Approaching *ab initio* Complete Basis Set Limit and Experimental Accuracy for Ions in the AMOEBA Polarizable Force Field

Chuanjie Wu[1], Pengyu Ren[2] and Jay W. Ponder[1],*

[1] *Department of Biochemistry and Molecular Biophysics,*
Washington University School of Medicine, St. Louis, MO 63110
[2] *Department of Biomedical Engineering,*
University of Texas, Austin, TX 78712-1062

*Corresponding Author

Abstract

Highly charged systems are generally difficult to model with fixed-charge force fields, especially when simulating trans-phase properties for these charged systems without considering polarization. AMOEBA force field for some typical cations and anions, including Na^+, K^+, Mg^{2+}, Ca^{2+}, Cl^-, Br^-, I^-, $HCOO^-$, CH_2COO^-, $CH_3CH_2COO^-$, NH_4^+, $[(CH_3)_xNH_{4-x}]^+$ (x=1,4), which are parameterized based on gas phase properties have been proved to yield high accurate hydration free energy. The gas phase ion-water dimer interaction energies have been compared to QM results from full-MP2, and regular frozen-core MP2 methods combined with basis sets from aug-cc-pVDZ to complete basis set limit (aug-cc-pV(T->Q)Z, or aug-cc-pv(Q-5)Z), and also corresponding available co-valence basis sets (aug-cc-pCVDZ to aug-cc-pCVQZ). QM results show that co-valence basis set with full-MP2 is required for metal ion related QM calculations. AMOEBA could re-produce MP2/CBS gas phase ion-water dimer interaction energies, and at the same time the simulated hydration free energies for these ions agree very well with experimental data. This further tells the good capability of AMOEBA force field for simulating trans-phase properties. This is another good case to show the necessity of introducing polarization in force fields.

1. Introduction

Accurately predicting the trans-phase properties is not only important for phase equilibrium related chemical engineering processes, but also fundamental for complex inhomogeneous biomolecular system simulations. Similar to other molecular simulation problems, good sampling strategy and reliable force fields are key to achieve this goal. Force field is determinant for trans-phase related properties. Without a reliable force field, no matter how good is the sampling method, there is no way to reach the desired accuracy. In fixed-charge force field, parameters have to be adapted for condensed phase property calculations by pre-including certain polarization effects in intermolecular interaction terms (1). When this force field is applied to another phase or trans-phase properties, the static pre-incorporated polarization will deviate physical properties by a certain systematic error. Mobley *et al*(2-5) found the hydration free energy of small organic molecules explicitly predicted with fixed-charge force fields are systematically too positive in comparison with the experimental results. Swope *et al (6, 7)* has pointed the systematic role polarization cost plays in comparing experimental heats of vaporization with ones computed using fixed charge force fields.

The transferability between different phases is one of biggest advantages of polarizable force fields. By introducing polarization, polarizable force fields are able to reflect the intermolecular electrostatic interaction difference when the chemical environment changes in different phases. To better represent the atomic polarization in force fields, there have been several efforts like fluctuating charges(8), adding off-site charges points with Drude-oscillator method(9, 10), calculating induced dipole in addition to monopole, monopole+dipole or monopole+dipole+quadrupole interactions(11-16).

These polarizable force fields form a new generation next to the classic force fields, like polarizable AMBER force field(17), Frisner-Berne Fluctuating Charging CHARMM force field(8), Drude Osicllator CHARMM force field(9, 10, 18-20), Polarizable Force Field (PFF)(21), SIBFA(22, 23), AMOEBA polarizable force field(13, 14, 16, 24). AMOEBA is the one using multipole through quadrupole to represent the "permanent" part of the electrostatic interaction and induced dipole to include the polarization effects. In AMOEBA, polarization is damped for avoiding polarization catastrophe. AMOEBA also applies buffered 14-7 van der Waals (vdW) potential(25) which is "softer" in the short range than Lenaard-Jones 12-6 vdW potential generally used in most classical force fields.

Hydration free energy is not only interested by many chemical and drug researchers, but also turns out to be a good physical properties for testing the quality of force fields. In absolute hydration free energy, one has to design a thermodynamic cycle cross gas phase and condensed phase, which is also good for testing the transferability of force field between phases. There have been many systematic prediction of hydration free energy of small molecules with computation chemistry methods. As mentioned above, fixed-charge model intends to give systematic too positive results, especially for polar systems. We predicted a couple of hydration free energies of some typical small molecules varying from alkane, alcohol, sulfide, amines, aromatic structures to amides with AMOEBA force field. The standard deviation reaches about 0.6 kcal/mol in comparison to the experimental values. At the same time, AMOEBA gives good hydrogen bonding structures and energies for gas phase dimers (homodimer and dimer with water). They can also produce high quality heat of vaporization and liquid density for pure liquid of those testing molecules. The parameters were not parameterized using hydration free energy experimental values. This tells the great

potential of using polarizable force field for predicting trans-phase properties. Table 1 shows the hydration free energy partition in electrostatics and vdW contributions calculated with AMOEBA. The vdW contribution to the free energy is about 0.9~3 kcal/mol in almost all the molecule. The more polar the molecule is, the more contribution to the free energy from electrostatics. The electrostatics contribution has a good correlation with the molecular dipole and quadrupoles. It is also directly related to the number of polar atoms (N, O, S) or charged sites (ionic sites). The aromatic structures like benzene has more significant out-of-plane quadrupoles than monopole and dipoles, which tells us the importance of using quadrupole to get proper spatial electrostatic distributions. Jiao *et al (26)* showed the contribution of electronic polarization to binding appears to be crucial in their prediction of ligand-protein binding free energy with AMOEBA force field.

Since the electrostatic part for free energy calculation is much more significant in charged molecules than that for neutral ones, a straightforward idea is to test the performance of electrostatic interaction term of a potential or solvation model by calculating the free energy of ions. Grossfield *et al (27)* calculated the solvation free energy of potassium, sodium and chloride ions with AMOEBA force field. The relative free energy between the ions and the salt free energy of NaCl and KCl agrees very well with the experimental values. They also mentioned that the reported solvation free energy of ions could be very different, depending on the reference experimental states. But generally the relative solvation free energies between ions and the free energies of the salts with both positive and negative charges are consistent between the results from different experimental references. This idea was widely recognized and also referred in the later ion solvation free energy calculations. Jiao *et al(28)* used a similar protocol for calculating the solvation free energies of calcium

and magnesium dications once they developed the parameters for these two ions. Kelly et al(29, 30) used the density functional solvation model SM6 to accurately predict absolute aqueous solvation free energies of typical cations and anions. They also presented a compilation of absolute aqueous solvation free energies for 121 unclustered ions and 147 conventional and absolute aqueous solvation free energies of 51 clustered ions containing from 1 to 6 water molecules, using a value of -265.9 kcal/mol for the absolute aqueous solvation free energy of the proton. The data from this compilation will be used as a main source of reference values. In this compilation, the 'conventional aqueous solvation free energyies' relative to the free energy of proton (-265.9 kcal/mol) are reported and considered to be consistent. Yu *et al (20)* introduced and parameterized polarizable models based on classical Drude ossiclator in conjunction with the SWM4- NDP water model for monatomic ions including Li+, Na+, K+, Rb+, Cs+, Mg2+, Ca2+, Sr2+, Ba2+, Zn2+, F-, Cl-, Br-, and I-. They found that the various gas-phase and condensed-phase properties are in reasonable agreement with the available experimental data and their provided *ab initio* calculations. They also noted that the models for divalent cations from their work are somewhat less accurate than those of the monovalent ions. The results from *ab initio* quantum chemistry calculations and the hydration free energy are not perfectly reproduced. They think that could be in part due to the neglect of charge-transfer effects and to the breakdown of linear response.

High quality metal ion–water interaction structure and energies are very improtatn reference for validating the transphase transferability of polarizable force fields and even some density functional method. It is worth to point that there is one important fact, which is generally not noticed by molecular simulation researches but has already been discovered

by quantum mechanics researchers. In metal ion-water dimer *ab initio* calculation, we need to use core-valence basis set and have all the electrons involved in X-order (x=2,4) wave function perturbation. Without considering the electrons in all the layers, it generally underestimates the binding energy between ion and water by 0.5-2 kcal/mol or 3~9%, which is big for *ab initio* calculations. Complete basis set (CBS) limit extension has enabled minimization of basis set superposition error (BSSE) in quantum mechanics calculations and theoretically it should eliminate BSSE when the CBS is ideally accurate. So for cation-water calculation, using core-valence basis set and CBS calculation with MP2 or CCSD(T) methods will give an extreme accurate *ab initio* reference value for force field based calculations. This has not caused enough attention in many molecular mechanics calculations, and will be applied in this paper.

In this paper, we will show the parameterization results for a serious of ions with AMOEBA potential model. The gas phase results will be calibrated with the CBS *ab initio* calculation results, and core-valence basis set with full second order Møller–Plesset perturbation applied for alkali metal ion related calculations. The abilities for predicting condensed phase properties of these parameters are verified by comparing the aqueous solvation free energy calculations to experimental collections.

2. Theory and Methods

2.1 AMOEBA Potential and Parameterization

AMOEBA force field has similar valence terms to those of most classical force fields, like harmonic bond and angle functional, Fourier dihedral series (from 1 to 6), out-of-plane bending term, stretch-bend cross term. PI-torsion term is used to keep the planar structure across a Pi-bond. With this term, the torsions across the Pi-bond will be calculated just once

instead of summing up four individual torsions during the energy calculation. Torsion-torsion corrections were applied through Spline fitting to the energy differences at the potential energy surfaces calculated with QM and AMOEBA before torsion-torsion correction. In AMOEBA force field, they are currently used for reproducing the Ramachandran map of peptides, and the potential energy surface of sugar rings in nucleic acids.

AMOEBA uses multipoles through quadrupole to represent the 'permanent' electrostatic potential and introduces polarization by self-consistently calculating the induce dipole. Van der Waals (VDW) interactions are calculated with Hagren's Buffered 14-7 potential (25), which is 'softer' than Lennard-Jones 12-6 functional in the short range.

Parameterization of AMOEBA generally starts from *ab initio* calculation, which will provide electron wave function coefficients for distributed multipole analysis with Stone's GDMA program (31, 32). Since the monoatomic ions have spherically symmetric electrostatic field, they only have monopoles (charges), polarizabilities and polarization dampling factor as their nonzero polarizability parameters. The dipole and quadrupole moments are all zeroed out. Atomic polarizabilities are fitted to molecular experimental polarizabilities. Monoatomic ion polarizabilities are from *ab initio* calculation and the damping factor was adjusted with ion-water dimmer interaction calculation results. Jiao *et al (28)* has reported the method details for obtaining the polarizability and dampling factors of Mg^{2+} and Ca^{2+}. After getting the distributed multipoles at the global coordinate frame system, we run TINKER POLEDIT (33) to rotate them to local frame systems defined by the center atom and the adjacent 2 to 3 reference atoms for molecules containing more than 3 atoms. Currently TINKER has fully implemented three types of local coordinate frame systems, which are named Z-then-X, Bisector, and Z-Bisector. In Z-then-X coordinate system,

Z and X-axis directions follow the vectors from center atom to the first and second reference atoms. Y-axis is defined with Right-hand rule. The Bisector coordinate system uses the bisector of the angle formed by the two vectors from center atom to the two reference atoms, and X-axis is point to either side of the reference atom. Z-Bisector is mainly useful for pucker structure like the –NH3 group in methylamine. It uses one center atom and three reference atoms. The Z-axis is following the vector from the center atom to a reference atom like N-C vector, and the X-axis is on the side of bisector of the other two reference atoms. As mentioned above, in the 'permanent' electrostatic parameters, monoatomic ions/molecules only have nonzero monopole. Diatomic particles have monopole and nonzero Z-components in dipole (d_z) and quadrupole (q_{xx}, q_{yy}, q_{zz}; $q_{xx} = q_{yy} = -q_{zz}/2$). "Z-then-X" and "Z-bisector" coordinate frames are symmetric by ZX plane, so the Y-components (d_y, q_{xy}, q_{yz}) need to be zeroed out. "Bisector" has the bisector vector as an additional symmetry operator besides ZX plane, so the multipoles for the center atom only have nonzero Z-components. When the "Z-then-X" coordinate system is applied to C3 system like –CH3 group, only Z-components are nonzero too for keeping the symmetric properties. Besides rotating the distributed multipoles from global frame system to local coordinate frames, POLEDIT will also removing inter-polarzation group polarization with applying Thole's model(34).

The electrostatic parameters are further optimized by fitting to the *ab initio* potential surface around the molecule with TINKER POETNIAL program. The potentials on the grids around a molecule are first calculated with *ab initio* method like MP2 with aug-cc-pVTZ and MP2 density in SCF calculation. Then the potentials are read into POTENTIAL program to be compared with the potentials calculated with AMOEBA. By numerically evaluating the gradients of RMS to multipole parameters, the RMS values are minimized with an optimally

conditioned variable metric nonlinear optimization routine without line searches. Grids are uniformly distributed with Saff's scheme (35) on 4 shells at a spacing of 0.35 angstrom and the first shell is 1.0 angstrom from atomic vdW surface. The AMOEBA potential are determined with a probe of point charge q (+1) following Equation 1.

$$ESP_{AMOEBA} = q \cdot T \cdot (M + \mu) \tag{1}$$

where q is the point charge; M is the 13-dimensional multipole vector (Equ. 2) and T is the 13×13 dimensional tensor matrix about atom separation r_{ij} (Equ. 3). Indexes i and j represent different atoms. The superscript t denotes the transpose. μ is the induce dipole vector [μ_{ix}, μ_{iy}, μ_{iz}].

$$M_i = \left[\; q_i, d_{ix}, d_{iy}, d_{iz}, Q_{ixx}, Q_{ixy}, Q_{ixz}, Q_{iyx}, Q_{iyy}, Q_{iyz}, Q_{izx}, Q_{izy}, Q_{izz} \; \right]^t_{(13\times1)} \tag{2}$$

$$T_{ij} = \begin{bmatrix} 1 & \dfrac{\partial}{\partial x_j} & \dfrac{\partial}{\partial x_j} & \dfrac{\partial}{\partial x_j} & \cdots \\[2mm] \dfrac{\partial}{\partial x_i} & \dfrac{\partial^2}{\partial x_i \partial x_j} & \dfrac{\partial^2}{\partial x_i \partial y_j} & \dfrac{\partial^2}{\partial x_i \partial z_j} & \cdots \\[2mm] \dfrac{\partial}{\partial y_i} & \dfrac{\partial^2}{\partial y_i \partial x_j} & \dfrac{\partial^2}{\partial y_i \partial y_j} & \dfrac{\partial^2}{\partial y_i \partial z_j} & \cdots \\[2mm] \dfrac{\partial}{\partial z_i} & \dfrac{\partial^2}{\partial z_i \partial x_j} & \dfrac{\partial^2}{\partial z_i \partial y_j} & \dfrac{\partial^2}{\partial z_i \partial z_j} & \cdots \\[2mm] \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}_{(13\times13)} \cdot \begin{pmatrix} \dfrac{1}{r_{ij}} \end{pmatrix} \tag{3}$$

Similar to classical force fields like OPLS, we fit the vdW parameters to experimental heat of vaporization and liquid density/liquid pressure for neutral molecules. As a polarizable force field, another advantage is that we can also reproduce the *ab initio* gas phase dimer interaction structural and energy results. So for ionic pieces, we derive the vdW parameters by fitting to ion-water dimer QM calculation results.

Valence parameters are fitted to the QM atomic forces, hessians, and frequencies with TINKER VALENCE program. Structural RMS is also included in the fitting procedure. Torsions are fitted to the potential surface energy differences between *ab initio* and AMOEBA results for multiple molecular conformations generated by rotating the specific torsion. Potential energies of the conformations are weighed with the slope of the potential energy plot at each conformer.

2.2 Basis Sets Including the Core-Valence Basis Set for Metal Ions

Core valence basis sets have been recognized as important for calculating metal ion related *ab initio* calculations. We got the core valence basis sets for Na+, K+, Mg2+ and Ca2+ from EMSL basis set exchange(36, 37). In the ion-dimer calculations, we applied aug-cc-pVXZ (X=D,T,Q) basis set for water atoms, and core-valence basis sets for ions. Full MP2 method (all the electrons are considered as valence electrons) was used for optimizing the dimer structures. Then we ran single point with MP2 electron density and BSSE correction to get the interaction energy between ion and water. Due to the availability of core valence basis sets from EMSL, we tried cc-pCVXZ(X=D,T,Q,5) and aug-cc-pCVXZ (X=D,T,Q) basis sets for sodium ion, CVXZ(X=D,T,Q) for potassium and magnesium ions, and tried cc-pCVXZ(X=D,T,Q,5) for calcium ion.

For the other ions/molecules involved in the paper, we used aug-cc-pVXZ basis sets for dimer optimization and interaction energy evaluations. For verification purposes, we also did MP2 calculation with core valence basis set aug-cc-pVXZ(X=D,T,Q,5) for chloride anion. FC-MP2 (frozen core MP2) and full MP2 were both done to Cl-, Br-, and I-. In the iodine related calculation, we also used the effective core potential (ECP) or pseudo-potential basis set. The dimer calculations for alkane ammonium ions and acetic acids are standard

MP2/aug-cc-pVXZ (X=T,Q) calculations. Ion-water interaction energies were also done with single point with turning on MP2 electron density after full dimer structural optimization.

2.3 Complete Basis Set Limit

Ideally extrapolating the *ab initio* energies to CBS limit results will 'eliminate' the basis set superposition error at the selected *ab initio* method level and is much more reliable as a reference value for polarizable force field parameterization. In this paper, we used MP2 to do all the calculations and extrapolated Dunning basis set series to CBS limit. A simple but rigorously correct extropolation scheme for dimer structural and intermolecular interaction energy calculations was used and shown as Equation 4.

$$Y_{CBS} = \frac{Y_N N^3 - Y_{N-1}(N-1)^3}{N^3 - (N-1)^3} \tag{4}$$

where Y is the intermolecular distance D or the interaction energy E. N is T, Q or 5. Here we mostly used T->Q extrapolation, and there is also Q->5 extrapolation for few of them (Na$^+$, Cl$^-$, Mg^{2+}, Ca^{2+}).

2.4 FEP/BAR Free Energy Calculations

Hydration free energy (ΔG_{hyd}) calculations were performed with FEP/BAR method in TINKER. The thermodynamics cycle for calculating ΔG_{hyd} was done in three steps: discharging the solute in vacuum ($\Delta G_{discharging}$) in 7 steps, growing the vdW of the solute in water ($\Delta G_{coupling}$ or ΔG_{vdW}) in 16 steps, and recharging the solute in water in 7 steps ($\Delta G_{recharging}$). Then $\Delta G_{hyd} = \Delta G_{discharging} + \Delta G_{coupling} + \Delta G_{recharging}$. The multipoles and polarizability parameters were perturbed in the vacuum and aqueous charging components with a stepped procedure of (0.000, 0.200, 0.447, 0.632, 0.774, 0.874, 1.000). The

electrostatic contribution to hydration free energy $\Delta G_{elec}=\Delta G_{discharging} + \Delta G_{recharging}$. The vdW

of the solute was grown gradually in the following steps (0.000, 0.067, 0.133, 0.200, 0.267,

0.333, 0.400, 0.467, 0.533, 0.600, 0.667, 0.733, 0.800, 0.867, 0.933, 1.000). Solute was

solvated in a water box with 850 molecules at the density of 1.0000 g/ml. NVT simulations

at 298K with Berendsen thermostat were done for all the perturbation steps. Systems were

equilibrated for 50 ps before the next 1 ns of collecting effective snapshots. 2000 snapshots

were picked from the whole production trajectory at a time interval of 0.5 ps. The simulation

time step is 0.1 fs. Then we used BAR method (Equ. 5) to evaluate the free energy results.

$$\Delta A(j)_{\lambda_i \rightarrow \lambda_{i+1}} = ln\frac{< 1/[1 + exp((U_{\lambda_i} - U_{\lambda_{i+1}} + C)/RT)] >_{\lambda_{i+1}}}{< 1/[1 + exp((U_{\lambda_{i+1}} - U_{\lambda_i} - C)/RT)] >_{\lambda_i}} + C \qquad (5)$$

$$C = \Delta A(j-1)_{\lambda_i \rightarrow \lambda_{i+1}}$$

where ΔA is the free energy between two perturbation windows λ_i to λ_{i+1}. U is the potential

energy of the specific snapshot. <> is the ensemble average and C is the constant. The

calculation needs to be done iteratively until C converges to the free energy results. Here j is

the iteration step. A soft-core buffered 14-7 potential (Equ. 6) was applied to the vdW

interaction between solute and solvent. The vdW was tapered from 0.9rc (rc is the vdW

cutoff) to rc to keep a smooth gradient with a polynomial functional.

$$U_{ij} = \lambda^5 \varepsilon_{ij} \frac{1.07^7}{\left[0.7(1-\lambda)^2 + \left(\rho_{ij} + 0.07\right)^7\right]} \left(\frac{1.12}{0.7(1-\lambda)^2 + \rho_{ij}^7 + 0.12} - 2\right) \qquad (6)$$

The long range correction (LRC) to this soft-core vdW potential was numerically

calculated with considering the tapering, see Equation 7.

$$U_{LRC} = \frac{1}{2} \sum_i \sum_j \left\{ (N_i \times N_j) - (1 - \lambda^5) \left[L_i \times N_j + (N_i - L_i) \times L_j \right] \right\} \frac{4\pi}{V} \int\limits_{r_c}^{\infty} U_{ij}(r) r^2 dr \qquad (7)$$

where i and j are atom class indexes. N is the number of a certain atom classes. L is the number of a certain ligand atom classes. U is the vdW potential energy between the pairs.

3. Results and Discussion

3.1 CBS and Gas Phase Dimer Calculation Results

We listed the QM and AMOEBA calculation results for the alkali metal ion-water and halogen ion-water dimers in Table 2. As shown in the Table, we have tried several different QM calculation combinations for Na$^+$-H2O dimers structure optimization and intermolecular interaction energy analysis. It is clear that without using core-valence basis set in full MP2 method, the geometry (O-Na+ distance) error is about 0.06 angstrom, while the interaction energy error is about 1.5~2 kcal/mol or 6%~8%, which is a 'huge' error for this interaction. The dimer intermolecular distance and interaction energy from regular optimization with MP2-Full/aug-cc-pVQZ results are completely wrong. Optimizations on counterpoise surface for MP2-full methods improves the results a little but still cannot fix the problem. These errors do not disappear when the basis set sizes are increased for the traditional frozen-core (FC-MP2) method, which tells that it is key to apply core-valence functional in the basis set and apply full MP2 to do the *ab initio* calculations for the Na+-water dimers and similar metal cation-water dimers.

Just for a testing we also did *ab initio* calculations with regular Cunning's basis set with full MP2, the results are clearly not good, as we could expect. The general problem is that the results are not converging very well. The basis set superposition errors are enlarged in comparison with the results from the combination of FC-MP2 and regular Cunning's basis

sets without core valence functionals. The application of MP2-Full/aVQZ even produced an obvious broken result. This kind of combination tests was also tested on the calculations of bromide anion-water and iodine anion-water dimers. The bigger BSSE can also be observed from the testing results. Please note that in the iondine calculation, we have also used pseudopotential basis sets and ECP method.

We also did calculations with relatively big Pople basis set 6-311++G(3df,3pd) in MP2 methods. These calculations are generally faster than the ones with Cunning basis sets like aug-cc-pVTZ, but the structure and energy results are quite comparable to the results from MP2/aug-cc-pVXZ (X=D,T,Q,5).

The core-valence basis set combined with full MP2 calculation was applied to Na^+-water, K^+-Water, Mg^{2+}-water, and Ca^{2+}-water dimers. In dication calculations, the core-valence calculation results improve the interaction energy by up to 3~4 kcal/mol. The geometry change (about 0.03 angstrom shorter) is only about half of that in monocations, but energy is much more significant, which makes sense if considering the charge difference.

We tested both regular basis sets and the basis sets with core valence correlation effects for Cl^- and other anions. The results do not show a strong favorite for core valence basis set. The distance change is only about 0.005 angstrom and the energy change is about 0.1 kcal/mol. So in the calculations for Br^--water and I^--water dimers calculations, we did not applied core valence basis sets. In iodine related calculations, there may be more errors from ECP method for such big atoms with big atom orders.

To make sure the parameterization basis is correct or closer to "ultimate correct results" is cricital for a polarizable force field. One of the reason is that it will give a very good testify of the force field performance; and the other second is that the single ions have simply

and pure parameter sets, which are useful to proved reference values for polarizable force field parameterization and validation. In the condensed phase solvation free energy prediction part of this paper, we will show how we use the free energy results from a single ion as reference value to test the parameters for other ions and organic systems.

Base on the single ion-water dimer calculations, we carefully parameterization AMOEBA force field for these ions based on the CBS results. The geometry and intermolecular energies are also shown in Table 2. It is clear that AMOEBA has approached the CBS results at very high level.

Table 2 shows the single ion (inorganic)-water dimers results calculated at difference *ab initio* calculation levels, while Table 3 is for the CBS extrapolations of several typical organic ion-water dimers. The calculations for organic ions are standard calculations. We did MP2/aug-cc-pVTZ and MP2/aug-cc-pVQZ calculations, and the extrapolated the results to CBS(T->Q) with Equ. (4). In Table 3, we chose amine ions to represent the organic cations and carboxylic acid ions to represent the organic anions. The positive and negative charges on these two kinds of molecules polarize water in different ways and strength. The cation acts as a strong hydrogen bond donor in the interactions with water, while the anion is a strong acceptor to the water hydrogen. These two extreme cases are also good to give more tests on the polarizable force field model. Even more we chose the amine ions with 0~4 methyl groups. The test of methylation effects will tell better about the quality about the force field. The tetramethyl ammonium ion does not have any explicit N-H hydrogen bond in the dimers, so we did not show the dimer results in Table 3, but we listed it in the Table because the hydration free energy for this molecule will also be calculated with transferring parameters from other amine ions.

We again parameterized the force field based on the CBS results, and get quite well results both for geometries and intermolecular energies. Since for organic molecules, there are more error sources from conformations and multi-atom environment, the results are a little less perfect than the results for single ion-water dimers. But in comparison with a CBS limit and empirical force field, they can still be considered very good. The distance and energy errors are generally below 3%.

3.2 Hydration Free Energy

In Table 4, we listed the hydration free energies calculated with AMOEBA and the comparison of relative hydration free energies with experimental values. The relative hydration free energies for cations use potassium as the reference value, while chloride anion is used for calculating the relative free energies for the other anions. The agreement between AMOEBA predicted values and experimental values are great. The signed average error is -0.213 kcal/mol and the unsigned error is 0.394kcal/mol for the ions in the Table. Especially AMOEBA is able to give very good prediction to the dication ions Ca^{2+} and Mg^{2+}, which is a big problem with fixed charge potential models. The prediction of iodine hydration free energy is also encouraging in consideration of the chemical properties of iodine ions and difficulty in describing it with quantum mechanical methods. Prediction of organic ions is more challenging since there are more force field variables affecting the hydration free energy. It requires very high quality parameters to properly describe the partition of enthalpy and entropy in its hydration free energy, also a good and efficient sampling is also important for predicting the hydration free energies of flexible molecules. AMOEBA errors for the three carboxylic acids are all less than 0.5 kcal/mol.

We run another calculation on a series of ammonium ions to see if the hydration free energy predicted with AMOEBA force field will reflect the methylation effects. Results are presented in Table 5. Clearly the predicted methylation effect agrees very well with the experimental values. All the prediction errors are below 4 kcal/mol. It is interesting that the methylation effects are 9, 8, 7 kcal/mol when increasing the number of methyl groups in ammoniums from 0 to 1, 2, 3. We have not find the experimental value for the tetramethylammonium ion, but AMOEBA has predicted the methylation effect is 6 from trimethylammonium to tetramethylammonium, which is following the existing trend.

When predicting hydration free energies for ions, either with experimental or theoretical methods, although the absolute free energies may differ a lot from method to method, but relative hydration free energies are usually consistent between the methods. The results in Table 5 show that the AMOEBA is able to reproduce the methylation effects on ammonium ions by accurately predicting the relative hydration free energy of these ions.

Another consistence in the hydration free energies for ions exists in the hydration free energies of the neutral salts composed with counter ions. We also calculated the hydration free energy of some salts and results are in Table 6. The salt free energies in Table 6 for AMOEBA were calculated by summing up the independent solvation free energies of ions in the salt. The salts in Table 6 are 4 chloride salts: sodium chloride, potassium chloride, ammonium chloride, and methylammonium chloride. The prediction errors are 0.01~0.09 kcal/mol (NaCl), 0.18~0.58 kcal/mol (KCl), 0.25 kcal/mol (NH4Cl) and 0.86 kcal/mol (CH3NH3Cl) respectively in comparison to the available experimental values. The percentage errors are all below 1%. The free energies for the salts here have been converted to use reference gas phase state 1M/L instead of 1atm. The difference is dG(1M/L, salt) =

Σ[dG(1atm, ion)-1.9] kcal/mol. So for the salts in Table 6, each salt has two ions, the difference between dG(1M/L) and dG(1atm) is -3.8 kcal/mol. The standard state 1M/L is also reference state for all the other hydration free energy calculation results in this article.

4. Conclusion

By introducing polarization and mulitpole interaction into electrostatic interaction, AMOEBA has shown the high transphase capabilities. In the gas phase, AMOEBA is able to reach the MP2/CBS limit accuracy. At the same time, the accuracy in free energy estimation for different molecules including highly charged pieces has been very well maintained.

The high accuracy of AMOEBA in different phases will benefit not only to easier parameterization for complicated molecular systems including highly charged systems which are usually lack of experimental condensed phase properties, but also to describe and calculate accurate inter/intramolecular non-bond interactions. A better non-bond interaction is essential for molecular structural prediction, drug design researches, and molecular physical properties predictions. Our QM results also further show that it is critical to choose co-valence basis sets for metal ion *ab initio* calculations.

References

1. William L. Jorgensen, David S. Maxwell, & Tirado-Rives J (1996) Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *Journal of the American Chemical Society* 118(45):11225-11236.
2. Mobley DL, Bayly CI, Cooper MD, & Dill KA (2009) Predictions of Hydration Free Energies from All-Atom Molecular Dynamics Simulations. (Translated from English) *Journal of Physical Chemistry B* 113(14):4533-4537 (in English).
3. Mobley DL, Bayly CI, Cooper MD, Shirts MR, & Dill KA (2009) Small Molecule Hydration Free Energies in Explicit Solvent: An Extensive Test of Fixed-Charge Atomistic Simulations. (Translated from English) *Journal of Chemical Theory and Computation* 5(2):350-358 (in English).
4. Mobley DL, Dumont E, Chodera JD, & Dill KA (2007) Comparison of charge models for fixed-charge force fields: Small-molecule hydration free energies in explicit solvent. *Journal of Physical Chemistry B* 111(9):2242-2254.
5. Nicholls A*, et al.* (2008) Predicting small-molecule solvation free energies: An informal blind test for computational chemistry. (Translated from English) *J. Med. Chem.* 51(4):769-779 (in English).
6. Swope WC, Horn HW, & Rice JE (2010) Accounting for Polarization Cost When Using Fixed Charge Force Fields. I. Method for Computing Energy. (Translated from English) *Journal of Physical Chemistry B* 114(26):8621-8630 (in English).
7. Swope WC HH, Rice JE. (2010) Accounting for polarization cost when using fixed charge force fields. II. Method and application for computing effect of polarization cost on free energy of hydration. *J Phys Chem B* 114(26):8631-8645.
8. Stern HA, Rittner F, Berne BJ, & Friesner RA (2001) Combined fluctuating charge and polarizable dipole models: Application to a five-site water potential function. *Journal of Chemical Physics* 115:15.
9. Lamoureux G & Roux B (2003) Modeling induced polarization with classical Drude oscillators: Theory and molecular dynamics simulation algorithm. *Journal of Chemical Physics* 119:15.
10. Anisimov VM*, et al.* (2005) Determination of Electrostatic Parameters for a Polarizable Force Field Based on the Classical Drude Oscillator. *Journal of Chemical Theory and Computation* 1:16.
11. Gao J, Habibollahzadeh D, & Shao L (1995) A Polarizable Intermolecular Potential Functions for Simulations of Liquid Alcohols. *Journal of Physical Chemistry* 99:8.
12. Xie W, Pu J, MacKerell AD, Jr., & Gao J (2007) Development of a Polarizable Intermolecular Potential Function (PIPF) for Liquid Amides and Alkanes. *Journal of Chemical Theory and Computation* 3:12.
13. Ren PY & Ponder JW (2003) Polarizable atomic multipole water model for molecular mechanics simulation. *Journal of Physical Chemistry B* 107(24):5933-5947.
14. Ponder JW & Case DA (2003) Force fields for protein simulations. *Protein Simulations* 66:27-+.
15. Wang JM*, et al.* (2011) Development of Polarizable Models for Molecular Mechanical Calculations II: Induced Dipole Models Significantly Improve Accuracy of Intermolecular Interaction Energies. (Translated from English) *Journal of Physical Chemistry B* 115(12):3100-3111 (in English).

16. Ponder JW WC, Ren P, Pande VS, Chodera JD, Schnieders MJ, Haque I, Mobley DL, Lambrecht DS, DiStasio RA Jr, Head-Gordon M, Clark GN, Johnson ME, Head-Gordon T (2010) Current status of the AMOEBA polarizable force field. *J Phys Chem B.* 114(8):2549-2564.
17. Cieplak P, Caldwell J, & Kollman P (2001) Molecular Mechanical Models for Organic and Biological Systems Going Beyond the Atom Centered Two Body Additive Approximation: Aqueous Solution Free Energies of Methanol and N-Methyl Acetamide, Nucleic Acid Base, and Amide Hydrogen Bonding and Chloroform/Water Partition Coefficients of the Nucleic Acid Bases. *Journal of Computational Chemistry* 22:1048-1057.
18. Jiang W, *et al.* (2011) High-Performance Scalable Molecular Dynamics Simulations of a Polarizable Force Field Based on Classical Drude Oscillators in NAMD. (Translated from English) *Journal of Physical Chemistry Letters* 2(2):87-92 (in English).
19. Vorobyov IV, Anisimov VM, & MacKerell AD (2005) Polarizable empirical force field for alkanes based on the classical drude oscillator model. (Translated from English) *Journal of Physical Chemistry B* 109(40):18988-18999 (in English).
20. Yu H, *et al.* (2010) A Combined Experimental and Theoretical Study of Ion Solvation in Liquid N-Methylacetamide. (Translated from English) *Journal of the American Chemical Society* 132(31):10847-10856 (in English).
21. Maple JR, *et al.* (2005) A Polarizable Force Field and Continuum Solvation Methodology for Modeling of Protein–Ligand Interactions. *Journal of Chemical Theory and Computation* 1(4):694-715.
22. Devereux M, van Severen MC, Parisel O, Piquemal JP, & Gresh N (2011) Role of Cation Polarization in holo- and hemi-Directed [Pb(H(2)O)(n)](2+) Complexes and Development of a Pb(2+) Polarizable Force Field. (Translated from English) *Journal of Chemical Theory and Computation* 7(1):138-147 (in English).
23. Piquemal JP, Cisneros GA, Reinhardt P, Gresh N, & Darden T (2006) Towards a force field based on density fitting. *Journal of Chemical Physics* 124(10).
24. Ren P, Wu C, & Ponder JW (2011) Polarizable Atomic Multipole-Based Molecular Mechanics for Organic Molecules. *Journal of Chemical Theory and Computation* 7(10):3143-3161.
25. Halgren TA (1992) Representation Of Vanderwaals (Vdw) Interactions In Molecular Mechanics Force-Fields - Potential Form, Combination Rules, And Vdw Parameters. *Journal of the American Chemical Society* 114(20):7827-7843.
26. Jiao D, Golubkov PA, Darden TA, & Ren P (2008) Calculation of protein-ligand binding free energy by using a polarizable potential. *Proceedings of the National Academy of Sciences of the United States of America* 105(17):6290-6295.
27. Grossfield A, Ren P, & Ponder JW (2003) Single ion solvation thermodynamics from simulations with a polarizable force field. *Journal of the American Chemical Society* 125:14671-14682.
28. Jiao D, King C, Grossfield A, Darden TA, & Ren PY (2006) Simulation of Ca2+ and Mg2+ solvation using polarizable atomic multipole potential. *Journal of Physical Chemistry B* 110(37):18553-18559.
29. Kelly CP, Cramer CJ, & Truhlar DG (2005) SM6: A Density Functional Theory Continuum Solvation Model for Calculating Aqueous Solvation Free Energies of Neutrals, Ions, and Solute-Water Clusters. *J. Chem. Theory Comput.* 1:1133-1152.

30. Kelly CP, Cramer CJ, & Truhlar DG (2006) Aqueous Solvation Free Energies of Ions and Ion-Water Clusters Based on an Accurate Value for the Absolute Aqueous Solvation Free Energy of the Proton. *J. Phys. Chem. B* 110:16066-16081.
31. Stone AJ (2005) Distributed multipole analysis: Stability for large basis sets. *Journal of Chemical Theory and Computation* 1(6):1128-1132.
32. Stone AJ (2008) Intermolecular potentials. (Translated from English) *Science* 321(5890):787-789 (in English).
33. Ponder JW (2009) TINKER 5.0. (Washington Univeristy School of Medicine, Saint Louis).
34. Thole BT (1981) Molecular Polarizabilities Calculated with a Modified Dipole Interaction. (Translated from English) *Chemical Physics* 59(3):341-350 (in English).
35. Saff EB & Kuijlaars ABJ (1997) Distributing Many Points on a Sphere. *The Mathmatical Inteligencer* 19(1):7.
36. Feller D (1996) The Role of Databases in Support of Computational Chemistry Calculations. *J. Comp. Chem.* 17(3):16.
37. Schuchardt KL*, et al.* (2007) Basis Set Exchange: A Community Database for Computational Sciences. *J. Chem. Inf. Model.* 47(3):8.
38. Schmid R, Miah AM, & Sapunov VN (2000) A new table of the thermodynamic quantities of ionic hydration: Values and some applications (enthalpy-entropy compensation and Born radii). *Physical Chemistry Chemical Physics* 2:97-102.
39. Friendman HL & Krishnan CV (1973) *Water-A Comprehensive Treatise* (Plenum Press, New York).
40. Tissandier MD*, et al.* (1998) The Proton's Absolute Aqueous Enthalpy and Gibbs Free Energy of Solvation from Cluster-Ion Solvation Data. *J. Phys. Chem. A* 102:7787-7794.
41. Rizzo RC, Aynechi T, Case DA, & Kuntz ID (2006) Estimation of absolute free energies of hydration using continuum methods: Accuracy of partial charge models and optimization of nonpolar contributions. *J. Chem. Theory Comput.* 2:128-139.

Table 1. Molecular dipoles, quadrupoles and decomposition of hydration free energies

Molecule	Dipole (Debye)	Principal Axes Quadrupole (Buckingham)			ΔG_{elec} (kcal/mol)	ΔG_{vdw} (kcal/mol)
		Q_{xx}	Q_{yy}	Q_{zz}		
methane	0.000	0.000	0.000	0.000	-0.462±0.017	2.164±0.114
ethane	0.000	-0.488	0.244	0.244	-0.472±0.016	2.116±0.138
propane	0.110	-0.516	0.431	0.085	-0.629±0.018	2.228±0.156
n-butane	1.365	-2.509	2.015	0.494	-1.396±0.031	2.427±0.176
methanol	1.875	-2.434	3.314	-0.880	-6.949±0.098	2.112±0.131
ethanol	1.828	-4.221	4.212	0.008	-6.968±0.108	2.207±0.146
propanol	1.900	-6.506	11.633	1.124	-7.141±0.105	2.189±0.166
isopropanol	1.540	-3.638	4.680	-1.042	-6.873±0.145	2.559±0.196
phenol	1.461	-8.788	6.185	2.603	-6.939±0.095	1.740±0.188
p-cresol	1.289	-7.997	11.163	-3.167	-7.511±0.098	1.740±0.211
methylether	1.335	-2.448	3.230	-0.782	-4.307±0.144	2.057±0.238
benzene	0.000	-7.973	3.987	3.987	-2.919±0.052	1.556±0.181
toluene	0.387	-8.430	4.654	3.776	-3.147±0.053	1.454±0.193
ethylbenzene	0.302	-8.271	4.263	4.008	-3.305±0.059	2.178±0.217
methylamine	1.384	-3.845	2.102	1.744	-7.596±0.111	2.081±0.135
ethylamine	1.203	-6.116	3.168	2.948	-6.516±0.091	2.109±0.150
dimethylamine	1.166	-3.099	2.065	1.034	-5.055±0.115	1.625±0.147
trimethylamine	0.698	-3.504	1.824	1.680	-4.612±0.096	1.860±0.168
imidazole	3.900	-6.111	8.058	-1.947	-11.520±0.124	1.168±0.177
N-methylacetamide	3.877	-4.271	-1.351	5.622	-10.691±0.124	1.917±0.176
acetic acid	1.721	-7.991	7.549	0.442	-7.655±0.103	1.965±0.149
hydrogen sulfide	1.108	-3.919	2.850	1.068	-2.272±0.050	1.815±0.115
methylsulfide	1.607	-3.337	3.585	-0.248	-3.188±0.093	1.670±0.180
ethylsulfide	1.638	-2.599	3.508	-0.909	-3.867±0.066	2.032±0.171
dimethylsulfide	1.593	-2.434	4.032	-1.597	-3.442±0.061	1.487±0.162
methylethylsulfide	1.567	-2.411	4.204	-1.793	-4.082±0.099	1.964±0.217
water	1.772	-2.167	0.000	-0.330	-7.908±0.103	2.032±0.088
formamide	3.960	-2.696	3.454	-0.758	-10.374±0.116	1.597±0.129
pyrrolidine	1.285	-6.394	4.463	1.931	-6.509±0.087	1.566±0.175
3-formylindole	5.078	-9.440	16.120	-6.680	-9.352±0.109	0.819±0.236
3-ethylindole	2.077	-12.983	12.225	0.758	-3.608±0.077	1.295±0.258
acetamide	3.948	-3.417	5.551	-2.135	-11.007±0.120	1.626±0.153
benzamidine (+)	6.377	-29.503	33.187	-3.684	-48.268±0.208	1.565±0224
ethylimidazole (+)	3.764	-18.862	15.801	3.062	-50.550±0.175	1.558±0.204
nitrogen	0.000	-1.581	0.790	0.790	-0.295±0.013	2.508±0.110
acetaldehyde	2.706	-3.198	1.746	1.452	-4.670±0.073	2.181±0.141
indole	2.219	-13.429	11.483	1.946	-5.717±0.075	0.769±0.226
ammonia	1.512	-2.460	1.230	1.230	-5.923±0.083	2.029±0.099
dimethylsulfoxide	4.181	-6.169	5.653	0.515	-11.703±0.123	2.078±0.169
acetonitrile	3.915	-2.506	1.253	1.253	-5.538±0.074	1.765±0.138

Table 2. QM and AMOEBA Calculation Results for alkali ion-water and halogen-water dimers

Ion	Method[a]	Ion-OW Distance (Å)	E_{inter} (kcal/mol)[b]
Na[+]	MP2-FC/aVDZ	2.276	-21.978(-22.270)
	MP2-FC/aVTZ	2.279	-22.199(-22.390)
	MP2-FC/aVQZ	2.279	-22.351(-22.443)
	MP2-FC/aV5Z	2.275	-22.471(-22.516)
	MP2-FC/CBS(T->Q) [c]	2.279	-22.462
	MP2-FC/CBS(Q->5)	2.271	-22.597
	MP2-FC/6-311++G(3df,3pd)	2.270	-22.293(-22.754)
	MP2-Full/aVDZ	2.182[d];2.264[e]	-21.926(-24.194) [d]; -21.814(-22.798) [e]
	MP2-Full/aVTZ	2.165[d];2.261[e]	-22.217(-24.952) [d]; -22.613(-24.660) [e]
	MP2-Full/aVQZ	1.816 [d];2.248 [e]	-8.195(-44.009) [d]; -23.009(-31.877) [e]
	MP2-Full/aV5Z	2.148 [d];2.413 [e]	-23.162(-31.002)[d];-22.235(-22.275)[e]
	MP2-Full/6-311++G(3df,3pd)	2.227[d];2.244[e]	-23.226(-24.385)[d];-23.236(-24.372)[e]
	MP2-Full/cc-pCVDZ+aVDZ	2.240	-22.412(-23.182)
	MP2-Full/cc-pCVTZ+aVTZ	2.227	-23.269(-23.881)
	MP2-Full/cc-pCVQZ+aVQZ	2.221	-23.756(-23.976)
	MP2-Full/cc-pCV5Z+aV5Z	2.215	-23.998(-24.185)
	MP2-Full/CBS(T->Q)	2.217	-24.111
	MP2-Full/CBS(Q->5)	2.209	-24.252
	MP2-Full/aug-cc-pCVDZ+aVDZ	2.240	-22.456(-23.263)
	MP2-Full/aug-cc-pCVTZ+aVTZ	2.227	-23.319(-23.888)
	MP2-Full/aug-cc-pCVQZ+aVQZ	2.221	-23.768(-24.014)
	MP2-Full/CBS(T->Q)	2.217	-24.096
	AMOEBA	2.213	-23.940
K[+]	MP2-FC/6-311++G(3df,3pd)	2.618	-17.271(-17.575)
	MP2-Full/CVDZ+aVDZ[f]	2.636	-16.093(-16.807)
	MP2-Full/CVTZ+aVTZ	2.581	-17.271(-18.630)
	MP2-Full/CVQZ+aVQZ	2.586	-17.794(-18.549)
	MP2-Full/CBS(T->Q)	2.589	-18.176
	AMOEBA	2.579	-17.804
Mg[2+]	MP2-FC/6-311++G(3df,3pd)	1.931	-77.798(-78.624)
	MP2-Full/cc-pCVDZ+aVDZ	1.934	-77.439(-79.048)
	MP2-Full/cc-pCVTZ+aVTZ	1.911	-80.517(-81.703)
	MP2-Full/cc-pCVQZ+aVQZ	1.905	-81.768(-82.168)
	MP2-Full/cc-pCV5Z+aV5Z	1.903	-82.152(-82.398)
	MP2-Full/CBS(T->Q)	1.901	-82.681
	MP2-Full/CBS(Q->5)	1.901	-82.555
	MP2-Full/aug-cc-pCVDZ+aVDZ	1.934	-77.472(-79.180)
	MP2-Full/aug-cc-pCVTZ+aVTZ	1.911	-80.595(-81.909)
	MP2-Full/aug-cc-pCVQZ+aVQZ	1.905	-81.796(-82.233)
	MP2-Full/CBS(T->Q)	1.901	-82.672
	AMOEBA	1.876	-82.149
Ca[2+]	MP2-FC/6-311++G(3df,3pd)	2.246	-54.875(-55.377)
	MP2-Full/cc-pCVDZ+aVDZ	2.256	-53.582(-54.920)
	MP2-Full/cc-pCVTZ+aVTZ	2.222	-55.737(-57.229)
	MP2-Full/cc-pCVQZ+aVQZ	2.215	-56.969(-58.309)
	MP2-Full/cc-pCV5Z+aV5Z	2.216	-57.022(-58.923)
	MP2-Full/CBS(T->Q)	2.210	-57.868
	MP2-Full/CBS(Q->5)	2.217	-57.078
	AMOEBA	2.217	-57.056

Cl⁻	MP2-FC/aVDZ	3.137	-14.289(-15.239)
	MP2-FC/aVTZ	3.095	-15.241(-16.054)
	MP2-FC/aVQZ	3.091	-15.474(-15.969)
	MP2-FC/aV5Z	3.089	-15.584(-15.934)
	MP2-FC/CBS(T->Q)	3.088	-15.644
	MP2-FC/CBS(Q->5)	3.087	-15.699
	MP2-Full/aug-cc-pCVDZ+aVDZ	3.129	-14.326(-15.529)
	MP2-Full/aug-cc-pCVTZ+aVTZ	3.086	-15.211(-16.316)
	MP2-Full/aug-cc-pCVQZ+aVQZ	3.084	-15.547(-16.073)
	MP2-Full/aug-cc-pCV5Z+aV5Z	3.083	-15.649(-16.019)
	MP2-Full/CBS(T->Q)	3.083	-15.792
	MP2-Full/CBS(Q->5)	3.082	-15.756
	AMOEBA	3.119	-15.993
Br⁻	MP2-FC/aVDZ	3.332	-12.253(-13.693)
	MP2-FC/aVTZ	3.261	-13.084(-14.787)
	MP2-FC/aVQZ	3.237	-13.340(-14.749)
	MP2-FC/aV5Z	3.238	-13.472(-14.584)
	MP2-FC/CBS(T->Q)	3.219	-13.529
	MP2-FC/CBS(Q->5)	3.239	-13.610
	MP2-Full/aVDZ	3.327	-12.228(-14.015)
	MP2-Full/aVTZ	3.212	-12.836(-16.486)
	MP2-Full/aVQZ	3.206	-13.216(-15.492)
	MP2-Full/CBS(T->Q)	3.202	-13.493
	AMOEBA	3.280	-13.881
I⁻	MP2-FC/aVDZ	3.574	-10.060(-11.738)
	MP2-FC/aVTZ	3.498	-10.810(-13.063)
	MP2-FC/aVQZ	3.440	-10.988(-13.852)
	MP2-FC/aV5Z	3.416	-10.964(-13.994)
	MP2-FC/CBS(T->Q)	3.398	-11.118
	MP2-FC/CBS(Q->5)	3.391	-10.939
	MP2-Full/aVDZ	3.572	-10.042(-11.813)
	MP2-Full/aVTZ	3.493	-10.753(-13.180)
	MP2-Full/aVQZ	3.438	-10.985(-13.874)
	MP2-Full/CBS(T->Q)	3.398	-11.154
	AMOEBA	3.550	-11.221

[a] aVXZ (X=D,T,Q,5) is the abbreviation of aug-cc-pVXZ (X=D,T,Q,5) and the basis set used for the atoms in water when the core-valence basis set is used for the ions or all the atoms in the dimer when there is no core-valence basis set. For an example, pCVTZ+aVTZ means the core valence basis set pCVTZ is applied on the ion and aug-cc-pVTZ is used for water. MP2-Full is the full MP2 and MP2-FC is the frozen core MP2. MP2-FC is the default MP2 method in Gaussian and most QM packages.

[b] The value in the parenthesis is the BSSE uncorrected interaction energy.

[c] CBS is the complete basis set limit extrapolation of the above two calculations.

[d] Regular structural optimization followed by BSSE single point correction using MP2 electron density.

[e] Structural optimization is done on counterpoise surface; Then BSSE correction using MP2 electron density is performed.

[f] The available core valence basis sets in EMSL are Feller Misc. CVXZ (X=D,T,Q) basis sets.

Table 3. QM and AMOEBA Calculation Results for ammonium ion-water and organic acid-water dimers

	N(O)-OW distance [a]		Intermolecular Energy (kcal/mol) [b]			
	MP2-FC/6-311++G(2d,2p)	AMOEBA	MP2-FC/aVTZ	MP2-FC/aVQZ	MP2-FC/CBS(T->Q)	AMOEBA
NH_4^+	2.714	2.704	-20.472(-21.023)	-20.770(-21.063)	-20.988	-20.316
$CH_3NH_3^+$	3.500	3.528	-18.329(-18.945)	-18.603(-18.913)	-18.802	-18.961
$(CH_3)_2NH_2^+$	1.747	1.756	-16.819(-17.508)	-17.065(-17.393)	-17.244	-17.078
$(CH_3)_3NH^+$	1.765	1.766	-15.856(-16.635)	-16.071(-16.431)	-16.235	-15.987
$(CH_3)_4N^+$ [c]	-	-	-	-	-	-
$HCOO^-$	2.656	2.725	-17.319(-18.170)	-17.650(-18.126)	-17.892	-18.193
CH_3COO^-	2.648	2.703	-17.391(-18.333)	-17.740(-18.245)	-17.996	-17.587
$CH_3CH_2COO^-$	2.651	2.698	-17.290(-18.235)	-17.628(-18.122)	-17.875	-17.184

[a] Dimer structures were optimized with MP2/6-311++G(2d,2p) and the we did single point calculation with basis sets aug-cc-pVTZ and aug-cc-pVQZ when MP2 electron density turned on for SCF cycles.
[b] The energies are BSSE corrected and the values in the parenthesis are the intermolecular energies before BSSE correction.
[c] $(CH_3)_3N^+$ was parameterized. Since there is not direct hydrogen bonding between this ion and water, the dimer structure and energies were not calculated. Parameters were transferred from $(CH_3)_3NH^+$ and then later used for hydration free energy calculations.

Table 4. Absolute and relative solvation free energies of some ions

Ion	AMOEBA Predicted	Relative HFE		
		AMOEBA	Expt.(30)	Diff.
K^+	-74.366 ± 0.322	0.000	0.000	0.000
Na^+	-92.058 ± 0.355	-17.692	-17.447	-0.245
Mg^{2+}	-438.658 ± 0.617	-364.292	-363.767	-0.525
Ca^{2+}	-361.830 ± 0.559	-287.464	-285.612	-1.852
Cl^-	-85.750 ± 0.355	0.000	0.000	0.000
Br^-	-79.619 ± 0.347	6.131	6.2	-0.069
I^-	-70.693 ± 0.337	15.057	14.6	0.457
$HCOO^-$	-87.006 ± 0.398	-1.256	-1.70	0.444
CH_3COO^-	-89.092 ± 0.433	-3.342	-3.10	-0.242
$CH_3CH2COO^-$	-87.551 ± 0.470	-1.801	-1.70	-0.101

Table 5. The methylation effects on hydration free energy of ammonium ions

	AMOEBA	Relative	Truhlar's Collections(29, 30)	Diff
NH4+	-73.429 ± 0.329	0.000	0.0	0.000
CH3NH3+	-64.325 ± 0.350	9.104	8.8	0.304
(CH3)2NH2+	-56.394 ± 0.344	17.035	17.4	-0.365
(CH3)3NH+	-49.336 ± 0.343	24.093	24.1	-0.007
(CH3)4N+	-43.215 ± 0.339	30.214	-	

Table 6. Hydration free energy of neutral salts

	AMOEBA	Experimental
NaCl	-177.808	-177.8(Schmid(38)) -177.9(Friedman(39)) -177.8(Tissandier(40))
KCl	-160.115	-160.3(Schmid(38)) -160.4(Friedman(39)) -160.6(Tissandier(40))
NH_4Cl	-159.247	-159.0 (Rizzo(41))
CH_3NH_3Cl	-150.143	-151.0 (Rizzo(41))