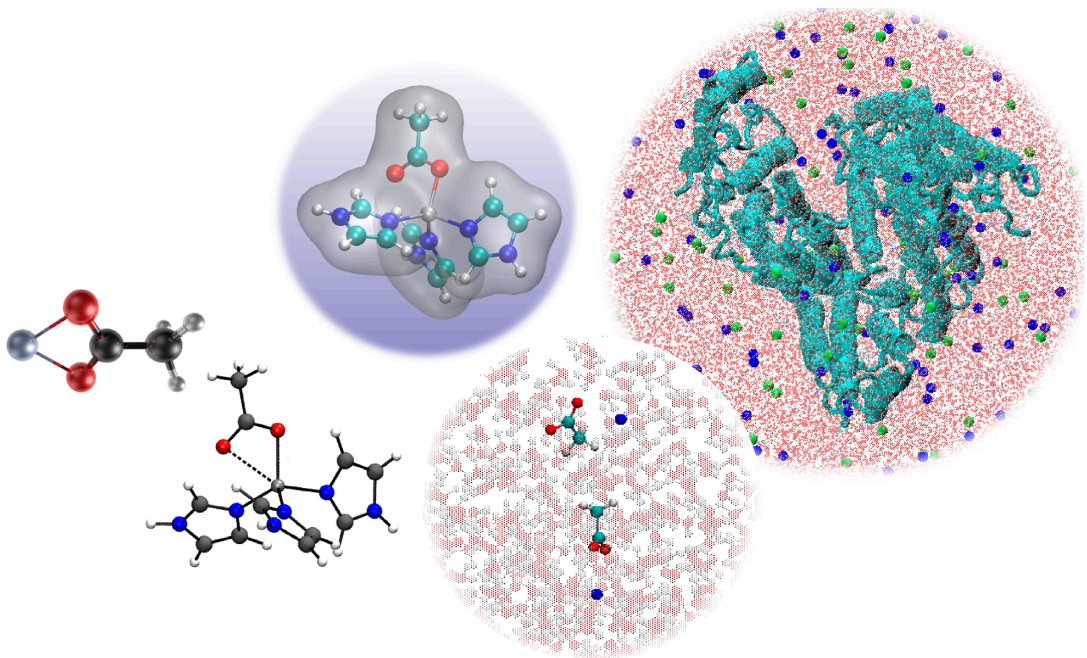
Linnaeus University Dissertations
Nr 312/2018

Emma Ahlstrand

Metal ions in life

towards accurate computer-aided studies of protein–ion interactions



LINNAEUS UNIVERSITY PRESS

Metal ions in life

towards accurate computer-aided studies of protein-ion interactions

Linnaeus University Dissertations

No 312/2018

METAL IONS IN LIFE

towards accurate computer-aided studies of protein-ion interactions

EMMA AHLSTRAND

LINNAEUS UNIVERSITY PRESS

Metal ions in life: towards accurate computer-aided studies of protein-ion interactions
Doctoral Dissertation, Department of Chemistry and Biomedical Sciences, Linnaeus University, Kalmar, 2018

ISBN: 978-91-88761-37-8 (print), 978-91-88761-38-5 (pdf)
Published by: Linnaeus University Press, 351 95 Växjö
Printed by: DanagårdLiTHO, 2018

Abstract

Ahlstrand, Emma (2018). *Metal ions in life: towards accurate computer-aided studies of protein-ion interactions,* Linnaeus University Dissertations No 312/2018, ISBN: 978-91-88761-37-8 (print), 978-91-88761-38-5 (pdf). Written in English.

The importance of ions in life sciences can not be overstated. The interaction between metal ions and proteins is vital because it is involved in a variety of biological processes. The ions contribute to stability and function of proteins. Moreover, they are relevant in disease progression.

Realistic computer simulations pave the way for drug development, through providing detailed insights into the dynamics of proteins and various biological processes that occur in the body. Such information can be impossible to achieve through experiments of living subjects in vivo or from test tube experiments in vitro alone. However, theoretical methods have to result in accurate predictions. In my thesis, I studied different ways to handle the ions in simulations. Since the systems contain thousands of atoms the calculations are demanding. Despite the availability of computer clusters, the com putational capacity is not sufficient. I have examined the simplified models used in simulations of larger systems (e.g., whole proteins) to pave the way for improvements of the simulation models.

Different ions have different effects on biochemical systems and it is important to be able to distinguish between them. Thus, from a biochemical point of view, it is central to be able to describe their unique characteristics. Their difference can be from vital to toxic to the body. Zinc is essential and present in more than 3000 proteins in our body and has a very flexible interaction with proteins. This property has proved to be hard to reproduce in computer simulations. Cadmium can replace zinc, but is toxic because it does not have the same catalytic ability. From a modelling perspective do these ions have similar characteristics as they have the same ionic charge. Inclusion of more realistic electron effects may be necessary to be able to simulate the difference.

With my studies, I have contributed towards a better understanding of the interactions between metal ions and proteins. I have pointed out a direction for further improvement of methods for simulations of large systems.

For the same purpose, I have also studied the frequently occurring ions sodium and potassium found as salts in all body fluids, but also lithium belonging to the same group in the periodic table and used in therapeutic purposes. The results show that potassium and sodium can be simulated by a commonly used computational approach, whereas more advanced methods are required to study lithium ions accurately.

Overall, the work within this thesis has explored ion-protein interactions and provided information about methods for energy calculations and models for molecular dynamics simulations for some of the most important ions within biochemistry.

Keywords: molecular dynamics, polarisable force fields, quantum chemistry, zinc, alkali ions

Populärvetenskaplig sammanfattning

Betydelsen av joner i biovetenskap kan inte överskattas. Interaktionen mellan metalljoner och proteiner är livsviktig eftersom den är essentiell i en rad olika biologiska processer. Joner bidrar till proteiners stabilitet och funktion. Därigenom har jonerna också inverkan i sjukdomsprocesser.

Realistiska datorsimuleringar underlättar vid utveckling av nya mediciner. Simuleringar kan ge detaljerad insikt i proteiners dynamik och olika biokemiska processer som sker i kroppen som inte är möjliga att erhålla enbart via experiment, varken i levande objekt (*in vivo*) eller på labbänken (*in vitro*). För detta krävs dock att de teoretiska metoderna resulterar i korrekta prediktioner.

I min avhandling studerar jag olika sätt att hantera jonerna på i simuleringar. Eftersom systemen består av tiotusentals atomer blir de beräkningar som krävs en utmaning. Trots stora kluster av sammankopplade datorer är beräkningskapaciteten inte tillräcklig. Jag har undersökt de förenklade modeller som används vid simuleringar av större system (till exempel hela proteiner) för att bana väg för förbättringar av dessa modeller.

Olika joner har olika effekter på de biokemiska systemen och det är angeläget att kunna särskilja mellan dem. Därför är det utifrån ett biokemiskt perspektiv centralt att kunna beskriva deras unika egenskaper, då skillnaden kan röra sig om att vara livsviktig och giftig för kroppen. Zink är ett livsnödvändigt ämne som förekommer i över 3000 proteiner i vår kropp och har en flexibel interaktion till proteiner. Denna unika egenskap har visat sig svår att återskapa i datorsimuleringar. Kadmium liknar zink i dess kemiska egenskaper och kan därför byta plats med zink. Dessvärre är kadmium giftigt för kroppen vilket bland annat beror på att proteinet mister sin katalytiska förmåga när kadmium ersätter zink. Utifrån ett modelleringsperspektiv har jonerna samma förutsättningar då de har samma laddning. Inräknande av elektronernas effekter kan utgöra skillnaden för om en simulering ska ge korrekt resultat. Jag har med mina studier bidragit till ökad förståelse av interaktionerna mellan joner och proteiner samt vilka metoder som är lämpliga. Jag har visat på i vilken riktning förbättringspotential för simulering av stora system finns.

I samma syfte har jag även studerat de vanligt förekommande jonerna natrium och kalium som finns som salter i all kroppsvätska. Jag har också undersökt litium som tillhör samma grupp i periodiska systemet och som används i terapeutiskt syfte. Här visar det sig att för natrium och kalium kan klassiska modeller användas, medan för litium-simuleringar krävs en mer avancerad beskrivning.

Sammanfattningsvis har arbetet inom denna avhandling utforskat jon-protein-interaktioner samt tillhandahållit information om metoder för energiberäkningar och modeller för molekyldynamiksimuleringar för några av de viktigaste jonerna inom biokemi.

To life, science and my family

"Science is a wonderful thing
if one does not have to earn one's living at it."
Albert Einstein

List of publications

(I) *E. Ahlstrand*, D. Spångberg, K. Hermansson, R. Friedman, Interaction energies between metal ions (Zn^{2+} and Cd^{2+}) and biologically relevant ligands, Int. J. Quant. Chem. 113, 2554, (2013).

(II) O. Becconi, *E. Ahlstrand*, A. Salis, R. Friedman, Protein-ion interactions: simulations of bovine serum albumin in physiological solutions of NaCl, KCl and LiCl, Israel J. Chem. 57, 403, (2017)

(III) *E. Ahlstrand*, K. Hermansson, R. Friedman, Interaction energies in complexes of Zn and amino acids: a comparison of ab initio and force field based calculations, J. Phys. Chem. A, 121, 2643, (2017).

(IV) *E. Ahlstrand*, J. Zukerman Schpector, R. Friedman, Computer simulations of alkali-acetate solutions: accuracy of the force fields in difference concentrations, J. Chem. Phys., 147, 194102, (2017).

In addition, the following article was published outside the scope of this thesis.

E. Ahlstrand, A Buetti-Dinh, R. Friedman, An interactive computer lab of the galvanic cell for students in biochemistry, BAMBED: Biochem. Molecular Biol. Edu., doi: 10.1002/bmb.21091 (2017).

Contents

CCSD(T)	Coupled-Cluster method, for Single- and Double-excitation with perturbative treatment of Triples
DFT	Density Functional Theory
EDA	Energy Decomposition Analysis
FF	Force Field
HF	Hartree-Fock
LJ	Lennard-Jones
MD	Molecular Dynamics
MM	Molecular Mechanics
MP2	Møller-Plesset second order perturbation
QM	Quantum Mechanics
PME	Particle-Mesh Ewald summation
RDF	Radial Distribution Function
vdW	van der Waals

Chapter 1

Metal ions in life

Ions are indispensable for life and many of the essential nutrients are in fact metal ions. Metal ions exist in different forms in biological systems. The ions can be hydrated, bound to membranes or proteins, as well as in association with anions, e.g., chlorides, phosphates and carboxylates, in the diverse heterogeneous buffer of cell fluids [1]. The far most abundant metal ions in the human body are alkali and alkaline earth metal ions, including sodium, potassium, calcium, and magnesium [1]. These ions are commonly involved in electrolyte balances, signalling processes, as enzyme regulators and as components in bones and teeth [2–4]. Sodium is the major component of the cations of extracellular fluid, while potassium is the principal cation of intracellular fluid [4].

Salts of monovalent ions are fundamental to life as they regulate homeostasis, electric potentials of cells and stabilise biomolecular structures. Protein surfaces have the ability to attract small charged molecules, to hold them near the protein surface for relatively long durations and

to shuttle them between surface residues [5] enabling reactions such as binding of a ligand to its receptor or insertion of an ion into an ion channel. Thus, the interactions depend on the specific composition of the electrolyte (cation–anion pairs) and on its interactions with the solvent molecules and with the macromolecular surfaces [3]. Specific interactions that depend on the nature of electrolytes rather than colligative properties of aqueous electrolyte solutions, are one subject that will be considered in the context of this thesis.

Multivalent metals are considered to play critical roles in biology as they are involved in cellular and subcellular functions. Among these enzymatic elements are metal ions such as magnesium, zinc, iron, copper and manganese [1]. These ions are referred to as trace elements or micronutrients although it is important to note that iron and zinc are often found in substantial amounts in the human body (2-4 g [6]). Zinc is one of the most important ions in human and it is found in all body tissues [7]. The human zinc proteome covers about 3000 proteins [8] and is therefore critical in many biological processes [7]. Three major functions of zinc are catalysis, protein regulation and structural maintenance where the ion ensures correct protein folding and prevents unfolding. The function of the ion is mainly dictated by the combinations of ligands in the complex formed with surrounding amino acid residues in the binding site, the so called coordination sphere [9, 10]. The characteristics of this Group XII metal enables interactions with different ligands. The most common interacting atoms are imidazole nitrogens from Histidines (His), carboxylate oxygens from the side chains of Glutamate (Glu) or Aspartate (Asp), and sulfurs from sulfhydryl groups of Cysteines (Cys), as well as oxygens of structural water molecules [11, 12].

In active sites, zinc has a flexible and dynamic coordination environments, facilitating its function as a catalyst where the ion participates directly in bond-making and/or bond-breaking [9]. Catalytic binding sites are typically four-coordinated, but may as well be five-coordinated depending on the enzyme mechanism [9], see Figure 1.1A. Structures containing one Glu or Asp are able to undergo a carboxylate shift mechanism where the binding mode changes between mono- and bidentate. This mechanism is involved in reactions catalysed by farnesyl transferase and some alcohol dehydrogenases for example [13, 14]. A characteristic catalytic four-coordinated zinc site contains three protein-supplied ligands, usually His or Cys side chains, and a water molecule [9, 11, 13, 15], see Figure 1.1B.

In structural sites zinc ions help to stabilise the conformation. For a small protein, such as metallothionein, the ion can be important for the structure of the entire protein [14]. The so called zinc finger is a common binding motif where zinc ions stabilise the fold of the protein, which consist of four ligands (His and/or Cys) [25]. These sites are often inaccessible to solvent and the metal ions have strong charge-charge interactions in the protein environment [10]. These stable structural sites are not discussed further in this thesis as the scope is to study the dynamics of the ions.

Zinc can be involved in metal-induced protein folding [10], as well as in intermolecular bridging. Such bridging binding sites located between proteins, with ligands provided from peptide chains from different protein monomers, have stabilising functions of dimers and oligomers. One example is the involvement of zinc ions in the formation of pathogenic amyloid aggregation [17, 26]. Small-sized, soluble amyloid-β (Aβ) oligomers formed during the early steps of peptide aggregation are the main cytotoxic agents in Alzheimer's disease [24, 27–29]. Figure 1.1C is

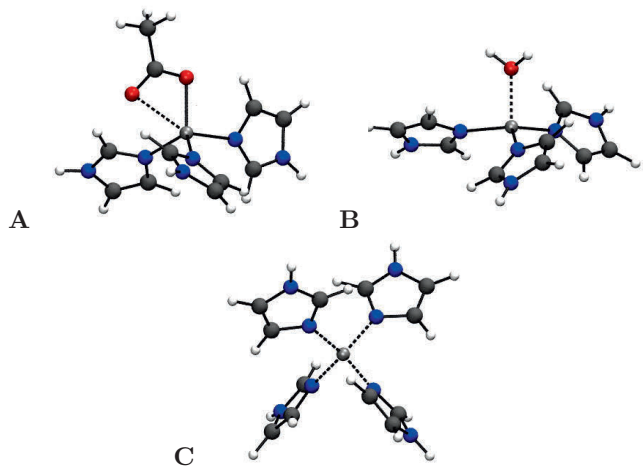


FIGURE 1.1: Examples of zinc binding sites. Neutral His is represented by imidazole (Imi) and the deprotonated Asp or Glu side chains by the acetate anion (Ace). Dashed lines show bonds with distances > 2 Å. **A)** Imi_3Ace. This combination is common in enzymes (e.g., in superoxide dismutase (SOD) [16], in the amyloid-β (Aβ) peptide [17] and in S100A12 [18]). **B)** Imi_3water is a very common active catalytic site found in carbonic anhydrase [14, 19–21]. **C)** Imi_4 which has been suggested to be involved in pathogenic zinc-catalysed amyloid aggregation where two His binding residues come from two adjacent Aβ peptides respectively [22–24]. (These are sites that are further discussed in my work, and does not serve to show the overall most frequently occurring binding sites).

an example of a zinc binding site that can act as intermolecular bridge between peptides in dimers or oligomers.

An interesting comment to this is that while zinc affects the associations of these proteins, one would also expect that the association of these proteins affects the availability of zinc ion. Ion concentrations needs to be strictly controlled to fulfil all the biological functions and to avoid unwanted side effects [14]. The flexible zinc-binding sites are

vulnerable to replacement by cations such as nickel, cadmium, mercury and lead, which may constitute one of the possible pathways for heavy-metal intoxication in living organisms [10]. The carboxylate side chains of Asp/Glu play an important role in shaping and fine-tuning the geometry of the binding site so that the protein can selectively sequester a particular metal co-factor from the cell fluids [10].

Chapter 2

Interactions and interaction energies

2.1 Intra and intermolecular interactions

Covalent *intramolecular* interactions that exist within molecules are strong, with enthalpies on the order of 100 kcal/mol [30]. The *intermolecular* forces that exist between molecules are significantly weaker than the bonding forces. Even though one individual interaction is modest, cumulatively the energies of intermolecular interactions are significant and influence the physical properties of matter, such as the boiling point, diffusion rate and viscosity.

The interaction energy ΔE^{int} is the difference between the energy of a complex AB (E_{AB}) and the energies of the fragments A and B (E_A and E_B). In a practical example used in my work one such equation

can look like this:

$$\Delta E^{int} = E_{complex} - (E_{Zn^{2+}} + E_{ligand}) \tag{2.1}$$

where $E_{complex}$ is the total energy (sum of intra- and intermolecular energy) of the complex, $E_{Zn^{2+}}$ is the energy of the Zn ion and E_{ligand} is the energy of the ligand. If several ligands are present the interaction energy can instead be calculated as:

$$\Delta E^{int} = E_{complex} - (E_{Zn^{2+}} + \sum_{all\ ligands} E_{ligand}) \tag{2.2}$$

Atoms in a molecule are held together by interactions of different origins. There is a sliding scale from electrons being shared between two atoms or exclusively held by one atom or the other. However, simplified rules are implemented to classify different types of bonds, i.e., covalent, dative, ionic and metallic bonds.

The interaction between two charged ions (or molecules) is referred to as an electrostatic interaction that can be described by the Coulomb law:

$$V_{el} = (z_1 z_2)/(4\pi\varepsilon_0 R) \tag{2.3}$$

where $z_1 z_2$ are the charges of the ions, R the inter-ionic distance and $1/(4\pi\varepsilon_0)$ is the Coulomb constant. These ionic electrostatic interactions are relatively strong and are able to act over long ranges. Metal ions are spherical and the Coulombic interactions they produce are isotropic.

Molecules without a net charge can still be involved in electrostatic interactions. The strength of these interactions depends on the distribution and properties of the electron clouds surrounding the nuclei. Electronegativity is the tendency of any atom to pull electrons towards

itself, and away from other atoms. Atoms with higher electronegativities can carry partial negative charges ($\delta-$) while the atoms with lower electronegativities carry partial positive charges ($\delta+$). Electrons are thus not shared equally which leads to that molecules (or parts thereof) can become dipoles. For molecules with more than two atoms, the three dimensional structure determines if the molecule becomes polar or not. The entire molecule will be non-polar if the dipole moments cancel out each other due to symmetry. Nonpolar molecules can however get induced dipole moment from an applied external field and thus become polar. This distortion of the electron distribution is called electronic polarisation, while the distortion of the positions of the nuclei, i.e. bending and stretching of molecules, is called distortion polarisation. The induced dipole moment is proportional to the strength of the applied field via the polarisability. Orientational polarisation is when permanent dipoles rotate according to the applied field, i.e, alignment of polar molecules.

The interaction between two charged particles in a dielectric medium can be described by the shielded Coulomb potential:

$$V_{el} = (z_1 z_2)/(4\pi\varepsilon_0\varepsilon_r R) \tag{2.4}$$

where ε_r is the dielectric constant of the medium or solvent. The dielectric constant is the relative static permittivity that is a temperature dependent material property. For a solvent it is a relative measure of its chemical polarity. Water that is very polar has ε_r of around 80 at room temperature (78.54 at 25°C). Water is very efficient at shielding charges, hence the electrostatic interactions between ions (or charged amino acids) are significantly decreased in aqueous solution. The interaction energy associated with the process of solvation of a solute by water molecules is called hydration energy.

Interactions between ions and dipoles are called ion-dipole interactions (or charge-dipole interactions). Dipoles also interact with other dipoles (dipole-dipole interactions), see Figure 2.1. The orientational forces acting between two permanent molecular dipoles are referred to as Keesom forces. Forces acting between a permanent dipole and an induced dipole, are called Debye or induction forces. The London dispersion force [31] or induced dipole-induced dipole interaction is a temporary attractive force that results from asymmetrical distribution of electrons about the nucleus in two adjacent atoms. A common name for these interactions is attractive van der Waals (vdW) interactions. Thus, there is not a clear differentiation between electrostatic interactions and vdW interactions, since both types involve interactions between charged molecules or molecules with a residual dipole moment.

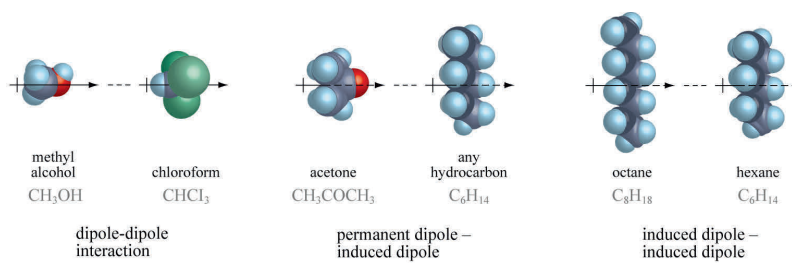


FIGURE 2.1: Electrostatic interactions between neutral molecules. The charge distribution are here described schematically with an arrow pointing in the direction of the molecule that carry a partial negative charge ($\delta-$). Figure from www.rusnano.com/wiki/article619.

A mathematical model that approximates the vdW interactions is the the Lennard Jones (LJ) potential (V_{LJ}), first proposed in 1924 [32, 33]:

$$V_{LJ} = 4\mathcal{E}_{ij}[(\sigma_{ij}/r)^{12} - (\sigma_{ij}/r)^{6}] \tag{2.5}$$

where $\mathcal{E}$ is the depth of the potential well, which correlates to the interaction strength, σ is the finite distance at which the inter-particle potential is zero, r is the distance between two particles i and j, see Figure 2.2. The parameter σ_i is also referred to as the vdW radii for particle i. The r^{-12} term represents the repulsive term that describes the short range repulsion (Pauli repulsion) due to overlapping electron orbitals. The r^{-6} term describes the long-range attractive interactions including the Debye induction, Keesom orientational and London dispersion interactions [3].

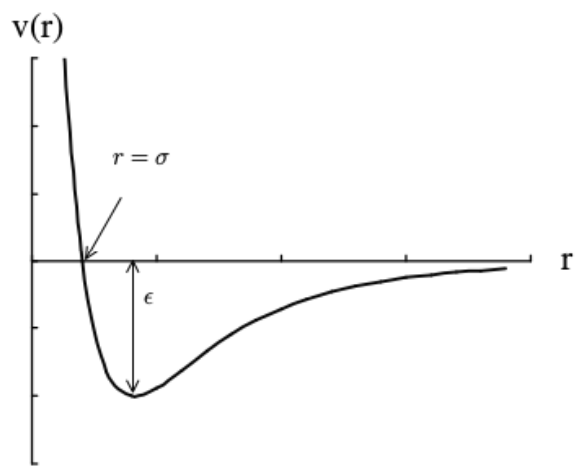


FIGURE 2.2: Figure of LJ 12-6 potential. The pair potential v(r) = 0 at r =σ, repulsive for r < σ and attractive for r > σ with a minimum of $v(r_0) = -\mathcal{E}$. at r_0. $r_0 = 2^{1/6}\ \sigma \approx 1.12\sigma$. For large distances the potential v(r) approaches zero.

Ions that have high electronegativity retain their valence electrons very strongly, have low polarisability and are called hard ions. Hard ions are generally of small size. Soft ions, e.g., Cu^{+}, Ag^{+}, Au^{+}, are soft due to the presence of several internal electrons that shield the nuclear attractive force from the outer valence shell electrons [3]. The different softness of the electron cloud is quantified by the polarisability. Zn^{2+}, Cu^{2+}, Ni^{2+}, Fe^{2+}, Co^{2+} are borderline ions, i.e. they are in-between soft and hard. The softness plays a role in determining which ligands the different ions prefer. The general rule of hard and soft acids and bases (HSAB) states that stable complexes result from interactions between hard acids and hard bases. Hard bases are oxygen interacting ligands, i.e. H_2O, hydroxyls and carboxylates. Nitrogen interacting ligands are in between hard and soft acids while ligands with sulfur (thiolates, Cys) as the interacting atom are soft bases. The diversity of the electronic states available for transition metal ions result in large coordination numbers and relatively labile chemical bonding for such complexes [34].

So called salt bridges, formed between paired anionic (e.g., Asp or Glu) and cationic (e.g., Lys or Arg) amino acid side chains are favourable electrostatic interactions that occur reasonably frequently in proteins. The charged groups are often around 3.0 to 5.0 Å from each other. A salt bridge is typically a combination of two noncovalent interactions: electrostatic interactions and hydrogen bonding.

A hydrogen bond is a special kind of interaction that involves a hydrogen atom covalently bound to a highly electronegative atom, such as an oxygen or nitrogen. The energies of hydrogen bonds vary depending on the donor-acceptor distance and angle and the temperature and are typically in the order of 1-7 kcal/mol [30]. Hydrogen bonds can occur

both inter- and intramolecularly. Hydrogen bonds are highly influential for the structure of macromolecules such as proteins.

2.2 Thermodynamic considerations for ionic interactions

Water is a polar molecule and its fluid is built up by a network of hydrogen bonds. This gives rise to special features of the fluid for example boiling temperature, melting temperature, volume and density of the solid compared to the liquid. When a solute is placed in liquid water, the water molecules arrange in a highly ordered cage-like structure around the solute, which is accompanied with a decrease in their entropy. According to Boltzmann the entropy can be expressed as a function of the number of microstates that can be populated and expressed as:

$$S = k \ln \Omega \tag{2.6}$$

where k is Boltzmanns constant and Ω is the number of microstates, i.e. possible arrangements.

The second law of thermodynamics states that the total entropy of a system and its surroundings always increases for a spontaneous process. The free energy change (dG) of a process is a result of the enthalpy and entropy changes for the system:

$$\mathrm{d}G = \mathrm{d}H - T\mathrm{d}S \tag{2.7}$$

where H is the enthalpy, T is the temperature and S is the entropy. The free energy change (dG) of complexation can be used to predict the binding energy. Complex formation is hence a balance between

entropical and enthalpical changes involved; a more negative dG can result from making dH more negative or from making dS more positive.

The chemical potential of a pure substance is defined as the free energy per mole substance. The chemical potential of a solute is related to the activity (a) of that solute. In an ideal solution the activity can be replaced by the molality. In ionic solutions there are significant electrostatic interactions and consequently, the behaviour of an electrolyte solution deviates considerably from an ideal solution. Non-ideality arises because ions of opposite charge attract each other, while ions of the same charge repel each other. On average, each ion is surrounded more closely by ions of opposite charge than by ions of like charge. The activity coefficient, γ, describes the deviation from ideality in concentration for a mixture:

$$a = \gamma \cdot C \tag{2.8}$$

where a is the activity that is proportional to concentration, C. The ion activity can be considered as the effective concentration.

Mean activity coefficients for monovalent ions in dilute solutions, on the order of 0.1 M or less, with complete dissociation can be predicted by the Debye-Hückel theory [35]. For moderately and highly concentrated salt solutions salt specificity emerges that arises from the nature of the ions [3]. As a consequence of specific ions effects different salts, depending on their concentrations, can be used to solubilise (salt-in) or precipitate (salt-out) proteins [36]. Thus, simple electrostatics cannot explain how anions or cations of the same nominal charge can behave differently. The electrostatic interaction involving ion-pairs is also balanced against the entropy-loss of their fixed positions. Ions induce an ordered local structure of the surrounding water molecules (in other words, they reduce the number of microstates that the solvating water

can adopt) which results in negative contribution to the entropy change for the ion solvation process.

Ion pairing describes the (partial) association of oppositely charged ions in electrolyte solutions [37]. The chemical equilibrium between free ions (metal ions, M^+ and anions, A^-) and ion-pairs (MA) is then stated in the association constant, K_a, as:

$$K_a = [MA]/([M^+][A^-]) \tag{2.9}$$

At equilibrium the reaction free energy change is equal to zero ($\Delta_r G = 0$) and:

$$\Delta_a G = -RTlnK_a. \tag{2.10}$$

where $\Delta_a G$ is the free energy for association.

Ion pairs exist as contact ion pairs (CIP) and solvent-shared ion pairs (SIP). Sometimes solvent-separated ion pairs (2SIP) are also discussed [38]. Species can be defined as ion pairs based on a specified cut-off distance between the ions. According to the Bjerrum approach [39] ions are considered as taking part in an ion pair when the distance between the ions is smaller than or equal to a certain length. The Bjerrum length is the distance at which the electrostatic interaction between two ions (2.4) is equivalent to the thermal energy k_BT (which at 25°C is 0.6 kcal/mol). The radial distribution function (RDF) can be a helpful tool to define this cut-off distance. The RDF is a plot of the average density distribution (ρ) of two-particle distances (r):

$$g(r) = \rho_r/\rho_{bulk} \tag{2.11}$$

Liquids are homogeneous systems with an uniform particle density on average (ρ_{bulk}). However, local density variations arises due to that

the particles form a shell structure around the central particle and the average density will show maxima and minima (ρ_r), see Figure 2.3. As the distance increases ρ_r approaches the bulk density and the RDF becomes uniform (g(r)=1). The integral of the first peak of the RDF is the average number of particles in the first coordination shell and is also called the coordination number. The first minimum of the RDF is a common criterion to define the first coordination shell. It is important to note that the probability to find a particle at a certain distance in a fluid does not depend only on the interactions between a pair of particles, but also on the interactions with all other particles in their surroundings.

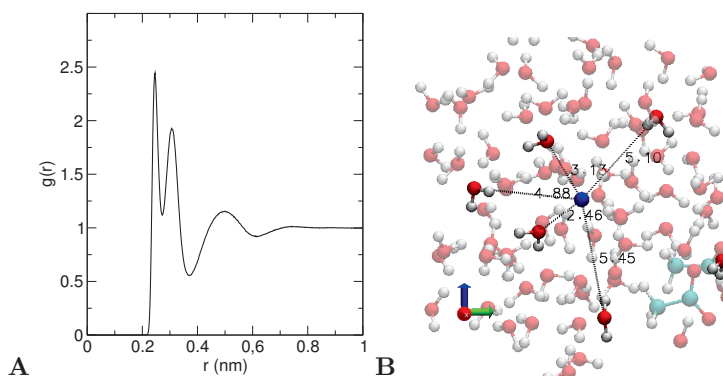


FIGURE 2.3: A) A RDF for water molecules around Na^+ in 1 M Na-acetate solution calculated from a molecular dynamics simulation. The first maximum appears at 2.4 Å which corresponds to the coordination shell of nearest neighbours around the Na^+. This is associated with the attractive minimum in the interaction. B) A snapshot from the simulation showing the interactions that corresponds to the peaks in the RDF.

Chapter 3

Computational methods

The methods within computational chemistry range from those that are fairly accurate, including all quantum effects, to more approximate, that consider atoms, or even groups of atoms, represented by a single interaction site. If applied correctly, these methods enable the researcher to investigate properties of the systems that are inaccessible through experiments [40]. The choice of method is often a compromise between accuracy and computational power and the description has to be suitable for the specific question to be answered.

The number of particles included in the system is a known critical factor for the computational cost. Classical molecular dynamics (MD) simulations, based on force fields, are generally performed for systems involving thousands of atoms or more. MD is a good way to get information on the atomistic-level events of proteins [41, 42], as many functions of proteins are governed by their dynamic character. Here it is important to consider that the simulations reach the time scales

relevant for the dynamics of the systems simulated. For the biologically relevant events such as protein-protein aggregation [43–46], protein folding [47], and drug binding [48, 49], the time scales are in the order from microseconds up to seconds. For long duration events an even more approximate solution than all-atom MD, a so-called coarse-grained force field (FF) is best suited to follow the process [45, 50–53]. In coarse-grained protein MD simulations several atoms (typically two to ten) are grouped together and represented by a single interaction site instead of treating each atom separately.

The simplistic nature of FFs unfortunately makes it impossible to directly study events involving changes to the electronic structures of molecules (e.g., bond formation/breaking or changes in protonation states). In such cases, other, more accurate methods are needed. In cases where the interesting chemistry and physics are localised to relatively few atoms of the system, the system can be divided into an electronically important region, e.g., the active site of an enzyme, and a surrounding region [54–56]. A quantum mechanical (QM) treatment is then applied to that part of the system that requires so, while classical molecular mechanical (MM) can be used for the surrounding. QM/MM is a very useful approach combining the strengths of QM methods and MM methods. The Nobel Prize in Chemistry 2013 was awarded jointly to Martin Karplus, Michael Levitt and Arieh Warshel who pioneered "the development of multiscale models for complex chemical systems" [57]. The main problem of QM/MM is the treatment of the boundary and the interaction between the subsystems. This is not straightforward considering that the transition from a MM region, described by classical potentials, to the QM region, with combined effects of nuclei and electrons, is not smooth. Another strategy to include electronic effects in simulations is to use a polarisable FF that includes some of the electronic interactions for the whole system.

In ab-initio QM-MD the forces are described based on both electrons and nuclei and the dynamics is captured in terms of classical dynamics and statistical mechanics [58]. Ab initio MD of ions in solutions for example have been reported (e.g., [59–63]). The inclusion of electrons increases the accuracy of the description of the system, but at the same time also increase the computational cost and is thus limited to rather small systems and short time scales. For biological systems, pure QM calculations are usually performed statically in vacuum or in a continuum solvent description.

3.1 Quantum mechanics

The optimal geometry of a molecule is found at the minimum of the potential energy surface. The energy can be described by the Schrödinger equation [64]. The Hamiltonian for a many-electron molecule contains a sum of the kinetic energies for the electrons and each nuclei respectively, the electrostatic attraction between electrons and nuclei, the electrostatic repulsion between electrons and the electrostatic repulsion between the nuclei. An approximate solution the time-independent Schrödinger equation can be obtained by use of the Born-Oppenheimer approximation [65]. The Born-Oppenheimer approximation is the assumption that the electronic and nuclear motions can be separated due to the large difference in the nuclear and electronic masses. Their motions are thus considered to occur on different time scales.

There are different methods to solve the time independent Shrödinger equation and to represent the electronic wavefunction (Ψ). The wavefunction is a mathematical function that can be used to calculate the electron distribution [66]. The major issue of all quantum chemistry methods is how to handle the interactions between electrons. The Pauli

exclusion principle requires that two electrons with the same spin cannot occupy the same region in space and the wavefunction must be consistent with that. The effect of indistinguishable particles on the electronic energy is known as the exchange energy. The electron-electron distances increase when the wave functions of indistinguishable particles overlap, thus the repulsion is lower than expected. In real systems, electron movements are also instantaneously correlated so that average electron-electron distances increase and the total repulsion decreases.

3.1.1 Hartree-Fock and post-HF

The common Hartree-Fock (HF) method uses a single Slater-determinant [67] wavefunction in which each row represents an electron coordinate and each column contains a one-electron basis function. The resulting molecular orbitals are spin orbitals derived from occupied spatial orbitals and spin functions. The antisymmetric Slater determinant fulfils the Pauli exclusion principle, but the HF theory fails to describe the electron correlation. The molecular orbitals of the HF method are optimised by evaluating the energy of an electron in each molecular orbital moving in the mean field of all other electrons, instead of the influence from individual electrons. The difference between the HF energy and the exact energy is a common definition of the correlation energy. The correlation energy is sometimes also divided into dynamical and non-dynamical, or static, correlation energy. The static correlation energy is important for molecules where the energy calculation needs to be based on more than a single-determinant. The common HF procedure uses the iterative self consistent field (SCF) method for finding the density matrix that gives the lowest energy.

Higher-order, more elaborate and accurate solutions to the Schrödinger equation can be built upon HF. There are different ways to determine

how the wavefunction should be expanded to include electronic correlation. Prominent examples of post-HF methods are configuration interaction (CI), multi-configurational self consistent field (MCSCF), Møller-Plesset second order perturbation (MP2) [68], and coupled cluster methods (e.g., CCSD(T) [69, 70]). In systems where several different electronic configurations are close in energy, e.g., transition metals with their open d-shells, the use of a multi-configuration method like MCSCF may improve the results [34]. This approach can also offer insights into intermediates and transition states in catalysis. Unfortunately, the cost of performing the calculations limits their use to rather small systems (10-20 atoms). In the full CI and complete active space SCF method (CASSCF, a MCSCF method) all possible excited determinants are taken into account, in addition to the ground state determinant, in the construction of the wavefunction. Thus, the MCSCF method takes account of the static correlation. Lower levels of these methods can also be performed, e.g., restricted active space SCF method (RASSCF) where the user has to choose which excited states to include. This makes it rather complicated to set up. In the MP2 theory and the coupled-cluster method, for single- and double-excitation with perturbative treatment of triples CCSD(T) dynamic electron correlation are introduced as correction terms (perturbations).

3.1.2 Density functional theory

In density functional theory (DFT) methods, the energy of the electronic system is represented by use of a functional of the electron density (ρ(r) [71]), without solving the wavefunction explicitly. DFT is less computationally demanding and therefore faster than post-HF methods [72]. The functionals can be constructed based on an exact solution, but are commonly constructed to fit experimental data. The

density can be expressed in terms of the occupied orbitals according to an approach developed by Kohn and Sham [73]. Kohn-Sham theory define the energy as a sum of (i) the non-interacting kinetic energy, (ii) the classical Coulomb energy of the electrons, consisting of their interaction with the nuclei, their self-Coulomb energy and (iii) an unknown exchange-correlation (xc) functional [74]. The total energy is minimised, with respect to the shape of the orbitals, self consistently by iterative methods starting from a tentative ρ(r). The exchange-correlation energy (E_{xc}) is the main source of error in DFT calculations and methods to improve this part have been developed in several stages.

The DFT methods include the interaction of the entire electron distribution and the self-interaction of electrons must be removed by the exchange-correlation functional. The local spin density approximation (LSDA) [67] was one of the earliest DFT methods. A local functional cannot completely remove the self-interaction. Therefore, the general gradient approximation (GGA) methods that are corrected to include spatial inhomogeneousity by including a term for the gradient of exchange and correlation of ρ(r) were developed. These were further developed into functionals called meta-GGA, which include some dependence on the kinetic energy density. Hybrid-GGA functionals (such as B3LYP [75–77] and B98 [78]) include a fraction of exact HF exchange energy in the functional. An energy density that depends on HF exchange can provide a better approximate functional by removing some of the self-interaction. There are also combined hybrid meta-GGA functionals (e.g., TPSSh [79] and M06 [80]). In the range-separated hybrid functionals a split depending on the inter-electronic distance is applied whereafter different amount of HF exchange correction for long range or short range self-interaction can be used (e.g., ωB97 [81]). A

long-range dispersion correction can also be applied to the GGA-type density functionals to improve the results (DFT+D3 [82]).

3.1.3 Basis sets

In quantum chemistry, a local basis set is a set of mathematical functions that is used to represent the electronic wavefunction for all electrons. The basis sets are either included in the calculation program packages or can be fetched from basis set archives such as for example the environmental molecular science laboratory (EMSL) basis set library [83, 84]. The use of a finite basis set in quantum chemical calculations is an approximation and the choice of an adequate basis set can have a huge impact on the quality of the results [85]. The chosen basis set has to be able to describe the features of the system that are important for the properties or behaviour of the problem of interest. Historically, the molecular orbitals were formed as a linear combination of atomic orbitals (LCAO). Later on, the orbitals were expressed as linear combinations of any set of mathematical functions that were convenient to manipulate and which in linear combination gave useful representations of molecular orbitals [66]. The basis sets are constructed by optimising beforehand a set of parameters for every shell and every element. Slater and Gaussian type functions are the simplest, and Slater-type orbitals (STO) were the first kind of basis functions used in quantum chemistry [66]. Contracted Gaussian type orbitals (GTO) are linear combination of Gaussians with different exponents [66].

The number of basis functions (contractions) used for representing a single Slater atomic orbital are denoted zeta (Z). The letter V denotes split valence basis sets, with only one contraction for inner orbitals,

and two or more contractions for valence orbitals, e.g., DZV or TZV. Addition of extra functions to the valence shell improves the bond description in different environments increasing the flexibility of the basis in molecules and improving the possibility to describe the intermolecular interactions. Polarisation functions are supplementing functions with angular momentum higher than the highest occupied atomic orbital in the isolated atom. For systems that have significant electron density at relatively large distances from the nucleus supplementary diffuse functions can be used. Diffuse functions are for example needed for lone-pair electrons, that on the average are at larger distances from the nuclei than core or bonding electrons [66].

In calculation of interaction energies of complexes an error called basis set superposition error (BSSE) arises from the fact that the basis set of the complex can describe the monomer charge density better than each monomer basis set. To remedy the BSSE a counterpoise correction procedure [86] can be applied, in which all energies are calculated in the same basis set. Such a correction does not necessarily improve the accuracy to all problems of interest though [74].

Calculations involving heavy elements can be performed using effective core potentials (ECPs) (also called pseudopotentials) where the chemically less relevant core electrons are replaced by a core potential function to reduce the required computational cost. Pseudopotentials combine the nuclear charge and the shielding effects of the core electrons into a single potential, eliminating the need of an explicit all-particle approach. Moreover, relativistic effects, that are important for heavy atoms, can be approximated by the ECP.

3.1.4 Atomic charges

QM methods can be used to estimate the partial charge distribution (magnitude and location) within a molecule. Atomic charges are useful for qualitative understanding of molecules and their interactions, e.g., for determination of dipole moment, and for the development of atom-types in FFs, that will be described in Section 3.2.1. The atomic charge is not a physical property of the atom but rather a representation of the electronic density in the molecule. The positions of the electrons around the nuclei are not fixed and they do not belong to a certain nucleus. There are several different calculation schemes for assigning the fraction of the electron distribution to each atom, e.g., Mulliken population analysis, [87], natural bond orbital (NBO) analysis [88, 89], the atoms in molecules (AIM) method [90], molecular electrostatic potential charge fitting via the Merz-Kollman-Singh (MKS) scheme [91] and the charges from electrostatic potentials using a grid-based (CHELPG) method [92].

Mulliken's method is widely used as it is simple; the electrons are partitioned to the atoms based on the linear combination of the atomic orbitals. However, its drawbacks include strong dependence on the basis set and its reliance on equal division of overlapping occupied orbitals even between atoms with differences in electronegativities. Natural bond analysis classifies and localises electrons to orbitals involved in bonding and anti-bonding (NBOs). With this method atomic partial charges converge to a stable values as the basis set size is increased. NBO tend to predict larger charges than several other population analysis methods. Both Mulliken and NBO charges are best used for comparing differences rather than determining absolute atomic charges. The AIM theory uses the gradient vector field of electron density to

partition the molecular electron density in basins of charge containing the nuclei.

In restrained electrostatic potential (RESP) [93] based methods, such as MKS and CHELPG, the population analysis is computed from the molecular electrostatic potential (MEP). The MEP is the force acting on a positive test charge at a given point of the molecule. The main difference between the MKS and CHELPG method can be found in the procedure used to fit the MEP. The CHELPG atomic charges are fitted to reproduce the MEP at a number of points around the molecule, while the MKS population analysis method calculates the MEP along several layers encompassing the molecule. The RESP approach has proved to be useful to derive atomic charge values for biomolecules [94].

3.1.5 Energy decomposition analysis

Energy decomposition analysis (EDA) can in a way provide a bridge between quantum chemical calculations and traditional bonding models. The interaction energy ΔE_{int} between two fragments in a molecule or between a metal ion and ligand(s) in a complex (Equation 2.1 and 2.2) can by use of different methods be divided into contributions from different origins. Interaction energies in reality are not separable, however in computational chemistry EDA can be performed. There exists several different schemes in the literature, starting from the pioneering work of Morokuma *et al.* [95]. Examples of schemes are symmetry adapted perturbation theory (SAPT) [96], natural energy decomposition analysis (NEDA) [97], fragment molecular orbital (FMO) method [98], restricted variational space (RVS) method [99] and localised molecular orbital EDA (LMOEDA) [100]. The LMOEDA scheme has been used in this thesis.

Within the LMOEDA, HF and DFT energies are divided into electrostatics, exchange, repulsion, polarisation (which includes charge transfer), and dispersion terms [100, 101], see Figure 3.1. Energy decomposition can as well be made in a solution (EDA-PCM) [102], in these cases an extra term named desolvation is included.

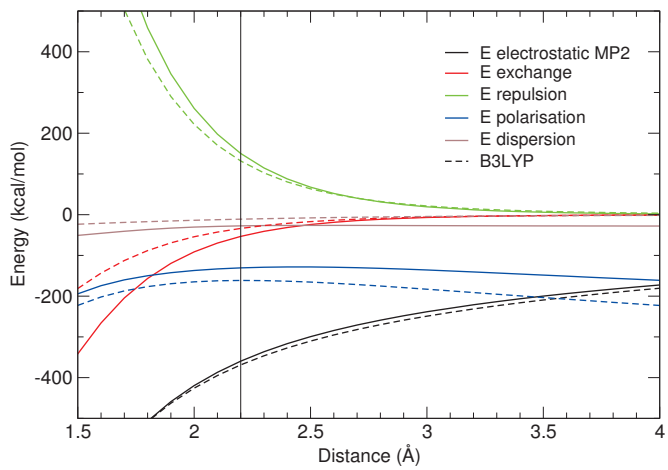


FIGURE 3.1: Diagram of attractive and repulsive energies in the interaction of Zn-methanthiolate (SCH_3) with the post-HF method MP2 and DFT with the B3LYP functional. The energies are divided into electrostatics, exchange, repulsion, polarisation (which includes charge transfer), and dispersion according to the LMOEDA-scheme [100]. The bond length where the total interaction energy has a minimum is at 2.2 Å.

3.1.6 Solution models

QM calculations of liquids and solutions can be done either by use of a hierarchical approach in which the most interesting part of the system is treated individually (explicitly) or the solvent can be represented as a implicit structureless polarisable bulk medium characterised

mainly by its dielectric constant ε. Several implicit approaches to introducing solvation into quantum chemistry have been developed, including the polarisable continuum model (PCM) [103], the conductor-like screening model (COSMO) [104] and solvation model based on density (SMD) [105]. A continuum representation of a protein environment or ionic solutions is a challenge, as discussed in Chapter 2.

3.2 Molecular mechanics and dynamics

Molecular mechanical (MM) methods neglect both electrons and the quantum aspects of the nuclear motion, and consider a rather simplified scheme of the interactions within a system. A molecular dynamics (MD) simulation generates a trajectory representing the time evolution of the particle positions and velocities. Coordinates and velocities are updated during each time step in the simulation based on calculations of a potential function derived from a FF.

3.2.1 Force fields

The force field (FF) is a set of parameters used to calculate the forces working on each particle in the system. Most FFs divide interactions into bonded and non-bonded parts. The bonded interactions refers to atoms that are involved in covalent bonds while non-bonded describe electrostatic and vdW forces. For the interactions between bonded atoms a spring-like interaction model is normally employed, in which the atoms are described as spheres of different masses, whereas the bonds are described as springs with a different stiffness to reproduce the vibrations including bond stretching, angle bending and bond torsion in molecules.

The non-bonded electrostatic interactions are based on the atomic point charges that interact according to the Coulomb law (Eq 2.3), while the van der Waals interactions are regularly described using LJ 6-12 potential (Eq 2.5). There are also other models such as the Buckingham potential [106] and the Born-Mayer [107] potential, however the LJ potential is faster to calculate [108]. The 12-6 LJ potential incorporates only two parameters (σ and $\mathcal{E}$) and is therefore relatively convenient to parameterise, which also explains its wide utilisation in molecular simulations [34]. The atomic parameters are included in the FF. *Combination rules* are used for interactions between two atoms of different types. The Lorentz (arithmetic) or the geometric combination rule can be used for σ, while the Berthelot (geometric) are used for $\mathcal{E}$. The Lorentz combination rule stems from the hard sphere assumption, in which particles are treated as rigid balls, while the geometric combining rule relies on a soft sphere assumption, which allows the atoms come a little closer to each other.

To reduce the number of calculations, a truncated function with a defined cut-off distance (r_c) above which the potential is set to zero (neglecting the interaction) is used in MD-simulations. Van der Waals forces can for example be truncated with a cut-off at distances of 1.0-2.0 nm [109, 110] depending on the system size and the computational effort. There are different truncation schemes available in the MD simulation softwares.

Three of the most common biomolecular FFs, based on the interactions divided into a bonded and non-bonded part, are chemistry at Harvard macromolecular mechanics (CHARMM) [111–114], the assisted model building and energy refinement AMBER [115–121] and the optimised potential for liquid simulations-all atom (OPLS-AA) [122–125].

The parameters and expressions for the forces are achieved and validated by use of experimental data and/or QM results. For CHARMM the parameterisation of the non-bonding terms was carried out using MEP-based charges. For the AMBER FF [121], partial atomic charges were determined using the restrained electrostatic potential (RESP) [93] methodology. In OPLS the non-bonded parameters were developed from Monte Carlo statistical mechanics simulations by computing thermodynamic and structural properties such as the enthalpy of vaporisation and density of liquids.

There are limitations with these classical FFs. The use of fixed partial atomic charges leaves the model molecule unable to be polarised by an external electric field or nearby molecules. Many body effects are due to the fact that every single atom influences and depends on the surrounding atoms. Explicit many-body polarisation effects have been shown to be a significant component of intermolecular forces [126]. However, in fixed charge based electrostatic models, the atomic partial charges are implicitly polarised for condensed phases in an averaged fashion. Biomolecular FF parameters are typically adjusted to properly describe the bulk properties of a high-dielectric medium, e.g., liquid water, and variations of the charge distribution with the dielectric environment are not considered in the classical FFs [127]. Neither is electronic screening for charged moieties explicitly treated in the classical FFs [128].

3.2.2 Polarisable force fields

Force fields that incorporate ability to account for electronic polarisability of the simulated system are undergoing rapid development. Ionisable residues in proteins play a major role in most biological processes including enzymatic reactions, proton pumps, and protein stability [129]. The interaction between ions and dipoles, the dipole-dipole

interaction and the London dispersion electronic reorganisation are examples of interactions that are important in biochemical systems, see Chapter 2.

Three commonly used methodologies for modelling electronic polarisation, i.e. where the potential functions explicitly treat electronic polarisability, are fluctuating charge methods [130, 131], classical Drude oscillators [132, 133], and induced point-dipoles [57, 134–136]. A schematic view of the different polarisable models are presented in Figure 3.2.

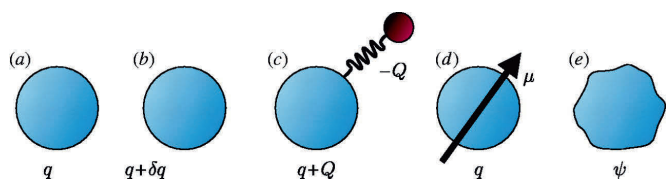


FIGURE 3.2: Qualitative description of approaches to modelling polarisation, a) fixed charge, b) fluctuating charge, c) Drude oscillator model, d) induced dipole and e) wave function (Ψ) description. Adapted from C. J. Illingworth, and C. Domene Proc. R. Soc. A 2009 [137].

In the fluctuating charge model, charges are allowed to vary by incorporation of an additional electrostatic energy term, representing the energy required to modify the charge of an individual atom. The total electrostatic energy is then minimised with respect to the individual atomic charges, with a set of constraints, for example keeping the overall charges on specific molecules, or of the system as a whole constant [137]. The polarisable CHARMM Drude FF [133] is based on the Drude oscillator method. In this model a massless charged particle (Drude particle) is attached to the nucleus via a short harmonic spring. The polarisability (α) is defined by the spring force constant, the charge

applied on the Drude particle, and the displacement of the massless particle creates the induced dipole. A damping scheme (Thole correction) is applied to the dipole-dipole interactions in order to avoid artificially high polarisabilities at short distances [138]. One advantage of the Drude model is that it preserves the simple particle-particle Coulomb electrostatic interaction employed in nonpolarisable FFs. The induced point-dipoles is a matrix based approach for modelling anisotropic polarisation by explicitly induced dipoles. The atomic multipole optimised energetics for biomolecular simulation (AMOEBA) [136], is based on the induced dipole approach.

The reason for the limited use of polarisable FFs is that these FFs are still under development and that inclusion of additional calculations is time-consuming. It is computationally expensive to achieve self-consistency of the dipoles at each MD time step [128].

3.2.3 Simulation set up

This section covers a description of the procedure and options available for the running of MD simulations. AMBER [139], CHARMM [111, 140], GROMACS [141–143] and NAMD [144, 145] are examples of software for MD-simulations. Regardless of the software used the basic steps of a MD simulation are: system preparation, energy minimisation, calibration of the system before simulation and finally the production run for data collection. Atomic coordinates for many biomolecular systems are provided in the Protein Data Bank (PDB) [146] that is a repository for experimentally defined macromolecular structures, captured form X-ray or NMR experiments. X-ray crystallography usually cannot be used to resolve the positions of hydrogen atoms. Hence, coordinates for hydrogen atoms of the structure are generated in the MD

software. Later, energy minimisation of the protein ensures their positions are reasonable. The structures are then commonly solvated with rigid water models such as three-site transferrable intermolecular potential (TIP3P) [147], simple point-charge (SPC) [148] or simple point charge extended (SPC/E) [149]. The next step of system preparation is to generate relevant salt concentration and to neutralise the charge of the solute, which is done by addition of ions in the solution.

The models are assigned initial velocities related to the temperature (or the average kinetic energy) of the system at the beginning of the simulation. To reproduce real world experiments the canonical ensemble in which the number of particles, volume and temperature are constant (NVT) or the isothermal-isobaric ensemble with constant number of particles, pressure and temperature (NPT) are usually employed. Different algorithms that maintain the constant temperature, volume or pressure are available. One of the earlier examples of temperature thermostat is the Berendsen thermostat algorithm that couples the simulated system to an external heath bath [150]. However, due to artefacts not reproducing a proper canonical ensemble more advanced thermostats such as the velocity-rescaling thermostat [151] where the velocities are lowered using a constant friction have been developed. In order to control the pressure in the simulation, the system can be coupled to an external bath with the Parrinello-Rahman algorithm [152]. The Nose-Hoover can be used as thermostat or barostat [153, 154].

During the MD-simulation different algorithms are also used to constrain the lengths of the bonds in molecules, e.g the linear constraint solver (LINCS) [155] or the SHAKE [156] algorithms; and to keep the water molecules rigid, e.g., the SETTLE algorithm [157]. These algorithms achieve computational efficiency as the number of explicit

calculations decreases and the time step between each calculation can be increased.

By far the most compute-intensive task of a MM-based MD simulation is the calculation of the non-bonded interactions. In MD simulations the main approach to this problem is the use of cut-offs, as described above. The electrostatic contribution extends over a larger distance than the vdW interactions, making these the most computationally demanding of the two. Long-range electrostatic interactions in MD simulations are most commonly treated with a particle-mesh Ewald (PME) summation method [158, 159].

Additionally, since the size of the simulated system is finite, boundary effects have to be accounted for in order to remove their influence on the calculated properties. This is usually done by applying periodic boundary conditions (PBC), which assume that each side of the simulation box is in contact with the opposite side. The use of PBC effectively copies the system in all three dimensions, and a particle that leaves the box through one side is replaced by another particle entering the box from the opposite side, thus keeping the total number of particles in the system constant.

Chapter 4

Modelling metal ions with force fields

Molecular dynamics (MD) can be used to study specific ion effects and ion-dependent phenomena around and in proteins [5, 41, 160–162], DNA [163, 164] and biological membranes [165]. However, these simulations are not without obstacles. Classical MD-simulations typically rely on pair potentials as described in Chapter 3, hence the atomic force field (FF) parameters that define the molecular interactions are critical. The influence of the ions and their parameters may well affect the results of the whole simulation [166]. To properly model ionic interactions the model needs to simultaneously reproduce several properties of the ions, including interactions with other ions, the solvent and other molecules or ligands over time.

There are today two main approaches for the modelling of ion-ligand interactions. One can either use a bonded or a non-bonded model. In

the *bonded model* the ion-ligand interactions are treated as covalent bonds. This is an effective way to retain the specific geometry of the metal-ligand coordination sphere. For some simulations, e.g., for the representation of structural metal binding sites that are very stable, this approach can be accurate. The bonded model prohibits the modification of interactions between the metal ion and its ligands, and can thus not be employed to simulate coordination number changes and ligand exchange processes that are involved in enzymatic processes. The *non-bonded model* treats the metal ion interaction with the surrounding environment without an explicit bonding term. In the non-bonded model, the coordination of the metal ion is flexible, which allows coordination number switching and ligand exchange to the metal ion. This model is commonly used both for ions in solution and for ions in proteins. Ions can be represented with a *dummy atom model*, in which charges are placed between the metal ions and surrounding residues for a more realistic description of metal-ion covalent bond formation. Non-bonded ion models are also used within polarisable FF. Inclusion of polarisation is important for example for reliable studies of the selectivity of biological ion channels, as the parameters have to describe the solvation of the ions both by water and by the protein environment [129].

4.1 Parameter derivation and validation strategies

FF parameterisation can be based on comparison to experimental target values such as thermodynamic and structural properties, or theoretical target values such as results from QM calculations. Derivation

based on QM results such as atomic charges and potential energy functions is an effective and rather straightforward method, however some validation and comparison against experimental data is valuable. The different parameterisation strategies are accompanied with benefits and drawbacks of which some will be discussed here.

A common strategy in parameter derivation of biological macromolecules is to develop parameters for small representative fragments. The parameters are then transferred and applied to the larger complexes or macromolecules where validation and optimisation are performed [167]. The advantages of starting with smaller models are that small systems can be treated by use of high level very accurate ab initio (QM) methods and that more experimental data are available for small systems. As the typical classical FF model is based on pair-functions, the simplest procedure for parameterisation of non-bonded terms is to fit the FF parameters against QM calculated interaction energies between pairs of ions and ligands. EDA from QM calculations can provide detailed information about the interaction that are helpful in the parameterisation process.

One drawback with interaction energies for small systems, of just the ion and one ligand, is that the ion-ligand binding is liable to be overestimated as a consequence of non-additivity, i.e., that the first ligand has a higher interaction energy than the second, and so on. For multivalent ions, single ligand calculations as references can also cause errors connected to charge transfer effects, where additional ligands would stabilise the charge on the ion [34]. QM calculations can give information on the charge transfer effects in complexes, as the partial charges on the atoms can be estimated. However, it is important to note that different methods may end up in dissimilar results as the atomic charge is not a defined entity. The pair potential can also be calculated for a

complex with more than one ligand, a multi-ligand complex. Further, recalculation upon substitution or removal of different ligands in such multi-ligand complexes yields valuable information for parameterisation [168–173]. To properly model biochemical metal ion interactions the model systems should ideally include also counterions such as phosphates [174] and chlorides [175].

Many-body effects are non-additive interactions that cannot be directly handled with pair potentials. Hence, parameters derived based on QM calculations performed in the gas phase may not yield proper results in the condensed phase. Optimisation and validation of the FF parameters to reproduce experimental values is then crucial [129, 167]. Experimental values are typically thermodynamic properties of condensed phase, such as free energies of aqueous solvation (hydration free energies) and coordination numbers [167]. One error-cancelling strategy, for incorporation of many-body effects into the pair potential is the use of an effective two-body potential. For this purpose the interaction energies can be calculated between an ion and different ligands within an implicit solvation model [176–178]. Polarisation effects when molecules are transferred from the gas phase to the condensed phase can be estimated by calculation of partial atomic charges for a molecule placed in an implicit solvation model [179]. In the parameterisation process of polarisable FF caution should be addressed to that the polarisability of a molecule or an ion is generally smaller in the condensed phase than in gas phase because of the presence of surrounding species [137]. An alternative to effective pair potential is addition of many-body terms [180] to the pair potential, however many-body potentials have the drawback that they are more complicated.

Validation of the developed potentials taking into account realistic interactions, e.g., with respect to experimental coordination number,

ion ligand distances and binding free energies is useful [181]. For ions in aqueous solution examples of targets for parameterisation are, experimental hydration free energy, ion-oxygen distances, coordination number of the first solvation shell, RDFs, diffusion coefficients, mean residence times, electric conductivity, activity and osmotic coefficients [182, 183]. Moreover, lattice constants and lattice energies can help in the size definition for the ion [34]. If the ion parameters are intended to be used also in combination with biomolecules it is important to refine the parameters so that they also reflect on the difference between the solvation of the cation in the protein and in water. It is possible in many cases to reproduce the solvation in water by adjusting the FF parameters, but it is important to validate the energies in proteins as well [129], which is more complicated due to lack of experimental data and the non-uniform environment in protein binding sites. Relative binding energies between different ions can be helpful in the parameterisation process [129].

As simulations in the area of biochemistry are performed in physiological salt solutions, a range of different conditions, dielectric constants, pH and salt concentrations have to be considered to conclude that a model is rather generally valid. For the ions and salts in solution, effective concentrations and cation–anion parameter combinations are important to consider. The simulation conditions such as water model, treatment of long-range electrostatic interactions, e.g., truncation and PME-set up, and truncation of vdW interactions are also important [175]. The consistency across the alkali ion series should also be considered if selectivity and discrimination between ions are important.

4.2 Bonded models

The bonded model is simple and has low risk for instability of the system. For use of the bonded model approach, a relatively conserved binding site has to be expected during the simulation. The bonded model can thus be applied for structural metal sites that are not of central interest, far from the catalytic site or the ligand binding site [184]. Parameterisation can be made upon comparison to experimental data or based on QM-calculations, by use of the Seminario approach [185]. Special metal site parameterisation tools fro combination with the AMBER FF have been developed [11, 186]. The bonded model can not be applied for flexible binding sites. Instead, a non-bonded treatment is needed where the metal coordination undergoes changes.

4.3 Non-bonded models

The simplest non-bonded models treat the metal ion as a sphere with a fixed integer charge and the interactions with other atoms are represented by Columbic and LJ terms (see Equations 2.3 and 2.5). The charge of the ion is normally set to the formal charge, i.e., an integer number according to their oxidation state. As the charge is fixed the parameters that can be adjusted are the σ and $\mathcal{E}$ in Equation 2.5. The LJ function is hence advantageous both in parameterisation and computation.

4.3.1 Alkali ions

In 1990 Åqvist parameterised alkali and alkaline-earth metal ions using the free energy perturbation (FEP) method to best reproduce experimental hydration free energies [182]. These values were not derived with the PBC model and PME, and inclusion of PME has been shown to affect the results [187]. Several other optimisations of LJ parameters for alkali metal has been performed, among other things to include PME effects, (e.g., [188–194]). The water models commonly used in simulations actually result in different bulk properties for liquid water [195]. The optimal ion parameters for each water model can thus vary substantially [175]. The Kirkwood-Buff (KB) theory has been used for parameterisation of alkali ion solutions [196–198]. The KB-theory [199] is a general statistical mechanical theory of solutions, in which solution structure are related to thermodynamic response functions [38]. For example has alkali ion-specific trends of ion pairing to carboxylates been related to experimental osmotic coefficients via the KB theory [183]. Parameters were hence optimised to reflect on the interactions between alkali ions and protein surface charges.

4.3.2 The zinc ion

Stote and Karplus parameterised the LJ potential for the Zn^{2+} ion with the TIP3P water model in 1995 [200]. LJ parameters for divalent metal cations including correction for PME, were developed based on the thermodynamic integration method [201]. These parameters were tested to reproduce hydration free energies. Pure non-bonded schemes have difficulties in maintaining the overall topology of metal coordination sphere in MD simulations. Although, zinc has a completely filled d shell and two electrons in a outermost s orbital, the zinc ion Zn^{2+} still

has an electron configuration that includes both p and d orbitals. This enables it to accommodate to various coordinations, but makes it more challenging to model Zn^{2+} than many other non-transition metal ions. Instead of four-ligand coordination as typical in proteins the Zn^{2+} ion shows a preference for higher coordination number values (e.g., 5 or 6) during simulations with a non-bonded model [202]. One reason for this can be that the LJ parameters for Zn^{2+} were parameterised in aqueous solution, where the Zn^{2+} cation is hexa-coordinated.

The fixed charge assumption may be troublesome, since charge transfer may well be a characteristic of some ligand complexes [180], especially for divalent ions. Zinc can adopt different coordination geometries at a relatively low free energy cost due to the availability of vacant metal orbitals that can accept charge from the ligands [10, 129]. As a consequence, the zinc ion tends to get close to any negatively charged amino acid residue [203]. Parameters with additional terms and correction functions have been developed to represent the electrostatic interactions between the Zn^{2+} ion and the ligands for zinc metalloenzymes [180, 204–207].

4.3.3 Dummy atom models

Among other models used for ions in proteins is the semi-bonded cationic dummy atom model [208]. In the cationic dummy atom approach the metal atom is substituted by a set of cationic dummy atoms placed around the metal nucleus, imposing a given coordination geometry to the complex, in agreement with the particular nature of the complex and of the metal atom. For example, for a standard Zn metalloenzyme, with a tetracoordinated tetrahedral coordination geometry [209, 210]. The charge of the metal atom (e.g., 2+ for Zn^{2+}) is distributed equally

through dummy atoms (four dummy atoms with a partial charge of +0.5 each for the example given), whereas vdW parameters and/or mass are attributed solely to central metal atom. The cationic dummy atom model requires an intricate parameterisation process due to the many empirical parameters i.e., the charge and vdW parameters on the metal ion and the charge and eventual vdW on each site as well.

4.3.4 Polarisable models

Treatment of ions and ionised groups is one of the main reasons for the need of a polarisable FF [129]. The electronic cloud around the metal ion can change and redistribute the charges in response to changes in the surrounding environment. Electrostatic effects, such as charge transfer, dispersion and many-body polarisation are critical to shaping ion-protein binding potential energy surfaces and thus determine the specificity and selectivity [211]. Ion channels are one example of a complex biological system where the ion chemical environment described by the mean-field approach might be inadequate [137].

Parameters for metal ions have been developed for the CHARMM Drude-2013 polarisable model [212–214] and for the AMOEBA FF [215–217]. The CHARMM Drude-2013 FF has been used in simulations to study differential impact of the monovalent cations Li^+, Na^+, K^+, and Rb^+ on DNA minor groove width [218]. The polarisable FF reproduce experimental solution X-ray scattering data more accurately than non-polarisable FF and thereby enables prediction of protein-DNA interactions. The CHARMM Drude polarisable FF have been used to evaluate structural and functional similarities between two sodium ion transporters [219]. The AMOEBA FF has been used to study zinc-containing matrix metalloproteinase complexes [220]. Matrix metalloproteinases have been widely investigated as potential drug targets

in a range of diseases. Here the authors find that the polarisation effect plays a determining role in Zn^2+ coordination geometry and that the polarising field of Zn^2+ exerts a strong influence on the relative affinities of the ligands.

Chapter 5

Objectives

The main objectives underlying this thesis were to enhance our knowledge about how to handle interactions of metal ions in simulations to serve an advancement of the field towards accurate simulations of protein-ion interactions. Computational studies of proteins using force fields parameters have been employed since several decades. However, the present generation of force fields may not be fully adequate to study proteins when metal ions are included. There are two aspects of cation interactions that are considered in this work: (I) the interactions of metal cofactors, primarily Zn^{2+}, with ligands in proteins, and (II) the interactions of electrolytic cations with negatively charged groups such as carboxylates in solutions. In this context, we aim to discuss cases where dynamical changes in metal-binding sites (e.g., shift of carboxyl atoms, or substitution of ligands) are enabled and improve our knowledge regarding simulations without neglecting the interactions with the solvent. For realistic simulations of any biomolecule, correctly accounting for macromolecules solvated in electrolyte solutions in the

physiological concentration interval are necessary. More specific aims were:

§ 1 To examine the appropriate methodology for QM calculations of gas-phase interaction energies for Zn and Cd-metalloprotein model complexes at equilibrium as well as non-equilibrium distances. Inclusion of EDA yielded physical insights into the interaction and the origin of the interaction energies and states where the different methods differed in their mathematical descriptions of the interactions.

§ 2 To extend the validity of the ab initio QM calculations, we computed interaction energies on systems with higher complexity, with four to six ligands, both in gas phase and in solution. The aim was to gather detailed information on the interaction between Zn and ligands by use of localised molecular orbital EDA and EDA in polarisable continuum model and to clarify the differences between gas phase and solution. Further objectives were to examined the suitability of current FF based MD methods to deal with metalloproteins and address the differences in available parameters specifically for the Zn^{2+} ion. We aim to show the relationships between different FFs for Zn protein binding sites, specifically sites located at the interface between the protein and the cellular fluid.

§ 3 Towards understanding of the biological role of alkali ions from simulations we performed a benchmark study of different parameters to examine and demonstrate in detail the interaction between acetate and alkali ions in solution. We used dissociation coefficients between ions and acetate as one of the means to test different ion parameters as they can be observed experimentally and calculated from simulations.

§ 4 Furthermore, the choice of protein force field, water model and in particular LJ parameters for the ions crucially affect any analysis of ion-binding, and somewhat also other features of protein simulations. To shed light on the difference between different alkali cations we present simulations of the protein bovine serum albumin (BSA) at physiological simulations of LiCl, NaCl, and KCl. We examine the influence on protein conformation and stability at physiological concentrations of the different salts.

Chapter 6

Research results

Four research articles were published within the scope of the doctoral studies. Here follows a presentation of the results from each of these studies. The presentation is formulated to link the studies, which means that the articles are not presented in number order. Also, the frame is to connect the findings, hence the results are not necessarily presented in the same order as in the original papers. Eventual limitations with the methods and further research suggestions are here mentioned in conjunction with each article summary.

Paper I. Interaction Energies Between Metal Ions (Zn^{2+} and Cd^{2+}) and Biologically Relevant Ligands

In **Paper I** we deliver a systematic comparison of the accuracy of several computational methods for calculations of interactions of the Zn^{2+} and Cd^{2+} ions in model systems representing amino acid side chains.

In order to consider the dynamics of the interactions, we computed the energies also at non-equilibrium distances. We conclude that MP2 can be regarded as a suitable high-level method for calculations Zn-interactions with amino acid ligands at equilibrium distances. At large ion-ligand separation (3.0 Å), however, the MP2 results diverged compared to CCSD(T), a fairly accurate post-HF method used as a reference. The RMSE was on average 4 kcal/mol and was particularly high when the ligand was an anion. **Paper I** points out some strengths and weaknesses of a couple of different density functional methods. In general, all functionals overestimate (in absolute values) the interaction energies compared to CCSD(T). However, the range-separated functional ωB97xD achieved results that agreed better to CCSD(T) than MP2 at large ion-ligand distances. EDA revealed that meta and hybrid DFT calculations overestimate the magnitude of the polarisation energy, making this contribution too favourable.

Additionally, in **Paper I** we compared the ability of the different methods to discriminate between Zn^{2+} and Cd^{2+}. Zn is one of the essential nutrients, while Cd is toxic [162]. The chemical differences in ligand binding were illustrated by QM calculations of interaction energies for the respective ions. Zn had more favourable interaction energies with ligands than Cd in a range of between 20 to 40 kcal/mol depending on the type of ligand according to the reference method CCSD(T). Our calculations show that the hybrid-GGA functional B98 had the best ability to discriminate between Zn and Cd interaction energies for different ligands. The overestimation of the interaction energies is cancelled out between the Zn and Cd, hence this functional is applicable for such a comparative purpose. The M06 functional, however, is not good in this respect because the calculated results of the Cd-ligand are systematically closer to CCSD(T), whereas the values for Zn are systematically more overbinding compared to CCSD(T), which makes the

energy difference too high.

Along with the benefit of the possibility to study small complexes with very accurate QM-methods, there are a number of limitations for FF development based on such interaction energy calculations. Firstly, a surrounding environment, solvent or protein, influence the pairwise interaction through many-body effects. Secondly, the metal-ligand distances and the charge distribution, and thus the interaction energy, is affected by the other ligands that are present in coordinated complex. Redistribution of the charges might well occur by inclusion of additional ligands. **Paper III** is a continuation of this study including systems with higher complexity.

Paper III. Interaction Energies in Complexes of Zn and Amino Acids: A Comparison of Ab Initio and Force Field Based Calculations

Many Zn binding sites are located at the interface between the protein and the cellular fluid, and consequently interactions with both proteins and the bulk solvent are involved [12, 169, 200]. In aqueous solutions, Zn ions adopt an octahedral coordination, while in proteins Zn can have different coordinations, with a tetrahedral conformation found most frequently [12, 15, 169]. In **Paper III** we aim to show the agreement or difference between the QM and FF calculated interaction energies in a series of related molecules. Hence, we report on QM calculated interaction energies for ten Zn complexes with coordinations that mimic protein binding sites, for example carbonic anhydrase, metallo matrix proteinases and in S100-family proteins. Among the ten complexes, we have models with coordination numbers four, five and six. We included Zn complexes with water and only one or two amino acid side chain

mimics. Such complexes shall yield information about the interactions for the ion in transition between solution and a protein binding site. As we found in **Paper I** that MP2 is an appropriate methodology for QM calculations of metal-ion complexes this method was used in **Paper III**. Based on EDA results, the ligands were found to have high impact on the relative strength of polarisation and electrostatic interactions. Interestingly, ligand-ligand interactions did not play a significant role in the binding of Zn. Rather, the interactions were almost exclusively dictated by the ion-ligand interaction.

The fact that FFs are not good at reproducing geometries and energies of Zn complexes is quite well established in the literature, though there is no optimal solution of how to handle the ion. In **Paper III** we examined the suitability of current MD methods. We reported on interaction energies in the ten Zn complexes using the different classical FFs (OPLS-AA, CHARMM, AMBER) and compared those to the ab initio QM results. The examination of Zn-parameters available for classical FF calculations revealed that it is not possible to obtain satisfactory results compared to MP2. The differences between the MM and MP2 interaction energies were 65 kcal/mol on average and the largest discrepancy was 170 kcal/mol. The interaction energies for the six systems that included imidazole-ligands were underestimated, while the interaction energies for the other three systems were overestimated by the FF based calculations. This relationship has to be improved to generate simulations that are more reliable. The different LJ-parameters that were applied in this study could not lift this imbalance. The three FFs displayed quite similar trends. The CHARMM27 FF resulted in interaction energy changes that were slightly closer to the MP2 results, than OPLS-AA and AMBER99SB. Regarding the distances between the Zn atom and its nitrogen ligands on imidazole, the OPLS-AA resulted in accurate equilibrium distance, while with the AMBER99SB

and CHARMM27 FFs the distance were $\sim$ 0.10 Å too long.

The similarity between the FFs can be attributed to the similarity in partial charges. In **Paper III**, the atomic charges for Zn in the ten different model systems with four to six ligands, were calculated by the CHELPG method. It was shown that the water molecules apparently polarised the Zn ion, which resulted in a +2.2 au charge for the Zn ion in fully hydrated Zn complex. In the systems with a single acetate and water ligands, the Zn charge was just over +1 au, whereas in the complexes that contained one to three imidazoles, the charge of the Zn ion ranged between +0.5 and +1 au. The results reveal that considerable charge transfer is present in all system except the fully hydrated system.

Classical FFs do not always perform well at the interface between high dielectric substances such as water and low dielectric matter, e.g., proteins and membranes, as polarisation is only implicitly accounted for by classical FFs. Moreover, the need to correctly model the electronic distribution in the binding site is evident, as it differs depending on the metal-binding ligands. Therefore, in addition to the classical FFs, we employed calculations with a novel polarisable FF in **Paper III**. The results with the CHARMM-Drude polarisable FF revealed a marked improvement over all of the nonpolarisable FFs. The MM to QM difference was more consistent compared to the nonpolarisable FFs and there was agreement on the ordering of the interaction energies between the MM and QM calculations.

In order to gain further insight into Zn-ligand interactions, the free energies of interaction were estimated by QM calculations with polarisable continuum model (PCM) as solvent representation, and energy decomposition analysis with the LMOEDA/EDA-PCM methods were employed to examine the characteristics of the different complexes.

Tetrahydrofuran (THF) PCM was used to represent interactions in proteins, as it has a dielectric constant that can be representative for a protein environment. The relative orders of the interaction energies for the ten complexes were not the same in gas phase and in solvent. The free energies of interactions for the complexes with charged ligands were 30 kcal/mol more negative (i.e., more stable) in THF than those in water. The difference for the systems with neutral ligands were much smaller, and for the fully hydrated ion, embedding in THF increased the free energy of interaction by as little as 1 kcal/mol.

The results presented in **Paper III** suggest that simulations involving ligand exchange may be possible by use of the polarisable FF. One limitation with this study is that it does not include Cystein ligands, which are commonly encountered in Zn binding sites. Moreover, to state the performance of the polarisable FF in the dynamic situation in a simulation a validation following a full simulation protocol for a variety of binding sites has to be performed.

Paper IV. Computer Simulations of Alkali-Acetate Solutions: Accuracy of the Force fields in Different Concentrations

In **Paper IV** we examine and demonstrate in detail the interaction between acetate and alkali ions in solution, in different concentrations, for different ion FF parameters. We compared the solution properties from simulations against experimental values for association constants and activity coefficients. For Na^+, K^+ and Cs^+, classical FF parameters could reveal reasonable results for 0.1-1 M salt concentrations. For Li^+ and Na^+ two different classical parameter sets were tested. Simulations with parameters for Li^+-acetate solutions did not achieve

an overall good agreement with the experiment. The Åqvist parameters with SPC/E water yielded activity coefficients that were close to the experimental values for 0.1 M solutions, however overbinding of Li-acetate were found for all classical FF parameters in 1 M solutions.

EDA calculations revealed that the contributions of the exchange interaction to the overall free energy of interaction between Li^+ and water or acetate were not higher than for other ions, whereas polarisation was important for Li^+-water interactions. The polarisable parameters reproduced all observables well when the solution concentration was 1 M, but not in lower concentrations (0.1 M and below). For Na^+ and K^+ the CHARMM Drude-2013 polarisable FF resulted in too strong binding and in the Na^+ simulations clusters were formed, which are not found in experiments. The reason for cluster formation may be too strong ion-acetate interactions with the specific parameters applied. Differences between the two tested non-polarisable FF parameters for Na^+ appeared to be due to the interactions between the cations and the water, and point out that further tuning the ion-water interactions may be important to achieve better agreement with the experiment when interactions with carboxylates are considered.

In **Paper IV** a couple of alkali ion parameters were selected. To be more comprehensive even more parameter sets could be tested. For the polarisable FF, it is recommended to tune the Li^+ parameters for low concentrations and adjust the pair-interactions between Na^+ and the carboxylate oxygens.

Paper II. Protein-Ion Interactions: Simulations of Bovine Serum Albumin in Physiological Solutions of NaCl, KCl and LiCl

In **Paper II** we report on our study of the interactions of alkali ions with bovine serum albumin (BSA) and how they affect the protein motion and protein structure. Three independent 100 ns simulations of BSA were run in 0.1 M salt solution of LiCl, NaCl and KCl at temperatures of 300 K and 350 K. The simulations reveal that the affinity of Li^+ ions to BSA was much higher than that of Na^+ and K^+. There were about 30 Li^+ ions that bound to the protein surface, whereas only a few Na^+ or K^+ ions bound at the same ionic strength. Na^+ and K^+ had very similar affinities to BSA, which is in contrast to the expected better binding of Na^+ to the protein surface compared with K^+.

Overall, the results in **Paper II** show that the ion concentrations do not affect BSA structure and dynamics in a noticeable manner.

Chapter 7

Conclusions and future perspectives

In the molecular simulations of ions the FF parameters describing the non-bonded interactions are critical. Traditionally, this has been controlled by adjustment of the LJ parameters σ and $\mathcal{E}$, as this is an efficient way to do the calculations in the computer programs. The disadvantage is that there is risk to push oneself in a corner where it appears that what is optimised for one type of interaction does not fit another type and it becomes difficult to use parameters for a wide variety of simulations. This might be the case for the divalent zinc ion where the non-bonded model is unable to give an accurate reproduction for a general metal site as the specificity of ion-ligand interactions results in a diversity of ion-protein environments. In particular unwanted exchange of ligands might be unavoidable with such a classical model. Our results indicate that a better description of the Zn-His interaction relative to the Zn-water and Zn-Asp/Glu interaction could

possibly improve the simulation, under the premise that it is manageable. The use of polarisable FFs such as CHARMM Drude-2013 for the same Zn-ligand model systems showed that interaction energies with correct relative order could be obtained with inclusion of polarisation effects.

Our investigations of different non-polarisable FF parameters for alkali ion-acetate interactions show that agreement with experimental activity coefficient and association constants can be achieved for K^+ and Na^+, while the Li-acetate interaction is more problematic. However, differences between the non-polarisable FF parameters for Na^+ were found. They were found to arise due to interactions with the water, suggesting that further tuning the ion-water interactions can be useful to achieve better agreement with the experiments when interactions with carboxylates are considered. Though, it may as well be needed to take into account polarisable effects for the Na^+. Unfortunately, the Drude-2013 polarisable parameters for Na^+, as it is now, could not reproduce the experimental observations of the acetate salt solution. Further development of the FF parameters could lead to better results. For the Li^+-acetate solution (1 M), more realistic results were obtained by use of the polarisable FF than by classical FF parameters.

The results thus show that inclusion of the effects of electrons on the interaction between ions and different ligands in biochemical systems improves the simulations. With more efficient computing this will be possible for larger and larger molecules. However, I see that, as always in history, the more you learn the more you have the opportunity to simplify. This enables two branches; one that examines everything as detailed as possible that provides information usable for the other branch that develops simpler but correct models with suitable assumptions to accurately study larger systems and longer times. I see that the

area within computational biochemistry is still in a strong development phase and that it will lead to many changes ahead.

Constants

Atom mass unit (u) = 1.66057 E-27 kg
Avogadros number (N_a) = 6.022045 E+23 mol^{-1}
Boltzmann constant (k_B) = 1.3806 E-23 $m^2 \cdot kg/(s^2 \cdot K)$ =
1.3806 E-23 J/K = 1.98719 E-3 kcal/(mol · K)
Coulomb constant (k_e) = 8.99 E+9 Nm^2/C^2 = 332 kcal Å/(mol · e^2)
Electron charge (e) = 1.6022 E-19 C
Electron mass (m_e) = 9.1094 E-31 kg
Faraday constant (F) = 9.6485 E+4 C/mol
Gas constant (R) = 8.3145 J/(mol · K) = 1.9872 E-3 kcal/(mol · K)
Thermal energy k_BT at 25°C = 0.59 kcal/mol

Unit conversions

Electric potential: 1 J/C = 1V
Electric power: 1 J/s = 1W
Electric current: 1 C/s = 1 A
Energy: 1 kcal = 4.184 kJ = 2.6114 E+22 eV = 4186.8 Nm
Force: 1 N
LJ-parameters: $R_{min} = 2^{(1/6)}\sigma \approx 1.22\sigma$
Polarisability: $C^2m^2J^{-1}$
Pressure: 1 atm = 1.01325 bar = 101325 Pa

Graphs

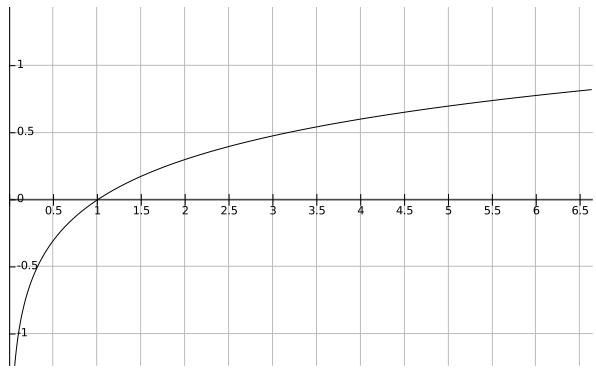


FIGURE 1: Graph of y=log(x). This function is applied for the association constant (K_a).

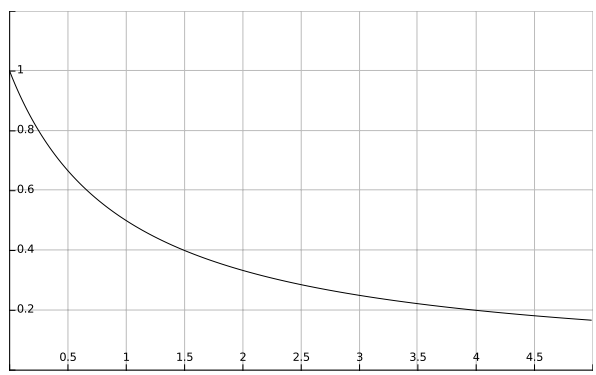


FIGURE 2: Graph of y=1/(1+x). In my work this function is applicable for estimation of the relation between the activity coefficient (γ), the association constant (K_a) and the concentration(C) in $\gamma = 1/(1 + K_a \cdot C)$.

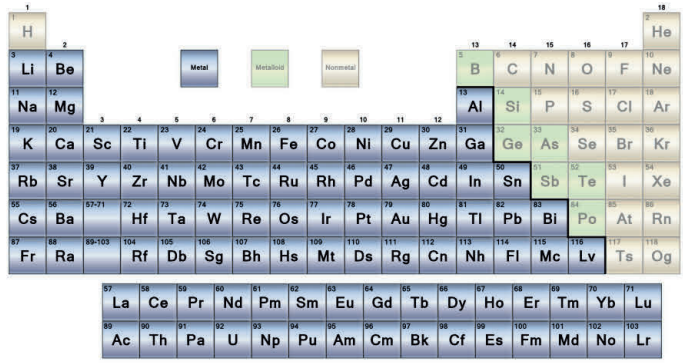


FIGURE 3: Periodic table of elements.
https://sciencenotes.org/list-metals/

Bibliography

[1] Carter, K. P., Young, A. M., and Palmer, A. E. Fluorescent sensors for measuring metal ions in living systems. *Chem. Rev.* **114**, 4564–4601 (2014).

[2] Ussing, H. H. *The alkali metal ions in isolated systems and tissues*, 1–195. Springer Berlin Heidelberg, Berlin, Heidelberg (1959).

[3] Lo Nostro, P. and Ninham, B. W. Hofmeister phenomena: An update on ion specificity in biology. *Chem. Rev.* **112**, 2286–2322 (2012).

[4] Stephanos, J. J. and Addison, A. W. *Chemistry of Metalloproteins: Problems and Solutions in Bioinorganic Chemistry*, 1–448. Somerset: Wiley. (2014).

[5] Friedman, R., Nachliel, E., and Gutman, M. Molecular dynamics of a protein surface: ion-residues interactions. *Biophys. J.* **89**, 768–81 (2005).

[6] Rink, L. Zinc and the immune system. *Proc. Nutr. Soc.* **59**, 541552 (2000).

[7] Chasapis, C. T., Loutsidou, A. C., Spiliopoulou, C. A., and Stefanidou, M. E. Zinc and human health: an update. *Archives of Toxicology* **86**, 521–534 (2012).

[8] Maret, W. *Zinc and the zinc proteome*, 479–501. Springer Netherlands, Dordrecht (2013).

[9] Deerfield, D. W., Carter, C. W., and Pedersen, L. G. Models for protein-zinc ion binding sites. II. The catalytic sites. *Int. J. Quant. Chem.* **83**, 150–165 (2001).

[10] Dudev, T. and Lim, C. Principles governing Mg, Ca, and Zn binding and selectivity in proteins. *Chem. Rev.* **103**, 773–788 (2003).

[11] Lin, F. and Wang, R. Systematic derivation of AMBER force field parameters applicable to zinc-containing systems. *J. Chem. Theory Comput.* **6**, 1852–1870 (2010).

[12] Sousa, S. F., Lopes, A. B., Fernandes, P. A., and Ramos, M. J. The zinc proteome: a tale of stability and functionality. *Dalton Trans.* , 7946–7956 (2009).

[13] Sousa, S. F., Fernandes, P. A., and Ramos, M. J. The carboxylate shift in zinc enzymes: a computational study. *J. Am. Chem. Soc.* **129**, 1378–1385 (2007).

[14] Maret, W. and Li, Y. Coordination dynamics of zinc in proteins. *Chem. Rev.* **109**, 4682–4707 (2009).

[15] Alberts, I. L., Nadassy, K., and Wodak, S. J. Analysis of zinc binding sites in protein crystal structures. *Protein Science* **7**, 1700–1716 (1998).

[16] Ferraroni, M., Rypniewski, W., Wilson, K., Viezzoli, M., Banci, L., Bertini, I., and Mangani, S. The crystal structure of the monomeric human SOD mutant F50E/G51E/E133Q at atomic resolution. The enzyme mechanism revisited. *J. Mol. Biol.* **288**(3), 413 – 426 (1999).

[17] Zirah, S., Kozin, S. A., Mazur, A. K., Blond, A., Cheminant, M., Sgalas-Milazzo, I., Debey, P., and Rebuffat, S. Structural changes of region 1-16 of the Alzheimer disease amyloid β-peptide upon zinc binding and in vitro aging. *J. Biol. Chem.* **281**, 2151–2161 (2006).

[18] Moroz, O. V., Blagova, E. V., Wilkinson, A. J., Wilson, K. S., and Bronstein, I. B. The crystal structures of human S100A12 in apo form and in complex with zinc: new insights into S100A12 oligomerisation. *J. Mol. Biol.* **391**, 536–551 (2009).

[19] Åqvist, J. and Warshel, A. Computer simulation of the initial proton transfer step in human carbonic anhydrase I. *J. Mol. Biol.* **224**, 7–14 (1992).

[20] McCall, K. A. and Fierke, C. A. Probing determinants of the metal ion selectivity in carbonic anhydrase using mutagenesis. *Biochemistry* **43**, 3979–3986 (2004).

[21] Frison, G. and Ohanessian, G. A comparative study of semiempirical, ab initio, and DFT methods in evaluating metal-ligand bond strength, proton affinity, and interactions between first and second shell ligands in Zn-biomimetic complexes. *J. Comput. Chem.* **29**, 416–433 (2008).

[22] Minicozzi, V., Stellato, F., Comai, M., Serra, M. D., Potrich, C., Meyer-Klaucke, W., and Morante, S. Identifying the minimal

copper- and zinc-binding site sequence in amyloid-beta peptides. *J. Biol. Chem.* **283**, 10784–10792 (2008).

[23] Istrate, A., Tsvetkov, P., Mantsyzov, A., Kulikova, A., Kozin, S., Makarov, A., and Polshakov, V. NMR solution structure of rat Aβ(1-16): toward understanding the mechanism of rats' resistance to Alzheimer's disease. *Biophys. J.* **102**, 136–143 (2012).

[24] Miller, Y., Ma, B., and Nussinov, R. Metal binding sites in amyloid oligomers: complexes and mechanisms. *Coord. Chem. Rev.* **256**, 2245 – 2252 (2012).

[25] Laity, J. H., Lee, B. M., and Wright, P. E. Zinc finger proteins: new insights into structural and functional diversity. *Curr. Opin. Struct. Biol.* **11**, 39 – 46 (2001).

[26] Tsvetkov, P. O., Kulikova, A. A., Golovin, A. V., Tkachev, Y. V., Archakov, A. I., Kozin, S. A., and Makarov, A. A. Minimal Zn^{2+} binding site of amyloid-β. *Biophys. J.* **99**, L84–L86 (2010).

[27] Klein, W., Jr., W. S., and Teplow, D. Small assemblies of unmodified amyloid β-protein are the proximate neurotoxin in Alzheimer's disease. *Neurobiol. Aging* **25**, 569 – 580 (2004).

[28] Yang, M. and Teplow, D. B. Amyloid β-protein monomer folding: Free-energy surfaces reveal alloform-specific differences. *J. Mol. Biol.* **384**, 450 – 464 (2008).

[29] Lee, C. and Ham, S. Characterizing amyloid-beta protein misfolding from molecular dynamics simulations with explicit water. *J. Comput. Chem.* **32**, 349–355 (2011).

[30] M., B. J., L., T. J., and L., S. *Chemical Bonds in Biochemistry.* W H Freeman, New York, 5th edition, (2002). Section 1.3.

[31] London, F. Zur theorie und systematik der molekularkräfte. *Zeitschrift für Physik* **63**, 245–279 (1930).

[32] Jones, J. E. On the determination of molecular fields. I. From the variation of the viscosity of a gas with temperature. *Proc. R. Soc. London A: Math., Phys. and Engineer. Sci.* **106**(738), 441–462 (1924).

[33] Jones, J. E. On the determination of molecular fields. II. From the equation of state of a gas. *Proc. R. Soc. London A: Math., Phys. Engineer. Sci.* **106**, 463–477 (1924).

[34] Li, P. and Merz, K. M. Metal ion modeling using classical mechanics. *Chem. Rev.* **117**, 1564–1686 (2017).

[35] Debye, P. Physik. In *25 Jahre Kaiser Wilhelm-Gesellschaft ur Förderung der Wissenschaften*, 46–56. Springer (1936).

[36] Hofmeister, F. Zur lehre von der wirkung der salze. *Arch. Exp. Pathol. Pharmakol.* **24**, 247–60 (1888).

[37] Marcus, Y. and Hefter, G. Ion pairing. *Chem. Rev.* **106**, 4585–4621 (2006).

[38] Van Der Vegt, N. F., Haldrup, K., Roke, S., Zheng, J., Lund, M., and Bakker, H. J. Water-mediated ion pairing: occurrence and relevance. *Chem. rev.* **116**(13), 7626–7641 (2016).

[39] Bjerrum, N. Ionic association. I. Influence of the ionic association on activity of ion at moderate degree of association. *Danske Vidensk. Selskab* **7**, 1–48 (1926).

[40] Banci, L. Molecular dynamics simulations of metalloproteins. *Curr. Opin. Chem. Biol.* **7**, 143 – 149 (2003).

[41] Friedman, R. Ions and the protein surface revisited: extensive molecular dynamics simulations and analysis of protein structures in alkali-chloride solutions. *J. Phys. Chem. B* **115**, 9213–9223 (2011).

[42] Friedman, R., Boye, K., and Flatmark, K. Molecular modelling and simulations in cancer research. *Biochim. Biophys. Acta (BBA) - Rev. Cancer* **1836**, 1–14 (2013).

[43] Friedman, R., Pellarin, R., and Caflisch, A. Amyloid aggregation on lipid bilayers and its impact on membrane permeability. *J. Mol. Biol.* **387**, 407–415 (2009).

[44] Friedman, R., Pellarin, R., and Caflisch, A. Soluble protofibrils as metastable intermediates in simulations of amyloid fibril degradation induced by lipid vesicles. *J. Phys. Chem. Lett.* **1**, 471–474 (2010).

[45] Friedman, R. and Caflisch, A. Surfactant effects on amyloid aggregation kinetics. *J. Mol. Biol.* **414**, 303–312 (2011).

[46] Friedman, R. Aggregation of amyloids in a cellular context: modelling and experiment. *Biochem. J.* **438** (2011).

[47] Li, W., Wang, J., Zhang, J., and Wang, W. Molecular simulations of metal-coupled protein folding. *Curr. Opin. Struct. Biol.* **30**, 25 – 31 (2015).

[48] Hernandez Maganhi, S., Jensen, P., Caracelli, I., Zukerman Schpector, J., Frohling, S., and Friedman, R. Palbociclib can overcome mutations in cyclin dependent kinase 6 that break hydrogen bonds between the drug and the protein. *Protein Science* **26**, 870–879 (2017).

[49] Friedman, R. The molecular mechanism behind resistance of the kinase FLT3 to the inhibitor quizartinib. *Proteins: Structure, Function, and Bioinformatics* **85**, 2143–2152 (2017).

[50] Derreumaux, P. and Mousseau, N. Coarse-grained protein molecular dynamics simulations. *J. Chem. Phys.* **126**, 025101–025101 (2007).

[51] Rojas, A., Liwo, A., Browne, D., and Scheraga, H. A. Mechanism of fiber assembly: treatment of Aβ peptide aggregation with a coarse-grained united-residue force field. *J. Mol. Biol.* **404**, 537–552 (2010).

[52] Riniker, S., Allison, J. R., and van Gunsteren, W. F. On developing coarse-grained models for biomolecular simulation: a review. *Phys. Chem. Chem. Phys.* **14**, 12423–12430 (2012).

[53] Rebič, M., Mocci, F., Uličný, J., Lyubartsev, A. P., and Laaksonen, A. Coarse-grained simulation of rodlike higher-order quadruplex structures at different salt concentrations. *ACS Omega* **2**, 386–396 (2017).

[54] Czene, A., Tóth, E., Németh, E., Otten, H., Poulsen, J.-C. N., Christensen, H. E., Rulíšek, L., Nagata, K., Larsen, S., and Gyurcsik, B. A new insight into the zinc-dependent DNA-cleavage by the colicin E7 nuclease: A crystallographic and computational study. *Metallomics* **6**, 2090–2099 (2014).

[55] Fernandes, H. S., Ramos, M. J., and Cerqueira, N. M. F. S. A. The catalytic mechanism of the pyridoxal-5′-phosphate-dependent enzyme, histidine decarboxylase: a computational study. *Chemistry - A Eur. J.* **23**, 9162–9173 (2017).

[56] Moreira, C., Ramos, M. J., and Fernandes, P. A. Clarifying the catalytic mechanism of human glutamine synthetase: a QM/MM study. *J. Phys. Chem. B* **121**, 6313–6320 (2017).

[57] Warshel, A. and Levitt, M. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **103**, 227–249 (1976).

[58] Marx, D. and Hutter, J. *Ab initio molecular dynamics: basic theory and advanced methods*. Cambridge University Press, (2009).

[59] Rempe, S. B., Pratt, L. R., Hummer, G., Kress, J. D., Martin, R. L., and Redondo, A. The hydration number of Li^+ in liquid water. *J. Am. Chem. Soc.* **122**, 966–967 (2000).

[60] Rempe, S. B. and Pratt, L. R. The hydration number of Na^+ in liquid water. *Fluid Phase Equilibria* **183-184**, 121 – 132 (2001).

[61] Leung, K., Rempe, S. B., and von Lilienfeld, O. A. Ab initio molecular dynamics calculations of ion hydration free energies. *J. Chem. Phys.* **130**(20), 204507 (2009).

[62] Rowley, C. N. and Roux, B. The solvation structure of Na^+ and K^+ in liquid water determined from high level ab initio molecular dynamics simulations. *J. Chem. Theory. Comput.* **8**, 3526–3535 (2012).

[63] Mehandzhiyski, A. Y., Riccardi, E., van Erp, T. S., Trinh, T. T., and Grimes, B. A. Ab initio molecular dynamics study on the interactions between carboxylate ions and metal ions in water. *J. Phys. Chem. B* **119**, 10710–10719 (2015).

[64] Schrödinger, E. Quantisierung als Eigenwertproblem. *Annalen der Physik* **384**, 361–376 (1926).

[65] Born, M. and Oppenheimer, R. Zur quantentheorie der molekeln. *Annalen der Physik* **389**, 457–484 (1927).

[66] Lewars, E. *Computational chemistry: introduction to the theory and applications of molecular and quantum mechanics*. Springer, (2003).

[67] Slater, J. C. A simplification of the Hartree-Fock method. *Phys. Rev.* **81**, 385–390 (1951).

[68] Møller, C. and Plesset, M. Note on an approximation treatment for many-electron systems. *Phys. Rev.* **46**, 0618–0622 (1934).

[69] Raghavachari, K., Trucks, G. W., Pople, J. A., and Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **157**, 479–483 (1989).

[70] Piecuch, P., Kucharski, S. A., Kowalski, K., and Musial, M. Efficient computer implementation of the renormalized coupled-cluster methods: the R-CCSD[T], R-CCSD(T), CR-CCSD[T], and CR-CCSD(T) approaches. *Comp. Phys. Commun.* **149**, 71 – 96 (2002).

[71] Hohenberg, P. and Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **136**, B864–B871 (1964).

[72] Sousa, S. F., Fernandes, P. A., and Ramos, M. J. General performance of density functionals. *J. Phys. Chem. A* **111**, 10439–10452 (2007).

[73] Kohn, W. and Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140**, A1133–A1138 (1965).

[74] Peverati, R. and Truhlar, D. G. Quest for a universal density functional: the accuracy of density functionals across a broad spectrum of databases in chemistry and physics. *Phil. Trans. R. Soc. A* **372**(2011), 20120476 (2014).

[75] Vosko, S. H., Wilk, L., and Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.* **58**, 1200–1211 (1980).

[76] Lee, C., Yang, W., and Parr, R. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B.* **37**, 785–789 (1988).

[77] Becke, A. D. A new mixing of hartree fock and local density-functional theories. *J. Chem. Phys.* **98**, 1372–1377 (1993).

[78] Schmider, H. and Becke, A. Optimized density functionals from the extended G2 test set. *J. Chem. Phys.* **108**, 9624–9631 (1998).

[79] Staroverov, V. N., Scuseria, G. E., Tao, J., and Perdew, J. P. Comparative assessment of a new nonempirical density functional: molecules and hydrogen-bonded complexes. *J. Chem. Phys.* **119**, 12129–12137 (2003).

[80] Zhao, Y. and Truhlar, D. G. Density functionals with broad applicability in chemistry. *Acc. Chem. Res.* **41**, 157–167 (2008).

[81] Chai, J.-D. and Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Phys. Chem. Chem. Phys.* **10**, 6615–6620 (2008).

[82] Grimme, S., Antony, J., Ehrlich, S., and Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **132**, 154104 (2010).

[83] Feller, D. The role of databases in support of computational chemistry calculations. *J. Comput. Chem.* **17**, 1571–1586 (1996).

[84] Schuchardt, K. L., Didier, B. T., Elsethagen, T., Sun, L., Gurumoorthi, V., Chase, J., Li, J., and Windus, T. L. Basis set exchange: a community database for computational sciences. *J. Chem. Info. Model.* **47**, 1045–1052 (2007).

[85] Hill, J. G. Gaussian basis sets for molecular applications. *Int. J. Quant. Chem.* **113**, 21–34 (2013).

[86] Boys, S. and Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.* **19**, 553–566 (1970).

[87] Mulliken, R. S. Electronic population analysis on LCAO-MO molecular wave functions. I. *J. Chem. Phys.* **23**, 1833–1840 (1955).

[88] Foster, J. and Weinhold, F. Natural hybrid orbitals. *J. Am. Chem. Soc.* **102**(24), 7211–7218 (1980).

[89] Reed, A. E., Curtiss, L. A., and Weinhold, F. Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. *Chemical Reviews* **88**(6), 899–926 (1988).

[90] Bader, R. F. *Atoms in molecules.* Wiley Online Library, (1990).

[91] Singh, U. C. and Kollman, P. A. An approach to computing electrostatic charges for molecules. *J. Comp. Chem.* **5**, 129–145 (1984).

[92] Breneman, C. M. and Wiberg, K. B. Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *J. Comput. Chem.* **11**, 361–373 (1990).

[93] Bayly, C. I., Cieplak, P., Cornell, W. D., and Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **97**, 10269–10280 (1993).

[94] Dupradeau, F.-Y., Czard, C., Lelong, R., Stanislawiak, ., Pcher, J., Delepine, J. C., and Cieplak, P. R.E.D D.B.: a database for RESP and ESP atomic charges, and force field libraries. *Nucleic Acids Research* **36**(suppl_1), D360–D367 (2008).

[95] Kitaura, K. and Morokuma, K. A new energy decomposition scheme for molecular interactions within the Hartree-Fock approximation. *Int. J. Quant. Chem.* **10**(2), 325–340 (1976).

[96] Jeziorski, B., Moszynski, R., and Szalewicz, K. Perturbation theory approach to intermolecular potential energy surfaces of van der Waals complexes. *Chem. Rev.* **94**(7), 1887–1930 (1994).

[97] Glendening, E. D. and Streitwieser, A. Natural energy decomposition analysis: an energy partitioning procedure for molecular interactions with application to weak hydrogen bonding, strong ionic, and moderate donor–acceptor interactions. *J. Chem. Phys.* **100**(4), 2900–2909 (1994).

[98] Fedorov, D. G., Kitaura, K., Li, H., Jensen, J. H., and Gordon, M. S. The polarizable continuum model (PCM) interfaced with the fragment molecular orbital method (FMO). *J. Comput. Chem.* **27**(8), 976–985 (2006).

[99] Stevens, W. J. and Fink, W. H. Frozen fragment reduced variational space analysis of hydrogen bonding interactions. application to the water dimer. *Chem. Phys. Lett.* **139**(1), 15–22 (1987).

[100] Peifeng, S. and Hui, L. Energy decomposition analysis of covalent bonds and intermolecular interactions. *J. Chem. Phys.* **131**, 014102 (2009).

[101] Su, P., Jiang, Z., Chen, Z., and Wu, W. Energy decomposition scheme based on the generalized Kohn-Sham scheme. *J. Phys. Chem. A* **118**(13), 2531–2542 (2014).

[102] Su, P., Liu, H., and Wu, W. Free energy decomposition analysis of bonding and nonbonding interactions in solution. *J. Chem. Phys.* **137**, 034111 (2012).

[103] Miertus, S., Scrocco, E., and Tomasi, J. Electrostatic interaction of a solute with a continuum - a direct utilization of ab-initio molecular potentials for the prevision of solvent effects. *Chem. Phys.* **55**, 117–129 (1981).

[104] Klamt, A. and Schüürmann, G. Cosmo: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans.* **2**(5), 799–805 (1993).

[105] Marenich, A. V., Cramer, C. J., and Truhlar, D. G. Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J. Phys. Chem. B* **113**(18), 6378–6396 (2009).

[106] Buckingham, R. A. The classical equation of state of gaseous helium, neon and argon. *Proc. R. Soc. London A: Math., Phys. Engineer. Sci.* **168**, 264–283 (1938).

[107] Born, M. and Mayer, J. E. Zur gittertheorie der ionenkristalle. *Zeitschrift für Physik* **75**, 1–18 (1932).

[108] White, D. N. A computationally efficient alternative to the Buckingham potential for molecular mechanics calculations. *J. Computer-Aided Molecular Design* **11**, 517–521, Sep (1997).

[109] Becconi, O., Ahlstrand, E., Salis, A., and Friedman, R. Protein-ion interactions: simulations of bovine serum albumin in physiological solutions of NaCl, KCl and LiCl. *Israel J. Chem.* **57**, 403–412 (2017).

[110] Ahlstrand, E., Hermansson, K., and Friedman, R. Interaction energies in complexes of Zn and amino acids: a comparison of ab initio and force field based calculations. *J. Phys. Chem. A* **121**, 2643–2654 (2017).

[111] Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.* **4**, 187–217 (1983).

[112] MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586–3616 (1998).

[113] MacKerell, A. D., Banavali, N., and Foloppe, N. Development and current status of the charmm force field for nucleic acids. *Biopolymers* **56**, 257–265 (2000).

[114] MacKerell, A. D., Feig, M., and Brooks, C. L. Improved treatment of the protein backbone in empirical force fields. *J. Am. Chem. Soc.* **126**, 698–699 (2004).

[115] Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., and Weiner, P. A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **106**, 765–784 (1984).

[116] Weiner, S. J., Kollman, P. A., Nguyen, D. T., and Case, D. A. An all atom force field for simulations of proteins and nucleic acids. *J. Comput. Chem.* **7**, 230–252 (1986).

[117] Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **117**, 5179–5197 (1995).

[118] Cheatham, T. E., Cieplak, P., and Kollman, P. A. A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *J. Biomol. Struct. Dynamics* **16**, 845–862 (1999).

[119] Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J., and Kollman, P. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **24**(16), 1999–2012 (2003).

[120] Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. Development and testing of a general AMBER force field. *J. Comput. Chem.* **25**, 1157–1174 (2004).

[121] Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. Comparison of multiple AMBER force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., and Bioinform.* **65**, 712–725 (2006).

[122] Jorgensen, W. L., Maxwell, D. S., and Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **118**, 11225–11236 (1996).

[123] Damm, W., Frontera, A., TiradoRives, J., and Jorgensen, W. L. OPLS all-atom force field for carbohydrates. *J. Comput. Chem.* **18**, 1955–1970 (1997).

[124] Kaminski, G., Friesner, R. A., Tirado-Rives, J., and Jorgensen, W. L. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B* **105**, 6474–6487 (2001).

[125] Robertson, M. J., Tirado-Rives, J., and Jorgensen, W. L. Improved peptide and protein torsional energetics with the OPLS-AA force field. *J. Chem. Theory Comput.* **11**, 3499–3509 (2015).

[126] Shi, Y., Ren, P., Schnieders, M., and Piquemal, J.-P. *Polarizable Force Fields for Biomolecular Modeling*, 51–86. John Wiley & Sons, Inc (2015).

[127] Halgren, T. A. and Damm, W. Polarizable force fields. *Curr. Opin. Struct. Biol.* **11**(2), 236 – 242 (2001).

[128] Leontyev, I. and Stuchebrukhov, A. Accounting for electronic polarization in non-polarizable force fields. *Physical Chemistry Chemical Physics* **13**(7), 2613–2626 (2011).

[129] Warshel, A., Kato, M., and Pisliakov, A. V. Polarizable force fields: history, test cases, and prospects. *J. Chem. Theory Comput.* **3**(6), 2034–2045 (2007).

[130] Rick, S. W., Stuart, S. J., and Berne, B. J. Dynamical fluctuating charge force fields: Application to liquid water. *J. Chem. Phys.* **101**(7), 6141–6156 (1994).

[131] Patel, S., Mackerell, A. D. J., and Brooks, C. L. r. CHARMM fluctuating charge force field for proteins: II protein/solvent properties from molecular dynamics simulations using a nonadditive electrostatic model. *J. Comput. Chem.* **25**, 1504–14 (2004).

[132] Drude, P. *The theory of optics*. Courier Corporation, (1925).

[133] Baker, C. M., Anisimov, V. M., and MacKerell, A. D. Development of CHARMM polarizable force field for nucleic acid bases based on the classical Drude oscillator model. *J. Phys. Chem. B* **115**, 580–596 (2011).

[134] Applequist, J., Carl, J. R., and Fung, K.-K. Atom dipole interaction model for molecular polarizability. Application to polyatomic molecules and determination of atom polarizabilities. *J. Am. Chem. Soc.* **94**(9), 2952–2960 (1972).

[135] Kaminski, G. A., Stern, H. A., Berne, B. J., Friesner, R. A., Cao, Y. X., Murphy, R. B., Zhou, R., and Halgren, T. A. Development of a polarizable force field for proteins via ab initio quantum chemistry: first generation model and gas phase tests. *J. Comput. Chem.* **23**, 1515–31 (2002).

[136] Ren, P., Wu, C., and Ponder, J. W. Polarizable atomic multipole-based molecular mechanics for organic molecules. *J. Chem. Theory Comput.* **7**, 3143–3161 (2011).

[137] Illingworth, C. J. and Domene, C. Many-body effects and simulations of potassium channels. *Proc. R. Soc. London A: Math., Phys. Eng. Sci.* (2009).

[138] Thole, B. Molecular polarizabilities calculated with a modified dipole interaction. *Chem. Phys.* **59**(3), 341 – 350 (1981).

[139] Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. The AMBER biomolecular simulation programs. *J. Comput. Chem.* **26**(16), 1668–1688 (2005).

[140] Brooks, B. R., Brooks, C. L., MacKerell, A. D., Nilsson, L., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A., Fischer, S., Gao, J., Hodoscek, M., Kuczera, K., Lazaridis, T., Ma, J., Paci, E., Pastor, R. W., Post, C. B., Schaefer, M., Tidor, B., Venable, R. W., Woodcock, H. L., Wu, X., and Karplus, M. CHARMM: the biomolecular simulation program. *J. Comp. Chem.* **30**, 1545–614 (2009).

[141] Berendsen, H. J. C., van der Spoel, D., and Vandrunen, R. Gromacs - a message-passing parallel molecular-dynamics implementation. *Comput. Phys. Commun.* **91**, 43–56 (1995).

[142] Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., and Berendsen, H. J. C. GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).

[143] Abraham, M. J., Murtola, T., Schulz, R., Pll, S., Smith, J. C., Hess, B., and Lindahl, E. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **12**, 19 – 25 (2015).

[144] Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kalé, L., and Schulten,

K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).

[145] Jiang, W., Hardy, D. J., Phillips, J. C., Mackerell, A. D., Schulten, K., and Roux, B. High-performance scalable molecular dynamics simulations of a polarizable force field based on classical Drude oscillators in NAMD. *J. Phys. Chem. Lett.* **2**, 87–92 (2011).

[146] Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. The protein data bank. *Eur. J. Biochem.* **80**, 319–324 (1977).

[147] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).

[148] Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., and Hermans, J. *Interaction models for water in relation to protein hydration.* Intermolecular Forces. D. Reidel Publishing Company Dordrecht, (1981).

[149] Berendsen, H. J. C., Grigera, J. R., and Straatsma, T. The missing term in effective pair potentials. *J. Phys. Chem.* **91**, 6269–6271 (1987).

[150] Berendsen, H. J. C., Postma, J. P. M., DiNola, A., and Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–90 (1984).

[151] Bussi, G., Donadio, D., and Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).

[152] Parrinello, M. and Rahman, A. Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).

[153] Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *J. chem. phys.* **81**, 511–519 (1984).

[154] Hoover, W. G. Canonical dynamics: equilibrium phase-space distributions. *Phys. rev. A* **31**, 1695 (1985).

[155] Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. LINCS: a linear constraint solver for molecular simulations. *J. Comp. Chem.* **18**, 1463–72 (1997).

[156] Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341 (1977).

[157] Miyamoto, S. and Kollman, P. A. SETTLE: an analytical version of the SHAKE and RATTLE algorithms for rigid water models. *J. Comp. Chem.* **13**, 952–962 (1992).

[158] Ewald, P. P. Die berechnung optischer und elektrostatischer gitterpotentiale. *Annalen der Physik* **369**, 253–287 (1921).

[159] Darden, T., York, D., and Pedersen, L. Particle mesh Ewald: an N-log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–92 (1993).

[160] Friedman, R., Nachliel, E., and Gutman, M. Protein surface - the dynamics of the interactions between protein, water and small solutes. *J. Biol. Phys.* **31**, 433–452 (2005).

[161] Project, E., Friedman, R., Nachliel, E., and Gutman, M. A molecular dynamics study of the effect of Ca^{2+} removal on calmodulin structure. *Biophys. J.* **90**, 3842–3850 (2006).

[162] Friedman, R. Structural and computational insights into the versatility of cadmium binding to proteins. *Dalton Trans.* **43**, 2878–2887 (2014).

[163] Mocci, F. and Saba, G. Molecular dynamics simulations of A·T-rich oligomers: sequence-specific binding of Na^+ in the minor groove of B-DNA. *Biopolymers* **68**, 471–485 (2003).

[164] Mocci, F. and Laaksonen, A. Insight into nucleic acid counterion interactions from inside molecular dynamics simulations is worth its salt. *Soft Matter* **8**(36), 9268–9284 (2012).

[165] Pineda De Castro, L. F., Dopson, M., and Friedman, R. Biological membranes in extreme conditions: anionic tetraether lipid membranes and their interactions with sodium and potassium. *J. Phys. Chem. B* **120**(41), 10628–10634, October (2016).

[166] Rebic, M., Laaksonen, A., Sponer, J., Ulicny, J., and Mocci, F. Molecular dynamics simulation study of parallel telomeric DNA quadruplexes at different ionic strengths: evaluation of water and ion models. *J. Phys. Chem. B* **120**, 7380–7391 (2016).

[167] Lopes, P. E. M., Guvench, O., and MacKerell, A. D. *Current Status of Protein Force Fields for Molecular Dynamics Simulations*, 47–71. Springer New York, New York, NY (2015).

[168] Rulíšek, L. and Havlas, Z. Ab initio calculations of monosubstituted (CH_3OH, CH_3SH, NH_3) hydrated ions of Zn^{2+} and Ni^{2+}. *J. Phys. Chem. A* **103**, 1634–1639 (1999).

[169] Dudev, T. and Lim, C. Tetrahedral vs octahedral zinc complexes with ligands of biological interest: a DFT/CDM study. *J. Am. Chem. Soc.* **122**, 11146–11153 (2000).

[170] Rulíšek, L. and Havlas, Z. Theoretical studies of metal ion selectivity. 1. DFT calculations of interaction energies of amino acid side chains with selected transition metal ions (Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+}, and Hg^{2+}). *J. Am. Chem. Soc.* **122**, 10428–10439 (2000).

[171] Rulíšek, L. and Havlas, Z. Theoretical studies of metal ion selectivity. 2. DFT calculations of complexation energies of selected transition metal ions (Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+}, and Hg^{2+}) in metal-binding sites of metalloproteins. *J. Phys. Chem. A* **106**, 3855–3866 (2002).

[172] Rulíšek, L. and Havlas, Z. Theoretical studies of metal ion selectivity. iii. A theoretical design of the most specific combinations of functional groups representing amino acid side chains for the selected metal ions (Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+}, and Hg^{2+}). *J. Phys. Chem. B* **107**, 2376–2385 (2003).

[173] Gutten, O. and Rulíšek, L. Predicting the stability constants of metal-ion complexes from first principles. *Inorganic Chemistry* **52**, 10347–10355 (2013).

[174] Rulíšek, L. and Šponer, J. Outer-shell and inner-shell coordination of phosphate group to hydrated metal ions (Mg^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+}) in the presence and absence of nucleobase. the role of nonelectrostatic effects. *J. Phys. Chem. B* **107**, 1913–1923 (2003).

[175] Joung, I. S. and Cheatham, T. E. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated

biomolecular simulations. *J. Phys. Chem. B* **112**, 9020–9041 (2008).

[176] Floris, F., Persico, M., Tani, A., and Tomasi, J. Free energies and structures of hydrated cations, based on effective pair potentials. *Chem. Phys.* **195**, 207 – 220 (1995).

[177] Chillemi, G., D'Angelo, P., Pavel, N. V., Sanna, N., and Barone, V. Development and validation of an integrated computational approach for the study of ionic species in solution by means of effective two-body potentials. the case of Zn^{2+}, Ni^{2+}, and Co^{2+} in aqueous solutions. *J. Am. Chem. Soc.* **124**(9), 1968–1976 (2002).

[178] Migliorati, V., Chillemi, G., and DAngelo, P. On the solvation of the Zn^{2+} ion in methanol: a combined quantum mechanics, molecular dynamics, and EXAFS approach. *Inorg. Chem.* **50**, 8509–8515 (2011).

[179] Jämbeck, J. P., Mocci, F., Lyubartsev, A. P., and Laaksonen, A. Partial atomic charges and their impact on the free energy of solvation. *J. Comput. Chemistry* **34**, 187–197 (2013).

[180] Sakharov, D. V. and Lim, C. Zn protein simulations including charge transfer and local polarization effects. *J. Am. Chem. Soc.* **127**, 4921–4929 (2005).

[181] Babu, C. S. and Lim, C. Empirical force fields for biologically active divalent metal cations in water. *J. Phys. Chem. A* **110**, 691–699 (2006).

[182] Åqvist, J. Ion-water interaction potentials derived from Free Energy Perturbation simulations. *J. Chem. Phys.* **94**, 8021–4 (1990).

[183] Hess, B. and van der Vegt, N. F. Cation specific binding with protein surface charges. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 13296–13300 (2009).

[184] Hu, L. and Ryde, U. Comparison of methods to obtain force-field parameters for metal sites. *J. Chem. Theory Comput.* **7**, 2452–2463 (2011).

[185] Seminario, J. M. Calculation of intramolecular force fields from second-derivative tensors. *Int. J. Quant. Chem. Quant. Chem. Symp.* **30**, 1271–1277 (1996).

[186] Peters, M. B., Yang, Y., Wang, B., Füsti-Molnár, L., Weaver, M. N., and Merz, K. M. Structural survey of zinc containing proteins and the development of the zinc AMBER force field (ZAFF). *J. Chem. Theory Comput.* **6**, 2935–2947 (2010).

[187] Darden, T., Pearlman, D., and Pedersen, L. G. Ionic charging free energies: spherical versus periodic boundary conditions. *J. chem. phys.* **109**, 10921–10935 (1998).

[188] Dang, L. X. Free energies for association of Cs^+ to 18-crown-6 ether in water. A molecular dynamics study including counter ions. *Chem. Phys. Lett.* **227**, 211 – 214 (1994).

[189] Dang, L. X. Mechanism and thermodynamics of ion selectivity in aqueous solutions of 18-crown-6 ether: a molecular dynamics study. *J. Am. Chem. Soc.* **117**, 6954–6960 (1995).

[190] Dang, L. X. Computational study of ion binding to the liquid interface of water. *J. Phys. Chem. B* **106**, 10388–10394 (2002).

[191] Jensen, K. P. and Jorgensen, W. L. Halide, ammonium, and alkali metal ion parameters for modeling aqueous solutions. *J. Chem. Theory Comput.* **2**, 1499–1509 (2006).

[192] Horinek, D., Mamatkulov, S. I., and Netz, R. R. Rational design of ion force fields based on thermodynamic solvation properties. *J. Chem. Phys.* **130**, 124507 (2009).

[193] Fyta, M., Kalcher, I., Dzubiella, J., Vrbka, L., and Netz, R. R. Ionic force field optimization based on single-ion and ion-pair solvation properties. *J. Chem. Phys.* **132**, 024911–024911 (2010).

[194] Fyta, M. and Netz, R. R. Ionic force field optimization based on single-ion and ion-pair solvation properties: going beyond standard mixing rules. *J. Chem. Phys.* **136**, 124103 (2012).

[195] Mark, P. and Nilsson, L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *J. Phys. Chem. A* **105**(43), 9954–9960 (2001).

[196] Weerasinghe, S. and Smith, P. E. A Kirkwood-Buff derived force field for sodium chloride in water. *J. Chem. Phys.* **119**, 11342–11349 (2003).

[197] Gee, M. B., Cox, N. R., Jiao, Y., Bentenitis, N., Weerasinghe, S., and Smith, P. E. A Kirkwood-Buff derived force field for aqueous alkali halides. *J. Chem. Theory Comput.* **7**, 1369–1380 (2011).

[198] Klasczyk, B. and Knecht, V. Kirkwood-Buff derived force field for alkali chlorides in simple point charge water. *J. Chem. Phys.* **132**, 024109–024109 (2010).

[199] Kirkwood, J. G. and Buff, F. P. The statistical mechanical theory of solutions. I. *J. Chem. Phys.* **19**, 774–777 (1951).

[200] Stote, R. H. and Karplus, M. Zinc binding in proteins and solution: a simple but accurate nonbonded representation. *Proteins* **23**, 12–31 (1995).

[201] Li, P., Roberts, B. P., Chakravorty, D. K., and Merz, K. M. Rational design of particle mesh Ewald compatible Lennard-Jones parameters for +2 metal cations in explicit solvent. *J. Chem. Theory Comput* **9**, 2733–2748 (2013).

[202] Gong, W., Wu, R., and Zhang, Y. Thiol versus hydroxamate as zinc binding group in hdac inhibition: an ab initio QM/MM molecular dynamics study. *J. Comput. Chem.* **36**, 2228–2235 (2015).

[203] Zhu, X., Lopes, P. E., and MacKerell, A. D. Recent developments and applications of the CHARMM force fields. *Wiley Interdisc. Rev.: Comput. Molecular Sci.* **2**(1), 167–185 (2012).

[204] Wu, R., Lu, Z., Cao, Z., and Zhang, Y. A transferable nonbonded pairwise force field to model zinc interactions in metalloproteins. *J. Chem. Theory. Comput.* **7**, 433–443 (2011).

[205] Zhu, T., Xiao, X., Ji, C., and Zhang, J. Z. H. A new quantum calibrated force field for zincprotein complex. *J. Chem. Theory Comput.* **9**, 1788–1798 (2013).

[206] Li, P. and Merz, K. M. Taking into account the ion-induced dipole interaction in the nonbonded model of ions. *J. Chem. Theory Comput.* **10**(1), 289–297 (2014).

[207] Gresh, N., El Hage, K., Perahia, D., Piquemal, J.-P., Berthomieu, C., and Berthomieu, D. Polarizable molecular mechanics studies of Cu(I)/Zn(II) superoxide dismutase: bimetallic binding site and structured waters. *J. Comput. Chem.* **35**(29), 2096–2106 (2014).

[208] Åqvist, J. and Warshel, A. Free energy relationships in metalloenzyme-catalyzed reactions. calculations of the effects of metal ion substitutions in staphylococcal nuclease. *J. Am. Chem. Soc.* **112**(8), 2860–2868 (1990).

[209] Pang, Y.-P. Successful molecular dynamics simulation of two zinc complexes bridged by a hydroxide in phosphotriesterase using the cationic dummy atom method. *Proteins: Structure, Function, and Bioinformatics* **45**, 183–189 (2001).

[210] Duarte, F., Bauer, P., Barrozo, A., Amrein, B. A., Purg, M., qvist, J., and Kamerlin, S. C. L. Force field independent metal parameters using a nonbonded dummy model. *The Journal of Physical Chemistry B* **118**(16), 4351–4362 (2014).

[211] Jing, Z., Qi, R., Liu, C., and Ren, P. Study of interactions between metal ions and protein model compounds by energy decomposition analyses and the AMOEBA force field. *J. Chem. Phys.* **147**, 161733 (2017).

[212] Yu, H., Whitfield, T. W., Harder, E., Lamoureux, G., Vorobyov, I., Anisimov, V. M., MacKerell, A. D., and Roux, B. Simulating monovalent and divalent ions in aqueous solution using a Drude polarizable force field. *J. Chem. Theory Comput.* **6**, 774–786 (2010).

[213] Luo, Y., Jiang, W., Yu, H., MacKerell, A. D., and Roux, B. Simulation study of ion pairing in concentrated aqueous salt solutions with a polarizable force field. *Faraday Discuss.* **160**, 135–149 (2013).

[214] Li, H., Ngo, V., Da Silva, M. C., Salahub, D. R., Callahan, K., Roux, B., and Noskov, S. Y. Representation of ion-protein interactions using the Drude polarizable force-field. *J. Phys. Chem. B* **119**, 9401–9416 (2015).

[215] Grossfield, A., Ren, P., and Ponder, J. W. Ion solvation thermodynamics from simulation with a polarizable force field. *J. Am. Chem. Soc.* **125**, 15671–82 (2003).

[216] Wu, J. C., Piquemal, J.-P., Chaudret, R., Reinhardt, P., and Ren, P. Polarizable molecular dynamics simulation of Zn(II) in water using the AMOEBA force field. *J. Chem. Theory Comput.* **6**, 2059–2070 (2010).

[217] Xiang, J. Y. and Ponder, J. W. A valence bond model for aqueous Cu(II) and Zn(II) ions in the AMOEBA polarizable force field. *J Comput. Chem.* **34**(9), 739–749 (2013).

[218] Savelyev, A. and MacKerell, A. D. Differential deformability of the DNA minor groove and altered BI/BII backbone conformational equilibrium by the monovalent ions Li^+, Na^+, K^+, and Rb^+ via water-mediated hydrogen bonding. *J. Chem. Theory Comput.* **11**(9), 4473–4485 (2015).

[219] Vergara-Jaque, A., Fong, P., and Comer, J. Iodide binding in sodium-coupled cotransporters. *J. Chem. Info. Model.* **57**(12), 3043–3055 (2017).

[220] Zhang, J., Yang, W., Piquemal, J.-P., and Ren, P. Modeling structural coordination and ligand binding in zinc proteins with a polarizable potential. *J. Chem. Theory Comput.* **8**, 1314–1324 (2012).

Acknowledgements

I would like to thank my supervisor Ran Friedman to let me get this position and for guidance during the work. I would also like to thank my co-supervisors Kersti Hermansson, Lynn Kamerlin and Daniel Spångberg for scientific discussions and inspiration.
A special thanks to all the people at Norrgård, Linnaeus University, for the creative environment and the happy mood in the building.