**Univerzita Karlova v Praze**

**Přírodovědecká fakulta**


Katedra fyzikální a makromolekulární chemie

Modelování chemických vlastností nano- a biostruktur





**Ing. Jana Hladílková**


**Ion Specific Hofmeister Effects on Peptides and Proteins**

**Iontově specifické hofmeisterovské efekty na peptidy a proteiny**


Disertační práce


Ústav organické chemie a biochemie, AV ČR

Centrum  biomolekul a komplexních molekulových systémů

Školitel: Prof. Pavel Jungwirth, DSc.


Praha, 2014

Poděkovat bych chtěla především svému školiteli, Pavlu Jungwirthovi, který mi umožnil změnit obor a věnovat se práci, která mě celoživotně nadchla. Velké díky patří Honzovi Heydovi, který mě zasvětil do tajů výpočetní chemie a s velkou trpělivostí mě naučil vše potřebné. V neposlední řadě děkuji svým kolegům, se kterými slovo spolupráce získalo úplně nový rozměr a zejména Aleně s Evou za přátelství a morální podporu.

Děkuji své rodině, která je mi  velkou oporou při plnění mých snů, zejména si vážím tolerance svého manžela, se kterou přijímá má životní rozhodnutí.

**Prohlášení:**

Prohlašuji, že jsem závěrečnou práci zpracovala samostatně a že jsem uvedla všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, 25.6.2014

Jana Hladílková

# Contents

Title: Ion Specific Hofmeister Effects on Peptides and Proteins

Author: Ing. Jana Hladílková

Department: Physical and Macromolecular Chemistry

Advisor: Prof. Pavel Jungwirth, DSc., IOCB AS CR

Advisor's email address: pavel.jungwirth@uochb.cas.cz

**Abstract:** Classical molecular dynamics simulations in combination with advanced methods of analysis were used to shed light on missing parts of our molecular understanding of the Hofmeister series. In tandem with various experimental techniques, real proteins as well as model systems were investigated in aqueous salt solutions in order to identify and quantify ion-protein interactions either leading or not leading to the canonical cationic and anionic Hofmeister ordering.

The potassium cation was found to significantly enhance the BHMT enzymatic activity in contrast to the rest of the common monovalent cations. In the quest to rationalize this behavior, a key potassium binding site in the vicinity of the active site was discovered and described. Moreover, the exceptionally strong effect of $K^+$ on the enzymatic activity was explained by hydration properties of the cations within the limited space of the active site in interplay with their attraction to the nearby negatively charged residues. By contrast, only a small and indirect influence, which follows the cationic Hofmeister series, was established for the LinB dehalogenase. The binding hot spot for all the cations was assigned at the mouth of the tunnel leading to the active site. This assumption was further supported by single point mutations at the tunnel mouth of this enzyme.

A systematic study of anion-peptide interactions was realized for a variety of model systems with aid of NMR experiments. Examination of the $(VPGVG)_{120}$ polypeptide revealed a dominant role of anion-backbone interactions in neutral biological systems. On top of that, it was shown that anions are not attracted to the nonpolar side chains of residues like valine. In order to quantify our previous findings, capped triglycine was investigated as a model system for the peptide bond in aqueous solutions of five sodium salts. It was confirmed that the more weakly hydrated the anion is, the more it interacts with the peptide backbone ($SCN^- > I^- > Br^- > Cl^- > SO_4^{2-}$). Consequently, thiocyanate and iodide act like salting-in agents, bromide and chloride are neutral in contrast with sulfate which is repelled from the backbone surface and shows a salting-out behavior.

To capture the effect of charged residues, anion-peptide binding sites for uncapped aqueous triglycine were explored. In this case, charge-charge interactions dominate, resulting in a reversed Hofmeister series ($SO_4^{2-} > Cl^- > Br^- > I^-$), i.e., the more strongly hydrated the anion is, the larger its affinity to the positively charged N-terminus. Interestingly, $SCN^-$ does not fully follow this rule as a consequence of a synergy between charge-charge and anion-backbone interactions of this non-spherical ion. A direct comparison of our results with experimental data published in the 1970s led to a discovery of an error in the original publication. It was proven that an inefficient synthetic procedure caused assignment of the measured salting-out constants to a fully capped triglycine molecule instead of a half-capped version, which matched with our computational results.

Finally, an innovative way of studying ion-protein interactions was demonstrated on the example of electrophoretic measurements and calculations of neutral model systems and electroosmotic flow markers.

**Keywords:** molecular dynamics, proteins, peptides, ions, Hofmeister series.

Název práce: Iontově specifické hofmeisterovské efekty na peptidy a proteiny

Autor: Ing. Jana Hladílková

Katedra: Fyzikální a makromolekulární chemie

Vedoucí doktorské práce: Prof. Pavel Jungwirth, DSc., ÚOCHB AV ČR

E-mail vedoucího: pavel.jungwirth@uochb.cas.cz

**Abstrakt:** V této dizertaci byla použita klasická molekulová dynamika s pokročilými metodami analýzy dat pro doplnění a zároveň objasnění jevů týkající se Hofmeisterovy řady iontů. V kombinaci se širokým spekterem experimentálních metod byly studovány jak reálné proteiny, tak i modelové systémy v roztocích hofmeisterovských solí. Cílem práce bylo identifikovat a kvantifikovat specifické interakce iontů s peptidy a proteiny, které způsobují výsledné hofmeistrovské řazení kationtů a aniontů, či které toto pravidlo porušují.

Příkladem nehofmeisterovského chování je zrychlení enyzmatické reakce BHMT transferázy pomocí draselných iontů, zatímco ostatní běžné monovalentní kationty nemají tento efekt. Molekulové simulace nám umožnily určit a detailně popsat vazebné místo poblíž aktivního místa enzymu, které bylo později krystalograficky potvrzeno. Specificita pro draselný kation byla vysvětlena na základě hydratačních vlastností jednotlivých kationtů a interakcí se záporně nabitými rezidui aktivního místa. Naproti tomu pouze malý efekt monovalentních kationtů na enzymatickou reakci, který odpovídá hofmeisterovskému řazení, byl pozorován pro LinB dehalogenázu. V tomto případě se kationty vážou u ústí tunelu vedoucího ke katalytické triádě, a proto pouze nepřímo ovlivňují rychlost reakce, což bylo podpořeno výpočty i měřením mutovaných variací tohoto enzymu.

Dále byla provedena systematická studie interakcí aniontů s peptidy na modelových strukturách různých velikostí s podporou dat z nukleární magnetické rezonance. Výsledky pro $(VPGVG)_{120}$ polypeptid nejen odhalily dominanci interakcí aniontů s páteří peptidu, ale navíc ukázaly, že anionty nejsou přitahovány k nepolárním částem reziduí, jakým je například valin. Následující analýza triglycinu s oběma chránícími skupinami, který byl zvolen jako model peptidové vazby, potvrdila, že čím slaběji je anion hydratován, tím silněji se váže na bílkovinovou páteř ($SCN^- > I^- > Br^- > Cl^- > SO_4^{2-}$). Ve výsledku se proto thiokyanát a jodid chovají jako vsolovací činidla, bromid a chlorid jako neutrální ionty, zatímco sulfát peptidy/proteiny vysoluje.

Efekt nabitých reziduí na interakce iontů s peptidy/proteiny byl kvantifikován na zwitteriontové (nechráněné) struktuře triglycinu. V tomto případě hraje hlavní roli interakce aniontů s pozitivně nabitým N-koncem molekuly, což vede k převrácení pořadí iontů v Hofmeisterově řadě ($SO_4^{2-} > Cl^- > Br^- > I^-$). Thiokyanát porušuje toto řazení, jelikož se jako jediný váže jak na pozitivně nabitý N-konec tripeptidu, tak i na peptidovou vazbu, což mu umožňuje jeho nekulový tvar. Při přímém srovnání našich výsledků s experimentálními daty publikovanými v 70. letech, se nám podařilo odhallit chybu v původní publikaci. Ukázali jsme, že otištěné vysolovací konstanty souhlasí s našimi výpočty pro molekulu triglycinu s jednou chránící skupinou na N-konci, zatímco se rozcházejí s výsledky pro autory uvedenou oboustranně chráněnou variantu. Náš předpoklad byl potvrzen nereprodukovatelností posledního syntetického kroku v publikovaném článku.
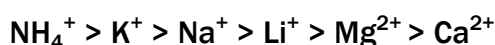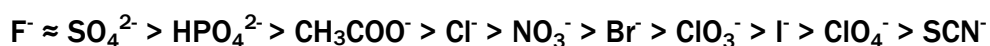
Jako ukázka nového přístupu jdoucí za jednoduché principy hofmeistrovského vázání kationtů a aniontů na peptidy a proteiny, byly uvedeny výpočty a měření elektroforetických dat pro neutrální molekuly včetně elektroforetických markerů.

**Klíčová slova**: molekulová dynamika, proteiny, peptidy, ionty, Hofmeisterova řada iontů.

# Chapter 1

# Introduction

Franz Hofmeister, a professor of pharmacology at the German University in Prague, is considered to be the first scientist who systematically studied the effects of ions on proteins[1]. In the late 19th century, he and his co-workers published a series of seven papers with a common title: 'Lessons on the effects of salts', all in German (for English translation see ref. 2). In the second article[3] entitled 'Concerning regularities in the protein-precipitating effects of salts and the relationship of these effects to the physiological behaviour of salts' Hofmeister published a series of experiments on precipitation of the egg-white proteins from aqueous salt solutions. From these measurements he was able to order the salts by their ability to precipitate the proteins. On top of that, he inventively separated anionic and cationic effects using salts with a common cation or anion, respectively, and the resulting series of ions got to be known as the lyotropic or Hofmeister series:

$$F^- \approx SO_4^{2-} > HPO_4^{2-} > CH_3COO^- > Cl^- > NO_3^- > Br^- > ClO_3^- > I^- > ClO_4^- > SCN^-$$

$$NH_4^+ > K^+ > Na^+ > Li^+ > Mg^{2+} > Ca^{2+}$$

In the following and probably the most significant paper[4] of the series with the title: 'On the water withdrawing effect of salts' Hofmeister attempted to provide an explanation for his observations and presented remarkable experimental data for ions interacting with colloidal particles and other proteins. He concluded that the ordering of the ions is caused by the water-absorbing effects of the salts, in other words, by their strength of hydration, which was in 1930s rephrased into the kosmotropes/chaotropes theory[5,6]. According to this theory, kosmotropes, the structure-making ions, have the ability to strongly organize water molecules further than in their immediate vicinity[7]. As a result, they effectively dehydrate the protein by 'stealing' water molecules, leading to a precipitation of the protein from the aqueous solution (= salting-out effect). In contrast, chaotropes, the structure-breaking ions, do not have the ability to strongly bind water molecules[8,9], therefore, the protein can be hydrated sufficiently and remains dissolved in the aqueous solution (= salting-in effect).

Upon a deeper look into this theory, we find several problems, which cannot be ignored. First of all, there is conclusive experimental and theoretical evidence that no long-range organization of water molecules around ions exists, particularly not around monovalent ones[10,11]. Secondly, the role of the protein is completely ignored in this theory and according to it the structural changes or mutations of the protein thus should not make any difference in the results. However, the numerous experimentally observed violations of the Hofmeister ordering cannot be explained by such a simple theory. A good example is the precipitation of the protein lysozyme. While at high pH and high ionic strength the ions follow the Hofmeister series, at low pH and moderate salt concentration a reversed Hofmeister series is observed, which means the weakly hydrated ions salt-out the lysozyme more effectively than the strongly hydrated ones[12,13].

Realizing the weak points in the explanations mentioned above, the main scientific interest shifted from the ion-water interactions to studies of the ion-protein interactions in the late 1950s. Experimental approaches were based on thermodynamic measurements (of viscosity, calorimetry, etc.) of small model systems, which were simple enough for interpretation and, at the same time, representative for the protein structure. As the simplest model of the protein backbone, N-methylacetamide (NMA) was examined[14,15], followed by a systematic study of solubilities of polyglycine oligomers[16,17], in which the salting-out constants of Hofmeister ions were obtained from the measurements. A general conclusion that weakly hydrated anions ($I^-$, $SCN^-$, etc.) and strongly hydrated cations (such as $Li^+$ or $Mg^{2+}$) interact with the peptide backbone unlike strongly hydrated anions ($F^-$, $SO_4^{2-}$, ...) and weakly hydrated cations ($Cs^+$, $NH_4^+$, and similar) followed from the data. At the end of the 1970s, first attempts of unraveling the whole protein structure and identifying the binding sites for ions were made using the X-ray structural methods and NMR spectroscopy[18]. The lack of sufficient resolution and sensitivity of the measurements, as well as the poor computational resources and techniques, resulted in the topic falling out of interest for the next two decades.

The first modern Hofmeister conference was organized in 2004, where the topic was reopened and, more importantly, the key questions were formulated. The original Hofmeister papers were translated into English[2] in a special journal issue[19], summarizing the current knowledge and the number of papers dedicated to this topic increased significantly, which was reflected in the term of 'Rennaisance for Hofmeister'[20]. The advanced experimental and computational techniques allowed scientists to study ion-specific effects in more complex systems, including a variety of proteins, DNA, biomembranes, as well as non-biological molecules and nanostructures.

While no universal solution of the whole problem exists yet, the goal of this thesis is to provide some of the missing pieces of the Hofmeister puzzle, focusing on the interactions of ions with proteins and peptides. The latest studies showed that the ion-protein interactions consist of two roughly equally important key components - ion-backbone and ion-side chain binding[21-24] and a reductionist approach can be applied. Within this thesis, we address using molecular dynamics simulations in close contact with experiments the following question – where exactly and how strongly the Hofmeister ions bind to the backbone, to the side chains, or at specific binding sites of proteins and peptides (i. e., enzymatic active site).

## *Chapter 2*

## **Synergy of Theory and Experiment**

To come up with explanations for the biologically relevant questions concerning ion-protein interactions, it is extremely important to employ both theoretical and experimental methods. Computational power has increased to a point, where we can simulate anything from a single atom almost to a whole cell with lower or higher degrees of simplifications. Living under the illusion that what appears on the screen "must be true", we might forget to justify our approximations and to look for external checks. Benchmarking of the theoretical methods should be the standard and, on top of that, comparison with experimental results should be requested. Then, if used with care and expertise, theoretical calculations can, on an atomistic scale, provide insights into experimental observations with a level of detail, which is hardly accessible to experimental techniques. When we understand the nature of experimental findings, we may be able to distinguish cause from consequence or even predict results yet unavailable to the experiment. The following section serves as a list of theoretical and experimental techniques, which were crucial for obtaining the results and reaching the conclusions presented in this thesis.

## **2.1 Computational simulations and analytical tools**

All theoretical results presented in this thesis were obtained from simulations using readily available molecular dynamics (MD) software packages – AMBER[25] and GROMACS[26]; except for quantum mechanics/molecular mechanics (QM/MM) calculations, where we used the Chemshell[27] package to couple ORCA[28] as QM software and DL_POLY[29] using AMBER force field as MM package. The following analyses of converged data sets were performed to obtain quantities which are directly or indirectly comparable with experiments or transformed by visualization techniques to display statistical properties that help understanding of the results.

### **2.1.1 Spatial distribution function**

The spatial distribution function (SDF), or more specifically the ion map around an investigated molecule, determines the three-dimensional density distribution of atoms

(specifically ions) in a coordinate system (Eq. 1):

$$g_{SDF}(\vec{r}) = \frac{\rho(\vec{r})}{\rho_{bulk}}$$ (1)

In this equation, $\rho_{bulk}$ stands for the ion bulk density and $\rho(\vec{r})$ indicates the ion number density at the position $\vec{r}$ defined in the coordinate system from the biomolecule. The SDF as a radial and angular distribution can be simplified in case of spherical particles, where only the relative distance from a defined center matters. Therefore, when we integrate over all the angles, we end up with a definition of the radial distribution function (RDF) as in Eq.2.

$$g(r) = \frac{1}{4\pi r^2} \int\limits_0^{2\pi} \int\limits_0^{\pi} g(r, \varphi, \theta) \, r^2 \, d\varphi \, d\theta$$ (2)

The 3D ion probability (SDF) can be visualized as a "cloud" with different iso-level values defining at least how many times must the local density be higher than the bulk density, to make the cloud visible (Fig. 1). The analogy for the iso-level is the iso-contour in the two-dimensional space.
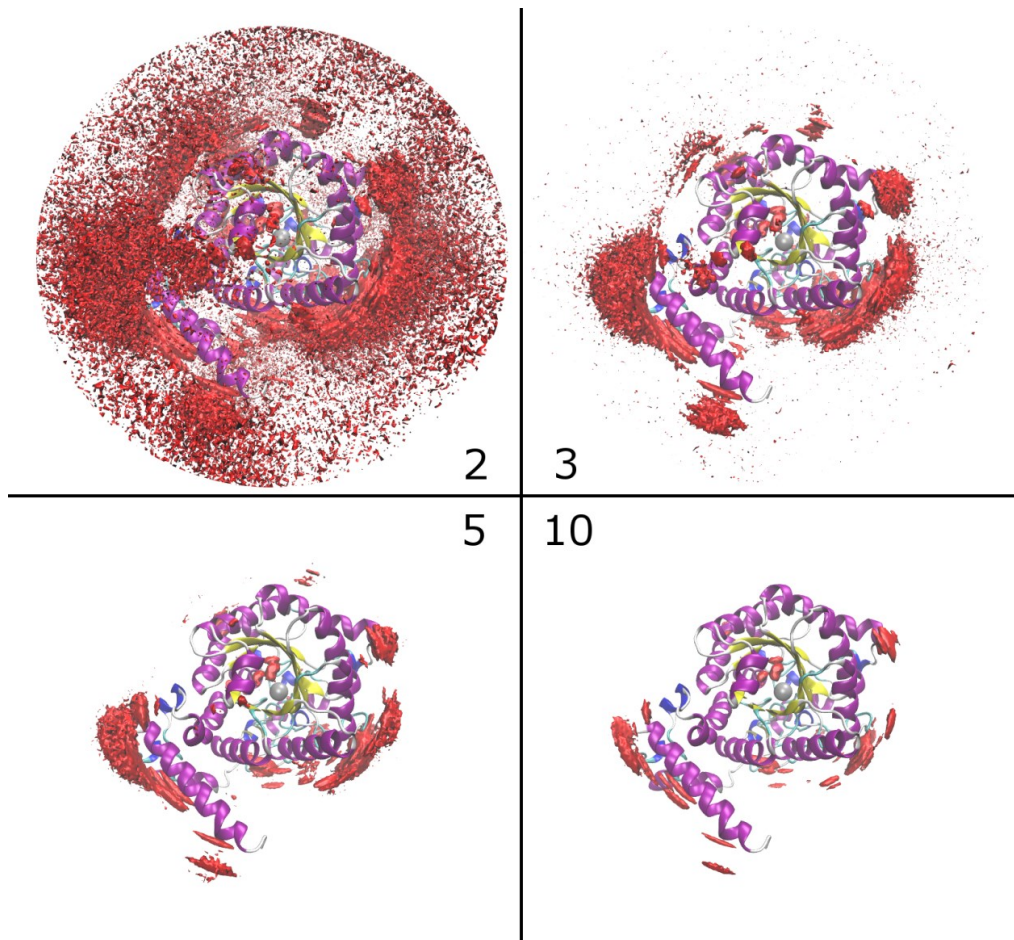


**Fig. 1**: Potassium density map around the BMHT enzyme shown with different values of iso-levels (2, 3, 5, and 10) with iso-level = 1 corresponding to the bulk density.

During the procedure of calculating ion maps, the biomolecule is in every snapshot centered and rotated to the same position as the reference one and the occurrence of all ions is detected and recorded into a three dimensional grid. The final map is then normalized to the bulk density (iso-level = 1). The difficulty of this approach arises from the fact that we select exactly three atoms, which become the referential triade for overlapping the biomolecule from every snapshot. Therefore, the more flexible the biomolecule is, the more blurred maps we get. To avoid this, we need to run longer simulations to obtain better statistics and then to divide the trajectory into several clusters, corresponding to characteristic biomolecular geometries, and calculate the density maps for each of them. The final ion density map can then be obtained by averaging over the clusters. The other option is to constrain the molecular movements and freeze the molecule in its (experimentally) most probable molecular conformation.

## 2.1.2 Proximal distribution function

In addition to the radial distribution function (RDF) it is convenient to define a proximal distribution function[29]. The need to introduce this new quantity originates in the limitations of the RDF, where the spherical symmetry is assumed, which for complex systems is not strictly applicable. When we want to distinguish between different types of interactions or different parts of the molecule interacting with ions, we need to divide the surrounding space into parts. Consequently, every defined group has its own interacting area and every ion in every snapshot is counted only once as belonging to the part of the space where it occurs at the given moment (Fig. 2).
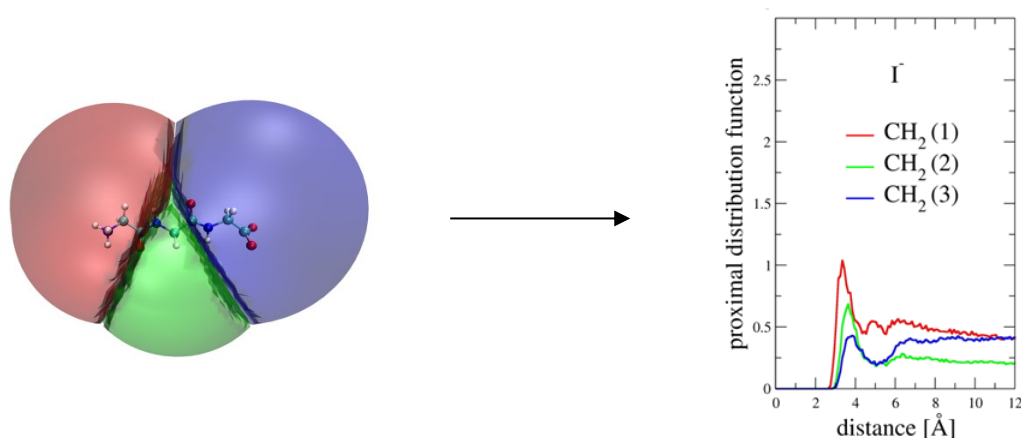


**Fig. 2**: Visualization of the space division around an uncapped glycine trimer and corresponding iodide proximal distribution functions in each part (the color notation is the same in both parts of the figure). Note that the sum of proximal distribution functions converges to one at 12 Å.

As a result, the proximal distribution function does not converge to one at an infinite distance from the defined group. Nevertheless, a sum of all calculated proximal distribution functions satisfies this criterion. A direct connection between proximal distribution functions and excess coordination numbers ($\Gamma$ values)[31] allows for direct comparison with experimental results.

### 2.1.3 Preferential binding coefficients

While looking at Fig. 2, one might argue that it cannot be a fair comparison, since the volumes of each spatial part differ in size. However, when we look below at the definition of the preferential binding coefficient ($\Gamma$), we find that the volume actually does not play a role. Starting from the Kirkwood-Buff theory[32], let's assume a ternary system, where the solvent (most often water) is denoted as 1, the biomolecule as 2, and the cosolvent (salt) as 3; then the preferential binding coefficient $\Gamma_{23}$ is defined by Eq. 3:

$$\Gamma_{23} = \rho_3 \left( G_{23} - G_{21} \right) \tag{3}$$

where $\rho_3$ is the salt concentration in the ternary mixture and $G_{ij}$ is the Kirkwood-Buff (KB) integral (Eq. 3), defined for open (grand canonical) systems. However, most of our simulations are performed in a canonical constant pressure - constant temperature (NPT) ensemble, therefore, the KB integrals have to be approximated[31] with the assumption that beyond a certain distance $R$ all the RDFs are converged to unity (Eq. 4).

$$G_{ij} = G_{ji} = 4\pi \int_0^\infty \left[ g_{ij}^{\mu VT}(r) - 1 \right] r^2 dr \approx 4\pi \int_0^R \left[ g_{ij}^{NPT}(r) - 1 \right] r^2 dr \tag{4}$$

In practice, we do not need to evaluate all the KB integrals. It suffices to directly count the number of water molecules ($N_{21}$) or the ions ($N_{23}$) in a distance $r$ from the biomolecule (center of mass or its surface), to be able to calculate preferential binding coefficients from Eq. 5.

$$\Gamma_{23}(r) = \langle N_{23}(r) - \rho_3 N_{21}(r) \rangle = \langle N_{23}(r) - \frac{N_3 - N_{23}(r)}{N_1 - N_{21}(r)} \times N_{21}(r) \rangle \tag{5}$$

$N_1$ and $N_3$ stand for the total number of water molecules and ions in our ternary system, respectively, and the notation $\langle \cdots \rangle$ represents the ensemble average. Now, we can apply this knowledge to evaluate the preferential binding coefficients from the proximal distribution functions. By integration of each of them, we obtain the number of ions and

water molecules at a distance $r$ from a defined *group* of our investigated molecule and by applying Eq. 6, we can assign the $\Gamma^{region}$ values for each spatial part:

$$\Gamma^{region}(r) = N_{proximal}^{ion-group}(r) - \frac{N_{total}^{ions} - N_{proximal}^{ion-group}(r)}{N_{total}^{water} - N_{proximal}^{water-group}(r)} \times N_{proximal}^{water-group}(r) \qquad (6)$$

We clearly see why the different sizes of the spatial parts do not influence the results: what we take into account is not the absolute number of bound ions but the relative number dependent on the number of water molecules in the same spatial section. Therefore, $\Gamma > 0$, when we observe preferential binding of ions and $\Gamma < 0$, when the preferential hydration occurs (depletion of ions). It must be emphasized here that the thermodynamic property we want to calculate and eventually compare with experiment is the limiting value at large distances (12 Å in our calculations), not the function $\Gamma(r)$ itself (see the example of iodide around the triglycine molecule in Fig. 3). The sums of $\Gamma$ values for cations and anions must be equal to each other to satisfy the electroneutrality condition. Therefore, if the target molecule is charged, we need to subtract the additional binding resulting from the extra charge $Z$ to get the preferential binding coefficient correctly (Eq. 7).

$$\Gamma_{23}(r) = \langle N_{23}(r) - \frac{N_3 - N_{23}(r)}{N_1 - N_{21}(r)} \times N_{21}(r) - Z \rangle \qquad (7)$$

This can also serve as an internal check of the chosen cut off distance and data convergence.
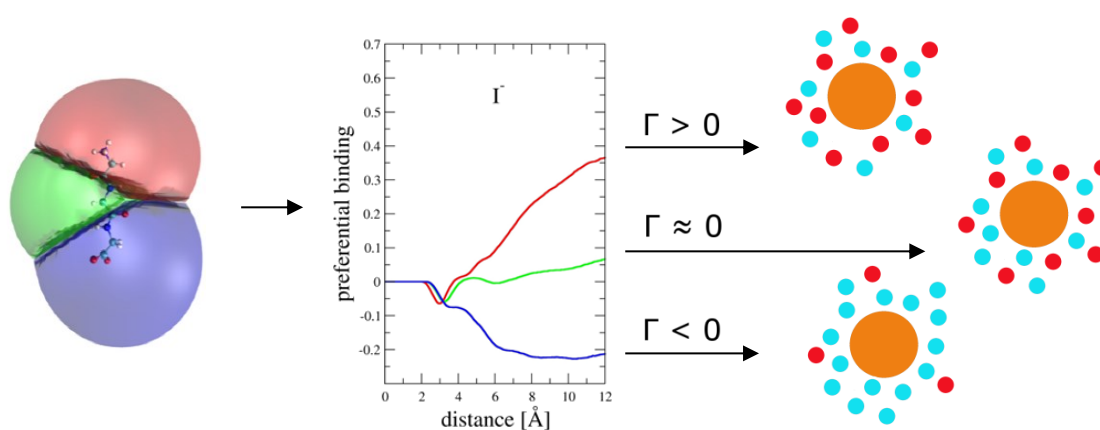


**Fig. 3**: Visualization of the spatial division around an uncapped glycine trimer and the corresponding iodide preferential binding coefficients for each spatial part (the color notation remains). Additionally, the meaning of the $\Gamma$ coefficient is explained by small cartoons showing the preferential binding of ions (in red) vs. the preferential hydration (water molecules in blue) of an investigated molecule (in orange).

### 2.1.4 Salting-out constants

Having the preferential binding coefficient Γ, we can now perform the last step and evaluate the salting-out constant $k_s$[33], i.e., the overall effect of the salt on the biomolecule (Eq. 8). The more positive $k_s$ value we calculate, the higher the salt exhibits a salting-out effect on the molecule, in other words, the less amount of salt we need for peptide precipitation from the aqueous phase. Contrary, the more negative $k_s$, the more salting-in behavior of the salt we observe.

$$k_s = -\frac{\Gamma_{salt}}{2.303 \times c_{salt}}$$
(8)

The experimental definition of salting-out coefficients (Eq. 9) is based on solubility measurements[17], where solubility of an investigated molecule is determined in pure water ($S_{water}$) and in a salt solution ($S_{salt}$) with concentration of $c_{salt}$.

$$k_s = \frac{\log(\frac{S_{water}}{S_{salt}})}{c_{salt}}$$
(9)

As a result, we can directly compare experimental and calculated $k_s$ values. This relation is valid on condition that the preferential binding coefficient Γ does not vary with the peptide concentration and an ideal behavior of the solution can be assumed. Therefore, we can get the most accurate results for molecules with low solubility, and often still good agreement for more soluble ones[34]. The additional value of the calculations lies in distinguishing between involved and not involved parts of the molecule in terms of ion binding, resulting in its overall molecular behavior in salt solutions.

### 2.1.5 LCST measurements

Another example of a thermodynamic experiment, where the overall effect of salt on the peptide structure is measured, is determination of the lower critical solution temperature (LCST). Some types of polypeptides (for example elastin-like polypeptide) undergo a reversible phase transition named "inverse temperature transition"[35,36]. Below the critical temperature, the polypeptide is highly soluble in water, while above this point the molecule experiences a hydrophobic collapse, forming a separate phase. Consequently, the originally transparent solution becomes turbid and, therefore, the critical point is called "clouding temperature". The LCST can be tuned by changing the composition of the

polypeptide or altered by varying the salts. The change of the clouding temperature as a function of salt concentration can then be monitored and linear or nonlinear dependence is obtained. The linear part of the curve can be subtracted and what remains is the ion binding isotherm (Eq. 10). From the whole curve the ion-polypeptide dissociation constant $K_D$ is calculated:

$$\Delta T = -c[M] + \frac{B_{max}[M]}{K_D + [M]}$$

(10)

where $c$ is a slope of the linear part (in units of temperature divided by molar concentration) and $[M]$ stands for the salt molar concentration. The binding isotherm consists of a fraction of $B_{max}$, which represents the maximum increase in the LCST (units of temperature) over the sum of the equilibrium dissociation constant, $K_D$ (units of concentration) and the salt molarity. It follows that the more nonlinear behavior we measure, the more ion accumulation on the peptide surface we can expect. For the purely linear dependences, we thus assume an infinitely weak binding interaction with the peptide or even ion depletion.

## 2.2 Molecular spectroscopy

In analogy to the LCST measurements, a peak shift $\Delta\delta$ in $^1$H NMR spectra can be monitored as a function of salt concentration. Also similarly, the curves are subsequently divided into linear and nonlinear contribution, as written in Eq. 11:

$$\Delta\delta = -c[M] + \frac{\Delta\delta_{max}[M]}{K_D + [M]}$$

(11)

In this case, the linear constant $c$ has different units than in Eq. 10 (i.e., chemical shift divided by molar concentration). $\Delta\delta_{max}$ represents the maximum change in chemical shift.

The main disadvantage of this method results from the fact that the relative chemical shifts lie at the edge of measurability (hundredth of ppm) which leads to high demands on the experimental equipment and precision of the measurement. The biggest advantage comes from the ability to distinguish between different types of protons in the structure. Therefore, it is possible to determine different dissociation constants for different parts of the molecule and suggest which of them are responsible for the ion binding. It should be stressed here, that this approach brings new insight into the Hofmeister topic, however, it cannot offer all the answers, since the exchangeable protons are problematic to measure and the determined effects are very local ($r^{-4}$ dependence)[37,38].

Infrared spectroscopy (IR) can be used in a similar fashion as described above. The changes in frequencies are measured and again, linear and nonlinear contributions are calculated by a fitting procedure. The local character of the measurement remains[39], but other functional groups than protons can be examined (carbonyl, carboxyl, etc.). The other advantage of the IR spectroscopy is that it allows us to monitor structural changes of the investigated molecules (for example, intramolecular vs. intermolecular hydrogen bonding).

Protein crystal geometries, obtained from the X-ray spectroscopy, serve as the first and key structures for most of the molecular dynamics simulations. The crystal is, however, a product of specific crystallization conditions, therefore, it does not have to truly match the situation in solvent. On top of that, it is well known that sodium cations can be misinterpreted as water molecules as they have similar electron densities. Nevertheless, the combination of all the above listed spectroscopic methods serves as a benchmark for theoretical calculations and leads to a better understanding of the nature of ion-protein and ion-peptide binding.

## 2.3 Electrophoresis

The last experiment we made use of in the context of this study, is electrophoresis, which is a chemical separation technique. By measuring the electro-osmotic flow (EOF)[40], we can establish net electrophoretic mobilites $(\Delta\mu^S)$ of charged solutes $(\mu^S)$ with respect to neutral EOF markers $(\mu^M)$[41] (Eq. 12).

$$\Delta\mu^S_{salt} = \mu^S_{salt} - \mu^M_{salt} = \frac{(t_M - t_S)l_d}{(t_{M2} - t_S)l_U} \cdot \frac{l_c}{U} \tag{12}$$

During the experiment, times when the solute $(t_S)$, marker $(t_M)$, and the second zone of marker $(t_{M2})$ pass the detector are measured at the distance $l_d$ applying the electric field U/l$_c$ for the time $t_U$. By employing this equation, we assume that all the EOF markers move exactly in the same way within the velocity of the EOF. However, the present ions of the background electrolyte can bind to the markers with different affinities and affect their mobilites[42,43], moreover, attraction to the capillary may occur[44]. Using molecular dynamics simulations, we attempted to quantify the ion-marker binding and asked the question, whether the electrophoretic technique could also be used for separation of neutral molecules.

For this purpose, we can calculate the electrophoretic mobility $\mu^S_{salt}$ from our simulations (Eq. 13), where we assume a spherical shape of a solute[45].

$$\mu_{salt}^{S}(c_{salt}) = \mu_{lim}^{S} + \frac{2}{3}\frac{e}{\eta} \sum_{i=X^{+},Y^{-}} z_{i}c_{salt} \int_{0}^{\infty} \left(g_{S}^{i}(r) - 1\right)r \, \mathrm{d}r \qquad (13)$$

We have to sum up the cationic ($X^{+}$) and anionic ($Y^{-}$) effects on the solute ($S$), which are estimated from the ion-solute radial distribution functions $g(r)$. $C_{salt}$ stands for the salt concentration, $z_i$ for the charge of individual ions, $e$ for the elementary charge, and η for the water viscosity coefficient. For neutral solutes, $\mu_{lim}^{S}$ , the mobility in neat water equals to zero, therefore, when we want to compare calculations with experimental data, we have to define the relative mobility of the solute as $\Delta\mu_{salt}^{S}(c_i) = \mu_{salt}^{S}(c_i) - \mu_{salt}^{M}(c_i)$.

At the end of this chapter, let us discuss important differences in the above described experimental approaches. While the NMR and IR spectroscopies sensitively report on the very local environment ($r^{-4}$ dependence) and cannot cover changes beyond the first solvation shell, electrophoresis belongs to methods with intermediate ($r$-dependent) interaction length-scale. In contrast, the thermodynamic regime pertinent to the clouding experiments, involves the long range ($r^{2}$-dependent) information. Therefore, the results of these methods may differ and be complementary to each other, since they give answers to questions relevant to different length scales.

## *Chapter 3*

# Ion – protein Interactions

This part of the results is dedicated to specific ionic effects on protein structures. Two types of enyzmes, the betaine-homocysteine *S*-methyltransferase (BHMT) and the haloalkane dehalogenase LinB, were investigated and the connection between ion binding and enzymatic activity was described. Additionally, a thorough investigation of the active site and side-products release from the LinB enzyme was performed. Presented data are primarily aimed at explaining experimental observations of the individual systems rather than deducing general conclusions concerning ion-specific effects.

## 3.1 Cationic activation of the BHMT enzyme

BHMT is an enzyme produced by liver and kidney tissues of monogastric animals[46]. It catalyses a methyl transfer from betaine to homocysteine (Hcy) at the active site, where a zinc ion is present. In this way, methionine is regenerated in the body serving as a critical component for the synthesis of *S*-adenosylmethionine, which is involved in hundreds of methylation reactions[47]. Previous research showed that Hcy binds before betaine approaches the active site, while after the reaction the newly-emerged methionine leaves the active site last[48,49]. The crystal structures of rat and human BHMT were solved, in the latter case with a substrate inhibitor (CBHcy), too[50,51,52].

Colleagues from our Institute performed a detailed experimental study on the influence of monovalent cations on the activity of BHMT[53]. Although sodium is known to bind most strongly to negatively charged residues[54], it is the potassium cation which enhances the activity significantly. Rubidium and ammonium cations influence the activity mildly, while the rest of the measured cations have no effect (Fig. 4).

Our goal was to investigate where the potassium ions bind to BMHT, how this can influence its enzymatic activity, as well as to explain why the regular cationic Hofmeister series[3] is not observed in this case. Experimental measurements showed that the $k_M$ value (with units of a substrate concentration) for Hcy decreases, while the $k_{cat}$ value (turn over number) does not vary significantly after adding the potassium salt. Therefore, we focused

on the potential ion binding site close to the active site where the ions could influence the substrate binding.
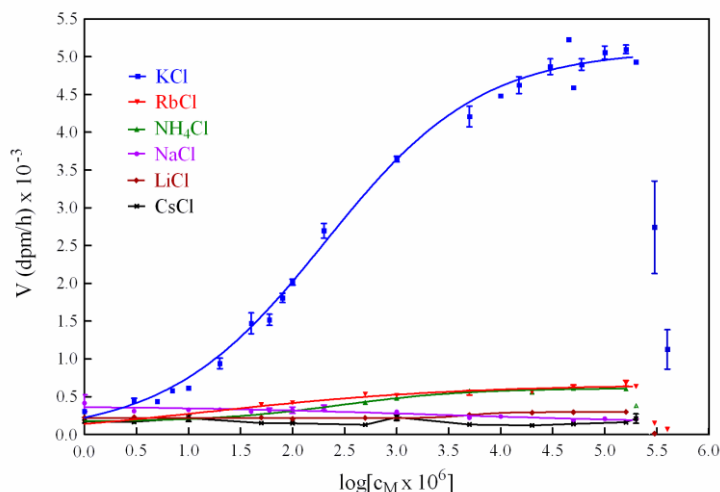


**Fig. 4**: The effect of monovalent cations on BHMT activity at 0-400 mM concentration of chloride salts. The experimental data points are averaged means of multiple measurements[53].

        Starting from the crystal structures of rat[50] and human[52] BHMT, we selected a single monomer assuming the ion binding would be similar in all units of the tetramer. From the collected data generated by classical MD simulations, we calculated the spatial distributions of sodium, potassium, and rubidium cations. We detected two regions with the highest probability of the cation occurrence at the active site, close to negatively charged residues Asp26 and Glu159 (see, as an example, the potassium spatial distribution function around BHMT in Fig. 5A). In addition, we calculated the radial distribution function for these cations with respect to the carboxylic carbon of Asp26 and by its integration we determined an average number of ions at the corresponding site (Fig. 6).

        The results suggest that due to the charge compensation one cation may be present at the active site. What differs is the position which the individual cations can occupy. The sodium cation keeps very strongly its solvation shell of six water molecules and, effectively, becomes the largest sphere from all and does not interact with carboxylic groups of this particular active site very efficiently. In contrast, the potassium cation easily loses two of its water molecules and binds strongly to the carboxylic groups of the active site oscillating between Asp26 and Glu159. Finally, the rubidium cation also has the ability to abandon two water molecules, however, being larger than potassium, the strength of the charge-charge interaction at the active site weakens.
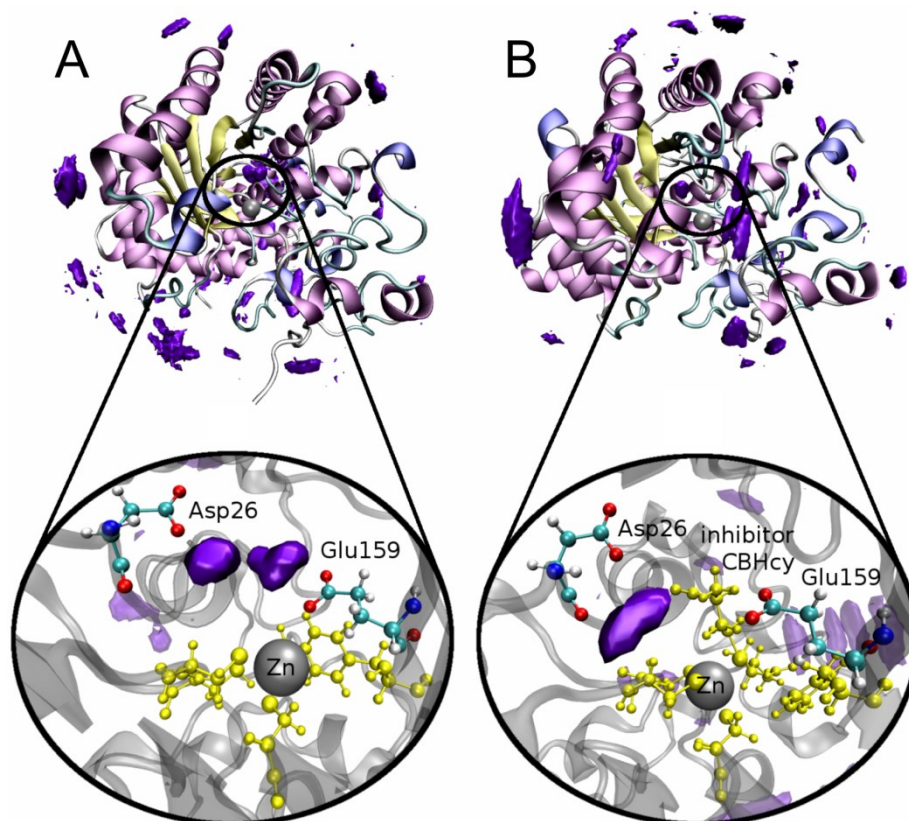
**Fig. 5**: Isosurface of the spatial distribution function of the potassium ion (in purple) calculated around the BHMT monomer with an isovalue=10. Part A shows the identified binding sites near the Asp26 and Glu159 residues. Part B shows how the binding site is shifted towards Asp26 when the inhibitor CBHcy is present at the active site. Zinc ion is coloured in grey and its ligand resudes (Cys247, Cys299, Cys300 and Tyr160) with inhibitor in yellow.
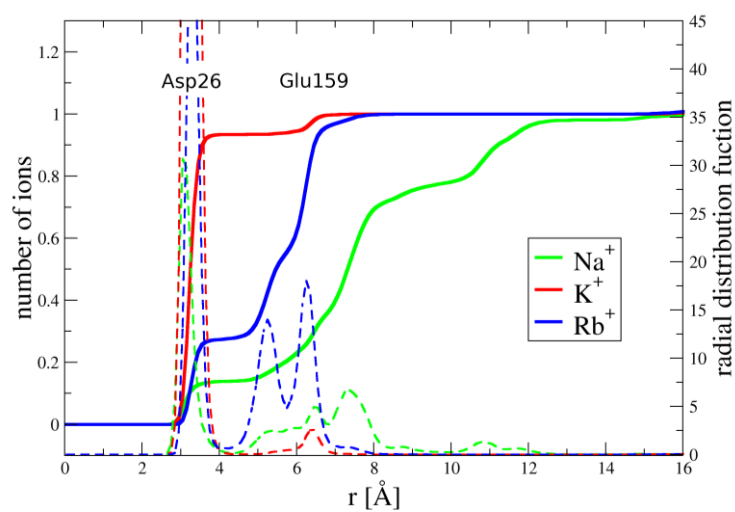


**Fig. 6**: Radial distribution functions (dashed lines) and average numbers of monovalent cations (full lines) occuring at the active site of BHMT.

To find out whether and how the identified potassium binding sites change when the substrate approaches the active site, we performed additional MD simulations with the bound CBHcy inhibitor. In Fig. 5B we can notice that the potassium ion preferably comes closer to the Asp26 and oscillate between the carboxyl and the carbonyl group of this residue. The original binding site in the vicinity of Glu159 is now sterically hindered by the inhibitor and slight residual rearrangements arise. A detailed picture of the described potassium binding based on the MD simulations is provided in Fig. 7A.

Independently of our research, a collaborating experimental group succeeded in crystallization and determination of the tetrameric structure of human BHMT in a complex with the *L*-Hcy ligand. They conclusively confirmed the $K^+$ binding site (Fig. 7B) predicted by our MD simulations (Fig. 7A) with the potassium coordination number of 8. One carboxylate (Asp26), three carbonyl oxygens (Asp26, Gly27, and Gly298), two amide oxygens (Gln72, Gln247) and two water molecules interact simultaneously with the potassium cation. On top of that, the water molecule attached to the $K^+$ ion mediates additional stabilization of the Hcy ligand as well as the CBHcy inhibitor. Both substrate and inhibitor were detected in similar positions bound to the zinc ion and are additionally stabilized by interactions with Glu159 (amino group), Phe29 and Val30 (via carboxy groups).
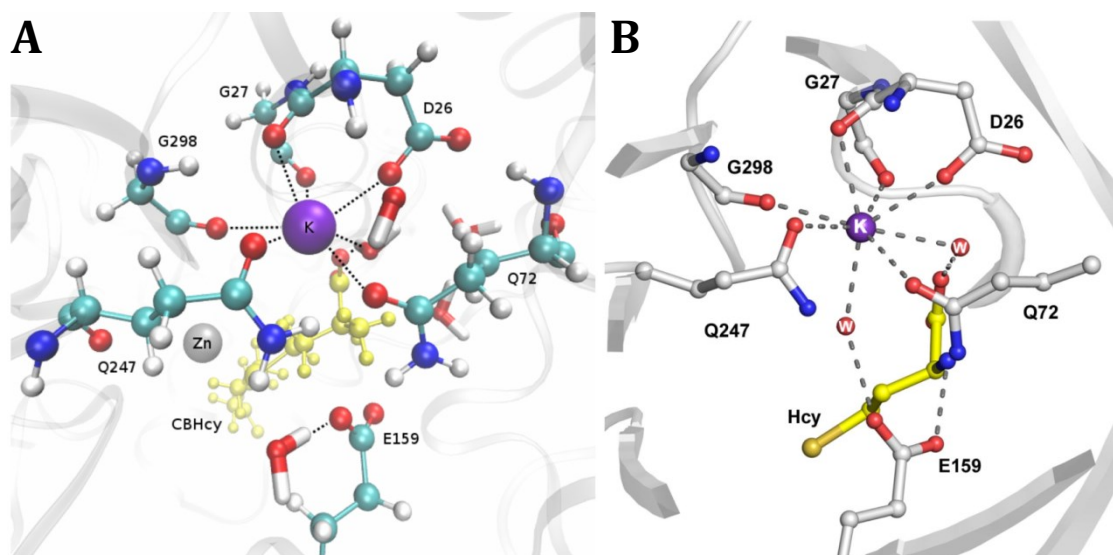


**Fig. 7**: Panel **A** provides a detailed picture of the modeled potassium binding at the active site of the BHMT monomer with the CBHcy inhibitor. Panel **B** shows a cartoon representation of the X-ray electron density map for the $K^+$ ion in BHMT including the *L*-Hcy ligand. The potassium ion is depicted in purple, while the substrate and the inhibitor in yellow. Residues in the closest vicinity to $K^+$ are shown in balls and sticks.

Finally, BHMT mutants were prepared to quantify the importance of individual residues responsible for potassium binding. All mutants, in which one of the residues

directly bound to K$^+$ (Gly27, Glu159 and Asp26) was altered, showed much lower activity than the wild type. On the other hand, the mutation of Gly28, which is a member of the conserved fingerprint sequence Asp26-Gly27-Gly28 (DGG) not directly involved in potassium binding, influenced the activity only mildly. The mutants of Gly27 and Glu159 were completely inactive, however, the very low yields of these mutants point out that these residues may be crucial for protein folding[53].

## 3.2 Cation specific effects on LinB

The LinB microbial enzyme belongs to the family of haloalkane dehalogenases and catalyses, in the absence of any cofactor, the conversion of alkylhalides to alcohols, releasing halide anions and protons as side products[55] (Fig. 8).
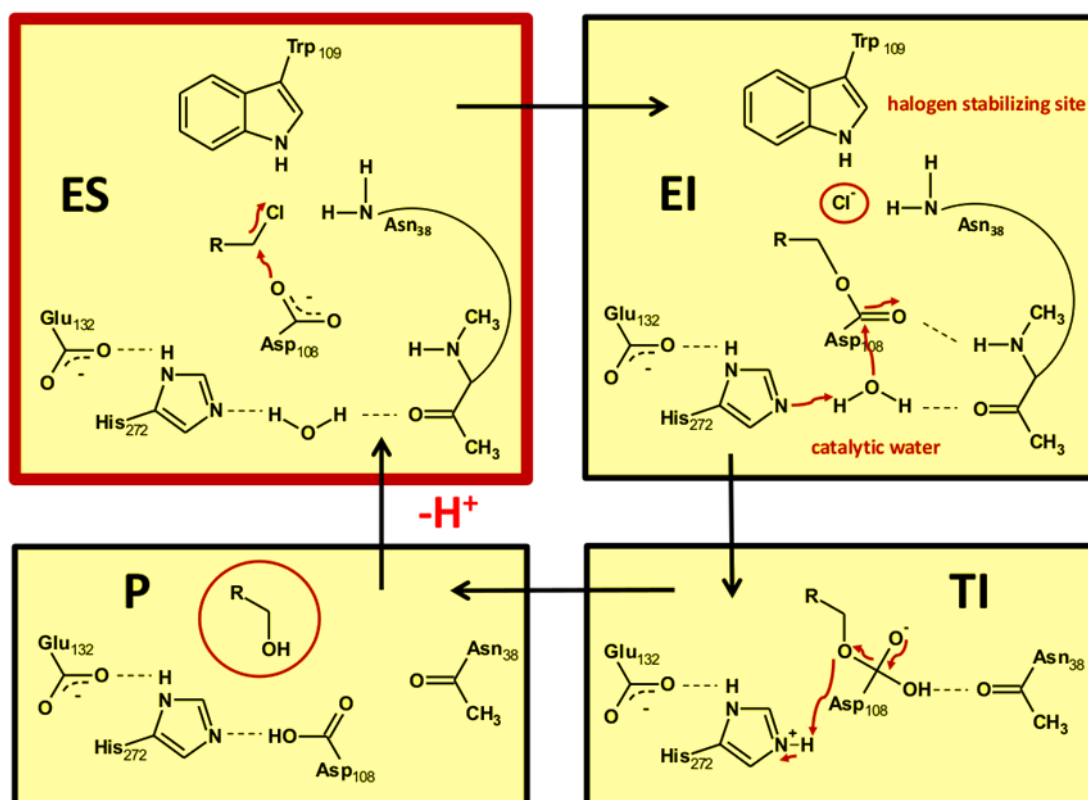


**Fig. 8**: Scheme of the catalytic reaction of the LinB enzyme. ES stands for the enzyme-substrate complex, EI and TI for the intermediates, and P for the product.

The active site of the enzyme is located at the bottom of a narrow tunnel and consists of a catalytic triade and two halogen-stabilizing residues. The catalytic triade is formed by a nucleophile (Asp108), a general base (His272), and a catalytic acid (Glu132). The halogen-stabilizing residues (Trp109 and Asn38) form a binding site for the halide anion produced during the enzymatic cycle[56,57].

Similarly to the BHMT enzyme, our goal was to investigate how the monovalent cations interact with the protein, and how this binding can influence the enzymatic activity. We performed classical MD simulations in 0.5 M salt solutions of NaCl, KCl, RbCl, and CsCl. The uncertainty about the protonation state at the active site forced us to run two sets of simulations with and without an extra proton on the catalytic histidine. It turned out that the trends were the same for both cases and, therefore, the bellow data are presented only for the protonated one[58].

From the spatial distribution functions of the alkali ions (Fig. 9A), we notice that sodium interacts the strongest, with all the cations following the Hofmeister ordering. Not surprisingly, most of the binding "clouds" occur at the vicinity of negatively charged residues (Asp and Glu) and carbonyl groups of the backbone, with that the overall charge of the protein is -10 e. Interestingly, none of the binding sites differs with the change of the cation.
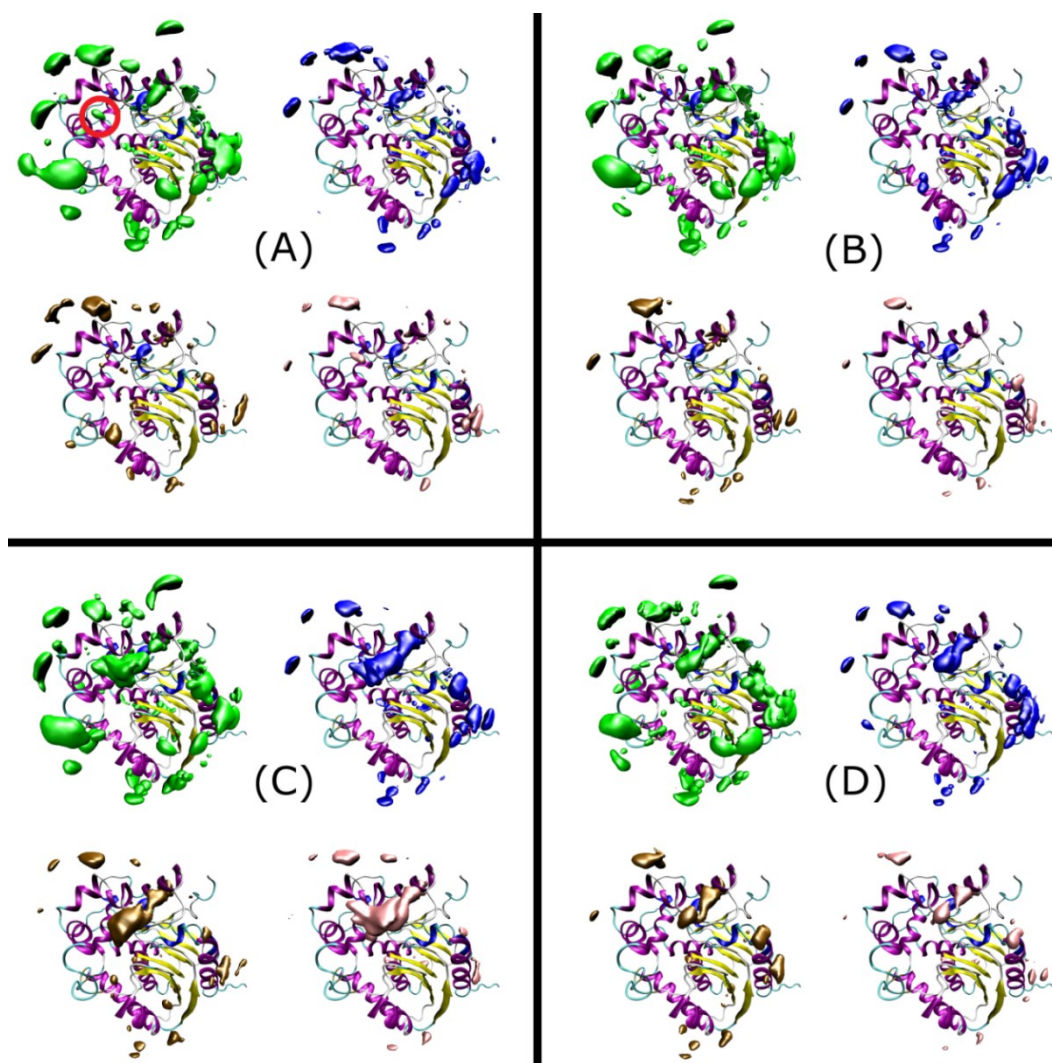


**Fig. 9**: The spatial distribution functions of sodium (green), potassium (blue), rubidium (brown), and cesium (pink) cations around the LinB enzyme are shown with a binding hotspot highlighted with a red circle. Part A represents the result for LinB wt, parts B-D for the LinB mutants D147N, L177D, and D147N-L177D, respectively.

This qualitative observation was then complemented by quantitative evaluation of the radial distribution functions and the ion coordination numbers. It follows from this analysis that the bigger the cation, the weaker binding we observe. The difference between sodium and potassium is significant, while only a modest change is observed between potassium and rubidium ions. Cesium exhibits the weakest affinity among of the investigated monovalent cations.

Although we detected several binding sites at the LinB wild type (wt) surface, the previous study of this enzyme suggested that the most important site for influencing the enzymatic activity should be at the tunnel mouth[56]. From the crystal structures, there was no evidence of cations entering into the tunnel and approaching the active site and we did not observe it in our MD simulations either. Therefore, several areas close to the tunnel mouth were monitored with the hotspot region located right above it (red circle in Fig. 9A). The radial distribution functions of the different cations with respect to the amino acids forming the hotspot were evaluated (see Fig. 10A) and the regular Hofmeister ordering was observed.
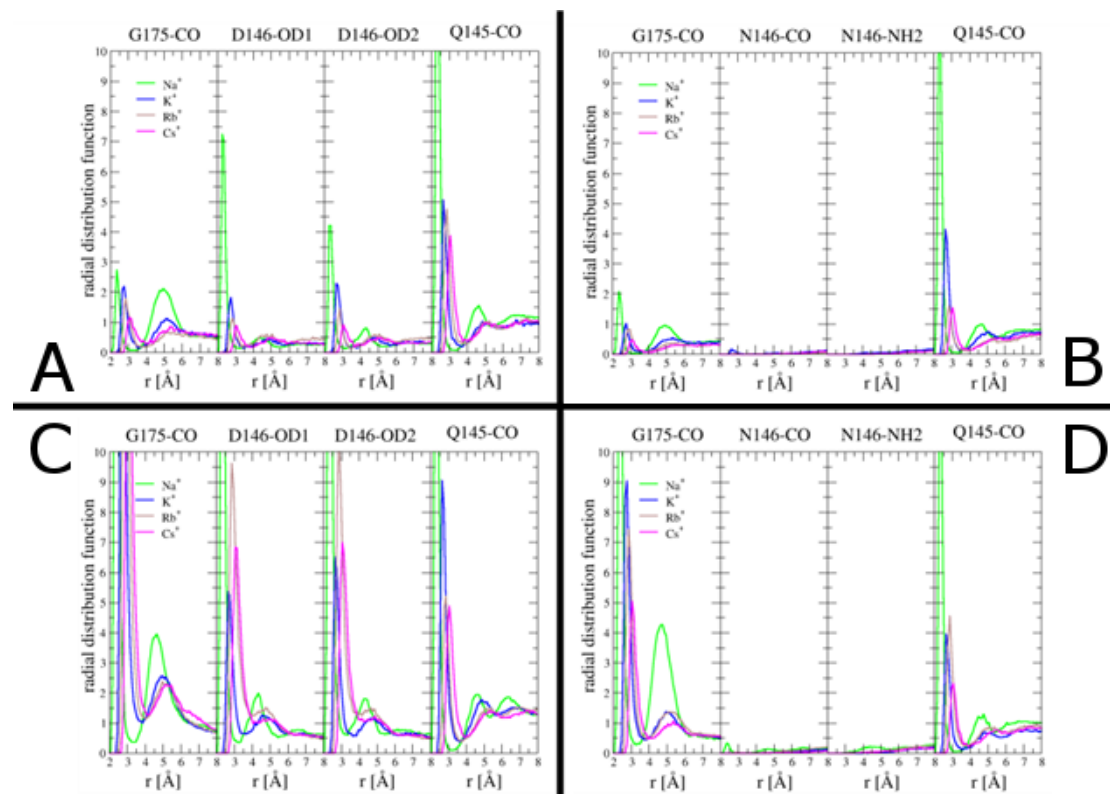


**Fig. 10**: The radial distribution functions (RDFS) of different cations at the hotspot region, formed by G176, D147, and Q146 residues. Part A represents the result for LinB wt, while parts B-D for the LinB mutants D147N, L177D, and D147N-L177D, respectively. When the mutation L177D was present, we added the RDFS to the new carboxyl group at the tunnel mouth.

To properly examine the nature of ion binding at the hotspot region, two single point mutants (D147N and L177D) and a double point mutant (D147N-L177D) of the LinB enzyme were simulated in order to see the influence of the local charge. The overall spatial (Fig. 9B-D) and radial distribution functions for LinB mutants differ very little from the wild type. In contrast, the local spatial (Fig. 9B-D) and radial (10B-D) distribution functions with respect to the hotspot region respond significantly to the charge variations. In Fig. 9B, we see that the reduced charge at the tunnel mouth (D147N) caused a disappearance of the corresponding cationic cloud (circled position in Fig. 9A). Inversely, the additional negative charge introduced by the LinB L177D mutation expanded the cationic clouds at the very same position (Fig. 9C), however, the cationic specificity vanished. The most interesting case, when the double mutant was examined, shows that despite the same resulting charge, the mutation D147 seems to be more important one and is not fully compensated by the L177D mutation (Fig. 9D). When we compare the same results in terms of the radial distribution functions (Fig. 10A-D), we get similar answers. Interestingly, for the double mutant (D147N-L177D) we find the effect to be very local. At larger distances the two mutations compensate each other and the RDFs to Q145 residue matches the ones of the wild type.

In order to compare our results with experiment, the single point mutants were expressed and the steady-state kinetics for them and the wild type was measured in aqueous solutions containing 0.5 M of NaCl, KCl, RbCl, or CsCl. Unfortunately, the double mutant couldn't be analyzed due to the low yield of its expression. The results revealed that a point charge mutation can influence the enzyme kinetics. The relative reaction rates of the investigated enzymes were obtained (Fig. 11). Based on substrate inhibition in a glycine buffer, the enzymes were ordered as LinB D147N > LinB wt > LinB L177D, which corresponds to the rate of cationic binding in our MD simulations. On the other hand, cationic specificity on individual enzymes was weak and uncertain due to the error of the measurements.
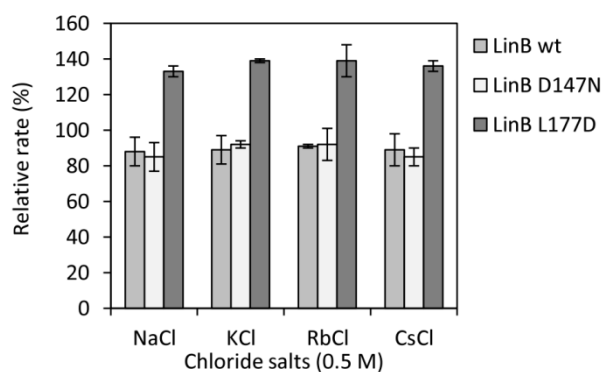


**Fig. 11**: The relative reaction rates of LinB wt, LinB D147N and LinB L177D expressed as a percentage of the reaction rate in glycine.

This is most likely because the monovalent cations do not penetrate the LinB tunnel leading to the active site and, therefore, are not able to directly influence either substrate binding or residues involved in the enzymatic reaction.

Interestingly, the enzymatic reaction did not follow a simple Michaelis-Menten kinetics and the velocity data had to be fitted by a more complex kinetic scheme, where two or three substrate molecules were employed. The fitted data implied that the binding of the third substrate molecule creating the ESSS complex is specifically affected by the cations occupying the tunnel mouth region. In agreement with cationic affinities to the tunnel mouth, formation and productivity of ESSS complex was inhibited strongly by sodium and potassium while rubidium and cesium had only a weak effect.

## 3.3 Release of side-products from the LinB active site

The general observation that cations do not penetrate the LinB tunnel raises a question, how large the barrier is and how it can be influenced by varying the charge on the catalytic histidine (His272). We used the umbrella sampling method[60] to calculate the free energy ($\Delta$G) profile of Na$^+$ pulled from the bulk through the tunnel to the LinB active site. The whole trajectory of the cation (26 Å) was divided into 52 parts (windows). A Na$^+$-pulling origin distance was sampled for 12 ns with a constant force employed to keep the ion in each window. Using the WHAM[59] procedure, data were processed and free energy profiles of Na$^+$ penetrating the tunnel with or without protonated His272 were obtained. Within this method, a non-Cartesian reaction coordinate is used, therefore, a volume entropy correction needs to be included. A general r$^2$ correction, valid for a radial coordinate in a bulk solution, cannot be applied here. The accessible volume for the ion receding from the pulling origin does not increase with the square of the distance from the origin, but it is the size of the tunnel which determines it. Therefore, we evaluated the actual normalized accessible volume *(V$_{accessible}$)* for the ion in each window and manually corrected the results using the following formula (Eq. 14):

$$G^{REAL} = G^{WHAM} + kT \times \frac{\ln(V_{accessible})}{4185} \tag{14}$$

The resulting free energy barriers changed significantly with the protonation state of the catalytic histidine. When the active site comprises of neutral His272, the barrier for Na$^+$ to achieve the active site is relatively small, ~3.5 kcal/mol. In contrast, the doubly protonated His272 increases the barrier to 14 kcal/mol. Taking into account that none of the

crystal structures show $Na^+$ at the LinB active site, we concluded that it is more probable for the proton to stay at the active site after the enzymatic reaction. However, a sodium cation is often misinterpreted as a water molecule in X-ray structures, therefore, additional investigation is requested.

At the active site in the vicinity of the catalytic triade, there are two halide-stabilizing residues (Trp109 and Asn38). During the enzymatic cycle, a halide ion is produced and immediately localized there (see Fig. 8: EI). It is known from the X-ray structures that the halide anion is released from the active site only after the main product, i. e., the alcohol. Therefore, the protonation state of the catalytic histidine may be influential in the halide release. A study on the $Br^-$ release from the LinB active site combining both theory and experiment was published[60]. However, it suffered from a significant mismatch between the MD predicted and the experimental results for the wild type enzyme. On top of that, only one protonation state of His272 was considered in this publication.

In collaboration with the same experimental group, our goal was to find out whether the releasing paths differ for different anions ($Cl^-$, $Br^-$, and $I^-$) and order the anions according to their free energies of binding in the LinB active site. The protonation effect of the catalytic triade on the results was also investigated. Stopped-flow fluorescence experiments[61] were conducted to establish the energetic difference ($\Delta\Delta G$) of a halide to occur in the bulk solution or to be bound at the active site between the halide-stablizing residues.The intrinsic fluorescence of Trp109 was used to obtain the $\Delta\Delta G$ values for all investigated halides ($Cl^-$, $Br^-$, and $I^-$) (Tab. 1).

| | calculated ΔΔG (neutral His272) (kcal/mol) | experimental ΔΔG (kcal/mol) | calculated ΔΔG (charged His272) (kcal/mol) |
|---|---|---|---|
| $Cl^-$ | -0.63 | -0.74 | +1.00 |
| $Br^-$ | -1.95 | -0.87 | -0.46 |
| $I^-$ | -2.07 | -1.71 | -0.49 |
| $Na^+$ | +1.70 | - | +15.00 |

**Tab. 1**: Energetic differences ΔΔG between a bound halide at the LinB active site and a free halide in the bulk solution for $Cl^-$, $Br^-$, and $I^-$ are shown as a comparison between experimental and theoretical results. Two protonation states of His272 were considered.

We employed the same methods as described above for $Na^+$ to calculate free energy profiles of a halide release from the active site for all three halide ions ($Cl^-$, $Br^-$, and $I^-$). Two possible protonation states of His272 were considered and two sets of profiles were thus calculated. We hoped that a direct comparison with the experimental ΔΔG values would

help us to identify the relevant protonation state of His272 (Tab. 1). Moreover, using this procedure, a detailed picture of the halide release is obtained, including the energetic barriers or identifying places of halide stabilization/destabilization in the tunnel (Fig. 12).
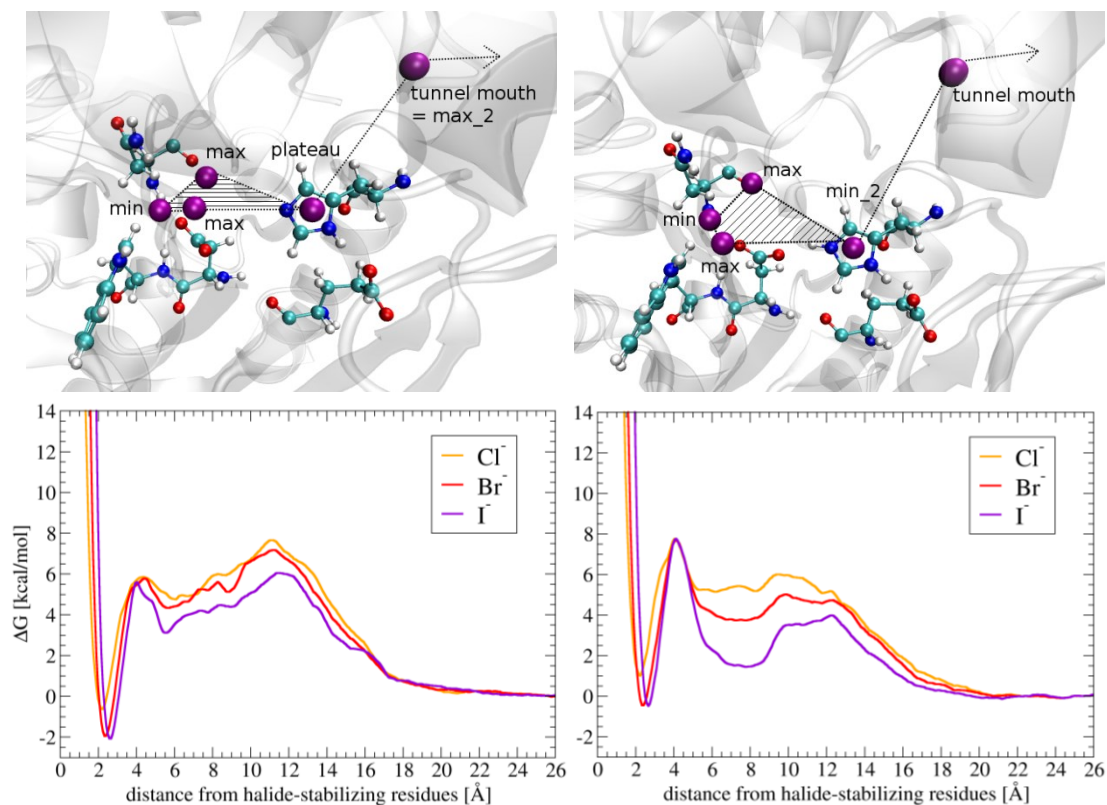


**Fig. 12**: Free energy profiles of the Cl⁻ (orange), Br⁻ (red) and I⁻ (violet) release from the LinB active site to the bulk. The left panel belongs to the case of neutral catalytic His272, while the right panel corresponds to the case of charged His272. A path of I⁻ from the halide-stabilizing site to the tunnel mouth is drawn for both cases above the free energy profiles.

We found that all the anions are released through the same LinB tunnel. On top of that, the shapes of the free energy profiles do not differ qualitatively for individual anions. Independently on the protonation state of His272, we observe the ion stabilization by the halide-stabilizing residues in the following order: I⁻ > Br⁻ >> Cl⁻ (see Fig. 12 and Tab. 1). The experimental ΔΔG values are in agreement with our results within the error of the umbrella sampling method (~1 kcal/mol) and reveal the following ordering: I⁻ >> Br⁻ > Cl⁻. Unfortunately, the differences between ΔΔG values in the two cases (protonated and neutral His272) are too small, therefore, we cannot make a definite conclusion about the protonation state from the data. Moreover, the experimental ΔΔG values lie in between the two theoretical predictions.

While the relative ΔΔG values are quite similar, the shapes of the ΔG profiles differ with the protonation state of His272 (Fig. 12). When the catalytic histidine is neutral, the

curves exhibit only one minimum at the halide-stabilizing site and a plateau region in the tunnel with a maximum at the tunnel mouth. By contrast, when His272 is doubly protonated, additional stabilization of the anions occurs nearby this residue (a second minimum in Fig. 12, right pannel). In this case the strength of the interaction orders the anions as $I^-$ >> $Br^-$ >> $Cl^-$. While a barrier to penetrate/leave the tunnel is decreased when His272 is doubly protonated, the highest barrier of the free energy profiles remains the same in both cases (~8 kcal/mol). It should be stressed here that the His272 deprotonation changes the arrangement of the catalytic triade (see Fig. 12) and influences the hydration of the active site, too.

To systematically examine this effect and determine the most probable time of the proton release from the LinB active site, we performed a series of MD and QM/MM calculations. The proton transfers from the catalytic His272 or Asp108 to a water molecule creating $H_3O^+$ were performed for four scenarios. First, to mimic the situation right after the enzymatic reaction, both products (butanol and chloride ion) were present at the active site. In the second case, only the chloride anion was present in between the halide-stabilizing residues. The third option stands for the purely hydrated active site without additional atoms/molecules there. To mimic the situation right before the reaction, the substrate (*n*-butyl chloride) was located at the active site as a fourth possibility. The QM part consisted of the catalytic triade (Asp108, His272, and Glu132) and 4-5 closest water molecules plus the substrate/product(s) if present. The DFT method with BP86 and B3LYP functionals was used in combination with def2SVP and def2TZVP basis sets.

The preliminary results show that when the substrate approaches the active site, the hydrophobicity of the area increases as the water channel leading to the active site is disconnected. This causes a disadvantage in formation of $H_3O^+$. Consequently, we found local minima only with the proton bound to the LinB residues (Asp108, His272). Virtually the same results were obtained when both products occupied the active site at the same time. In contrast, when there is no additional molecule at the active site, the area is perfectly hydrated and a direct water chain through the tunnel is detected. In spite of that, we did not identify a local minimum for $H_3O^+$ in this area. Using proton transfer scans from His272 or Asp108 to the nearest water molecule, no local minimum was detected. When we started from the MD snapshot with equilibrated $H_3O^+$ molecule at the active site, the proton moved back to one of the LinB residues during the minimization procedure of the QM/MM approach.

The only case, when we found local minima for $H_3O^+$ close to the catalytic site, was the one with chloride anion at the halide-stabilizing residues. The active site is then

perfectly hydrated and, on top of that, the negatively charged chloride helps to stabilize the hydronium cation after the proton transfer from His272 or Asp108. Additional check of a base and method/functional dependence is being realized now, as well as attempts to find the corresponding transition states of the proton transfer and perform evaluation of the barriers. The work is still in progress and the final results will be published in due time. So far, we can conclude that most likely the proton is released from the active site when halide anion is still present at the site and no other molecule occupies it. However, if it is transferred via the water channel directly to the bulk or whether it moves as an ion pair of $H_3O^+$ - $Cl^-$ is still an open question.

*Chapter 4*

# Ion – peptide Interactions

In contrast to the previous chapter, which focused on interactions of ions at specialized active sites of enzymes, we present here a systematic study on the binding patterns of the Hofmeister ions at the peptide surfaces in order to identify general trends and explain the nature of these interactions. To achieve this goal, we explore mainly model peptide systems of different sizes and compositions, to be able to distinguish individual features of the complex matter of specific ion interactions.

## 4.1 Neutral binding sites of elastin-like polypeptide

The largest system, we had experimental data available from our American colleagues, is a biopolymer $(VPGVG)_{120}$, known as the elastin-like polypeptide (ELP). As described in the methods section of this thesis, the molecule undergoes an inverse phase transition as a function of temperature and, therefore, the lower critical solution temperature (LCST) can be measured. The LCST is known to be decreased by sodium salts in the following anionic order[62]:

$$SO_4^{2-} > S_2O_3^- > F^- > Cl^- > Br^- > NO_3^- > I^- > ClO_4^- > SCN^-$$

Our colleagues performed the LCST measurements as a function of salt concentration for three sodium salts – $Na_2SO_4$, NaCl, and NaSCN. Results confirm that $Na_2SO_4$ decreases the LCST significantly, while NaCl only mildly, and NaSCN even increases it. On top of that, a nonlinear concentration dependence of the LCST can be observed for NaSCN, in contrast to the linear behavior of $Na_2SO_4$ and NaCl. The assumption that the nonlinear part can be matched with a binding isotherm leads to the conclusion that specific binding sites for thiocyanate exist at the peptide surface[63].

Additionally, proton NMR experiments were conducted below the LCST with the same three sodium salts. A change of the individual proton chemical shifts as a function of salt concentration was monitored (Fig. 13).
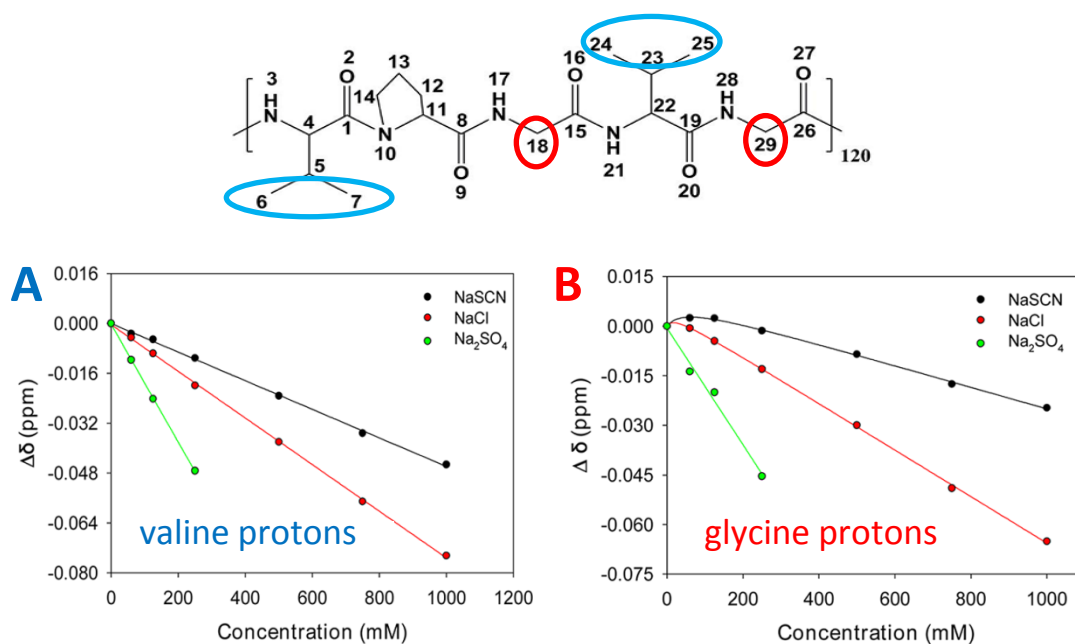
**Fig. 13:** Chemical shift change as a function of salt concentration for valine protons in part A and glycine protons in part B. The corresponding protons are highlighted in the molecular formula depicted above the graphs.

The protons of the valine side chains (in blue circles) display only linear changes of their chemical shifts for all three salts (Fig. 13A). By contrast, the character of the changes of the NMR shift for the $C_\alpha$-protons of the glycine residues (in red circles) is highly dependent on added salt (Fig. 13B). While only a linear behavior is observed in case of $Na_2SO_4$, a notably nonlinear dependence is recorded for NaSCN and, to a much lesser extent, also for NaCl. Peptide bond protons show very similar dependences as glycine $C_\alpha$-protons. Proline ring protons shifts are nonlinearly influenced only in the NaSCN solution, but the calculated dissociation constant $K_D$ is much higher (~270 mM) compared to $K_D$ of $C_\alpha$-protons (50 mM).

To probe molecular mechanisms of the ion binding to the ELP surface, we performed classical MD simulations of the VPGVG monomer in the same salt solutions ($Na_2SO_4$, NaCl, and NaSCN). To avoid artificial binding resulting from the charged ends of the monomer, we acetylated the N-terminus and methylated the C-terminus of the molecule. It is known experimentally that the ELP structure is unfolded and extended below the LCST[64]. Therefore, we picked one of the observed extended geometries and ran additional simulations with such a constrained geometry of the monomer. In this way, we aimed at mimicking better the experimental conditions and all the presented results are thus obtained from the constrained 100 ns runs. Nevertheless, when we compare the ion distributions of the constrained and the unconstrained pentapeptide, there are no

significant differences. The simulation box consisted of a single VPGVG monomer and a 1 M aqueous solution of NaSCN or NaCl for which polarizable force fields were used in combination with POL3 water model. In order to maintain the same ionic strength for the third system, we calculated the ion distribution around VPGVG in a 0.33 M solution of $Na_2SO_4$. The problem of an artificially overestimated cation-anion binding was solved using scaled charges for sodium and sulfate ions[65].

The interaction regions of the VPGVG monomer can be divided into several groups according to the NMR response. In Fig. 14, we show how we segmented the space into four parts – peptide bonds, nonlinearly and linearly dependent regions (see the above discussion of NMR results), and terminal groups (acetyl~ and ~N-methyl amide). As the capping groups are irrelevant for comparison with the experiment, we exclude them from our further analysis.
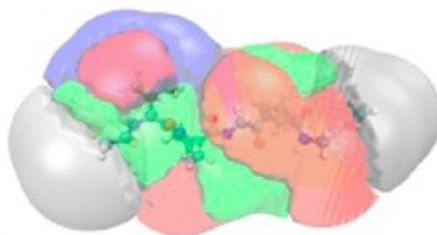


**Fig. 14**: Interaction space division of the VPGVG monomer surface according to the NMR response – peptide bonds in green, nonlinearly dependent groups in red, linearly dependent groups in blue, and terminal groups in grey.

The ion-peptide interactions were analyzed in terms of proximal distribution functions of $SCN^-$, $Cl^-$, and $SO_4^{2-}$ as well as their spatial distribution. From the ion distribution maps (Fig. 15A), we see that $SCN^-$ occupies more regions around the peptide and in a more non-uniform way than $Cl^-$ and $SO_4^{2-}$. The proximal distributions functions can quantify this observation (Fig. 15B-D) and we can conclude that $SCN^-$ interacts mainly with the NH group of the peptide bonds and with the methylene groups of the glycine residues, while the interaction of $SCN^-$ with the proline ring only marginally contributes to the curvature of nonlinear regions. In contrast to $SCN^-$, all the interactions are significantly reduced for $Cl^-$ and even more for $SO_4^{2-}$.

Probably the best way how to rationalize the experimental results is to evaluate ion preferential binding coefficients $\Gamma$ with respect to water[7]. From the definition of this quantity (Chapter 2: Eq. 6), the curve should converge to a limiting value at large distances (12Å in our case). Positive $\Gamma$ values then signify ion attraction and negative $\Gamma$ values repulsion
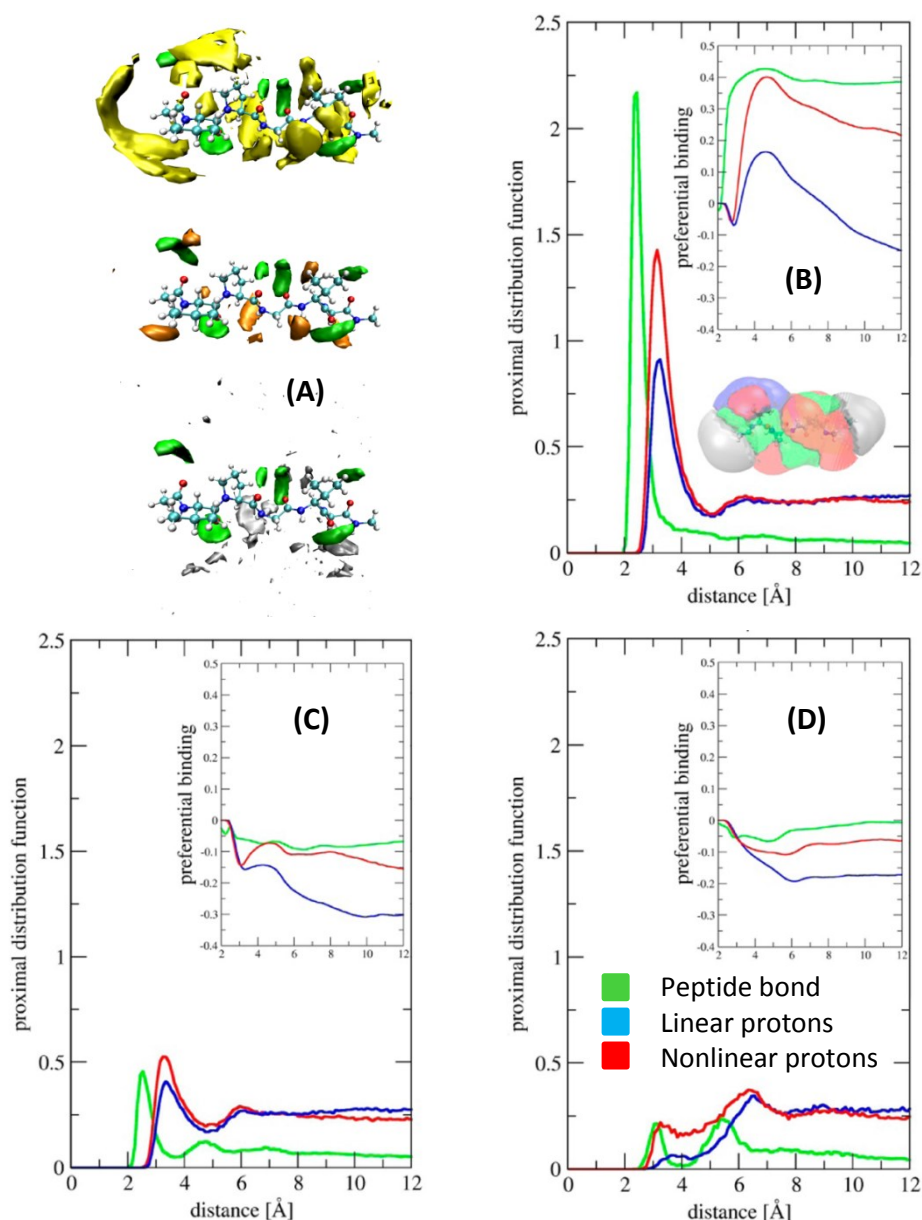
**Fig. 15**: Spatial distribution functions of SCN$^-$ (in yellow), Cl$^-$ (in orange), SO$_4^{2-}$ (in grey), and counter ion Na$^+$ (in green) in panel A. Proximal distribution functions of the same anions are shown in panels B-D with respect to the peptide bond (in green), nonlinearly responding regions (in red), and linearly dependent regions (in blue). Insets in panel B-D show preferential binding coefficients $\Gamma$ of the three anions with respect to the same three spatial groups in the same color notation.

from the peptide surface. When we closely inspect the results for SCN$^-$ (inset of Fig. 15B), we conclude that the most probable binding sites for this anion are the NH groups and the adjacent methylene groups of the peptide backbone, as both corresponding $\Gamma$ values are positive. In contrast, SCN$^-$ is on average repelled from the valine side chains. The preferential binding coefficients of Cl$^-$ and SO$_4^{2-}$ indicate ion repulsion from all the spatial parts of the molecule, notably, most strongly from the valine residues (insets of Fig. 15C-D).

Both experimental and MD results demonstrate that the weakly hydrated SCN⁻ ions strongly interact with the peptide backbone, while the neutral side chains do not play a significant role in salting–in behavior of NaSCN. By contrast, strongly hydrated anions are depleted from the neutral peptide surface, which consequently leads to the hydrophobic collapse of the polypeptide (salting-out effect).

## 4.2 Backbone vs. charged side chain interactions

Building on our gained knowledge, we chose a representative system of anion-backbone interactions only. N-methylacetamide (NMA) serves as the simplest and widely studied model of a peptide bond[17,66,67], however, it does not contain any alpha carbon in its structure. Therefore, we focus our research on the more representative triglycine (GGG) molecule, where both terminal groups are capped. In this way, the backbone model does not contain additional binding sites for the ions.

Similarly to our previous study[63], we combined NMR experiments with MD simulations to verify the general conclusion that anion-backbone interactions follow the Hofmeister series. We investigated a fully flexible capped triglycine molecule in five different sodium salt solutions (NaSCN, NaI, NaBr, NaCl, and $Na_2SO_4$)[68]. As there are three NMR distinguishable methylene protons in the GGG structure, we divided the surrounding space into the three corresponding parts. We applied the same MD data analysis and, as a result, spatial and proximal distribution functions with $\Gamma$ values were obtained. The left half of Fig. 16 shows the spatial distribution functions of all studied anions around the terminated (capped) GGG tripeptide.

We see that the more weakly the hydrated anion is, the more pronounced is its interaction with the backbone, which follows a direct Hofmeister series: $SO_4^{2-} < Cl^- < Br^- < I^- <$ $SCN^-$. From the corresponding preferential binding coefficients (an example for sulfate is shown in Fig. 17), we draw the conclusion that sulfate exhibits no attraction to the backbone binding sites (i.e., NH and $CH_2$ groups); as a matter of fact, it is repelled from the capped triglycine surface. Chloride and bromide behave as Hofmeister-neutral ions with no net preferential binding. Iodide becomes the first anion of the series, which is considerably attracted to the backbone and, finally, thiocyanate exhibits the strongest affinity and acts as a strong salting-in agent.
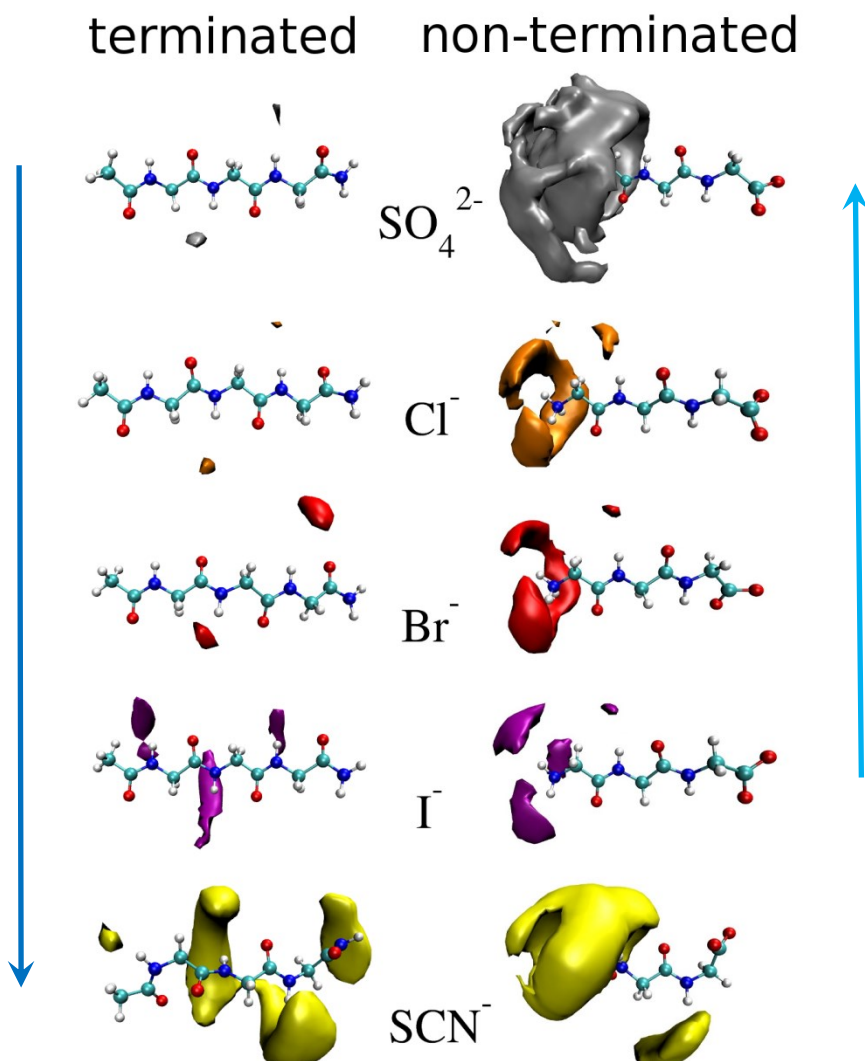
**Fig. 16**: Density maps of $SO_4^{2-}$, Cl⁻, Br-, I⁻, and SCN⁻ around terminated (left column) and non-terminated (right column) triglycine at an isovalue of 3.
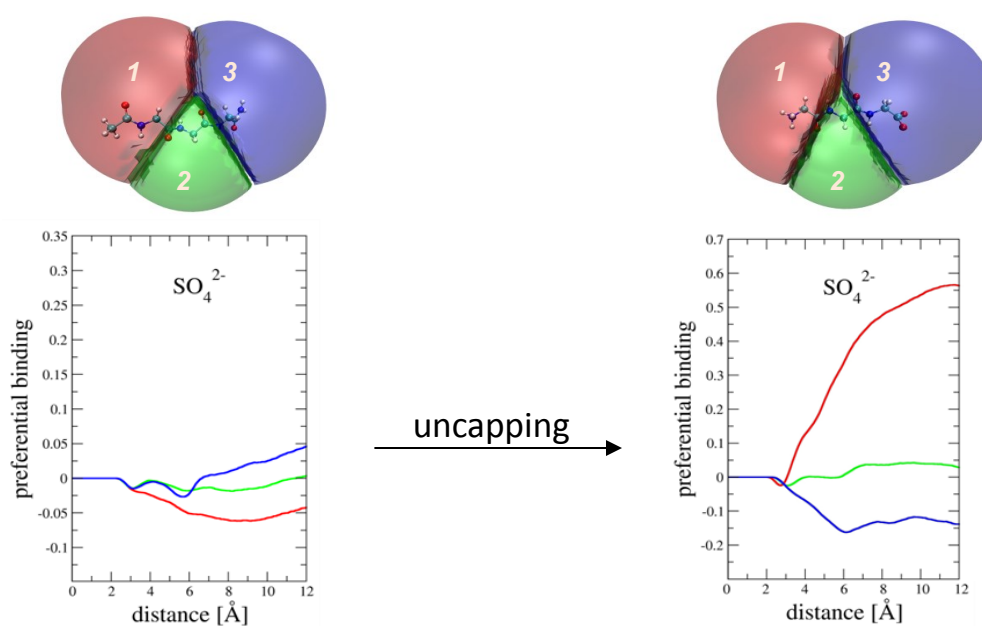


**Fig. 17**: Preferential binding coefficients Γ for sulfate anion interacting with capped (left) and uncapped (right) triglycine. Individual curves correspond to the spatial parts depicted above.

When we compare our results with measured dissociation constants $K_D$ for individual ions, we find a good agreement. Similarly to the polypeptide study, changes in NMR shifts of the three types of methylene protons were measured as a function of the different salt concentrations. Linear parts of the data were subtracted and the resulting curves were matched with the binding isotherm. We see in Tab. 2 that vanishing binding constants were established for $SO_4^{2-}$, $Cl^-$, and $Br^-$. Iodide is the first anion with measurable $K_D$, while $SCN^-$ binds the strongest and at a larger area than iodide.

| $Na^+$ counter-ion | Ac-GGG-NH$_2$ | | | | GGG | | |
|---|---|---|---|---|---|---|---|
| | $K_d$ (mM) | | | | $K_d$ (mM) | | |
| | α-proton 1 | α-proton 2 | α-proton 3 | | α-proton 1 | α-proton 2 | α-proton 3 |
| $SO_4^{2-}$ | - | - | - | | 70 | - | - |
| $Cl^-$ | - | - | - | | 290 | - | - |
| $Br^-$ | - | - | - | | 790 | - | - |
| $I^-$ | 1400 | - | - | | 1500 | - | - |
| $SCN^-$ | 1200 | 1800 | - | | 960 | 1600 | - |

**Tab. 2**: Fitted values for dissociation constants $K_d$ of methylene protons 1, 2, and 3 of a terminated (left part) and non-terminated (right part) GGG molecule in 5 sodium salts.

As charged residues are frequently present in protein stuctures, we investigated also the uncapped version of triglycine in order to unravel how the presence of a charge influences the anionic binding. The uncapped N-terminus serves as a model binding site for positively charged residues (Lys), while the uncapped C-terminus represents negatively charged residues (Asp or Glu).

From the spatial distribution functions (right half of Fig. 16) and preferential binding coefficients to the non-terminated GGG (an example for sulfate is presented in Fig. 17, right half) we find a reversal of the anionic Hofmeister series. In contrast to capped GGG, where no significant interactions with terminal groups are observed, the positively charged N-terminus dominates the anionic binding and direct charge-charge interactions play a key role. As a result, the anion with the biggest charge density, $SO_4^{2-}$, exhibits the strongest affinity, followed by $Cl^-$ > $Br^-$ > $I^-$. The more complex situation for $SCN^-$ arises from the non-spherical shape of the anion, as well as from the fact that both charge-charge and charge-backbone interaction are involved. As a result, thiocyanate does not exactly follow

the reversed Hofmeister series having stronger affinity for the uncapped triglycine structure than the halide ions.

The reversed anionic Hofmeister series is confirmed by the fitted dissociation constants obtained from NMR measurements. Sulfate displays the tightest and ioide the weakest $K_D$ value for the methylene group closest to the charged N-terminus (Tab. 2). Interestingly, thiocyanate is the only ion dissociation constant of which is non-vanishing for both of the methylene groups (1 and 2). This is in agreement with MD observations and, therefore, we can conclude $SCN^-$ becomes the outlier of the series. The $K_D$ values, however, put $SCN^-$ only before iodide, resulting in the following order: $SO_4^{2-} > Cl^- > Br^- > SCN^- > I^-$. This small difference between MD and NMR results may be due to the fact that NMR monitors changes at the methylene groups but not directly at the rest of the binding sites. It is also possible that we somewhat overestimate the $SCN^-$ binding by the employed empirical force field.

A general fact that the Hofmeister series can be reversed by capping/uncapping of the tripeptide gives us a powerful tool for tuning ion-peptide interactions. It follows from our observations that we have to consider not only anions themselves or their hydration properties, but also the peptide structure itself, as the latter dominates the interactions with individual anions. Consequently, pH titration or mutation of charged amino acids can be systematically used to obtain desirable peptide properties, such as solubility or enzymatic activity based on specific ion interactions.

## 4.3 Theory challenges experiment

For the purpose of ordering Hofmeister ions interacting with the peptide bond, Nandi and Robinson established salting-out constants $k_s$ of oligoglycines based on their solubility in various salt solutions already in the 1970s. As their study[17] included capped triglycine, the possibility of comparison with our results was tempting. The experimental $k_s$ values are summarized in Tab. 3 together with the data for uncapped triglycine[69]. We see that all the measured salts exhibit salting-in behavior with uncapped triglycine. By contrast, the $k_s$ values for capped GGG follow the Hofmeister series. Salting-out effect ($k_s > 0$) is observed for $Na_2SO_4 > NaF$, no effect ($k_s \approx 0$) for NaCl and NaBr, and salting-in behavior ($k_s < 0$) for NaI < NaSCN $\approx$ NaClO$_4$. This in general qualitatively matches our previous conclusions.

| salt | CH$_3$CO-GGG-COOC$_2$H$_5$[17] | | GGG[69] | |
|:---:|:---:|:---:|:---:|:---:|
| | solubility | k$_s$ (M$^{-1}$) | solubility | k$_s$ (M$^{-1}$) |
| water | 0.648 | 0 | 6.41 | 0 |
| Na$_2$SO$_4$ | 0.18 | 0.55 | --- | --- |
| NaF | 0.38 | 0.23 | --- | --- |
| NaCl | 0.55 | 0.07 | 9.09 | -0.15 |
| NaBr | 0.57 | 0.05 | 9.48 | -0.17 |
| NaI | 0.87 | -0.13 | --- | --- |
| NaSCN | 1.05 | -0.21 | --- | --- |
| NaClO$_4$ | 1.05 | -0.21 | --- | --- |

**Tab. 3**: Experimental[17,69] solubilities (g of peptide/100g of solution) and salting-out constant k$_s$ for acetyltriglycine ethyl ester and uncapped triglycine in 1 M salt solutions of Hofmeister salts (except for 0.33 M Na$_2$SO$_4$).

In order to directly compare experimental and theoretical salting-out constants, we sampled the capped and the uncapped triglycine molecule in seven different salt solutions (Na$_2$SO$_4$, NaF, NaCl, NaBr, NaI, NaSCN, and NaClO$_4$). After evaluation of the preferential binding coefficients $\Gamma_s$ for each salt individually, we applied a direct relation k$_s$ = -$\Gamma_s$/(2.303*c$_s$) to obtain the salting-out constants k$_s$ from MD simulations (Fig. 18 top and middle part).

In case of uncapped GGG, we observe salting-in behavior of all ions with a bell-shape dependence when looking at the results from the strongly hydrated to the weakly hydrated anions. The explanation lies in the dual character of anionic binding to the charged triglycine structure. While the charge-charge interactions play a key role for strongly hydrated anions, the interactions with the backbone dominate for weakly hydrated anions. As a result, the most effective salting-in agents occur at both edges of the Hofmeister series.

When we compare our results[70] for capped GGG with those ones of Nandi and Robinson[17], for most ions we find a good agreement considering the accuracy of the applied empirical force fields. However, for the strongly hydrated anions (SO$_4^{2-}$ and F$^-$), there is a clear mismatch between experimental and theoretical k$_s$ values. While MD predicts that strongly hydrated anions interact very weakly with the backbone resulting in slightly positive salting-out constants, experimental values indicate a strong repulsion (high positive k$_s$ values) from the peptide surface. This disagreement led us to look more deeply at the original experimental data. Capped triglycine was synthetized in two steps from the uncapped version with a half-capped intermediate (Fig. 19).

**Fig. 18**: Salting-out coefficients $k_s$ obtained from MD simulations for uncapped (top), capped (middle), and half-capped (bottom) triglycine in comparison with experimental results from Nandi and Robinson[16].



uncapped GGG                    half-capped GGG

capped GGG (ester)

**Fig. 19**: A two-step synthesis of capped triglycine presented by Nandi and Robinson[17].

Therefore, for the completeness of our investigation, the half-capped triglycine structure was simulated in the same seven salt solutions. Consequently, the resulting salting-out constants were calculated and compared with experimental results. We see from Fig. 14 (bottom) that the $k_s$ values match quantitatively the experimental values for all the investigated salts. The extra negative charge in the structure enhances the sodium binding; on the other hand, it increases the anionic repulsion from this site. The higher the charge density of the anion, the more significant depletion is observed and it dominates over additional cationic attraction. In contrast, the weakly hydrated anions interact with the backbone and are thus less repelled from the carboxylic group. Therefore, the enhanced cation binding can to some extent counterbalance this repulsion and the resulting salting-out constants do not differ much from the fully-capped structure. This conclusion nicely explains why the major differences become evident for $Na_2SO_4$ and NaF only.

Eventually, our suggestion that Nandi and Robinson most probably did not succeed in preparation of capped GGG and measured solubility of the half-capped version instead, was experimentally proven. The published esterification process[17] turned out not to be reproducible and, as a result, an alternative capping route was suggested[70].

## 4.4 Electrophoresis of neutral molecules

As mentioned above, there has been a long-lasting debate, whether NMA is an appropriate model for the peptide backbone[66] or whether we need to investigate oligoglycine as the simplest relevant model. To address this question, we chose a non-standard approach for neutral molecules, namely electrophoresis which is usually used for separation of charged molecules only[71]. However, when the neutral molecule interacts with cations and anions of the background electrolyte (BGE) in an unequal way, it becomes effectively a charged species with a non-zero electrophoretic mobility. From our previous results, we would then expect that by varying the cations or anions of the BGE, the electrophoretic mobilities of NMA must differ. Moreover, we would be able to quantify these effects and directly connect cationic and anionic Hofmeister series by finding such an ion pair, where the binding of the cation and the anion is the same resulting in a zero mobility of the neutral molecule.

Therefore, we asked our collaborators to measure electrophoretic mobilities of NMA in different electrolytes, altering both cations and anions in the BGE salts. The standard way of reporting electrophoretic mobilites of any molecule is to measure the mobility relative to an electrophoretic (EOF) marker ($\Delta\mu$)[72] rather than to establish the

absolute value ($\mu$). Nevertheless, the commonly used markers as dimethylsulfoxide (DMSO) or thiourea (TUR) can in principle interact with the BGE, too. In this case, the assumption widely used in the electrophoretic community that all the EOF markers behave like immobile species independently on the used electrolyte, would not be strictly valid. We wanted to quantify this effect and find cases with significant ion-marker interactions.

To check on the BGE dependent mobility of the marker molecules, our colleagues performed experiments, where a mixture of NMA, DMSO, and TUR was subjected to the electric field in a sodium methylsulphonate/sodium acetate solution. The presence of the acetate buffer is crucial for the technique as such and is pertinent to all the experiments. We observed a clear separation of all three molecules (Fig. 20), which implies different interactions of the markers with the BGE.
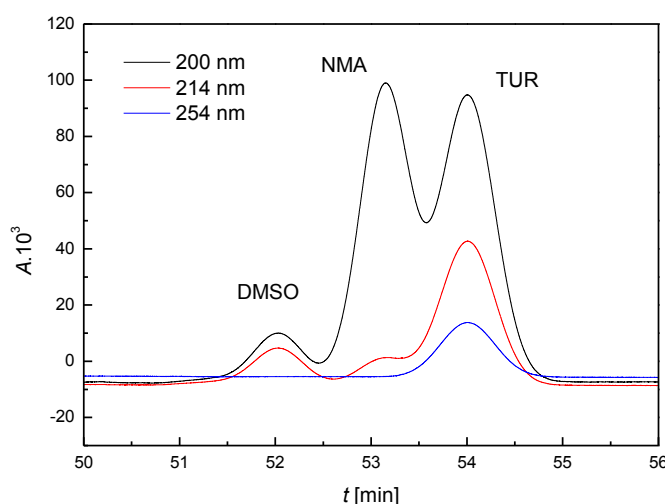


**Fig. 20**: Electropherogram from the separation of NMA/DMSO/TUR mixture in the solution of $NaCH_3SO_3$/ $CH_3COONa$ 4:1. The UV spectra shows clear separation of analytes at 200 nm, 214 nm, and 254 nm.

Based on these findings, the experiment was extended and the electrophoretic mobilities of the neutral molecules (NMA, DMSO, and TUR) were measured in a wide range of Hofmeister salts[73]. In addition, the mobilites of NMA relative to DMSO $\Delta\mu^{NMA-DMSO}$ or to thiourea $\Delta\mu^{NMA-TUR}$ were calculated. For a better interpretation of the experimental data, we needed an atomistic insight. Therefore, we simulated all three investigated molecules in twelve salt solutions – four acetate salts and eight sodium salts as presented in Tab. 4. From the classical molecular dynamic simulations, the preferential binding coefficients Γ were obtained and by applying Eq. 13 (see Chapter 2), the individual electrophoretic mobilities Δμ were calculated. Due to statistical sampling issues, different salt concentrations were used in the simulations (1.3 M) than in the experiment (0.25 M). Due to this fact, the final

comparison is made in terms of the dΔμ/dc slope as the formula for the electrophoretic mobility assessment is linear with respect to the salt concentration. Both experimental and computational results are summarized in Tab. 4 for N-methylacetamide with respect to thiourea.

| | $d\Delta\mu_{XAc}^{NMA,TUR}/dc_{XAc}$ | | | $d\Delta\mu_{NaX}^{NMA,TUR}/dc_{NaX}$ | |
| | $[10^{-9}\cdot m^2\cdot V^{-1}\cdot s^{-1}\cdot mol^{-1}\cdot l]$ | | | $[10^{-9}\cdot m^2\cdot V^{-1}\cdot s^{-1}\cdot mol^{-1}\cdot l]$ | |
| salt | experiment | simulation | salt | experiment | Simulation |
|------|-----------|-----------|------|-----------|-----------|
| CsAc | -1.6 | -1.0 | $Na_2SO_4$ | 5.0 | 0.1 |
| KAc | --- | -1.0 | $CH_3COONa$ | 0.0 | 0.1 |
| NaAc | 0.0 | 0.1 | NaF | 2.0 | 1.3 |
| LiAc | 1.5 | 1.5 | NaCl | 1.5 | -1.1 |
| | | | $CH_3SO_3Na$ | 1.0 | 1.2 |
| | | | $NaClO_4$ | -1.1 | -1.4 |
| | | | NaI | --- | -2.9 |
| | | | NaSCN | --- | -0.6 |

**Tab. 4**: Relative electrophoretic mobilities of NMA with respect to TUR as measured and calculated in a variety of acetate salts (on the left) and sodium salts (on the right).

The changes of electrophoretic mobilities by varying the cations of the BGE can be interpreted using the radial distribution function of these cations (Fig. 21). While lithium is attracted to the NMA surface and, consequently, is in excess in the vicinity of the molecule, cesium is repelled from this area. Electrophoretic mobility of NMA in LiAc is close to zero (Fig. 21, left-bottom), which means that $Li^+$ and $Ac^-$ occupy the NMA surroundings to a similar extent. Consequently, NMA migrates to the anode very slowly behaving like a slightly negatively charged species. It also results from Fig. 21 that NMA in CsAc moves faster to the anode having a fairly negative value of the electrophoretic mobility. In this case, $Ac^-$ is dominating the NMA-ion interactions, while cesium is depleted from the NMA surface.

To compare our results with the experiment, we need to subtract the mobilites of the EOF markers. To demonstrate our approach[73], we discuss here the thiourea marker only. We see in Fig. 21 (right half) that all three cations are attracted to the TUR surface, with $Cs^+$ exhibiting the strongest attraction among the cations. At the same time, acetate is attracted to the surface even more as the resulting $\mu_{salt}^{TUR}$ are negative for all acetate salts. Subtracting the corresponding mobilites for NMA and TUR, we get the final $\Delta\mu_{salt}^{NMA-TUR}$ values, which are directly comparable with the experiment.
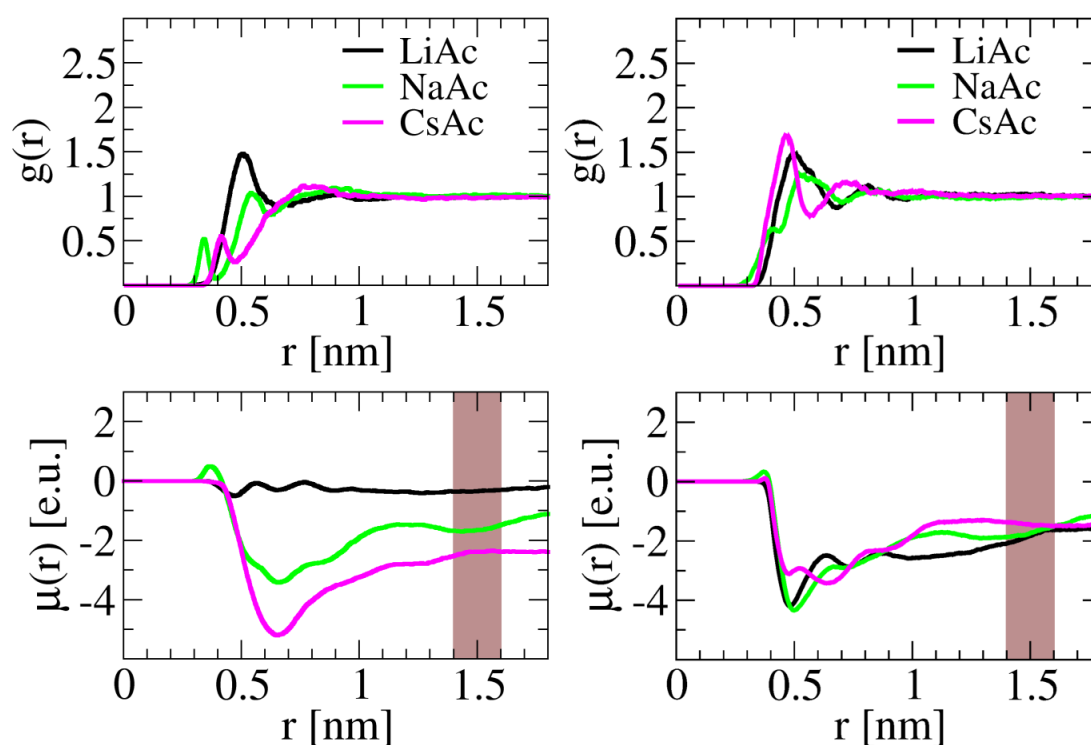
**Figure 21:** Radial distribution functions g(r)  of cations (Li[+] - black, Na[+] – green, Cs[+] - pink) with respect to NMA (left) and TUR (right) are presented in the upper part. The corresponding electrophoretic mobilites μ(r) for NMA and TUR in three acetate salts are shown in the lower part of the figure. The μ(r)  value is read from the highlighted area.

The resulting $\Delta\mu_{NaAc}^{NMA-TUR}$ value is very close to zero in a NaAc solution, indicating that in average NMA and TUR migrate in the electric field with the same speed, and consequently, are inseparable under these conditions (Tab. 4). This means that the ratio of cationic/anionic binding must be very similar for both of them. Our calculations provide an additional information that the overall negative charges of the migrating molecules (NMA and TUR) have different origin. The first results from the lack of sodium cations in the NMA vicinity, while the letter is caused by the enhanced attraction of acetate to TUR. The positive $\Delta\mu_{LiAc}^{NMA-TUR}$ value in LiAc (Tab. 4) can be explained by different binding ratios of the BGE. TUR migrates faster than NMA in the electric field because NMA exhibits an overall charge close to zero, while TUR migrates as an anion. Finally, in the CsAc solution NMA gets to the UV detector of the electrophoretic setup earlier than TUR. The depletion of Cs[+] from the NMA surface leads to faster migration of this molecule as an anion than in the case of TUR. All the calculated numbers are in a good agreement with experimental results (Tab. 5, left part).

Similarly, the effect of the anionic binding on the NMA relative mobility $\Delta\mu_{salt}^{NMA-TUR}$ was examined. The general observations follow our previous results for capped triglycine:

the more weakly hydrated anion, the more it binds to the NMA molecule. In particular, fluoride is repelled by far the most from the NMA surface which leads to the migration of NMA towards the cathode, in other words, acting like a cation. In contrast, perchlorate has the highest affinity for NMA resulting in the migration of NMA towards the anode.

When we calculated the ionic affinities for the markers (TUR and DMSO) and evaluated the relative electrophoretic mobilites $\Delta\mu_{salt}^{NMA-TUR}$, the data were in a good agreement with the experimental results (Tab. 5, right part). However, two mismatched values were obtained for $Na_2SO_4$ and (surprisingly) for NaCl. The first disagreement may be explained by the strong concentration dependent non-ideality of the $Na_2SO_4$ solution, as well as by a nonlinear grow of the ion affinities to NMA and TUR in relatively high salt concentrations. Explaining the second finding is not straightforward, as for the NaCl solution all the qualitative features are captured properly. According to our previous results, interactions with NMA follow the regular Hofmeister ordering $F^- < Cl^- < I^-$. In contrast, H-bonds play the key role in anionic binding for TUR and, therefore, the opposite trend is expected and also observed: $F^- > Cl^- > I^-$. Unfortunately, the quantitative balance of these opposing trends is probably not properly captured in calculations for the chloride anion.

The major contribution of this combined computational and experimental project was not only the verification of the simulation force fields for ions and small molecules, but also rationalization of the measured electrophoretic data. We show that neutral markers do not, in general, migrate as purely neutral molecules, but rather exhibit electrolyte specific electrophoretic mobilities. To fulfill our original ambition concerning models of the peptide backbone, we plan to investigate the capped triglycine molecule as our next target and compare how much it differs from the NMA molecule in terms of interaction with the BGE ions.

## *Chapter 5*

## Conclusions

The principle goal of the present thesis was to show how efficient a combined theoretical and experimental approach can be for obtaining complex molecular-level answers for biologically relevant questions. We applied primarily classical molecular dynamic simulations and advanced data analysis to provide atomistic insight into the issue of the Hofmeister series. Taking into account both cationic and anionic series of ions and by investigating real proteins as well as model systems, we were able to identify, quantify, and systematize variations of ion-protein interactions.

Cations bind in general more weakly to the proteins than anions, however, they can influence enzymatic activity as well. The effect can be rather moderate and follows the cationic Hofmeister series, which was demonstrated for the LinB dehalogenase[58]. In this case, cations do not enter the active site, therefore, the enzymatic activity is enhanced/suppressed indirectly at the mouth of the tunnel leading to the catalytic residues. It follows from our simulations that the higher the cationic charge density, the stronger the observed interaction with negatively charged residues and carbonyl groups of the backbone.

In contrast to the modest cationic effects on dehalogenases, we showed a dramatic enhancement of the BHMT enzymatic activity by the potassium cation[53]. Despite the fact that all the negatively charged sites at the protein surface attract the cations in a Hofmeister fashion, at the active site of the enzyme potassium is bound the strongest among the monovalent cations, defying thus the Hofmeister series. We were able to explain this observation by the interplay between the cationic hydration and the interaction with negatively charged residues at the active site. While potassium easily loses two of its water molecules, sodium is most of the time fully hydrated, which disadvantages it with respect to the ion-Asp/Glu interactions in a limited space of the active site.

A systematic study[63] of anion-peptide interactions was at the same time performed on variety of model systems. By investigation of the $(VPGVG)_{120}$ polypeptide, we proved that the backbone dominates most of the anionic binding with neutral residues. Moreover, the strength of the anion-backbone interaction follows the anionic Hofmeister series, i. e., the more weakly hydrated the anion is, the tighter is its binding to the backbone. On top of that, we showed that anions are not attracted to the nonpolar side chains of residues like valine.

To further investigate the anion-peptide binding, capped triglycine was examined as a model system for the peptide backbone[68]. We expanded the range of investigated Hofmeister salts and essentially validated our results from the previous study. Namely, $SO_4^{2-}$ is depleted from the backbone, resulting in salting-out behavior of sulfate salts, while $Cl^-$ and $Br^-$ act like neutral agents with no or only minor interactions with the backbone. $I^-$ and even more so $SCN^-$ bind strongly to the capped GGG peptide leading to a salting-in effect.

In order to mimic the effects of positively and negatively charged residues commonly present in the peptide structures the uncapped triglycine molecule in the same solutions of Hofmeister salts was considered. Now, the charge-charge interactions play the key role and, most importantly, a reversed Hofmeister series is observed ($SO_4^{2-} > Cl^- > Br^- > I^-$). The more strongly hydrated the anion, the more attraction to the positively charged N-terminus and the more repulsion from the negatively charged C-terminus is observed. Interestingly, $SCN^-$ is an outlier of the reversed series as in this case there may be a synergic effect of the charge-charge and anion-backbone interactions.

In order to validate our results by comparison with experimental data, we used the concept of salting-out constants[33]. The observed discrepancy between computational and experimental results led us to finding an error in data published by Nandi and Robinson[17] in the 1970s. We proved that they presented salting-out constants for half-capped instead of fully capped triglycine, which was their original intention. Their synthetic procedure was shown as inefficient resulting in a half-caped triglycine molecule the behavior of which very well matched with our results[70].

The conclusions of this thesis demonstrate also the strength of a reductionist approach[17,74] which divides ion effects into backbone and side-chain contributions. The fact that the Hofmeister series can be reversed by uncapping of the GGG oligopeptide indicates that the desired peptide properties or enzymatic activity can be achieved by single point mutations or pH titration, i.e., by influencing the ion binding by chemically modifying amino acid side chains at the biomolecular surface.

Future development should go beyond the simplifying concept of separate anionic and cationic Hofmeister series. Not only ion-protein interactions but also cation-anion attraction both in the bulk solution and at the protein surface should be invoked to explain the variety of experimental observations[75,76]. On top of that, specifically tailored ion-binding sites of enzymes (i.e. BHMT) or membrane ion channels are known to challenge simple Hofmeister rules. Last but not least, innovative experimental set-ups in combination with a computational insight should provide a fresh look into the problems as we demonstrated in the case of electrophoresis of neutral markers[73].

# Bibliography

1. Jungwirth, P., Cremer, P.: *Nature Chem.* **6**, 261-263 (2014).

2. Kunz, W., Henle, J., Ninham B. W.: *Curr. Opin. in Colloid & Interf. Sci.* **9**, 19-37 (2004).

3. Hofmeister, F.: *Arch. Exp. Pathol. Pharmakol. (Leipzig)* **24,** 247–260 (1888).

4. Hofmeister, F.: *Arch. Exp. Pathol. Pharmakol. (Leipzig)* **25,** 1–30 (1888).

5. Cox, W. M. & Wolfenden, J. H.: *Proc. R. Soc. London A* **145,** 475–488 (1934).

6. Gurney, R. W.: *Ionic Processes in Solution* (McGraw-Hill, 1953).

7. Marcus, Y.: *Chem. Rev.* **109,** 1346–1370 (2009).

8. Hribar, B., Southall, N. T., Vlachy, V. & Dill, K. A.: *J. Am. Chem. Soc.* **124,** 12302–12311 (2002).

9. Thomas, A. S. & Elcock, A. H.: *J. Am. Chem. Soc.* **129,** 14887–14898 (2007).

10. Omta, A. W., Kropman, M. F., Woutersen, S. & Bakker, H. J.: *Science* **301,** 347–349 (2003).

11. Funkner, S. *et al.: J. Am. Chem. Soc.* **134,** 1030–1035 (2012).

12. Rieskautt, M. M. & Ducruix, A. F.: *J. Biol. Chem.* **264,** 745–748 (1989).

13. Zhang, Y. J. & Cremer, P. S.: *Proc. Natl Acad. Sci. USA* **106,** 15249–15253 (2009).

14. Bello, J. & Bello, H. R.: *Nature* **190,** 440–441 (1961).

15. Bello, J. & Bello, H. R.: *Nature* **194,** 681–682 (1962).

16. Robinson, D. R. & Jencks, W. P.: *J. Am. Chem. Soc.* **87,** 2470–2479 (1965).

17. Nandi, P. K. & Robinson, D. R.: *J. Am. Chem. Soc.* **94,** 1299–1308 (1972).

18. Matthews, B. W.: *Ann. Rev. of Physical Chem* **27**, 493 (1976).

19. Kunz, W., Lo Nostro, P. & Ninham, B. W.: *Curr. Opin. Colloid Interf. Sci.* **9,** 1–18 (2004).

20. Wilson, E. K.: *Chem. Eng. News* **85,** 47–49 (2007).

21. Baldwin, R. L.: *Biophys. J.* **71,** 2056–2063 (1996).

22. Street, T. O., Bolen, D. W. & Rose, G. D.: *Proc. Natl Acad. Sci. USA* **103,** 13997–14002 (2006).

23. Pegram, L. M. & Record, M. T.: *J. Phys. Chem. B* **112,** 9428–9436 (2008).

24. Kalcher, I., Horinek, D., Netz, R. R. & Dzubiella, J.: *J. Phys. Condens. Matter* **21,** 424108 (2009).

25. D.A. Case, V. Babin, J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T.S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K.M. Merz, F. Paesani, D.R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu and P.A. Kollman: AMBER 14, University of California, San Francisco (2014).

26. van der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., Berendsen, H. J. C.: *J. Comp. Chem*. **26**, 1701–1718 (2005).

27. Sherwood, P., de Vries, A. H., Guest, M. F.,  Schreckenbach, G., Catlow, C. R. A., French, S. A. , Sokol, A. A., Bromley, S. T., Thiel, W., Turner, A. J., Billeter, S., Terstegen, F. , Thiel, S., Kendrick, J., Rogers, S. C., Casci, J., Watson, M., King, F., Karlsen, E., Sjøvoll, M. , Fahmi, A., Schäfer, A., Lennartz, Ch.:  *J. Mol. Struct. (Theochem.)* **632**, 1 (2003).

28. Neese, F. : *WIREs Comput. Mol. Sci.* **2** , 73-78 (2012).

29. Kästner, J., Carr, J.M., Keal, T.W., Thiel, W., Wander, A., Sherwood, P.: *J. Phys. Chem. A* **113**, 11856-11865 (2009).

30. P. K. Mehrotra, P. K.,  Beveridge, D. L.: *J. Am. Chem. Soc.* **102**, 4287–4294 (1980).

31. Pierce, V., Kang, M., Aburi, M., Weerasinger, S., Smith, P.E.: *Cell Biochem. Biophys*. **50**, 1–22 (2008).

32. Kirkwood, J. G., Buff, F. P.:  *J. Chem. Phys.* **19**, 774-777 (1951).

33. Smith, P. E., Mazo, R. A.: *J.of Phys.Chem. B*  **112**, 7875-7884 (2008).

34. Urry D.W. et al.: *Prog. Biophys. Mol. Biol.* **57**, 23–57 (1992).

35. Meyer, D.E., Chilkoti, A.: *Nature Biotechnology*, **17**, 1112-1115 (1999).

36. Cho, Y.H., Zhang, Y.J., Christensen, T., Sagle, L.B. Chilkoti, A., Cremer, P.S.: *J. Phys. Chem. B*. **112**, 13765-13771 (2008).

37. Dengler, S., Klaus, A., Tiddy, G. J. T., Kunz, W.: *Faraday Discuss.* **160**, 121-133 (2012).

38. Holz, M., Grunder, R., Sacco, A., Meleleo, A.:  *J. Chem. Soc., Faraday Trans.* **89**, 1215-1222 (1993).

39. Okur, H. I., Kherb, J., Cremer, P. S.:  *J. Am. Chem. Soc*. **135**, 5062-5067 (2013).

40. Ghosal, S.: *Electrophoresis* **25**, 214-228 (2004).

41. Williams, B. A., Vigh, C.: *Anal. Chem.* **68**, 1174-1180 (1996).

42. Beckers, J. L., Everaerts, F. M., Ackermans, M. T.: *J. Chromatogr.* **537**, 407-428 (1991).

43. Ibrahim, A., Allison, S. A., Cottet, H.: *Anal. Chem.* **84**, 9422-9430 (2012).

44. Koval, D. A., Kasicka, V., Zuskova, I.: *Electrophoresis*  **26**, 3221-3231 (2005).

45.  Altenberger, A. R., Friedman, H. L.:  *J. Chem. Phys.* **78**, 4162-4173 (1983).

46. Delgado-Reyes, C., Wallig, M. A., Garrow, T. A.: *Arch. Biochem. Biophys.* **393**, 184-186 (2001).

47. Petrossian, T. C., Clarke, S. G.: *Mol. Cel. Proteomics* **10**, 1-12 (2011).

48. Fromm, H. J., Nordlie, R. C.: *Arch. Biochem. Biophys.* **81**, 363-376 (1959).

49. Finkelstein, J. D., Harris, B. J., Kyle, W. E.: *Arch. Biochem. Biophys.* **153**, 320-324 (1972).

50. Gonzáles, B., Pajares, M.A., Matínez-Ripoll M.M., Blundell, T.L., Sanz-Aparicio J.: *J. Mol. Biol.* **338**, 771-782 (2004).

51. Evans, J. C., Huddler, D. P., Jiracek, J., Castro, C., Millian, N., Garrow, T. A., Ludwig, M. L.: *Structure* **10**, 1159-1071 (2002).

52. Jiracek, J., Collinsova, M., Rosenberg, I., Budesinsky, M., Protivinska, E., Netusilova, H., Garrow, T. A.: *J. Med. Chem.* **49**, 3982-3989 (2006).

53. Mladkova, J., Hladilkova, J., Diamond, C. E., Tryon, K., Yamada, K., Garrow, T. A., Jungwirth, P., Koutmos, M., Jiracek, J.: *Proteins*, in press (2014).

54. Vrbka, L., Vondrasek, J., Jagoda-Cwiklik, B., Vacha, R., Jungwirth, P.: *Proc. Natl. Acad. Sci. U. S. A.* **103**, 15440-15444 (2006).

55. Damborsky, J., Rorije, E., Jesenska, A., Nagata, Y., Klopman, G., Peijnenburg, W.: *Environ. Toxicol. Chem.* **20**, 2681-2689 (2001).

56. Chaloupkova, R., Sykorova, J., Prokop, Z., Jesenska, A., Monincovaa, M., Pavlova, M., Tsuda, M., Nagata, Y., Damborsky, J.: *J. Biol. Chem.* **278**, 52622-52628 (2003).

57. Pavlova, M., Klvana, M., Prokop, Z., Chaloupkova, R., Banas, P., Otyepka, M., Wade, R. C., Tsuda, M., Nagata, Y., Damborsky, J.: *Nat. Chem. Biol.* **5**, 727-733 (2009).

58. Stepankova, V., Paterova, J., Damborsky, J., Jungwirth, P., Chaloupkova, R., Heyda, J.: *J. Phys. Chem. B* **117**, 6394-6402 (2013).

59. Kumar, S., Bouzida, D., Swendsen, R. H., Kollman, P. A., Rosenberg, J. M.: *J. Comput. Chem.* **13**, 1011-1021 (1992).

60. Biedermannova, L., Prokop, Z., Gora, A., Chovancova, E., Kovacs, M., Damborsky, J., Wade, R. C.: *J. Biol. Chem.* **287**, 29062-29074 (2012).

61. Hladilkova, J., Prokop, Z., Chaloupkova, R., Damborsky, J., Jungwirth, P.: *J. Phys. Chem. B* **117**, 14329-14335 (2013).

62. Cho, Y.H., Zhang, Y.J., Christensen, T., Sagle, L.B. Chilkoti, A., Cremer P.S.: *J. Phys. Chem. B.* **112**, 13765-13771 (2008).

63. Rembert, K., Paterova, J., Heyda, J., Hilty, C., Jungwirth, P., Cremer, P. S.: *J. Am. Chem. Soc.* **134**, 10039-10046 (2012).

64. Cho, Y.H., Sagle, L.B., Iimura, S., Zhang, Y.J., Kherb, J., Chilkoti, A., Scholtz, J.M., Cremer, P.S.: *J. Am. Chem. Soc.* **131**, 15188-15193 (2009).

65. Pegado, L., Marsalek, O., Jungwirth, P., Wernersson, E.: *Phys. Chem. Chem. Phys.* **14**, 10248-57 (2012).

66. Bolen, D. W., Rose, G. D.: *Annu. Rev. Biochem.* **77**, 339–362 (2008).

67. Heyda, J., Vincent, J. C., Tobias, D. J., Dzubiella, J., Jungwirth, P.: *J. Phys. Chem. B* **114**, 1213–1220 (2010).

68. Paterova, J., Rembert, K., Heyda, J., Kurra, Y., Okur, H. I., Liu, W. R., Hilty, C., Cremer, P. S., Jungwirth, P.: *J. Phys. Chem. B* **117**, 8150-8158 (2013).

69. Venkatesu, P., Lee, M. J., Lin, H. M.: *Biochem. Engineering J.* **32**, 157-170 (2006).

70. Hladilkova, J., Heyda, J., Rembert, K., Okur, H. I., Kurra, Y., Liu, W. R., Hilty, C., Cremer, P. S., Jungwirth, P.: *J. Phys. Chem. Lett.* **4**, 4069-4073 (2013).

71. Ghosal, S.: *Electrophoresis* **25**, 214–228 (2004).

72. Williams, B. A., Vigh, C.: *Anal. Chem.* **68**, 1174–1180 (1996).

73. Krizek, T., Kubickova, A., Hladilkova, J., Coufal, P., Heyda, J., Jungwirth, P.: *Electrophoresis* **35**, 617-624 (2014).

74. Kalcher, I., Horinek, D., Netz, R. R. & Dzubiella J.: *J. Phys. Cond. Matt.* **21**, 424108 (2009).

75. Schneider, C. P., Shukla, D. & Trout, B. J.: *J. Phys. Chem. B* **115**, 7447-7458 (2011).

76. Dempsey, C.E., Mason, P.E. & Junwirth, P.: *J. Am. Chem. Soc.* **133**, 7300-7303 (2011).

# List of Abbreviations

| | |
|---|---|
| Asn | asparagine amino acid |
| Asp | aspartic amino acid |
| BGE | background electrolyte, used in electrophoresis |
| BHMT | betaine—homocysteine S-methyltransferase |
| CBHcy | $S$-(δ-carboxybutyl)-dl-homocysteine; inhibitor of BHMT |
| Cys | cysteine amino acid |
| DCC | sequence of aspartic acid, glycine and glycine amino acids |
| DMSO | dimethyl sulfoxide |
| ELP | elastine-like polypeptide |
| EOF | electro-osmotic flow |
| GGG | triglycine |
| Gln | glutamine amino acid |
| Glu | glutamic amino acid |
| Gly | glycine amino acid |
| Hcy | Homocysteine |
| His | histidine amino acid |
| KB | Kirkwood-Buff |
| LCST | lower critical solution temperature |
| MD | molecular dynamics |
| NMA | N-methylacetamide |
| NMR | nuclear magnetic resonance |
| NPT | isothermal-isobaric ensemble |
| QM/MM | combination of quantum mechanics and molecular mechanics |
| RDF | radial distribution function |
| SDF | spatial distribution function |
| Trp | tryptophane amino acid |
| TUR | thiourea |
| Tyr | tyrosine amino acid |
| VPGVG | valine-proline-glycine-valine-glycine sequence of amino acids |
| WHAM | weighted histogram analysis method |
| wt | wild type of an enzyme |

# List of Attached Publications

1. Rembert, K.; Paterová, J.; Heyda, J.; Hilty, C.; Jungwirth, P.; Cremer, P. S.: The Molecular Mechanisms of Ion-Specific Effects on Proteins. *Journal of the American Chemical Society*, **2012**, *134*, 10039.

2. Štěpánková, V; Paterová, J.; Damborský, J.; Jungwirth, P.; Chaloupková, R.; Heyda, J.: Cation Specific Effects on Enzymatic Catalysis Driven by Interactions at the Tunnel Mouth. *Journal of Physical Chemistry B*, **2013**, *117*, 6394.

3. Paterová, J.; Rembert, K.; Heyda, J.; Kurra, Y.; Okur, H. I.; Liu, W. R.; Hilty, C.; Cremer, P. S.; Jungwirth, P.: Reversal of the Hofmeister Series: Specific Ion Effects on Peptides. *Journal of Physical Chemistry B*, **2013**, *117*, 8150.

4. Hladílková, J.; Prokop, Z.; Chaloupková, R.; Damborský, J.; Jungwirth, P.: Release of Halide Ions From the Buried Active Site of the Haloalkane Dehalogenase LinB Revealed by Stopped-Flow Fluorescence Analysis and Free Energy Calculations. *Journal of Physical Chemistry B*, **2013**, 117, 14329.

5. Hladílková, J.; Heyda, J.; Rembert, K.; Okur, H. I.; Kurra, Y.; Liu, W. R.; Hilty, C.; Cremer, P. S.; Jungwirth, P.: Effects of End-Group Termination on Salting-Out Constants for Triglycine. *Journal of Physical Chemistry Letters* **2013**, *4*, 4069.

6. Křížek, T.; Kubíčková, A.; Hladílková, J.; Coufal, P.; Heyda, J.; Jungwirth, P.: Electrophoretic mobilities of neutral analytes and electroosmotic flow markers in aqueous solutions of Hofmeister salts. *Electrophoresis* **2014**, *35*, 617.

7. Mládková, J.; Hladílková, J.; Diamond, C. E.; Tryon, K.; Yamada, K.; Garrow, T. A.; Jungwirth, P.; Koutmos, M.; Jiráček, J.: Specific Potassium Ion Interactions Facilitate Homocysteine Binding to Betaine-Homocysteine *S*-Methyltransferase. *Proteins,* **2014**, in press (DOI: 10.1002/prot.24619).