Developing Ion Parameters and Investigating Ion Specific Effects in Biological Systems

im Fachbereich Physik der Freien Universität Berlin
eingereichte Dissertation

zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.)

vorgelegt von
Sadra Kashef Ol Gheta

Berlin, 2017

Erster Gutachter :	Prof. Dr. Roland Netz
Zweiter Gutachter:	Prof. Dr. Petra Imhof
Tag der Disputation:	08. September 2017

Abstract

Biomolecular processes often involve ions in aqueous solutions. Accurate force-field parameters are thus required for these ions, if we are to use molecular simulations to investigate the roles they play. In the course of this PhD study, we addressed multiple outstanding issues concerning the development of force fields for ions in aqueous solution. We optimized parameters for monoatomic ions for the TIP5P water model, and show that the activity derivative is often not sufficient to optimize ion-ion interactions, since it reaches a plateau near the experimental target value. We propose using coordination numbers at high salt concentrations, as another target for parameterization.

Furthermore, we optimized a set of Lennard-Jones parameters for classical, all-atom models of acetate and several alkylated and non-alkylated forms of sulfate, sulfonate and phosphate ions. The parameters were optimized to reproduce their interactions with water and with the physiologically relevant sodium, ammonium and methylammonium cations. The parameters are suitable for the TIP3P water model and were developed based on experimental information, with ab initio, gas phase calculations being used for the cases where experimental information is missing. The parameterization scheme is different from previous approaches because it combines existing experimental data and quantum mechanical calculations, to find parameters when the experimental data are sparse for specific ions.

Finally, we used our optimized parameters to investigate ion specific effects of alkyl-anions towards $CH_3NH_3^+$, using known experimental data on polyanionic dendritic polymers (dPGs) interactions with L-selectin as reference. We show that several factors, such as presence of Na^+, number of free anionic sites and protonation state, are involved in these interactions, and these factors can potentially influence the dPG-selectin affinity trend. Although we could not unequivocally identify the origin of the order magnitude of differences reported experimentally, we show that some experimental trends can be captured if the mentioned factors are considered.

Zusammenfassung

Da biomolekulare Prozesse in wässrigen Lösungen in der Regel in der Anwesenheit von Ionen stattfinden, wird dafür in Molekulardynamik-Simulationen die Anwendung entsprechend präziser Kraftfeldparameter, die diese berücksichtigen, notwendig. Im Rahmen der vorliegenden Doktorarbeit sind wir daher noch bestehende Problemstellungen bezüglich der Entwicklung solcher Kraftfelder für Ionen in wässriger Lösung angegangen. Wir haben die Parameter für einatomige Ionen für das TIP5P-Wasser-Modell mit Hinblick auf die freie Solvatisierungsenergie und Lösungsaktivität als experimentelle Zielwerte optimiert. Wir zeigen, dass die Ableitung der Aktivität oftmals nicht für die Optimierung der Ion-Ion-Wechselwirkungen ausreicht, da diese nahe den Zielwerten ein Plateau erreicht. Für die Optimierung der Ionen-Parameter beabsichtigen wir die Koordinationszahlen bei hohen Salzkonzentrationen als anderen Zielwert zu benutzen. Darüber hinaus haben wir eine Anzahl von Lennard-Jones-Parametern für klassische, alle Atome umfassende Modelle für Azetat-, mehrere alkylierte und nicht-alkylierte Arten von Sulfat-, Sulfonat- und Phosphationen optimiert. Die Parameter wurden auf die Reproduzierbarkeit derer Wechselwirkungen mit Wasser und physiologisch relevanten Natrium-, Ammonium- und Methylammoniumkationen optimiert. Die Parameter wurden auf Grundlage experimenteller Informationen entwickelt und sind für das TIP3P Wassermodel geeignet. Hierbei wurden ab initio Gasphasenberechnungen genutzt in den Fällen, in denen keine experimentellen Daten zur Verfügung standen. Das Parametrisierungschema unterscheidet sich von vorangegangenen Ansätzen, jedoch kombiniert es bestehende experimentelle Daten mit quantenmechanischen Berechnungen, um Parameter zu bestimmen, falls experimentelle Daten für bestimmte Ionen rar sind. Wir haben abschließend unsere Parameter für die Untersuchung ionenspezifischer Effekte des Alkyl-Anions gegenüber $CH_3NH_3^+$benutzt unter Anwendung von bekannten experimentellen Messdaten zur Wechselwirkungen zwischen polyanionischen dendritischen Polymeren (dPGs) mit L-Selectin als Referenz. Wir zeigen, dass mehrere Faktoren, wie die Anwesenheit von Na^+, die Anzahl freier anionischer Stellen, Protonierungstellen, usw. in diese Wechselwirkungen involviert sind und dass diese Faktoren möglicherweise die Tendenz der dPG-Selectin-Affinität beeinflussen. Auch wenn wir die Ursache der Unterschiede um mehrere Größenordnungen, die experimentell bestimmt wurden, nicht zweifelsfrei feststellen konnten, können wir zeigen, dass einige dieser experimentell bestimmten Tendenzen unter Berücksichtigung der genannten Faktoren erfasst werden können.

Contents

List of Figures

Chapter 1

Introduction

1.1 Ions in Biological Systems

Monoatomic and polyatomic ions are key players in many biological processes, such as DNA folding [1], blood coagulation and anti-coagulation [2, 3], protein stability [4] and protein crystallization [5]. Furthermore, the structure and function of biomolecules such as proteins, nucleic acids and lipid bilayer membranes are regulated or affected by interactions with ions [4, 6, 7], and chemical reactions in aerosol droplets critically depend on their salt concentration [8].

Ions, inorganic and organic, show specific trends in their solution properties as well as in their interaction with biological molecules[9]. The Hofmeister series ranks the ions in the order of their ability to precipitate the proteins but similar series have been observed when ranking ions by their impact on other properties. The effect of anions and cations on the solubility of the proteins are usually ordered as the following: [10, 11, 12]

$$CO_3^{2-} < SO_4^{2-} < H_2PO_4^- < F^- < Cl^- < Br^- < I^- < ClO_4^- < SCN^-$$

and

$$N(CH_3)_4^+ < NH_4^+ < Cs^+ < Rb^+ < K^+ < Na^+ < Li^+ < Mg^{2+} < Ca^{2+}$$

Monoatomic and polyatomic ions have been vastly investigated for their specific effects. Although much progress has been made in understanding the role of ions in solution, the origin of many specific ion effects remains unknown.

1.1.1 Specific ion effects involving monoatomic ions

Sodium, potassium, magnesium and calcium along with chloride are the most abundant ions in the human body and play an important role in many physiological processes[13]. Each of the cations above can form ion-pairs with chloride or bind with negative sites on proteins[4]. These interactions are specific for each ion-protein pair and depend on other conditions such as pH and temperature, resulting in selective protein salt-out or salt-in behaviour[4].

The interactions of ions with the cell membrane and the lipid bilayers are also of great importance. Sodium is abundant in extracellular media whereas potassium is present inside the cells with similar concentration as of the external sodium[14].

Other alkali and alkali earth metals are present in much lower concentrations, however their interactions with the proteins cannot be neglected for medical and toxicological purpose[15, 16, 17, 18]. Previous studies suggest that the binding contant of the alkali metals with the negatively charged membranes is consistent with the Hofmeister series[19, 14].

The alkali and alkali earth metal ions as well as chloride, interact selectively with ion channel proteins. These ion channels are in charge of selectively permitting ions in and out of the cells. Several ion channels have been identified and categorized based on their ion selectivity; chloride channels(CLC family)[20, 21], sodium channels[22], potassium channels[23, 24] calcium channels [25], etc. The proteins in these channels have been thoroughly studied over the past years however the mechanism through which these channels distinguish between the ions has remained a hot topic in the field of structural biology.

1.1.2 Specific ion effects involving polytomic ions

Polyatomic ions such as phosphates, sulfates and carboxylates are abundant in physiological context and play important roles in a number of processes, such as cell adhesion, protein-DNA interaction or energy transduction. Many of the molecular scale details regarding the interaction of polyatomic ions with biological polymers are not understood, in part due to the lack of good quality molecular models for these ions. An important examples of specific ion effects involving polyatomic ions is the interaction of polyanions with selectin proteins. Selectins are well known for their role in the adhesion of leukocytes and platelets to the endothelium that takes place, e.g., during inflammation. Adhesion to the endothelium must occur before paracellular and transcellular transmigration of the leukocytes towards inflamed tissue[26]. Recent studies also highlight the key role of selectins on cancer metastasis. Metastasis is the cause of 90% of fatalities associated with cancer and happens when migrated tumor cells attach to and penetrate healthy organs[27]. The key step in cancer metastasis is the selectin-mediated adhesion of invading cells to the endothelium of other organs. Because of the important biological role played by selectins, much effort has been put into characterizing them, into finding artificial ligands that effectively compete with the natural ones, and into characterizing both types of ligands to understand their molecular scale interactions with selectins. Given the abundance of sulfate in several natural ligands of selectins, several studies have further investigated the interactions between sulfated ligands and positive amino acids directly in selectins[28, 29, 30]. The Haag group investigated the potential of dendrimeric polyglycerol (dPG) polymers functionalized with various anionic functional groups as possible inhibitors of L-selectin interactions with natural ligands[31, 28] For dendrimers of similar diameter and with similar density of functionalized groups, they found that the affinity of dendrimers for selectins depends strongly on the nature of the anionic group, increasing in the order carboxylate $<$ phosphate $<$ phosphonate $\sim$ sulfonate $<$ bisphosphonate $\lll$ sulfate, with the affinity between sulfate and selectin being two or more orders of magnitude larger than for the other anions[31]. This anionic series indicates that interactions between L-selectin and these anions cannot be understood in terms of

simple considerations based on electrostatic interactions or the acidity of the anionic groups. Instead, this series suggests that complex and competitive microscopic intermolecular and intramolecular interactions (e.g. hydrogen bonding) between the binding sites of selectin and dPGs might play a role. Understanding the molecular scale origin of this series requires that we first understand the intrinsic interactions between individual anionic groups and basic amino acids. Once these are understood, we should then investigate supramolecular effects arising from the presence of multiple anionic sites on the dendrimer and multiple cationic sites on the protein.

1.2 Using molecular simulations to investigate specific ion effects

The molecular mechanisms by which ions act in biological systems are currently incompletely understood, in part because inferring molecular scale details from the signals detected in experiment cannot be unambiguously done. In contrast, molecular scale details are directly accessible from simulations using atomistic models, making the development of such models of critical importance. Because proteins and other biomolecules are typically large (50-3000 kDa), only simulations using classical, non-polarizable force fields give access to the long length and time scales relevant for many processes taking place in bio-systems. To be useful, force fields must strike the correct balance between the various intermolecular interactions: protein-water, protein-protein, water-water, water-ion, ion-counterion and ion-protein interactions; in this last case, the main contribution would come from the interaction between the ions in solution and the charged amino acids of the protein. Moreover, for many applications it is indispensable that the force field captures the intrinsic directionality that characterizes short-range interactions in hydrogen-bonded systems; in contrast, explicit consideration of polarizability is often unnecessary and fixed-charge models perform well under many conditions [32, 33, 34, 35, 36, 37, 38]. Short-range directionality naturally arises in all-atom models, where each atom is explicitly represented, and many all-atom, fixed-charge force fields have been developed for biomolecules, water and ions [37]. Despite several decades of development of all-atom force fields, optimizing the balance between all interactions remains an on-going effort which, because of its magnitude, proceeds stepwise and iteratively. Atomistic, fixed-charged force fields for proteins in water are now quite successful at predicting binding affinities[39] and the folded structure of small proteins [40, 41, 42]. These force fields were parameterized against reliable experimental observables that reflect the balance between ion-water, protein-water and water-water interactions (*e.g.*, free energies of hydration) [36, 43]. More recently, force fields for aqueous solutions of alkali, alkali-earth and halide ions were developed that also show reasonable anion-cation interactions, and are thus able to reproduce experimental osmotic pressures or solution activities at salt concentrations below 1 M [44, 45, 46, 47, 48]. Simulations based on these force fields have already offered unprecedented insight into the molecular scale origin of ion-specific effects in biological systems [49]. In contrast to the relatively advanced state of development of force fields for monoatomic ions, force fields for polyatomic

ions in biomolecular systems are currently still largely untested or non-existent.

1.2.1 Monoatomic ions in TIP5P water

Of the various classical, all-atom, water models with functional forms that are in principle compatible with the most popular force fields for biomolecules [37], the TIP5P model is of particular interest because it reproduces a number of properties of water (*e.g.*, the liquid density, temperature of maximum density and diffusivity as a function of temperature and pressure) substantially better than the TIP3P model used to develop the widely used AMBER and CHARMM force fields for proteins [36, 35]. Recent studies show that the TIP5P water model [50] is of particular interest for simulations using the GLYCAM06h force field for sugars: this water model results in particularly good agreement with experimental data such as diffusion coefficients, free energies of hydration and osmotic pressure for solutions of disaccharides [38, 51]. However, widespread usage of the TIP5P water model in simulations of biomolecular systems is currently hindered by the absence of a force field for inorganic ions that has been optimized for this water model: to the best of our knowledge, the closest parameter set that has been presented was developed for the TIP5P/Ew water model and is limited to halide ions and sodium [52].

Replacing the original water model associated with a biomolecular force fied by TIP5P, or by any other water model offering better descriptions of the structure and the dynamics of water [37], is sometimes acceptable without further modifications [53], using careful consideration. However, to obtain optimal results, further tuning of the protein and ion force fields [54, 55, 56, 44] is indispensable. In this doctoral project we have addressed this issue: we have developed parameters for monoatomic specifically optimized for the TIP5P water.

1.2.1.1 Difficulties optimizing Lennard-Jones parameters for monoatomic anions

An adequate force field must correctly reflect the balance between water-water, water-solute and solute-solute interactions.

Many force field were originally developed to reproduce properties that mostly reflect self-interactions and interactions with solvent molecules, such as density and enthalpy or free energy of solvation, as well as structural properties of biomolecules. However, matching single molecule properties does not reflect the balance of solute and solvent interactions, which may require additional tuning [57, 58, 51].

For this work, we chose to optimize parameters based on the solvation free energy and, for chosen salts, the solution activity derivative, because these properties can be accurately determined in both experiment and simulation. Using solvation free energies for parameterization has in the past been problematic [59], because of the large uncertainty associated with experimental values of single-ion solvation free energies, and the dramatic dependency of the calculated single ion solvation free energies on the simulation methodology, mostly associated with the treatment of electrostatic interactions in MD simulations. However, these issues have by now been mostly resolved,

as summarized in the methods section. Using activity derivatives as the target data to develop parameters has also been problematic because simulations are typically run in closed systems, whereas the theory that allows the calculation for activity derivatives from the Kirkwood-Buff integrals is developed for open systems. Recent theoretical calculations [60], summarized in section 2.4, which allow the adequate estimate of the Kirkwood-Buff integrals from simulations in closed systems have largely eliminated this issue.

Another difficulty associated with using the activity derivative as a target value to reproduce the ion-ion interaction, as we show in this work, is that it often reaches a plateau near the target experimental value, for a wide range of parameter values. The plateau emerges from the interconversion between different types of ion pairs, so parameters leading to equally good agreement with the target solution activity or activity derivative yield very different solution structures. Others have also observed plateaus in the activity derivative as a function of anion-cation interaction parameters, and these plateaus have been similarly linked to interconversion between different types of ion pairs [61]. The importance of lifting the indetermination in the parameters introduced by these plateaus by comparison with other properties that reflect ion-ion interactions, however, has not always been recognized. Instead, in some instances authors have arbitrarily chosen a particular scaling factor without physically justifying this choice (*e.g.*, parameters for $CaCl_2$ and $CaBr_2$ in ref. [61]). Here we show that by optimizing parameters to simultaneously satisfy the activity derivatives at intermediate concentration and the number of contact ion pairs at high concentration it is possible to obtain a reliable set of anion-cation parameters.

1.2.1.2 Proposed approach to develop parameters for monoatomic ions in TIP5P water

Our aim is to develop a force field for the Li^+, Na^+, K^+, Rb^+, Cs^+, F^-, Cl^-, Br^-, I^-, Mg^{2+} and Ca^{2+} ions in TIP5P water that adequately reflects the balance of ion-water and ion-ion interactions. To achieve this aim, we separately optimize the Lennard-Jones parameters for ion-water and anion-cation interactions to reproduce experimental observables that reflect those interactions. We opt to develop the ion-water parameters against the solvation free energies of ions at infinite dilution (done by our collaborators); for the case of NaCl, KCl, $MgCl_2$ and $CaCl_2$, ion-pair interactions are adjusted based on the molar activity derivative at 0.5 m concentration where m stands for molal concentration as well as the ion-ion coordination numbers. Our results show that adjusting ion-pair interactions based solely on the activity derivative – a state-of-the-art approach to optimize these interactions – is often insufficient, and that to obtain parameters that yields the correct solution structures requires that other experimental properties reflecting ion-ion interactions, are also used for parameterization. Here we use the anion-cation coordination number at high concentration of the salts to resolve this issue.

1.2.2 Polyatomic ions in TIP3P water

In contrast to the advanced state of development of classical molecular models for alkali halide ions for explicit water simulations, there are currently only a few models of polyatomic anions that have been parameterized to reflect the correct balance between their interactions with water, with cationic amino acids or with commonly used monoatomic counterions, such as Na^+ [62]. Such models are of interest because polyatomic anions such as sulfates, phosphates or acetates are present in the physiological context in both inorganic and alkylated form, and the molecular scale details regarding their interaction with biomolecules such as proteins are not yet understood[31, 63, 64]. Simulations of biomolecular systems performed so far with these ions have simply used general parameters based on commonly used biomolecular force fields such as CHARMM or AMBER [36]: the partial charges are obtained from low level quantum mechanical calculations and RESP fitting[65], and the Lennard-Jones parameters for each element are the same as those used in the chosen biomolecular force field. Parameters obtained this way are typically used for simulations without further tests – either comparison with experimental data or with ab initio calculations – under the assumption that the ion parameters must be reliable because they were obtained with the established protocol used to develop the biomolecular force field. However, these protocols have only been proven reasonable for neutral biomolecules [36, 43] and they are known to fail for charged molecules, *e.g.*, they yield hydration free energies substantially more negative than those experimentally measured [36].

In this work, we aim to overcome the above limitations of existing classical, fixed-charge models of polyatomic ions for simulations of biomolecular systems in explicit water. We focus specifically on the polyatomic anions methyl-sulfonate ($CH_3SO_3^-$), methyl-sulfate ($CH_3SO_4^-$), acetate (CH_3COO^-), dimethyl-phosphate ($(CH_3)_2PO_4^-$) and methyl-phosphate ($CH_3PO_4^{2-}$) (shown in Figure A.2); these anions were selected because they are good analogues of anionic polymers with longer alkyl chains. We also develop parameters for the non-methylated versions of the sulfate and phosphate ions (SO_4^{2-}, HSO_4^-, HPO_4^{2-}, $H_2PO_4^-$).

Figure 1.1 Oxoanions parameterized in this study.

The parameterization approach we present ensures that the ion parameters reflect the correct balance of the ion interactions with water, with the commonly used counterion Na^+ and with an analogue (methylammonium; $CH_3NH_3^+$) of the amino acid lysine, while ensuring that the parameters for all the ions remain internally consistent. Because often ammonium (NH_4^+) is taken instead of methylammonium as an analogue of lysine, we also parameterize the interactions of the anions with ammonium, to assess whether this cation is indeed a reasonable analogue of the side chain of lysine. The parameters developed here are compatible with widely used force fields

for molecular simulations: the AMBER [36] force field for proteins, the accompanying TIP3P [66] water model, and the sodium model of Young and Cheatham which is often used as counterion[44].

Intramolecular parameters for the anions already exist; the main issue is to parameterize their intermolecular interactions. In the context of the AMBER [36] force field for proteins, intermolecular interactions depend on the Lennard-Jones parameters and on the partial charges of the atoms composing the interacting species. We obtain the partial charges of the ions following the standard procedure used in AMBER [36] – low level ab initio calculations followed by RESP fitting – so only the Lennard-Jones parameters of the ions must be optimized.

1.2.2.1 Difficulties optimizing Lennard-Jones parameters for polyatomic anions

To parameterize intermolecular Lennard-Jones potentials, it would be desirable to reproduce high quality thermodynamic data that strongly depends on intermolecular interactions: *e.g.*, free energies of hydration, which reflect ion-water interactions, or solution activities, which reflect anion-cation interactions. This approach was followed by several groups to develop better models of alkali halides [44, 48, 45]. However, experimental data exists only for a few of the cases of interest here. For example, hydration free energies exist for CH_3COO^-, but not for for $CH_3SO_3^-$, $CH_3SO_4^-$, $CH_3PO_4^{2-}$, $(CH_3)_2PO_4^-$. [67] Osmotic pressure or solution activity data for solutions containing the cationic amino acids arginine and lysine, or suitable small-molecule analogues, and polyatomic anions are also not available. For solutions involving polyatomic anions and Na^+, data exists only for $CH_3SO_3^-$ and CH_3COO^-.

Given that there is no complete data set of experimental information that we can use to obtain an internally consistent set of parameters for all the anions mentioned above, an alternative would be to develop classical models solely based on ab initio calculations. One possibility would be to develop parameters to match observables (*e.g.*, the potential of mean force between two ions) calculated from ab initio molecular dynamics (MD) in explicit solvent [68]. There are two main shortcomings to this approach that make it unsuitable, though: i) currently, ab initio MD can only access timescales of tens of picoseconds, so the calculated observables have unacceptably high statistical uncertainty; ii) a low level of theory must necessarily be used in these calculations, possibly introducing systematic errors.

Another possibility would be to develop parameters based solely on single-point ab initio calculations, either of small hydrated systems (*e.g.*, an ion-pair or a single ion surrounded by its first hydration layer), of water-ion dimers or of a dry ion-pair; these last two types of system are here referred to as systems in vacuum. Multiple studies, however [43, 69], show that these approaches fail to produce parameters that are adequate for simulations of aqueous systems: the ion-water or anion-cation intermolecular interaction energies calculated using ab initio methods differ significantly from the energies of the same systems calculated using classical models optimized to reproduce properties of *aqueous* systems. How, then, can we obtain an internally consistent set of parameters for the ions in question, given the above described limitations

in experimental data and in ab initio methods?

1.2.2.2 Proposed approach

We propose an approach to develop non-bonded parameters for ionic species in aqueous systems that combines experimental data and ab initio simulations, in a manner that yields an internally consistent set of parameters for the anion-cation and ion-water interactions. We describe the approach with a general example. Let us consider species A, B and B'. B and B' are similar, *i.e.*, they have identical charge and similar composition (*e.g.*, $CH_3SO_4^-$ and CH_3COO^-). AB and AB' denote dimers of these species interacting non-covalently. Lennard-Jones parameters for the intermolecular interactions in AB *in water* have been developed based on experimental data (Figure 1.2, Step 1), but analogous experimental data for the AB' system does not exist. Using high level ab initio calculations, we calculate the intermolecular energies in the AB and AB' systems in the vacuum, as a function of the intermolecular distance (Figure 1.2, Step 2). We also perform the analogous potential energy scan for the

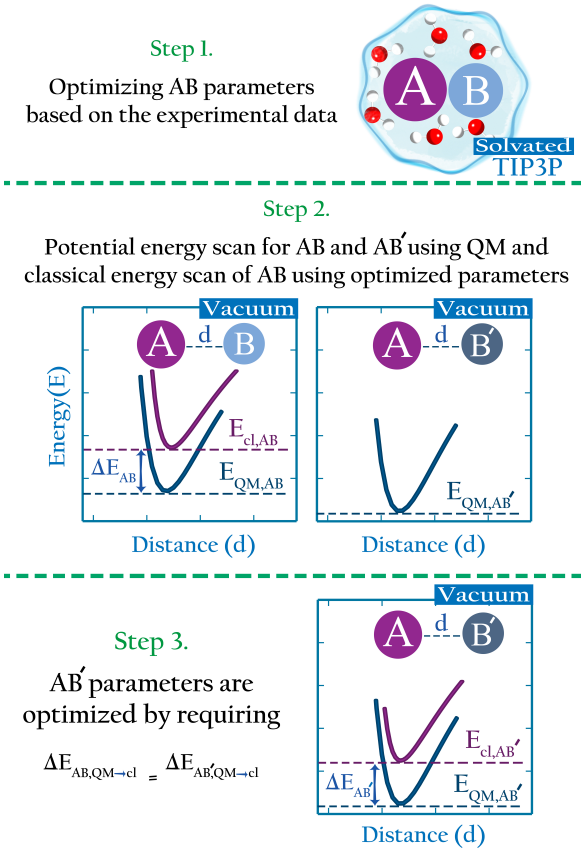

Figure 1.2 The parameterization approach in this study.

AB system using the classical, optimized parameters. As illustrated in cartoon form in Figure 1.2 (Step 2), the interaction energy for system AB from ab initio calculations differs from that calculated using the classical model. This difference reflects

the fact that the classical parameters (Lennard-Jones) were optimized for the intermolecular interactions of AB in water, and that the polarization state of molecules or ions in a dimer in the vacuum is very different from their polarization state in water [70, 43]. Therefore, the AB parameters offer a poor description of the intermolecular interactions of this dimer in the vacuum. We then assume that similar species should experience similar perturbations in their polarization state relative to the vacuum when immersed in water. If this assumption is true, then the *difference* in the minimum interaction energy between the classical and ab initio methods for AB, $E_{QM,AB}$-$E_{cl,AB}$, should be very similar to the same difference for the AB' system, $E_{QM,AB'}$ -$E_{cl,AB'}$. In this article we show that this assumption indeed holds. This finding implies that classical parameters for the intermolecular interactions of the AB' system can be determined by requiring that

$$\Delta E_{AB',QM\rightarrow cl} = \Delta E_{AB,QM\rightarrow cl} \tag{1.1}$$

where

$$\Delta E_{AB,QM\rightarrow cl} = E_{cl,AB} - E_{QM,AB} \tag{1.2}$$

is the difference in the energy minimum of the indicated dimer calculated using ab initio and that calculated using the optimized classical parameters, as illustrated in Figure 1.2 (Step 3). A similar approach has proven reasonable to develop parameters for neutral compounds like alcohols [69, 43] but has not been applied to charged species.

We use this approach to develop a set of classical parameters for polyatomic anions which is internally consistent and which is fully compatible with the AMBER [36] force field for proteins, the TIP3P [66] water model and the sodium ion model of Joung and Cheatham [44]. The experimental properties used for parameterization here are the solution activity derivatives and the hydration free energies. Solution activities exist only for $NaCH_3SO_3$, $NH_4CH_3SO_3$, $NaCH_3COO$, Na_2SO_4 and $(NH_4)_2SO_4$; reliable hydration free energies exist only for HSO_4^-, CH_3COO^- and SO_4^{2-}. We optimize the parameters for the ions based on this experimental information, and then use ab initio calculations and the approach described above to obtain parameters for the cases for which experimental information is not available. The final parameters are presented in the SI as .top/.itp files for GROMACS users; those files also the values of the parameters in AMBER format, indicated as comments.

1.2.3 Interactions between selectins and anionic dendrimeric polymers

As summarised in section 1.1.2, the specific ion effects observed in the interactions between dendrimeric polyanions and selectins (cationic proteins) are not understood. Investigating these effects requires that we first understand how individual anions and cationic amino acids interact; subsequently, supramolecular effects arising from the presence of multiple anionic sites in the dendrimer and multiple cationic sites in the protein should also be investigated. In this thesis we address the first of these issues only. Having developed parameters for small molecule analogues of the most

important anions (sulfate, phosphate, carboxylate) and the most abundant cationic amino acid in selectins (methylammonium as a mimic of lysine), we perform molecular simulations to investigate the mechanisms by which the anion identity affects anion-cation binding strength as a function of distance, and we characterized the molecular details of these interactions. Our results suggest that interactions between proteins and charged polymers strongly depend on the concentration of the counterion (i.e. Na^+). Counterions should be considered when interpreting experimental results, since strong interactions between the protein and the charged polymer may be greatly decreased by even stronger interactions between the polymer and the counterion.

Chapter 2

Methods

Here we introduce the various methods applied during the course of this work. We first describe in general how we parameterized both ion-water and anion-cation interactions for monoatomic ions in TIP5P water (section 2.1), and subsequently how we parameterized the same types of interactions for various polyatomic ions (section 2.2). To parameterize ions, we take the advantage of the widely used free energy perturbation method (section 2.3) and Kirkwood-Buff (KB) theory (section 2.4) to calculate the solvation free energy of ions and the activity derivatives of salt solutions, respectively. Furthermore, the parameters for the ions for which limited experimental data was available were obtained using the ab initio methods described in detail in section 2.2.6, following the approach described in section 1.2.1.2. In the final stages of this work, we quantify the binding affinity between the various anions and cations by calculating the potential of mean force as a function of distance using umbrella sampling (section 2.5). The potential of mean force is then used to compute the anion-cation binding constant (section 2.6).

2.1 Monoatomic ions in TIP5P water

The aim of this part of the work was to develop parameters for the ions composing the salts NaCl, KCl, $MgCl_2$ and $CaCl_2$, to simulate their behavior in solution using the TIP5P water model. This required the optimization of both ion-water and anion-cation interactions to reproduce experimental observables, as detailed below.

2.1.1 The TIP5P water model

The TIP5P [50] water model is a 5-site model, where the two hydrogens carry a positive charge of $+0.241\,e$, and the oxygen carries no charge. In contrast to the 3-site water models such as SPC [71] and TIP3P [66], TIP5P has two virtual sites, with -0.241$e$charge, to represent the oxygen's lone pairs[50]; see Figure 2.1.

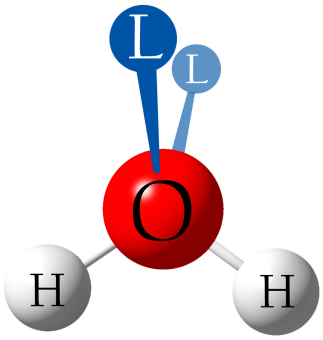

Figure 2.1 TIP5P water model. L represents the negatively charged virtual sites.

2.1.2 Optimizing the Lennard-Jones parameters for ion-water and anion-cation interactions

The non-bonded Lennard-Jones (LJ) potential form

$$V(r_{ij}) = \left(\frac{C_{ij}^{(12)}}{r_{ij}^{12}} - \frac{C_{ij}^{(6)}}{r_{ij}^{6}} \right) \tag{2.1}$$

is used, where $C_{ij}^{(12)}$ and $C_{ij}^{(6)}$ are LJ potential parameters describing the interactions between species i and j, and the superscripts between parenthesis are identifiers and not mathematical exponents. The r^{12} dependence of short range interactions is arbitrary, but the r^6 dependence of long range interactions is physically motivated [72]. We take advantage of this physical basis to obtain the homoionic $C^{(6)}$ coefficients for each ion species using the Slater-Kirkwood equation for homoionic dispersion interactions [73, 74]:

$$C_{ii}^{(6)} = C_{SK} \tilde{\alpha}^{3/2} n^{1/2} \tag{2.2}$$

where $C_{SK} = 23.97$ kJ·mol^{-1}· nm$^{3/2}$, $\tilde{\alpha} = \alpha/(4\pi\epsilon_0)$ is the static dipole polarizability volume of the ion, α is the static dipole polarizability, and n is the effective number of electrons for the isoelectronic noble gas. The static polarizability values for the cations are taken from the gas phase polarizabilities [75, 76, 77]; for the anions, the values for ion polarizabilities in solution are used [78]. The effective number of electrons are estimated based on the noble gas polarizabilities and dispersion coefficients [79, 80].

We use the $C^{(12)}$ coefficients as tuning parameters to reproduce experimental observables. The $C^{(12)}$ coefficient associated with ion-water interactions ($C_{iO}^{(12)}$), was parameterized by our collaborator to reproduce the experimental values of the intrinsic solvation free energy. Details of these calculations are not given here. Subsequently, we parameterized the $C^{(12)}$ coefficient associated with anion-cation interactions for some of the salts, to reproduce the activity derivative of solutions at 0.5 m concen-

tration. The activity derivative from simulation was calculated using Kirkwood-Buff theory, as detailed in section 2.4.

The Lennard-Jones coefficients, $C_{ij}^{(n)}$ (where $n = 6, 12$) for interactions between different species that are not explicitly parameterized are obtained based on the geometric average of the corresponding self-interaction parameters [50], $C_{ii}^{(n)}$ and $C_{jj}^{(n)}$, *i.e.*,

$$C_{ij}^{(n)} = (C_{ii}^{(n)} C_{jj}^{(n)})^{1/2} \tag{2.3}$$

Equation 2.3 is also used to obtain the $C_{ii}^{(12)}$ term for the interaction between two identical ions from the $C^{(12)}$ coefficients for water-water ($C_{OO}^{(n)}$) and ion-water ($C_{iO}^{(n)}$) interactions.

The initial choice of the anion-cation $C_{ij}^{(12)}$ parameters that are optimized against the activity derivative is the geometric average obtained with Equation 2.3, of the associated homoionic parameters for ions i and j.

For two cases (NaCl and KCl), the solution activity derivative by itself proved to be insufficient to uniquely determine the optimum anion-cation interaction parameters. For this reason, for these two salts the optimum parameters were found by requiring that they simultaneously reproduce the solution activity derivative at 0.5 m and the experimentally-determined number of contact ion pairs at high concentration.

2.1.3 General simulation details

Simulations were performed using the GROMACS simulation package [81, 82] and the TIP5P water model [50]. We perform two types of simulations: simulations of NaCl, KCl, $MgCl_2$ and $CaCl_2$ solutions at 0.5 m concentration, used to calculate the activity derivatives of these solutions; and simulations of NaCl at 6.18 m and KCl at 4.56 m, to calculate anion-cation radial distribution functions and coordination numbers.

The simulation boxes are cubic and periodic boundary conditions are applied in all directions. The initial configurations are assembled with the `editconf` module in Gromacs and the Packmol package [83]. Energy minimization is performed by the steepest descent algorithm. Non-bonded potentials are cut off at 1.4 nm; electrostatic interactions beyond that cutoff distance are calculated using the particle mesh Ewald (PME) method [84]. The leap-frog algorithm [85] is employed to integrate the equations of motion every 2 fs. Simulations are performed at 298.15 K; the temperature is kept constant using the Nose-Hoover scheme [86, 87] with a time constant of 0.1 ps. Simulations in the NPT ensemble are done with a target pressure of 1 bar, using the Parrinello-Rahman [88] approach with a time constant of 0.5 ps. All systems are equilibrated for 100 ps in the NVT ensemble followed by 100 ps in the NPT ensemble. For the calculation of the activity derivatives, salt solutions are equilibrated in the NPT ensemble for 10 ns. For each salt, we select three distinct configurations from the equilibration run as the starting configurations for 3 independent NVT production runs. The starting configurations have a box size close to the average box size during equilibration. Each production run lasts 50 ns, totaling 150 ns simulation time for each salt, with configurations saved every 2 ps.

2.2 Polyatomic ions in TIP3P water

This part of the work focuses specifically on polyatomic anions: HSO_4^-, $CH_3SO_4^-$, $CH_3SO_3^-$, $H_2PO_4^-$, $CH_3HPO_4^-$, $(CH_3)_2PO_4^-$, CH_3COO^-, SO_4^{2-}, HPO_4^{2-} and $CH_3PO_4^{2-}$. We optimize their interactions with TIP3P water, and with the cations Na^+, NH_4^+ and $CH_3NH_3^+$.

2.2.1 The TIP3P water model

The Lennard-Jones (LJ) parameters of the TIP3P oxygen are $\varepsilon_{OO} = 0.6364$ kJ/mol and $\sigma_{OO} = 0.315061$ nm, with the partial charge of $q_O = -0.8340e$ [66]. The partial charge for each of the two hydrogens is $q_H = +0.4170e$, however there are no LJ parameters assigned to the two hydrogens in TIP3P water [66].

2.2.2 Parameters for the ions

The parameters for NH_4^+ and $CH_3NH_3^+$ are those from GAFF. The hydrogen atoms connected to the nitrogen in these cations are indicated as H_N. The sodium and chloride models used here are those from Joung and Cheatham [44].

We introduce a general notation for the anions, A^{n-}, here investigated: $(RO)_kXO_m^{n-}$ (where X={S,P,C} and R={H,CH_3}). The anions have up to three different oxygen types, represented as follows: atom type O2 represents terminal oxygens, atom type OH represents acidic oxygens and atom type OS represents X-O-X oxygens (*e.g.*, C-O-P in $(CH_3)_2PO_4^-$). We use $\mathcal{O}$ to represent any of the three oxygen types: $\mathcal{O}$={O2,OH,OS}. Most bonded parameters for the anions are from GAFF. The only exceptions are the force constants associated with the HO-OH-X, OH-X-OH and OH-X-O2 angles in the $HSO_4^-$and $H_2PO_4^-$ions, which were changed from their original values to 90 kcal mol^{-1} rad^2, to prevent unphysical collapse of the hydroxyl hydrogen on other oxygens of the oxoanion. The new values of the force constants were taken from ref. [89]. The partial charges for the atoms composing the anions are obtained using the standard RESP fitting protocol used in the AMBER [36] force field, as described in section 2.2.6. The Lennard-Jones parameters ε or σ for the oxygens of each of the anions are optimized in this study, as described in the Results section. For the remaining atoms of the anions, the Lennard-Jones parameters of the GAFF [90, 91, 36, 92] force field are used.

2.2.2.1 Optimizing the Lennard-Jones parameters for ion-water and anion-cation interactions

The ε and σ parameters determine the Lennard-Jones interaction energy of two species i and j as

$$U_{ij} = 4\varepsilon_{ij}\left(\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right) \tag{2.4}$$

Unless otherwise noted, the Lorentz-Berthelot combination rules are used to obtain ε_{ij} and σ_{ij} from the self-interaction parameters, denoted by the subscripts ii and jj, as follows:

$$\varepsilon_{ij} = \sqrt{\varepsilon_{ii}\varepsilon_{jj}}, \quad \sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2} \tag{2.5}$$

The optimization of the oxygen parameters proceeded in four stages. First, the self-interaction ε or σ parameters of the oxygens in HSO_4^-, CH_3COO^- and SO_4^{2-} were optimized to reproduce the experimentally determined difference in solvation free energy between those anions and Cl^-:

$$\Delta\Delta G_{solv} = \Delta G_{solv}(A^{n-}) - n \times \Delta G_{solv}(Cl^-) \tag{2.6}$$

We choose to optimize parameters to reproduce the experimental value of $\Delta\Delta G_{solv}$ instead of ΔG_{solv} because of the lower uncertainty [93] associated with experimental values of $\Delta\Delta G_{solv}$ relative to ΔG_{solv}, and because we aim to create a force field compatible with the Na^+ and Cl^- ions developed by Joung and Cheatham [44]. The parameters obtained this way yielded the reference values used in the second parameterization stage to optimize the self-interaction parameters of the oxygen atoms of the remaining anions following approach described in the introduction. The self-interaction oxygen parameters optimized in stages 1 and 2 should be used in equations 2.5 to obtain the ε and σ parameters for the interaction of the anions with water and with any other species in solution, except the cations Na^+, NH_4^+ and $CH_3NH_3^+$, for which we derive specific anion-cation parameters.

For the anion-cation interactions investigated here, the combination rules shown in equation 2.5 proved insufficient, as described in the Results section. For this reason, the Lennard-Jones ε or σ parameters corresponding to the interaction between the oxygens of the anions and Na^+, as well as with N and H_N in NH_4^+ or $CH_3NH_3^+$, were parameterized in stage 3 to reproduce the activity derivative of 0.5 m solutions of salts for which that information was available. In stage 4, we obtained optimized parameters for the interactions of the remaining anions and Na^+, $CH_3NH_3^+$ and NH_4^+ using the ab initio approach described in the Introduction, and using the parameters optimized in stage 3 as reference. Each parameterization stage is described in detail in the Results section; the final scaling factors to the GAFF parameters are shown in Tables 3.3, 3.4, 3.5, 3.7 and 3.8.

2.2.3 Solvation free energy

The principles of free energy perturbation are given in section 2.3; here we describe only the specific implementation details that were used.

The starting configuration for the solvation free energy calculations is prepared by placing a single ion in a simulation box of dimensions $\sim 4.0 \times 4.0 \times 4.0$ nm^3 and solvating it in TIP3P water [66]. Periodic boundary conditions are applied in all directions. The free energy is obtained along the path coordinates of λ_{LJ} and λ_C, which are the scaling factor to the Lennard-Jones and the electrostatic potentials, respectively. These scaling factors range from 0 to 1 (1 for the decoupled state

and 0 for the fully coupled state). To ensure adequate ensemble overlap, we use 21 steps for each of λ_{LJ} and λ_C; the steps are evenly spaced for the λ_C; for the λ_{LJ} spacing we use 21 unevenly spaced points, $\lambda_{LJ} \in \{0.00\ 0.06\ 0.12\ 0.18\ 0.24\ 0.30\ 0.36\ 0.42\ 0.46\ 0.50\ 0.52\ 0.54\ 0.56\ 0.58\ 0.60\ 0.64\ 0.68\ 0.72\ 0.76\ 0.80\ 1.00\}$. Soft core potentials are used during the LJ coupling. For each λ we perform an equilibration for 100 ps followed by a production run of 1 ns in NPT ensemble.The total solvation free energies are calculated from the partial results along the path coordinates using Bennett's acceptance ratio (BAR) [94], as implemented in Gromacs [81, 82]. We set a non-bonded potential cutoff at 1.2 nm distance. Beyond this cutoff distance, the electrostatic interactions are calculated with the particle mesh Ewald (PME) method [84]. Lennard-Jones interactions smoothly shifted to zero between 1.0 and 1.2 nm. A leap-frog stochastic dynamics (SD) integrator [95] is used to integrate the equations of motion in all simulations with the temperature fixed at 298 K. For NPT simulations we maintained the pressure at 1 bar using the Berendsen barostat [96] with a coupling time of 1.0 ps. All bonds with H-atoms were restrained using the LINCS algorithm [97]. We used a 2 fs time step.

2.2.3.0.1 Correction terms to the solvation free energies: The free energy values calculated from simulations (ΔG^{sim}) are known to vary markedly depending on the method by which electrostatic interactions are calculated [98, 99, 59, 100, 73]. To avoid this method dependence it is necessary to apply multiple correction terms to the calculated values so that they can be compared with experimental free energies of solvation. The largest of these terms is the Wigner correction term for the ion in the gas phase; this term is already included in the free energy values calculated using Gromacs [81, 82].

Another important correction term eliminates the error incurred on by using particle-based summation schemes when calculating the long-range electrostatic interactions as Gromacs [81, 82] does, instead of using the more adequate molecule-based summation schemes. This term is not applied here, however, because it varies linearly with the charge of the ion and, because we are concerned only with *differences* in hydration free energy (see equation 2.6), this term cancels out.

The free energies calculated using our chosen setup are intrinsic: they do not contain a contribution of crossing the macroscopic air-water interface [93]. However, often compilations of hydration free energies from experiment include this contribution – they are referred to as real free energies of hydration of single ions [93]. Because the surface potential term is linear in the charge of the ion and thus cancels out in equation 2.6, it can be ignored in our approach, independently of which experimental data set is used for parameterization.

The free energy values calculated from simulations correspond to the transfer of an ion in a given volume in the gas to the same volume of liquid. In contrast, the experimental values correspond to a process where the ion is transferred from an ideal gas state at 1 atm to an ideal aqueous solution at 1 M concentration, *i.e.*, the experimental values of the solvation free energy contain an entropic term that is

absent from simulation. This entropic term, ΔG_{press}, has the form:

$$\Delta G_{press} = -T\Delta S_{press} = -k_B T \ln\left(\frac{V_l}{V_g}\right) \tag{2.7}$$

where $V_l = 1$ l/mol is the molar volume of an ideal aqueous solution of concentration 1 M and $V_g = 24.465$ l/mol is that of an ideal gas at 1 atm. This term will not cancel out in the difference shown in equation 2.6 for multivalent ions, so $\Delta G_{press} = 1.888$ kcal/mol must be added to the solvation free energy values calculated using simulations before that equation can be applied.

Transferring a molecule from the gas phase to the solution phase changes its electronic density. A polarization correction is also included in calculating solvation free energies to take this change into consideration. This correction (ΔG_{pol}) was calculated using ab initio methods, as described in detail in section 2.2.6.2.

All the correction terms that do not cancel out when using equation 2.6 were applied in this work to obtain the calculated ΔG_{solv}^{Comp}, following equation 2.8:

$$\Delta G_{solv}^{Comp} = \Delta G_{cav} + \Delta G_{chg} + \Delta G_{pol} + \Delta G_{press} \tag{2.8}$$

where ΔG_{cav} and ΔG_{chg} are the Lennard-Jones and the electrostatic components of the hydration free energy yielded directly by Gromacs.

2.2.4 Solution activity derivative

The principles behind the calculation of activity derivatives are summarized in section 2.4; here we describe only the particular implementation details that were used.

All simulations are performed using the Gromacs simulation package [81, 82]. All simulation boxes are cubic with periodic boundary conditions applied in the XZY directions. The simulation box for the calculation of the activity derivative was prepared by putting 72 anions and 72 cations (144 when using divalent anions) in a cubic box with dimensions $4.0 \times 4.0 \times 4.0$ nm^3 and solvating the system by adding 7996 TIP3P [66] water molecules using editconf and solvate tools form Gromacs [81, 82]. The salt concentration simulated is thus 0.5 m. We select three distinct configurations from the equilibration run, with a box volume similar to the average box volume obtained during the equilibration. Each of these configurations is then used as the initial state for a 50 ns simulation in the NVT ensemble. The total simulation time for each production run is thus 150 ns; this large simulation time is necessary to provide enough sampling to calculate activity derivatives. The remaining simulation details are identical to those used to calculate free energies of solvation (section 2.2.3).

2.2.5 Potential of mean force

Here we present the details pertaining to the umbrella sampling simulations used to calculate the potential of mean force for various ion pairs. The principles of the technique are summarized in section 2.5. We run the umbrella sampling for the CH_3COO^-, $CH_3SO_3^-$, $CH_3SO_4^-$, $H_2PO_4^-$, $(CH_3)_2PO_4^-$, $CH_3HPO_4^-$ and $CH_3PO_4^{2-}$ ions

with each of the Na^+ and $CH_3NH_3^+$. We set up all simulations by placing a single anion-cation pair inside a cubic box of size $\sim$3.4×3.4×3.4 nm^3 and solvate the molecules using TIP3P water. The systems are then equilibrated for 200 ps in the NPT ensemble at 298.15 K. The temperature is kept constant using the Nose-Hoover scheme [86, 87] with a time constant of 0.1 ps. Simulations are done with a pressure of 1 bar, using the Parrinello-Rahman [88] approach with a time constant of 0.5 ps. The remaining simulation details are identical to those used to calculate free energies of solvation (section 2.2.3).

Running umbrella sampling requires generating several configurations along a reaction coordinate. We take the advantage of steered molecular dynamics simulations (SMD) and the implemented pull code in Gromacs to pull ion-pairs apart, along the cation-anion direction. We set the pull rate to 0.001 nm/ps and pull force constant to 5000 kJ mol^{-1} nm^{-2}. The initial distance between cation and anion is set to be zero using *pull_ init1 = 0*, to force the two molecules to approach as close as possible. We run the pulling simulations for 1.5 ns, which corresponds to a maximum anion-cation separation of 15 Å, sufficient for the purpose of our study. We select multiple configurations with anion-cation separations between the minimum and the maximum separation spanned during the SMD simulation; each selected configuration is the starting point for an umbrella sampling window. For each window we run an independent NPT simulation for 10 ns. We use the WHAM [101] analysis program implemented in Gromacs to compute the potential of mean force; see section 2.5 for details.

2.2.6 Ab Initio Calculations

2.2.6.1 Charge Fitting

The classical charges for the new species are obtained consistently with the AMBER [36] force field for proteins. The molecules are first optimized using MP2/6-31G* level of theory to find the global minimum geometries. Thereafter we use the HF/6-31G* level of theory to calculate the Merz-Kollman [102] atomic charges fitted to the electrostatic potential. All calculations are performed using the Gaussian 03 software package [103]. The Gaussian output is later plugged into the Antechamber module of the AMBER [36] software suit to generate the classical atomic point charges using RESP [65] fitting.

2.2.6.2 Polarization correction to the solvation free energy

The polarization energies were calculated in Gaussian 03. The structures are first optimized using the IEF-PCM [104, 105, 106, 107, 108] model with B3LYP [109, 110, 111]/aug-cc-pVTZ level of theory followed by a single point energy calculation in vacuum, using the same level of theory. This energy corresponds to a solvated structure (more precisely, electron density, ρ_{solv}), measured in gas phase, $E^{gas}[\rho_{solv}]$. Thereafter, we re-optimize the molecules in the gas phase to obtain the minimum energies in the vacuum, $E^{gas}[\rho_{gas}]$. Subtracting the former from the latter energy will give us good estimates of the polarization energies [89].

$$\Delta G_{Pol} = E^{gas}[\rho_{solv}] - E^{gas}[\rho_{gas}] \tag{2.9}$$

The default values in Gaussian 03 for the dielectric constant, fast response dielectric constant (ε_{inf}) and probe radius (R_{solv}) are 78.355, 1.776 and 1.385 Å, respectively for this model.

2.2.6.3 Potential Energy Scan

To optimize the parameters of the interaction of the anions with water and with cations for the cases where experimental data is unavailable, we require rigid-body potential energy scans of the binding energy in anion-water or anion-cation systems in the vacuum, as described in the Introduction and illustrated schematically in Figure 1.2. The potential energy curves as a function of the intermolecular distance are obtained for each dimer using the MP2 [112] level of theory with augmented-cc-pvt z[113] basis set in the Gaussian 03 [103] package. The structure of each isolated species is first optimized with the same level of theory. After optimization, we perform the gas phase 1-D potential energy while holding the structure of the molecules fix during the scan. Typical configurations for the anion-cation and anion-water potential energy scans performed are shown in Figure 2.2.

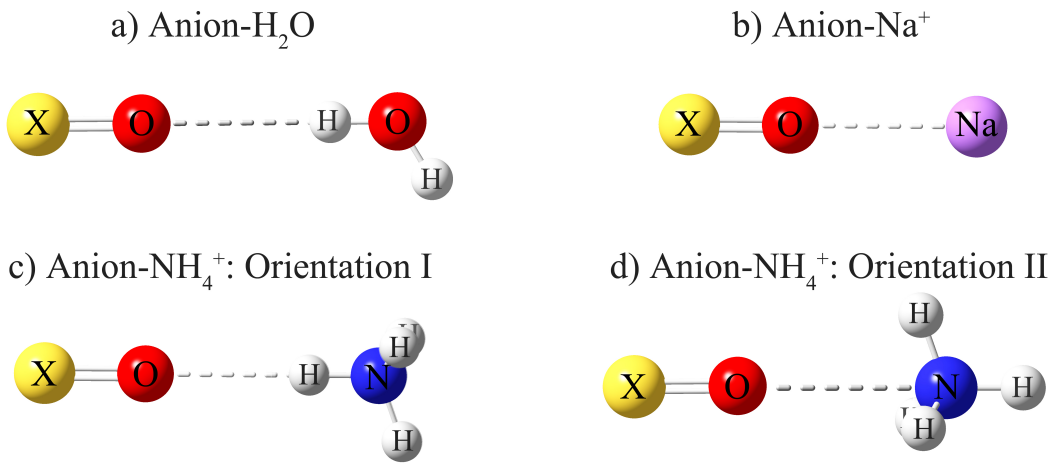

Figure 2.2 Example configurations for the 1-D potential energy scans of anions: a) with H_2O. b) with Na^+, c) with NH_4^+ in orientation I (strong hydrogen bond), d) with NH_4^+ in orientation II (weak hydrogen bond).

2.2.6.3.1 Basis set superposition error When studying ion-ion interactions with a finite basis set, the basis set superposition error (BSSE) [114] could become significantly large. The BSSE occurs when two species A and B are approaching each other and form a complex A-B. The resulting A-B complex is additionally stabilized due to the fact that each of the A and B are using extra basis functions from one another [115]. However, the overlap integral decreases as the distance between A

and B increases, and becomes zero when the two species are beyond a few angstroms apart; therefore, the total basis functions used vary as a function of the intermolecular distance between A and B. Under these circumstances, the binding energy is calculated as $\Delta E_{Binding} = E^{AB}(AB) - E^{A}(A) - E^{B}(B)$, where the superscripts are the basis functions used for the energy calculations. Using varying basis functions during 1D potential energy scans thus leads to artifacts in the binding energy and the minimum distance between the two ions. To resolve this problem we use the Boys and Bernardi [116] counterpoise corrections (CP) which is shown to correct for the basis set superposition error by using the AB basis to calculate the energies for each of A and B. The expression for the binding energy used in this work, which includes the CP correction, is thus:

$$E_{QM,AB} = E^{AB}(AB) - E^{AB}(A) - E^{AB}(B) \tag{2.10}$$

2.3 Calculating solvation free energies using free energy perturbation

The free energy of solvation of a single molecule can be calculated using free energy perturbation method. The aim is to find a way to calculate the differences in free energy between state 1 (solvated) and state 0 (non-solvated), using the connection between thermodynamics and statistical mechanics in the canonical ensemble [117]:

$$A = -\beta^{-1} \ln Q(N, V, T) \tag{2.11}$$

where A is the Helmholtz free energy, $\beta = (k_B T)^{-1}$, k_B is the Boltzmann constant and $Q(N, V, T)$ is the canonical partition function. The difference in free energy between the states 0 and 1 of the system can be expressed as the ratio between their partition functions [117]:

$$\Delta A = -\beta^{-1} \ln \frac{Q_1}{Q_0} \tag{2.12}$$

The canonical partition function for a system consist of N-particles is written as [117]:

$$Q = \frac{1}{h^{3N} N!} \int \int \exp[-\beta \mathcal{H}(x, p_x)] dx dp_x \tag{2.13}$$

where h is the Planck constant and $\mathcal{H}(x, p_x)$ is the Hamiltonian of the system and is a function of Cartesian coordinates, x, and the momentum, p_x. The $N!$ term is to factor out the configurational repetition. Substituting Eq. 2.13 into Eq. 2.11 gives:

$$\Delta A = \frac{-1}{\beta} \ln \frac{\int \int \exp[-\beta \mathcal{H}_1(x, p_x)] dx dp_x}{\int \int \exp[-\beta \mathcal{H}_0(x, p_x)] dx dp_x} \tag{2.14}$$

Using the relations between the target and the reference Hamiltonian [117]:

$$\mathcal{H}_1(x, p_x) = \mathcal{H}_0(x, p_x) + \Delta \mathcal{H}(x, p_x) \tag{2.15}$$

and the probability density function of finding the reference system in a state defined by x and p_x

$$P_0(x, p_x) = \frac{\exp[-\beta \mathcal{H}_0(x, p_x)] dx dp_x}{\int \int exp[-\beta \mathcal{H}_0(x, p_x)] dx dp_x} \tag{2.16}$$

the Eq. 2.14 becomes:

$$\Delta A = \frac{-1}{\beta} \ln \int \int \exp[-\beta \Delta \mathcal{H}(x, p_x)] P_0(x, p_x) dx dp_x \tag{2.17}$$

We write the equation 2.17 in an ensemble average form, over all reference state configurations [117]:

$$\Delta A = \frac{-1}{\beta} \ln \langle \exp[-\beta \Delta \mathcal{H}(x, p_x)] \rangle_0 \tag{2.18}$$

If the masses of the particles are the same in systems 0 and 1, the kinetic terms in the partition function would be identical for both systems and therefore we can re-write the equation 2.18 as the following [117]:

$$\Delta A = \frac{-1}{\beta} \ln \langle \exp[-\beta \Delta U(x, p_x)] \rangle_0 \tag{2.19}$$

where ΔU is the potential energy difference between systems 0 and 1.

Equation 2.19 is a good estimate of the free energy changes if the perturbation is small. For large perturbations such as the free energy change due to solvation of a molecule, the configurations with no water-solute overlap are poorly sampled in the reference system and therefore equation 2.19 cannot be used directly. Instead, we must use intermediate states between the reference and the target systems. We can define N-2 intermediate states and re-write equation 2.19 as [117]:

$$\Delta A = \sum_{i=1}^{N-1} \Delta A_{i,i+1} = \frac{-1}{\beta} \sum_{i=1}^{N-1} \ln \langle \exp[-\beta \Delta U_{i,i+1}] \rangle_0 \tag{2.20}$$

We can generalize this formalism by writing the Hamiltonian as a linear function of λ values [117]

$$\mathcal{H}(\lambda_i) = \lambda_i \mathcal{H}_1 + (1 - \lambda_i) \mathcal{H}_0 = \mathcal{H}_0 + \lambda_i \Delta \mathcal{H} \tag{2.21}$$

where $0 < \lambda_i < 1$. The change in Hamiltonian between states i and $i+1$ is therefore given as

$$\Delta \mathcal{H}_i = \mathcal{H}(\lambda_{i+1}) - \mathcal{H}(\lambda_i) = (\lambda_{i+1} - \lambda_i) \Delta \mathcal{H} = \Delta \lambda_i \Delta \mathcal{H} \tag{2.22}$$

We can therefore write the total change in the free energy during solvation of a molecule as [117]

$$\Delta A = \sum_{i=1}^{N-1} \Delta A_{i,i+1} = \frac{-1}{\beta} \sum_{i=1}^{N-1} \ln \langle \exp[-\beta \Delta \lambda_i \Delta U_{i,i+1}] \rangle_{\lambda_i} \tag{2.23}$$

Note that we used the assumption of constant masses for the particles again and factored out the kinetic term in the Hamiltonian [117].

To calculate the changes in free energy between each pair of consecutive states with least error, several numerical methods were developed. Here we use the Bennett Acceptance Ratio (BAR) method [94] to achieve this goal. At each step the change in the free energy of the forward and backward path, gives a different result and averaging the two is not a good estimate of the ΔA since the accuracy of the free energy difference calculation is not the same for the two opposite directions. This is mainly due to the sampling differences between the reference and the target systems resulting in different probability distributions $\mathrm{P}_i(\Delta \mathrm{U}_{i,i+1})$ and $\mathrm{P}_{i+1}(\Delta \mathrm{U}_{i+1,i})$ [117]. Therefore, averaging over the forward and backward calculations might have less accuracy than each of the individual ones. The idea of BAR method is to find a point C between i and $i+1$ and calculate the forward free energy difference from i to C and backward calculation from $i+1$ to C. The differences between the two free energy calculations gives $\Delta \mathrm{A}_{i,i+1}$. The C is chosen so that the error in estimating the $\Delta \mathrm{A}_{i,i+1}$ is minimized [117, 94].

$$\Delta A_{i,i+1} = \frac{-1}{\beta} \ln \frac{\langle \{1 + \exp[\beta(\Delta U_{i,i+1} - C)]\}^{-1} \rangle_i}{\langle \{1 + \exp[-\beta(\Delta U_{i,i+1} - C)]\}^{-1} \rangle_{i+1}} + C \tag{2.24}$$

The constant C is defined

$$C = \Delta A_{i,i+1} + \frac{-1}{\beta} \ln \frac{n_i}{n_{i+1}} \tag{2.25}$$

where, n_i and n_{i+1} are the sample size in states i and $i+1$. Note by that defining the equation 2.24 this way, the point C does not need to be sampled and we only need to ensemble average over the states in i and $i+1$. Therefore, knowing the potential energy of each intermediate states, the BAR method can be applied to post-process the result. Since both Δ $\mathrm{A}_{i,i+1}$ and C values are unknown, the equations 2.24 and 2.25 should be solved together using iterative methods.

For the purpose of calculating free energy of solvation, we split the solvation process into the free energy changes due to the electrostatic interactions, V_c, and the changes due to van der Waals interactions, V_{LJ}. We define λ values for each term [118]

$$V_c(\lambda) = \frac{1}{4\pi\varepsilon_0\varepsilon_r r_{ij}} [(1-\lambda) q_i^0 q_j^0 + \lambda q_i^1 q_j^1] \tag{2.26}$$

where ε_r is the relative permittivity and ε_0 the vacuum permittivity.

$$V_{LJ}(\lambda) = \frac{(1-\lambda)C_{12}^0 + \lambda C_{12}^1}{r_{ij}^{12}} - \frac{(1-\lambda)C_6^0 + \lambda C_6^1}{r_{ij}^6} \tag{2.27}$$

where the 0 and 1 denote the states $\lambda = 0$ and $\lambda = 0$, respectively. C_6 and C_{12} are $4\varepsilon\sigma^6$ and $4\varepsilon\sigma^{12}$, respectively, and r_{ij} is the distance between two particles [118].

Turning on ($\lambda = 0$) or off ($\lambda = 1$) the electrostatic or van der Waals terms of the potential using the linear interpolation, might result in a large fluctuation in the vicinity of λ=0 and λ=1. This is due the weak interactions between the particles that allows them to closely approach each other and result in extreme repulsive or attractive forces when calculating potential energy of the following λ state.

To avoid unwanted instabilities near extreme values of λ we use soft core potentials,V_{sc}, for the Lennard-Jones interactions [118]

$$V_{sc}(r) = (1-\lambda)V^0(r_0) + \lambda V^1(r_1) \tag{2.28}$$

where

$$r_0 = (\alpha\sigma_0^6\lambda^p + r^6)^{\frac{1}{6}}$$

$$r_1 = (\alpha\sigma_1^6(1-\lambda)^p + r^6)^{\frac{1}{6}}$$

V^0 and V^1 are the normal van der Waals potentials of the system in states λ=0 and λ=1. σ is the radius of the interaction, equivalent to $(\frac{C_{12}}{C_6})^{1/6}$, α is the soft-core parameter and p is the soft-core λ power. The interactions in a soft-core potential converge to a constant value, rather than blowing up, as $r \to 0$ [118].

2.4 Calculating activity derivatives using Kirkwood-Buff theory

The solution activity derivatives were used to optimize anion-cation interaction parameters for both monoatomic ions in TIP5P water and polyatomic ions in TIP3P water. Details specific to each of these applications are given in Sections 2.1 and 2.2; here we provide a short summary of the most relevant aspects of the Kirkwood-Buff theory and its use in simulations of closed systems.

The molar activity derivative is defined as,

$$a_{cc} = \frac{\partial \ln a_c}{\partial \ln \rho_c}\Big|_{P,T} \tag{2.29}$$

where a_c is the solution molar activity and ρ_c is the number density of the solution, defined as $\rho_c = n_c/V = (n_+ + n_-)/V$; n_+ and n_- are the number of cations and anions, respectively, in the solution volume V. Using simulations, this derivative can be easily calculated from the radial distribution functions (rdfs) in dilute solutions, as formulated in the Kirkwood-Buff (KB) theory and described in detail elsewhere [119, 120, 48, 121, 122, 45].

The original KB theory is developed for the grand canonical ensemble, μVT; in this ensemble, the KB integrals are defined as:

$$G_{ij} = 4\pi \int_0^\infty r^2[g_{ij}^{\mu VT}(r) - 1]dr \tag{2.30}$$

where $g_{ij}(r)$ are the rdfs of the two species i and j and the r is the distance between them. Grand canonical ensemble simulations of dense systems are challenging to perform, because the probability of inserting particles is very low. To avoid this difficulty, we use an alternative formulation of the KB integrals, defined in terms of rdfs, $g_{ij}^{NVT}(r)$, obtained in the canonical ensemble [60, 123]:

$$G_{ij} = \int_0^{2R} [f_{ij}g_{ij}^{NVT}(r) - 1]4\pi r^2(1 - 3x/2 + x^3/2)dr \tag{2.31}$$

where $x = r/(2R)$ and f_{ij} is a concentration-dependent correction factor to the radial distribution function [124, 125, 126]. The interactions in the simulation box can be divided into two subsets: water(w)-co-solvent(c) interactions and co-solvent(c)-co-solvent(c) interactions. The corresponding KB integral terms are given by:

$$G_{cc} = \frac{1}{4}(2G_{+-} + G_{++} + G_{--}) \tag{2.32}$$

and

$$G_{cw} = G_{w+} + G_{w-} \tag{2.33}$$

Using these expressions for KB integrals the molar activity derivative of the solution is given by:

$$a_{cc} = \frac{1}{1 + \rho_c(G_{cc} - G_{cw})} \tag{2.34}$$

The activity derivative obtained this way can be compared to the molar activity derivative obtained from experiment, calculated using equation 2.29. Experimental activity derivatives are typically reported in the molal scale, so a correction factor must be applied to the molar activity derivatives to allow for comparisons with the equivalent molal quantity. At the low (0.5 m) salt concentration at which we work, however, the correction factor is very small and can safely be ignored.

2.5 Calculating potentials of mean force using umbrella sampling

The potential of mean force along a coordinate ξ, is defined from the average distribution function $\langle\rho(\xi)\rangle$ [127],

$$\mathcal{W}(\xi) = \mathcal{W}(\xi^*) - k_B T \ln\left[\frac{\langle\rho(\xi)\rangle}{\langle\rho(\xi^*)\rangle}\right] \tag{2.35}$$

where ξ^* and $\mathcal{W}(\xi^*)$ are some arbitrary constants. The average distribution function along ξ is defined as [127]

$$\langle\rho(\xi)\rangle = \frac{\int d\boldsymbol{R}\, \delta(\xi'[\boldsymbol{R}] - \xi)\, \exp(-\beta U(\boldsymbol{R}))}{\int d\boldsymbol{R}\, \exp(-\beta U(\boldsymbol{R}))} \tag{2.36}$$

where $\xi'[\boldsymbol{R}]$ is a function of $\boldsymbol{R}$ that only depends on a few or several degrees of freedom, such as distance, angle, etc. $U(\boldsymbol{R})$ is the potential energy of the system as a function of coordinates $\boldsymbol{R}$. For example we can find the average distribution function when the ξ set to be the distance between two particles i and j. The integrand in the numerator of equation 2.36 survives only if the distance function $\xi'[\boldsymbol{R}]$ is equal to the ξ value [127].

Calculating the PMF is not always a straightforward procedure, since the presence of high energy barriers along the ξ coordinate might make it impossible to sufficiently sample the configurational space. One way to overcome this issue is to apply some biased window potential, w_i, to facilitate the configurational sampling in the vicinity

of a chosen ξ_i value. The potential energy of the system would therefore be $[U(\boldsymbol{R}) + w_i(\xi)]$ [127]. This biasing potential would help to restraint the system around ξ_i value, so that a sufficient configurational sampling is obtained. In this study we use a harmonic function in the form $w_i(\xi)=\frac{1}{2}(\xi - \xi_i)^2$ to restrain the ξ coordinate, the cation-anion distance in this study, around consecutive ξ_i values. w_i denotes the bias potential in each window i. To obtain the final estimate of the PMF, the results of the successive windows are unbiased and combined together. The biased distribution function can be obtained by substituting the $U(\boldsymbol{R})$ with $[U(\boldsymbol{R}) + w_i(\xi)]$ in equation 2.36, as the following [127]:

$$\langle \rho(\xi) \rangle_{(i)}^{biased} = \exp(-\beta w_i(\xi)) \ \langle \rho(\xi) \rangle \ \langle exp(-\beta w_i(\xi)) \rangle^{-1} . \tag{2.37}$$

The unbiased PMF can be obtained by substituting equation 2.37 into equation 2.35, which gives

$$\mathcal{W}_i(\xi) = \mathcal{W}(\xi^*) - k_B T \ln \left[\frac{\langle \rho(\xi) \rangle_{(i)}}{\langle \rho(\xi^*) \rangle} \right] - w_i(\xi) + F_i \tag{2.38}$$

where F_i is the free energy resulted from the biased potential and defined as

$$\exp(-\beta F_i) = \langle \exp(-\beta w_i(\xi)) \rangle \tag{2.39}$$

Several approaches have been proposed to find the F_i values and calculating the PMF functions [127]. At the core of these approaches is the fact that F_i can be obtained by requiring that the unbiased potential, $\mathcal{W}_i(\xi)$, should be the same in the overlapping regions of adjacent windows. Therefore, one can adjust the $\mathcal{W}_i(\xi)$ values by changing the F_is until they match, which can be done using least-squares methods. The PMF along the ξ coordinate can be obtained by connecting the PMFs of the consecutive windows. Most of the approaches proposed to find the F_i values, however, require significant overlap between the two adjacent windows and therefore are computationally expensive [127].

The weighted histogram analysis method (WHAM) is intended to enhance the efficiency and the accuracy of the unbiasing process and PMF calculations [101, 127]. For an umbrella sampling with N_w biased windows, the optimal estimate of the unbiased distribution function is the weighted sum over all the N_w unbiased distribution functions [101, 127]

$$\langle \rho(\xi) \rangle = \sum_{i=1}^{N_w} [\langle \rho(\xi) \rangle]_{(i)}^{unbiased} \times \left[\frac{n_i \ \exp(-\beta(w_i(\xi) - F_i))}{\sum_{j=1}^{N_w} n_j \ \exp(-\beta(w_j(\xi) - F_j))} \right] \tag{2.40}$$

where n_i is the number of independent data points at each window that builds up the biased distribution function. Substituting equation 2.39 into equation 2.37, we find that the individual unbiased distribution function can be write as [101, 127]

$$\langle \rho(\xi) \rangle_{(i)}^{unbiased} = \exp(+\beta w_i(\xi)) \ \langle \rho(\xi) \rangle_{(i)}^{biased} \ \langle \exp(-\beta F_i) \rangle . \tag{2.41}$$

Plugging this into the equation 2.40, we have

$$\langle \rho(\xi) \rangle = \sum_{i=1}^{N_w} n_i \langle \rho(\xi) \rangle_{(i)}^{biased} \times \left[\sum_{j=1}^{N_w} n_j \exp(-\beta(w_j(\xi) - F_j)) \right]^{-1} \tag{2.42}$$

Assuming we have the optimal estimate of the distribution function we can write the equation 2.39 as

$$\exp(-\beta F_i) = \int d\xi \exp(-\beta w_i(\xi)) \langle \rho(\xi) \rangle \tag{2.43}$$

The WHAM equations 2.42 and 2.43 should be solved self-consistently by assigning initial values to F_is and finding an estimate for the unbiased distribution function $\langle \rho(\xi) \rangle$ using equation 2.42. By plugging the $\langle \rho(\xi) \rangle$ into the equation 2.43 we find better estimates of F_i values. The iteration between the two equations is repeated until they converge [101, 127].

2.6 Calculating the anion-cation binding constants from PMFs

Even though the formation of ion pairs in solution is not a chemical reaction, it is useful to apply the mass action formalism to describe it, and to define an equilibrium binding constant, K_{eq}, that characterizes the strength of binding between two ions.

Consider the following equilibrium, where two ions, A and B, form a non-covalently bound ion pair, AB:

$$\mathrm{A} + \mathrm{B} \rightleftharpoons \mathrm{AB} \tag{2.44}$$

$$\tag{2.45}$$

At equilibrium, the concentrations of each of the species must obey the following relation:

$$K_{eq} = \frac{[AB]}{[A][B]} \tag{2.46}$$

where K_{eq} is the equilibrium binding constant and depends on temperature onlyčiteSilber.

For simulations with many copies of A and B in a simulation box, Equation 2.46 may be applied directly to estimate the binding constant. For simulations at infinite dilution, where only single copies of A and B exist in a simulation box of volume V, however, equation 2.46 must be modified. Calculating the binding constant from simulations at infinite dilution using Equation 2.46, where the concentration of each species is obtained from the fraction of saved configurations where each species is present, is incorrect.

To derive the correct expression for the binding constant in the limit of a single pair of molecules, it is useful to re-write equation 2.46 as:

$$K_{eq} = \frac{n_{AB} V n_{Av}}{n_A n_B} \tag{2.47}$$

where n_A and n_B are the number of A or B molecules free in solution, n_{AB} is the number of AB species, n_{Av} is Avogadro's number and V is the volume of the solution. $n_A n_B$ thus defines the number of *pairs* of *free* A and B molecules. It follows that, for simulations at infinite dilution, $n_A n_B$ is simply equal to the number of saved configurations where the unbound state is observed, and n_{AB} is equal to the number of configurations where the bound state is present. Normalizing the numerator and the denominator in Equation 2.47 by the total number of saved configurations, we obtain the correct expression to calculate binding constants from simulations at infinite dilution [128]:

$$K_{eq} = V n_{Av} \frac{P_b}{P_u} \tag{2.48}$$

where P_b and P_u are the probability of being in the bound and unbound state, respectively.

We use Equation 2.48 to calculate the binding constants that characterize the interactions of the polyatomic anions under study here and Na^+, NH_4^+ and $CH_3NH_3^+$. We extract P_b and P_u from the potential of mean force curves, $\Delta G(r)$, calculated for individual anion-cation pairs. The probability of existing at a separation r is proportional to $\exp(-\beta \Delta G(r))$. It follows that P_b is the cumulative probability obtained by integrating the normalized probability density function between $r = 0$ and a given cutoff distance d, and P_u is $1 - P_b$. The volume used in equation 2.48 is that which is accessible to the molecules during the umbrella sampling simulation: it is the volume of a sphere of radius r_m, $4/3\pi r_m^3$. r_m is the maximum anion-cation separation considered, *i.e.*, the separation at which the cumulative probability is 1.

The binding constant necessarily depends on the cutoff distance d. We evaluate this dependance by calculating the binding constants assuming that the bound state of ions is only the contact ion pair configuration, and assuming also that the bound state encompasses both contact and solvent shared ion pair configurations.

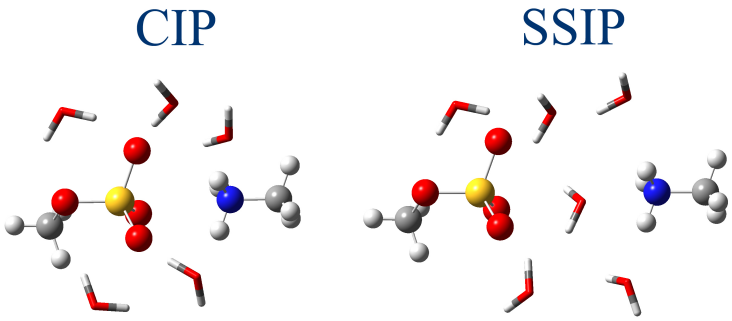

Figure 2.3 Contact ion pair(CIP) vs solvent separated ion pair(SSIP).

Chapter 3

Result & Discussion

3.1 Monoatomic ions in TIP5P water

Anion-cation interaction parameters for NaCl, KCl, $MgCl_2$ and $CaCl_2$ are optimized to reproduce the experimental activity derivative, a_{cc}^{exp}, of 0.5 molal concentration solutions. These parameters for the ion pair interactions are shown in Table 3.1. The optimized parameters are obtained by scaling the geometric average $C_{ij,GA}^{(12)}$, given by Equation 2.3, by a factor of f_{C12}:

$$C_{ij}^{(12)} = f_{C12} \times C_{ij,GA}^{(12)} \tag{3.1}$$

The results show that the geometric average is not a good estimate of anion-cation interactions.

Figure 3.1 shows the calculated activity derivatives as a function of the scaling factor f_{C12} applied to the geometric average $C_{ij,GA}^{(12)}$. For all salts, the values of a_{cc} calculated for $C_{ij,GA}^{(12)}$ lie far below the experimental target values. The low activity derivatives reflect the overly strong short-range anion-cation interactions, which are also reflected in the anion-cation rdfs shown in Figure 3.2. A sharp contact ion pair peak (CIP) peak at ≈ 3 Å as well as a broad solvent shared ion pair one (SIP) at ≈ 5 Å, where the ions are separated by a single water layer, are observed when f_{C12} is close to 1.

When f_{C12} is increased from 1, a_{cc} increases rapidly at first, but much more slowly, in the vicinity of the target values. For the example of NaCl (Figure 3.1 a), a_{cc} remains in the range of $0.89-0.93$, close to the target value $a_{cc} = 0.93$, for $f_{C12} = 1.2-2.5$. This suggests that one can choose the optimal scaling factor as the one that exactly matches the experimental value of a_{cc}, as has been done in state of the art force fields [48]. However, we observe that the choice of the scaling factor significantly changes the ion pairing patterns reflected in the rdfs shown in Figure 3.2 a for NaCl solutions in this interval of f_{C12}. Similar trends are also present at a higher concentration of 6.18 m as shown in Figure 3.3 a. For a scaling factor of $f_{C12} = 1.2$, we observe CIPs with a sharp peak around 3 Å at both low and high concentrations, while this peak vanishes for $f_{C12} = 2.5$, because the ions migrate to the second shell and form SIPs. Thus, the activity derivative remains almost constant despite parameters that lead to widely

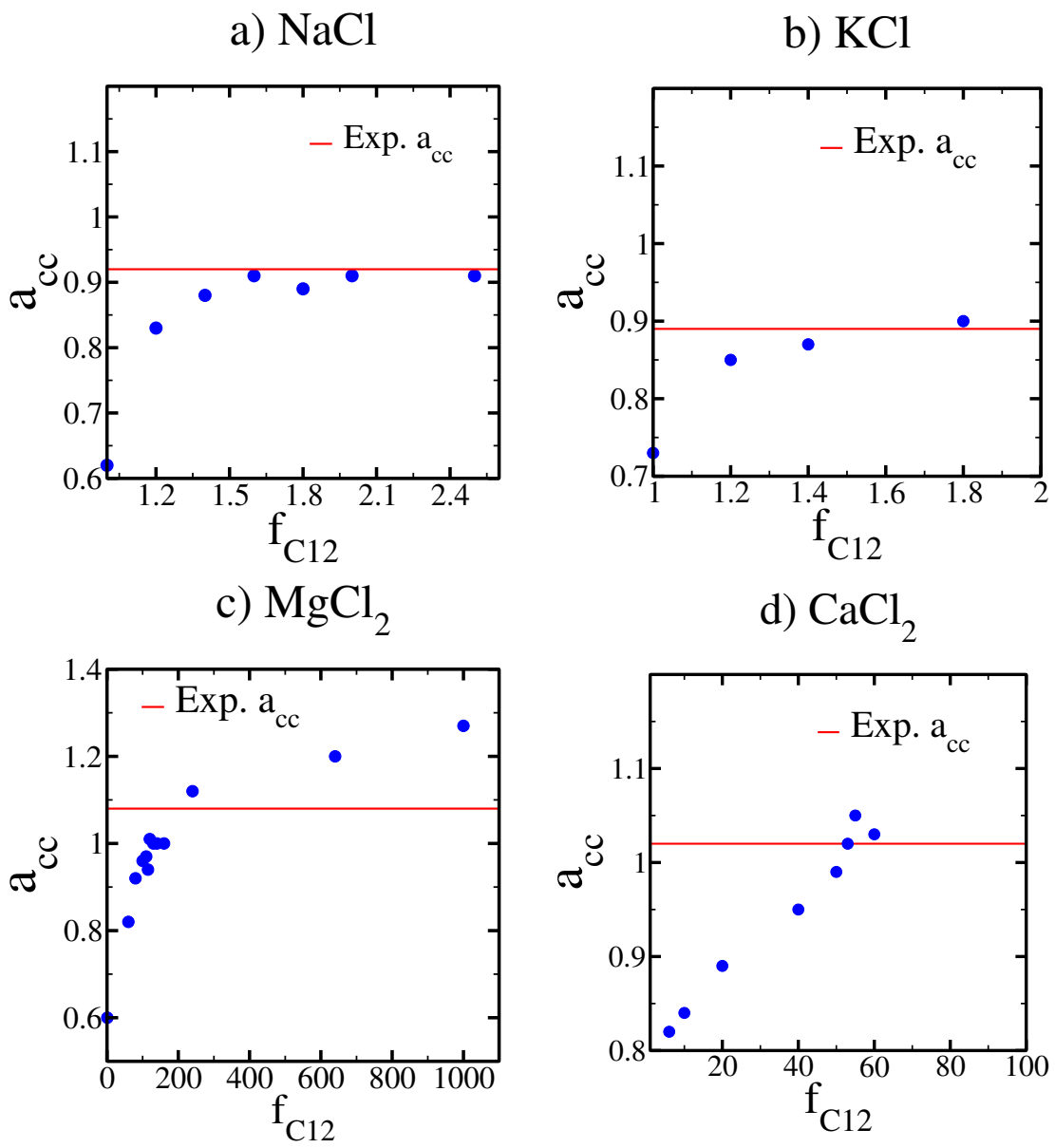

Figure 3.1 The calculated activity derivative with respect to f_{C12} for 0.5 m solutions of the indicated salts. The factor f_{C12} determines the deviation of the ion-pair $C_{ij}^{(12)}$ parameter from that obtained using a geometric average, $C_{ij,GA}^{(12)}$ (see Equation 3.1).

Table 3.1 Optimized ion-ion $C_{ij}^{(12)}$ parameters for the indicated salts; for completeness, the $C_{ij}^{(6)}$ calculated using Equation 2.2 are shown, as well as the scaling factors f_{C12}. The activity derivatives from experiment (a_{cc}^{exp}) and calculated using optimized parameters (a_{cc}^{sim}) are also given; values in parentheses are obtained using anion-cation parameters given by the geometric average (Equation 2.3).

salt	$C_{ij}^{(6)}$ [10^{-3} kJ·mol^{-1}·nm^6]	$C_{ij}^{(12)}$ [10^{-7} kJ·mol^{-1}·nm^{12}]	f_{C12}	a_{cc}^{sim}	a_{cc}^{exp}
NaCl	0.9870719	38.835145	1.6	0.91 (0.62)	0.93
KCl	3.9962779	117.888	1.2	0.85 (0.73)	0.89
$MgCl_2$	0.547327	1719.172854	240	1.12(0.60)	1.08
$CaCl_2$	2.687875	2445.089304	53	1.02(0.35)	1.02

different solution structures, and in this case should not be used as the only target property against which to parameterize anion-cation interactions. Another target property is required, to uniquely determine the optimum anion-cation interaction parameters. Here we use the coordination number C_n – the number of Cl^- ions around Na^+ ions, in contact ion pair configuration – obtained from x-ray diffraction (XRD) [129]: $C_n \approx 0.3$ at a concentration of 6.18 m. We find that, for $f_{C12} = 1.6$, the computational C_n is 0.32 at this concentration, as shown in Figure 3.3. The calculated activity derivative for this scaling factor is 0.91, in agreement with experiment. It is worth noting, that contrary to what was reported in other computational studies [48, 130], at low NaCl concentration (0.5 m), we observe almost no CIPs ($C_n = 0.01$) and most ions form SIPs ($C_n = 0.45$), using our optimized parameters.

Similar behavior is found for KCl solutions. The final scaling factor $f_{C12} = 1.2$ reproduces both the low concentration (0.5 m) activity derivative, which is 0.85 as compared with a target experimental value of 0.89, and C_n at high concentration (4.5 m), which is 0.55 and very close to the experimental value $C_n = 0.6$. At low concentration, the KCl solution shows few CIPs, at an anion-cation distance ≈ 3 Å ($C_n \approx 0.1$), but abundant SIPs at ≈ 5 Å ($C_n \approx 0.5$), indicating that only 10% of the Cl^- ions are in direct contact with K^+ ions. Larger scaling factors would also satisfy the experimental activity derivative, as shown in Figure 3.1 b but would lead to no contact ion pairs at low and high concentrations (Figures 3.2 b and 3.3 b), in disagreement with experiment. These results for KCl again emphasize that the activity derivative may not be sufficient to optimize the ionic interactions in some salts.

Others have also observed plateaus in the activity derivative as a function of anion-cation interaction parameters, and these plateaus have been similarly linked to interconversion between different types of ion pairs [61]. The importance of lifting the indetermination in the parameters introduced by these plateaus by comparison with other properties that reflect ion-ion interactions, however, has not always been recognized. Instead, in some instances authors have arbitrarily chosen a particular

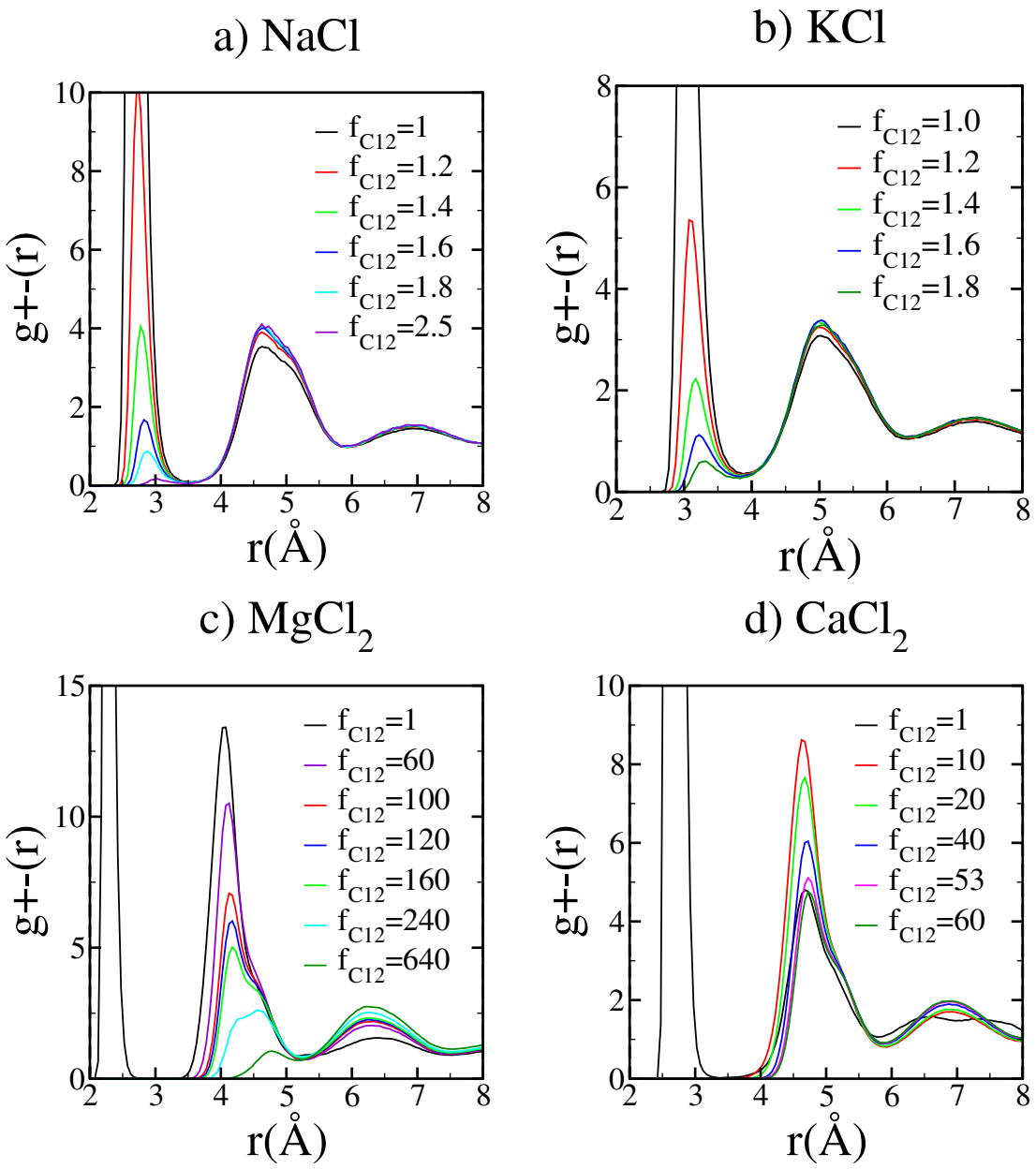

Figure 3.2 Anion-cation radial distribution function for 0.5 m solutions of the indicated salts. The factor f_{C12} determines the deviation of the ion-pair $C_{ij}^{(12)}$ parameter from the geometric average $C_{ij,GA}^{(12)}$ (see Equation 3.1).

scaling factor without physically justifying this choice (*e.g.*, parameters for $CaCl_2$ and $CaBr_2$ in ref. [61]). Our results demonstrate that this arbitrary choice can be easily avoided by the introduction of other target properties, and results in a more reliable description of solution structure.

For the divalent salts, $MgCl_2$ and $CaCl_2$, the scaling factors for $MgCl_2$ and $CaCl_2$ that reproduce the experimental activity derivatives are much larger, $f_{C12} = 240$ and $f_{C12} = 53$, respectively, than for the 1:1 salts. Neither of the salt solutions show CIPs

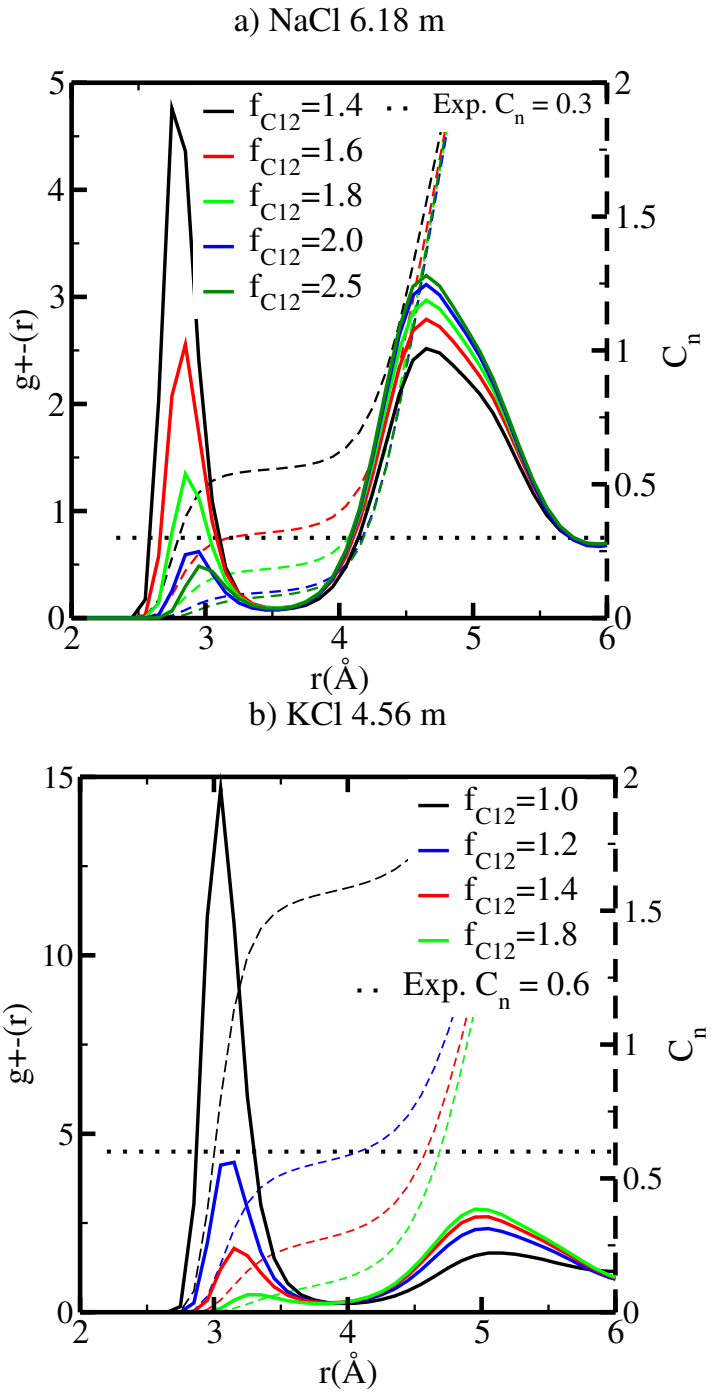

Figure 3.3 Anion-cation radial distribution function, left axis (solid-lines), and coordination number C_n, right axis(dashed-lines), for high concentration solutions of the indicated salts. The factor f_{C12} determines the deviation of the ion-pair $C_{ij}^{(12)}$ parameter from the geometric average $C_{ij,GA}^{(12)}$ (see Equation 3.1).

at 0.5 m, in agreement with the x-ray absorption fine spectroscopy (XAFS)/XRD data [131]. The activity derivative *v.s.* f_{C12} plots for $CaCl_2$ shows a pseudo-plateau in the

vicinity of the experimental target value; however, in contrast to our observations on NaCl and KCl solutions described above, the ionic distribution depends only weakly on the scaling factor in this region. Therefore good parameters can be obtained by using the activity derivative as the only target property for parameterization.

3.1.0.1 The effect of sampling time

To ensure that we have sufficient sampling time for the calculation of the activity derivatives of the salt solutions, we obtained the corresponding activity derivatives using a variety of time intervals, see Table 3.2. The results show that the activity derivative obtained from radial distribution functions averaged over shorter time intervals are very similar to those obtained using data from the entire simulation. These results confirm that the 150 ns of sampling is sufficient to yield reliable results.

Table 3.2 Activity derivative of salt solutions in different time scales.

Ion	Sampling Time ns	a_{cc}
	0-50	0.89
NaCl	50-100	0.91
f=1.6	100-150	0.92
	0-150	0.91
	0-50	0.84
KCl	50-100	0.88
f=1.2	100-150	0.84
	0-150	0.85
	0-50	1.05
$CaCl_2$	50-100	1.02
f=53	100-150	0.98
	0-150	1.02
	0-50	1.12
$MgCl_2$	50-100	1.13
f=240	100-150	1.10
	0-150	1.12

3.2 Polyatomic ions in TIP3P water

The optimized parameters are presented in the Appendix B, in .top/.itp format for gromacs users; the original values are shown as comments, to facilitate comparisons. Amber-format parameters are given in the same files in the form of comments.

3.2.1 Determining self-interaction parameters for the oxygens of the anions

3.2.1.1 Parameters developed based on solvation free energies

The computational solvation free energies (SolvFEs) are calculated as described in section 2.2.3. We attempt to reproduce the experimental target data ($\Delta\Delta G_{solv}^{Exp}$; see equation 2.6) by scanning a range of ε_{O2-O2} or σ_{O2-O2} with values between $f = 0.8$ and $f = 1.2$ times the original GAFF [90, 91, 36, 92] parameters; for HSO_4^-, scanning of the ε_{OH-OH} or σ_{OH-OH} parameters would in principle also be necessary but in practice was not done because the original GAFF [90, 91, 36, 92] parameters proved adequate.

Given that only one target property was selected for optimization, an infinite number of combinations of ε and σ values could yield equally good agreement with the target, possibly at the expense of the performance of the model regarding other observables. Because the original GAFF parameters are derived to reproduce the density and heats of vaporization of organic liquids [90, 91, 36, 92], we choose to retain them for either ε or σ, and optimize the remaining parameter only. Preliminary results (not shown) indicate that the free energy of hydration is more sensitive to changes in σ than to changes in ε. For this reason, we choose to optimize σ for the cases where the original GAFF [90, 91, 36, 92] parameters led to deviations of tens of kcal/mol from the target hydration free energies, but we optimize ε for the cases where the deviation was lower than 10 kcal/mol.

For each anion for which experimental values of the hydration free energies exist, we plot the calculated values of the solvation free energy difference, $\Delta\Delta G_{solv}^{Comp}$, as a function of the scaling factor, f, applied to ε_{O2-O2} or σ_{O2-O2} parameters, and fit the result with a linear function. The optimal ε_{O2-O2} or σ_{O2-O2} parameters are found by solving the linear function for the experimental SolvFE difference, $\Delta\Delta G_{solv}^{Exp}$. The procedure is illustrated in Figure 3.4 for SO_4^{2-}; the strongly linear dependence of $\Delta\Delta G_{solv}^{Comp}$ on σ_{O2O2} is apparent.

The optimized parameters for HSO_4^-, SO_4^{2-} and CH_3COO^-, as well as their experimental and computational SolvFEs, are given in Table 3.3. We recommend that the parameters given in Table 3.3 for O2 are used for the interactions of the anions with water and with other species in solution using combination rules (equation 2.5) for all cases except those which are specifically parameterized.

The GAFF [90, 91, 36, 92] ε and σ parameters lead to a -0.6 kcal/mol deviation from the target value for HSO_4^-. Due to the uncertainty associated with measurements of the solvation free energy of single ions[59], the experimental error in measuring the SolvFE of HSO_4^- is at least of the order of this deviation. Therefore we do not

Table 3.3 Self-interaction parameters for O2 derived against experimental solvation free energies. The parameters are expressed in terms of a scaling factor relative to the GAFF values (shown in the Supporting Information).

Anion	**Parameters**	**n**×$\Delta G_{solv}^{Comp}(Cl^-)$ kcal/mol	$\Delta G_{solv}^{Comp}(A^-n)$ kcal/mol	$\Delta\Delta G_{solv}^{Comp}$ kcal/mol	$\Delta\Delta G_{solv}^{Exp}$ kcal/mol	**Deviation**[d] kcal/mol
HSO_4^-	..[b]	1×(-87.3)	-85.1	2.2	2.8[c]	-0.6
CH_3COO^-	$0.77 \times \varepsilon_{O2-O2}$	1×(-87.3)	-93.0	-5.7(-3.9)[a]	-6.2[c]	0.5 (+2.3)[a]
SO_4^{2-}	$1.17 \times \sigma_{O2-O2}$	2×(-87.3)	-270.0	-95.4 (-122.5)[a]	-94.6[c]	-0.8 (-27.9)[a]

[a]GAFF parameters; [b] No scaling; [c] reference [67]; [d] $\Delta\Delta G_{solv}^{Comp}$ - $\Delta\Delta G_{solv}^{Exp}$; $\Delta\Delta G_{solv}$ is defined in Equation 2.6

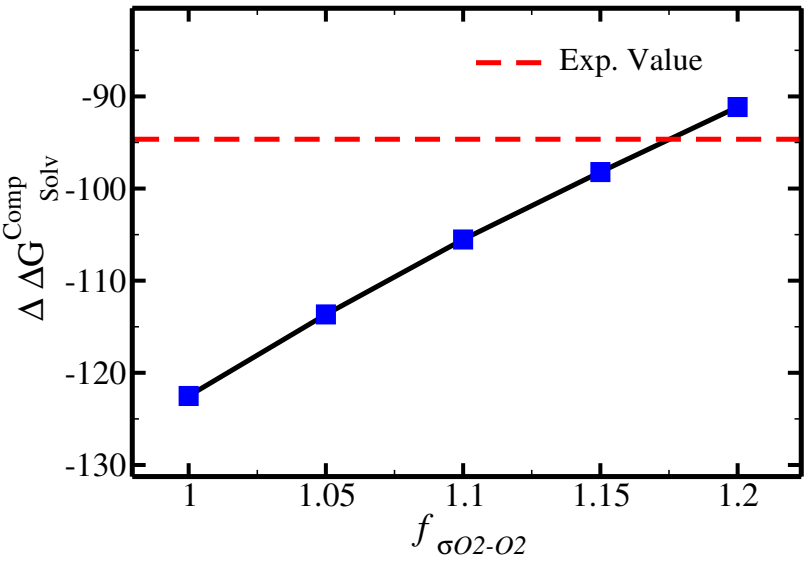

Figure 3.4 Difference in solvation free energy, $\Delta\Delta G_{solv}^{Comp} = \Delta G_{solv}^{Comp}(SO_4^{2-}) - 2 \times \Delta G_{solv}^{Comp}(Cl^-)$, as a function of the scaling factor $f_{\sigma_{O2-O2}}$ for the SO_4^{2-} anion. The red line shows the target experimental value, $\Delta\Delta G_{solv}^{Exp}$.

further parameterize HSO_4^- against the SolvFE and we use the GAFF [90, 91, 36, 92] Lennard-Jones parameters for this anion.

For acetate (CH_3COO^-), the original ε_{O2-O2} needs to be scaled by a factor of 0.77 to retrieve the experimental SolvFE for this ion. The optimized parameter thus reproduces the difference in hydration free energies between Cl^- and CH_3COO^- better than the GAFF [90, 91, 36, 92] parameters. We note, however, that the deviation from experiment using the GAFF [90, 91, 36, 92] parameters is only +2.3 kcal/mol, so using GAFF parameters to investigate the interactions of this anion with water is acceptable.

In contrast, GAFF [90, 91, 36, 92] parameters yield a far too low hydration free energy for SO_4^{2-}: the deviation from experiment is −27.9 kcal/mol. A scaling factor of 1.17 to the original σ_{O2-O2} parameter is required to reproduce the target $\Delta\Delta G_{solv}^{Exp}$

of the sulfate. GAFF [90, 91, 36, 92] parameters dramatically fail to reproduce the SolvFE of multivalent ions, emphasizing the importance of parameter optimization with respect to SolvFE for those anions.

3.2.1.2 Parameters developed using ab initio calculations

To optimize the oxygen parameters for $CH_3SO_4^-$, $CH_3SO_3^-$, HPO_4^{2-} and $CH_3PO_4^{2-}$, for which reliable hydration free energies measured experimentally could not be found in the literature, we use the ab initio approach described in the Introduction.

3.2.1.2.1 Validation of approach for ion-water systems. As described generically in the introduction, this approach is based on the assumption that, for similar anions, the difference, $\Delta E_{QM\to cl}$, between the energy minimum of anion-water dimers calculated using ab initio and the equivalent energy calculated using optimized classical parameters should be similar for anions of similar charge and structure (see Figure 1.2). To evaluate whether this assumption holds, we perform potential energy scans as a function of anion-water distance for the CH_3COO^--H_2O and the HSO_4^--H_2O systems. These are the only systems containing comparable anions for which parameters were optimized based on the free energies of hydration. The scans are performed using ab initio and using the optimized parameters, in the system configuration shown in Figure 2.2a; these scans are shown in Figure 3.5 . We find (see Table 3.4) that $\Delta E_{QM\to cl} = -1.3$ kcal/mol for HSO_4^-, and $\Delta E_{QM\to cl} = -1.2$ kcal/mol for CH_3COO^-; the similarity in the $\Delta E_{AB,QM\to cl}$ values for these two anions confirms that the ab initio approach proposed here holds for anion-water dimers.

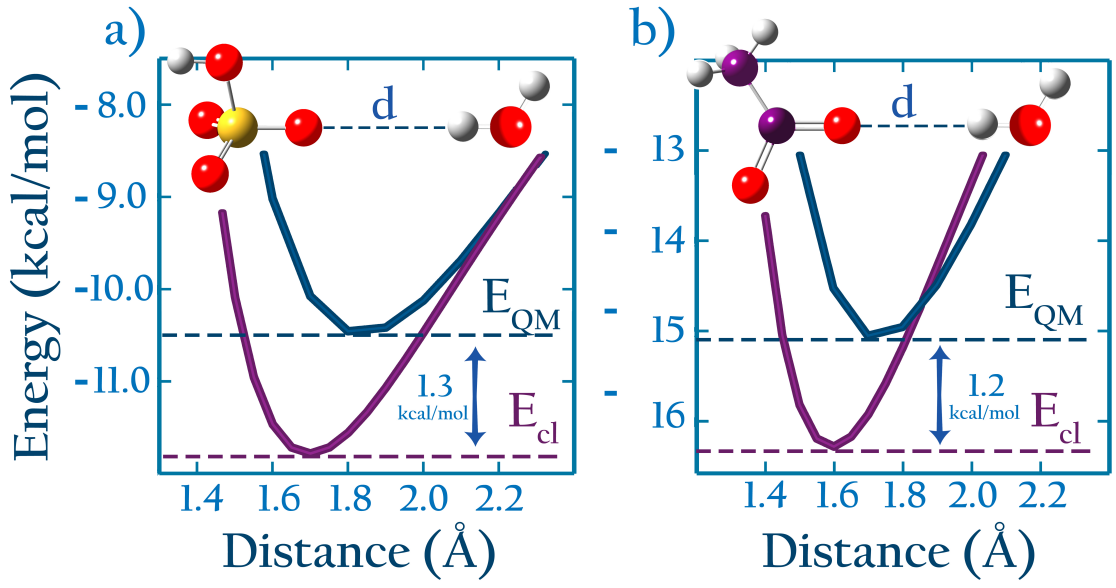

Figure 3.5 a)HSO_4^--H_2O b)CH_3COO^--H_2O, 1D energy scan with both QM and classical methods. The curves in (a) are averaged over scans of different symmetries, as described in the Appendix A.

The scans also illustrate that developing parameters by requiring that the classical potential energy curve overlaps with the ab initio is unsuitable. Those parameters would greatly overestimate anion-water interactions.

Table 3.4 Differences between the classical and quantum potential energy scans for all anion-water dimers, and values of the optimized anion-water parameters.

Anion	Optimization[c]	Parameters[d]	$\Delta \mathbf{E}^{e}_{QM\rightarrow cl}$ kcal/mol	$\Delta \mathbf{R}^{f}_{QM\rightarrow cl}$ Å
HSO_4^-	Exp	.. [b]	-1.3[a]	-0.14[a]
$CH_3SO_4^-$	Comp	.. [b]	-1.4[a]	-0.15[a]
$CH_3SO_3^-$	Comp	.. [b]	-1.7[a]	-0.15[a]
$H_2PO_4^-$	Comp	.. [b]	-1.4[a]	-0.12[a]
$(CH_3)_2PO_4^-$	Comp	.. [b]	-1.1[a]	-0.11[a]
CH_3COO^-	Exp	$0.77 \times \varepsilon_{\mathcal{OO}}$	-1.2(-0.8[a])	-0.14(-0.10[a])
SO_4^{2-}	Exp	$1.17 \times \sigma_{\mathcal{OO}}$	4.8(1.0[a])	0.22(-0.06[a])
HPO_4^{2-}	Comp	$1.135 \times \sigma_{\mathcal{OO}}$	4.8(1.1[a])	0.16(-0.05[a])
$CH_3PO_4^{2-}$	Comp	$1.135 \times \sigma_{\mathcal{OO}}$	4.8(1.3[a])	0.16(-0.05[a])

[a] GAFF parameters; [b] No scaling; [c] Parameters optimized based on hydration free energies (Exp), repeated here to facilitate comparisons, or based on the ab initio approach (Comp); [d] Expressed as a scaling factor relative to the original GAFF parameters (given in the Supporting Information); [e] $\Delta E_{QM\rightarrow cl}$ is defined in Equation 1.2; [f] $\Delta R_{QM\rightarrow cl}$: the difference between the position of the dimer energy minimum obtained using ab initio calculations and using classical parameters.

3.2.1.2.2 Parameter optimization. For each of the $CH_3SO_4^-$, $CH_3SO_3^-$, HPO_4^{2-} and $CH_3PO_4^{2-}$ anions, we perform an ab initio gas phase potential energy scan of the anion-water dimer, as described in section 2.2.6.3. For parameter optimization, we require the same type of potential energy scan to be performed for a reference species for which classical parameters already exist: we use SO_4^{2-} as the reference compound for $CH_3PO_4^{2-}$ and HPO_4^{2-}, and HSO_4^- as the reference compound for $CH_3SO_4^-$, $CH_3SO_3^-$, $H_2PO_4^-$ and $(CH_3)_2PO_4^-$ because of their similarities in total charge and structure. For the reference species, we calculate also the analogous potential energy scans using the classical parameters previously derived using hydration free energies. Finally, we optimize the oxygen parameters for the anions by testing many $\varepsilon_{\mathcal{OO}}$ or $\sigma_{\mathcal{OO}}$ values until the value that best fulfills the condition indicated in Equation 1.1 is found.

The optimized self-interaction parameters for the oxygens in all the anions are shown in Table 3.4. The same Table also shows $\Delta R_{QM\rightarrow cl}$, the difference in the position of the potential energy minimum calculated using ab initio and using classical parameters. For the divalent anions, the optimized parameters result in larger anion-water separation at the energy minimum of the dimer than the original GAFF [90, 91, 36, 92] parameters. This increase has been observed in other parameterizations of multivalent oxoanions based on experimental data [89], and illustrates the inadequacy

of applying GAFF [90, 91, 36, 92] parameters, derived for neutral molecules, to heavily charged species.

The original GAFF parameters for $CH_3SO_4^-$ yield $\Delta E_{QM\rightarrow cl} = -1.4$ kcal/mol; for $CH_3SO_3^-$, $H_2PO_4^-$ and $(CH_3)_2PO_4^-$ they yield $\Delta E_{QM\rightarrow cl} = (-1.7, -1.4, -1.1)$ kcal/mol (see Table 3.4), respectively. Tuning the oxygen parameters of these anions to reproduce the target $\Delta E_{QM\rightarrow cl} = -1.3$ kcal/mol obtained for HSO_4^- leads to changes in hydration free energy lower than 2 kcal/mol. Because this difference is slight, we opt to retain the original GAFF [90, 91, 36, 92] parameters for $CH_3SO_4^-$, $H_2PO_4^-$ and $CH_3SO_3^-$.

The original GAFF [90, 91, 36, 92] parameters lead to a $\Delta E_{QM\rightarrow cl} = +1.1$ kcal/mol for HPO_4^{2-}, and to $\Delta E_{QM\rightarrow cl} = +1.3$ kcal/mol for $CH_3PO_4^{2-}$. Because these values differ markedly from the target $\Delta E_{QM\rightarrow cl} = +4.8$ kcal/mol obtained for SO_4^{2-}, the oxygens of the two divalent ions are optimized to reproduce the target $\Delta E_{QM\rightarrow cl}$. The value of $\Delta E_{QM\rightarrow cl}$ depends predominantly on the parameters of the O2 oxygens and is essentially independent of the parameters of the other oxygens; however, for consistency, the σ parameter of the OH oxygens in HPO_4^{2-} is scaled by the same prefactor used to scale σ_{O2}; following the same logic, the σ parameter of the OS oxygens in $CH_3PO_4^{2-}$ is also scaled by the same scaling factor applied to σ_{O2} of that species. Applying these scaling factors to OH or OS is recommended, despite the fact that they should lead to relatively small changes in the overall interaction between these species and water; *e.g.*, scaling or not the OH oxygen in HPO_4^{2-} leads to changes of $\sim$2 kcal/mol in SolvFE.

The calculated uncorrected ($\Delta G_{cav}+\Delta G_{chg}$) SolvFE of HPO_4^{2-} are -311.8 kcal/mol and -285.8 kcal/mol, for GAFF [90, 91, 36, 92] and our parameters, respectively. This large change in solvation free energy illustrates the critical importance of optimizing ion-water interactions for divalent anions. We do not compare these results to the only available experimental value for SolvFE of HPO_4^{2-} because that value has very high associated uncertainty [132]: that value was not directly measured, but instead was calculated based on rough estimates of lattice energies.

3.2.2 Developing specific anion-cation parameters

3.2.2.1 Parameters developed based on solution activity derivatives

After finding optimal self-interaction parameters for the anions based on ion-water interactions, we evaluate whether those parameters can be used with the combination rules shown in Equation 2.5 to obtain reasonable interactions between the anions and two cations of interest: the commonly used counterion Na^+, and an analogue, NH_4^+, of the terminal side chain group of the amino acid lysine. We assess the quality of those parameters by calculating, as described in section 3.5, the activity derivatives of 0.5 m solutions of salts for which experimental solution activities exist. These results are given in Table 3.5, under column $a_{cc}^{Calc,b}$ and are referred to in the text as arising from using unoptimized anion-cation parameters. For comparison, we also calculate activity derivatives using exclusively the original set of GAFF [90, 91, 36, 92] parameters for the anions (denoted as $a_{cc}^{Calc,c}$ in Table 3.5).

Table 3.5 Solution activity derivatives from experiment (a_{cc}^{Exp}) and from simulation (a_{cc}^{Calc}). The optimal anion-cation parameters are expressed in terms of the values calculated from Equation 2.5 using the original GAFF values, to facilitate comparisons.

Anion	Cation	Parameters	$a_{cc}^{Calc,a}$	$a_{cc}^{Calc,b}$	$a_{cc}^{Calc,c}$	$a_{cc}^{Exp,d}$
$CH_3SO_3^-$	Na^+	$1.1 \times \varepsilon_{\mathcal{O}Na}$	0.95	0.89	0.89	0.93
	NH_4^+	$1.1 \times \{ \begin{matrix} \varepsilon_{\mathcal{O}N} \\ \varepsilon_{\mathcal{O}H_N} \end{matrix}$	0.89	0.85	0.85	0.90
CH_3COO^-	Na^+	$1.022 \times \sigma_{\mathcal{O}Na}$	0.99	0.35	0.22	1.00
SO_4^{2-}	Na^+	$1.13 \times \sigma_{\mathcal{O}Na}$	0.64	*	*	0.62±0.02
	NH_4^+	$1.13 \times \{ \begin{matrix} \sigma_{\mathcal{O}N} \\ \sigma_{\mathcal{O}H_N} \end{matrix}$	0.60	*	*	0.62±0.02

* Could not be calculated because of aggregation; [a] Optimized $\sigma_{\mathcal{OO}}$, $\varepsilon_{\mathcal{OO}}$ and anion-cation parameters; [b] Optimized $\sigma_{\mathcal{OO}}$ and $\varepsilon_{\mathcal{OO}}$, with anion-cation parameters derived from combination rules (for some salts, this set of parameters coincides with the original GAFF parameters); [c] Original GAFF parameters; [d] reference [133]

The unoptimized anion-cation parameters fail, often dramatically, to adequately describe anion-cation interactions in all cases; the original set of GAFF [90, 91, 36, 92] parameters also fails. These failures point to the need of specifically optimizing the anion-cation interactions. Large modifications to the anion-cation parameters (see Table 3.5) are indispensable to attain agreement with experiment for solutions of Na_2SO_4 or $(NH_4)_2SO_4$: for both salts, using either unoptimized anion-cation parameters or using the original GAFF [90, 91, 36, 92] parameters led to dramatic aggregation, as illustrated in Figure 3.6a for Na_2SO_4. Such large aggregation is clearly unphysical, because our simulations are done at 0.5 m, far below the solubility limit of Na_2SO_4 (21.94 g/100 g of solution) [134], and of $(NH_4)_2SO_4$ (43.3 g/100 g of solution) [135, 80]. As shown in Figure 3.7, the optimized anion-cation parameters lead to very few contact ion-pairs formed per ion, with both ions forming primarily solvent shared ion-pairs, *i.e.*, configurations where the ions share their first hydration layer.

For $NaCH_3COO$, using unoptimized anion-cation parameters leads to a solution activity derivative, $a_{cc}^{Calc} = 0.35$, which is far lower than the target value of $a_{cc}^{Exp} = 1.00$; the original GAFF [90, 91, 36, 92] parameters perform even worse, yielding $a_{cc}^{Calc} = 0.22$. To achieve good agreement with experiment, it is necessary to modify $\sigma_{\mathcal{O}Na}$. The optimized parameters lead to a number of contact ion-pairs forming per ion that is only 1/5 of that obtained using the original GAFF [90, 91, 36, 92] parameters, as shown in Figure 3.8b. Our results thus indicate that the interactions between Na^+ and acidic amino acids are markedly overestimated when using the sodium model of Joung and Cheatham [44] in the AMBER [36] force field.

For $NaCH_3SO_3$ and $NH_4CH_3SO_3$, the activity derivative calculated without optimizing anion-cation interactions differs less than 5% from experiment. Despite this seemingly small difference, we provide optimized anion-cation parameters for these salts (see Table 3.5) because they lead to quantitative differences in the number of ion-pairs: Figure 3.7 show that the optimized anion-cation parameters lead to 25%

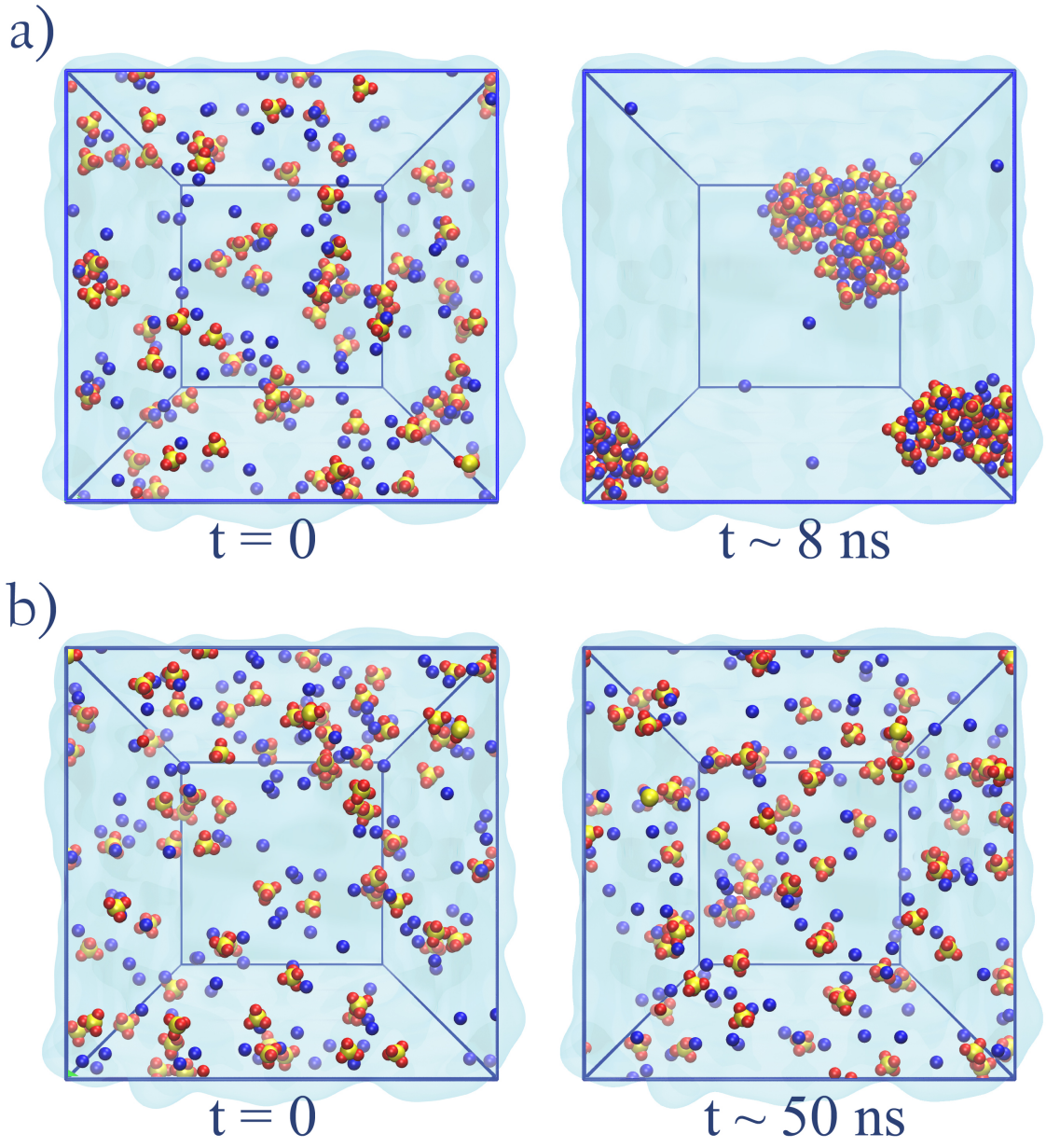

Figure 3.6 Simulations of Na_2SO_4. a) Original GAFF parameters: aggregation becomes visible at 500 ps and large aggregates are seen at 8 ns. b) Optimized anion-cation parameters: aggregation is not observed within the time scope of the simulation.

fewer contact ion-pairs – defined as ions in direct contact with each other – than the non-optimized ones.

3.2.2.2 Parameters developed using ab initio calculations

For several of the salts of interest here, the solution activities are not available in the literature or have high associated uncertainty. High uncertainty is often found for salts

of amphiprotic species like phosphate, because solutions of those salts simultaneously contain $H_2PO_4^-$ and HPO_4^{2-}, making it difficult to deconvolute the contribution of each species to the solution activity [133]. For these salts, we optimize anion-cation interactions again following the approach described in the introduction and using the optimized anion-cation parameters developed in the previous section.

3.2.2.2.1 Determining optimal anion-cation orientations to be used in gas-phase potential energy scans. In ab initio calculations, contact ion-pairs in the vacuum benefit from extra delocalization stability due to the overlap of cation and anion orbitals; this extra stabilization is sometimes referred to as intermolecular hyperconjugation energy [136]. However, quantum effects such as electron delocalization and charge transfer between two ions are not expected to be dominant in aqueous solution; in this case the charge transfer preferentially occurs between the ions and the water molecules in the first hydration shell. [137, 138, 139]

The fact that quantum effects stabilize ion-ion interactions in the vacuum but not in solution represents a problem when using our approach to develop classical parameters for ions in aqueous solution. Given that the magnitude of the hyperconjugation energy varies amongst contact ion-pairs – depending on the electronic structure and the relative orientation of the two ions – our ab initio approach might introduce artifacts if the two ion-pairs being compared benefit from different hyperconjugation stabilities. To resolve this issue, we must ensure that the anions for which we want to make the ab initio comparison have similar interactions with the same cation. We assess the similarity of interactions by performing natural bond orbital (NBO)[136] analysis. This analysis proved to well-capture the electron density transfers,[140, 141, 142] on the geometries shown in Figure 2.2a,b,c and at distances corresponding to the minimum energy on the 1-D potential curves for each ion-pair. To perform this analysis we use the geometries from MP2 calculations and perform B3LYP[109, 110, 111]/aug-cc-PVTZ calculations to obtain the electron densities. Using the B3LYP method is necessary to obtain the well-defined one-electron densities required by NBO calculations.[140].

The results of the NBO analysis are shown in Table 3.6. In ion-pairs with Na^+, the electron density transfers from the lone pair (LP) of the oxygen in the anion to the unfilled valence-shell (LP*) in Na^+. The hyperconjugation energies vary in a very small range, between 6.5-10.0 kcal/mol for monovalent and between 11.0-11.5 kcal/mol for divalent anions, suggesting that the donor-acceptor interactions are similar for all ion-pairs involving Na^+. Therefore, the configuration shown in Figure 2.2b can be used for the ab initio calculations needed to derive optimized anion-cation parameters.

Table 3.6 Second Order Perturbation Theory Analysis of Fock Matrix in NBO Basis for anion-cation pairs

Ion-Pair	Donor(i)	Acceptor (j)	E(2)† kcal/mol	Sum
Anion-Na^+				
$CH_3SO_3^- \cdots Na^+$	LP(1-3)O	LP*(1-3)Na	7.5	
$CH_3SO_4^- \cdots Na^+$	LP(1-3)O	LP*(1-3)Na	6.5	
$CH_3COO^- \cdots Na^+$	LP(1-3)O	LP*(1-3)Na	10.0	
$H_2PO_4^- \cdots Na^+$	LP(1-3)O	LP*(1-3)Na	7.0	
$(CH_3)_2PO_4^- \cdots Na^+$	LP(1-3)O	LP*(1-3)Na	7.0	
$SO_4^{2-} \cdots Na^+$	LP(1-3)O	LP*(1-3)Na	11.5	
$CH_3PO_4^{2-} \cdots Na^+$	LP(1-3)O	LP*(1-3)Na	11.0	
Anion-NH_4^+: Orientation I				
$CH_3SO_3^- \cdots NH_4^+$	LP(2)O	BD*(1)H-N	48.0	88.0
	BD*(2)S-O	BD*(1)H-N	40.0	
$CH_3SO_4^- \cdots NH_4^+$	LP(2)O	BD*(1)H-N	45.0	94.0
	BD*(2)S-O	BD*(1)H-N	49.0	
$CH_3COO^- \cdots NH_4^+$	LP(2)O	BD*(1)H-N	51.0	
$SO_4^{2-} \cdots NH_4^+$	LP(2)O	BD*(1)H-N	77.0	130.0
	BD*(2)S-O	BD*(1)H-N	53.0	
$CH_3PO_4^{2-} \cdots NH_4^+$	LP(2)O	BD*(1)H-N	85.5	233.5
	BD*(2)S-O	BD*(1)H-N	148.0	
Anion-NH_4^+: Orientation II				
$CH_3SO_3^- \cdots NH_4^+$	LP(1)O	BD*(1)N-H(1-3)	4.5	
$CH_3SO_4^- \cdots NH_4^+$	LP(2)O	BD*(1)N-H(1-3)	4.5	
$CH_3COO^- \cdots NH_4^+$	LP(2)O	BD*(1)N-H(1-3)	6.0	
$H_2PO_4^- \cdots NH_4^+$	LP(1)O	BD*(1)N-H(1-3)	5.0	
$(CH_3)_2PO_4^- \cdots NH_4^+$	LP(1)O	BD*(1)N-H(1-3)	5.0	
$SO_4^{2-} \cdots NH_4^+$	LP(2)O	BD*(1)N-H(1-3)	8.0	
$CH_3PO_4^{2-} \cdots NH_4^+$	LP(2)O	BD*(1)N-H(1-3)	8.5	

LP = Lone Pair
LP*= Unfilled Valence-Shell
BD*= Valence Anti-bonding
† Stabilization energy E(2) estimated as $E(2)=\Delta E_{ij}= q_i \frac{F(i,j)^2}{\varepsilon_j-\varepsilon_i}$, where q_i is the donor orbital occupancy, ε_i, ε_j are diagonal elements (orbital energies) and F(i,j) is the off-diagonal NBO Fock matrix element.[143]

In contrast to the ion-pairs involving Na^+, those involving NH_4^+ in configuration I (see Figure 2.2c) exhibit two different electron density transfers, one from the LP of the oxygen to the valence antibonding (BD*) orbital of (H-N) and the other from BD* of (S-O) to BD* of (H-N). The total hyperconjugation energies lie between 51.0-94.0 kcal/mol for monovalent and 130.0-283.5 kcal/mol for divalent ions. This wider range of hyperconjugation energies arises from different donor-acceptor interactions in these ion-pairs, making impossible the type of comparison necessary for our parameterization approach. To resolve this problem we searched for different relative orientations of the ions involved in NH_4^+ ion pairs, to find an orientation in which hyperconjugation energies are both minimized and are similar for the various ion pairs. For orientation II (See Figure 2.2d), the the hyper-conjugation energies are in the narrow range of 4.5-6.0 kcal/mol for monovalent anions and 8.0-8.5 kcal/mol for divalent anions. We use this configuration to find the parameters for the ion-pairs containing ammonium and methyl ammonium.

3.2.2.2.2 Validation of approach for anion-cation systems. Similarly to what we describe above to validate the approach for anion-water systems, we perform potential energy scans as a function of the anion-cation distance for the $CH_3SO_3^- \cdots Na^+$ and the $CH_3COO^- \cdots Na^+$ systems, the only ones that are comparable and for which anion-cation interaction parameters based on experimental data exist. The potential energy scans (not shown) are done using quantum mechanics and using the classical models with optimized parameters, in the configuration shown in Figure 2.2b. We find, as shown in Table 3.7, that $\Delta E_{QM \to cl} = +5.1$ kcal/mol for $CH_3SO_3^-$ and $\Delta E_{QM \to cl} = +5.3$ kcal/mol for CH_3COO^-. The two values are very similar, confirming that the approach holds also for anion-cation systems.

3.2.2.2.3 Parameter optimization. For parameter optimization, we require potential energy scans for reference species for which classical parameters already exist: we use $NaCH_3SO_3$ as the reference for $NaCH_3SO_4$, NaH_2PO_4 and $Na(CH_3)_2PO_4$; Na_2SO_4 as the reference for $Na_2CH_3PO_4$; $NH_4CH_3SO_3$ as the reference for NH_4^+ or $CH_3NH_3^+$ salts of all monovalent anions; $(NH_4)_2SO_4$ as the reference for NH_4^+ or $CH_3NH_3^+$ salts of all divalent anions. The Lennard-Jones parameters describing the interactions of the oxygen atoms of the anions with the N and H_N atoms in NH_4^+ and $CH_3NH_3^+$ and with Na^+ are optimized by testing many values until the one that best fulfills the condition indicated in Equation 1.1 to within 0.5 kcal/mol is found. This procedure is similar to that which was followed to optimize water-ion interactions.

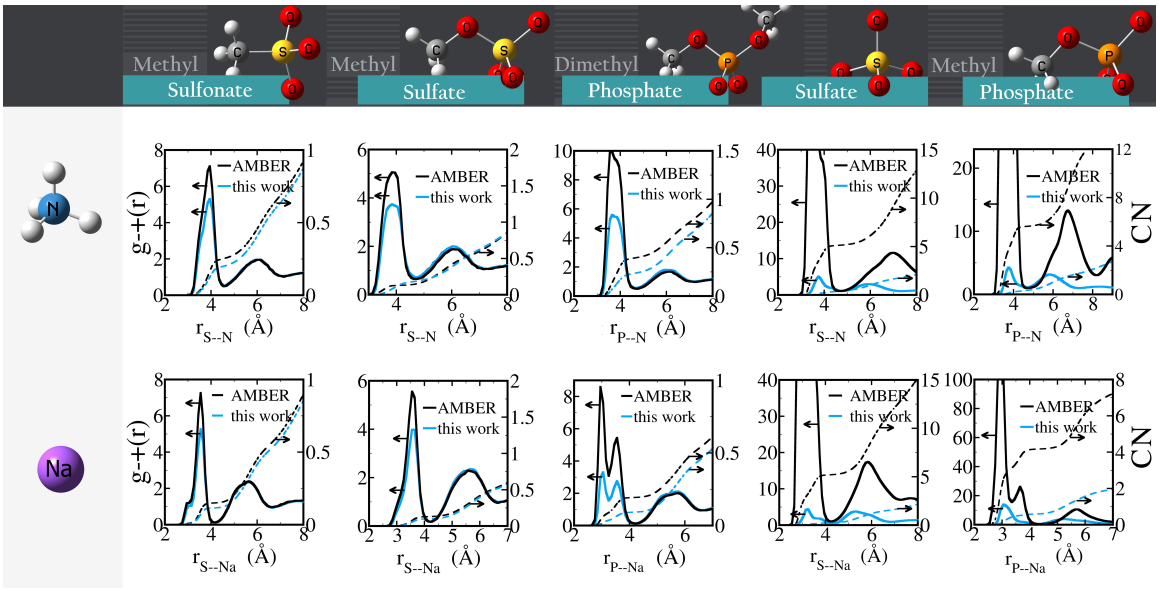

Figure 3.7 Radial distribution functions of anion-cation pairs, $g-+(r)$ (solid lines, left y-axes) and corresponding number integral (CN; dashed lines, right y-axes) using AMBER/GAFF parameters and using the optimized parameters.

Table 3.7 Differences between the classical and quantum potential energy scans for all anion-Na^+ dimers, and values of the optimized anion-Na^+ parameters.

Anion	Optimization[b]	Parameters[c]	$\Delta E^d_{QM\rightarrow cl}$ kcal/mol	$\Delta R^e_{QM\rightarrow cl}$ Å
CH_3COO^-	Exp	$1.022 \times \sigma_{ONa}$	5.3(3.9[a])	0.08(0.06[a])
$CH_3SO_3^-$	Exp	$1.1 \times \varepsilon_{ONa}$	5.1(4.6[a])	0.06(0.05[a])
$CH_3SO_4^-$	Comp	$1.1 \times \varepsilon_{ONa}$	5.1(4.7[a])	0.06(0.05[a])
$H_2PO_4^-$	Comp	$1.02 \times \sigma_{ONa}$	5.2(3.0[a])	0.09(0.05[a])
$(CH_3)_2PO_4^-$	Comp	$1.01 \times \sigma_{ONa}$	4.9(3.8[a])	0.07(0.05[a])
SO_4^{2-}	Exp	$1.13 \times \sigma_{ONa}$	22.4(5.1[a])	0.35(0.07[a])
$CH_3PO_4^{2-}$	Comp	$1.132 \times \sigma_{ONa}$	22.0(3.1[a])	0.35(0.06[a])

[a] GAFF parameters; [b] Parameters optimized based on activity derivatives (Exp), repeated here to facilitate comparisons, or based on the ab initio approach (Comp); [c] Expressed as a scaling factor relative to the original GAFF parameters, calculated using equation 2.5 from the GAFF self-interaction parameters given in the Supporting Information; [d] $\Delta E_{QM\rightarrow cl}$ is defined in Equation 1.2; [e] $\Delta R_{QM\rightarrow cl}$: the difference between the position of the dimer energy minimum obtained using ab initio calculations and using classical parameters.

The optimized anion-cation interaction parameters are shown in Table 3.7 and Table 3.8. In all cases, the parameters optimized based on ab initio lead to $\Delta E_{AB,QM\rightarrow cl}$ differences within 0.5 kcal/mol of the target values. The anion-cation distances for monovalent anions are slightly larger using the optimized parameters than using GAFF [90, 91, 36, 92] . The same trend is observed for the divalent anions, although

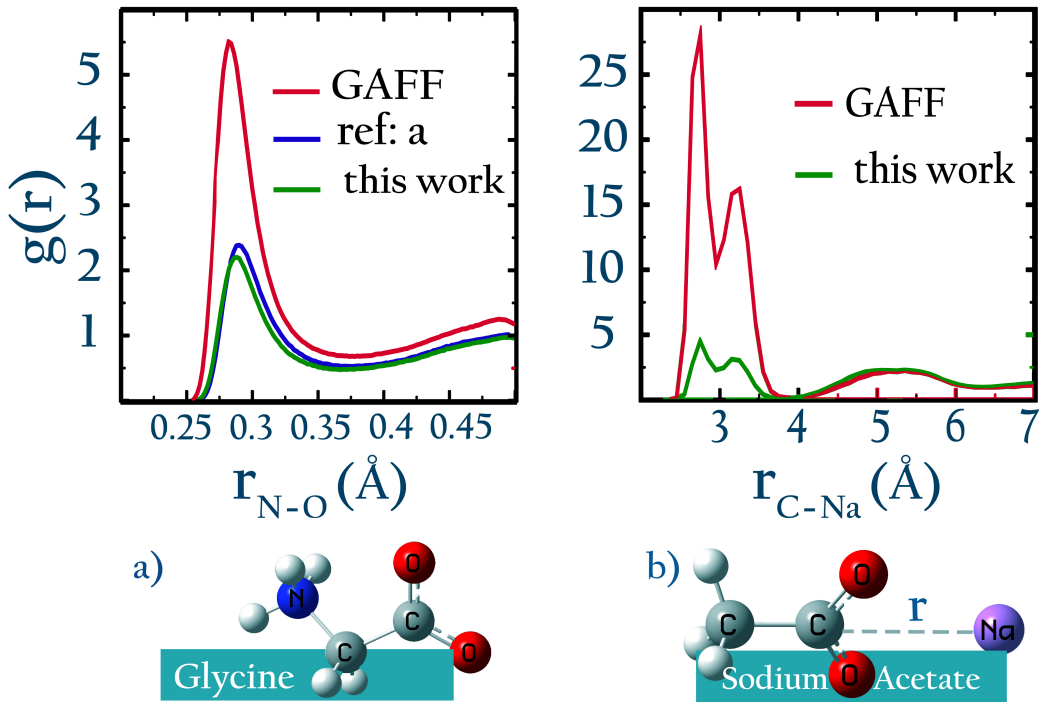

Figure 3.8 a) 3 m solution of Glycine (a = ref. [62]), b) 0.5 m solution of sodium acetate

for those systems the increase in anion-cation distance is substantially larger than for the monovalent anions. Such an increase is not surprising: as mentioned above, it reflects the fact that Lennard-Jones parameters optimized for neutral species are not appropriate for charged species.

Our parameters for the $(CH_3)_2PO_4^- \cdots CH_3NH_3^+$ interaction are very close to recently proposed parameters based that reproduce osmotic pressure measurements made on DNA arrays, again confirming the suitability of our ab initio approach: we propose that $\sigma_{ON} = 1.04 \times \sigma_{ON,GAFF}$, which corresponds to a +0.124 Å increase in the σ_{ON} value. This value is very close to the 0.14 Å increase proposed by Aksimentiev and Yoo [62].

Our results show that optimum parameters for the NH_4^+-anion interactions differ markedly from those for $CH_3NH_3^+$-anion interactions. Developing parameters for primary amines should be done using $CH_3NH_3^+$ as the reference species, and not NH_4^+, which is a poor mimic of primary amines.

The radial distribution functions for some of the salts for which the anion-cation interactions are optimized using ab initio are shown in Figure 3.7. Similarly to the systems for which parameters are optimized based on experimental data, we find that the optimized parameters reduce anion-cation attraction relative to GAFF [90, 91, 36, 92] parameters. These results suggest that GAFF [90, 91, 36, 92] parameters in general overestimate cation-anion interactions and may lead to aggregation, particularly in solutions with higher salt concentration. The most dramatic changes are seen in solutions of divalent ions: GAFF [90, 91, 36, 92] parameters lead to very rapid (< 1 ns)

Table 3.8 Differences between the classical and quantum potential energy scans for all anion-NH_4^+ and anion-$CH_3NH_3^+$ dimers, and values of the optimized anion-cation parameters.

Anion	Cation	Optimization[b]	Parameters[c]	$\Delta E^{d}_{QM\to cl}$ kcal/mol	$\Delta R^{e}_{QM\to cl}$ Å
$CH_3SO_3^-$	NH_4^+	Exp.	$1.10\times \{ \varepsilon_{\mathcal{O}N}, \varepsilon_{\mathcal{O}H_N}$	10.1(9.7[a])	0.25(0.23[a])
	$CH_3NH_3^+$	Comp.	$1.05\times \{ \sigma_{\mathcal{O}N}, \sigma_{\mathcal{O}H_N}$	10.2(6.8[a])	0.34(0.21[a])
$CH_3SO_4^-$	NH_4^+	Comp.	$1.20\times \{ \varepsilon_{\mathcal{O}N}, \varepsilon_{\mathcal{O}H_N}$	9.9(9.1[a])	0.27(0.24[a])
	$CH_3NH_3^+$	Comp.	$1.05\times \{ \sigma_{\mathcal{O}N}, \sigma_{\mathcal{O}H_N}$	10.6(7.4[a])	0.35(0.21[a])
$H_2PO_4^-$	NH_4^+	Comp.	$1.02\times \{ \sigma_{\mathcal{O}N}, \sigma_{\mathcal{O}H_N}$	10.4(8.9[a])	0.26(0.20[a])
	$CH_3NH_3^+$	Comp.	$1.04\times \{ \sigma_{\mathcal{O}N}, \sigma_{\mathcal{O}H_N}$	10.0(7.0[a])	0.30(0.19[a])
$(CH_3)_2PO_4^-$	NH_4^+	Comp.	$1.01\times \{ \sigma_{\mathcal{O}N}, \sigma_{\mathcal{O}H_N}$	10.1(9.5[a])	0.22(0.21[a])
	$CH_3NH_3^+$	Comp.	$1.04\times \{ \sigma_{\mathcal{O}N}, \sigma_{\mathcal{O}H_N}$	10.4(7.5[a])	0.30(0.19[a])
CH_3COO^-	NH_4^+	Comp.	..[f]	10.2(10.9[a])	0.20(0.23[a])
	$CH_3NH_3^+$	Comp.	$1.02\times \{ \sigma_{\mathcal{O}N}, \sigma_{\mathcal{O}H_N}$	9.9(8.9[a])	0.24(0.21[a])
SO_4^{2-}	NH_4^+	Exp.	$1.13\times \{ \sigma_{\mathcal{O}N}, \sigma_{\mathcal{O}H_N}$	30.3(16.9[a])	0.59(0.26[a])
	$CH_3NH_3^+$	Comp.	$1.14\times \{ \sigma_{\mathcal{O}N}, \sigma_{\mathcal{O}H_N}$	30.0(15.8[a])	0.61(0.25[a])
$CH_3PO_4^{2-}$	NH_4^+	Comp.	$1.13\times \{ \sigma_{\mathcal{O}N}, \sigma_{\mathcal{O}H_N}$	30.0(16.0[a])	0.58(0.25[a])
	$CH_3NH_3^+$	Comp.	$1.15\times \{ \sigma_{\mathcal{O}N}, \sigma_{\mathcal{O}H_N}$	30.7(14.7[a])	0.62(0.23[a])

[a] GAFF parameters; [b] Parameters optimized based on activity derivatives (Exp) or based on the ab initio approach (Comp); [c] Expressed as a scaling factor relative to the original GAFF parameters, calculated using equation 2.5 from the GAFF self-interaction parameters given in the Supporting Information; [d] $\Delta E_{QM\to cl}$ is defined in Equation 1.2; [e] $\Delta R_{QM\to cl}$: the difference between the position of the dimer energy minimum obtained using ab initio calculations and using classical parameters; [f] No scaling of anion-cation interactions.

aggregation of the ions and crystal formation, whereas our parameters yield solubilized systems. For all solutions containing monovalent anions, the optimized parameters decrease the number of contact ion pairs while leaving the number of solvent-shared ion-pairs unchanged. The overestimation of anion-cation interactions is particularly marked for $Na^+\cdots CH_3COO^-$, $NH_4^+\cdots(CH_3)_2PO_4^-$and $Na^+\cdots(CH_3)_2PO_4^-$.

GAFF [90, 91, 36, 92] also particularly overestimates the interactions between $CH_3NH_3^+$ and CH_3COO^-. This effect is illustrated in Figure 3.8a, by comparing the radial distribution of a 3 m solution of glycine zwitterion using GAFF and using optimized parameters. Overestimation of interactions between acidic and cationic side-chains, *i.e.*, salt-bridge formation, has also been observed by others. [144] This artifact can be prevented by using our optimized anion-cation interaction parameters.

In Figure 3.8a we also compare the anion-cation rdf for glycine obtained using our parameters with that obtained using parameters optimized from osmotic pressure measurements [62]. The similarity of the results obtained from our parameters – optimized using our ab initio approach– with those obtained from parameters optimized to reproduce experimental data [62] further demonstrates the success of our approach.

3.3 Interactions between oxoanionic polymers and cationic proteins

As mentioned in the introduction (section 1.1.2), we would like to determine and compare the affinities between single anion-cations pairs in our simulations, and to compare our results to the experimental data from Haag's group, with the ultimate aim of understanding the origin of the experimental trends. The experimentally measured affinity of the anionic dendritic polyglycerols towards the cationic protein L-selectin has the following trend:
carboxylate $<$ phosphate $<$ phosphonate $\sim$ sulfonate $<$ bisphosphonate $\lll$ sulfate
To explain the observed differences in the anionic dPGs, previous studies have considered the chemical differences in the anions, such as charge density, acidity, etc., to the macroscopic differences in the affinity of dPG molecules towards L-selectin. These simple characteristics, however, proved insufficient to understand the experimental results.

3.3.1 Characterizing monovalent anion-cation interactions

To better characterize the energetics of the anion-cation interactions, with the hope of understanding experimental trends regarding ion-specific effects involving multivalent oxoanions, we performed umbrella sampling and calculated the potential of mean force (PMF) for the interaction between each of Na^+ and $CH_3NH_3^+$ with $CH_3SO_3^-$, $CH_3SO_4^-$, $CH_3HPO_4^-$, CH_3COO^- and $CH_3PO_4^{2-}$. The result of the umbrella sampling is plotted in Figures 3.9 and 3.10, without and with the entropic contribution, $-2k_BT\ln(\xi)$, respectively. In Figure 3.9 the black and red curves corresponds to the

1-1 and 1-2 Coulomb potential, respectively. The PMFs are shifted to converge to the corresponding Coulomb potential at the 14 Å separation.

When the entropic $2k_BT\ln(r)$ contribution is factored in, it is clear that neither CIP nor SSIP states are favored for most ions; substantial association is only to be expected between $CH_3PO_4^{2-}$ and Na^+. The intrinsic binding strength (without the $2k_BT\ln(r)$ contribution; see Figure 3.9) is also low for most ion pairs: the difference in binding free energy between CIP relative to configurations at large separation is of order 1 kcal/mol. For $CH_3NH_3^+$ and $CH_3SO_3^-$ in particular, the CIP configuration is actually disfavoured, having a free energy that is more positive than the free energy at large separation. Only Na^+ and $CH_3PO_4^{2-}$ show substantial binding strength, of multiple kcal/mol. The minimum binding energy free energy between Na^+ and $CH_3SO_4^-$, approximately 0.5 kcal/mol, is consistent with the experimentally measured impact of a single sulfate-cationic amino acid interaction on the binding affinity between L-selectin and a sulfated ligand [30]. This agreement with experiment suggests that our models capture SO_4^{2-}-$CH_3NH_3^+$ interactions reasonably well.

To better quantify the anion-cation affinity in our simulations, and thereafter the trend of interaction of anions with $CH_3NH_3^+$, we calculated anion-cation binding constants using the relations given in section 2.6. The anion-cation binding constants, as well as the distance corresponding to the probability maximum (R_{max}) and corresponding to the maximum anion-cation separation defining a CIP or a SSIP, are listed in Table 3.9. We considered the contribution of contact ion pairs (CIP), as well as the overall contribution of CIP and solvent shared ion pairs (SSIP), and calculated the corresponding binding constants, K(CIP) and K(CIP+SSIP).

All the anion-cation pairs show binding constants in the M^{-1} range, consistent with the expected high solubility of these salts: acetate salts are generally soluble; so are ammonium salts, including ammonium sals with the various forms of the phosphate anion [80]. In contrast, sodium dihydrogen phosphate has a much lower solubility than sodium hydrogen phosphate [80], a trend which is consistent with the substantially large differences in the binding constants of Na^+-$CH_3HPO_4^-$ (K(CIP)=0.295 M^{-1}) and Na^+-$CH_3PO_4^{2-}$ (K(CIP)=339 M^{-1}) with that we obtain from our optimized parameters.

Based on the K(CIP) results, the anions show the following trend in their affinity towards $CH_3NH_3^+$(see Figure 3.11):

$$CH_3SO_3^- < CH_3PO_4^{2-} < CH_3SO_4^- < CH_3HPO_4^- < CH_3COO^-$$

Considering K(CIP+SSIP), however, the trend changes dramatically with the phosphate having the highest binding constant and the monovalent ions having very similar relative (see Figure 3.11):

$$CH_3SO_3^- \sim CH_3COO^- \sim CH_3SO_4^- \sim CH_3HPO_4^- < CH_3PO_4^{2-}$$

The similarity of the K(CIP+SSIP) values between the monovalent ions suggests that these anions show specific effects when they are in direct contact with the cationic groups. At large distances, non-specific effects dominate their interactions. However, for divalent ions, such as $CH_3PO_4^{2-}$, due to the higher charges on the anions, their specific effect might decay slower than the monovalent ions.

However, neither K(CIP) nor K(CIP+SSIP) are consistent with the experimental affinity of dendrimeric anionic polymers for selectins, nor are these binding constants

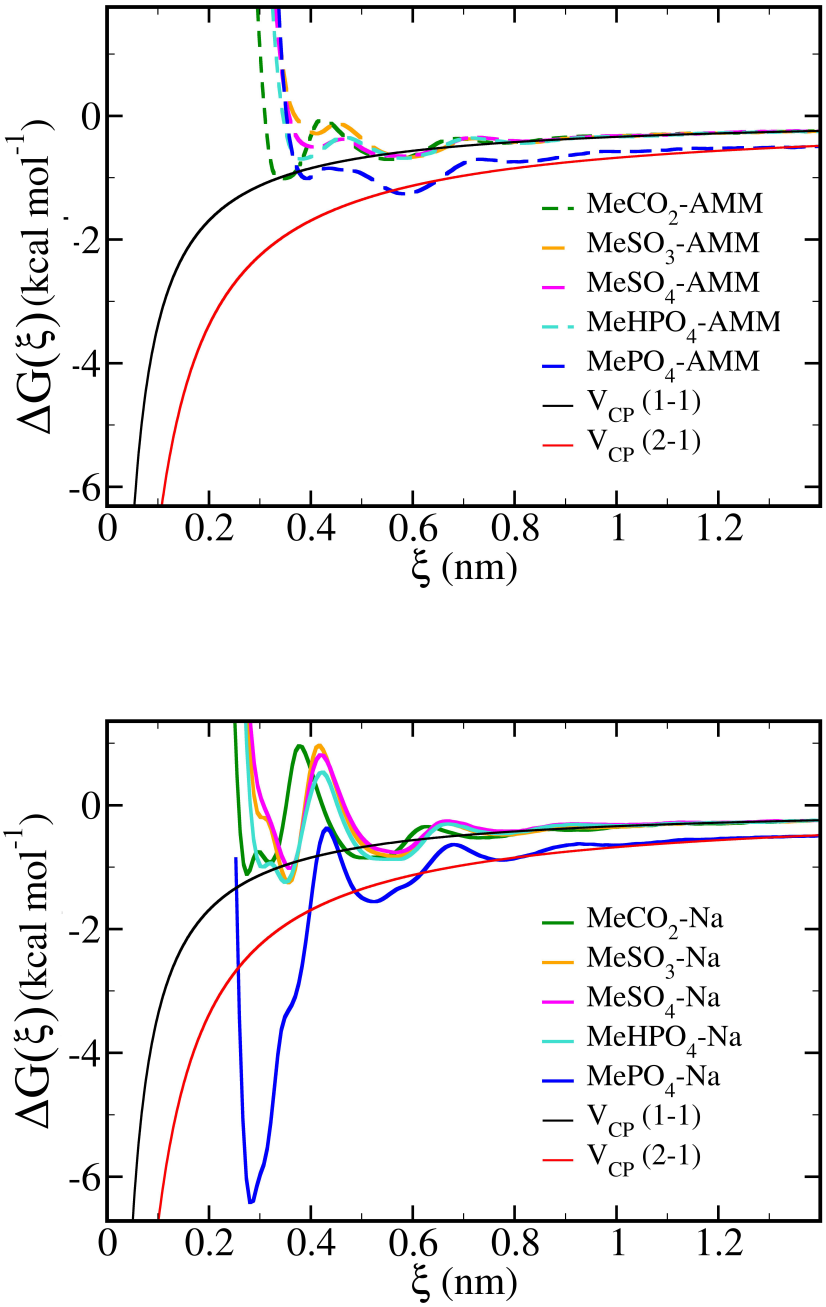

Figure 3.9 Free energy as a function of anion-cation distance for the indicated anion-cation pairs (dashed lines). The entropic contribution, $-k_BT2\ln(r)$, is not included. The expected limiting behavior at large ξ, corresponding to the electrostatic interaction between two monovalent ions (V_CP(1-1)) or divalent ions (V_CP(2-1)) of opposite charge, is shown. The top figure shows the PMFs of the various anions with $CH_3NH_3^+$ (AMM), the bottom one with Na^+.

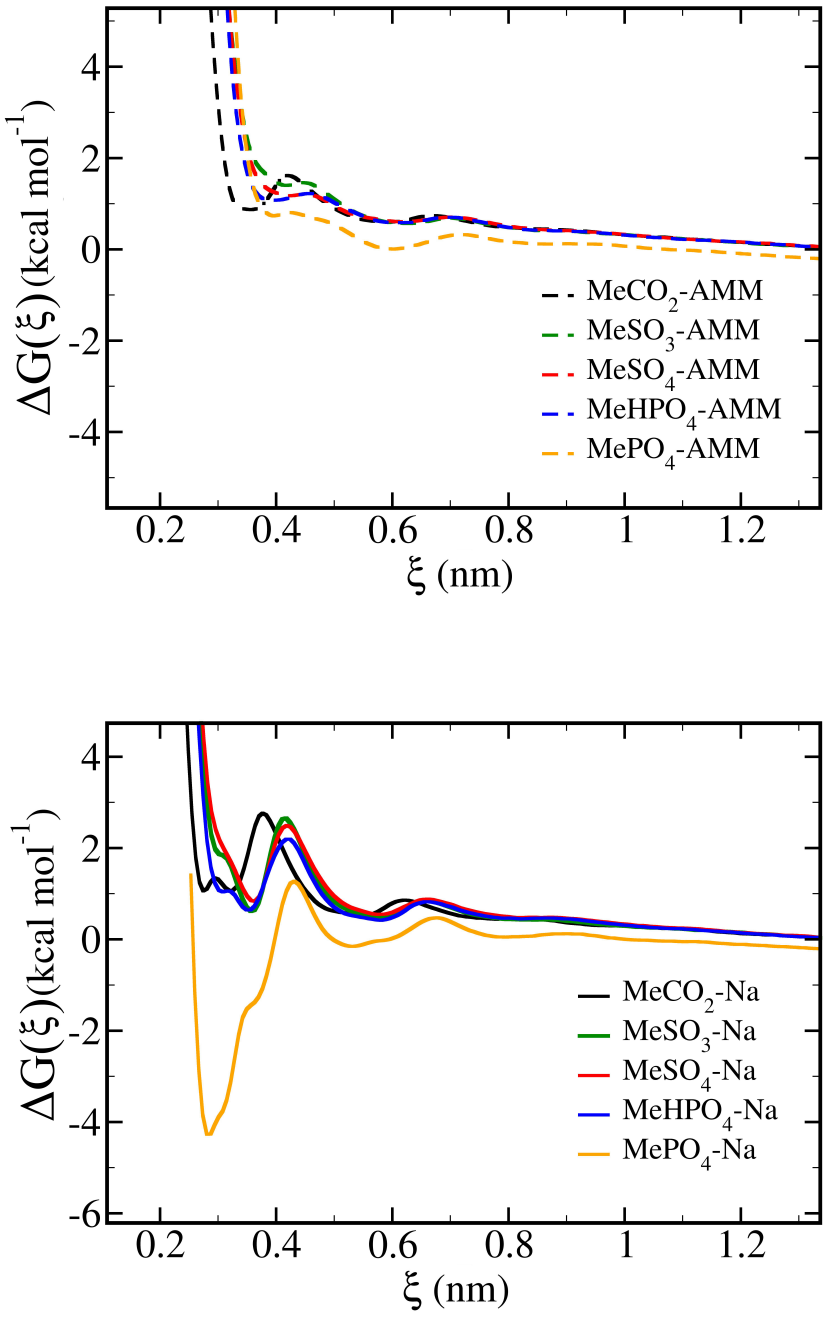

Figure 3.10 Free energy as a function of anion-cation distance for the indicated anion-cation pairs (dashed lines), including the entropic contribution, $-k_B T 2 \ln(r)$. The top figure shows the PMFs of the various anions with $CH_3NH_3^+$ (AMM), the bottom one with Na^+.

Table 3.9 Anion-cation binding constants calculated from the potential of mean force curves including the $-2k_BT\ln(\xi)$ contribution. Shown are also the position of the CIP and SSIP configurations with lowest free energy (R_{min}(CIP), R_{min}(SSIP)), and the maximum anion-cation separation defining a CIP or a SSIP, (R_{max}(CIP), R_{max}(SSIP)).

Anion	Cation	R_{max}(CIP) nm	R_{min}(CIP) nm	K(CIP) M^{-1}	R_{max}(SSIP) nm	R_{min}(SSIP) nm	K(CIP+SSIP) M^{-1}
$CH_3SO_3^-$	Na^+	0.356	0.417	0.210	0.580	0.662	1.18
	$CH_3NH_3^+$	0.419	0.445	0.094	0.630	0.695	1.15
$CH_3SO_4^-$	Na^+	0.363	0.417	0.170	0.574	0.662	1.070
	$CH_3NH_3^+$	0.422	0.448	0.153	0.620	0.700	1.27
$H_2PO_4^-$	Na^+	0.353	0.422	0.199	0.577	0.658	1.28
	$CH_3NH_3^+$	0.403	0.456	0.222	0.626	0.698	1.37
$CH_3HPO_4^-$	Na^+	0.356	0.417	0.295	0.581	0.656	1.40
	$CH_3NH_3^+$	0.396	0.455	0.213	0.606	0.692	1.27
$(CH_3)_2PO_4^-$	Na^+	0.351	0.419	0.315	0.581	0.655	1.27
	$CH_3NH_3^+$	0.384	0.456	0.176	0.621	0.694	1.18
CH_3COO^-	Na^+	0.318	0.376	0.149	0.555	0.619	0.98
	$CH_3NH_3^+$	0.358	0.419	0.235	0.594	0.662	1.20
CH_3COOH	Na^+	0.349	0.403	0.0875	0.598	0.639	0.574
	$CH_3NH_3^+$	0.387	0.446	0.212	0.652	0.659	0.757
$CH_3PO_4^{2-}$	Na^+	0.280	0.431	339	0.534	0.678	415
	$CH_3NH_3^+$	0.395	0.426	0.144	0.601	0.727	2.11

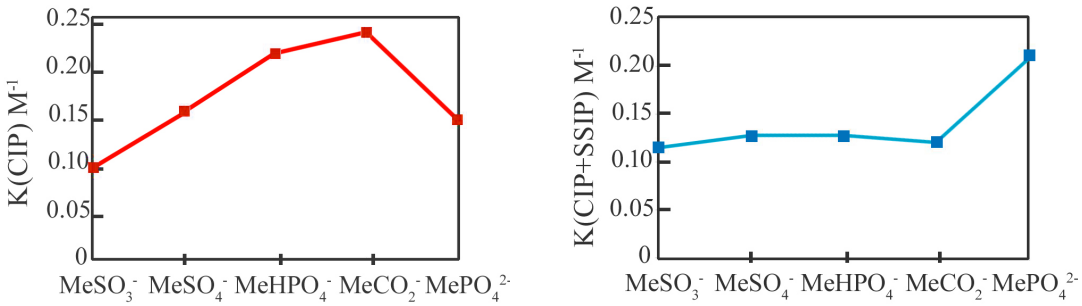

Figure 3.11 Binding constants (M^{-1}) for $CH_3NH_3^+$ and various anions in (a) CIP and (b) CIP+SSIP configurations, calculated from the potential of mean force curves including the $-2k_BT\ln(\xi)$ entropic contribution.

consistent with the order of magnitude differences in the affinity of the different anionic dPGs for selectins. Therefore, we tested the following parameters, that might affect the affinity trend in the anions, to shed light on the origin of the difference in the molecular level: a) counter ion effect, b) multivalent anion-cation interactions and c) protonation state of the anionic bases.

3.3.2 The impact of competition with Na^+ on the fraction of free anions

Na^+ is abundant in the extracellular physiological environment. In Haag's study the concentration of Na^+ was kept at the standard extracellular level, 0.15 M, which is significantly higher than the dPGs' or the selectin concentrations. Therefore, one should not neglect the effect of the presence of Na^+ in the solution. We obtained the K(CIP) and K(CIP+SSIP) binding constant of Na^+-anion, and the results are listed in the table 3.9. The results show that the binding affinity of sodium varies a lot towards different anions and therefore can play a significant role in changing the trend of the anionic interactions with the $CH_3NH_3^+$. This idea is further investigated by solving the binding equilibrium equations with Na^+ and $CH_3NH_3^+$, for each anion (A^{n-}):

$$\mathrm{Na^+ + A^{n-} \rightleftharpoons NaA^{(n-1)-}} \qquad K_{NaA} = \frac{[NaA^{(n-1)-}]}{[Na^+][A^{n-}]} \tag{3.2}$$

$$\mathrm{CH_3NH_3{}^+ + A^{n-} \rightleftharpoons NaA^{(n-1)-}} \qquad K_{CH_3NH_3A} = \frac{[CH_3NH_3A^{(n-1)-}]}{[CH_3NH_3^+][A^{n-}]} \tag{3.3}$$

Using the equations above, the percentage of bound anions to each of Na^+ and $CH_3NH_3^+$ was calculated for a given concentration of ions. This calculation was performed for ion concentrations of 1 and 3 M. These high concentrations correspond to local concentrations near the surface of the dendritic polymer and the selectin. For the surface coverage and dendritic polymer radius used in some experiments, the typical distances between anionic groups at the surface of the dendritic polymer range between 4 Åand 20 Å, which correspond to the local anion concentrations we used for our calculations. Comparable or even higher local cationic concentrations can be expected for selectin, following the same argument. These high charge densities necessarily increase the local concentration of the Na^+ ions near the dendritic polymer to much beyond the bulk concentration of 0.15 M, so, for simplicity, the local Na^+ concentration was considered to be equal to the local anion concentration at the surface of the dendritic polymer.

Figure 3.13 shows that the presence of Na^+ will significantly impact the binding pattern of $CH_3PO_4^{2-}$, making it a very poor ligand for $CH_3NH_3^+$, in either of CIP or SSIP regimes: the strong affinity between Na^+ and $CH_3PO_4^{2-}$ will effectively result in very little binding between that anion and $CH_3NH_3^+$. Na^+ interacts similarly with all monovalent anions, therefore we do not expect that the presence of Na^+ will

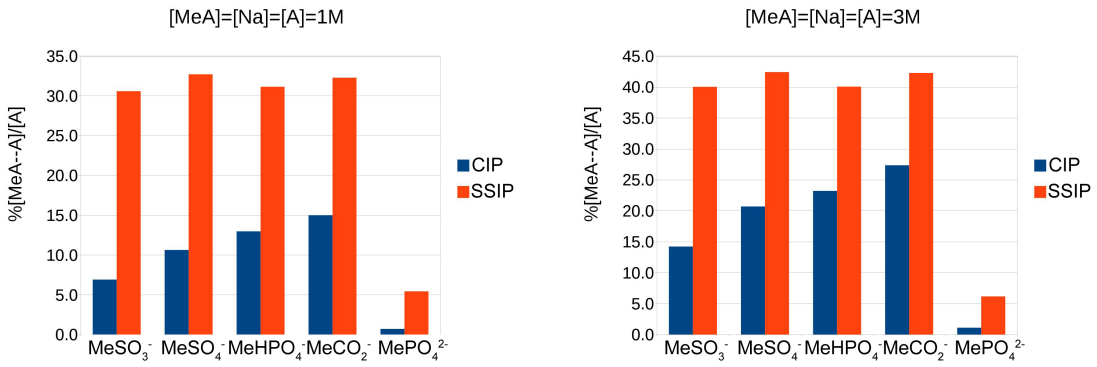

Figure 3.12 Percentage of $CH_3NH_3^+$-anion CIP (blue) and CIP+SIP (red) ion pairs for the indicated concentrations. The concentrations used here reflect local concentrations at the surface of the selectin and the dendritic polymer, as detailed in the text.

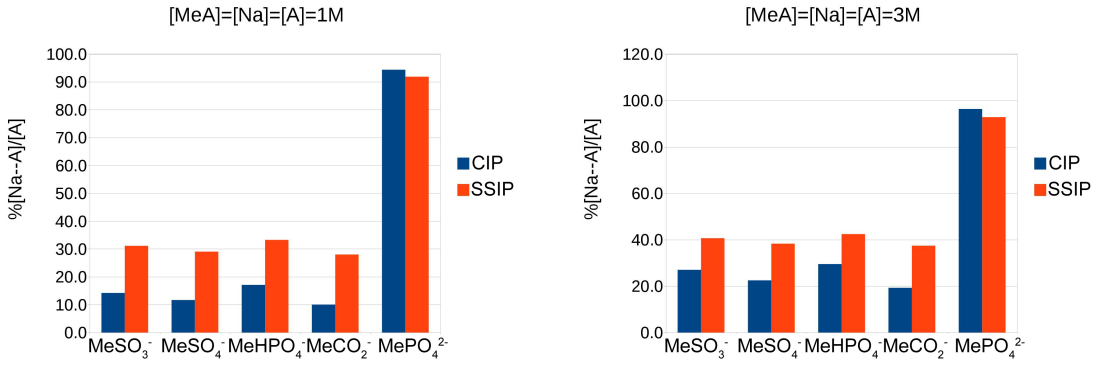

Figure 3.13 Percentage of Na^+-anion CIP (blue) and CIP+SIP (red) ion pairs for the indicated concentrations. The concentrations used here reflect local concentrations at the surface of the selectin and the dendritic polymer, as detailed in the text.

Table 3.10 Binding constants between two rigid species assuming n-site interaction between them, considering either CIP or CIP+SSIP anion-cation configurations at each site.

Anion	$K^{(1)}$		$K^{(2)}$		$K^{(3)}$		$K^{(4)}$		$K^{(5)}$	
	CIP	CIP+SSIP	CIP	CIP+SSIP	CIP	CIP+SSIP	CIP	CIP+SSIP	CIP	CIP+SSIP
$CH_3SO_3^-$	0.094	1.15	0.073	1.70	0.056	2.63	0.042	4.18	0.030	6.75
$CH_3SO_4^-$	0.153	1.27	0.182	1.95	0.213	3.029	0.242	4.799	0.260	7.69
$CH_3HPO_4^-$	0.213	1.27	0.307	1.97	0.450	3.16	0.647	5.15	0.889	8.457
CH_3COO^-	0.235	1.20	0.563	2.20	1.443	4.489	3.654	10.07	8.766	24.61
$CH_3PO_4^{2-}$	0.144	2.11	0.205	4.909	0.247	12.1	0.236	30.9	0.188	80.7

result in very different fractions of available anionic sites on the dendritic polymer. Nevertheless, high concentrations of Na^+ should redicue the percentage of anions available for binding to $CH_3NH_3^+$.

3.3.3 Impact of multivalency

We next investigated how the differences in the binding affinity of single anions and cations would propagate in the context of multivalent interactions, which are present between the dendritic anions and the selectins. We assume completely rigid species (protein and dPG) and we further assume that each ion pair is sufficiently separated from the adjacent ion-pairs so that they do not cross-interact. Under these assumptions, the binding free energy of n ion-pairs can be constructed by excluding the entropic $2k_BT\ln(r)$ contribution from $n-1$ single ion-pair binding free energies [145]. The binding affinity arising from n-site interactions between 2 rigid species is therefore obtained from summing over $n-1$ entropy excluded free energies and 1 original ion-pair free energy.

The binding constants assuming n-site interactions between 2 rigid species are listed in table 3.10; the corresponding binding free energy, $\Delta G_{binding}$, normalized to the 1 M reference state [146] are shown in Figure 3.14.

Considering multivalent interactions magnifies the binding affinity differences between the anions; in particular, for $CH_3SO_3^-$, the binding affinity arising from CIP contacts decreases for increasing n, a consequence of the unfavourable nature of that state. The presence of multivalent interactions, however, does not alter the trend visible in the monovalent interactions, *i.e.*, the experimental trend is not mimicked in this multiple binding model.

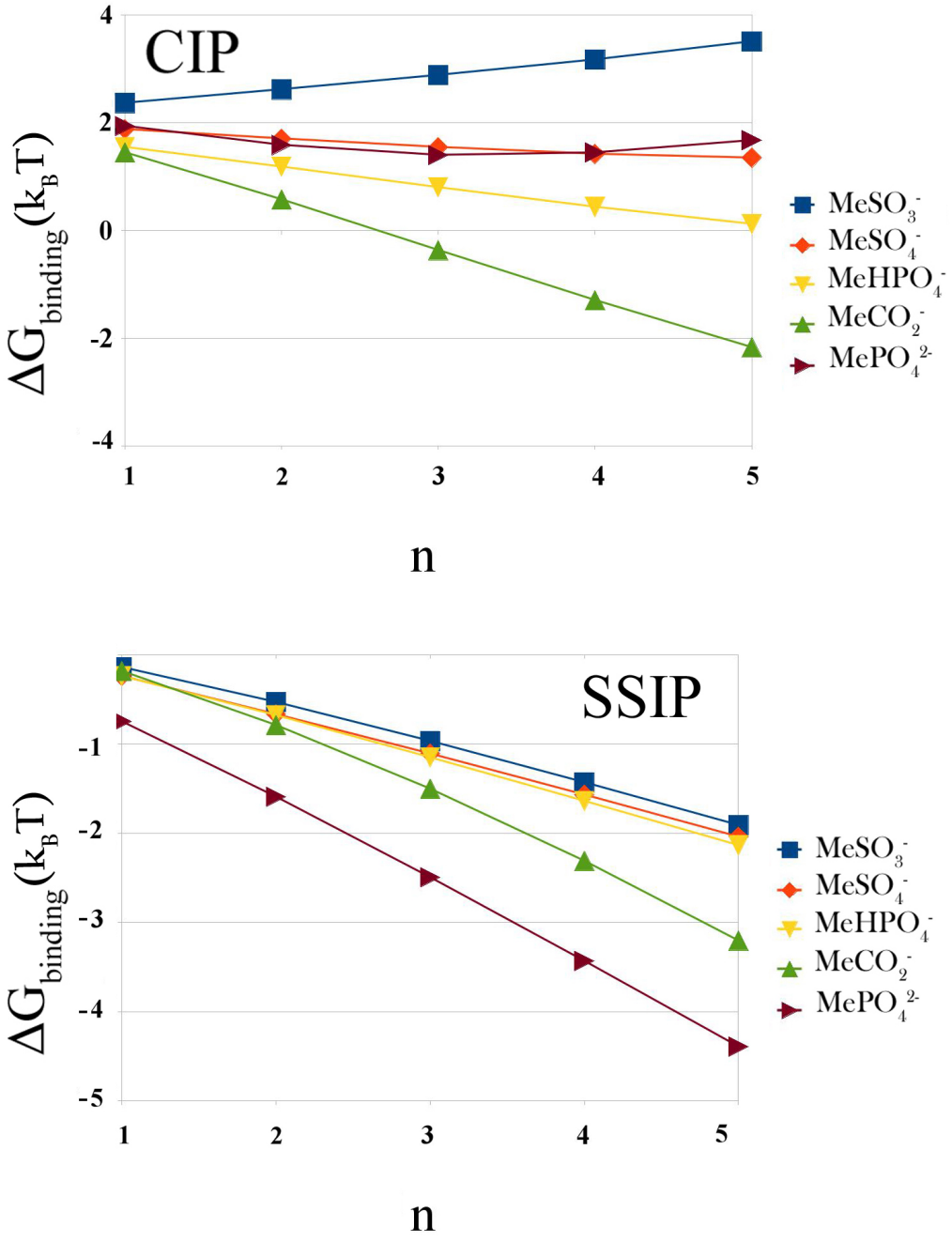

Figure 3.14 Free energy of binding assuming n-site interactions between 2 rigid species; the cationic sites are mimicked as $CH_3NH_3^+$, the anionic ones as the indicated anions. (a) CIP interactions; (b) CIP+SSIP interations.

3.3.4 Multivalency, pKa and competition with Na^+- impact on binding affinity

As a result of interaction with Na^+ some of the anionic sites will not be available to interact with the selectins. Since the percent of blockage depends on the concentration of Na^+ and on the identity and concentration of the anion itself, we investigate the effect of Na^+ blockage by considering various local concentrations of anionic head groups and Na^+. We assume that the interaction area of a dPG is about the size of the accessible surface of a selectin, and is about 600 $Å^2$. We considered various numbers of anionic head groups, **N**, ranging from 1 to 5, within this interaction area in dPG. Having the area and the number of head groups we calculated the local anion density corresponding to each **N** value, and finally calculated the number of anionic sites that are not bound to Na^+, and are thus available for binding with selectin. In this calculation, we considered both CIP or CIP+SSIP Na^+-anion configurations. The number of free anionic sites calculated this way are shown in tables 3.11 and 3.12.

Table 3.11 Number of free anionic sites (NF) on a dPG, given a maximum of N sites, assuming that each site is blocked when it forms a CIP with Na^+.

N	Number of Free Sites (**NF**)				
	$CH_3SO_3^-$	$CH_3SO_4^-$	$CH_3HPO_4^-$	CH_3COO^-	$CH_3PO_4^{2-}$
1	1.0	1.0	1.0	1.0	0.3
2	1.9	1.9	1.8	1.9	0.2
3	2.7	2.7	2.6	2.8	0.2
4	3.4	3.5	3.3	3.6	0.2
5	4.1	4.2	3.9	4.3	0.2

Table 3.12 Number of free anionic sites (NF) on a dPG, given a maximum of N sites, assuming that each site is blocked when it forms either CIP or SSIP pairs with Na^+.

N	Number of Free Sites(**NF**)				
	$CH_3SO_3^-$	$CH_3SO_4^-$	$CH_3HPO_4^-$	CH_3COO^-	$CH_3PO_4^{2-}$
1	0.9	0.9	0.9	0.9	0.3
2	1.5	1.6	1.5	1.6	0.2
3	2.0	2.1	1.9	2.1	0.2
4	2.4	2.5	2.3	2.6	0.2
5	2.8	2.8	2.6	2.9	0.2

The number of free anionic sites within the given interaction area might be further reduced if the local pKa of the anionic sites is altered relative to the expected value at low surface density. To examine this effect we consider the pKa of the conjugated acids of anions, shown in Table 3.13.

Table 3.13 pKa of various anions.

Anions Conjugate Acid	pKa
$\mathbf{MeSO_3H}$	-2.6
$\mathbf{MeSO_4H}$	-3.5
$\mathbf{MePO_4H_2}$	1.54
$\mathbf{CH_3COOH}$	4.76
$\mathbf{MePO_4H^-}$	6.31

The pKa of the conjugate acids of CH_3COO^- and $CH_3PO_4^{2-}$ are not far from the physiological pH=7 Therefore they can be partially protonated, which will further reduce the number of free sites, NF, from the values listed in Tables 3.12 and 3.11. This effect is of great importance especially considering the fact that the anionic head groups may be closely packed at the surface of the dPGs, and therefore strong internal hydrogen bonding as well as high charge density might affect the degree of protonation. Marked changes in pKa have been reported previously for dicarboxylic acids such as succinic acid, malonic acid, etc.

If the corresponding pKa of each CH_3COO^- group on a dPG raises, for example to pKa=6.5, the free sites will be reduce to %76 of the original value (see table 3.14) and an CH_3COO^- dPG with N=4 would only have 3 sites available. This effect would reduce the affinity of the acetate-dPG for selectin relative to that measured for other dPGs with different anions, which might alter the relative affinities of differently functionalized dPGs for selectin.

Table 3.14 Percentage ratio of CH_3COO^- to the sum of protonated and non-protonated states.

pKa	$\%CH_3COO^-$
4.76	99
5.00	99
5.50	97
6.00	91
6.50	76
7.00	50

We now consider the scenario of a substantial change in the pKa of acetate, from 4.76 to 6.5, and consider also the impact of multivalency and competition with Na^+ and ask how it would change the binding affinity between two species, in the context of a 5-site model. We further consider that, because of its pKa, 83% of the phosphate sites will be in the form of $CH_3HPO_4^-$ AND 17% in the form of $CH_3PO_4^{2-}$. In Table 3.15 we show the binding constants quantifying the affinity of the various anionic dPGs for the cationic 5-site species, calculated under these conditions.

The results show that considering the impact of pKa and of sodium competition simultaneously, in the context of multivalent binding between two species has a substantial effect on the relative binding affinity between two species. Although, we do

Table 3.15 Binding constants assuming a maximum number of 5 binding sites between 2 rigid species, and considering the impact of competition with Na^+ and of the pKa on the number of available anionic binding sites.

N=5	$CH_3SO_3^-$	$CH_3SO_4^-$	$CH_3HPO_4^-$	CH_3COO^-	$CH_3PO_4^{2-}$
pKa applied and blocked by Na^+ in CIP form					
NF	∼4	∼4	∼3	∼3	∼0
$K^{(n)}$(CIP)	0.042	0.242	0.450	1.443	0
$K^{(n)}$(SSIP+SSIP)	4.18	4.80	3.16	4.489	0
pKa applied and blocked by Na^+ in CIP+SSIP forms					
NF	∼3	∼3	∼2	∼2	∼0
$K^{(n)}$(CIP)	0.056	0.213	0.307	0.563	0
$K^{(n)}$(CIP+SSIP)	2.63	3.29	1.97	2.20	0

not see as large differences as in the real dPGS, It shows that the anionic binding is more sensitive to the environmental factors such as pH and counter ion presence than the chemical properties of the anions themselves.

3.3.5 Binding affinity between 3-site anionic and cationic polymers - comparison between simulation and the multivalent model

To assess the quality of the simple analytical multivalent binding model we employ here, we performed umbrella sampling and PMF calculations for model multivalent systems, with 3 anionic and cationic groups, illustrated in Figure 3.15. Each species consists of a branched alkane core and 3 anionic or cationic ($CH_3NH_3^+$) termini.

The binding constants obtained from the PMFs for the 3-site polymers are shown in Table 3.16. For comparison, we repeat in the same table the corresponding binding constants obtained from the 3-site analytical binding model. The results show that the simple multivalent model used here has strong limitations. Firstly, this model substantially underestimates the binding between the 3-site polymers. This underestimation is consistent with the presence of the hydrophobic effect in the simulations, which is not accounted for in the model, as well as the absence of the cross-interaction between individual ion-pairs. More importantly, however, the model fails to predict the relative binding affinity between the various systems, where the trend of interaction changes going from 1-1 to 3-3 binding sites in a real system. These results show that the simple analytical model we develop is useful to explore the potential impact of the identity of the anion, competition with Na^+ and local pKa of the anion on the binding affinity between selectins and ligands in a qualitative way only. The binding constants for interaction between the 3-site anionic and cationic polymers suggest that factors such as the relative orientation of anions and cations, as well as the possibility of multidentate interactions(e.g. one anion interacting with more

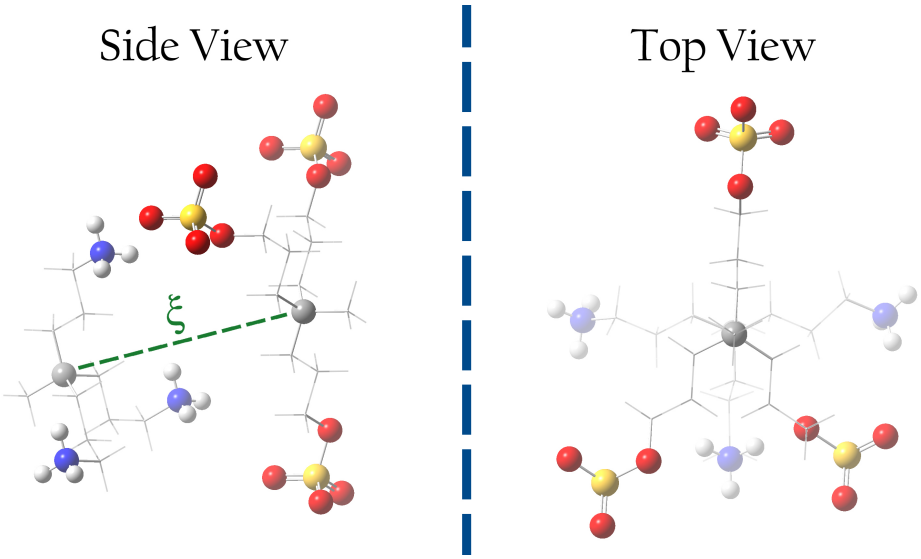

Figure 3.15 Example configuration of 3-site system for which the PMF as a function of distance is calculated.

than one cations simultaneously), may substantially alter the binding affinity between multivalent species.

3.4 Conclusions and outlook

We developed parameters for the alkali and halide monovalent ions, and the divalent cations Ca2+ and Mg2+, for the TIP5P water model for which no ion parameters are available, using the free energy of solvation and the activity derivative of the salt solutions as the experimental target values. We show that the activity derivative is a great target value for optimize the ion-ion interactions, however, it is poorly correlated to the structural properties of the ions in water. We proposed that the activity derivative should be used along with anion-cation coordination number, obtained from x-ray data of solutions in high concentration, to optimize for ionic interactions and the structural properties at the same time. Furthermore, We have developed a new approach to obtain ion parameters for all-atom molecular dynamics simulations of biomolecular systems when experimental data to parameterize the ions are sparse. The approach consists of combining the existing experimental data with high level QM calculations. The parameters for $CH_3SO_3^-$, $CH_3SO_4^-$, $CH_3HPO_4^-$, CH_3COO^-, $CH_3PO_4^{2-}$as well as the non-methylated forms of these ions are developed to capture ion-water, anion-$Na^+$and anion-$CH_3NH_3^+$interactions. These optimized parameters made the grounds for the second part of the project, where we investigated the interaction of the individual anions with the head groups of the side chains of cationic amino acids (CH3NH3+ and guanidinium) as well as with Na+, which is abundant in biological environments and buffer solutions as electrolyte. Our potential of mean force curves as a function of interionic distance, obtained using umbrella sampling, suggest that the interaction with Na^+has a great effect on affinity of the anions to-

Table 3.16 Binding constants calculated from PMF curves for the interaction between 3-site anionic and cationic polymers.

Anion	K(CIP) n=3 model M^{-1}	K(CIP) 3-3 Umbrella Sampling M^{-1}	$\frac{K_{Umb.\,Sampling}}{K_{model}}$
$CH_3SO_3^-$	0.056	4.52	76
$CH_3SO_4^-$	0.213	14.26	67
$CH_3HPO_4^-$	0.450	12.6592	28
CH_3COO^-	1.443	17.0	12

wards $CH_3NH_3^+$, by decreasing the number of free sites, as well as influencing the structural properties of binding(i.e. CIP and SSIP forms). We also investigate the impact of multivalency and the protnation status on the binding affinity of anions in presence of Na^+, using rigid n-site models. The results show that these factors are of great importance and can change the affinity trend drastically. The binding constant results from the multivalent models were compared to the ones from the umbrella sampling and PMF calculations of 3-site polymers.

To achieve a thorough understanding of the interaction of the anionic dPGs and selectin, one should first use the small anionic analogues to map out all the possible electrostatically interactive sites on selectins. Furthermore the effect of multivalent binding on the interaction with selectin should be investigated using small polymer chains. In order to have a realistic comparison between different anioncs, the effect of anionic local density on the protonation state of the polymers should be also addressed accordingly.

Appendix A

Symmetries and Orientations

A.1 Symmetry and QM Potential Scan

Depending on the symmetry of the anions, different anion-oxygens could experience different chemical environments. Figure A.1 illustrates molecules with different symmetry point groups. When performing potential energy scans for the interaction of these ions with each of water, sodium, NH_4^+ and $CH_3NH_3^+$, the three oxygens of a C3v molecule, Figure A.1a, would yield the same result. However, for the molecule with Cs symmetry, Figure A.1b, two of the oxygens are similar and the third oxygen interacts differently with water and the counter ions. However, in classical force fields, all terminal oxygens are treated as equal and are assigned identical LJ parameters. To resolve this discrepancy between quantum and classical descriptions of these anions, we performed the potential energy scans for every distinct oxygen in the anions. The final $\Delta E_{QM\rightarrow cl}$ and $\Delta R_{QM\rightarrow cl}$ values are obtained as an average of different scans.

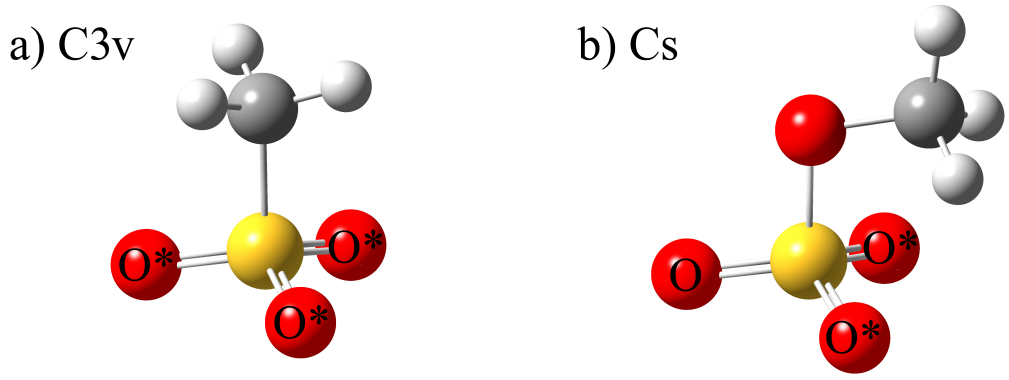

Figure A.1 The three oxygens in a) experience similar chemical environments due to C3v symmetry. The molecules with geometries similar to b) have only two similar oxygens and the third one is distinguishable, due to lower symmetry. Oxygens labeled with * experience similar chemical environments within that molecule.

A.2 Effect of Orientation on Water-Ion Interactions

We examined whether our proposed approach would yield consistent results if potential energy scans in other orientations were used. The comparison between the two different orientations, listed in Table A.1, shows that water-ion interactions could be optimized regardless of the orientation of water around the ions, using our approach. The values of $\Delta E_{QM\rightarrow cl}$ for the reference species (HSO_4^-) are essentially independent of the orientation (I or II; see Figure A.2), and the same is observed for the other species ($H_2PO_4^-$and $(CH_3)_2PO_4^-$) parameterized based on the reference species. It is important to note that the absolute minimum energies are different in different orientations. However, since the method is developed to capture the differences in the energies ΔE, the parameterization could be done regardless of the orientation.

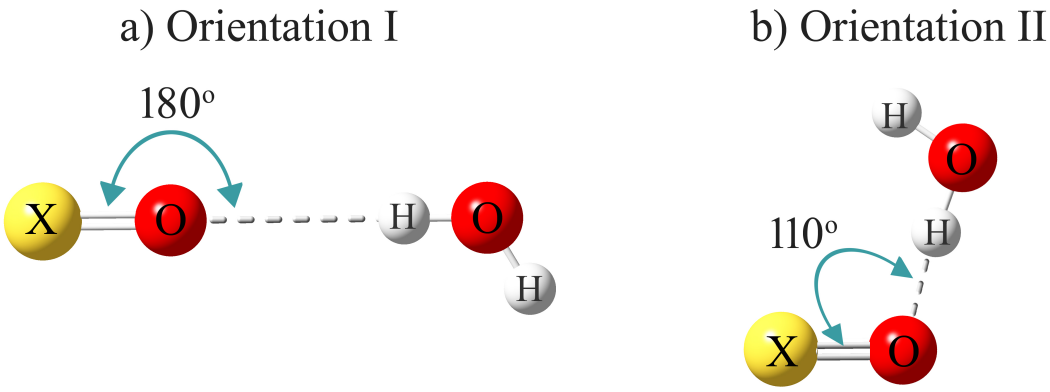

Figure A.2 The two orientations of water around anions. X=S,P

Table A.1 Differences between the classical and quantum potential energy scans for the mentioned anion-water dimers, in two different orientations.

Orientation	Anion	Optimization[c]	Parameters[d]	QM kcal/mol	Classic kcal/mol	$\Delta E^{e}_{QM\rightarrow cl}$ kcal/mol	$\Delta R^{f}_{QM\rightarrow cl}$ Å
I	HSO_4^-	Exp	.. [b]	-10.4	-11.8	-1.3[a]	-0.14[a]
	$H_2PO_4^-$	Comp	.. [b]	-13.2	-14.6	-1.4[a]	-0.12[a]
	$(CH_3)_2PO_4^-$	Comp	.. [b]	-13.4	-14.5	-1.1[a]	-0.11[a]
II	HSO_4^-	Exp	.. [b]	-12.4	-13.6	-1.2[a]	-0.14[a]
	$H_2PO_4^-$	Comp	.. [b]	-14.1	-15.2	-1.1[a]	-0.14[a]
	$(CH_3)_2PO_4^-$	Comp	.. [b]	-14.6	-15.7	-1.1[a]	-0.14[a]

[a] GAFF parameters; [b] No scaling; [c] Parameters optimized based on hydration free energies (Exp), repeated here to facilitate comparisons, or based on the ab initio approach (Comp); [d] Expressed as a scaling factor relative to the original GAFF parameters (given in the Supporting Information); [e] $\Delta E_{QM\rightarrow cl}$ is defined in Equation M1.2; [f] $\Delta R_{QM\rightarrow cl}$: the difference between the position of the dimer energy minimum obtained using ab initio calculations and using classical parameters. * X=S,P

Appendix B

Parameters and Guidelines

B.1 AMBER users

The parameters are provided in the *.top files, which is a gromacs format file. These files contain the original as well as the optimized values in both gromacs and amber formats. For amber users, the amber atom types are given as comments in the *.top files (";" character is used to comment). You should ignore our suggested names for atom types, since they were chosen to be compatible with gromacs naming conventions. Keep in mind that in many cases, the amber users need to define a new LJtype instead, as described below.

All the parameters for the bonds, angles and dihedrals are retained from original amber force field. The charges and masses are given in the *.itp files of the corresponding ions, under the [*atoms*] section.

Amber users are first supposed to prepare their own *.prmtop and *.inpcrd files using LEaP or similar programs, and then load the *.prmtop file into the topology file editor, ParmEd. You can use ParmEd to modify the self-interactions, called *SinglType*, as well as the pairwise interactions, called *PairType*.

Warning!!! This study is intended to provide parameters for specific anions and cations. If you have atoms in neutral groups or atoms in ionic groups other that those optimized in this study, and they share the same atom types (i.e O, O2, OH, OS, N3, H) with the anions and cations that we parameterized, you first need to create new van der Waals atom types, using "addLJType" command in the LEaP, to avoid unnecessary complications. As an example, to add new LJtype for atoms 3, 5, 6 and 9, the command line should read:

addLJType @3,5,6,9

where @ is a keyword to select an atom number, in this case atoms 3,5,6 and 9. The command will assign r and epsilon values of the first selected atom, here atom 3, to the rest of atoms, and give them a new LJtype.

To change the self interaction parameters, you can use the "changeLJSingleType" command. This will change the interaction of a <mask> (atom selection) with itself as well as with all other atomtypes. Note that, all the atoms in the <mask> should have the same atom types, otherwise it will raise an error. As an example, to change the self-interaction of atoms 3, 5, 6 and 9, to which we just assigned a new atom type, the command line should look like:

changeLJSingleType @3 r(Å) epsilon(kcal/mol)

where r and epsilon are the radius and Lennard-Jones well depth, respectively, and are given in the *.top files provided by us, under the [*atomtypes*] section. It is important to note that, we only need to select atom 3, to apply the change to the rest of the atoms with the same LJType. Therefore, here applying the change to atom 3 would effectively apply the change to the atoms 5, 6 and 9.

To change the pairwise interactions (i.e. anion-cation interactions in this study), one can use "changeLJPair" command. As an example, to change the pairwise inter-

action between atoms @3,5,6,9 – all of which have the same LJType – and all sodium ions, the command line should read:

changeLJPair @3 :Na+ R_{ij}(Å) $epsilon_{ij}$(kcal/mol)

where ":Na+" is to select all Na+ residues, R_{ij} is the minimum distance between i and j, and $epsilon_{ij}$ is the LJ well depth between atoms i and j. The R_{ij} and $epsilon_{ij}$ values are given in the *.top files provided by us, under the [*nonbond_params*] section.

Detailed instructions on how to use ParmEd can be found in the amber manual.

B.2 Gromacs users

The original and optimized parameters are provided in *.top files along with the corresponding *.itp files. Since this study is intended to provide parameters for specific anions and cations, we should avoid unintended changes to the atoms in neutral groups or atoms in ionic groups other that those optimized in this study, that share the same atom types (i.e O, O2, OS, OH, N3, H) with the anions and cations that we parameterized. Therefore, for each optimized parameter a new atom type is introduced (e.g. O is replaced with OSUL for SO_4^{2-}). We also changed the atom types N3 and H of NH_4^+ and $CH_3NH_3^+$ to (NAMO, HAMO) and (NAMM, HAMM) respectively, to distinguish between the two cations. We recommend that you change the atom type of amine head groups in your simulation to (NAMM, HAMM), since often there are some other atoms with type N3 or H, which should not be changed (e.g. H type in guanidinium).

B.3 Parameters

Acetate.top

```
; optimized parameters for acetate ion with respect to Na+ , NH4+ and CH3NH3+;

[ atomtypes ]
;name   bond_type     mass     charge   ptype   sigma nm   epsilon kJ/mol   rAMBER; A   epsilonAMBER; kcal/mol
 CT       CT         0.00000  0.00000   A     3.39967e-01   4.57730e-01    ;    1.9080         0.1094
 HC       HC         0.00000  0.00000   A     2.64953e-01   6.56888e-02    ;    1.4870         0.0157
 C        C          0.00000  0.00000   A     3.39967e-01   3.59824e-01    ;    1.9080         0.0860
 OACE     OACE       0.00000  0.00000   A     2.95992e-01   6.76553e-01    ;    1.6612         0.1617
;OACE     OACE       0.00000  0.00000   A     2.95992e-01   8.78640e-01    ;    1.6612         0.2100
 NA+      NA+        0.00000  0.00000   A     2.43928e-01   3.65846e-01    ;    1.3690         0.0874
 NAMO     NAMO       0.00000  0.00000   A     3.25000e-01   7.11280e-01    ;    1.8240         0.1700
 HAMO     HAMO       0.00000  0.00000   A     1.06908e-01   6.56888e-02    ;    0.6000         0.0157
 NAMM     NAMM       0.00000  0.00000   A     3.25000e-01   7.11280e-01    ;    1.8240         0.1700
 HAMM     HAMM       0.00000  0.00000   A     1.06908e-01   6.56888e-02    ;    0.6000         0.0157
 HP       HP         0.00000  0.00000   A     2.64953e-01   6.56888e-02    ;    1.4870         0.0157


[ nonbond_params ]
;i         j          func            sigma       epsilon         RijAMBER; A    epsilon_ijAMBER, kcal/mol
 OACE      NA+         1            2.75899e-1 4.97508e-1    ;    3.0969          0.1189    ;
;OACE      NA+         1            2.69960e-1 5.66963e-1    ;    3.0302          0.1355    ;
```

```
 OACE      NAMM         1              3.16706e-1 6.93699e-1    ;    3.5549          0.1658    ;
;OACE      NAMM         1              3.10496e-1 7.90543e-1    ;    3.4852          0.1889    ;
 OACE      HAMM         1              2.05479e-1 2.10813e-1    ;    2.3064          0.0504    ;
;OACE      HAMM         1              2.01450e-1 2.40243e-1    ;    2.2612          0.0574    ;
```

Dihydrogen_Phosphate.top

```
; optimized parameters for Dihydrogen Phosphate H2PO4- ion with respect to Na+ , NH4+ and CH3NH3+;


[ atomtypes ]
;name    bond_type     mass     charge   ptype    sigma nm    epsilon kJ/mol   rAMBER; A    epsilonAMBER; kcal/mol
 OHH2P     OHH2P       0.00000  0.00000   A      3.06647e-01   8.80314e-01    ;    1.7210         0.2104
 HO        HO          0.00000  0.00000   A      0.00000e+00   0.00000e+00    ;    0.0000         0.0000
 p5        p5          0.00000  0.00000   A      3.74177e-01   8.36800e-01    ;    2.1000         0.2000
 OH2P      OH2P        0.00000  0.00000   A      2.95992e-01   8.78640e-01    ;    1.6612         0.2100
 NA+       NA+         0.00000  0.00000   A      2.43928e-01   3.65846e-01    ;    1.3690         0.0874

 NAMO      NAMO        0.00000  0.00000   A      3.25000e-01   7.11280e-01    ;    1.8240         0.1700
 HAMO      HAMO        0.00000  0.00000   A      1.06908e-01   6.56888e-02    ;    0.6000         0.0157
 CT        CT          0.00000  0.00000   A      3.39967e-01   4.57730e-01    ;    1.9080         0.1094
 NAMM      NAMM        0.00000  0.00000   A      3.25000e-01   7.11280e-01    ;    1.8240         0.1700
 HAMM      HAMM        0.00000  0.00000   A      1.06908e-01   6.56888e-02    ;    0.6000         0.0157
 HP        HP          0.00000  0.00000   A      2.64953e-01   6.56888e-02    ;    1.4870         0.0157


[ nonbond_params ]
;i          j           func             sigma        epsilon          RijAMBER; A    epsilon_ijAMBER, kcal/mol
 OH2P      NA+          1              2.75359e-1 5.66963e-1    ;    3.0908          0.1355
;OH2P      NA+          1              2.69960e-1 5.66963e-1    ;    3.0302          0.1355
 OHH2P     NA+          1              2.80793e-1 5.67503e-1    ;    3.1518          0.1356
 OHH2P     NA+          1              2.75287e-1 5.67503e-1    ;    3.0900          0.1356


 OH2P      NAMO         1              3.16706e-1 7.90543e-1    ;    3.5549          0.1889
;OH2P      NAMO         1              3.10496e-1 7.90543e-1    ;    3.4852          0.1889
 OH2P      HAMO         1              2.05479e-1 2.40243e-1    ;    2.3064          0.0574
;OH2P      HAMO         1              2.01450e-1 2.40243e-1    ;    2.2612          0.0574

 OHH2P     NAMO         1              3.22139e-1 7.91296e-1    ;    3.6159          0.1891
;OHH2P     NAMO         1              3.15823e-1 7.91296e-1    ;    3.5450          0.1891
 OHH2P     HAMO         1              2.10912e-1 2.40472e-1    ;    2.3674          0.0575
;OHH2P     HAMO         1              2.06777e-1 2.40472e-1    ;    2.3210          0.0575


 OH2P      NAMM         1              3.22916e-1 7.90543e-1    ;    3.6246          0.1889
;OH2P      NAMM         1              3.10496e-1 7.90543e-1    ;    3.4852          0.1889
 OH2P      HAMM         1              2.09508e-1 2.40243e-1    ;    2.3516          0.0574
;OH2P      HAMM         1              2.01450e-1 2.40243e-1    ;    2.2612          0.0574

 OHH2P     NAMM         1              3.28456e-1 7.91296e-1    ;    3.6868          0.1891
;OHH2P     NAMM         1              3.15823e-1 7.91296e-1    ;    3.5450          0.1891
 OHH2P     HAMM         1              2.15048e-1 2.40472e-1    ;    2.4138          0.0575
;OHH2P     HAMM         1              2.06777e-1 2.40472e-1    ;    2.3210          0.0575
```

Dimethyl_Phosphate.top

```
; optimized parameters for Dihydrogen Dimethyl Phosphate CH32PO4- ion with respect to Na+ , NH4+ and CH3NH3+;

[ atomtypes ]
;name   bond_type     mass     charge   ptype   sigma nm   epsilon kJ/mol   rAMBER; A   epsilonAMBER; kcal/mol
 p5       p5         0.00000  0.00000   A     3.74177e-01   8.36800e-01    ;    2.1000        0.2000
 OM2P     OM2P       0.00000  0.00000   A     2.95992e-01   8.78640e-01    ;    1.6612        0.2100
 OSM2P    OSM2P      0.00000  0.00000   A     3.00001e-01   7.11280e-01    ;    1.6837        0.1700
 CT       CT         0.00000  0.00000   A     3.39967e-01   4.57730e-01    ;    1.9080        0.1094
 H1       H1         0.00000  0.00000   A     2.64953e-01   6.56888e-02    ;    1.4870        0.0157
 NA+      NA+        0.00000  0.00000   A     2.43928e-01   3.65846e-01    ;    1.3690        0.0874
 NAMO     NAMO       0.00000  0.00000   A     3.25000e-01   7.11280e-01    ;    1.8240        0.1700
 HAMO     HAMO       0.00000  0.00000   A     1.06908e-01   6.56888e-02    ;    0.6000        0.0157
 NAMM     NAMM       0.00000  0.00000   A     3.25000e-01   7.11280e-01    ;    1.8240        0.1700
 HAMM     HAMM       0.00000  0.00000   A     1.06908e-01   6.56888e-02    ;    0.6000        0.0157
 HP       HP         0.00000  0.00000   A     2.64953e-01   6.56888e-02    ;    1.4870        0.0157


[ nonbond_params ]
;i         j          func            sigma       epsilon          RijAMBER; A    epsilon_ijAMBER, kcal/mol
 OM2P      NA+         1             2.72660e-1 5.66963e-1    ;    3.0605          0.1355
;OM2P      NA+         1             2.69960e-1 5.66963e-1    ;    3.0302          0.1355
 OSM2P     NA+         1             2.74685e-1 5.10117e-1    ;    3.0832          0.1219
;OSM2P     NA+         1             2.71965e-1 5.10117e-1    ;    3.0527          0.1219


 OM2P      NAMO        1             3.13601e-1 7.90543e-1    ;    3.5200          0.1889
;OM2P      NAMO        1             3.10496e-1 7.90543e-1    ;    3.4852          0.1889
 OM2P      HAMO        1             2.03464e-1 2.40243e-1    ;    2.2838          0.0574
;OM2P      HAMO        1             2.01450e-1 2.40243e-1    ;    2.2612          0.0574

 OSM2P     NAMO        1             3.15625e-1 7.11280e-1    ;    3.5428          0.1700
;OSM2P     NAMO        1             3.12500e-1 7.11280e-1    ;    3.5077          0.1700
 OSM2P     HAMO        1             2.05488e-1 2.16155e-1    ;    2.3065          0.0517
;OSM2P     HAMO        1             2.03454e-1 2.16155e-1    ;    2.2837          0.0517


 OM2P      NAMM        1             3.22916e-1 7.90543e-1    ;    3.6246          0.1889
;OM2P      NAMM        1             3.10496e-1 7.90543e-1    ;    3.4852          0.1889
 OM2P      HAMM        1             2.09508e-1 2.40243e-1    ;    2.3516          0.0574
;OM2P      HAMM        1             2.01450e-1 2.40243e-1    ;    2.2612          0.0574

 OSM2P     NAMM        1             3.25000e-1 7.11280e-1    ;    3.6480          0.1700
;OSM2P     NAMM        1             3.12500e-1 7.11280e-1    ;    3.5077          0.1700
 OSM2P     HAMM        1             2.11592e-1 2.16155e-1    ;    2.3750          0.0517
;OSM2P     HAMM        1             2.03454e-1 2.16155e-1    ;    2.2837          0.0517
```

Methyl_Phosphate.top

```
; optimized parameters for Methyl Phosphate CH3PO42- ion with respect to Na+ , NH4+ and CH3NH3+;


[ atomtypes ]
;name   bond_type     mass     charge   ptype   sigma nm   epsilon kJ/mol   rAMBER; A   epsilonAMBER; kcal/mol
 OMPO     OMPO       0.00000  0.00000   A     3.35951e-01   8.78640e-01    ;    1.8855        0.2100
;OMPO     OMPO       0.00000  0.00000   A     2.95992e-01   8.78640e-01    ;    1.6612        0.2100
 p5       p5         0.00000  0.00000   A     3.74177e-01   8.36800e-01    ;    2.1000        0.2000
 OSMPO    OSMPO      0.00000  0.00000   A     3.40501e-01   7.11280e-01    ;    1.9110        0.1700
;OSMPO    OSMPO      0.00000  0.00000   A     3.00001e-01   7.11280e-01    ;    1.6837        0.1700
 CT       CT         0.00000  0.00000   A     3.39967e-01   4.57730e-01    ;    1.9080        0.1094
 H1       H1         0.00000  0.00000   A     2.64953e-01   6.56888e-02    ;    1.4870        0.0157
 NA+      NA+        0.00000  0.00000   A     2.43928e-01   3.65846e-01    ;    1.3690        0.0874
```

```
 NAMO      NAMO        0.00000  0.00000   A     3.25000e-01   7.11280e-01    ;    1.8240        0.1700
 HAMO      HAMO        0.00000  0.00000   A     1.06908e-01   6.56888e-02    ;    0.6000        0.0157
 NAMM      NAMM        0.00000  0.00000   A     3.25000e-01   7.11280e-01    ;    1.8240        0.1700
 HAMM      HAMM        0.00000  0.00000   A     1.06908e-01   6.56888e-02    ;    0.6000        0.0157
 HP        HP          0.00000  0.00000   A     2.64953e-01   6.56888e-02    ;    1.4870        0.0157


[ nonbond_params ]
;i         j           func             sigma       epsilon         RijAMBER; A    epsilon_ijAMBER, kcal/mol
 OMPO      NA+          1              3.05055e-1 5.66963e-1    ;    3.4241          0.1355
;OMPO      NA+          1              2.69960e-1 5.66963e-1    ;    3.0302          0.1355
 OSMPO     NA+          1              3.07320e-1 5.10117e-1    ;    3.4495          0.1219
;OSMPO     NA+          1              2.71965e-1 5.10117e-1    ;    3.0527          0.1219


 OMPO      NAMO         1              3.50860e-1 7.90543e-1    ;    3.9383          0.1889
;OMPO      NAMO         1              3.10496e-1 7.90543e-1    ;    3.4852          0.1889
 OMPO      HAMO         1              2.27638e-1 2.40243e-1    ;    2.5551          0.0574
;OMPO      HAMO         1              2.01450e-1 2.40243e-1    ;    2.2612          0.0574

 OSMPO     NAMO         1              3.53125e-1 7.11280e-1    ;    3.9637          0.1700
;OSMPO     NAMO         1              3.12500e-1 7.11280e-1    ;    3.5077          0.1700
 OSMPO     HAMO         1              2.29903e-1 2.16155e-1    ;    2.5806          0.0517
;OSMPO     HAMO         1              2.03454e-1 2.16155e-1    ;    2.2837          0.0517


 OMPO      NAMM         1              3.57070e-1 7.90543e-1    ;    4.0080          0.1889
;OMPO      NAMM         1              3.10496e-1 7.90543e-1    ;    3.4852          0.1889
 OMPO      HAMM         1              2.31667e-1 2.40243e-1    ;    2.6004          0.0574
;OMPO      HAMM         1              2.01450e-1 2.40243e-1    ;    2.2612          0.0574

 OSMPO     NAMM         1              3.59375e-1 7.11280e-1    ;    4.0338          0.1700
;OSMPO     NAMM         1              3.12500e-1 7.11280e-1    ;    3.5077          0.1700
 OSMPO     HAMM         1              2.33972e-1 2.16155e-1    ;    2.6262          0.0517
;OSMPO     HAMM         1              2.03454e-1 2.16155e-1    ;    2.2837          0.0517
```

Methyl_Sulfate.top

```
; optimized parameters for Methyl_Sulfate ionCH3SO4- ion with respect to Na+ , NH4+ and CH3NH3+;


[ atomtypes ]
;name    bond_type     mass     charge   ptype   sigma         epsilon           rAMBER    epsilonAMBER
 OMSO4     OMSO4       0.00000  0.00000   A     2.95992e-01   8.78640e-01    ;    1.6612        0.2100
 s6        s6          0.00000  0.00000   A     3.56359e-01   1.04600e+00    ;    2.0000        0.2500
 OSMSO     OSMSO       0.00000  0.00000   A     3.00001e-01   7.11280e-01    ;    1.6837        0.1700
 CT        CT          0.00000  0.00000   A     3.39967e-01   4.57730e-01    ;    1.9080        0.1094
 H1        H1          0.00000  0.00000   A     2.64953e-01   6.56888e-02    ;    1.4870        0.0157
 NA+       NA+         0.00000  0.00000   A     2.43928e-01   3.65846e-01    ;    1.3690        0.0874
 NAMO      NAMO        0.00000  0.00000   A     3.25000e-01   7.11280e-01    ;    1.8240        0.1700
 HAMO      HAMO        0.00000  0.00000   A     1.06908e-01   6.56888e-02    ;    0.6000        0.0157
 NAMM      NAMM        0.00000  0.00000   A     3.25000e-01   7.11280e-01    ;    1.8240        0.1700
 HAMM      HAMM        0.00000  0.00000   A     1.06908e-01   6.56888e-02    ;    0.6000        0.0157
 HP        HP          0.00000  0.00000   A     2.64953e-01   6.56888e-02    ;    1.4870        0.0157


[ nonbond_params ]
;i         j           func             sigma       epsilon         RijAMBER; A    epsilon_ijAMBER, kcal/mol
 OMSO4     NA+          1              2.69960e-1 6.23659e-1    ;    3.0302          0.1491
;OMSO4     NA+          1              2.69960e-1 5.66963e-1    ;    3.0302          0.1355
 OSMSO     NA+          1              2.71965e-1 5.61129e-1    ;    3.0527          0.1341
;OSMSO     NA+          1              2.71965e-1 5.10117e-1    ;    3.0527          0.1219
```

```
 OMSO4     NAMO         1             3.10496e-1 9.48652e-1    ;    3.4852          0.2267
;OMSO4     NAMO         1             3.10496e-1 7.90543e-1    ;    3.4852          0.1889
 OMSO4     HAMO         1             2.01450e-1 2.88292e-1    ;    2.2612          0.0689
;OMSO4     HAMO         1             2.01450e-1 2.40243e-1    ;    2.2612          0.0574

 OSMSO     NAMO         1             3.12500e-1 8.53536e-1    ;    3.5077          0.2040
;OSMSO     NAMO         1             3.12500e-1 7.11280e-1    ;    3.5077          0.1700
 OSMSO     HAMO         1             2.03454e-1 2.59386e-1    ;    2.2837          0.0620
;OSMSO     HAMO         1             2.03454e-1 2.16155e-1    ;    2.2837          0.0517


 OMSO4     NAMM         1             3.26021e-1 7.90543e-1    ;    3.6595          0.1889
;OMSO4     NAMM         1             3.10496e-1 7.90543e-1    ;    3.4852          0.1889
 OMSO4     HAMM         1             2.11522e-1 2.40243e-1    ;    2.3743          0.0574
;OMSO4     HAMM         1             2.01450e-1 2.40243e-1    ;    2.2612          0.0574

 OSMSO     NAMM         1             3.28125e-1 7.11280e-1    ;    3.6830          0.1700
;OSMSO     NAMM         1             3.12500e-1 7.11280e-1    ;    3.5077          0.1700
 OSMSO     HAMM         1             2.13627e-1 2.16155e-1    ;    2.3978          0.0517
;OSMSO     HAMM         1             2.03454e-1 2.16155e-1    ;    2.2837          0.0517
```

Methyl_Sulfonate.top

```
; optimized parameters for Methyl_Sulfonate ion CH3SO3- ion with respect to Na+ , NH4+ and CH3NH3+;

[ atomtypes ]
;name   bond_type     mass     charge   ptype   sigma         epsilon           rAMBER    epsilonAMBER
 OMSO3     OMSO3       0.00000  0.00000   A     2.95992e-01   8.78640e-01    ;    1.6612        0.2100
 s6        s6          0.00000  0.00000   A     3.56359e-01   1.04600e+00    ;    2.0000        0.2500
 CT        CT          0.00000  0.00000   A     3.39967e-01   4.57730e-01    ;    1.9080        0.1094
 H1        H1          0.00000  0.00000   A     2.64953e-01   6.56888e-02    ;    1.4870        0.0157
 NA+       NA+         0.00000  0.00000   A     2.43928e-01   3.65846e-01    ;    1.3690        0.0874

 NAMO      NAMO        0.00000  0.00000   A     3.25000e-01   7.11280e-01    ;    1.8240        0.1700
 HAMO      HAMO        0.00000  0.00000   A     1.06908e-01   6.56888e-02    ;    0.6000        0.0157
 NAMM      NAMM        0.00000  0.00000   A     3.25000e-01   7.11280e-01    ;    1.8240        0.1700
 HAMM      HAMM        0.00000  0.00000   A     1.06908e-01   6.56888e-02    ;    0.6000        0.0157
 HP        HP          0.00000  0.00000   A     2.64953e-01   6.56888e-02    ;    1.4870        0.0157


[ nonbond_params ]
;i         j           func          sigma       epsilon         RijAMBER; A    epsilon_ijAMBER, kcal/mol
 OMSO3     NA+          1            2.69960e-1 6.23659e-1    ;    3.0302          0.1491
;OMSO3     NA+          1            2.69960e-1 5.66963e-1    ;    3.0302          0.1355

 OMSO3     NAMO         1            3.10496e-1 8.69598e-1    ;    3.4852          0.2078
;OMSO3     NAMO         1            3.10496e-1 7.90543e-1    ;    3.4852          0.1889
 OMSO3     HAMO         1            2.01450e-1 2.64267e-1    ;    2.2612          0.0632
;OMSO3     HAMO         1            2.01450e-1 2.40243e-1    ;    2.2612          0.0574

 OMSO3     NAMM         1            3.26021e-1 7.90543e-1    ;    3.6595          0.1889
;OMSO3     NAMM         1            3.10496e-1 7.90543e-1    ;    3.4852          0.1889
 OMSO3     HAMM         1            2.11522e-1 2.40243e-1    ;    2.3742          0.0574
;OMSO3     HAMM         1            2.01450e-1 2.40243e-1    ;    2.2612          0.0574
```

Sulfate.top

```
; optimized parameters for sulfate SO42- ion with respect to Na+ , NH4+ and CH3NH3+;
```

```
[ atomtypes ]
;name    bond_type     mass     charge   ptype   sigma nm    epsilon kJ/mol   rAMBER; A   epsilonAMBER; kcal/mol
 OSUL      OSUL        0.00000  0.00000   A     3.46311e-01   8.78640e-01    ;    1.9436        0.2100
;OSUL      OSUL        0.00000  0.00000   A     2.95992e-01   8.78640e-01    ;    1.6612        0.2100
 s6        s6          0.00000  0.00000   A     3.56359e-01   1.04600e+00    ;    2.0000        0.2500
 NA+       NA+         0.00000  0.00000   A     2.43928e-01   3.65846e-01    ;    1.3690        0.0874
 CT        CT          0.00000  0.00000   A     3.39967e-01   4.57730e-01    ;    1.9080        0.1094
 NAMO      NAMO        0.00000  0.00000   A     3.25000e-01   7.11280e-01    ;    1.8240        0.1700
 HAMO      HAMO        0.00000  0.00000   A     1.06908e-01   6.56888e-02    ;    0.6000        0.0157
 NAMM      NAMM        0.00000  0.00000   A     3.25000e-01   7.11280e-01    ;    1.8240        0.1700
 HAMM      HAMM        0.00000  0.00000   A     1.06908e-01   6.56888e-02    ;    0.6000        0.0157
 HP        HP          0.00000  0.00000   A     2.64953e-01   6.56888e-02    ;    1.4870        0.0157


[ nonbond_params ]
;i         j           func             sigma       epsilon          RijAMBER; A    epsilon_ijAMBER, kcal/mol
 OSUL      NA+          1              3.05055e-1 5.66963e-1    ;    3.4241          0.1355
;OSUL      NA+          1              2.69960e-1 5.66963e-1    ;    3.0302          0.1355

 OSUL      NAMO         1              3.50860e-1 7.90543e-1    ;    3.9383          0.1889
;OSUL      NAMO         1              3.10496e-1 7.90543e-1    ;    3.4852          0.1889
 OSUL      HAMO         1              2.27638e-1 2.40243e-1    ;    2.5551          0.0574
;OSUL      HAMO         1              2.01450e-1 2.40243e-1    ;    2.2612          0.0574

 OSUL      NAMM         1              3.53965e-1 7.90543e-1    ;    3.9731          0.1889
;OSUL      NAMM         1              3.10496e-1 7.90543e-1    ;    3.4852          0.1889
 OSUL      HAMM         1              2.29653e-1 2.40243e-1    ;    2.5778          0.0574
;OSUL      HAMM         1              2.01450e-1 2.40243e-1    ;    2.2612          0.0574
```

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Dissertationsschrift mit dem Titel

Developing Ion Parameters and Investigating Ion Specific Effects in Biological Systems

selbständig angefertigt und hierfür keine anderen als die angegebenen Hilfsmittel verwendet habe. Die Arbeit ist weder in einem früheren Promotionsverfahren angenommen noch als ungenügend beurteilt worden.

Berlin, den 10. April 2017

Publications

Publications related to this thesis:

- "Is the solution activity derivative sufficient to parameterize ion-ion interactions in aqueous solution?-ions for TIP5P water"[147].

Parts of this thesis and the appendices have previously been published. Reprinted (adapted) with permission from ([147]) Copyright (2017) American Chemical Society.

- "Developing force fields when experimental data is sparse: AMBER/GAFF-compatible parameters for inorganic and alkyl oxoanions"[148].

Parts of this thesis and the appendices have previously been published. Reprinted (adapted) with permission from ([148]) - Published by the PCCP Owner Societies.

Acknowledgements

I would like to thank my supervisor, Ana Vila Verde, for her guidance and support during my PhD studies. Thank you for your patience, kindness and the great feedbacks and advice. I thank my official advisor, Roland Netz, for accepting me in his group and for the informative course works, as well as the great teaching opportunity he offered me. I also thank Petra Imhof, for accepting to be my second referee.

I thank Narges Moradi, for her help and support. Thank you for being there for me, during these three years.

I am very thankful to my office mate Joaõ Robalo for the great discussions, sincere help and most importantly the coffee breaks.

I thank my friends and family and beloved ones, specially my sister and brother, Sara and Sina, and my parents, Fatemeh Samadi and Nassir Kashef Ol Gheta, who passionately supported me throughout my studies.

Bibliography

[1] H. G. Garg, L. Yu, C. A. Hales, T. Toida, T. Islam, and R. J. Linhardt, "Sulfation patterns in heparin and heparan sulfate: Effects on the proliferation of bovine pulmonary artery smooth muscle cells," *Biochimica et Biophysica Acta - Molecular Basis of Disease*, vol. 1639, no. 3, pp. 225–231, 2003.

[2] S. A. Smith, N. J. Mutch, D. Baskar, P. Rohloff, R. Docampo, and J. H. Morrissey, "Polyphosphate modulates blood coagulation and fibrinolysis.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 4, pp. 903–908, 2006.

[3] J. T. Gallagher, "Heparin - A Century of Progress," *Handbook of Experimental Pharmacology*, vol. 207, pp. 347–360, 2012.

[4] R. Baldwin, "How Hofmeister ion interactions affect protein stability," *Biophysical Journal*, vol. 71, no. 4, pp. 2056–2063, 1996.

[5] K. D. Collins, "Ions from the Hofmeister series and osmolytes: Effects on proteins in solution and in the crystallization process," *Methods*, vol. 34, no. 3, pp. 300–311, 2004.

[6] C. F. Anderson and M. T. Record, "Salt-Nucleic Acid Interactions," *Annual Review of Physical Chemistry*, vol. 46, no. 1, pp. 657–700, 1995.

[7] S. Mclaughlin, "The electrostatic properties of membranes.," *Annual review of biophysics and biophysical chemistry*, vol. 18, pp. 113–136, 1989.

[8] E. M. Knipping, "Experiments and Simulations of Ion-Enhanced Interfacial Chemistry on Aqueous NaCl Aerosols," *Science*, vol. 288, no. 5464, pp. 301–306, 2000.

[9] F. Hofmeister, "Zur Lehre von der Wirkung der Salze - Zweite Mittheilung," *Arch. für Exp. Pathol. und Pharmakologie*, vol. 24, no. 4-5, pp. 247–260, 1888.

[10] W. Kunz, *Specific Ion Effects*. World Scientific Publishing Co Pte Ltd, 2009.

[11] N. Schwierz, D. Horinek, and R. R. Netz, "Anionic and cationic hofmeister effects on hydrophobic and hydrophilic surfaces," *Langmuir*, vol. 29, no. 8, pp. 2602–2614, 2013.

[12] Y. Zhang and P. S. Cremer, "Chemistry of Hofmeister anions and osmolytes.," *Annual review of physical chemistry*, vol. 61, pp. 63–83, 2010.

[13] A. Sigel, H. Sigel, and R. K. Sigel, *Interrelations between Essential Metal Ions and Human Diseases*, vol. 13. 2013.

[14] B. Klasczyk, V. Knecht, R. Lipowsky, and R. Dimova, "Interactions of alkali metal chlorides with phosphatidylcholine vesicles," *Langmuir*, vol. 26, no. 24, pp. 18951–18958, 2010.

[15] P. Melnikov and L. Z. Zanoni, "Clinical effects of cesium intake," *Biological Trace Element Research*, vol. 135, no. 1-3, pp. 1–9, 2010.

[16] Meltzer H. L., Taylor R. M., Platman S. R., and Fieve R. R., "Rubidium: a Potential Modifier of Affect and Behaviour," *Nature*, vol. 223, no. 5203, pp. 321–322, 1969.

[17] F. N. Johnson, "The psychopharmacology of lithium," *Neuroscience and Biobehavioral Reviews*, vol. 3, no. 1, pp. 15–30, 1979.

[18] H. Aral and A. Vecchio-Sadus, "Toxicity of lithium to humans and the environmental literature review," *Ecotoxicology and Environmental Safety*, vol. 70, no. 3, pp. 349–356, 2008.

[19] M. Eisenberg, T. Gresalfi, T. Riccio, and S. McLaughlin, "Adsorption of monovalent cations to bilayer membranes containing negative phospholipids," *Biochemistry*, vol. 18, no. 23, pp. 5213–5223, 1979.

[20] D. N. Sheppard and M. J. Welsh, "Structure and function of the CFTR chloride channel.," *Physiological reviews*, vol. 79, no. 1 Suppl, pp. S23–S45, 1999.

[21] R. Dutzler, E. B. Campbel, M. Cadene, B. T. Chait, and R. MacKinnon, "X-ray structure of a ClC chloride channel at 3.0A reveals the molecular basis of anion selectivity," *Nature*, vol. 415, no. January, pp. 287–294, 2002.

[22] A. L. Goldin, R. L. Barchi, J. H. Caldwell, F. Hofmann, J. R. Howe, J. C. Hunter, R. G. Kallen, G. Mandel, M. H. Meisler, Y. B. Netter, M. Noda, M. M. Tamkun, S. G. Waxman, J. N. Wood, and W. A. Catterall, "Nomenclature of Voltage-Gated Sodium Channels," *Neuron*, vol. 28, no. 2, pp. 365–368, 2000.

[23] H. Hibino, A. Inanobe, K. Furutani, S. Murakami, I. Findlay, and Y. Kurachi, "Inwardly Rectifying Potassium Channels: Their Structure, Function, and Physiological Roles," *Physiology Reviews*, vol. 90, pp. 291–366, 2010.

[24] C. Vergara, R. Latorre, N. V. Marrion, and J. P. Adelman, "Calcium-activated potassium channels," *Current Opinion in Neurobiology*, vol. 8, pp. 321–329, jun 1998.

[25] T. F. McDonald, S. Pelzer, W. Trautwein, and D. J. Pelzer, "Regulation and modulation of calcium channels in cardiac, skeletal, and smooth muscle cells," *Physiol Rev*, vol. 74, no. 2, pp. 365–507, 1994.

[26] S. D. Rosen and C. R. Bertozzi, "Two selectins converge on sulphate. Leukocyte adhesion.," *Current biology : CB*, vol. 6, no. 3, pp. 261–264, 1996.

[27] P. Mehlen and A. Puisieux, "Metastasis: a question of life or death.," *Nature reviews. Cancer*, vol. 6, no. 6, pp. 449–58, 2006.

[28] M. Weinhart, D. Gröger, S. Enders, S. B. Riese, J. Dernedde, R. K. Kainthan, D. E. Brooks, and R. Haag, "The Role of Dimension in Multivalent Binding Events: Structure-Activity Relationship of Dendritic Polyglycerol Sulfate Binding to L-Selectin in Correlation with Size and Surface Charge Density," *Macromolecular Bioscience*, vol. 11, no. 8, pp. 1088–1098, 2011.

[29] R. Malhotra, N. R. Taylor, and M. I. Bird, "Anionic phospholipids bind to L-selectin (but not E-selectin) at a site distinct From the Carbohydrate-Binding Site," *Structure*, vol. 303, pp. 297–303, 1996.

[30] A. L. Woelke, C. Kuehne, T. Meyer, G. Galstyan, J. Dernedde, and E. W. Knapp, "Understanding selectin counter-receptor binding from electrostatic energy computations and experimental binding studies," *Journal of Physical Chemistry B*, vol. 117, no. 51, pp. 16443–16454, 2013.

[31] M. Weinhart, D. Gröger, S. Enders, J. Dernedde, and R. Haag, "Synthesis of dendritic polyglycerol anions and their efficiency toward L-selectin inhibition.," *Biomacromolecules*, vol. 12, no. 7, pp. 2502–11, 2011.

[32] Y. Tong, A. Vila Verde, and R. K. Campen, "The free OD at the air/D2O interface is structurally and dynamically heterogeneous.," *The journal of physical chemistry. B*, vol. 117, no. 39, pp. 11753–64, 2013.

[33] A. Vila Verde, P. G. Bolhuis, and R. K. Campen, "Statics and dynamics of free and hydrogen-bonded OH groups at the air/water interface.," *The journal of physical chemistry. B*, vol. 116, no. 31, pp. 9467–81, 2012.

[34] A. Vila Verde, P. J. Beltramo, and J. K. Maranas, "Adsorption of homopolypeptides on gold investigated using atomistic molecular dynamics.," *Langmuir : the ACS journal of surfaces and colloids*, vol. 27, no. 10, pp. 5918–26, 2011.

[35] a. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus, "All-atom empirical potential for molecular modeling and dynamics studies of proteins.," *J. Phys. Chem. B*, vol. 102, no. 18, pp. 3586–3616, 1998.

[36] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, "A Second Generation Force Field for the Simulation of Proteins , Nucleic Acids , and Organic Molecules," *Journal of American Chemical Society*, no. 6, pp. 5179–5197, 1995.

[37] C. Vega, J. L. F. Abascal, M. M. Conde, and J. L. Aragones, "What ice can teach us about water interactions: a critical comparison of the performance of different water models," *Faraday Discuss.*, vol. 141, pp. 251–276, 2009.

[38] J. Sauter and A. Grafmüller, "Solution Properties of Hemicellulose Polysaccharides with Four Common Carbohydrate Force Fields.," *Journal of chemical theory and computation*, vol. 11, no. 4, pp. 1765–74, 2015.

[39] W. L. Ash, M. R. Zlomislic, E. O. Oloo, and D. P. Tieleman, "Computer simulations of membrane proteins," *Biochim. Biophys. Acta - Biomembr.*, vol. 1666, no. 1-2, pp. 158–189, 2004.

[40] D. E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R. O. Dror, M. P. Eastwood, J. A. Bank, J. M. Jumper, J. K. Salmon, Y. Shan, and W. Wriggers, "Atomic-Level Characterization of the Structural Dynamics of Proteins," *Science*, vol. 330, no. 6002, pp. 341–346, 2010.

[41] K. Lindorff-Larsen, N. Trbovic, P. Maragakis, S. Piana, and D. E. Shaw, "Structure and Dynamics of an Unfolded Protein Examined by Molecular Dynamics Simulation," *Journal of the American Chemical Society*, vol. 134, no. 8, pp. 3787–3791, 2012.

[42] C. Zhang and J. Ma, "Enhanced sampling and applications in protein folding in explicit solvent," *The Journal of Chemical Physics*, vol. 132, no. 24, p. 244101, 2010.

[43] A. D. Mackerell and M. Karplus, "Importance of Attractive van der Waals Contribution in Empirical Energy Function Models for the Heat of Vaporization of Polar Liquids," *The Journal of Physical Chemistry*, vol. 2, no. 6, pp. 10559–10560, 1991.

[44] I. S. Joung and T. E. Cheatham, "Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations.," *The journal of physical chemistry. B*, vol. 112, no. 30, pp. 9020–41, 2008.

[45] B. Klasczyk and V. Knecht, "Kirkwood-Buff derived force field for alkali chlorides in simple point charge water," *J. Chem. Phys.*, vol. 132, no. 2, 2010.

[46] D. Horinek, S. I. Mamatkulov, and R. R. Netz, "Rational design of ion force fields based on thermodynamic solvation properties," *Journal of Chemical Physics*, vol. 130, no. 12, p. 124507, 2009.

[47] M. Fyta, I. Kalcher, J. Dzubiella, L. Vrbka, and R. R. Netz, "Ionic force field optimization based on single-ion and ion-pair solvation properties.," *The Journal of chemical physics*, vol. 132, no. 2, p. 024911, 2010.

[48] M. Fyta and R. R. Netz, "Ionic force field optimization based on single-ion and ion-pair solvation properties: Going beyond standard mixing rules," *Journal of Chemical Physics*, vol. 136, no. 12, pp. 0–11, 2012.

[49] M. Kanduc, A. Schlaich, E. Schneck, and R. R. Netz, "Water-Mediated Interactions between Hydrophilic and Hydrophobic Surfaces," *Langmuir*, vol. 32, pp. 8767–8782, 2016.

[50] M. W. Mahoney and W. L. Jorgensen, "A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions," *The Journal of Chemical Physics*, vol. 112, no. 20, pp. 8910–8922, 2000.

[51] J. Sauter and A. Grafmüller, "Predicting the Chemical Potential and Osmotic Pressure of Polysaccharide Solutions by Molecular Simulations.," *Journal of chemical theory and computation*, vol. 12, no. 9, pp. 4375–84, 2016.

[52] I. Gladich, P. Shepson, I. Szleifer, and M. Carignano, "Halide and sodium ion parameters for modeling aqueous solutions in TIP5P-Ew water," *Chemical Physics Letters*, vol. 489, no. 1, pp. 113–117, 2010.

[53] D. R. Nutt and J. C. Smith, "Molecular dynamics simulations of proteins: Can the explicit water model be varied?," *J. Chem. Theory Comput.*, vol. 3, no. 4, pp. 1550–1560, 2007.

[54] P. S. Nerenberg, B. Jo, C. So, A. Tripathy, and T. Head-Gordon, "Optimizing solute-water van der Waals interactions to reproduce solvation free energies.," *The journal of physical chemistry. B*, vol. 116, no. 15, pp. 4524–34, 2012.

[55] P. S. Nerenberg and T. Head-Gordon, "Optimizing Protein-Solvent Force Fields to Reproduce Intrinsic Conformational Preferences of Model Peptides.," *Journal of chemical theory and computation*, vol. 7, no. 4, pp. 1220–30, 2011.

[56] I. S. Joung and T. E. Cheatham, "Molecular dynamics simulations of the dynamic and energetic properties of alkali and halide ions using water-model-specific ion parameters.," *The journal of physical chemistry. B*, vol. 113, no. 40, pp. 13279–90, 2009.

[57] S. Karunaweera, M. B. Gee, S. Weerasinghe, and P. E. Smith, "Theory and Simulation of Multicomponent Osmotic Systems.," *Journal of chemical theory and computation*, vol. 8, no. 10, pp. 3493–3503, 2012.

[58] E. A. Ploetz, N. Bentenitis, and P. E. Smith, "Developing force fields from the microscopic structure of solutions," *Fluid Phase Equilibria*, vol. 290, no. 1, pp. 43–47, 2010.

[59] P. Hünenberger and M. Reif, *Single-ion solvation - experimental and theoretical approaches to elusive thermodynamic quantities*. RSC Theoretical and Computational Chemistry Series, RSC, 2011.

[60] P. Krüger, S. K. Schnell, D. Bedeaux, S. Kjelstrup, T. J. H. Vlugt, and J. M. Simon, "Kirkwood-Buff integrals for finite volumes," *J. Phys. Chem. Lett.*, vol. 4, no. 2, pp. 235–238, 2013.

[61] S. Mamatkulov, M. Fyta, and R. R. Netz, "Force fields for divalent cations based on single-ion and ion-pair properties.," *The Journal of chemical physics*, vol. 138, no. 2, p. 024505, 2013.

[62] J. Yoo and A. Aksimentiev, "Improved Parameterization of Amine-Carboxylate and Amine-Phosphate Interactions for Molecular Dynamics Simulations Using the CHARMM and AMBER Force Fields," *Journal of Chemical Theory and Computation*, vol. 12, no. 1, pp. 430–443, 2016.

[63] K. D. Collins, "Why continuum electrostatics theories cannot explain biological structure, polyelectrolytes or ionic strength effects in ion-protein interactions," *Biophysical Chemistry*, vol. 167, pp. 33–49, 2012.

[64] Y. Zhang and P. S. Cremer, "Interactions between macromolecules and ions: the Hofmeister series," 2006.

[65] C. C. I. Bayly, P. Cieplak, W. D. Cornell, and P. a. Kollman, "A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model," *The Journal of Physical Chemistry*, vol. 97, pp. 10269–10280, 1993.

[66] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," *The Journal of Chemical Physics*, vol. 79, no. 2, p. 926, 1983.

[67] Y. Marcus, *Ion properties*. New York, USA: Marcel Dekker, Inc., 1997.

[68] E. Pluharova, O. Marsalek, B. Schmidt, and P. Jungwirth, "Ab Initio Molecular Dynamics Approach to a Quantitative Description of Ion Pairing in Water," *Journal of Physical Chemistry Letters*, vol. 4, no. 23, pp. 4177–4181, 2013.

[69] W. L. Jorgensen, "Optimized intermolecular potential functions for liquid alcohols," *J. Phys. Chem.*, vol. 90, no. 7, pp. 1276–1284, 1986.

[70] I. Leontyev and A. Stuchebrukhov, "Accounting for electronic polarization in non-polarizable force fields," *Phys. Chem. Chem. Phys.*, vol. 13, no. 7, pp. 2613–2626, 2011.

[71] E. b. B. Pullman, *Intermolecular Forces*. Springer-Science+Business Media, B.V., volume 14 ed., 1981.

[72] R. E. Johnson, *Introduction to Atomic and Molecular Collisions*. New York: Springer US, 1982.

[73] M. M. Reif and P. H. Hünenberger, "Computation of methodology-independent single-ion solvation properties from molecular simulations. IV. Optimized Lennard-Jones interaction parameter sets for the alkali and halide ions in water.," *The Journal of chemical physics*, vol. 134, no. 14, p. 144104, 2011.

[74] J. C. Slater and J. G. Kirkwood, "The Van Der Waals Forces in Gases," *Physical Review*, vol. 37, no. 6, pp. 682–697, 1931.

[75] I. S. Lim, J. K. Laerdahl, and P. Schwerdtfeger, "Fully relativistic coupled-cluster static dipole polarizabilities of the positively charged alkali ions from Li^+ to 119^+," *The Journal of Chemical Physics*, vol. 116, no. 1, p. 172, 2002.

[76] I. Miadokova, V. Kellö, and A. J. Sadlej, "Standardized basis sets for high-level-correlated relativistic calculations of atomic and molecular electric properties in the spin-averaged Douglas-Kroll (no-pair) approximation I. Groups Ib and IIb," *Theor. Chim. Acta*, vol. 96, pp. 166–175, 1997.

[77] I. S. Lim and P. Schwerdtfeger, "Four-component and scalar relativistic Douglas-Kroll calculations for static dipole polarizabilities of the alkaline-earth-metal elements and their ions from Ca^n to Ra^n (n=0, +1, +2)," *Physical Review A*, vol. 70, no. 6, p. 062501, 2004.

[78] N. C. Pyper, C. G. Pike, and P. P. Edwards, "The polarizabilities of species present in ionic solutions," *Mol. Phys.*, vol. 76, no. January 2015, pp. 353–372, 1992.

[79] A. D. Becke and E. R. Johnson, "Exchange-hole dipole moment and the dispersion interaction: high-order dispersion coefficients.," *The Journal of chemical physics*, vol. 124, no. 1, p. 14104, 2006.

[80] William M. Haynes, *CRC Handbook of Chemistry and Physics*. CRC Press, 2016.

[81] H. J. C. Berendsen, D. van der Spoel, and R. van Drunen, "GROMACS: A message-passing parallel molecular dynamics implementation," *Computer Physics Communications*, vol. 91, no. 1-3, pp. 43–56, 1995.

[82] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen, "GROMACS: fast, flexible, and free.," *Journal of computational chemistry*, vol. 26, no. 16, pp. 1701–18, 2005.

[83] L. Martínez, R. Andrade, E. G. Birgin, and J. M. Martínez, "PACKMOL: a package for building initial configurations for molecular dynamics simulations.," *Journal of computational chemistry*, vol. 30, no. 13, pp. 2157–64, 2009.

[84] T. Darden, D. York, and L. Pedersen, "Particle mesh Ewald: An N.log(N) method for Ewald sums in large systems," *The Journal of Chemical Physics*, vol. 98, no. 12, p. 10089, 1993.

[85] N. Goga, A. J. Rzepiela, A. H. de Vries, S. J. Marrink, and H. J. C. Berendsen, "Efficient Algorithms for Langevin and DPD Dynamics.," *Journal of chemical theory and computation*, vol. 8, no. 10, pp. 3637–49, 2012.

[86] S. Nosé, "A molecular dynamics method for simulations in the canonical ensemble," *Molecular Physics*, vol. 52, no. 2, pp. 255–268, 1984.

[87] W. G. Hoover, "Canonical dynamics: Equilibrium phase-space distributions," *Physical Review A*, vol. 31, no. 3, pp. 1695–1697, 1985.

[88] M. Parrinello, "Polymorphic transitions in single crystals: A new molecular dynamics method," *Journal of Applied Physics*, vol. 52, no. 12, p. 7182, 1981.

[89] T. Steinbrecher, J. Latzer, and D. a. Case, "Revised AMBER parameters for bioorganic phosphates," *J. Chem. Theory Comput.*, vol. 8, no. 11, pp. 4405–4412, 2012.

[90] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, "Development and testing of a general amber force field," *J. Comput. Chem.*, vol. 25, no. 9, pp. 1157–1174, 2004.

[91] J. Wang, P. Cieplak, and P. A. Kollman, "How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules?," *J. Comput. Chem.*, vol. 21, no. 12, pp. 1049–1074, 2000.

[92] W. L. Jorgensen and J. Tirado-Rives, "The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin," *J. Am. Chem. Soc.*, vol. 110, no. 6, pp. 1657–1666, 1988.

[93] P. Hünenberger, "Thermostat algorithms for molecular dynamics simulations," in *Advanced Computer Simulation Approaches for Soft Matter Sciences I*, vol. 173, pp. 105–147, 2005.

[94] C. H. Bennett, "Efficient estimation of free energy differences from Monte Carlo data," *Journal of Computational Physics*, vol. 22, no. 2, pp. 245–268, 1976.

[95] W. F. van Gunsteren and J. C. Berendsen, "A Leap-frog Algorithm for Stochastic Dynamics," *Molecular Simulation*, vol. 1, no. May 2012, pp. 173–185, 1988.

[96] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, a. DiNola, and J. R. Haak, "Molecular dynamics with coupling to an external bath," *The Journal of Chemical Physics*, vol. 81, no. 1984, pp. 3684–3690, 1984.

[97] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije, "LINCS: A linear constraint solver for molecular simulations," *Journal of Computational Chemistry*, vol. 18, no. 12, pp. 1463–1472, 1997.

[98] M. A. Kastenholz and P. H. Hünenberger, "Computation of methodology-independent ionic solvation free energies from molecular simulations. II. The hydration free energy of the sodium cation.," *The Journal of chemical physics*, vol. 124, no. 22, p. 224501, 2006.

[99] M. A. Kastenholz and P. H. Hünenberger, "Computation of methodology-independent ionic solvation free energies from molecular simulations. I. The electrostatic potential in molecular liquids.," *The Journal of chemical physics*, vol. 124, no. 12, p. 124106, 2006.

[100] M. M. Reif and P. H. Hünenberger, "Computation of methodology-independent single-ion solvation properties from molecular simulations. III. Correction terms for the solvation free energies, enthalpies, entropies, heat capacities, volumes, compressibilities, and expansivities of solvated ion," *The Journal of chemical physics*, vol. 134, no. 14, p. 144103, 2011.

[101] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman, "The weighted histogram analysis method for free energy calculations on biomolecules. I. The method," *Journal of Computational Chemistry*, vol. 13, no. 8, pp. 1011–1021, 1992.

[102] U. C. Singh and P. a. Kollman, "An approach to computing electrostatic charges for molecules," *J. Comput. Chem.*, vol. 5, no. 2, pp. 129–145, 1984.

[103] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, "Gaussian 03, Revision C.02."

[104] C. E, B. Mennucci, and J. Tomasi, "A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics," *The Journal of Chemical Physics*, vol. 107, no. 8, p. 3032, 1997.

[105] B. Mennucci, E. Cancès, and J. Tomasi, "Evaluation of Solvent Effects in Isotropic and Anisotropic Dielectrics and in Ionic Solutions with a Unified Integral Equation Method: Theoretical Bases, Computational Implementation, and Numerical Applications," *The Journal of Physical Chemistry B*, vol. 101, no. 49, pp. 10506–10517, 1997.

[106] J. Tomasi, B. Mennucci, and E. Cancès, "The IEF version of the PCM solvation method: an overview of a new method addressed to study molecular solutes at the QM ab initio level," *Journal of Molecular Structure (Theochem)*, vol. 464, pp. 211–226, 1999.

[107] S. Miertuš, E. Scrocco, and J. Tomasi, "Electrostatic interaction of a solute with a continuum. A direct utilizaion of AB initio molecular potentials for the prevision of solvent effects," *Chemical Physics*, vol. 55, no. 1, pp. 117–129, 1981.

[108] S. Miertuš and J. Tomasi, "Approximate evaluations of the electrostatic free energy and internal energy changes in solution processes," *Chemical Physics*, vol. 65, no. 2, pp. 239–245, 1982.

[109] A. D. Becke, "Density-functional thermochemistry. III. The role of exact exchange," *The Journal of Chemical Physics*, vol. 98, no. 7, p. 5648, 1993.

[110] a. D. Becke, "Density-functional exchange-energy approximation with correct asymptotic behavior," *Physical Review A*, vol. 38, no. 6, pp. 3098–3100, 1988.

[111] C. Lee, W. Yang, and R. G. Parr, "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density," *Phys. Rev. B*, vol. 37, no. 2, pp. 785–789, 1988.

[112] C. Møller and M. S. Plesset, "Note on an approximation treatment for many-electron systems," *Phys. Rev.*, vol. 46, no. 7, pp. 618–622, 1934.

[113] T. H. Dunning Jr, "Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen," *J. Chem. Phys.*, vol. 90, no. 1989, p. 1007, 1989.

[114] B. Liu and a. D. Mclean, "Accurate calculation of the attractive interaction of two ground state helium atoms," *Journal of Chemical Physics*, vol. 59, no. 8, pp. 4557–4558, 1973.

[115] C. Sherrill, "Counterpoise Correction and Basis Set Superposition Error," pp. 1–5, 2010.

[116] S. Boys and F. Bernardi, "The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors," *Molecular Physics*, vol. 19, pp. 553–566, 1970.

[117] C. Chipot and A. Pohorille, "Free Energy Calculations Theory and Applications in Chemistry and Biology," in *Springer Series in chemical physics*, p. volume 86, Springer, 2007.

[118] M. Abraham, D. van der Spoel, E. Lindahl, , B. Hess, and t. G. team Development, "GROMACS User Manual version 5.0.4,," 2014.

[119] J. G. Kirkwood and F. P. Buff, "The statistical mechanical theory of solutions. I," *Journal of Chemical Physics*, vol. 19, no. 1951, pp. 774–7, 1951.

[120] P. G. Kusalik and G. N. Patey, "The thermodynamic properties of electrolyte solutions: Some formal results," *J. Chem. Phys.*, vol. 86, no. 9, pp. 5110–5116, 1987.

[121] E. H. Rubensson, "Nonmonotonic recursive polynomial expansions for linear scaling calculation of the density matrix," *Journal of Chemical Theory and Computation*, vol. 7, no. 5, pp. 1233–1236, 2011.

[122] M. B. Gee, N. R. Cox, Y. Jiao, N. Bentenitis, S. Weerasinghe, and P. E. Smith, "A Kirkwood-Buff derived force field for aqueous alkali halides," *J. Chem. Theory Comput.*, vol. 7, no. 5, pp. 1369–1380, 2011.

[123] A. Vila Verde, M. Santer, and R. Lipowsky, "Solvent-shared pairs of densely charged ions induce intense but short-range supra-additive slowdown of water rotation," *Phys. Chem. Chem. Phys.*, vol. 18, pp. 1918–1930, 2016.

[124] J. L. Lebowitz and J. K. Percus, "Long-Range Correlations in a Closed System with Applications to Nonuniform Fluids," *Physical Review*, vol. 122, no. 6, pp. 1675–1691, 1961.

[125] A. P. Lyubartsev and S. Marčelja, "Evaluation of effective ion-ion potentials in aqueous electrolytes," *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.*, vol. 65, no. 4, pp. 1–6, 2002.

[126] J. D. Ramshaw, "Functional derivative relations for a finite non-uniform molecular fluid in the canonical ensemble," *Mol. Phys.*, vol. 41, no. 1, pp. 219–227, 1980.

[127] B. Roux, "The calculation of the potential of mean force using computer simulations," *Comput. Phys. Commun.*, vol. 91, pp. 275–282, 1995.

[128] J. Hu, R. Lipowsky, and T. R. Weikl, "Binding constants of membrane-anchored receptors and ligands depend strongly on the nanoscale roughness of membranes," *Proceedings of the National Academy of Sciences*, vol. 110, no. 38, pp. 15283–15288, 2013.

[129] H. Ohtaki and N. Fukushima, "A structural study of saturated aqueous solutions of some alkali halides by X-ray diffraction," *J. Solution Chem.*, vol. 21, no. 1, pp. 23–38, 1992.

[130] I. Kalcher and J. Dzubiella, "Structure-thermodynamics relation of electrolyte solutions," *The Journal of Chemical Physics*, vol. 130, no. 13, p. 134507, 2009.

[131] V.-T. Pham and J. L. Fulton, "Ion-pairing in aqueous CaCl2 and RbBr solutions: simultaneous structural refinement of XAFS and XRD data.," *The Journal of chemical physics*, vol. 138, no. 4, p. 044201, 2013.

[132] P. George, R. J. Witonsky, M. Trachtman, C. Wu, W. Dorwart, L. Richman, W. Richman, F. Shurayh, and B. Lentz, ""Squiggle-H2O". An enquiry into the importance of solvation effects in phosphate ester and anhydride reactions," *BBA - Bioenergetics*, vol. 223, no. 1, pp. 1–15, 1970.

[133] K. S. Pitzer, "Activity Coefficients in Electrolyte Solutions," pp. 75–153, CRC Press, 2 ed., 1991.

[134] R. W. Potter II and M. A. Clynne, "Solubility of highly soluble salts in aqueous media - Part 1, NaCl, KCl, CaCl2, Na2SO4, and K2SO4 solubilities to 100C," *Journal of Research of the U.S. Geological Survey*, vol. 6, no. 6, pp. 701–705, 1978.

[135] J. Barthel, "O. Söhnel und P. Novotny: Densities of Aqueous Solutions of Inorganic Substances, Vol. 22 aus der Reihe: Physical Science Data, Elsevier, Amsterdam, and New York 1985. 335 Seiten, Preis: Dfl. 175,-," *Berichte der Bunsengesellschaft für physikalische Chemie*, vol. 89, pp. 722–723, jun 1985.

[136] A. E. Reed, L. a. Curtiss, and F. Weinhold, "Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint.," *Chem. Rev. (Washington, DC, United States)*, vol. 88, no. 6, pp. 899–926, 1988.

[137] C. P. Petersen and M. S. Gordon, "Solvation of Sodium Chloride: An Effective Fragment Study of NaCl(H2O)n," *The Journal of Physical Chemistry A*, vol. 103, no. 21, pp. 4162–4166, 1999.

[138] M. Soniat, *Incorporation of Charge Transfer into Classical Molecular Dynamics Force Fields with Applications in Physical Chemistry.* PhD thesis, University of New Orleans, 2014.

[139] M. Soniat and S. W. Rick, "The effects of charge transfer on the aqueous solvation of ions," *The Journal of Chemical Physics*, vol. 137, no. 4, p. 44511, 2012.

[140] E. Mrázková and P. Hobza, "Hydration of sulfo and methyl groups in dimethyl sulfoxide is accompanied by the formation of red-shifted hydrogen bonds and improper blue-shifted hydrogen bonds: An ab initio quantum chemical study," *J. Phys. Chem. A*, vol. 107, no. 7, pp. 1032–1039, 2003.

[141] P. Hobza and Z. Havlas, "Improper, blue-shifting hydrogen bond," *Theoretical Chemistry Accounts*, vol. 108, no. 6, pp. 325–334, 2002.

[142] B. J. Van der Veken, W. a. Herrebout, R. Szostak, D. N. Shchepkin, Z. Havlas, and P. Hobza, "The nature of improper, blue-shifting hydrogen bonding verified experimentally," *J. Am. Chem. Soc.*, vol. 123, no. 49, pp. 12290–12293, 2001.

[143] E. D. Glendening, J. K. Badenhoop, A. E. Reed, J. E. Carpenter, J. A. Bohmann, C. M. Morales, C. R. Landis, and F. Weinhold, "NBO 6.0.," 2013.

[144] S. Sukenik, Y. Boyarski, and D. Harries, "Effect of salt on the formation of salt-bridges in beta-hairpin peptides," *Chemical Communications*, vol. 50, no. 60, pp. 8193–8196, 2014.

[145] M. Mammen, S.-K. Choi, and G. M. Whitesides, "Polyvalent interactions in biological systems: implications for design and use of multivalent ligands and inhibitors.," *Angewandte Chemie, International Edition*, vol. 37, pp. 2755–2794, 1998.

[146] Z. S. Hendsch and B. Tidor, "Do salt bridges stablize proteins - a continuum electrostatic analysis," *Prot. Sci.*, vol. 3, no. 2, pp. 211–226, 1994.

[147] V. Satarifard, S. Kashefolgheta, A. Vila Verde, and A. Grafmüller, "Is the solution activity derivative sufficient to parametrize ion-ion interactions? ions for tip5p water," *Journal of Chemical Theory and Computation*, vol. 13, no. 5, pp. 2112–2122, 2017. PMID: 28394606.

[148] S. Kashefolgheta and A. Vila Verde, "Developing force fields when experimental data is sparse: Amber/gaff-compatible parameters for inorganic and alkyl oxoanions," *Phys. Chem. Chem. Phys.*, vol. 19, pp. 20593–20607, 2017.