

Ion-Hydroxyl Interactions: From High-Level Quantum Benchmarks to Transferable Polarizable Force Fields

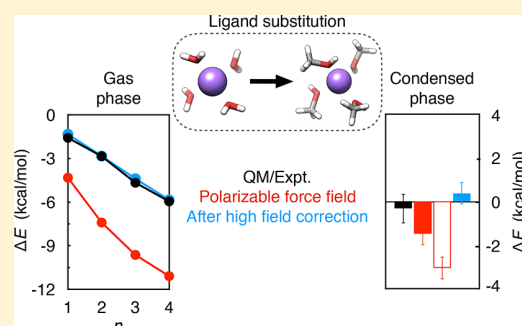
Vered Wineman-Fisher,[†] Yasmine Al-Hamdani,[‡] Iqbal Addou,[†] Alexandre Tkatchenko,[‡] and Sameer Varma^{*,†}

[†]Department of Cell Biology, Microbiology and Molecular Biology, University of South Florida, Tampa, Florida 33620, United States

[‡]Physics and Materials Science Research Unit, University of Luxembourg, 162a avenue de la Fiancerie, Luxembourg City, L-1511, Luxembourg

Supporting Information

ABSTRACT: Ion descriptors in molecular mechanics models are calibrated against reference data on ion–water interactions. It is then typically assumed that these descriptors will also satisfactorily describe interactions of ions with other functional groups, such as those present in biomolecules. However, several studies now demonstrate that this transferability assumption produces, in many different cases, large errors. Here we address this issue in a representative polarizable model and focus on transferability of cationic interactions from water to a series of alcohols. Both water and alcohols use hydroxyls for ion-coordination, and, therefore, this set of molecules constitutes the simplest possible case of transferability. We obtain gas phase reference data systematically from “gold-standard” quantum Monte Carlo and CCSD(T) methods, followed by benchmarked vdW-corrected DFT. We learn that the original polarizable model yields large gas phase water → alcohol transferability errors – the RMS and maximum errors are 2.3 and 5.1 kcal/mol, respectively. These errors are, nevertheless, systematic in that ion-alcohol interactions are overstabilized, and systematic errors typically imply that some essential physics is either missing or misrepresented. A comprehensive analysis shows that when both low- and high-field responses of ligand dipole polarization are described accurately, then transferability improves significantly – the RMS and maximum errors in the gas phase reduce, respectively, to 0.9 and 2.5 kcal/mol. Additionally, predictions of condensed phase transfer free energies also improve. Nevertheless, within the limits of the extrathermodynamic assumptions necessary to separate experimental estimates of salt dissolution into constituent cationic and anionic contributions, we note that the error in the condensed phase is systematic, which we attribute, at least, partially to the parametrization in long-range electrostatics. Overall, this work demonstrates a rational approach to boosting transferability of ionic interactions that will be applicable broadly to improving other polarizable and nonpolarizable models.



INTRODUCTION

Ions are ubiquitous in nature and vital to numerous physiological processes.¹ They participate primarily by either interacting directly with biomolecules and modulating their activities or serving as charge carriers in electrical responses of cells and tissues. Mechanistic understanding of many of these processes, such as the transport of ions through membranes,² requires molecular details of how ions bind to and unbind from biomolecules. Consequently, an understanding of such processes requires an understanding of the differences between an ion’s hydrated and biomolecule-bound states.

Molecular mechanics (MM) simulations hold tremendous potential in yielding such mechanistic insight.^{3,4} They are now used routinely to simulate events at the nano-to-microsecond time scale and for deriving relationships between the structures, dynamics, and functions of biomolecules. Nevertheless, the reliability of their predictions depends on their underlying potential energy functions. Several different

potential energy functions or force field models have been developed and have also been strategically refined over the past decades, significantly improving their predictive capabilities.^{5–8}

Systematic improvements have also been made to force field models of ionic interactions.^{9–27} However, most effort has been dedicated to improving descriptions of ion–water interactions. As such, calibrating ion descriptors against reference data on ion–water interactions does not guarantee meaningful predictions of ion interactions with other chemical groups. In fact, a compilation of recent studies shows that it is this transferability assumption that breaks down for many fundamental test cases.^{17,20,24,26,28–32} This is not surprising for nonpolarizable models that do not utilize explicit functions for describing induced effects and rely on the assumption that mean field approximations of induced effects in water are

Received: November 28, 2018

Published: March 4, 2019

transferable. Certainly, inclusion of explicit polarization improves performance,^{20,25–27} even in water;^{9–16} however, large transferability errors remain.^{20,24–26,28–32}

A typical approach to address transferability in both polarizable and nonpolarizable MM models consists of calibrating separate sets of nonbonded (NB) descriptors for every distinct pair of ion and its coordinating chemical group.^{18,21–27} This “NB-fix” strategy is elegant in that it does not require recalibration of interactions of the coordinating chemical groups with other chemical groups. In addition, its implementation is straightforward and does not sacrifice computational efficiency. However, the NB-fix strategy assigns all error corrections to one single interaction term, which is typically the Lennard-Jones term, although there is no supporting information of this term being the source of error. While NB-fix does improve transferability, at least for small ionic clusters, for larger clusters where many-body effects are important the error remains high.^{18,21–27} For example, even after NB-fix in a recent study,²⁵ the root-mean-square (RMS) error for ethanol and *N*-methylacetamide clusters of cations remained greater than 5 kcal/mol, and the maximum error exceeded 10 kcal/mol. It is important to note here that these errors were estimated by comparing against a density functional approximation (DFA) that was not benchmarked for ionic interactions, and, therefore, the error estimates themselves remain debatable. This, in fact, points to a pertinent issue with addressing transferability: access to reference data for ionic clusters is limited. As such, obtaining reference information from theoretical methods has been challenging in the past, but recent developments now make this possible. Additionally, DFAs can potentially predict a wide spread of ion-ligand binding energies, and, therefore, predictions from DFAs need to be validated against higher-level methods or experiment. Evidently, both force field and DFA development will benefit from having accurate reference information for benchmarking.

In this work, we adopt a systematic approach to address transferability. This begins with obtaining high level reference data to serve as benchmarks for a van der Waals inclusive DFA,^{33,34} which is then used to obtain gas phase reference data for ionic clusters, including those that are computationally demanding for treatment with “gold-standard” quantum methods. Benchmarking of small clusters is first carried out against coupled cluster theory with single, double, and perturbative triple excitations (CCSD(T)). Since CCSD(T) scales poorly (N^7) with systems size (N), we use diffusion Monte Carlo (DMC), which scales better (N^4) and is not affected adversely by shortfalls in basis sets, for benchmarking larger clusters. Indeed, DMC has been demonstrated to be as accurate as CCSD(T) for uncharged clusters,^{35–39} and here we address key challenges in DMC calculations for charged clusters and show that it reproduces CCSD(T) energies within stochastic errors.

We then use gas phase data from the benchmarked DFA^{33,34} and condensed phase data from experiments^{40,41} to analyze the performance of a representative MM model, where we methodically seek error sources and derive subsequent potential solutions. We use the polarizable AMOEBA model^{11,42} as a representative and focus on transferability of cationic (Na^+ and K^+) interactions from water to a series of alcohols. Both water and alcohols use hydroxyls for ion-coordination, and, therefore, this set of small molecules constitutes the simplest possible case of transferability. Yet,

as we note in the **Results** section, this representative MM model yields large water \rightarrow alcohol transferability errors—the RMS error is 2.3 kcal/mol and the maximum error exceeds 5 kcal/mol. Similar errors in water \rightarrow ethanol transferability were also reported for another popular polarizable model,²⁵ despite NB-fix corrections. Given that hydroxyl groups serve as key coordinators of cations in many proteins,^{43,44} getting water \rightarrow alcohol transferability right is also important from the perspective of studying ion-driven processes in proteins. Additionally, sticking to a single coordinator chemistry in this work allows us to derive systematic trends and relationships.

METHODS

Reference Energies. Diffusion Monte Carlo (DMC) can be used to compute noncovalent interaction energies with accuracy better than 1 kcal/mol.^{35–39,45–47} Typically, this level of accuracy is attributed to CCSD(T), but since CCSD(T) scales poorly with respect to system size and computer processors, DMC provides a suitable alternative for large clusters. To date, DMC has been used to compute benchmark interactions for a variety of uncharged systems. The DMC reference interaction energies that we compute here are for positively charged clusters containing K^+ and Na^+ cations.

To establish a converged protocol and also test the underlying approximations within DMC, we first compare ion–water interaction energies from DMC against those recently reported using complete basis set (CBS) extrapolated CCSD(T).⁴⁸ Guiding wave functions for DMC are computed from all-electron trial wave functions using ORCA⁴⁹ and the LDA and PBE exchange–correlation functionals. The Ahlrichs def2-QZVPP basis set⁵⁰ is used for all atoms. Note that we initially tested the use of the standard Trail and Needs pseudopotentials,⁵¹ and since they contain one explicit valence electron for Na and K, they left bare pseudopotentials for their cationic forms. This yielded overestimated interaction energies of K^+ with water (~ 2 kcal/mol) (see the **Supporting Information** for further details). We, therefore, switched to more expensive all-electron DMC, which yields agreement with CCSD(T) within the stochastic error of DMC results. The CASINO v2.13.610 code⁵² is used for computing DMC interaction energies. We use variational Monte Carlo (VMC) for optimizing trial wave functions with a Jastrow factor containing explicit electron–electron, electron–nucleus, and electron–electron–nucleus terms. We note that the use of all-electron wave functions also requires a particularly small time-step in DMC—using the locality approximation and following careful convergence tests we find the K^+ -water interaction energy converges with a 0.001 au time-step (see the **Supporting Information**). The fixed-node approximation is a well-known error source in DMC, which we also probe here by comparing HF, B3LYP, and LDA starting trial wave functions (see the **Supporting Information**). The resulting DMC interaction energies agree within 0.3 kcal/mol. For each system in this study, we perform DMC until stochastic errors of less than 1 kcal/mol are reached in interaction energies. To further ascertain the DMC results, we compare K^+ -methanol and Na^+ -methanol interaction energies from DMC against those obtained from CCSD(T). For details of the CCSD(T) calculations, see the **Supporting Information**. The all-electron DMC results agree with those obtained from these CCSD(T) calculations.

Table 1. Cluster Binding Energies (in kcal/mol) from First-Principles Methods: CCSD(T) (from Ref 48), DMC, and PBE0+TS

Na ⁺ /#ligands	H ₂ O			CH ₃ OH	
	CCSD(T)	DMC	PBE0+TS	DMC	PBE0+TS
1	-24.1	-24.5 ± 0.2	-24.7	-26.5 ± 0.3	-26.3
2	-22.9	-23.0 ± 0.3	-23.5	-24.4 ± 0.6	-25.0
3	-21.3	-21.6 ± 0.4	-22.0	-22.9 ± 0.6	-23.4
4	-19.8		-20.4	-21.1 ± 1.2	-21.7
K ⁺ /#ligands	H ₂ O			CH ₃ OH	
	CCSD(T)	DMC	PBE0+TS	DMC	PBE0+TS
1	-18.0	-17.9 ± 0.3	-18.2	-19.0 ± 0.3	-19.3
2	-17.0	-17.1 ± 0.3	-17.2	-18.1 ± 0.4	-18.1
3	-16.4	-15.9 ± 0.5	-16.3	-16.8 ± 0.5	-17.2
4	-15.7	-15.3 ± 0.4	-15.4	-16.6 ± 0.6	-16.3

Since DMC is significantly more expensive than DFT, we use the reference information from CCSD(T) and DMC to benchmark a vdW-corrected DFT exchange-correlation functional, namely PBE0+TS.^{33,34} The PBE0 hybrid functional contains 25% exact exchange and is supplemented by Tkatchenko-Scheffler (TS) corrections for dispersion. Exact exchange is particularly important in hydrogen bonded and charge transfer systems since it alleviates the delocalization error in DFT based approximations. All PBE0+TS calculations are performed using the FHI-AIMS package⁵³ with 'really tight' basis sets. Total energies are converged to within 10⁻⁶ eV, and electron densities are converged to within 10⁻⁵ electrons. Geometry optimizations are carried out with force criterion of 10⁻³ eV/Å and the PBE0+TS functional. The starting configurations for optimizations are taken from our previous studies,^{31,54} where they were optimized using the B3LYP density functional. The ion-ligand cluster geometries used in DMC are those obtained from PBE0+TS optimizations.

Solvation Free Energies. Condensed phase solvation free energies of ions are determined using Bennett's acceptance ratio (BAR),⁵⁵ and the configurational ensembles needed for BAR are obtained from molecular dynamics. For a system sampled equally in two different states *A* and *B*, Bennett shows that the free energy difference between them (ΔF) satisfies the following condition

$$e^{-\beta(\Delta F - c)} = \frac{\langle f(\beta(U_B - U_A - c)) \rangle_A}{\langle f(\beta(U_A - U_B + c)) \rangle_B} \quad (1)$$

In the equation above, $f(x) = 1/(1 + e^x)$ is the Fermi-Dirac distribution that satisfies the detailed balance condition. U_A and U_B are the potential energies obtained for the same configuration but computed using functions describing, respectively, states *A* and *B*. The triangular brackets indicate averages over configurational space sampled in states, *A* or *B*, indicated by subscripts. c is the energy offset, which is determined iteratively such that $c \approx \Delta F$.

The variance in ΔF depends inversely on sampling and also on the overlap of configurational space between states.⁵⁵ Therefore, instead of computing solvation free energies in one step, we compute them in 20 steps. We first scale down the ion's electrostatic parameters in ten steps, and then in the next ten steps we scale down the van der Waals parameters.^{11,29} In total, we carry out 21 separate MD simulations under NVT conditions ($T = 298$ K), each of which is 200 ps long, a time scale 1–2 orders in magnitude greater than the residence times of water in the first shells of Na⁺ and K⁺ ions.⁵⁶ In 11 of these

simulations, the ion's van der Waals parameters are set to their original assigned values, while its charge and polarizability are assigned different values, scaled from their original values to zero according to $q(\lambda) = \lambda q_0$ and $\alpha(\lambda) = \lambda \alpha_0$ with $\lambda = \{1, 0.9, \dots, 0.1, 0\}$. In the remaining ten simulations, the ion's charge and polarizability are set to zero, but its van der Waals parameters are assigned different values, scaled from their actual values (ϵ_0, r_0) to zero according to $r(\lambda) = 1 + \lambda(r_0 - 1)$, and $\epsilon(\lambda) = \lambda \epsilon_0$ with $\lambda = \{0.9, 0.8, \dots, 0.1, 0\}$. The first 100 ps of each of these simulations are discarded as equilibration, and the final 100 ps are used for computing ΔF .

The starting configuration of each of the 21 simulations is identical and is taken from an equilibrated trajectory of the ion in the condensed phase. Specifically, an ion is first placed in a pre-equilibrated cubic box of solvent molecules by removing one solvent molecule and replacing the solvent molecule with the ion. This box is then equilibrated under NPT conditions for 500 ps, with $P = 1$ atm and $T = 298$ K. The final 100 ps of the trajectory are used to obtain the average box length. The snapshot that has box dimensions closest to the average is selected as the starting configuration for free energy calculations. The pure water box contains 500 molecules, and its starting configuration is taken from the TINKER package.⁴² Each of the pure alcohol boxes contains 512 molecules and is constructed by periodic replication of an individual alcohol. Each of the pure solvent boxes is equilibrated for 5 ns under NPT conditions until densities stabilize. The final densities of methanol, ethanol, and propanol are, respectively, 0.77, 0.77, and 0.79 g/cm³, which are consistent with experiment.^{57–59}

The reported statistical errors in solvation free energies are the sum of errors in the 20 steps (windows). The statistical error in each window is obtained using a bootstrap Monte Carlo procedure.¹¹ In this approach, for a set of N data points, the same number of points (N) is chosen but randomly (duplication allowed), and the free energy difference is computed. The statistical uncertainty is essentially estimated as the standard deviation of the free energy difference over 10⁵ Monte Carlo trials.

All MD simulations are carried out using TINKER version 7.1, and the following control functions and parameters are chosen to be different from defaults. Integration is carried out using the RESPA algorithm with an outer time step of 1 fs.⁴² Temperature is regulated using an extended ensemble approach⁶⁰ and with a coupling constant of 0.1 ps. In isobaric simulations, pressure is regulated using a Monte Carlo

approach^{61,62} with a coupling constant of 0.1 ps. Electrostatic interactions are computed using particle mesh Ewald with a direct space cutoff of 9 Å. van der Waals interactions are computed explicitly for interatomic distances smaller than 9 Å. The convergence cutoff for induced dipoles is set at 0.01 D.

RESULTS AND DISCUSSION

We first report CCSD(T), DMC, and PBE0+TS reference gas-phase interaction energies for ion–ligand clusters. We then use these reference energies along with experimental solvation free energies to methodically evaluate the performance of a representative molecular mechanics polarizable model, AMOEBA,^{11,42} in predicting water → alcohol transferability. In this process, we also introduce physically motivated parameters into the model, which are shown to boost transferability.

Reference Energies for Ion–Water and Ion–Alcohol Clusters. Table 1 shows that ion–water cluster energies computed using DMC are in excellent agreement with those computed from basis set extrapolated CCSD(T). As also noted in the Methods section, the K⁺–methanol and Na⁺–methanol binding energies from DMC are also in agreement with those that we obtain from basis set extrapolated CCSD(T).

Table 1 also compares ion–water and ion–methanol cluster energies computed using PBE0+TS against DMC. PBE0+TS predictions are in agreement with DMC in terms of not only the interaction energies per ligand but also the trend with respect to cluster-size. We had noted in an earlier study³¹ that within a harmonic approximation, PBE0+TS also predicts gas phase ion–water cluster enthalpies and free energies consistent with experiment. We, therefore, utilize PBE0+TS to obtain reference energies for the remaining ion–alcohol clusters.

Transferability of Original Model. To evaluate transferability, we first determine gas phase water → alcohol substitution energies $\Delta E = E_{AX_n} - nE_X - E_{AW_n} + nE_W$ for the reactions below



where A refers to the ion, W refers to water, and X refers to one of the three alcohols, methanol, ethanol, and propanol. Figure 1 compares the substitution energies computed using the original molecular mechanics (MM) model against those computed from PBE0+TS. The RMS error is 2.3 kcal/mol, and the maximum error is greater than 5 kcal/mol. If the errors are estimated for progressive addition of methyl groups or methylene bridges, that is for substitution reactions $W \rightarrow M$, $M \rightarrow E$, and $E \rightarrow P$, then the RMS error in substitution energies is even higher at 3.2 kcal/mol. We note that these transferability errors are not related to topological differences between cluster geometries optimized using the original MM model and PBE0+TS, as the cluster geometries obtained from the two methods are similar.

Next, we use BAR to compute condensed phase transfer free energies of ions. For each ion we first determine $\Delta F_{water \rightarrow alcohol} = \Delta F_{alcohol} - \Delta F_{water}$, where ΔF_{water} and $\Delta F_{alcohol}$ are the solvation free energies of an ion in condensed phases of water and alcohols, respectively. Then, using $\Delta F_{water \rightarrow alcohol}$ we also determine the difference between the transfer free energies of K⁺ and Na⁺ ions, $\Delta F_{water \rightarrow alcohol}^{K \rightarrow Na} = \Delta F_{water \rightarrow alcohol}^{Na} - \Delta F_{water \rightarrow alcohol}^{K}$. We compare these predictions against experimental estimates and find clear discrepancies (data plotted in Figure 4).

It is important to note that estimations of single ion $\Delta F_{water \rightarrow alcohol}$ from experiments involve extrathermodynamic

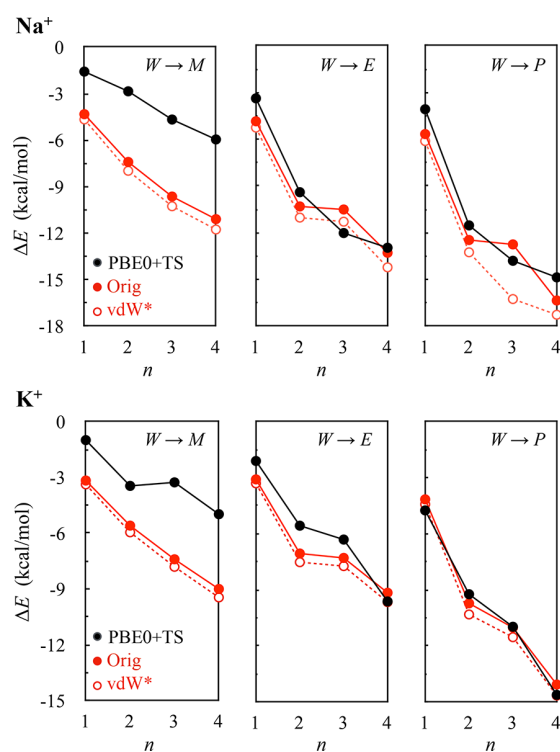


Figure 1. Comparison of substitution energies ΔE obtained from PBE0+TS and the original (MM) and recalibrated (MM–vdW*) molecular mechanics model. The ΔE values are estimated for the substitution reactions given by eq 2, and they are computed using geometries relaxed separately at the respective levels of theory. The letters W, M, E, and P refer, respectively, to water, methanol, ethanol, and propanol.

assumptions necessary to separate out the energetics of salt dissolution into their constituent cationic and anionic contributions.^{63,64} However, estimations of $\Delta F_{water \rightarrow alcohol}^{K \rightarrow Na}$ from experiments do not depend on extrathermodynamic assumptions.⁶⁵ We compare our predictions against experimental estimates based on two different extrathermodynamic assumptions, the TATB assumption^{40,63} and CPA.^{41,66} TATB assumes that the solvation properties of the large tetraphenylarsonium (TA) cation and the large tetraphenylborate anion (TB) are roughly equal.⁶³ CPA refers to the cluster-pair approximation⁶⁶ in which the energetics of cations and anions in their individual solvent clusters are expected to converge toward each other rapidly following a monotonous trend, although recent studies suggest a more complex convergence.⁶⁷

Overall, we find that the original MM model yields water → alcohol transferability errors in both gas and condensed phases. Nevertheless, the errors are systematic in that the interactions of ions with alcohols are overstabilized with respect to their corresponding interactions with water. This suggests that some essential physics is either missing or misrepresented, but what could be the source(s) of error? First, it is possible that the NB descriptors of ions are not calibrated appropriately against water, which may lead to an incorrect baseline for computing transferability. Second, we recognize that the ligands in this MM model, and, in fact, in all other MM models, are parametrized in the absence of ions. Consequently, we can expect that MM models, in general, will describe well the response at low electric fields near dipoles. As such, the water–water, alcohol–alcohol, and water–alcohol dimerization

Table 2. Ion–Water Binding Energies (ΔE in kcal/mol) and Optimum Distances (d in Å) Prior to (Orig¹¹) and after Recalibration (vdW*) of vdW Parameters against PBE0+TS^a

ion	# waters	PBE0+TS		Orig		vdW*	
		ΔE	d	ΔE	d	ΔE	d
Na ⁺	1	-24.7	2.23	-23.2	2.23	-25.0	2.17
	2	-47.2	2.25	-43.8	2.26	-47.2	2.20
	3	-66.1	2.27	-61.2	2.30	-65.6	2.24
	4	-81.9	2.30	-75.3	2.34	-80.6	2.28
K ⁺	1	-18.4	2.63	-17.3	2.60	-18.0	2.57
	2	-34.8	2.67	-32.7	2.64	-34.6	2.58
	3	-49.4	2.70	-46.4	2.67	-49.1	2.60
	4	-62.1	2.72	-58.2	2.70	-61.7	2.63
MAE				6.7%	0.9%	0.7%	2.5%

^aThe ion–water binding energy is defined as $\Delta E = E_{\text{cluster}} - E_{\text{ion}} - nE_{\text{W}}$, where E_{cluster} , E_{ion} , and E_{W} are the energies of the cluster, ion, and water following independent energy optimizations, and n is the number of water molecules in the cluster. d is the distance between the ion and the oxygen atom of water. MAE is an abbreviation for mean absolute error.

energies of this MM model are in excellent agreement with the Møller–Plesset second order perturbation (MP2) theory⁶⁸ (RMS error = 0.23 kcal/mol).^{69,70} However, electric fields near ions are significantly stronger than those near dipoles, and so it is possible that although the MM model describes well the response at low dipolar fields, it perhaps yields inaccurate responses for the stronger ionic fields. We now examine these two error sources and evaluate the effects of correcting them.

Recalibrating Baseline — Ion–Water Interactions.

The original vdW descriptors of Na⁺ and K⁺ ions were calibrated, in typical strategy, against ion–water pair interaction energies and condensed phase solvation energies.¹¹ We note first that the ion–water interaction energies used for calibration¹¹ are different from the reference energies obtained in Table 1. For example, Na⁺–water binding energy used for the original calibration was $\Delta E = -23.6$ kcal/mol, whereas the reference value obtained above is $\Delta E = -24.7$ kcal/mol. Second, we note that only two-body interaction energies were used as targets for calibration, as is done typically, which does not necessarily ensure that the calibrated parameters will also reproduce binding energies for larger clusters. In fact, as shown in Table 2, the mismatch between reference energies and those predicted from the originally calibrated model grows with cluster size. For a four-water cluster, the error is greater than 6 kcal/mol. It is, therefore, plausible that errors result, at least partially, from lack of many-body interactions in the target set.⁷¹

To examine this, we recalibrate the vdW descriptors of ions against the new reference data that also includes many-body interactions. In this model, vdW interactions are described using a buffered 14-7 function

$$U_{vdw} = \epsilon_{ij} \left(\frac{1.07}{\rho_{ij} + 0.07} \right)^7 \left(\frac{1.12}{\rho_{ij}^7 + 0.12} - 2 \right) \quad (3)$$

where ϵ_{ij} in kcal/mol is the potential well depth, and $\rho_{ij} = r_{ij}/r_{ij}^0$, where r_{ij} in Ångströms is the distance between sites i and j and r_{ij}^0 is the minimum energy distance. Reoptimizing vdW against the new target data, which includes all of the PBE0+TS data listed in Table 2, reduces the MAE against target data from 3.8% to 1.6%. The recalibrated descriptors (ϵ , r^0), which we refer to as the vdW* model, are (0.48, 2.50) for Na⁺ and (0.59, 3.51) for K⁺. The original vdW descriptors of Na⁺ were (0.26, 3.02), and those of K⁺ were (0.35, 3.71).

While recalibration of ion NB descriptors improves interactions with water, it does not improve water \rightarrow alcohol transferability. Additionally, the transferability errors remain systematic. The gas phase substitution energies for the vdW* model are shown in Figure 1, and condensed phase transfer free energies are shown in Figure 4.

Recalibrating High Field Response. Small molecules, including water and alcohols, in MM models are parametrized in the presence of dipolar fields and not the stronger ionic fields. Consequently, it is possible that although they describe well low field effects, they may not describe or capture high field responses satisfactorily.

To examine this, we determine using MP2 theory the induced dipoles of water and alcohols in the presence of a unit point charge (+1) and compare them against those estimated from the MM model. Figure 2 shows the component of the induced dipole along the vector connecting the unit charge to the oxygen atoms of the molecules. We note that the induced dipoles of all molecules are consistently underestimated at short distances from the point charge, but the error decreases with increasing distance from the point charge. In other words, while the polarization model performs well at low electric fields, its error increases at stronger fields that are present in an ion's first coordination shell.

To understand this field-dependent error, we examine the polarization model. Induced dipoles in this model are computed self-consistently, that is, the induced dipole produced at a given site is allowed to polarize other sites, and this mutual induction is iterated until the induced dipoles at all sites converge. Additionally, a Thole approach is employed to prevent polarization catastrophe, wherein electrostatic interactions are damped in the short-range. Damping is applied to only one of the two sites of an interaction pair using $\rho = \frac{3a}{4\pi} e^{-ar_{ij}^3/\sqrt{\alpha\alpha_j}}$, where r_{ij} is the distance between two sites with atomic polarizabilities α , and “ a ” is a dimensionless width parameter of the damped charge distribution that controls the damping strength.

The performance of this polarization model will, therefore, depend on the values chosen for atomic polarizabilities and damping factors. It is important to note that these parameters are not calibrated against high electric field responses — α are tuned to reproduce experimental values of molecular polarizabilities,^{42,69,72–74} and the damping factors for water and

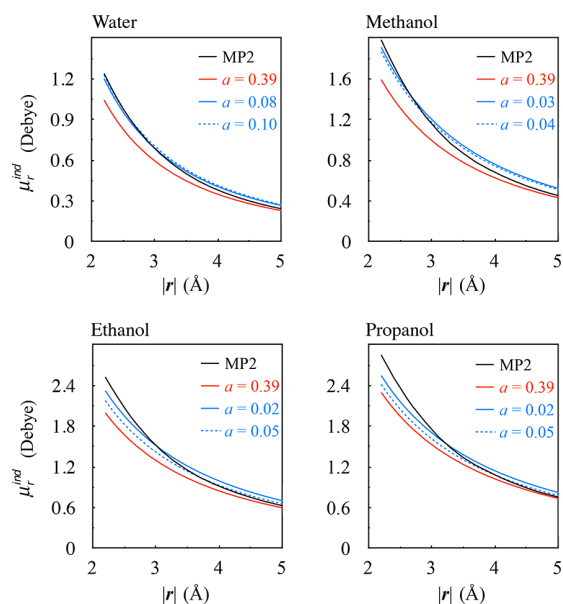


Figure 2. Comparison of induced dipoles obtained from the MM model against those obtained from MP2. μ_r^{ind} is the component of the induced dipole parallel to the vector connecting the unit point charge to the oxygen atoms of the molecules. Induced dipoles obtained using the original MM model with damping factor $a = 0.39$ are shown in red, and those obtained using modified damping factors are shown in blue. The solid blue lines show induced dipoles obtained using damping factors fitted against MP2 values in the distance range 2.2–3.1 Å, and the dashed blue lines show induced dipoles obtained using damping factors fitted against MP2 values in the distance range 2.6–3.6 Å for water and 2.6–3.65 Å for alcohols. The two distance ranges essentially correspond to the first coordination shells of Na^+ and K^+ ions determined from radial distribution functions. The geometries used for these calculations are essentially derived from the geometries of the molecules optimized in the presence of a Na^+ ion but after replacing the coordinates of the ion with a point charge and translating the molecule along the vector connecting the point charge and the oxygen atom of the molecule. We note that reoptimization of geometries following translation and with constraints on non-hydrogen atoms has a negligible effect on the reported results. MP2 electron densities are estimated using the aug-cc-pVDZ basis set, and we note that the effect of using a larger basis set (aug-cc-pVTZ) is negligible on the electron density.

alcohols are set at $a = 0.39$ to reproduce cluster energies of water and other molecules.^{73,74}

Figure 2 shows that the damping factor can be recalibrated to fit the field response at short distances. We define short distances as the range adopted by molecules in an ion's first shell in the condensed phase. This range is 2.2–3.1 Å for water and alcohols interacting with Na^+ , 2.6–3.5 Å for waters interacting with K^+ , and 2.6–3.65 Å for alcohols interacting with K^+ . These ranges are estimated by simulating ions in condensed phases under NVT conditions – we note that these distance ranges are not affected by the different sets of parameters used in this study. We also note in Figure 2 that the calibration of damping factor against field response yields not only different values for different distance-ranges but also different values for different molecules.

We also note from Figure 2 that when damping factors are chosen to fit the field response at short distances, then the error in field response increases at larger distances. Essentially, within the limit of the functional construct of the damping

function, we find that a single set of parameters is unable to capture simultaneously both short- and long-range field responses of a point charge. Note that in the computations above, we assigned the same damping factor to all the atomic centers in the molecule. Alternatively, assigning and calibrating individual damping factors assigned to each of the three functional groups, hydroxyls, methylenes, and methyls, do not change the overall conclusion.

But does a recalibration of damping factors improve transferability? To test this, we note first that the damping factor of water is modified to reproduce the high field response. If the new damping factor is now used to compute ion–water binding energies, then the computed values will be different from the reference energies in Table 2. We, therefore, recalibrate the vdW descriptors of the two ions again and use the same strategy as before but after setting the water damping factors to their new values. The recalibrated descriptors, which we refer to as the vdW-Pol[†] model, are (0.22, 2.8) for Na^+ and (0.49, 3.43) for K^+ .

Using these new ion descriptors, we recompute water → alcohol substitution energies for the reactions given by eq 2 in the gas phase (Figure 3). We find that without any other

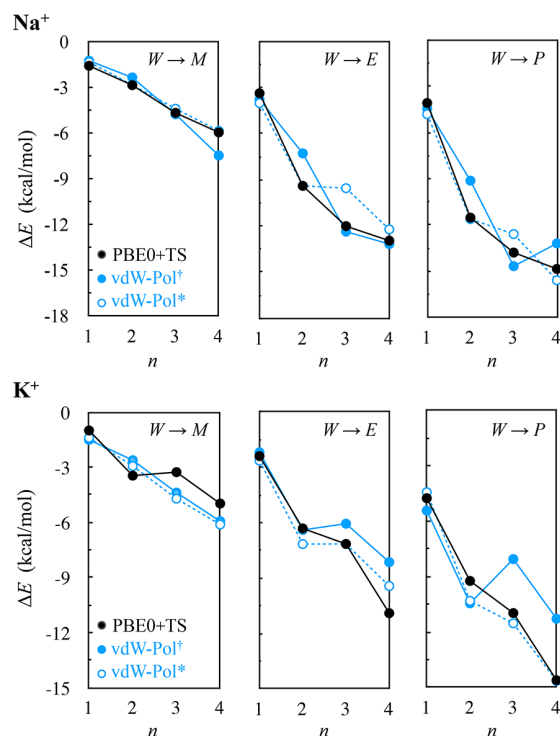


Figure 3. Comparison of substitution energies ΔE obtained from PBE0+TS and the recalibrated MM models, vdW-Pol[†] and vdW-Pol*. The ΔE values are estimated for the substitution reactions given by eq 2, and they are computed using geometries relaxed separately at the respective levels of theory. The letters *W*, *M*, *E*, and *P* refer, respectively, to water, methanol, ethanol, and propanol.

calibration, the RMS error in the water → alcohol substitution energies reduces from 2.3 to 1.4 kcal/mol. Additionally, when errors are computed for progressive addition of methyl groups or methylene bridges, that is for substitution reactions $W \rightarrow M$, $M \rightarrow E$, and $E \rightarrow P$, improvement is even more significant—errors reduce from 3.2 to 1.3 kcal/mol.

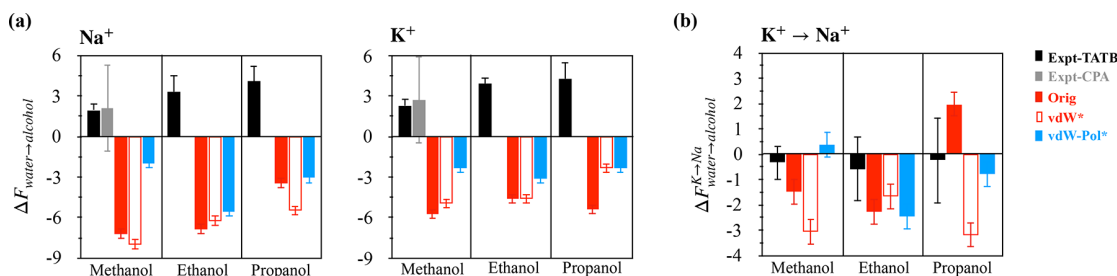


Figure 4. Predicted values of condensed phase transfer free energies (kcal/mol) from the Orig, vdW*, and vdW-Pol* models compared against experiment. (a) $\Delta F_{\text{water} \rightarrow \text{alcohol}}$ is the transfer free energy of an ion from water \rightarrow alcohol, and (b) $\Delta F_{\text{water} \rightarrow \text{alcohol}}^{K \rightarrow Na}$ is the difference between the transfer free energies of K^+ and Na^+ ions. While estimation of $\Delta F_{\text{water} \rightarrow \text{alcohol}}$ from experiments involves extrathermodynamic assumptions,⁶³ the estimation of $\Delta F_{\text{water} \rightarrow \text{alcohol}}^{K \rightarrow Na}$ does not.⁶⁵ Expt-TATB⁴⁰ and Expt-CPA⁴¹ refer to two different experimental estimates based on two different extrathermodynamic assumptions (see text for details). The errors bars on the experimental estimates indicate the upper limits of uncertainties reported in the respective studies.^{40,41}

We note, however, that since ligand–ligand dimerization energies of this MM model are in excellent agreement with MP2 theory, modifications to the ligand’s polarization model will introduce errors in ligand–ligand interactions. For the ion–ligand clusters examined here, changes in the ligand’s polarization model introduce only small errors (<1 kcal/mol), but these errors can be expected to grow with ligand packing in larger ion–ligand clusters or in the condensed phase. Consequently, the use of such a ligand-centric correction may not be a viable solution in the condensed phase.

Since damping in this polarization model is applied to only one of the two sites of an interaction pair, polarization correction can alternatively be introduced by modifying ion damping factors, without changing the underlying physical basis for the correction. Furthermore, since polarization contributions are effectively local, such an ion-centric correction will influence only local ion–ligand interactions. In fact, the polarization contribution to ion–water interaction energy is less than 6.1% beyond 5 Å. Consequently, such an ion-centric approach will fix high-field response but without adversely affecting the distant low-field response or ligand–ligand interactions. Figure 3 shows that similar improvements in gas phase transferability can be achieved by altering ion damping factors. We refer to this model as the vdW-Pol* model, and its RMS error in gas phase transferability is 0.9 kcal/mol. Note that we do not modify the damping factors of ions when ions interact with water – they are retained at their original value of $a = 0.39$. The best fit damping factors of Na^+ (K^+) ions are 0.04 (0.07), 0.08 (0.11), and 0.08 (0.39) for methanol, ethanol, and propanol, respectively. We make note here that such a numerical fix has been proposed previously for divalent cations,^{75,76} and our findings above essentially provide a physical basis for it.

Figure 4 shows the condensed phase water \rightarrow alcohol transfer free energies of ions. We find that the polarization corrections in the vdW-Pol* model do improve transferability predictions significantly. Now, if we assume that the discrepancies in $\Delta F_{\text{water} \rightarrow \text{methanol}}$ between simulations and experiments do not emanate from extrathermodynamic assumptions, then it would appear that simulations still overestimate solvation free energies in methanol. Since the vdW-Pol* model describes transferabilities in local interactions correctly (error <1 kcal/mol), and ligand–ligand interactions also match reference data, this residual errors in condensed phase must be due to errors in ion–ligand interactions beyond the ion’s first shell. To examine this, we compare distance-

dependent ion–ligand pair interaction energies obtained from the vdW-Pol* model against reference values from PBE0+TS (Figure 5a). In the short-range, we find that interaction

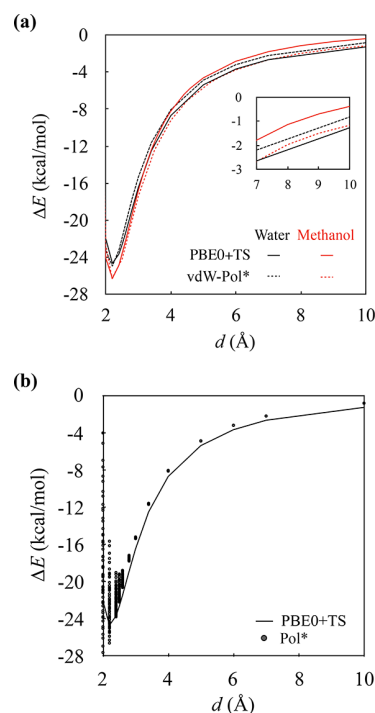


Figure 5. (a) Comparison of distance-dependent Na^+ –ligand pair interaction energies obtained from the vdW-Pol* model against reference values from PBE0+TS. (b) Effect of altering Na^+ vdW parameters on Na^+ –water interaction energies.

energies from the vdW-Pol* match those from PBE0+TS, but at distances greater than 4 Å, estimates from vdW-Pol* and PBE0+TS diverge. Importantly, in the long-range the vdW-Pol* model predicts ion–water interactions to be stronger than ion–methanol interactions (>1 kcal/mol), which is the exact opposite of what is expected from PBE0+TS. It is plausible that this error will add up in the condensed phase and make solvation of ions in alcohols more favorable than water. But what could be the source of the error in the long-range? We know that at long-range (>5 Å), the contributions from induced moments are small, as also noted above. We also do not find any combinations of vdW descriptors that correct this

long-range error (Figure 5b). Consequently, the long-range errors must be due to errors in the permanent multipoles of ligands. Note, however, that this finding does not suggest a breakdown of the multipole model. Instead, it recommends that static multipole assignments for ligands, which are always made under isolated conditions, perhaps need to be adjusted to enhance the accuracy of long-range electrostatics, and perhaps this should be included among the initial steps in the development of force field parameters.

CONCLUSIONS

The reliability of molecular mechanics simulations in describing biomolecular ion-driven processes depends on their ability to accurately model interactions of ions with both water and biomolecular functional groups. Here we address this issue in a representative polarizable MM model and focus on transferability of cationic interactions from water → alcohols. We specifically deal with two key issues. The first pertains to insufficiency in reference data, which is necessary for both testing and calibrating transferability. We overcome this by benchmarking a dispersion-corrected DFA against CCSD(T) and DMC. With respect to DMC, we circumvent key challenges for its application to charged clusters and demonstrate that it reproduces CCSD(T) energies within stochastic errors. The second issue pertains to large systematic errors in transferability, which, in fact, are also present in other popular MM models. We minimize transferability errors by strategically seeking and correcting error sources. We show first that the transferability errors are not due to errors in ion–water interactions – recalibration against the new reference data significantly improves ion–water interactions but does not improve water → alcohol transferability error. Recognizing that ligands in MM model are parametrized in the absence of ions and that electric fields near ions are much stronger than those used in parametrization, we systemically analyze field responses of ligands. We find that while the low field response of ligands matches reference values, the high field response, which ligands will exhibit in direct coordination with cations, is overdamped. Additionally, the extent of overdamping varies with ligand chemistry. Recalibrating polarization parameters against a high field response improves gas phase transferability significantly and without need for any other calibration. We then show that when both low- and high-field responses of ligand dipole polarization are described accurately, then transferability also improves significantly in the condensed phase. We attribute the residual mismatch between predictions and experiments in the condensed phase to errors in long distance interactions originating from parametrization of static multipoles of ligands and/or to the underlying extrathermodynamic assumptions involved in the estimation of single ion transfer free energies from experimental data. The corrections we introduce are directly transferable to functional groups in biomolecules. Overall, this work demonstrates a rational approach to boosting transferability of ionic interactions that will be applicable broadly to improving other polarizable and nonpolarizable models.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.8b01198.

Details of the convergence tests for quantum Monte Carlo calculations, and details regarding CCSD(T) calculations (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: svarma@usf.edu.

ORCID

Vered Wineman-Fisher: 0000-0002-5124-5900

Funding

The authors acknowledge the use of over 7 million CPU hours from Research Computing at USF and UL. Funding for this work was provided by NIH grant number R01GM118697.

Notes

The authors declare no competing financial interest.

REFERENCES

- (1) Alberts, B.; Johnson, A. D.; Lewis, J.; Morgan, D.; Raff, M.; Robert, K.; Walter, P. *Molecular Biology of the Cell*; W. W. Norton & Company: 2014; p 1464.
- (2) Hill, B. *Ion Channels of Excitable Membranes*; Oxford University Press: 2001; 814pp.
- (3) Karplus, M.; McCammon, J. A. Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* **2002**, *9*, 646–652.
- (4) Dror, R. O.; Dirks, R. M.; Grossman, J.; Xu, H.; Shaw, D. E. Biomolecular Simulation: A Computational Microscope for Molecular Biology. *Annu. Rev. Biophys.* **2012**, *41*, 429–452.
- (5) Beauchamp, K. A.; Lin, Y.-S.; Das, R.; Pande, V. S. Are Protein Force Fields Getting Better? A Systematic Benchmark on 524 Diverse NMR Measurements. *J. Chem. Theory Comput.* **2012**, *8*, 1409–1414.
- (6) Lopes, P. E.; Guvench, O.; MacKerell, A. D. Current Status of Protein Force Fields for Molecular Dynamics. *Methods Mol. Biol. (N. Y., NY, U. S.)* **2015**, *1215*, 47–71.
- (7) Yue, S.; Pengyu, R.; Michael, S.; JeanPhilip, P. *Reviews in Computational Chemistry*; Wiley-Blackwell: 2015; Vol. 28, pp 51–86, DOI: 10.1002/9781118889886.ch2.
- (8) Senftle, T. P.; Hong, S.; Islam, M. M.; Kylasa, S. B.; Zheng, Y.; Shin, Y. K.; Junkermeier, C.; Engel-Herbert, R.; Janik, M. J.; Aktulga, H. M.; Verstraelen, T.; Grama, A.; van Duin, A. C. T. The ReaxFF reactive force-field: development, applications and future directions. *Npj Computational Materials* **2016**, *2*, 15011–15023.
- (9) Lybrand, T. P.; Kollman, P. A. Waterwater and waterion potential functions including terms for many body effects. *J. Chem. Phys.* **1985**, *83*, 2923–2933.
- (10) Perera, L.; Berkowitz, M. L. Many body effects in molecular dynamics simulations of $\text{Na}^+(\text{H}_2\text{O})_n$ and $\text{Cl}^-(\text{H}_2\text{O})_n$ clusters. *J. Chem. Phys.* **1991**, *95*, 1954–1963.
- (11) Grossfield, A.; Ren, P.; Ponder, J. W. Ion Solvation Thermodynamics from Simulation with a Polarizable Force Field. *J. Am. Chem. Soc.* **2003**, *125*, 15671–15682.
- (12) Carrillo-Tripp, M.; Saint-Martin, H.; Ortega-Blake, I. A comparative study of the hydration of Na^+ and K^+ with refined polarizable model potentials. *J. Chem. Phys.* **2003**, *118*, 7062–7073.
- (13) Spångberg, D.; Hermansson, K. Many-body potentials for aqueous Li^+ , Na^+ , Mg^{2+} , and Al^{3+} : Comparison of effective three-body potentials and polarizable models. *J. Chem. Phys.* **2004**, *120*, 4829–4843.
- (14) Lamoureux, G.; Roux, B. Absolute Hydration Free Energy Scale for Alkali and Halide Ions Established from Simulations with a Polarizable Force Field. *J. Phys. Chem. B* **2006**, *110*, 3308–3322.
- (15) Whitfield, T. W.; Varma, S.; Harder, E.; Lamoureux, G.; Rempe, S. B.; Roux, B. Theoretical Study of Aqueous Solvation of K^+ Comparing ab Initio, Polarizable, and Fixed-Charge Models. *J. Chem. Theory Comput.* **2007**, *3*, 2068–2082.
- (16) Lee Warren, G.; Patel, S. Hydration free energies of monovalent ions in transferable intermolecular potential four point

fluctuating charge water: An assessment of simulation methodology and force field performance and transferability. *J. Chem. Phys.* **2007**, *127*, No. 064509.

(17) Joung, I. S.; Cheatham, T. E. Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *J. Phys. Chem. B* **2008**, *112*, 9020–9041.

(18) Baker, C. M.; Lopes, P. E. M.; Zhu, X.; Roux, B.; MacKerell, A. D. Accurate Calculation of Hydration Free Energies using Pair-Specific Lennard-Jones Parameters in the CHARMM Drude Polarizable Force Field. *J. Chem. Theory Comput.* **2010**, *6*, 1181–1198.

(19) Luo, Y.; Roux, B. Simulation of Osmotic Pressure in Concentrated Aqueous Salt Solutions. *J. Phys. Chem. Lett.* **2010**, *1*, 183–189.

(20) Varma, S.; Rogers, D. M.; Pratt, L. R.; Rempe, S. B. Design principles for K⁺ selectivity in membrane transport. *J. Gen. Physiol.* **2011**, *138*, 279–279.

(21) Yoo, J.; Aksimentiev, A. Improved Parametrization of Li⁺, Na⁺, K⁺, and Mg²⁺ Ions for All-Atom Molecular Dynamics Simulations of Nucleic Acid Systems. *J. Phys. Chem. Lett.* **2012**, *3*, 45–50.

(22) Fyta, M.; Netz, R. R. Ionic force field optimization based on single-ion and ion-pair solvation properties: Going beyond standard mixing rules. *J. Chem. Phys.* **2012**, *136*, No. 124103.

(23) Mamatkulov, S.; Fyta, M.; Netz, R. R. Force fields for divalent cations based on single-ion and ion-pair properties. *J. Chem. Phys.* **2013**, *138*, No. 024505.

(24) Savelyev, A.; MacKerell, A. D. Balancing the Interactions of Ions, Water, and DNA in the Drude Polarizable Force Field. *J. Phys. Chem. B* **2014**, *118*, 6742–6757.

(25) Li, H.; Ngo, V.; Da Silva, M. C.; Salahub, D. R.; Callahan, K.; Roux, B.; Noskov, S. Y. Representation of Ion-Protein Interactions Using the Drude Polarizable Force-Field. *J. Phys. Chem. B* **2015**, *119*, 9401–9416.

(26) Savelyev, A.; MacKerell, A. D. Competition among Li⁺, Na⁺, K⁺, and Rb⁺ Monovalent Ions for DNA in Molecular Dynamics Simulations Using the Additive CHARMM36 and Drude Polarizable Force Fields. *J. Phys. Chem. B* **2015**, *119*, 4428–4440.

(27) Jing, Z.; Qi, R.; Liu, C.; Ren, P. Study of interactions between metal ions and protein model compounds by energy decomposition analyses and the AMOEBA force field. *J. Chem. Phys.* **2017**, *147*, No. 161733.

(28) Warren, G. L.; Patel, S. Comparison of the Solvation Structure of Polarizable and Nonpolarizable Ions in Bulk Water and Near the Aqueous Liquid-Vapor Interface. *J. Phys. Chem. C* **2008**, *112*, 7455–7467.

(29) Varma, S.; Rempe, S. B. Multibody Effects in Ion Binding and Selectivity. *Biophys. J.* **2010**, *99*, 3394–3401.

(30) Rogers, D. M.; Beck, T. L. Quasichemical and structural analysis of polarizable anion hydration. *J. Chem. Phys.* **2010**, *132*, No. 014505.

(31) Rossi, M.; Tkatchenko, A.; Rempe, S. B.; Varma, S. Role of methyl-induced polarization in ion binding. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 12978–12983.

(32) Baldauf, C.; Kevin, P.; Stephan, W.; von Helden, G.; Beate, K.; Volker, B.; Matthias, S. How Cations Change Peptide Structure. *Chem. - Eur. J.* **2013**, *19*, 11224–11234.

(33) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158–6170.

(34) Tkatchenko, A.; Scheffler, M. Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data. *Phys. Rev. Lett.* **2009**, *102*, No. 073005.

(35) Dubecký, M.; Jurečka, P.; Derian, R.; Hobza, P.; Otyepka, M.; Mitas, L. Quantum Monte Carlo methods describe noncovalent interactions with subchemical accuracy. *J. Chem. Theory Comput.* **2013**, *9*, 4287–4292.

(36) Al-Hamdani, Y. S.; Rossi, M.; Alfè, D.; Tsatsoulis, T.; Ramberger, B.; Brandenburg, J. G.; Zen, A.; Kresse, G.; Grüneis, A.; Tkatchenko, A.; Michaelides, A. Properties of the water to boron

nitride interaction: From zero to two dimensions with benchmark accuracy. *J. Chem. Phys.* **2017**, *147*, No. 044710.

(37) Al-Hamdani, Y. S.; Alfè, D.; Von Lilienfeld, O. A.; Michaelides, A. Water on BN doped benzene: A hard test for exchange-correlation functionals and the impact of exact exchange on weak binding. *J. Chem. Phys.* **2014**, *141*, No. 18C530.

(38) Dubecký, M.; Derian, R.; Jurečka, P.; Mitas, L.; Hobza, P.; Otyepka, M. Quantum Monte Carlo for noncovalent interactions: an efficient protocol attaining benchmark accuracy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 20915–20923.

(39) Dubecký, M.; Mitas, L.; Jurečka, P. Noncovalent Interactions by Quantum Monte Carlo. *Chem. Rev.* **2016**, *116*, 5188–5215.

(40) Marcus, Y. Thermodynamic functions of transfer of single ions from water to nonaqueous and mixed solvents: Part I - Gibbs free energies of transfer to nonaqueous solvents. *Pure Appl. Chem.* **1983**, *55*, 977–1021.

(41) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. Single-Ion Solvation Free Energies and the Normal Hydrogen Electrode Potential in Methanol, Acetonitrile, and Dimethyl Sulfoxide. *J. Phys. Chem. B* **2007**, *111*, 408–422 PMID: 17214493 .

(42) Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J. W.; Ren, P. Polarizable Atomic Multipole-Based AMOEBA Force Field for Proteins. *J. Chem. Theory Comput.* **2013**, *9*, 4046–4063.

(43) Glusker, J. P. In *Metalloproteins: Structural Aspects*; Anfinsen, C., Edsall, J. T., Richards, F. M., Eisenberg, D. S., Eds.; *Advances in Protein Chemistry*; Academic Press: 1991; Vol. 42, pp 1–76, DOI: 10.1016/S0065-3233(08)60534-3.

(44) Page, M. J.; Di Cera, E. Role of Na⁺ and K⁺ in Enzyme Function. *Physiol. Rev.* **2006**, *86*, 1049–1092.

(45) Zen, A.; Brandenburg, J. G.; Klimeš, J.; Tkatchenko, A.; Alfè, D.; Michaelides, A. Fast and accurate quantum Monte Carlo for molecular crystals. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 1724–1729.

(46) Benali, A.; Shulenburger, L.; Romero, N. A.; Kim, J.; Von Lilienfeld, O. A. Application of diffusion Monte Carlo to materials dominated by van der Waals interactions. *J. Chem. Theory Comput.* **2014**, *10*, 3417–3422.

(47) Grossman, J. C. Benchmark quantum monte carlo calculations. *J. Chem. Phys.* **2002**, *117*, 1434–1440.

(48) Riera, M.; Götz, A. W.; Paesani, F. The *i*-TTM model for ab initio-based ion–water interaction potentials. II. Alkali metal ion–water potential energy functions. *Phys. Chem. Chem. Phys.* **2016**, *18*, 30334–30343.

(49) Neese, F. The ORCA program system. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2012**, *2*, 73–78.

(50) Weigend, F. Accurate Coulomb-fitting basis sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.

(51) Trail, J.; Needs, R. Smooth relativistic Hartree–Fock pseudopotentials for H to Ba and Lu to Hg. *J. Chem. Phys.* **2005**, *122*, 174109.

(52) Needs, R.; Towler, M.; Drummond, N.; Ríos, P. L. Continuum variational and diffusion quantum Monte Carlo calculations. *J. Phys.: Condens. Matter* **2010**, *22*, 023201.

(53) Blum, V.; Gehrke, R.; Hanke, F.; Havu, P.; Havu, V.; Ren, X.; Reuter, K.; Scheffler, M. Ab initio molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Commun.* **2009**, *180*, 2175–2196.

(54) Varma, S.; Rempe, S. B. Structural transitions in ion coordination driven by changes in competition for ligand binding. *J. Am. Chem. Soc.* **2008**, *130*, 15405–15419.

(55) Bennett, C. H. Efficient estimation of free energy differences from Monte Carlo data. *J. Comput. Phys.* **1976**, *22*, 245–268.

(56) Varma, S.; Rempe, S. B. Coordination numbers of alkali metal ions in aqueous solutions. *Biophys. Chem.* **2006**, *124*, 192–199 Ion Hydration Special Issue .

(57) Machado, J. R.; Streett, W. B. Equation of state and thermodynamic properties of liquid methanol from 298 to 489 K and pressures to 1040 bar. *J. Chem. Eng. Data* **1983**, *28*, 218–223.

- (58) Khattab, I. S.; Bandarkar, F.; Fakhree, M. A. A.; Jouyban, A. Density, viscosity, and surface tension of water+ ethanol mixtures from 293 to 323K. *Korean J. Chem. Eng.* **2012**, *29*, 812–817.
- (59) Wandschneider, A.; Lehmann, J. K.; Heintz, A. Surface tension and density of pure ionic liquids and some binary mixtures with 1-propanol and 1-butanol. *J. Chem. Eng. Data* **2008**, *53*, 596–599.
- (60) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, No. 014101.
- (61) Chow, K.-H.; Ferguson, D. M. Isothermal-isobaric molecular dynamics simulations with Monte Carlo volume sampling. *Comput. Phys. Commun.* **1995**, *91*, 283–289.
- (62) Åqvist, J.; Wennerström, P.; Nervall, M.; Bjelic, S.; Brandsdal, B. O. Molecular dynamics simulations of water and biomolecules with a Monte Carlo constant pressure algorithm. *Chem. Phys. Lett.* **2004**, *384*, 288–294.
- (63) Marcus, Y. Single ion Gibbs free energies of transfer from water to organic and mixed solvents. *Rev. Anal. Chem.* **1980**, *5*, 53137.
- (64) Pliego, J. R.; Miguel, E. L. M. Absolute Single-Ion Solvation Free Energy Scale in Methanol Determined by the Lithium Cluster-Continuum Approach. *J. Phys. Chem. B* **2013**, *117*, 5129–5135 PMID: 23570440 .
- (65) Asthagiri, D.; Dixit, P.; Merchant, S.; Paulaitis, M.; Pratt, L.; Rempe, S.; Varma, S. Ion selectivity from local configurations of ligands in solutions and ion channels. *Chem. Phys. Lett.* **2010**, *485*, 1–7.
- (66) Tissandier, M. D.; Cowen, K. A.; Feng, W. Y.; Gundlach, E.; Cohen, M. H.; Earhart, A. D.; Coe, J. V.; Tuttle, T. R. The Protons Absolute Aqueous Enthalpy and Gibbs Free Energy of Solvation from Cluster-Ion Solvation Data. *J. Phys. Chem. A* **1998**, *102*, 7787–7794.
- (67) Vlcek, L.; Chialvo, A. A.; Simonson, J. M. Correspondence between Cluster-Ion and Bulk Solution Thermodynamic Properties: On the Validity of the Cluster-Pair-Based Approximation. *J. Phys. Chem. A* **2013**, *117*, 11328–11338 PMID: 24093538 .
- (68) Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618–622.
- (69) Ren, P.; Wu, C.; Ponder, J. W. Polarizable Atomic Multipole-Based Molecular Mechanics for Organic Molecules. *J. Chem. Theory Comput.* **2011**, *7*, 3143–3161.
- (70) Rackers, J. A.; Wang, Q.; Liu, C.; Piquemal, J.-P.; Ren, P.; Ponder, J. W. An optimized charge penetration model for use with the AMOEBA force field. *Phys. Chem. Chem. Phys.* **2017**, *19*, 276–291.
- (71) Pliego, J. R.; Riveros, J. M. On the calculation of the absolute solvation free energy of ionic species: application of the extrapolation method to the hydroxide ion in aqueous solution. *J. Phys. Chem. B* **2000**, *104*, 5155–5160.
- (72) Thole, B. T. Molecular polarizabilities calculated with a modified dipole interaction. *Chem. Phys.* **1981**, *59*, 341–350.
- (73) Ren, P.; Ponder, J. W. Polarizable atomic multipole water model for molecular mechanics simulation. *J. Phys. Chem. B* **2003**, *107*, 5933–5947.
- (74) Ren, P.; Ponder, J. W. Temperature and pressure dependence of the AMOEBA water model. *J. Phys. Chem. B* **2004**, *108*, 13427–13437.
- (75) Piquemal, J.; Perera, L.; Cisneros, G.; Ren, P.; Pedersen, L.; Darden, T. Towards accurate solvation dynamics of divalent cations in water using the polarizable amoeba force field: From energetics to structure. *J. Chem. Phys.* **2006**, *125*, 054511.
- (76) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S. Current status of the AMOEBA polarizable force field. *J. Phys. Chem. B* **2010**, *114*, 2549–2564.

Supplementary Material: Ion-hydroxyl interactions: From high-level quantum benchmarks to transferable polarizable force fields

Vered Wineman-Fisher,[†] Yasmine Al-Hamdani,[‡] Iqbal Addou,[†] Alexandre
Tkatchenko,[‡] and Sameer Varma^{*,†}

[†]*Department of Cell Biology, Microbiology and Molecular Biology, University of South
Florida, Tampa, FL 33620*

[‡]*Physics and Materials Science Research Unit, University of Luxembourg, 162a avenue de
la Financier, Luxembourg City, L-1511, Luxembourg*

E-mail: svarma@usf.edu

Abstract

This material contains additional details regarding the reference calculations of the main article. Herein, details of the convergence tests for the diffusion Monte Carlo calculations and additional coupled cluster calculations can be found.

Diffusion Monte Carlo Convergence Tests

Given the various approximations that one has to take into account to compute diffusion Monte Carlo (DMC) energies, we report here the binding energies of Na^+ and K^+ with water. The important aspects to consider are the orbitals used to construct the guiding trial wavefunctions for DMC; the use of pseudopotentials; and the time-step. In Table 1, we compare the results of various combinations to CCSD(T)-F12 results from Riera *et al.*¹

Table 1: Binding energies (E^b) of Na^+ and K^+ with a water molecule. For comparison we show the CCSD(T)-F12 binding energies computed by Riera *et al.* The information for each DMC method corresponds to the type of orbitals used (*e.g.* LDA), whether Trail and Needs pseudopotentials were used (TN PP), and finally the time-step used in atomic units (a.u.). All energies and stochastic errors are given in kcal/mol. We have highlighted in bold the chosen setup for the results reported in the main study.

Methods	$E_{\text{K}^+/\text{H}_2\text{O}}^b$	stochastic error	$E_{\text{Na}^+/\text{H}_2\text{O}}^b$	stochastic error
CCSD(T)-F12 ¹	-17.97		-24.14	
DMC: LDA, TN PP, 0.005	-20.90	0.09	-24.68	0.07
DMC: LDA, TN PP, 0.002	-20.83	0.16	-24.69	0.14
DMC: PBE0, TN PP, 0.002	-20.84	0.15	-24.95	0.17
DMC: HF, TN PP, 0.002	-20.84	0.12	-24.99	0.13
DMC: LDA, 0.01	-21.55	0.12		
DMC: LDA, 0.005	-18.40	0.22		
DMC: LDA, 0.003	-18.56	0.24		
DMC: LDA, 0.001	-17.93	0.29	-24.467	0.229
DMC: B3LYP, 0.005	-19.07	0.13		
DMC: B3LYP, 0.001	-18.10	0.27		

As can be seen from Table 1, the DMC binding energies for Na^+ /water are in much closer agreement with CCSD(T)-F12 than K^+ /water. The ~ 2 kcal/mol discrepancy between DMC and CCSD(T)-F12 is alleviated by the use all-electron DMC, which is far more

computationally expensive and also requires a smaller time-step. Indeed, Table 1 shows that a time-step of 0.001 a.u. is needed in all-electron calculations to find agreement with CCSD(T)-F12, as well as between DMC binding energies with different orbitals. In particular, using a 0.001 a.u. time-step B3LYP and LDA orbitals yield DMC binding energies that are in agreement within the stochastic error bars. In addition, the use of HF, PBE0 or LDA orbitals, in conjunction with Trail and Needs pseudopotentials, led to the same DMC binding energies, indicating that the fixed-node approximation is not noticeably detrimental to the accuracy of the DMC binding energies.

Methanol-Cation Binding Energies

Due to the scarcity of reference information for methanol clusters around Na^+ and K^+ cations, we computed CCSD(T) interaction energies for comparison with DMC. Using the ORCA v4.0.1 package and the def2/J basis set for K^+ /methanol, we obtain a binding energy of -19.1 kcal/mol. In comparison, using a basis set of def2-QZVP for K, cc-pVTZ for CH_3 atoms, and aug-cc-pVTZ for OH atoms, we obtain -19.0 kcal/mol for the binding energy. The agreement amongst these results indicate reasonable convergence. In addition, the all-electron DMC binding energy for this system is -19.0 ± 0.3 kcal/mol. Thus, excellent agreement is found across these methods. As cross comparison, we found the CCSD(T) binding energy of Na^+ /methanol to also be in good agreement between all-electron DMC (-26.5 ± 0.3 kcal/mol) and CCSD(T) with def2/J basis set (-26.8 kcal/mol).

References

- (1) Riera, M.; Götz, A. W.; Paesani, F. *Physical Chemistry Chemical Physics* **2016**, *18*, 30334–30343.