. Calculation of interaction energy and many-body energy

The total interaction energy is calculated as $E_{\text{int}} = E_{\text{total}} - \sum_i E_i$ where the summation is over all fragments and all the energies are evaluated at the same geometry. In this work, one fragment is not necessarily one molecule. For example, to calculate the interaction between a metal ion and a pocket, the entire pocket consisting of multiple ligands is treated as one fragment. The many-body energy is calculated as

$$E_{\text{MB}} = E_{\text{int}} - \sum_{i,j} E_{\text{int}}(i,j) = E_{\text{total}} - \sum_i E_i - \sum_{i,j} (E_{\text{total}}(i,j) - E_i - E_j), \quad (2)$$

i.e., the many-body energy equals the total interaction energy minus the sum of two-body interaction energies between all pairs of fragments in the complex.

D. Optimization of force field parameters

The original vdW and damping parameters for Mg and Ca were based on Mg/Ca...water dimer interaction energies at the MP2/aug-cc-pVTZ or MP2/6-311++G(3df,3pd) level,⁶² which underestimate the magnitude of the binding energies

between Mg/Ca and water by as much as 5 kcal/mol (less negative). As a result, the AMOEBA interaction energies were also consistently underestimated. In addition, due to limited simulation time, the solvation free energies have significant statistical errors. A new set of parameters were obtained by comparing to Mg/Ca water dimer interaction energies at the MP2/CBS level.

The POLTYPE program⁶³ was used to obtain initial parameters for acetate and acetamide. The multipoles from GDMA were fitted to the electrostatic potential. The acetate multipoles obtained from electrostatic potential fitting gave too negative potentials along the O–C–O bisector, which affects the relative energy between bidentate and monodentate binding. By setting the quadrupoles on carboxylate C and O to 0, the correct potentials for bidentate and monodentate can be reproduced (Fig. S2 of the [supplementary material](#)). We note that the quadrupoles have no noticeable effect on hydration free energy. The carboxylate O vdW parameters were subsequently optimized based on acetate···water dimer interactions. The parameters were validated by comparing the solvation free energy with the experimental result. The pairwise vdW parameters between acetamide-O and Mg/Ca were also optimized according to the Mg/Ca···acetamide interaction energy at the MP2/CBS level.

E. Comparison between computed and experimental ionic solvation free energies

The single-ion solvation free energies cannot be measured directly by experiments since the solution must remain neutral. Once the absolute solvation free energy of the proton is set, the solvation free energies of all the other ions can be obtained by

$$\Delta G_{\text{solv}}^o = \Delta G_{\text{solv}}^{o,\text{conv}} + z\Delta G_{\text{solv}}^o(H^+), \quad (3)$$

where ΔG_{solv}^o is the standard solvation free energy of a single ion using an ideal gas at 1 atm as the standard state, $\Delta G_{\text{solv}}^{o,\text{conv}}$ is the conventional solvation free energy which can be calculated from measurable quantities, and z is the charge of the ion.⁶⁴ It can be seen from Eq. (3) that $\Delta G_{\text{solv}}^o(H^+)$ does not affect the sum of absolute solvation free energies of neutral ion pairs. Following previous theoretical work,⁶⁵ the standard state of an ideal gas at a concentration of 1 mol/L will be used, and the corresponding free energies will be denoted by a superscript asterisk. At 298 K, $\Delta G_{\text{solv}}^* = \Delta G_{\text{solv}}^o - 1.9$ kcal/mol. A summary of approaches used to derive $\Delta G_{\text{solv}}^o(H^+)$ can be found in the work of Grossfield *et al.*⁶⁶ There are large discrepancies between the reported values of $\Delta G_{\text{solv}}^o(H^+)$, and most of them fall within two categories that differ by ~8–15 kcal/mol.^{67–69} This difference has been attributed to the water surface potential.⁶⁸ For example, the extrathermodynamic assumptions of Schmid *et al.*⁶⁴ and Marcus⁷⁰ do not include the surface potential, while the cluster-pair approximation of Tissandier *et al.*⁷¹ does. Using the proton solvation free energy of Schmid *et al.*, the solvation free energies from various sources can be converted to values that do not include the surface potential.

The AMOEBA solvation free energies of Mg^{2+} , Ca^{2+} , and acetate were calculated by alchemically growing a single ion in a periodic box of water. The long-range electrostatics was handled by the particle mesh Ewald method (PME). No corrections for the finite size effects were applied since they are very

small for a system containing 500 or more water molecules.^{69,72,73} The calculated results do not include the surface potential, so they can be directly compared to the experimental solvation free energies using the reference proton solvation free energy of Schmid *et al.* This scheme for comparing computed and experimental ionic solvation free energies has been adopted by several studies,^{68,72,73} and it is considered to be one of the most reliable ways.⁷²

The surface potential of the AMOEBA water model was also computed in order to derive the “real” solvation free energies. The surface potential can be obtained by integrating the contributions of atomic charges and atomic dipoles from a simulation of a periodic slab geometry of water,

$$\phi_q(z) - \phi_q(z_0) = - \int_{z_0}^z E_z(z') dz', \quad (4)$$

where $\phi_q(z)$ is the contribution of charge to the electric potential and $E_z(z')$ is the average z -component of the electric field at position z' due to point charges,

$$\phi_\mu(z) - \phi_\mu(z_0) = \frac{1}{\epsilon_0} \int_{z_0}^z \rho_\mu(z') dz', \quad (5)$$

where $\rho_\mu(z')$ is the z -component of the dipole density on the plane parallel to the liquid-air interface at position z' , which includes permanent dipoles and induced dipoles. The average contribution of the traceless atomic quadrupoles in AMOEBA is 0.

F. Computational details

Dimers of Mg/Ca and ligands including water, acetate, acetamide, and imidazole were used in this study. These small molecules are model compounds for Asp/Glu and His side chains and protein backbones, which are the most common ligands in Mg/Ca binding pockets. The dimers were first minimized at the MP2/cc-pVTZ level, and then the interaction energy curves were scanned by keeping the geometry fixed and varying the intermolecular distance from 0.8 to 1.5 equilibrium distances. Models for monodentate binding of acetate to Mg/Ca were also built by setting the C–O–M angle to 150°, where C, O, and M denote the carboxylate C and O atoms and metal ions. The structures of the dimers are shown in Fig. 1. The interaction energies were evaluated at the MP2/aug-cc-pVQZ and MP2/aug-cc-pV5Z levels of theory and extrapolated to MP2/CBS by the two-point Helgaker extrapolation scheme.⁷⁴ In addition, the CCSD(T)/CBS (coupled cluster with a full treatment of singles and doubles and an estimate to the connected triples contribution, together with the complete basis set) interaction energies

FIG. 1. Structures of the metal ion-ligand dimers. White, cyan, blue, red, pink, and green spheres are H, C, N, O, Mg, and Ca atoms.

at the equilibrium distances were obtained by $E_{\text{CCSD(T)/CBS}} = E_{\text{MP2/CBS}} + E_{\text{CCSD(T)/aTZ}} - E_{\text{MP2/aTZ}}$. SAPT2+/aug-cc-pVTZ and ALMO/ ω B97x-D/def2-TZVPPD analyses were carried out at the same geometries. The Gaussian 09 software⁷⁵ was used for geometry optimization, the Psi4 package⁷⁶ was used for DF-MP2 and SAPT2+ calculations, and Q-Chem 4.2⁷⁷ was used for the ALMO analysis.

Dimer structures away from the optimized orientations were also constructed by randomly placing metal ions around the ligating atoms within a sphere of radius 3.5 Å. For acetate, QM calculations on Ca/Mg...acetate dimers in vacuum predict +1 charge on the metal ion and 0 charge on acetate at long separations,¹² which correspond to different species from the force field models. So only Mg/Ca...O distances shorter than 2.8 Å were considered. 60 structures for each dimers of Mg/Ca...water/acetamide/imidazole and 20 structures for Mg/Ca...acetate were obtained from the above procedures. The angles of ligands in these dimer structures were also modified to generate distorted structures. The O–C–O angle in acetate was fixed at 110°/130°, the C–C–O angle in acetamide was fixed at 110°/130°, and the C–N–C angle (around the pyridine N) in imidazole was fixed at 100°/110°, respectively. The interaction energies of these dimers were calculated at the MP2/CBS level with the same scheme as mentioned above.

EF hands were chosen as the model compound for metal ion-pocket interaction, which include four Asp/Glu side chains and one backbone carbonyl group in the first coordination shell. The coordinating ligands were represented by acetate and acetamide molecules. The geometries of the bidentate and monodentate Ca-pockets and monodentate Mg-pocket were taken from PDB (5cpv, 1b8l, and 1b8c, respectively), without further optimization. The geometry for the bidentate Mg-pocket was obtained from a short AMOEBA MD simulation at 300 K starting from the 5cpv PDB structure with Ca replaced by Mg. The interaction energies of the ion-pocket binding are calculated by MP2/aug-cc-pVTZ, and the charge transfer energies are evaluated by ALMO/HF/aug-cc-pVDZ. Comparison of different methods and basis sets for the calculation of charge transfer energies is shown in Table S1 of the [supplementary material](#).

The Ca, Mg, and acetate solvation simulations were performed on the TINKER-OpenMM package.^{78,79} The Bennett acceptance ratio method was employed. In each simulation, one metal ion is solvated in a cubic simulation box with a box length of 24.6 Å and a density of 1.0 g/cm³, which corresponds to ~500 water molecules. Additionally, larger simulation boxes of 1600 and 4000 water molecules for Mg solvation were used to verify if the finite-size effects are negligible. The non-bonded interactions of the ion were coupled to a parameter λ where $\lambda=0$ corresponds to no interaction and $\lambda=1$ corresponds to full interaction. The vdW interaction was first increased from 0 to 1, and then the electrostatics and polarization interaction was increased from 0 to 1 with the vdW term equal to the full interaction. A standard softcore modification for vdW as implemented in TINKER-OpenMM was used to allow for a smoother transition when λ is close to 0. There are 20 λ states in total, and for each λ state, a 10-ns NVT simulation at 300 K was carried out with a 2-fs time step and coordinates were saved every 5 ps for analysis. A

periodic box containing 1600 water molecules with dimensions $36.342 \times 36.342 \times 100$ Å³ was used for the surface potential calculation. The total simulation time is 2 ns and coordinates were saved every 1 ps for analysis. All the other simulation parameters, including the vdW cutoff, the polarization PME parameters, and thermostat, are standard and can be found elsewhere.^{80,81}

III. RESULTS AND DISCUSSIONS

A. Force field parameters

The optimized force field parameters for Mg²⁺, Ca²⁺, acetate O, and acetamide O are summarized in Table I. Compared to AMOEBA09, Mg²⁺ and Ca²⁺ both become smaller to make the Mg/Ca...water interaction energy lower. The solvation free energies of Mg²⁺, Ca²⁺, and acetate are listed in Table II. The experimental solvation free energies were converted using a standard state of 1 mol/L in gas phase and a reference proton hydration free energy of $\Delta G_{\text{solv}}^*(H^+) = -251.4$ kcal/mol.⁶⁴ The sum of proton and hydroxide solvation free energies is also shown for comparison between experimental results. There is an excellent agreement between AMOEBA and experimental solvation free energies for Mg²⁺ and Ca²⁺. For acetate, the AMOEBA result coincides with the upper bound of experimental values. Considering the large discrepancies between experimental solvation free energies of acetate and the uncertainties of each value (~2 kcal/mol),

TABLE I. Optimized force field parameters for Mg, Ca, acetate, and acetamide.

Polarization and vdW				
	α	a	R_0 (Å)	ϵ (kcal/mol)
Mg ²⁺	0.08	0.1150	2.90	0.280
Ca ²⁺	0.55	0.1800	3.59	0.350
Acetate O	1.20	0.3900	3.55	0.095
Pairwise vdW				
			R_0 (Å)	ϵ (kcal/mol)
Acetamide O	Mg ²⁺		3.00	0.153
Acetamide O	Ca ²⁺		3.27	0.173

TABLE II. Experimental and computed ionic solvation free energies (kcal/mol). The standard state is 1 mol/L in gas phase. The proton solvation free energy of Schmid *et al.* $\Delta G_{\text{solv}}^*(H^+) = -251.4$ kcal/mol is taken as the reference. The standard deviations of the last digit of the calculated AMOEBA results are shown in parentheses. Note: Marcus *et al.* incorrectly used $\Delta G_{\text{solv}}^* = \Delta G_{\text{solv}}^o + 1.9$ kcal/mol, and therefore -3.8 kcal/mol is added to the values reported therein.

	Mg ²⁺	Ca ²⁺	Acetate	H ⁺ + OH ⁻
AMOEBA	-434.8(3)	-356.9(3)	-86.9(3)	
Schmid ⁶⁴	-434.9	-356.8		-370.7
Marcus ⁷⁰	-435.4	-357.7	-94.0	-360.9
Pliego ^{82,83}			-87.1	-372.8
			-93.7	-372.8
Truhlar ⁶⁵			-92.1	-370.6

this agreement is satisfactory. The acetate... water interaction also matches the MP2/CBS values with errors of less than 2 kcal/mol (Fig. S3 of the [supplementary material](#)). Although it is possible to further lower the AMOEBA solvation free energy by optimizing the parameters, the interaction between acetate and water/metal ions would become too favorable. Therefore, this set of acetate parameters achieves a balance between matching experimental and QM reference data. The AMOEBA solvation free energies of Mg^{2+} from simulations with 500, 1600, and 4000 water molecules are -434.8 ± 0.3 , -434.8 ± 0.3 , and -434.9 ± 0.3 kcal/mol, respectively. This confirms that the finite size effects are negligible.

The surface potential of the AMOEBA water model is -0.352 ± 0.002 V, which translates to -8.1 kcal/mol for $+1 e$ charge. Combined with the intrinsic proton solvation free energies of Schmid *et al.* and Marcus *et al.*, this leads to estimated real proton solvation energies of -259.5 and -262.4 kcal/mol, respectively. These values are not compatible with the popular value of -265.9 kcal/mol from Truhlar *et al.*, but are close to a recent value of -262.2 kcal/mol estimated from large ion-water clusters.^{69,84} We note that the surface potential of the AMOEBA water model is smaller in magnitude than the values of other water models around -0.5 V.^{67,85} Quantum mechanical calculations also suggest that the AMOEBA geometry has a smaller surface potential than the geometries obtained from fixed charge models.⁸⁶ The AMOEBA solvation free energy results that include this surface potential contribution are tabulated and compared with experimental references with $\Delta G_{\text{solv}}^*(H^+) = -265.9$ kcal/mol and -262.2 kcal/mol in Table S2 of the [supplementary material](#).

The binding energies and equilibrium distances for all dimers are summarized in Table III. With the exception of the $\text{Mg}\cdots\text{acetate}$ dimer, the errors of the MP2 results are less than 0.5 kcal/mol. Therefore, MP2 results will be used as a benchmark for the total interaction energies. The percentage errors of ALMO and SAPT total interaction energies are around 2%, with SAPT interaction energies being lower than the ALMO interaction energies. Since the magnitude of the $\text{Mg}/\text{Ca}\cdots\text{acetate}$ interaction energies is larger than the others, the absolute errors are also larger with a maximum of ~ 10 kcal/mol. The binding energies can be reproduced by AMOEBA with a maximum percentage error of 6% for $\text{Ca}\cdots\text{acetamide}$ and $\text{Ca}\cdots\text{imidazole}$. To reach this

agreement, all equilibrium distances are underestimated except for $\text{Mg}/\text{Ca}\cdots\text{acetate}$ dimers. The dimer interaction energies and the component breakdown will be analyzed in detail.

B. Metal ion-ligand dimer interactions

1. $\text{Mg}/\text{Ca}\cdots\text{water}$

With the optimized parameters for Mg and Ca, good agreement was obtained for the $\text{Mg}/\text{Ca}\cdots\text{water}$ interaction energies (the top panels of Fig. 2). The equilibrium interaction energies can be well reproduced with an error of ~ 1 kcal/mol. The equilibrium distance for the $\text{Ca}\cdots\text{water}$ dimer also agrees with the MP2/CBS result. For $\text{Mg}\cdots\text{water}$, a compromise must be made between the interaction energy and the equilibrium distance, so the AMOEBA equilibrium distance is ~ 0.1 Å shorter than the reference value. The interaction energy curve of the $\text{Ca}\cdots\text{water}$ dimer is accurate in the entire range. The absolute error with respect to MP2/CBS is less than 1 kcal/mol in the repulsive region; the equilibrium to mid-range suffer from relatively large errors, with the biggest error being 3.1 kcal/mol at a $\text{Ca}\cdots\text{O}$ distance of 2.7 Å. For the $\text{Mg}\cdots\text{water}$ dimer, the interaction is reasonably good in the repulsive region, with an error of -3.9 kcal/mol at the shortest separation of 1.55 Å. Similar to Ca, the largest error is seen at the mid-range with a distance of ~ 2.5 Å.

The performances of ALMO and SAPT total interaction energies are almost identical to each other for the $\text{Ca}\cdots\text{water}$ dimer. Both methods are less attractive than the MP2/CBS benchmark. The error is 1-3 kcal/mol across the separation distances. Similarly, ALMO and SAPT underestimate the interaction energies of the $\text{Mg}\cdots\text{water}$ dimer, while the error is more pronounced than that of the $\text{Ca}\cdots\text{water}$ dimer. The error is 2-5 kcal/mol in the equilibrium to the repulsive region for SAPT and is 5-9 kcal/mol in the same region for ALMO. The errors can be attributed to the limited basis-set size for the SAPT calculation and the quality of the density functional used for ALMO. It has been shown that a newer version of the wB97 family of functionals, wB97X-V, achieves higher accuracy for the ion-water systems.

A comparison of individual components of the total interaction energy is helpful for understanding the source of errors for AMOEBA. The energy decomposition results

TABLE III. Dimer binding energies and equilibrium distances. The equilibrium distances are defined as the distance between $\text{Mg}^{2+}/\text{Ca}^{2+}$ and the closest O, C, O, and N atoms in water, acetate, acetamide, and imidazole, respectively.

System	Binding energy (kcal/mol)					Distance (Å)	
	AMOEBA	CCSD(T)	MP2	SAPT2+	ALMO	AMOEBA	MP2
$\text{Mg}\cdots\text{water}$	-83.54	-82.75	-82.70	-80.79	-80.28	1.82	1.93
$\text{Ca}\cdots\text{water}$	-56.02	-56.98	-57.39	-55.64	-53.96	2.18	2.25
$\text{Mg}\cdots\text{acetate}$	-390.44	-382.70	-380.91	-390.61	-382.31	2.25	2.25
$\text{Ca}\cdots\text{acetate}$	-325.21	-325.65	-325.50	-335.40	-320.89	2.54	2.51
$\text{Mg}\cdots\text{acetamide}$	-139.59	-138.87	-137.53	-140.34	-143.58	1.64	1.80
$\text{Ca}\cdots\text{acetamide}$	-98.29	-104.21	-103.92	-105.63	-105.42	1.94	2.06
$\text{Mg}\cdots\text{imidazole}$	-140.44	-136.86	-136.10	-138.55	-133.71	1.82	1.97
$\text{Ca}\cdots\text{imidazole}$	-102.30	-96.04	-96.40	-99.21	-90.79	2.11	2.25

FIG. 2. Interaction energy and energy decomposition of the Mg/Ca...water dimer as a function of the Mg/Ca...O distance.

from ALMO, SAPT, and AMOEBA are shown in the middle and bottom panels of Fig. 2. For both Mg and Ca, the contributions from Frz and Ind are comparable at equilibrium and longer distances, while the repulsive Frz interaction dominates at the short range. The ALMO total induction energy (Pol + CT) agrees well with the SAPT induction energy. The largest difference for the Mg...water dimer is ~ 2 kcal/mol at the equilibrium distance, which is acceptable considering their difference in the total interaction energy. For the Ca...water dimer, the SAPT induction is slightly softer than the ALMO induction at short range, and the discrepancy vanishes at equilibrium and long ranges. Overall this good agreement indicates that the partition of total interaction into frozen orbital energy and induction energy is reliable. In Secs. III B 2–III B 4, we will adopt a composite approach where the total induction energy is separated by ALMO and the frozen density energy is separated by SAPT.

The magnitude of charge transfer energy can be identified by the difference between ALMO polarization energy (Pol) and ALMO total induction energy (Pol + CT). The charge transfer energy for the Ca...water dimer given by ALMO (-7.5 kcal/mol) is significantly larger than the charge transfer energy for the Mg...water dimer (-2.7 kcal/mol). This result is consistent with a second generation ALMO

analysis,⁸⁷ but is contrary to the chemical trend. Mg has larger ionization energies than Ca, and in the ion-water systems, Mg transfers more charge to water than Ca does.⁸⁸ Therefore, Mg-complexes should generally have larger charge transfer energies than Ca-complexes. Interestingly, it has also been shown that the trend of charge-transfer energies between a monovalent ion and water given by ALMO/wB97X-D3 is $\text{Li} > \text{K} > \text{Na}$, whereas the correct chemical trend should be $\text{Li} > \text{Na} > \text{K}$.⁶¹ Therefore, the unphysical behavior of charge transfer energies of second- and third-period metal ions is a general problem with ALMO.

By comparing the AMOEBA polarization energy with ALMO total induction and polarization for the Ca...water dimer, it can be found that AMOEBA polarization agrees well with ALMO polarization. On the other hand, the AMOEBA Frz energy agrees better with ALMO Frz + CT than with ALMO Frz. AMOEBA polarization tends to be much more attractive at the short range and less attractive beyond the equilibrium distance. The error is more noticeable for the Mg...water dimer. The performance of AMOEBA Frz energy, when compared to ALMO Frz + CT, is reasonably good for both Mg...water and Ca...water over the entire range.

Further decomposition of the Frz energy into electrostatic and vdW contribution (see Fig. 3) reveals that the AMOEBA

FIG. 3. Electrostatics and vdW interaction of the Mg/Ca...water dimers.

vdW interaction for the Mg...water dimer agrees very well with SAPT vdW. The AMOEBA electrostatics interaction is accurate and only starts to deviate from SAPT electrostatics at the very short range, with a maximum error of 11 kcal/mol (less attractive). In sharp contrast, the AMOEBA electrostatics interaction for Ca...water has a maximum error of 31 kcal/mol. This error is compensated mainly by the less repulsive buffered 14-7 vdW interaction. Therefore, the accuracy of AMOEBA Frz energy for Ca...water relies on the softened vdW interaction to capture the charge-penetration and charge transfer effects. We expect that a carefully parameterized charge-penetration model⁸⁹ will also capture the charge-penetration effect and have better transferability. For Mg...water, the good agreement is the result of the accurate electrostatics and the vdW interaction. It should be noted that the close agreement between AMOEBA Frz energy and ALMO Frz + CT may be a coincidence due to the improperly assigned CT energy. If the CT energy for Ca...water is reliable, then the Mg...water CT energy should be even larger, and the AMOEBA Frz energy will be much more repulsive than the true Frz + CT energy.

2. Mg/Ca...acetate

The interactions between Mg/Ca and acetate are more relevant to ion-protein binding, since Asp and Glu are the most abundant residues in Mg/Ca binding pockets. The interaction energies for Mg/Ca...acetate dimers, each with two orientations representing bidentate and monodentate binding, are shown in Fig. 4. The equilibrium distances and equilibrium interaction energies can be reproduced by AMOEBA, with errors of -5.2 and -8.1 kcal/mol for bi- and monodentate binding with Mg and -0.6 and +3.3 kcal/mol for

bi- and mono-dentate binding with Ca. The errors are small compared to the total interaction energies of -300 to -400 kcal/mol. Similar to the ion-water systems, the long-range behavior of the total interaction energy is better for Ca...Ac⁻ than it is for Mg...Ac⁻. Besides, the monodentate binding from AMOEBA for both Mg and Ca is better than their bidentate counterpart. At short ranges, both Mg...Ac⁻ and Ca...Ac⁻ interaction energies are underestimated by AMOEBA.

The AMOEBA polarization energy for Mg...Ac⁻ is steeper than the ALMO polarization energy and is overly attractive at short ranges. The error for bidentate binding is about twice as much as that for monodentate binding. The fact that AMOEBA polarizes more in the bidentate orientation is easy to understand—the metal ion approaches two polarizable O sites in the bidentate orientation compared to one site in the monodentate orientation. However, the difference between ALMO polarization energies in the two orientations is less significant. This may be explained by the charge redistribution when the metal ion is close to one O atom. Since the electrostatics interaction is almost perfectly matched, the error in the polarization energy is due to the inherent functional form of the polarization model.

The Ca...Ac⁻ polarization energy is better described by AMOEBA than the Mg...Ac⁻ polarization. In the bidentate configuration, the AMOEBA polarization coincides with the ALMO polarization at 2.1 and 2.9 Å and is too attractive around the equilibrium distance. In the monodentate configuration, the AMOEBA polarization energy agrees with ALMO at the equilibrium and longer distances and becomes less attractive at the short range. By comparing bi- and monodentate binding, it can be found that ALMO polarization is about the same in the two

FIG. 4. Interaction energy and energy decomposition of the $\text{Mg/Ca} \cdots \text{Ac}^-$ dimer as a function of the $\text{Mg/Ca} \cdots \text{C}$ distance. The first and third columns are the bidentate orientation and the second and fourth columns are the monodentate orientation.

configurations; the AMOEBA polarization energies are about the same for the two configurations at the longest distance, but the slope for the bidentate configuration is steeper. This comparison is not quantitative since the distances are varied along different directions, $\text{Ca}-\text{C}$ in the bidentate configuration and $\text{Ca}-\text{O}$ in the monodentate configuration, and the scales of the distances are different. Nevertheless, the different behaviors for the two orientations indicate that it is difficult to accurately describe the relative binding energy between them. The relative binding energy between bi- and monodentate binding is crucial for modeling ion-protein binding, since it can affect the coordination number and therefore specificity for metal ions.

An examination of electrostatics and vdW interactions (see Fig. 4) shows the clear difference between Mg and Ca. For Mg, the AMOEBA electrostatics agrees well with the SAPT electrostatics for both bi- and mono-dentate binding, while

the vdW interaction is too repulsive at short ranges for bidentate binding. The vdW interaction goes up sharply at the short range, which makes the total interaction harder than the QM result. In this case, the vdW interaction in AMOEBA serves as a correction to the over-polarization. On the contrary, for Ca, neither of the AMOEBA electrostatics and vdW components agrees with ALMO. The AMOEBA electrostatics interactions only converge to the ALMO result at the long range and are less attractive at the short range. The AMOEBA vdW interaction is too soft even when compared to vdW + CT energy from SAPT and ALMO. As a result, the vdW component is able to capture both the charge penetration and the charge transfer effects for $\text{Ca} \cdots \text{Ac}^-$.

3. $\text{Mg/Ca} \cdots \text{acetamide}$

The protein backbone carbonyl group is also an important component of the metal ion-binding pocket. Here acetamide

FIG. 5. Interaction energy and energy decomposition of the Mg/Ca...acetamide dimer as a function of the Mg/Ca...O distance.

is chosen as a model compound. The interaction energies and component breakdown of Mg/Ca...acetamide dimers are shown in Fig. 5. The agreement is less satisfactory compared to the cases of Mg/Ca...water and Mg/Ca...Ac⁻ dimers. The Mg...O and Ca...O equilibrium distances are underestimated by 0.2 and 0.1 Å, respectively. The Mg...acetamide equilibrium interaction energy is 1.3 kcal/mol higher than the MP2 result, and the error is 8.4 kcal/mol for Ca...acetamide. The AMOEBA interaction energies are less repulsive at the short range and less attractive at the long range. The discrepancies at the long range are caused by the less favorable polarization and Frz energies.

The energy decomposition for the Ca...acetamide dimer is very similar to the Ca...water complex, with CT contributing to 13%-15% of the total interaction energy at equilibrium distance, and the remaining part almost evenly distributed among Frz and Pol contributions. Therefore, the charge-transfer energy is well captured by the vdW interactions from the original parameter (Fig. S4 of the [supplementary material](#)). However, both the permanent electrostatics and polarization in AMOEBA are less favorable than the ALMO counterparts at all the distances. Starting from the equilibrium distance onward, the errors of electrostatics and polarization are almost constant. The trend also holds true for Mg. Since the vdW interaction is used to compensate for the errors in electrostatics, the

interaction energy does not improve much at the long range and becomes too attractive at the short range, which explains why the equilibrium distance must be shorter in order to match the interaction energy.

The discrepancy between AMOEBA and QM electrostatics at the long range is an indication of problems with the multipoles, not the function form itself. Possible reasons for the problems with the multipoles include the quality of the QM electrostatics potential calculated at a limited basis set and the accuracy of the distributed multipole analysis⁹⁰ for determining the multipoles. Other multipole decomposition methods, such as iterated stockholder atoms,⁹¹ may have better convergence properties, and they will be investigated in the future.

4. Mg/Ca...imidazole

The behavior of the AMOEBA Mg/Ca...imidazole total interactions (Fig. 6) is similar to that of Mg/Ca...acetamide—less repulsive at the short range and less attractive at the long range, although no pairwise vdW parameters are used. The equilibrium distances are underestimated by 0.1-0.2 Å while the errors of the total interaction energy are about 2%. As for the individual components, the AMOEBA electrostatics is more positive than the ALMO values. For Mg...imidazole, the AMOEBA polarization is too positive at the long range

FIG. 6. Interaction energy and energy decomposition of the Mg/Ca...imidazole dimer as a function of the Mg/Ca...N distance.

and the magnitude increases too fast at the equilibrium to short range, which is the main source of error. The Ca polarization has a similar behavior but suffers from much less error. The vdW interactions for Mg and Ca are both softened to make up for the errors of electrostatics near equilibrium, but the interactions are not balanced at shorter and longer distances.

Comparing the polarization energies of Mg/Ca...imidazole and Mg/Ca...water, it can be found that the distance-dependences are almost identical. However, the errors for imidazole are much larger. Since the polarization parameters are mainly determined by the interaction with water, this indicates that when an inaccurate polarization model is used, the errors can be magnified when the parameters are transferred between different systems.

5. Dimers with non-ideal geometries

The interaction energies of dimers with non-ideal geometries are shown in Fig. 7. AMOEBA is accurate for metal ion-water dimers with various orientations. For dimers containing acetate, acetamide, and imidazole, most of the interactions energies in AMOEBA are less favorable than QM interaction energies. The errors are ~ 10 kcal/mol for Ca, ~ 20 kcal/mol for Mg, and can be as large as 50 kcal/mol

in the case of Mg...acetate dimers. The root mean square error (RMSE) for water, acetate, acetamide, and imidazole are 4.2, 23.6, 15.3, and 20.9 kcal/mol, respectively. The discrepancies can be attributed to the missing charge-transfer effects in AMOEBA, which has been discussed above. Specifically, the Mg...acetate interaction in AMOEBA weakens much faster than in QM as the distance becomes longer, which leads to the largest discrepancies in all systems. This is because QM calculations on Mg/Ca...acetate dimers have non-vanishing charge-transfer energy even when the two molecules are far from each other.¹² There are some structures for which the AMOEBA interaction energies are lower than the QM interaction energies. These structures are all located at the short-separation region, and the results are consistent with that of the ideal geometries. Despite these errors, a correlation between AMOEBA and QM interaction energies can be identified, and the dimers containing Ca are more accurately described by AMOEBA than the Mg-ligand dimers. The interaction energies of dimers with distorted ligands are shown in Fig. S5 of the [supplementary material](#). The errors in the repulsive region are exaggerated, while the overall performance is similar to that of the optimized ligands, with RMSEs of 21.0, 10.9, and 20.5 kcal/mol for acetate, acetamide, and imidazole, respectively.

FIG. 7. Comparison of MP2 and AMOEBA interaction energies for dimers with non-ideal geometries. The metal ions were randomly placed around the ligating atoms within a sphere of radius 3.5 Å.

C. Many-body effects in ion-pocket interactions

The metal ion binding pocket generally has multiple negatively charged Asp/Glu residues, and the many-body effects are expected to be strong. The structures of four model compounds for the Mg/Ca pocket are shown in Fig. 8. The total interaction energies and the many-body energies are plotted in Fig. 9. Overall, good agreement is found between AMOEBA and MP2 total interaction energies, although the AMOEBA interaction is too strong when there are multiple acetate

FIG. 8. Structures of model compounds for ion-binding pockets. (a) and (b) are Mg-binding pockets while (c) and (d) are Ca-binding pockets. White, cyan, blue, red, pink, and green spheres are H, C, N, O, Mg, and Ca atoms. All these four structures consist of four acetate molecules and one acetamide, and [(b)–(d)] have an additional water molecule. (a) and (c) each have one bidentate acetate, and all the other acetates bind to the metal ion with one oxygen atom. These four models are denoted as Mg-Bi, Mg-Mono, Ca-Bi, and Ca-Mono, respectively.

molecules. There is an almost linear relationship between AMOEBA and MP2 total interaction energies for Ca, and the agreement for Mg is worse. The total interaction energies have a strong correlation with the net charge of the pocket. However, the pockets with -3 charges have the lowest interaction energy, followed by pockets with -4 and -2 charges, i.e., the addition of the fourth carboxylate group into the binding pocket is not favorable. Therefore, the presence of four Asp/Glu residues in many Ca/Mg binding pockets must be stabilized by the protein backbone.

The result that the pockets with -4 charges on the ligands is less stable than the pockets with -3 charges seems trivial, since the strong repulsion between negatively charged residues makes it unlikely to have -4 charges. However, the many-body effects also play an important role in the relative stabilities of the pockets. As shown in the middle row in Fig. 9, the additive two-body interaction energies (total interaction energy minus many-body energy) are lowest when the pocket has -4 charges. This is contrary to the trend of total interaction energies since the both charge transfer and polarization are unfavorable for highly charged pockets. The difference between the lowest two-body energies for Mg and Ca is 178 kcal/mol, making it very favorable for Mg-binding. The difference between the total interaction energies for the same two complexes is 73 kcal/mol, which is very close to the relative solvation free energies between Mg and Ca. This result emphasizes the importance of many-body interactions in determining the specificity for Mg and Ca.

The many-body energies are explicitly shown in the bottom row of Fig. 9. Mg-pocket binding has more positive many-body energy than Ca-pocket binding, both in terms of the absolute value and the ratio to the total interaction energy. AMOEBA consistently underestimates the many-body energies. The correlation between AMOEBA and QM many-body energy for the Ca-pocket is much better than the correlation for the Mg-pocket. The almost linear relationship between

FIG. 9. Mg/Ca···pocket total interaction energies, two-body interaction energies, and many-body energies. The left column is Mg-pocket interactions and the right column is Ca-pocket interactions. Top, middle, and bottom rows are total interactions, two-body interactions, and many-body interactions, respectively. Systems with different net charges of the ligands (q) or numbers of ligands (n) are marked by different symbols and colors. The total interaction energies and two-body interaction energies correlate with q , while the many-body energies correlate with n .

AMOEBA and QM many-body energies indicates that a modification of the polarization model may capture the many-body effects. The problem with the Mg-pocket can be attributed to the charge transfer effect and the insufficient polarization model. Examination of the Mg-pocket many-body energies reveals that all the complexes with large errors (>30 kcal/mol) contain acetamide, which has already been shown to have incorrect dimer interaction and polarization energies at the long range.

To quantify the contribution of charge-transfer to the many-body effect, the charge-transfer energy between a metal ion and pocket is calculated. The total and additive two-body interactions between a metal ion and pocket, as well as the many-body energies, are shown in Fig. 10. Note that these interactions are all between the pocket and metal ions, and

the interactions between ligands are excluded. We decided to calculate the charge-transfer between metal ions and pockets because (1) currently the many-body EDA has not been implemented in common software and (2) the charge-transfer between ligands is much smaller than the charge transfer between ligands and metal ions. The many-body effect of CT is significant. For Ca, the total CT energies are $\sim 30\%$ of the sum of two-body CT energies and are comparable to the CT energies between Ca and one acetate or acetamide at the same geometry (from -4.9 to -10.8 kcal/mol). The total CT energies for Mg predicted by ALMO are only 3%-5% of the two-body CT energies, which are even smaller than the CT energies between Mg and water at the same geometry. This unphysical behavior may be an artifact of the limited size of the basis set. The total many-body energies and many-body

FIG. 10. Many-body effect of charge-transfer for ion-pocket binding.

energies without CT contribution are compared to the AMOEBA many-body energies. It can be seen that for Ca, the many-body energies from CT (~30 kcal/mol) are more than 10% of the total many-body energies. The ratio of CT to total many-body energies for Mg is expected to be similar or even larger if the chemical trend is preserved, so the many-body CT energies for Mg should be 40 kcal/mol or more. For Ca, the AMOEBA many-body energies are close to the QM many-body energies with the CT contribution deducted. Since the charge-transfer of Ca-complexes is included in the two-body vdW interactions in AMOEBA, the many-body effect of CT is missing. The many-body energies should be well produced with the addition of a correct charge-transfer term. For Mg, CT only makes up ~50% of the errors of AMOEBA many-body energies. This can be attributed to the inadequacy of the current polarization model for Mg and the missing CT term.

IV. CONCLUSION

A new set of parameters for Mg and Ca ions is presented, which is able to reproduce the ion-water dimer interaction energies at the MP2/CBS level as well as the experimental solvation free energies.

Energy decomposition analyses of complexes containing Mg/Ca and water or model compounds for common amino acids show that the electrostatics interactions for Ca are not well described, due to the missing charge-penetration term in AMOEBA, but in most cases, the vdW term is able to capture both the charge penetration and charge-transfer effects. For Mg, the electrostatics in AMOEBA is accurate since the charge penetration effect is less significant. However, the polarization energy is too attractive at the short range and less attractive at ~1.5 times the equilibrium distances, which makes it difficult to match both the equilibrium distance and equilibrium interaction energy using the current functional form of AMOEBA. The charge transfer energies of Mg-complexes should be larger than the charge-transfer energies of Ca-complexes, as opposed to ALMO predictions. The missing charge-transfer effect and the inaccurate polarization model are sources of the poorer performance of AMOEBA for Mg dimer interactions as compared to Ca.

The many-body effect is important for the pocket stability and the specificity for metal ions. Mg-complexes have much more positive many-body energies than Ca-complexes, which significantly reduces the relative binding energies between Mg and Ca. Without the many-body effects, Mg will always be favored by the EF hands. Therefore, an empirical force field must include the many-body effect in order to predict the binding free energies of metal ions. AMOEBA can reasonably reproduce the many-body effects in metal ion-ligand systems; however, it is consistently not positive enough.

Based on ALMO energy decomposition, the charge-transfer effect has a non-negligible contribution to the many-body effect in Ca-complexes. Since AMOEBA absorbs the charge-transfer energy in Ca-complexes into its additive vdW term, the many-body effect of charge-transfer is missing, which accounts for most of the errors of AMOEBA many-body energies. For Mg, it is still difficult to quantify the charge-transfer effect. A consistent EDA method that yields

the correct trend is needed to develop and evaluate charge-transfer models.

SUPPLEMENTARY MATERIAL

See [supplementary material](#) for comparisons of original and new force field parameters, the interaction between acetate and water, effects of different methods for the calculation of charge-transfer energy, and a full list of the optimized force field parameters.

ACKNOWLEDGMENTS

We appreciate support from the Robert A. Welch Foundation (No. F-1691) and the National Institutes of Health (Nos. R01GM106137 and R01GM114237). We thank Ms. S. Y. Cheng for careful reading of the manuscript.

- ¹M. Rouffet and S. M. Cohen, *Dalton Trans.* **40**(14), 3445–3454 (2011).
- ²R. J. White, P. S. Margolis, J. Trias, and Z. Yuan, *Curr. Opin. Pharmacol.* **3**(5), 502–507 (2003).
- ³L. Deng, S. Sundriyal, V. Rubio, Z. Z. Shi, and Y. Song, *J. Med. Chem.* **52**(21), 6539–6542 (2009).
- ⁴R. Levy, V. Sobolev, and M. Edelman, *Human Mutat.* **32**(11), 1309–1318 (2011).
- ⁵L. Breydo and V. N. Uversky, *Metallomics* **3**(11), 1163–1180 (2011).
- ⁶A. Gaeta and R. C. Hider, *Br. J. Pharmacol.* **146**(8), 1041–1059 (2005).
- ⁷T. Dudev and C. Lim, *Chem. Rev.* **114**(1), 538–556 (2014).
- ⁸T. Dudev and C. Lim, *Chem. Rev.* **103**(3), 773–788 (2003).
- ⁹M. J. Stevens and S. L. B. Rempe, *J. Phys. Chem. B* **120**(49), 12519–12530 (2016).
- ¹⁰S. Sikdar, M. Ghosh, M. De Raychaudhury, and J. Chakrabarti, *J. Phys. Chem. B* **119**(46), 14652–14659 (2015).
- ¹¹M. Lepšík and M. J. Field, *J. Phys. Chem. B* **111**(33), 10012–10022 (2007).
- ¹²A. Y. Mehandzhyski, E. Riccardi, T. S. van Erp, H. Koch, P.-O. Åstrand, T. T. Trinh, and B. A. Grimes, *J. Phys. Chem. A* **119**(40), 10195–10203 (2015).
- ¹³E. Ahlstrand, K. Hermansson, and R. Friedman, *J. Phys. Chem. A* **121**(13), 2643–2654 (2017).
- ¹⁴M. Kohagen, M. Lepšík, and P. Jungwirth, *J. Phys. Chem. Lett.* **5**(22), 3964–3969 (2014).
- ¹⁵S. M. Yannone, S. Hartung, A. L. Menon, M. W. Adams, and J. A. Tainer, *Curr. Opin. Biotechnol.* **23**(1), 89–95 (2012).
- ¹⁶J. Zhang, W. Yang, J. P. Piquemal, and P. Ren, *J. Chem. Theory Comput.* **8**, 1314–1324 (2012).
- ¹⁷J. C. Wu, J. P. Piquemal, R. Chaudret, P. Reinhardt, and P. Y. Ren, *J. Chem. Theory Comput.* **6**(7), 2059–2070 (2010).
- ¹⁸W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, *J. Am. Chem. Soc.* **117**(19), 5179–5197 (1995).
- ¹⁹X. Zhu, P. E. M. Lopes, and A. D. MacKerell, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2**(1), 167–185 (2012).
- ²⁰W. L. Jorgensen, D. S. Maxwell, and J. TiradoRives, *J. Am. Chem. Soc.* **118**(45), 11225–11236 (1996).
- ²¹M. M. Reif, P. H. Hünenberger, and C. Oostenbrink, *J. Chem. Theory Comput.* **8**(10), 3705–3723 (2012).
- ²²A. D. MacKerell and N. K. Banavali, *J. Comput. Chem.* **21**(2), 105–120 (2000).
- ²³P. Ren, J. Chun, D. G. Thomas, M. J. Schnieders, M. Marucho, J. Zhang, and N. A. Baker, *Q. Rev. Biophys.* **45**(4), 427–491 (2012).
- ²⁴Y. Shi, P. Ren, M. Schnieders, and J.-P. Piquemal, *Reviews in Computational Chemistry Volume 28* (John Wiley & Sons, Inc., 2015), pp. 51–86.
- ²⁵J. Huang, P. E. M. Lopes, B. Roux, and A. D. MacKerell, *J. Phys. Chem. Lett.* **5**(18), 3144–3150 (2014).
- ²⁶G. A. Cisneros, M. Karttunen, P. Ren, and C. Sagui, *Chem. Rev.* **114**(1), 779–814 (2014).
- ²⁷O. Engkvist, P. O. Åstrand, and G. Karlstrom, *Chem. Rev.* **100**(11), 4087–4108 (2000).
- ²⁸T. A. Halgren and W. Damm, *Curr. Opin. Struct. Biol.* **11**(2), 236–242 (2001).

- ²⁹Y. Shi, Z. Xia, J. J. Zhang, R. Best, C. J. Wu, J. W. Ponder, and P. Y. Ren, *J. Chem. Theory Comput.* **9**(9), 4046–4063 (2013).
- ³⁰H. A. Stern, F. Rittner, B. J. Berne, and R. A. Friesner, *J. Chem. Phys.* **115**(5), 2237–2251 (2001).
- ³¹J. Chen, D. Hundertmark, and T. J. Martínez, *J. Chem. Phys.* **129**(21), 214113 (2008).
- ³²S. Patel and C. L. Brooks, *J. Comput. Chem.* **25**(1), 1–16 (2004).
- ³³S. Patel, A. D. MacKerell, and C. L. Brooks, *J. Comput. Chem.* **25**(12), 1504–1514 (2004).
- ³⁴P. E. M. Lopes, B. Roux, and A. D. MacKerell, *Theor. Chem. Acc.* **124**(1), 11–28 (2009).
- ³⁵P. E. M. Lopes, J. Huang, J. Shim, Y. Luo, H. Li, B. Roux, and A. D. MacKerell, *J. Chem. Theory Comput.* **9**(12), 5430–5449 (2013).
- ³⁶J. A. Lemkul, J. Huang, B. Roux, and A. D. MacKerell, *Chem. Rev.* **116**(9), 4983–5013 (2016).
- ³⁷G. Lamoureux, A. D. MacKerell, and B. Roux, *J. Chem. Phys.* **119**(10), 5185–5197 (2003).
- ³⁸H. Li, V. Ngo, M. C. Da Silva, D. R. Salahub, K. Callahan, B. Roux, and S. Y. Noskov, *J. Phys. Chem. B* **119**(29), 9401–9416 (2015).
- ³⁹N. Gresh, G. A. Cisneros, T. A. Darden, and J.-P. Piquemal, *J. Chem. Theory Comput.* **3**(6), 1960–1986 (2007).
- ⁴⁰P. Cieplak, F.-Y. Dupradeau, Y. Duan, and J. Wang, *J. Phys.: Condens. Matter* **21**(33), 333102 (2009).
- ⁴¹W. Xie and J. Gao, *J. Chem. Theory Comput.* **3**(6), 1890–1900 (2007).
- ⁴²P. Ren and J. W. Ponder, *J. Phys. Chem. B* **107**(24), 5933–5947 (2003).
- ⁴³P. Ren, C. Wu, and J. W. Ponder, *J. Chem. Theory Comput.* **7**(10), 3143–3161 (2011).
- ⁴⁴C. Liu, R. Qi, Q. Wang, J. P. Piquemal, and P. Ren, *J. Chem. Theory Comput.* **13**, 2751 (2017).
- ⁴⁵R. Kumar, F.-F. Wang, G. R. Jenness, and K. D. Jordan, *J. Chem. Phys.* **132**(1), 014309 (2010).
- ⁴⁶R. A. Christie and K. D. Jordan, in *Intermolecular Forces and Clusters II*, edited by D. J. Wales (Springer, Berlin, Heidelberg, 2005), pp. 27–41.
- ⁴⁷D. V. Sakharov and C. Lim, *J. Am. Chem. Soc.* **127**(13), 4921–4929 (2005).
- ⁴⁸M. Soniat, L. Hartman, and S. W. Rick, *J. Chem. Theory Comput.* **11**(4), 1658–1667 (2015).
- ⁴⁹N. Gresh, K. El Hage, D. Perahia, J. P. Piquemal, C. Berthomieu, and D. Berthomieu, *J. Comput. Chem.* **35**(29), 2096–2106 (2014).
- ⁵⁰R. J. Deeth, *Coord. Chem. Rev.* **212**(1), 11–34 (2001).
- ⁵¹C. R. Landis, D. M. Root and T. Cleveland, *Reviews in Computational Chemistry* (John Wiley & Sons, Inc., 2007), pp. 73–148.
- ⁵²T. K. Firman and C. R. Landis, *J. Am. Chem. Soc.* **123**(47), 11728–11742 (2001).
- ⁵³A. E. Anastasi and R. J. Deeth, *J. Chem. Theory Comput.* **5**(9), 2339–2352 (2009).
- ⁵⁴I. Tubert-Brohman, M. Schmid, and M. Meuwly, *J. Chem. Theory Comput.* **5**(3), 530–539 (2009).
- ⁵⁵J.-P. Piquemal, B. Williams-Hubbard, N. Fey, R. J. Deeth, N. Gresh, and C. Giessner-Prettre, *J. Comput. Chem.* **24**(16), 1963–1970 (2003).
- ⁵⁶J. Y. Xiang and J. W. Ponder, *J. Chem. Theory Comput.* **10**(1), 298–311 (2014).
- ⁵⁷D. Suarez, V. M. Rayon, N. Diaz, and H. Valdes, *J. Phys. Chem. A* **115**(41), 11331–11343 (2011).
- ⁵⁸T. Dudev, M. Devereux, M. Meuwly, C. Lim, J.-P. Piquemal, and N. Gresh, *J. Comput. Chem.* **36**(5), 285–302 (2015).
- ⁵⁹J. G. McDaniel and J. R. Schmidt, *Annu. Rev. Phys. Chem.* **67**(1), 467–488 (2016).
- ⁶⁰A. J. Misquitta, *J. Chem. Theory Comput.* **9**(12), 5313–5326 (2013).
- ⁶¹K. U. Lao and J. M. Herbert, *J. Chem. Theory Comput.* **12**(6), 2569–2582 (2016).
- ⁶²D. Jiao, C. King, A. Grossfield, T. A. Darden, and P. Ren, *J. Phys. Chem. B* **110**(37), 18553–18559 (2006).
- ⁶³J. C. Wu, G. Chattree, and P. Ren, *Theor. Chem. Acc.* **131**(3), 1138 (2012).
- ⁶⁴R. Schmid, A. M. Miah, and V. N. Sapunov, *Phys. Chem. Chem. Phys.* **2**(1), 97–102 (2000).
- ⁶⁵C. P. Kelly, C. J. Cramer, and D. G. Truhlar, *J. Phys. Chem. B* **110**(32), 16066–16081 (2006).
- ⁶⁶A. Grossfield, P. Ren, and J. W. Ponder, *J. Am. Chem. Soc.* **125**(50), 15671–15682 (2003).
- ⁶⁷G. Lamoureux and B. Roux, *J. Phys. Chem. B* **110**(7), 3308–3322 (2006).
- ⁶⁸I. S. Joung and T. E. Cheatham, *J. Phys. Chem. B* **112**(30), 9020–9041 (2008).
- ⁶⁹H. Zhang, Y. Jiang, H. Yan, C. Yin, T. Tan, and D. van der Spoel, *J. Phys. Chem. Lett.* **8**(12), 2705–2712 (2017).
- ⁷⁰Y. Marcus, *J. Chem. Soc., Faraday Trans.* **87**(18), 2995–2999 (1991).
- ⁷¹M. D. Tissandier, K. A. Cowen, W. Y. Feng, E. Gundlach, M. H. Cohen, A. D. Earhart, J. V. Coe, and T. R. Tuttle, *J. Phys. Chem. A* **102**(40), 7787–7794 (1998).
- ⁷²T. Darden, D. Pearlman, and L. G. Pedersen, *J. Chem. Phys.* **109**(24), 10921–10935 (1998).
- ⁷³G. L. Warren and S. Patel, *J. Chem. Phys.* **127**(6), 064509 (2007).
- ⁷⁴A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen, and A. K. Wilson, *Chem. Phys. Lett.* **286**(3–4), 243–252 (1998).
- ⁷⁵M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, GAUSSIAN 09, Revision A.02, Gaussian, Inc., Wallingford, CT, 2009.
- ⁷⁶J. M. Turney, A. C. Simmonett, R. M. Parrish, E. G. Hohenstein, F. A. Evangelista, J. T. Fermann, B. J. Mintz, L. A. Burns, J. J. Wilke, M. L. Abrams, N. J. Russ, M. L. Leininger, C. L. Janssen, E. T. Seidl, W. D. Allen, H. F. Schaefer, R. A. King, E. F. Valeev, C. D. Sherrill, and T. D. Crawford, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2**(4), 556–565 (2012).
- ⁷⁷Y. Shao, Z. Gan, E. Epifanovsky, A. T. B. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Kuš, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M. Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. Woodcock III, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. O. Beran, Y. A. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Didenhofen, R. A. DiStasio, Jr., H. Do, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. D. Hanson-Heine, P. H. P. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyayev, J. Kim, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. D. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S.-P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, E. Neuscamman, C. M. Oana, R. Olivares-Amaya, D. P. O'Neill, J. A. Parkhill, T. M. Perrine, R. Peverati, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, S. M. Sharada, S. Sharma, D. W. Small, A. Sodt, T. Stein, D. Stück, Y.-C. Su, A. J. W. Thom, T. Tschimochi, V. Vanovschi, L. Vogt, O. Vydrov, T. Wang, M. A. Watson, J. Wenzel, A. White, C. F. Williams, J. Yang, S. Yeganeh, S. R. Yost, Z.-Q. You, I. Y. Zhang, X. Zhang, Y. Zhao, B. R. Brooks, G. K. L. Chan, D. M. Chipman, C. J. Cramer, W. A. Goddard III, M. S. Gordon, W. J. Hehre, A. Klamt, H. F. Schaefer III, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J.-D. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C.-P. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. V. Voorhis, J. M. Herbert, A. I. Krylov, P. M. W. Gill, and M. Head-Gordon, *Mol. Phys.* **113**(2), 184–215 (2015).
- ⁷⁸P. Eastman, M. S. Friedrichs, J. D. Chodera, R. J. Radmer, C. M. Bruns, J. P. Ku, K. A. Beauchamp, T. J. Lane, L.-P. Wang, D. Shukla, T. Tye, M. Houston, T. Stich, C. Klein, M. R. Shirts, and V. S. Pande, *J. Chem. Theory Comput.* **9**(1), 461–469 (2013).
- ⁷⁹M. Harger, D. Li, Z. Wang, K. Dalby, L. Lagardère, J.-P. Piquemal, J. Ponder, and P. Ren, *J. Comput. Chem.* **38**(23), 2047–2055 (2017).
- ⁸⁰J. R. Abella, S. Y. Cheng, Q. Wang, W. Yang, and P. Ren, *J. Chem. Theory Comput.* **10**(7), 2792–2801 (2014).
- ⁸¹D. R. Bell, R. Qi, Z. Jing, J. Y. Xiang, C. Mejias, M. J. Schnieders, J. W. Ponder, and P. Ren, *Phys. Chem. Chem. Phys.* **18**(44), 30261–30269 (2016).
- ⁸²J. R. Pliego, Jr. and J. M. Riveros, *Phys. Chem. Chem. Phys.* **4**(9), 1622–1627 (2002).

- ⁸³J. R. Pliego and J. M. Riveros, *Chem. Phys. Lett.* **332**(5), 597–602 (2000).
- ⁸⁴L. Vlcek, A. A. Chialvo, and J. M. Simonson, *J. Phys. Chem. A* **117**(44), 11328–11338 (2013).
- ⁸⁵L. X. Dang and T.-M. Chang, *J. Phys. Chem. B* **106**(2), 235–238 (2002).
- ⁸⁶T. L. Beck, *Chem. Phys. Lett.* **561–562**, 1–13 (2013).
- ⁸⁷Y. Mao, O. Demerdash, M. Head-Gordon, and T. Head-Gordon, *J. Chem. Theory Comput.* **12**(11), 5422–5437 (2016).
- ⁸⁸A. K. Katz, J. P. Glusker, S. A. Beebe, and C. W. Bock, *J. Am. Chem. Soc.* **118**(24), 5752–5763 (1996).
- ⁸⁹Q. Wang, J. A. Rackers, C. He, R. Qi, C. Narth, L. Lagardere, N. Gresh, J. W. Ponder, J.-P. Piquemal, and P. Ren, *J. Chem. Theory Comput.* **11**(6), 2609–2618 (2015).
- ⁹⁰A. J. Stone and M. Alderton, *Mol. Phys.* **56**(5), 1047–1064 (1985).
- ⁹¹A. J. Misquitta, A. J. Stone, and F. Fazeli, *J. Chem. Theory Comput.* **10**(12), 5405–5418 (2014).