Electrostatic-Determined Conformations – Exploring Binding Specificity of Selectins and Proton Gating in Cytochrome *c* Oxidase

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

eingereicht im Fachbereich Biologie, Chemie, Pharmazie
der Freien Universität Berlin

vorgelegt von

Anna Lena Wölke

aus Berlin

September 2014

Die vorliegende Arbeit wurde unter Anleitung von Prof. Dr. E.-W. Knapp im Zweitraum vom 01.04.2011 bis 18.09.2014 am Institut für Chemie der Freien Universität Berlin im Fachbereich Biologie, Chemie, Pharmazie durchgeführt.

1. Gutachter: Prof. Dr. Ernst-Walter Knapp, Freie Universität Berlin

2. Gutachter: Prof. Dr. Joachim Heberle, Freie Universität Berlin

Disputation am 13.11.2014

Acknowledgements

I would like to thank Prof. Dr. Ernst-Walter Knapp for your supervision and your support. I have greatly benefited from your knowledge and experience.

I would like to thank Prof. Dr. Joachim Heberle for fruitful discussions and our productive collaboration and also for reviewing this work.

I would like to acknowledge every of my co-workers in the AG Knapp, who supported me scientifically and who also created a really pleasant working atmosphere.

I would like to thank my family and friends for your support and your belief in me.

Table of Contents

List of Publications

This cumulative PhD thesis comprises the following publications:

A.L. Woelke, C. Kuehne, T. Meyer, G. Galstyan, J. Dernedde, E.W. Knapp, Understanding Selectin Counter-Receptor Binding from Electrostatic Energy Computations and Experimental Binding Studies, J. Phys. Chem. B, 117 (2013) 16443-16454.

A.L. Woelke, G. Galstyan, A. Galstyan, T. Meyer, J. Heberle, E.W. Knapp, Exploring the Possible Role of Glu286 in CcO by Electrostatic Energy Computations Combined with Molecular Dynamics, J. Phys. Chem. B, 117 (2013) 12432-12441.

A.L. Woelke, G. Galstyan, E.W. Knapp, Lysine 362 in Cytochrome c Oxidase Regulates Opening of the K-Channel via Changes in pKA and Conformation, Biochim. Biophys. Acta, Accepted for publication, August 2014.

A.L. Woelke, A. Wagner, G. Galstyan, T. Meyer, E.W. Knapp, Proton Transfer in the K-Channel Analogue of B-Type Cytochrome c Oxidase from *Thermus thermophilus*, Biophysical Journal, Accepted for publication, September 2014.

During the same period, I participated in the following additional publications:

A.L. Woelke*, M.S. Murgueitio*, R. Preissner, Theoretical Modeling Techniques and Their Impact on Tumor Immunology, Clin. Dev. Immunol., 2010 (2010) 271794.

A.L. Woelke, J. von Eichborn, M.S. Murgueitio, C.L. Worth, F. Castiglione, R. Preissner, Development of Immune-Specific Interaction Potentials and Their Application in the Multi-Agent-System VaccImm, PLoS One, 6 (2011) e23257.

J. von Eichborn*, A.L. Woelke*, F. Castiglione, R. Preissner, VaccImm: Simulating Peptide Vaccination in Cancer Therapy, BMC Bioinformatics, 14 (2013) 127.

* Joint first authorship

1. Introduction

Electrostatic interactions determine a great number of important processes in biology. A very prominent example is protein folding. The architecture of native protein structures is mainly determined by hydrogen bonds that are a sub class of electrostatic interactions. The stability of different protein conformations or protein complexes may often be accurately compared solely based on electrostatics. But also various enzymatic reactions, binding events or energy conversions are driven by electrostatic interactions.

Electrostatic forces act in a directed manner and have a long range decreasing only linearly with 1/r. These two properties may indicate why electrostatic forces influence or even dominate such a large number of biological processes. We can use this fact to explore complex reactions via electrostatic calculations, which are based on simple physical laws.

In this work, I investigate two biological research questions using electrostatic calculations: (i) which components drive the complex formation of selectins and their counter receptor on a molecular level, and (ii) what are the gating elements and their mechanisms of the proton input channels in A-type and B-type cytochrome *c* oxidase.

1.1 Calculating Electrostatic Energies

Electrostatic interactions may simply be deduced to the fact that opposite charges are attracting each other and equal charges are repelling each other. The mathematical description for this interaction is defined by Coulomb's law explaining that the force F that one sample point charge Q exerts onto another test charge q in vacuum at distance r is:

1.

$$\vec{F} = \frac{1}{4\pi\varepsilon_0} * \frac{q * Q}{\vec{r}^2}$$

with ε_0 being the vacuum permittivity. Therefore, the electric potential ϕ of this sample charge may be formulated as Coulomb's potential:

2.

$$\phi_q = \frac{1}{4\pi\varepsilon_0} * \frac{Q}{r}$$

However, in real-life applications the charge pattern is more complex. Proteins consist of thousands of atoms and reside mainly in aqueous solution including ions. Therefore, a more general definition of electrostatic interactions is described in this section.

Generally, a charge distribution can be described by a three-dimensional vector function of volume charge density $\rho(\vec{r})$. Charges produce an electric field, which is given by:

3.

$$\vec{E} = \iiint \frac{\vec{r} - \vec{r}'}{\|\vec{r} - \vec{r}'\|^3} \rho(\vec{r}') d^3r' = -\vec{\nabla}\phi$$

As shown, the electric field is the negative gradient of the electric potential.

Now, the behaviour of an electric field in a certain volume is taken into consideration. One of the Maxwell equations of classical electromagnetism is Gauss's law, which states that the electric flux through a closed surface $\oint \vec{E}\, d\vec{A}$ is proportional to the enclosed charge:

4.

$$\oint \vec{E}\, d\vec{A} = \frac{Q}{\varepsilon_0}$$

Using the divergence theorem Gauss's law can be rewritten as:

5.

$$\vec{\nabla}\, \vec{E} = \frac{\rho(\vec{r})}{\varepsilon_0}$$

Combining the electric field definition (3) and Gauss's law (5), the Poisson equation follows:

6.

$$\nabla^2 \phi(\vec{r}) = -\frac{\rho(\vec{r})}{\varepsilon_0} = -4\pi\rho(\vec{r})$$

When the Poisson equation is applied in protein science, some approximations are introduced to account for the effects of neutral dipoles and ions in solution.

Proteins are typically solvated in aqueous solution and water molecules consist of very mobile dipoles that adapt to the charge distribution inside the protein. This adaptation of dipoles to the electric field is called "dielectric screening" and it reduces the strength of electrostatic interaction compared to vacuum. However, most electrostatic calculations are more robust, when the solvent is considered only implicitly by its dielectric property than surrounding the protein with explicit water molecules. Replacing water by implicit solvent typically uses a dielectric constant of ε = 80, while inside of the protein volume the dielectric constant is set to values between 2 and 20, but most commonly to ε = 4, as it has been done throughout this study. The Poisson equation can be adapted for an inhomogeneous dielectric volume to:

7.

$$\vec{\nabla}\varepsilon(\vec{r}) * \vec{\nabla}\phi(\vec{r}) = -4\pi\rho(\vec{r})$$

An additional effect that needs to be taken into consideration is the effect of salt. Debye and Hückel have developed a description for the non-uniform distribution of ions in solution. From Debye-Hückel theory, ions in aqueous solution are surrounded by their counter ions and

form ionic clouds. Debye and Hückel introduced a statistical Boltzmann term to describe their electrostatic effect, which leads to the Poisson-Boltzmann equation (PBE):

8.

$$\vec{\nabla}\varepsilon(\vec{r}) * \vec{\nabla}\phi(\vec{r}) = -4\pi\left[\rho(\vec{r}) + \kappa^2 \frac{kT}{e_c} \upsilon(\vec{r}) \sinh\left(\frac{e_c\phi(\vec{r})}{kT}\right)\right]$$

with

$$\kappa = \sqrt{\frac{8\pi N_A e_c^2 I_s}{kT}}$$

The Debye-Hückel parameter κ takes into account the ionic strength I_S and scales it using the elementary charge e_c, Avogadro's number N_A, temperature T and the Boltzmann constant k. The statistical Boltzmann distribution of ions is introduced via $\sinh\left(\frac{e_c\phi(\vec{r})}{kT}\right)$. The volume exclusion function $\upsilon(\vec{r})$ describes the solvent accessibility and is equal to one in the solvent and equal to zero in protein or other molecules of investigation.

For complex geometries, as they are typically observed for proteins, the PBE cannot be solved analytically. At low charge densities the PBE can be replaced by a linear approximation based on Taylor series leading to the linearized PBE (LPBE):

9.

$$\vec{\nabla}\varepsilon(\vec{r}) * \vec{\nabla}\phi(\vec{r}) + \kappa^2\upsilon(\vec{r})\phi(\vec{r}) = -4\pi\rho(\vec{r})$$

There are several methods to solve the LPBE numerically. A very common technique, also applied in this work, is the method of Finite Differences. For this purpose, the molecule of interest is placed into a rectangular grid with a fixed grid constant and all atomic charges are interpolated onto the grid points. Then, a dedicated solver can solve the LPBE on every of the grid points; within this work, the "Adaptive-Poisson-Boltzmann Solver" (APBS [1]) was used.

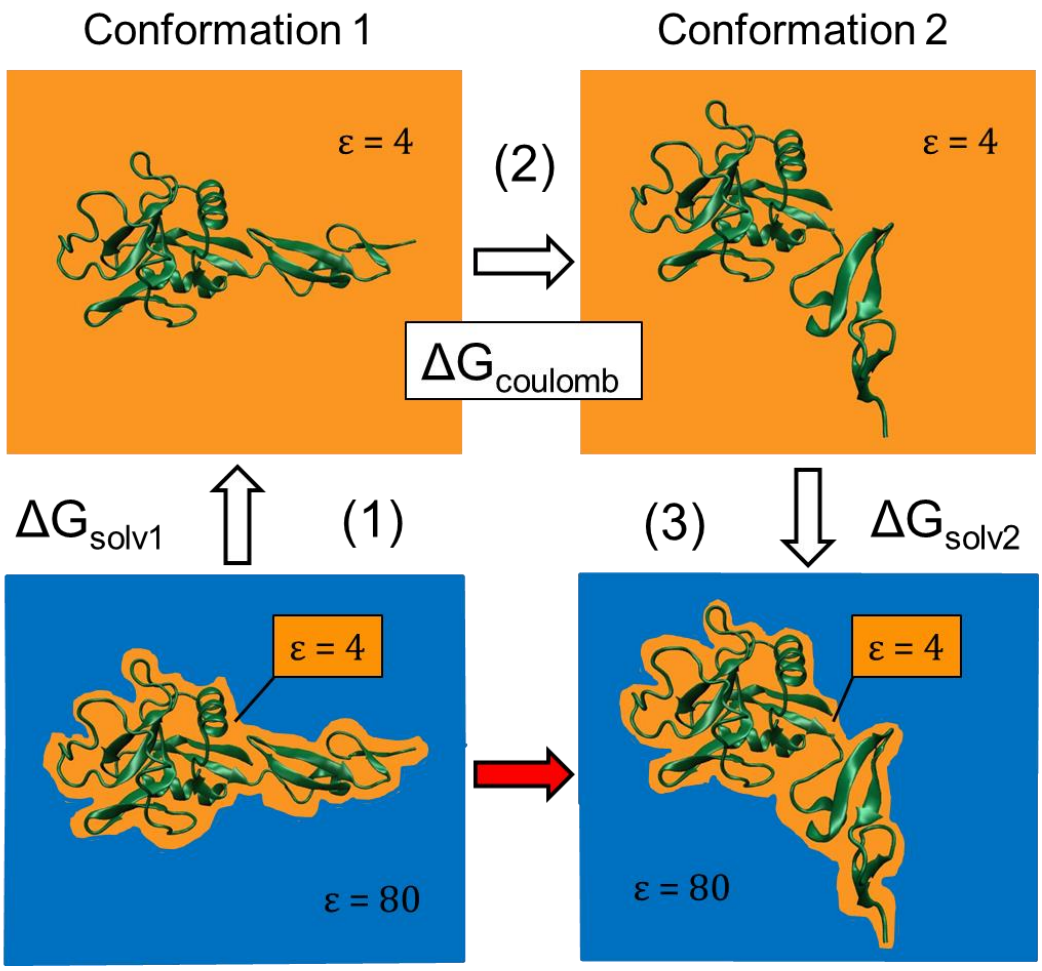

Figure 1: Thermodynamic cycle for electrostatic calculations. Aqueous solution is considered implicitly by a dielectric constant of ε = 80, protein atoms are kept explicitly and their conformational variability is represented by a dielectric constant of ε = 4 inside the protein volume. For calculating the electrostatic energy of e.g. a conformational change of a protein in solution (red arrow), the thermodynamic cycle is used. The energy sum is taken over (1) desolvation of conformation 1 from a heterogeneous dielectric medium with ε = 4 inside the protein volume and ε = 80 inside outside of the protein to a homogeneous dielectric medium of ε = 4, (2) the conformational change inside the homogeneous dielectric medium of ε = 4, and (3) solvation of conformation 2 from the homogeneous dielectric medium to the heterogeneous dielectric medium. This thermodynamic cycle may be used in a similar form for formation of a protein complex or a protonation event.

To calculate the electrostatic energy difference for two geometries of interest, e.g. a conformational change of a protein or a ligand binding event, the thermodynamic cycle (Fig. 1) is used. The molecule of interest is represented by point charges for all atoms. For both geometries, water is replaced by a dielectric constant of $\varepsilon = 80$ and within the molecule the dielectric constant is set to $\varepsilon = 4$. In a homogeneous dielectric medium of $\varepsilon = 4$, the interaction of point charges can be calculated by Coulomb's law (1). The so-called solvation energy is calculated by moving the molecule from the homogeneous dielectric medium of $\varepsilon = 4$ to a heterogeneous dielectric medium with $\varepsilon = 4$ inside the molecule surrounded by $\varepsilon = 80$. This solvation energy is calculated by solving the LPBE on a finite grid using APBS [2]. The sum of desolvating the molecule in geometry 1, changing the geometry inside a homogeneous dielectric, and solvating the molecule in geometry 2, results in the electrostatic energy of the conformational change (Fig. 1).

1.2 Calculating pK_A Values

For the calculation of electrostatic energies, it is crucial that all titratable residues are in their correct protonation state. For proton transfer reactions, it is even the main question of interest, whether a certain residue is protonated and how much energy is required for protonation. This information is described by the pK_A value. Here, the pK_A mainly refers to titratable amino acids, but the formulas hold generally for any type of titratable group. The pK_A value of a titratable group is equal to the pH value, where half of the titratable group is protonated and half is deprotonated. It may be derived from a simple generic description of an acid called HA, which dissociates to the deprotonated acid A^- and a proton H^+:

10.

$$HA \rightleftharpoons H^+ + A^-$$

The equilibrium constant for this dissociation is K_A:

11.

$$K_A = \frac{a(H^+)\, a(A^-)}{a(HA)} \approx \frac{c(H^+)\, c(A^-)}{c(HA)}$$

When introducing the logarithm, the Henderson-Hasselbalch equation is derived:

12.

$$pK_A = pH - \log\left(\frac{c(A^-)}{c(HA)}\right)$$

Depending on the question of interest, it might be more useful to refer to the protonation probability, which is described as:

13.

$$\langle x \rangle = \frac{e^{-\ln(10)[pH - pK_A]}}{1 + e^{-\ln(10)[pH - pK_A]}}$$

$$pK_A = pH - \log\left(\frac{1 - \langle x \rangle}{\langle x \rangle}\right)$$

Accordingly, the Gibbs free energy change of deprotonation of the acid is given by:

14.

$$\Delta G(HA \rightarrow A^{-}) = -\ln(10)\ RT\ (pH - pK_A)$$

In principle, the energy for protonation or deprotonation may be calculated from the thermodynamic cycle (Fig. 1). But in contrast to conformational changes, a protonation event involves formation of a new chemical bond. In this case, quantum chemical effects play a non-negligible role and pure electrostatic calculations are insufficient to describe this event. Instead, experimentally-measured model pK_A values are used from standard amino acids in aqueous solution. Accordingly, step 2 in the thermodynamic cycle (Fig. 1) uses a different calculation including the model pK_A of the amino acid and step 1 and step 3 are still calculated electrostatically. Thus, the pK_A of amino acid μ inside the protein is derived from its model pK_A from experiment and the free energy difference of moving the protonated residue out of the protein into water ($\Delta G_{\mu 1}$) and the deprotonated residue from water back into the protein ($\Delta G_{\mu 3}$):

15.

$$pK_A^P = pK_A^M + \frac{\Delta\Delta G_\mu^P}{RT\ln(10)}$$

The Gibbs free energy difference of an amino acid inside and outside of the protein may be calculated as generally described in the previous section for conformational changes of molecules. In this work the software karlsberg+ [3, 4] was used to calculate pK_As and this software further divides the Gibbs free energy difference for technical and performance reasons. The Gibbs free energy difference is composed of the solvation/desolvation energy of the amino acid $\Delta\Delta G_\mu^{solv}$ and of the interaction energy with all other charges of the protein $\Delta\Delta G_\mu^{inter}$:

16.

$$\Delta\Delta G_\mu^P = \Delta\Delta G_\mu^{solv} + \Delta\Delta G_\mu^{inter}$$

The solvation energies ΔG_μ^{solv} are calculated by moving the amino acid in the protonated and deprotonated state from the protein with a dielectric constant of ε = 4 inside into water with a dielectric constant of ε = 80. The electrostatic interaction energy with background charges

ΔG_{μ}^{inter} is a sum of all pair-wise interactions between the atoms of the amino acid and the atoms of the protein. This sum is simple to be calculated when the amino acid of interest is the only titratable group in the protein. However, most proteins contain several titratable groups and their protonation states depend on each other. For this reason, ΔG_{μ}^{inter} is further subdivided into the background charges of the protein ΔG_{μ}^{back} that do not change within the titration and the charges of titratable groups $\Delta G_{\mu}^{titrate}$. Now, ΔG_{μ}^{back} is calculated as the sum of pair-wise interactions, when all titratable residues are in their non-charged reference state. Hence, $\Delta G_{\mu}^{titrate}$ is the energy difference between the reference state with all titratable groups being uncharged and the "real" state with all titratable groups being in their actual protonation state. For the calculation of $\Delta G_{\mu}^{titrate}$ an interaction matrix ΔΔW is build up showing the energy of all pair-wise interactions between titratable groups. In principle, all possible combinations of protonation states in the protein would need to be evaluated and the energetically most favourable pattern would be chosen. However, this combinatorial space increases exponentially with the number of titratable groups and is computationally too demanding even for a protein of moderate size. Therefore, the utilized software karlsberg+ [3, 4] incorporates a Metropolis-Monte-Carlo (MMC) algorithm that iteratively changes the protonation pattern based on an energy evaluation and a probability factor. Usually, this algorithm finds the energetically most favourable protonation pattern and converges several orders of magnitude faster than a rigorous computation considering all possible protonation states.

pK_A values are typically calculated for one protein conformation, e.g. from the crystal structure. In many cases, this procedure is sufficient to obtain reasonable results. However, several residues undergo a local or induce even a global change in the protein structure upon changing their protonation state. karlsberg+ may also take this structural change into account. When the pK_A calculation is based on several alternative protein structures, which may be obtained e.g. from modelling or molecular dynamics simulation, an additional conformation-dependent term is added to the energy calculation. The overall electrostatic energy may be calculated for each conformation and the additional energy term ΔG_{conf} weights for the relevance of this conformation.

pK_A values computed for specific conformations may deviate from the measured value. In a typical pK_A measurement, a population of protein molecules is titrated all being in their individual conformation and the measured pK_A value is an average over all molecules. In

contrast, theoretical pK_A calculations may accurately determine the pK_A for specified conformations, but the amount of proteins in a certain conformation is mostly unknown. When an average pK_A should be calculated, there are ways to overcome this drawback; e.g. via sampling of conformations. Nevertheless, for certain questions of interest, this drawback of theory may become an advantage. To characterize the energetics of a proton within a reaction pathway, it is often more useful to determine the pK_A of a residue in a specific conformation that only arises under non-equilibrium conditions than having an averaged pK_A value. This conformation-dependent non-equilibrium pK_A will be called "action pK_A" in the following.

1.3 Molecular Dynamics Simulation

Molecular dynamics (MD) simulation may imitate the real-time movement of molecules on an atomistic level. In 2013, the most recent Nobel prize in chemistry was awarded to Martin Karplus, Michael Levitt and Arieh Warshel for their work in developing and extending the method of MD simulation. MD simulation incorporates Newton's classical physics on an atomistic level. Atoms and bonds are represented by balls and springs that swing and propagate their movement based on a force field. This force field is composed of bonded and non-bonded interactions. The bonded interactions may be divided into interactions of bonds, angles, dihedrals, and improper dihedrals, where the latter is used e.g. for ring planarity and chiral carbon atoms. The non-bonded interactions are composed of electrostatic and Van-der-Waals interactions, while the latter is approximated with the Lennard-Jones potential. Thus, the potential energy of a molecular system is the sum over all bonded and non-bonded interactions between all atoms:

17.

$$E_{pot} = \sum E_{bond} + \sum E_{angle} + \sum E_{dihedral} + \sum E_{improper} + \sum E_{elec} + \sum E_{VDW}$$

This potential energy function may be minimized to find a more stable conformation or it can be combined with a kinetic energy component to simulate the actual dynamic of molecules. For this purpose, kinetic impulses are introduced randomly onto the atoms to simulate molecular movement at a certain temperature, which is then propagated based on the kinetic and potential energy. To ensure stability of the system, the algorithm keeps certain parameters constant by adjusting them at a defined frequency of simulation steps, while one constant parameter in classical MD is the number of atoms (N). There are three main ensembles of MD settings with constant parameters, (i) the canonical ensemble (NVT) using volume and temperature, which is equivalent to the kinetic energy, (ii) the microcanonical ensemble (NVE) using volume and energy, which is the sum of potential and kinetic energy, and (iii) the isothermal-isobaric ensemble (NPT) using pressure and temperature. In all our MD simulations, the NPT-ensemble was used.

A critical factor in MD simulation is the long-range interaction and how far it propagates or how interactions are treated at the boundaries of the system. For non-bonded interactions, a certain cut-off is commonly defined, behind which the interaction is neglected. A switching

function is introduced to gradually turn off interactions. The system is commonly simulated in a rectangular box with periodic boundary conditions. Practically, all atoms leaving the simulation box at one side are introduced at the other side. This method may account for reasonable boundaries, if the simulation box is large enough that there is no significant self-interaction of the simulated molecule at the boundaries.

Parameterization in MD aims mainly at highest possible agreement with experimental or quantum mechanical data and secondly tries to increase computational efficiency. Parameters for increased efficiency should be chosen not to affect the simulation results and are not further discussed in this work. Parameterization of the force field is a demanding task that requires adjustment of simulation results to quantum mechanical or experimental data, which is mainly generated by vibrational spectroscopy. There exist several approaches to generate a force field for MD simulation, here I use the CHARMM force field [5-8], which is one of the most common force fields. In this force field, there exist force constants and equilibrium values for all bonded and non-bonded interactions. However, the present available parameterization is limited to the most abundant molecules, as e.g. standard amino acids, simple sugars or lipids. Exceptional cofactors, as e.g. the heme-copper center in cytochrome *c* oxidase, are not parameterized in the available CHARMM force field [5-8]. Therefore, one part of my work was to parameterize the force field for these cofactors and ligands. Bonded parameters were adopted from similar atom types out of the CHARMM force field [5-8] and partial charges were generated by quantum chemistry being done in collaboration.

1.4 Selectins

The immune system protects the body against invading microorganisms and toxic substances. For this purpose, immune cells, called leukocytes, are constantly travelling throughout the body via the blood stream to sites of infection or specialized tissue of the immune system. For leaving the blood stream, leukocytes need to attach to, roll on and then traverse the blood vessel wall, a process that is called extravasation. Extravasation is initiated by binding of selectins to their counter-receptors (C-Rs) that causes the leukocytes to tether to and start rolling on the blood vessel wall (Fig. 2A). Thus, this initial selectin-C-R binding is a key event in directing leukocytes to their correct destination and selectin malfunctioning may cause severe diseases [9]. Up-regulated selectin expression can lead to autoimmune diseases and enhances tumor cell metastasis, while down-regulation may be the cause for recurrent bacterial infection and persistent inflammation. Therefore, selectin binding is an interesting target in drug development [10], which requires further understanding of the binding event on a molecular level.

The interaction between selectin and its C-R is exceptional in several aspects making it well-suited for its specific function; (i) the binding free energy is dominated by electrostatic interaction, not by hydrophobic effects as many other protein-protein interactions, making it a promising candidate for my investigation using electrostatic calculations, (ii) due to the electrostatic effects, the selectin C-R attraction has a long range, which is needed for the initial tethering of leukocytes to the blood vessel wall [11], (iii) the on- and off rates of selectin C-R binding are remarkably fast (up to $k_{on} = 10^6\ M^{-1}s^{-1}$) [11] enabling the leukocyte to roll, while constantly forming and breaking selectin-C-R complexes, (iv) binding free energy of a single selectin-C-R complex is relatively weak (below 50 kJ/mol [12-15]), but a leukocyte presents up to thousands of selectin molecules that can simultaneously interact with their C-R creating a multivalent binding effect that is sufficiently strong, (v) in contrast to regular slip bonds, selectin binding exhibits a catch bond behaviour [11], meaning that binding becomes more strong when binding partners are pulled apart, e.g. under shear flow in the blood stream. The most popular theories ascribe the catch bond behaviour either to an allosteric change of the selectin molecule or to recurrent sliding-rebinding events triggered via pulling [11].

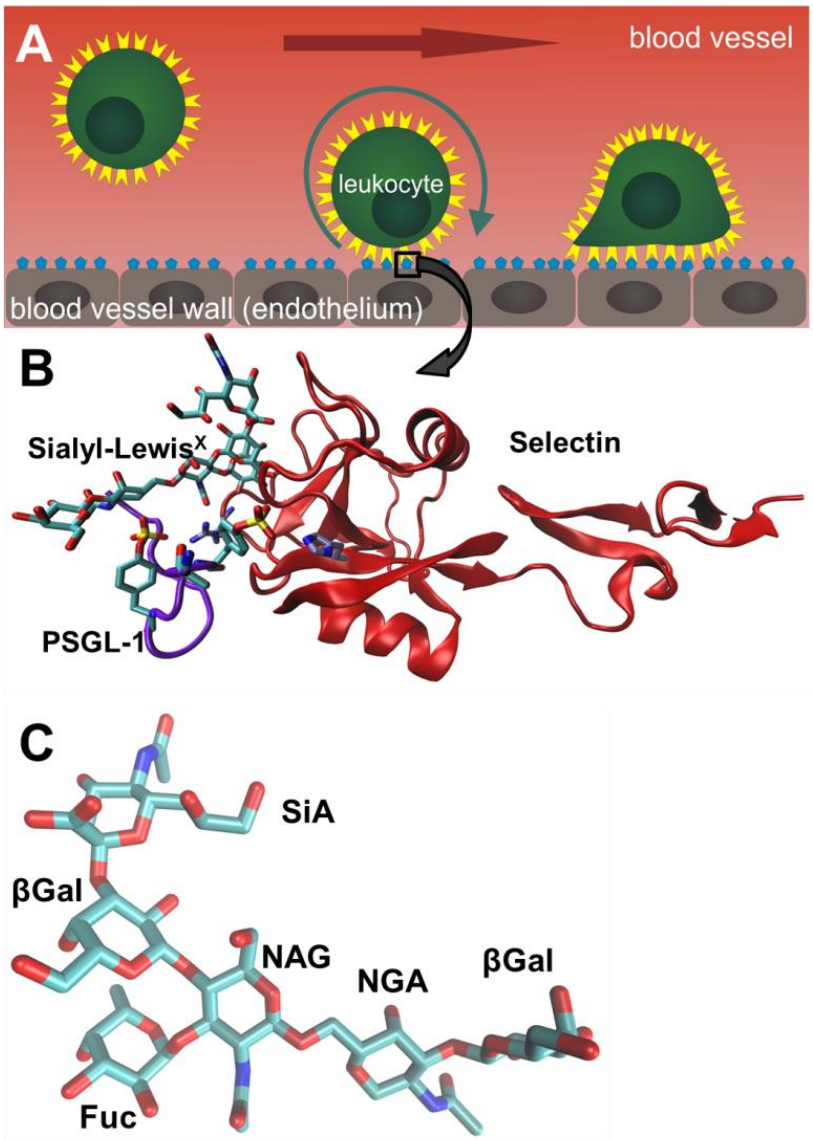

Figure 2: Extravasation process initiated by selectins binding to their C-R. (A) In the extravasation process, leukocytes tether and adhere to the blood vessel wall and then leave the blood stream into the tissue. The first step in extravasation is mediated by selectin-C-R binding. (B) Crystal structure (pdb-id: 1G1S [16]) of the co-crystallized complex of P-selectin (red) and the PSGL-1 ligand (purple rubber band and stick model for SLe^X). Selected residues important for binding are shown as stick model. (C) Sialyl-Lewis-X (SLe^X) is part of the six-subunit glycosylation of PSGL-1: two β-galactose (βGal), N-acetylglucosamin (NAG), N-acetylgalactosamin (NGA), fucose (Fuc) and sialic acid (SiA). SLe^X is a four-subunit carbohydrate composed of SiA, βGal, Fuc and NAG.

The selectin protein family consists of three types, platelet and endothelial cell (P), leukocyte (L), and endothelial cell (E) selectin. P-selectin is usually not presented on the blood vessel wall under resting conditions, but stored in Weibel-Palade bodies of endothelial cells and in α-granules of platelets and is only transported to the cell surface after pro-inflammatory stimulation. In contrast, E-selectin is newly synthesized in endothelial cells upon stimulation, which is regulated on the transcriptional and translational level. Only L-selectin is expressed on the counter-part of the blood vessel wall, which is the leukocyte. L-selectin mediates leukocyte rolling and neutrophil activation and may be shed from the leukocyte cell surface upon stimulation (for review see [11, 17, 18]).

All selectins exhibit the same domain structure and show high sequence homology [9]. Selectins are C-type lectins, which are characterized by the ability to bind specific carbohydrates in a Ca^{2+}-dependent fashion. The N-terminal lectin domain of selectins (Fig. 2B) includes Ca^{2+} and reaches out in the extracellular space being responsible for binding to the C-R. The lectin domain is followed by an EGF domain and two to nine consensus repeats, then selectins have a transmembrane part and a short cytoplasmic domain, which likely has signalling function [9].

The C-Rs of selectins are characterized by carrying certain carbohydrate motifs, mostly consisting of Sialyl-Lewis-X (SLe^X) [17]. SLe^X is a tetrasaccharide (Fig 2C) composed of fucose (Fuc), N-acetylglucosamine (NAG), β-galactose (βGal) and sialic acid (SiA). It may be attached posttranslational to protein scaffolds that in conjunction with SLe^X form the C-R for selectins. There exist several C-Rs with probably different function and distribution, but only P-selectin glycoprotein ligand 1 (PSGL-1) has been extensively characterized including a crystal structure of the P-selectin-PSGL-1 complex (Fig. 2B) [9]. PSGL-1 is a large homodimer being expressed on most leukocytes and also on certain endothelial cells with chronic inflammation [18]. Besides carrying a large amount of O-glycans, PSGL-1 undergoes another posttranslational modification, which is sulfation of tyrosines (TyS). PSGL-1 is not only responsible for ~90 % of P-selectin binding *in vivo*, but it is also the most important C-R for L-selectin in inflammatory settings and may bind to E-selectin as well [9]. This overlap in selectin binding specificity is not surprising when looking at the three-dimensional structure, which is highly similar in the C-R binding lection domain [16]. Nevertheless, there are distinct differences in affinity of PSGL-1 binding for all three selectins. For properly

glycosylated and sulfated PSGL-1, P-selectin is the strongest binder and E-selectin the weakest, while on opposite SLe^X alone is bound most strongly by E-selectin and most weakly by P-selectin [12-15]. Understanding these differences in binding affinity on a molecular level is an important step in exploring the complex network of selectin-C-R interactions and it is the prerequisite to search for molecules that modulate selectin-C-R binding and may serve as drug candidates.

1.5 Cytochrome *c* Oxidase

Cytochrome *c* oxidase (C*c*O) is the terminal enzyme of the respiratory chain in mitochondria and a central component of the aerobic metabolism in many bacteria. This membrane-spanning protein catalyses the reduction of molecular oxygen to water and utilizes the resulting energy to pump protons and build up an electrochemical gradient across the membrane. In the C*c*O reaction cycle, electrons are taken up from soluble cytochrome *c* at the positively charged P-side of the membrane. They are transported via the bimetallic copper A center (Cu_A) to heme a and finally to the binuclear center (BNC), which is composed of heme a_3 and copper B (Cu_B). The protons required for C*c*O function are taken up from the negatively charged N-side of the membrane. They can be differentiated into chemical protons consumed in the reaction and pumped protons, which are transported across the membrane. The existence of a proton loading site (PLS) is widely accepted, which is a titratable group that drastically changes its pK_A during the reaction cycle and thereby accomplishes proton pumping. The identity of the PLS remains elusive, propionic acids of heme a_3 or one of the Cu_B-ligating histidines have been suggested.

In the BNC, electrons and chemical protons are transferred onto molecular oxygen gaining two water molecules after one reaction cycle (Fig. 3). The state of the BNC with one oxygen molecule bound is called compound A, which converts to state P_M after the oxygen-splitting reaction. In each of the following four steps, one electron and one proton are consumed in the BNC and up to one proton is pumped, depending on the type of C*c*O and the reaction conditions (see next paragraph). The steps from P_M-state via F-state to O_H-state are called the oxidative half of the cycle and the steps from O_H-state via E_H-state to R-state are called the reductive half. In the R-state, C*c*O is ready to take up a new oxygen molecule to form compound A (for review see [19, 20]).

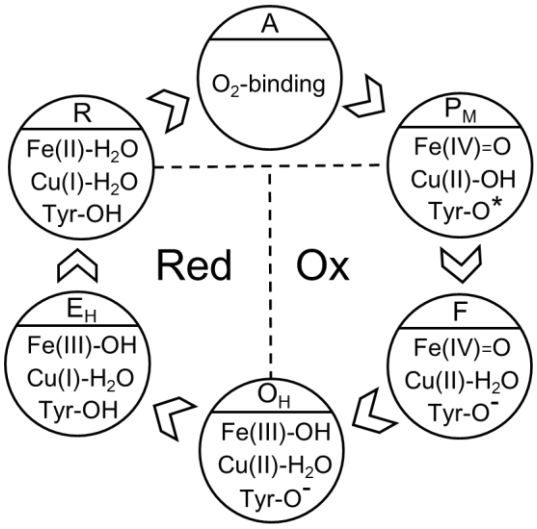

Figure 3: Catalytic cycle of C*c*O. Molecular oxygen may bind to the reduced state R forming compound A and is split forming state P_M. Then, the reaction proceeds in four steps, in all of which one electron and one chemical proton is taken up and one proton is pumped. In the oxidative first half (Ox) of the cycle, chemical protons are provided by the D-channel, while in the reductive second half (Red), the K-channel transports the chemical protons. Cycle adapted from [20].

The C*c*O protein family is divided in three types, A-, B-, and C-type. The most abundant A-type C*c*O is constitutively expressed in mitochondria and several bacteria [19, 20], whereas B- and C-type C*c*O are expressed in specialized bacteria under low oxygen conditions [21]. One well-studied example of A-type C*c*O stems from the bacterium *Rhodobacter sphaeroides*, which was also investigated in this study. This enzyme is composed of four subunits (Fig. 4), however the two subunits A and B retain full functionality [22] and therefore only this two-subunit enzyme was investigated here to increase computational efficiency. A-type C*c*Os have a pumping stoichiometry of 1 H^+/e^- [20]. The most studied B-type C*c*O is expressed in the thermophilic bacterium *Thermus thermophilus* being composed of three subunits and heme b instead of heme a. B-type C*c*Os presumably pump with a stoichiometry of 0.5 H^+/e^- [21], but there is evidence that they may also pump 1 H^+/e^-, when the proton gradient is eliminated [23]. C-type C*c*O is more distant to the other two types and will not be discussed in this study.

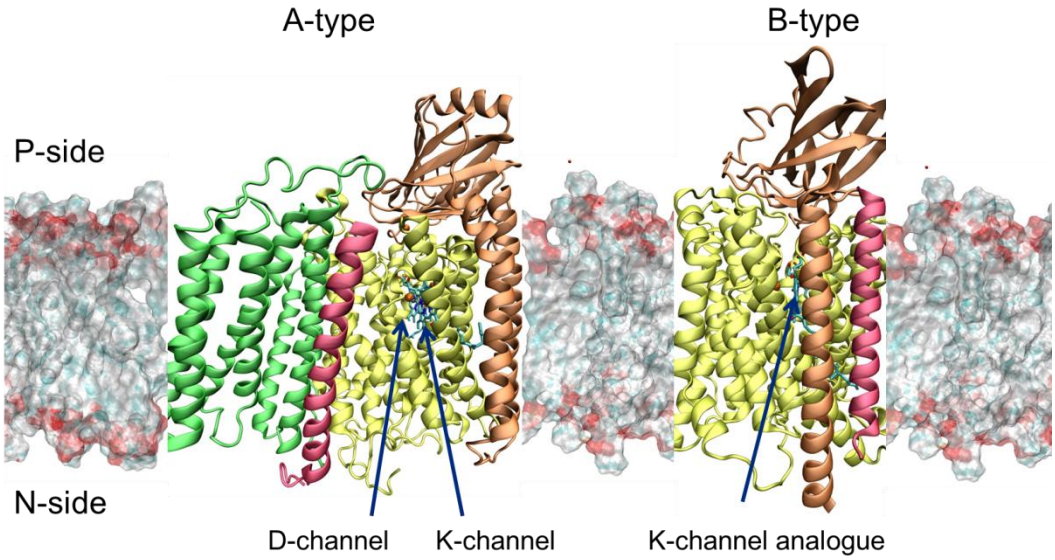

Figure 4: C*c*O of A- and B-type. The four-subunit A-type C*c*O from *Rhodobacter sphaeroides* (pdb-id: 1M56 [24]) is composed of chain A (yellow), B (orange), C (green) and D (red). The three-subunit B-type C*c*O from *Thermus thermophilus* (pdb-id 3S8F [25]) is composed of chain A (yellow), B (orange) and IIA (red). For both enzymes, all redox-active cofactors are located in chain A, except the Cu_A center being located in chain B. The approximate location of the proton input channels is indicated by arrows.

One of the most striking differences between A- and B-type C*c*O is the composition of the proton input channels. In A-type C*c*O, the protons are conducted via two different pathways – the D-channel named after its entrance residue aspartate D132 and the K-channel containing the central residue lysine K362. The D-channel is filled with several water molecules and ends at a conserved glutamate Glu286 (Fig. 5A) [26]. Glu286 points toward the D-channel in all crystal structures [26-28] and is situated between the hemes and above Glu286 there is a hydrophobic cavity, which may connect Glu286 to Cu_B of the BNC and to propionate D of heme a_3 (PRD_{a3}).

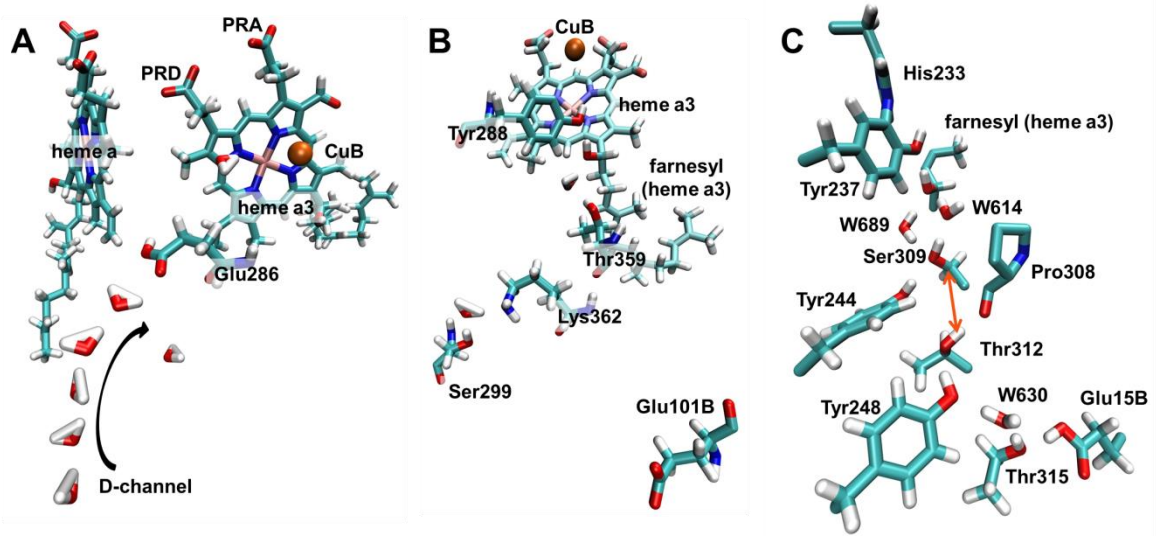

Figure 5: Proton input channels of A- and B-type C*c*O. (A) End of the D-channel of A-type C*c*O showing the environment of Glu286 with heme a and heme a_3. (B) The K-channel of A-type C*c*O starts possibly at Glu101B, runs via Ser299, Lys362 and Thr359 to the farnesyl OH-group and ends at Tyr288. Only two water molecules are resolved inside the K-channel. (C) The K-channel analogue of B-type C*c*O runs from Glu15B to Tyr237 with a continuous H-bond chain. It has only one gap of 4.9 Å in the H-bond chain between Thr312 and Ser309, which is indicated by an orange arrow. Geometries are taken from *R. sphaeroides* for A-type C*c*O (pdb-id 2GSM [26]) and from *T. thermophilus* for B-type C*c*O (pdb-id 3S8F [25]).

In contrast, the K-channel appears disconnected in the crystal structure (Fig. 5B) [26]. The channel was proposed to start at Glu101B [29], which is 12 Å away from the central Lys362. Lys362 has a narrow hydrophobic environment and points to the N-side. Lys362 is at 14 Å distance to Tyr288, which is part of the BNC. Tyr288 is covalently bound to His284 one of the Cu_B ligands and it constitutes the end of the K-channel. Between Lys362 and Tyr288 there is only one threonine Thr359 and one water molecule observed in the crystal structure (Fig. 5B) [26]. It is well established that all pumped protons are taken up via the D-channel, while two chemical protons enter via the D-channel and two via the K-channel [20]. In contrast, the B-type C*c*O was shown to transport all chemical and physical protons via the same input channel (Fig. 5C) [30], which is located at a similar position as the K-channel of the A-type C*c*O [25, 26], without sharing significant sequence homology. This K-channel analogue in B-type C*c*O is composed of a continuous H-bond chain of threonines, tyrosines and serine and a water molecule with only one 4.9 Å gap in between Thr312 and Ser309 (Fig. 5C). Only

glutamate and tyrosine at entrance and terminus, respectively, are conserved among the K-channels of A- and B-type C*c*O.

In summary, for A-type C*c*O all pumped protons are transported via the D-channel and the chemical protons are splitted, two via the D-channel and two via the K-channel, while in B-type C*c*O all chemical and pumped protons are transported via the K-channel analogue. This distinct proton distribution and their transport against a gradient must be regulated in terms of timing and direction. When starting this work, one proton gating element had been discovered being Glu286 inside the D-channel of A-type C*c*O [31, 32]. In MD simulation, Glu286 had been shown to adapt two different conformations [31, 32], either down-position toward the D-channel as in the crystal structures or up-position pointing toward the P-side. The proposed gating function of Glu286 was based on the rapid down-flipping of deprotonated Glu286 in MD simulation [31, 32]. This would hinder Glu286 from receiving protons from upwards and let them leak back toward the D-channel. But this gating function had been discussed controversially, the groups of Wikström [31] and Voth [32] demonstrated this gating behavior in MD simulations, while Cui and coworkers [33] doubted the rapid down-flipping of deprotonated Glu286 and observed even bridging positions, where Glu286 connects the upward cavity with the downward D-channel. Therefore, the exact function of Glu286 required further investigation. The K-channel in A-type C*c*O and the K-channel analogue in the B-type were not well characterized when starting this work.

The aim of this work is to identify and characterize the proton gating elements of the different channels in A- and B-type C*c*O.

1.6 Aim of this work

The aim of this work is to use electrostatic calculations to investigate two biologically relevant questions, one in the field of immunology, one in the field of energy metabolism.

The immunological study addresses the binding and complex formation of selectins and their C-Rs. The first step when leukocytes are leaving the blood stream – a process called extravasation – is mediated via binding of selectins to their C-Rs. Malfunctioning of selectins may lead to chronical infection, autoimmune diseases or cancer; therefore the selectin binding specificity is of great medical interest. Besides few point mutations known from humans, the complex formation of selectins with their C-Rs had not been characterized on a molecular level before this study. Here, I aim for an analysis of the different molecular components involved in selectin binding. There are three types of selectins – P-, L-, and E-selectin – binding with different affinity to their C-R PSGL-1. PSGL-1 is composed of a peptidic part including sulfated tyrosines and a carbohydrate part called Sialyl-Lewis-X. In the present study, it is investigated what is the contribution of electrostatic interaction to binding affinity of different amino acids and carbohydrate subunits of the selectins and PSGL-1. Also, the difference in binding affinity of the three selectins shall be explained. This systematic analysis of selectin complex formation is the basis for further binding studies and possible drug design attempts.

The study concerning energy metabolism focusses on the proton transport mechanism of C*c*O. C*c*O is the final enzyme of the aerobic chain and it catalyses the reaction of molecular oxygen to water. The reaction energy is used to pump protons across the membrane that C*c*O resides in thereby establishing an electrochemical gradient. The conversion of energy by C*c*O is more efficient than any man-made machine; thus it is desirable to understand the mechanism of C*c*O not only for biological interest, but also from a technological point of view.

For the reaction cycle of C*c*O, so-called chemical protons are consumed in the reaction with molecular oxygen and pumped protons are transported across the membrane. There exist defined proton input channels inside C*c*O that should fulfil the functions of controlled access to the reaction center and ensuring unidirectional proton transport. The location of the proton input channels is known from crystallographic and mutational experiments. In A-type C*c*O, the D-channel serves all pumped and two chemical protons and the K-channel transports the two remaining chemical protons, while in B-type C*c*O the K-channel analogue is used for all

chemical and pumped protons. However, the mechanism of regulated proton transport inside the channels was only partially known, when starting this work. One gating element was proposed to be glutamate 286 at the end of the D-channel, however, being controversially discussed in literature. Gating elements of the K-channel and the K-channel analogue were not identified. The aim of this work is to identify and characterize the gating elements of all three channels of A- and B-type C*c*O, which is an important step towards elucidating the full reaction mechanism of C*c*O.

2. Publications

2.1 Selectin Binding to Their Counter Receptor

Publication: A.L. Woelke, C. Kuehne, T. Meyer, G. Galstyan, J. Dernedde, E.W. Knapp, Understanding Selectin Counter-Receptor Binding from Electrostatic Energy Computations and Experimental Binding Studies, J. Phys. Chem. B, 117 (2013) 16443-16454.

doi-link: http://dx.doi.org/10.1021/jp4099123

Own contribution:

- Development of research question
- Force field parameterization of complex sugars and sulfated tyrosine
- Conduction of all electrostatic and pK_A calculations
- Analysis of results
- Preparation of manuscript

Other contributions:

- Supervision of project and publication (Ernst-Walter Knapp)
- Quantum chemical calculations of partial charges (Gegham Galstyan)
- Technical assistance with software karlsberg+ (Tim Meyer)
- Experimental validation (Christian Kühne and Jens Dernedde)

In this publication, the selectin binding specificity is characterized on an atomistic level by calculating the electrostatic binding energy of selectin-C-R complexes. The three selectin types, P-, L-, and E-selectin all bind to the glycosylated and sulfated PSGL-1 and also to the carbohydrate structure SLe^X alone, but with different affinity [12-15]. Explaining these differences by the different amino acid composition of the three selectins is the aim of this study.

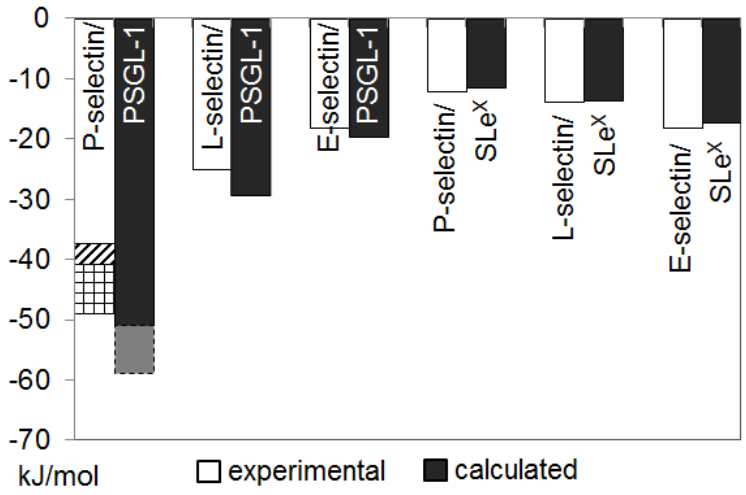

Figure 6: Binding free energies for six selectin ligand complexes in units of [kJ/mol]. Comparison of experimental (white bars) and theoretical values (black bars) based on electrostatic energy computations for the three types of selectins with the PSGL-1 ligand or only SLe^X. Experimental values are taken from literature [12-15], calculations are based on static structures being modelled as explained in the publication.

All my electrostatic binding calculations are based on atom coordinates of the selectin-C-R complex from crystallography if available. For all selectin types, crystal structures of the lectin domain are available [16, 34], but only P-selectin was co-crystallized with the tail of PSGL-1 [16] (PSGL-1 ligand in the following). Also, crystal structures of P- and E-selectin soaked with SLe^X are available [16]. All missing complexes were modelled by homology modelling, which is a very suitable technique in this case as the lectin domains of all three selectins share high three-dimensional similarity. For all given complexes, the electrostatic binding energy was calculated from the energy of the complex and the energy of selectin and C-R alone using a variant of the thermodynamic cycle (Fig. 1). This technique was first applied to the wild-type structures and compared to experimental data from literature [12-15] to confirm that the approach works appropriately. Indeed, calculated and experimental binding energies are in good agreement (Fig. 6). In the complex of P-selectin with PSGL-1 ligand, a protonation change is observed for $His114_{P\text{-selectin}}$, which is deprotonated in the unbound state, but must protonate for binding to PSGL-1 ligand. This energy for protonation must be spent upon binding and therefore it is subtracted from the binding energy of this complex.

To analyse the different components of binding the wild-type complex structures were modified by dividing the SLe^X ligand into different parts or by introducing alanine mutations either in selectin or in the PSGL-1 peptide. I could show that the Fuc subunit is most important for SLe^X binding in all selectins, while the SiA subunit is bound most strongly by L-selectin. E-selectin has a glutamate at position Glu107, while L-selectin exhibits an aspartate Asp107, and I found that this is the main reason that the latter exhibits weaker SLe^X binding. For binding of the PSGL-1 ligand by P- and L-selectin, all three TyS are important, while TyS48 and TyS51 contribute the most to binding energy. In contrast, E-selectin does not bind to the peptidic part of the PSGL-1 ligand because of its missing positive charge, which is in agreement with experiment and explains why E-selectin is the weakest PSGL-1 binder. In addition, I could show that the soaked crystal structures do not exhibit the appropriate conformation for SLe^X binding. Only in the co-crystallized P-selectin-PSGL-1 structure, selectin is in the correct conformation for SLe^X binding.

Inspired by calculation, selectin mutants were designed and their binding affinity was tested with surface plasmon resonance (SPR) measurements. All theoretical predictions could be confirmed by SPR.

2.2 Characterizing Glu286 inside the D-Channel of A-Type C*c*O

Publication: A.L. Woelke, G. Galstyan, A. Galstyan, T. Meyer, J. Heberle, E.W. Knapp, Exploring the Possible Role of Glu286 in CcO by Electrostatic Energy Computations Combined with Molecular Dynamics, J. Phys. Chem. B, 117 (2013) 12432-12441.

doi-link: http://dx.doi.org/10.1021/jp407250d

Own contribution:

- Development of research question
- Force field parameterization of cofactors
- Conduction of all MD simulations, electrostatic and pK_A calculations
- Analysis of results
- Preparation of manuscript

Other contributions:

- Supervision of project and publication (Ernst-Walter Knapp)
- Support in design of the study and analysis of results (Joachim Heberle)
- Quantum chemical calculations of partial charges (Gegham and Artur Galstyan)
- Technical assistance with software karlsberg+ (Tim Meyer)

In this publication, the role of Glu286 as a potential gate inside the D-channel of C*c*O is investigated. The literature concerning that topic is contradictory [31-33], thus this work intends to clarify whether Glu286 does have gating function.

MD simulations and electrostatic or pK_A calculations were performed. Before starting the MD, the system had to be set up by parameterizing the cofactors of C*c*O and embedding the protein from crystal structure [26] into an explicit membrane. MD simulations gained reasonably low RMSD values of below 1.1 Å for backbone atoms compared to the crystal structure.

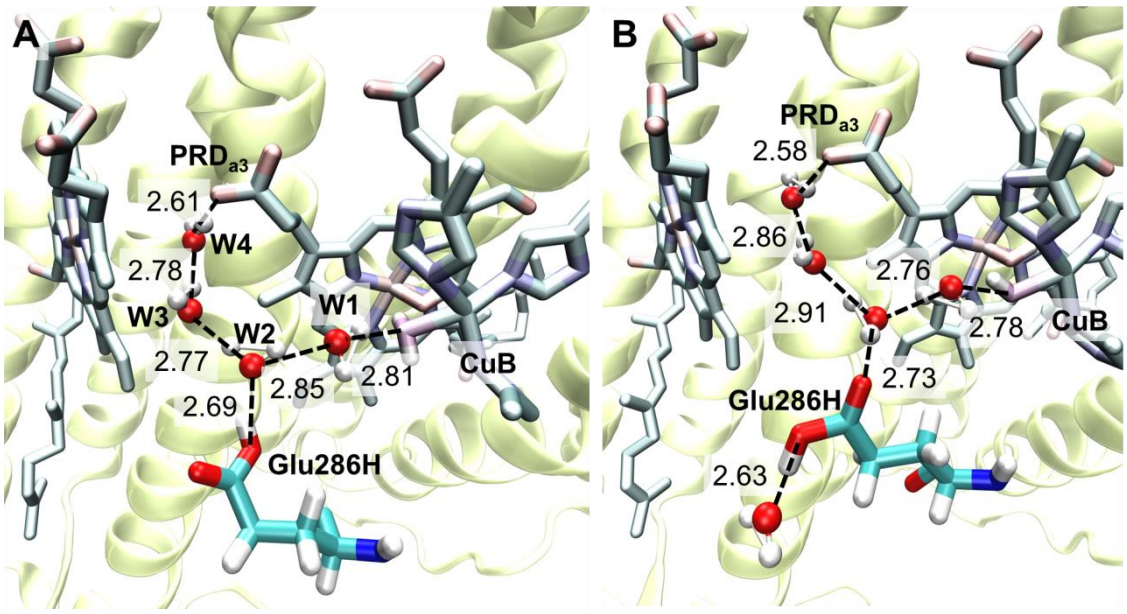

Figure 7: Glu286H in proton-conducting and leaking position. (A) Modelled position of Glu286H pointing upwards to the hydrophobic cavity, which accommodates four modelled water molecules. This conformation is well suited to transport protons to the BNC or to the potential transient PLS PRD_{a3}. The water chain is stable for 2.4 ns in MD simulation. (B) Conformation observed in MD simulation, where Glu286H is in H-bonding distance to the D-channel and to the hydrophobic cavity at the same time having its proton pointing downwards. This conformation is vulnerable to proton-leakage towards the D-channel.

Different initial setups were chosen for MD simulation. In the crystal structure [26], Glu286 points to the D-channel and the hydrophobic cavity above is empty (Fig 5A). However, other studies indicated that the hydrophobic cavity might be dried out and accommodate a water chain in the native enzyme [33] and that Glu286 can also point upwards to the hydrophobic cavity [31]. The upward position of Glu286 and the water chain in the cavity composed of four water molecules were modelled (Fig. 7A). Water positions are verified by calculating their favourable electrostatic energy compared to bulk water. Using this setup, different combinations were examined of a modelled water chain, Glu286 positions and protonation states of Glu286 or protonation states of PRD_{a3} and PRA_{a3} both being a potential PLS.

MD simulations reveal that deprotonated Glu286 (Glu286-) is repelled by the negative charge of PRD_{a3}- and flips back to the D-channel, as had been observed before [31, 32]. When protonating PRD_{a3} this repulsion is not observed and Glu286- remains in an upward position,

which could lead to proton leakage. However, when PRA_{a3} is protonated instead, Glu286- again flips down rapidly. Thus, leakage could only occur when the proton is expelled from PRD_{a3}, but not when it is transported further to PRA_{a3}.

The modelled water chain inside the hydrophobic cavity is stable for up to 2.4 ns only with protonated Glu286 (Glu286H) in an upward position (Fig. 7A) and for Glu286- the water chain loses H-bonding contact immediately. Thus, only Glu286H has a stable H-bond connection to the BNC and to PRD_{a3}.

The pK_A of Glu286 under physiological conditions is calculated to be above 12, but this value decreases to about 9 under increased water accessibility of Glu286, which is in accordance with experiment [35]. Therefore, Glu286- will reprotonate soon after donating the proton, which may help to prevent proton leakage.

Nevertheless, I also observed a conformation of Glu286H with H-bond connection to the hydrophobic cavity and the D-channel at the same time. Also, Glu286H has its proton pointing down to the D-channel (Fig. 7B), which could lead to back-transport of a proton.

My conclusion from this study is that Glu286 does not constitute the proton gating element of the D-channel in A-type C*c*O alone, but it is only a part of it. The high pK_A of Glu286, its flipping behaviour and its effect on the water chain of the hydrophobic cavity will surely support unidirectional proton transport, but also the hemes with their propionic acids are needed to prevent proton leakage completely.

2.3 Characterizing Lys362 inside the K-Channel of A-Type C*c*O

Publication: A.L. Woelke, G. Galstyan, E.W. Knapp, Lysine 362 in Cytochrome c Oxidase Regulates Opening of the K-Channel via Changes in pKA and Conformation, Biochim. Biophys. Acta, Accepted for publication, August 2014.

doi-link: http://dx.doi.org/10.1016/j.bbabio.2014.08.003

Own contribution:

- Development of research question
- Force field parameterization of cofactors
- Conduction of all MD simulations, electrostatic and pK_A calculations
- Analysis of results
- Preparation of manuscript

Other contributions:

- Supervision of project and publication (Ernst-Walter Knapp)
- Quantum chemical calculations of partial charges (Gegham Galstyan)

In this publication, the mechanism of the K-channel in A-type C*c*O is investigated by MD simulation and pK_A calculation. The setup is similar to the parameterization from the previous study about C*c*O concerning the D-channel. When starting this work, there was little knowledge about the K-channel. Experiments had shown that the two chemical protons in the reductive half of the catalytic reaction cycle are conducted via the K-channel and that Lys362 inside the channel is essential for its function [36, 37]. In the crystal structure [26], the K-channel appears disconnected having Glu101B at the entrance, Tyr288 at the end, and Lys362 in the middle with a very hydrophobic environment and without continuous H-bond connection throughout the channel (Fig. 5B). Lys362 points downward to the N-side in all crystal structures [26-28], but there was experimental evidence that Lys362 may move upward in certain states of the reaction cycle [38, 39], e.g. in P_R state, which is observed from a fully reduced C*c*O.

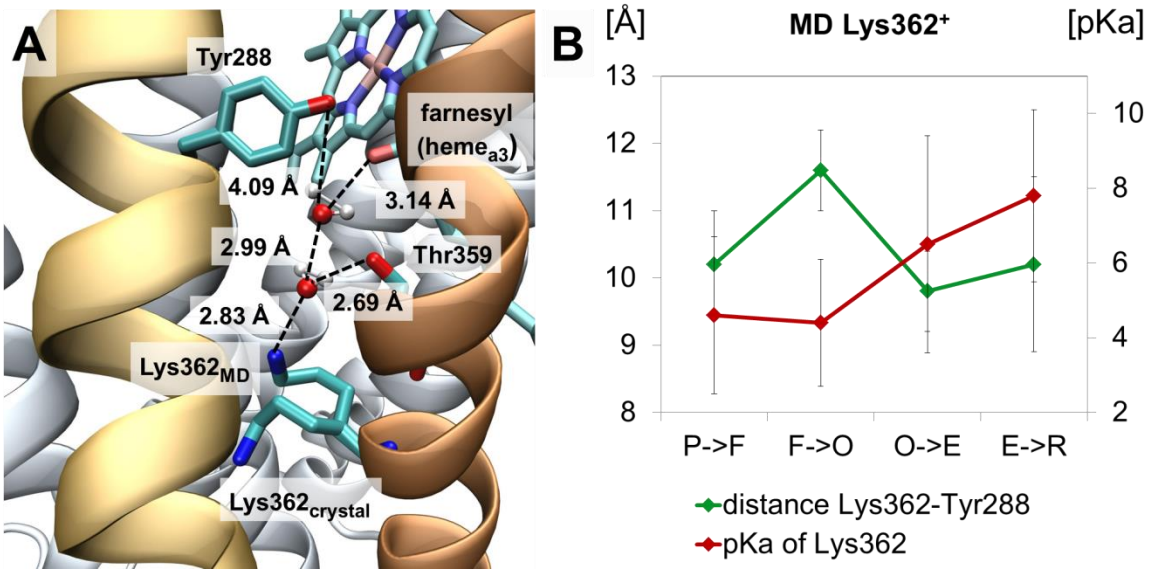

Figure 8: Lys362^{+} changes conformation and pK_A during the catalytic cycle. (A) Lys362 is pointing downwards in the crystal structure (pdb-id: 2GSM [26]), but moves upwards and becomes connected to the BNC via water molecules, when simulated as protonated Lys362. Coordinates of Lys362$_{down}$, cofactors and backbone atoms are taken from crystal structure (2GSM), coordinates of Lys362$_{up}$ and the two water molecules are taken from an MD snapshot (O->E-state simulation after 9.3 ns). (B) Average distance between Lys362 and Tyr288 and average pK_A of Lys362 for MD simulation with protonated Lys362. Averages are taken over 50 ns simulations, standard deviations are indicated by error bars.

In this study, MD simulations explicitly demonstrate that Lys362 moves up towards the BNC when Lys362 is protonated. This movement decreases the distance between Lys362 and Tyr288 being the end of the K-channel from 14 Å to up to 9 Å (Fig. 8B). Protonated Lys362 also becomes H-bond connected to Tyr288 via two water molecules (Fig. 8A). In contrast, deprotonated Lys362 remains at the crystal structure position pointing down towards the N-side. To determine under which circumstances Lys362 may protonate, its pK_A value was calculated from MD frames. While the Lys362 pK_A values from MD with deprotonated Lys362 are far below seven, values from MD with protonated Lys362 may reach above seven depending on the state in the catalytic cycle of C*c*O. The Lys362 pK_A is below five for the transition states of the oxidative half of the reaction cycle (states P->F and F->O), but it is around seven for the transition states of the reductive half (states O->E and E->R) (Fig. 8B).

From these results, I developed a model explaining how transport of the chemical protons via two different channels is regulated. In the oxidative half of the reaction cycle, Lys362 has a pK_A below five, therefore it cannot protonate and the K-channel remains closed; the chemical protons are transported via the D-channel. In the reductive half of the reaction cycle, the pK_A of Lys362 rises to seven or higher and Lys362 may protonate, thereby open the K-channel and the chemical protons are transported via the K-channel.

In this study, I also observed that in the P_R state Lys362 has a high pK_A and thus is protonated, which is in accordance with experimental evidence [38, 39]. Previously, it was thought that the P_R state is equal to the P->F intermediate under physiological conditions. However, this would result in an open K-channel for the P->F transition state. Therefore, I conclude that in the physiological P->F intermediate the additional electron resides on $heme_{a3}$ and not on Cu_B, as in the P_R state.

2.4 Characterizing the K-Channel Analogue of B-Type C*c*O

Publication: A.L. Woelke, A. Wagner, G. Galstyan, T. Meyer, E.W. Knapp, Proton Transfer in the K-Channel Analogue of B-Type Cytochrome c Oxidase from *Thermus thermophilus*, Biophysical Journal, 107 (2014) 2177-2184.

doi-link: http://dx.doi.org/10.1016/j.bpj.2014.09.010

Own contribution:

- Development of research question
- Force field parameterization of cofactors
- Conduction of most MD simulations, electrostatic and pK_A calculations
- Supervision of corresponding diploma thesis
- Analysis of results
- Preparation of manuscript

Other contributions:

- Supervision of project and publication (Ernst-Walter Knapp)
- Preliminary calculations during diploma thesis (Anke Wagner)
- Quantum chemical calculations of partial charges (Gegham Galstyan)
- Technical assistance with software karlsberg+ (Tim Meyer)

In this publication, the properties of the K-channel analogue in B-type C*c*O are investigated. By mutational studies it had been shown that the B-type C*c*O uses only one proton input channel [30], which is the K-channel analogue starting at Glu15B and ending at Tyr237. Therefore, all chemical and pumped protons must be transported via this channel. Besides this information, there was little knowledge about the mechanism of the K-channel analogue when starting this work.

The MD setup was performed analogously to the previous work about the D-channel of C*c*O. MD simulations gained reasonably low RMSD values of below 1.1 Å for backbone atoms compared to the crystal structure.

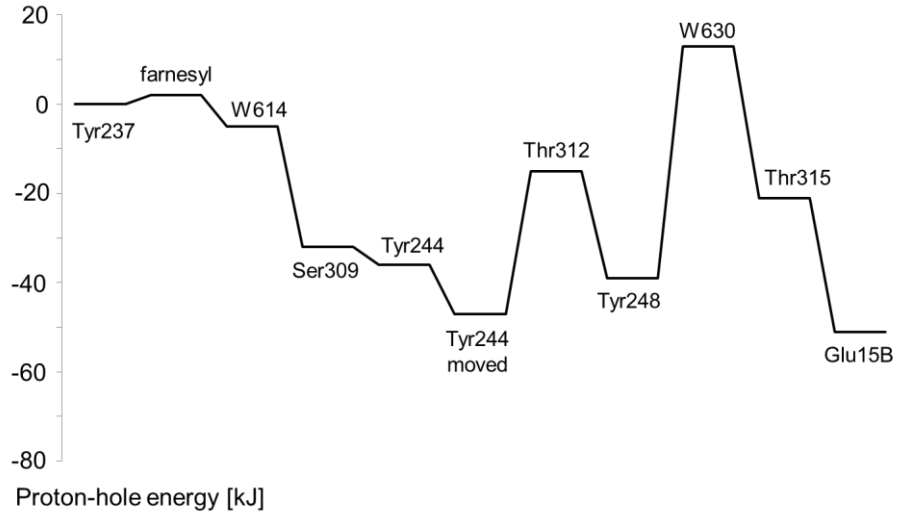

Figure 9: Electrostatic energy landscape of the proton-hole inside the K-channel analogue of B-type C*c*O. Proton-hole transport via the K-channel analogue starts with deprotonated Tyr237, which then receives a proton from the farnesyl-OH-group and the proton hole is transported through the channel to Glu15B. After deprotonation of Tyr244, this residue moves inside the gap between Ser309 and Thr312. The electrostatic energies before and after this conformational change are depicted. Electrostatic energies for all deprotonation steps were calculated using the software karlsberg+ [3, 4] as described in the publication.

The K-channel analogue in B-type C*c*O uses proton-hole transfer for conducting protons, as the proton-wire in the channel is composed of several hydrogen-bonded OH-groups being protonated in the resting-state and shuttling a proton hole when the channel is active. MD simulations were performed for different protonation states of the K-channel analogue to identify the function of the different residues. Glu15B at the entrance of the channel is demonstrated to have the role of a proton reservoir. In the resting state, Glu15B points inward the K-channel analogue H-bonded to Thr315. In this position, Glu15B has a high pK_A, thus it is likely protonated. This conformation is adapted in the crystal structure [25] and it remains stable in MD simulation, when Glu15B is protonated. In contrast, deprotonated Glu15B moves out of the K-channel analogue towards bulk water, where it may receive a new proton. This swinging-out motion promotes entry of new protons into the channel.

Tyr244 in the middle of the K-channel analogue is identified as the valve ensuring unidirectional proton transfer. The only gap in H-bonds of the K-channel analogue is seen

between Thr312 and Ser309 in the crystal structure [25]. In the protonated resting state of Tyr244 it forms a hydrogen bond towards Ser309 without connection to Thr312, which is observed in the crystal structure and in MD simulations with protonated Tyr244. In contrast, deprotonated Tyr244 immediately moves inside the gap between Thr312 and Ser309 and accepts H-bonds from both residues. This behaviour renders Tyr244 a valve for proton transfer as it connects the upper and lower part of the K-channel analogue only after donating its proton upwards to Ser309.

Tyr237, which is covalently bound to one of the Cu_B ligands, is the branching point of the K-channel analogue, where chemical and pumped protons take different routes. Its properties are discussed in the publication and I observed that it is likely protonated in the crystal structure.

Finally, the electrostatic energy landscape for all proton-hole transfer steps in the K-channel analogue was calculated (Fig. 9). This energy calculation demonstrates that proton-hole transport via the K-channel analogue does not exhibit large energy barriers.

3. Conclusion and Outlook

In this work, electrostatic calculations have been successfully applied to different biological questions in the fields of immunology and energy metabolism.

The binding strengths of selectin-C-R complexes were predicted with an error of only few kJ/mol or less.

The effect of several selectin mutants on binding strength was verified experimentally. This work is a valuable basis to understand the selectin-C-R interaction and to use this knowledge for the design of new drugs that control this interaction. In addition, this work provides an example on the applicability of electrostatic calculations for estimating binding strengths of biological complexes. Here, not only protein complexes were investigated, but also the binding affinities of protein-carbohydrate complexes involving SLe^X were determined with high accuracy. This is a difficult task as carbohydrates usually have low binding affinity because of their high solubility. However, when all their hydrogen bonds toward the protein are oriented correctly, it is possible to estimate their binding affinity using electrostatic calculations, as exemplified in this work. Generally, the correct complex geometry is the most critical aspect in electrostatic calculations. This fact may be utilised to identify incorrect geometries, as has been shown here for the soaked crystal structures not being suited for SLe^X binding. Still, the accuracy of electrostatic calculations is limited to the accuracy of the atom positions. One approach to overcome this problem is to use MD simulations for sampling a broader coordinate space. I tested this approach with 10 ns MD simulations for the wild-type selectin-C-R complexes. I could not improve the agreement with experiment, but the results are similar to the results from static structures (data not shown). However, this technique might improve predictions for mutant complexes, whose geometry differs strongly from the known wild-type structure. Nevertheless, the quality of MD simulations is determined by the quality of the force field. The complex carbohydrate SLe^X and sulfated tyrosines were not parameterized in the CHARMM force field [5, 6] when starting this work. Charges were taken from own QM calculation, which is a very accurate method, and all bonding parameters were adapted from similar compounds. For electrostatic calculations from static structures, this procedure has no significant limitation; however for long MD simulations the bonding force field parameters might need improvement.

In C*c*O, proton gating elements have been identified and characterized for the three proton input channels of A- and B-type C*c*O. For Glu286 inside the D-channel of A-type C*c*O, this analysis was preceded by several contradictory works indicating that Glu286 acts as gating element or not. The present work could explain the cause for some of the contradictions of previous studies. However, I could show that Glu286 is indeed part of the gating mechanism ensuring unidirectional proton transport, but also the action of both hemes and the water chain inside the hydrophobic cavity is needed. To explore this interplay between Glu286 and the hemes further, the use of electrostatic calculations is limited and QM calculations would be needed to represent all features of the hemes properly.

For both the K-channel in A-type C*c*O and the K-channel analogue in B-type C*c*O, the components ensuring proton gating were not known before, but are discovered in this work. The K-channel of A-type C*c*O does only transport two chemical protons and no pumped protons, thus a gate for unidirectional proton transport is of less importance. However, the challenge for the K-channel is to regulate access to the BNC. In principle, all chemical protons could probably be transported via the K-channel – as seen for its analogue in B-type C*c*O – but only two chemical protons are actually conducted via the K-channel in A-type C*c*O. In this work, the gating element of the K-channel was identified as Lys362, which was found to be the key residue for opening and closing of the K-channel. This action is mediated by a direct electrostatic effect of the electron charge distribution in the BNC for the different states of the reaction cycle. Our newly proposed mechanism can be a valuable starting point for further exploration and experimental validation.

B-type C*c*O is of great scientific interest, because (i) it pumps protons although having only a K-channel analogue and no D-channel, which transports pumped protons in A-type C*c*O, and (ii) pumping stoichiometry is likely only 0.5 H^+/e^- at physiological conditions [21]. Nevertheless, the chemical reaction inside the BNC and the identity of the PLS is expected to be the same in A- and B-type C*c*O as their three-dimensional structures of the BNC are virtually identical. To understand the route that protons take inside B-type C*c*O, it is a reasonable starting point to investigate the first segment of their pathway that is the proton input channel. In this work, several functions could be assigned to the residues in the pathway being the proton reservoir, the gating element and the branching point for chemical and pumped protons. These results are in line with previous experimental results on C*c*O mutants [30, 40].

All gating mechanisms may also be investigated using kinetic master equations. In general, a kinetic gate ensuring unidirectional movement must enable the forward movement to be significantly faster than the backward movement. This behavior could be explored by kinetic equations for the proton input channels of C*c*O types in the future.

A general conclusion and outlook of this work is that electrostatic calculations rely on correct atomic coordinates for all employed reaction steps. Therefore, improvements for all the structure-determining techniques as crystallography, molecular modelling and modern variants of MD simulations will help to improve the accuracy of electrostatic calculations. In addition, experimentally achieved knowledge about the protein of interest greatly helps to focus calculations on the relevant settings.

4. Abstract

Electrostatic interactions play a major role in determining protein conformations. They can drive large conformational changes or protein complex formation as well as side chain rotations of amino acids and also the protonation or deprotonation of titratable groups, characterized as pK_A values. In this work, electrostatic and pK_A calculations are used to address biologically relevant questions in two different fields of research namely immunology and energy metabolism.

The immunological study concerns the molecular basis for complex formation between selectins and their counter receptors (C-Rs). The first step when leukocytes are leaving the blood stream – a process called extravasation – is mediated via selectins. There exist three types of selectins, P-, L-, and E-selectin, which all have different binding affinities to their C-R PSGL-1. PSGL-1 is composed of a peptidic part and a carbohydrate part, which includes the four-subunit motif Sialyl-Lewis-X (SLe^X). In this study, the contribution of electrostatic interaction to binding affinity of the molecular components of the selectin-PSGL-1 complex was analysed systematically by calculating the electrostatic binding energy. The most important residues for binding were identified. The available crystal structures of P- and E-selectin soaked with SLe^X were shown to exhibit a geometry not appropriate for binding, in contrast to the co-crystallized structure of P-selectin with PSGL-1. E-selectin was found not to bind to the peptidic part of PSGL-1 in agreement with experimental findings. The most important calculations were verified in collaboration with experimentalists using the surface plasmon resonance technique.

The study on energy metabolism addresses the proton gating mechanism of cytochrome *c* oxidase (C*c*O). C*c*O is the terminal enzyme of the respiratory chain catalysing the reaction of molecular oxygen to water and pumping protons across the membrane. A-type C*c*O has two different input channels for protons, the K-channel conducting two chemical protons being consumed in the reaction and the D-channel conducting the remaining two chemical protons and all four pumped protons. In contrast, B-type C*c*O has only one proton input channel, the K-channel analogue, for all chemical and pumped protons. In this study, the molecular components mediating the proton gating mechanism were identified and characterized for all three types of channels.

In the D-channel of A-type C*c*O, the gating function of glutamate 286 (Glu286) was analysed. Glu286 supports unidirectional proton transport with its high pK_A value, rapid down-flipping of the deprotonated Glu286 and stabilizing an important water chain above Glu286 only in the protonated case. However, Glu286 does not exclusively circumvent back-transport of protons as it may adapt conformations, where it bridges water molecules upstream and downstream of Glu286.

In the K-channel of A-type C*c*O, its central residue lysine (Lys362) was identified as a gate for opening the K-channel. In one half of the C*c*O reaction cycle, Lys362 remains deprotonated pointing away from the reaction center, while in the other half Lys362 may protonate and conduct protons. From these results, a model was proposed explaining how C*c*O switches between channels to transport the chemical protons.

The mechanism of proton-hole transport inside the K-channel analogue of B-type C*c*O was also analysed. Specific residues were found to fulfil a distinct function; glutamate 15B at the channel entrance serves as a proton reservoir, tyrosine 244 in the middle of the channel has a gating function supporting unidirectional proton transport and tyrosine 237 at the end of the channel is the branching point, where the routes of chemical and pumped protons diverge. Electrostatic energy calculations could demonstrate that the K-channel analogue exhibits only small energy barriers for proton-hole transfer supporting rapid transport.

5. Zusammenfassung

Elektrostatische Interaktionen bestimmen zu einem Großteil die Geometrie von Protein Konformationen. Sie können sowohl große Protein-Konformationsänderungen und Komplexbildung hervorrufen als auch das Drehen von Aminosäureseitenketten initiieren oder eine Protonierung/Deprotonierung verursachen, ausgedrückt als pK_A Wert. In dieser Arbeit werden elektrostatische und pK_A Rechnungen verwendet um zwei wissenschaftliche Fragestellungen zu bearbeiten aus den Bereichen Immunologie und Energiestoffwechsel.

Die immunologische Fragestellung betrifft die Binding von Selektinen und ihren Gegenrezeptoren. Der erste Schritt in der Extravasations-Kaskade, bei der Leukozyten die Blutbahn verlassen und ins Gewebe wandern, wird durch Selektine bewerkstelligt. Es gibt drei Klassen von Selektinen, die P-, L- und E-Selektine, die alle mit unterschiedlicher Affinität an ihren Gegenrezeptor PSGL-1 binden. PSGL-1 besteht aus einem peptidischen Anteil mit sulfatierten Tyrosinen und aus einem Kohlenhydrat-Anteil, der das Sialyl-Lewis-X (SLe^X) Motif enthält. In der vorliegenden Arbeit wurden die einzelnen molekularen Komponenten der Selektin-PSGL-1-Binding systematisch analysiert mithilfe elektrostatischer Berechnung der Bindungsenergie. Die wichtigsten molekularen Gruppen für die Bindung konnten identifiziert werden. Es zeigte sich, dass die Kristallstrukturen von E-Selektin und P-Selektin, die nachträglich mit SLe^X getränkt wurden, eine für die SLe^X-Bindung ungeeignete Konformation aufweisen, im Gegensatz zur kokristallisierten Struktur von P-Selektin mit PSGL-1. Außerdem konnte gezeigt werden, dass E-Selektin nicht an den peptidischen Teil von PSGL-1 bindet. Die wichtigsten Ergebnisse der Elektrostatik-Rechnungen wurden in Kollaboration mit Experimentatoren mithilfe der Surface Plasmon Resonance Technik überprüft.

Die wissenschaftliche Fragestellung zum Energiestoffwechsel befasst sich mit dem Schleusenmechanismus der Protonenleitung in der Cytochrom *c* Oxidase (C*c*O). Die C*c*O ist das terminale Enzym der Atmungskette, welches molekularen Sauerstoff zu Wasser umwandelt und dabei Protonen über die Membran pumpt. Die C*c*O des Typs A hat zwei Eingangskanäle für Protonen: den K-Kanal, durch den zwei der chemischen, in der Reaktion verbrauchten Protonen geleitet werden, und den D-Kanal, der die verbleibenden zwei chemischen Protonen und alle vier gepumpten Protonen transportiert. Die C*c*O des Typs B hat

im Gegensatz dazu nur einen Protonen-Eingangskanal, den K-Analog-Kanal, der alle chemischen und gepumpten Protonen leitet.

Im D-Kanal der A-Typ C*c*O wurde die Schleusenfunktion des Glutamats 286 (Glu286) analysiert. Glu286 trägt dazu bei, dass der Protonentransport unidirektional erfolgt, mit seinem hohen pK_A Wert, einem schnellen Umflippen des deprotonierten Glu286 und der Stabilisierung einer wichtigen Wasserkette durch das protonierte Glu286. Jedoch kann Glu286 allein keine vollständige Schleusenfunktion übernehmen, da es sich auch teilweise in Konformationen befinden kann, in denen es die Wassermoleküle davor und dahinter miteinander verbindet und einen Protonen-Rücktransport nicht verhindern kann.

Für den K-Kanal der A-Typ C*c*O wurde das zentrale Lysin 362 (Lys362) als Schaltelement identifiziert, welches das Öffnen und Schließen des K-Kanals reguliert. In der einen Hälfte des Reaktionszyklus von C*c*O bleibt Lys362 durchgängig deprotoniert und zeigt in Richtung des Kanal-Eingangs. In der anderen Hälfte des Reaktionszyklus dagegen kann Lys362 protonieren und damit Protonen leiten. Aus diesen Ergebnissen wurde ein Model entwickelt, das erklären kann, wie C*c*O im Laufe des Reaktionszyklus zwischen den beiden Protonen-Kanälen wechselt.

Für den K-Analog-Kanal der C*c*O Typ B wurde der Protonen-Loch-Transport analysiert. Einige Aminosäuren des Kanals haben eine spezialisierte Funktion: Glutamat 15B am Kanaleingang wirkt als Protonen-Reservoir, Tyrosin 244 in der Kanalmitte sorgt als Schleuse für einen unidirektionalen Protonentransport und Tyrosin 237 am Kanalende ist die Verzweigung, an der sich die Wege der chemischen und der gepumpten Protonen trennen. Mit elektrostatischen Rechnungen konnte zudem gezeigt werden, dass der K-Analog-Kanal keine hohen Energiebarrieren aufweist und somit eine schnelle Protonenlochleitung fördert.

6. References

1. Baker, N.A., et al., *Electrostatics of Nanosystems: Application to Microtubules and the Ribosome.* Proc. Natl. Acad. Sci. U S A, 2001. **98**(18): p. 10037-41.
2. Baker, N.A., et al., *Electrostatics of Nanosystems: Application to Microtubules and the Ribosome.* Proc. Natl. Acad. Sci. USA, 2001. **98**(18): p. 10037-41.
3. Kieseritzky, G. and E.W. Knapp, *Optimizing pKa Computation in Proteins with pH Adapted Conformations.* Proteins, 2008. **71**(3): p. 1335-48.
4. Rabenstein, B. and E.W. Knapp, *Calculated pH-dependent population and protonation of carbon-monoxy-myoglobin conformers.* Biophys. J., 2001. **80**(3): p. 1141-50.
5. Brooks, B.R., et al., *CHARMM: the Biomolecular Simulation program.* J. Comput. Chem., 2009. **30**(10): p. 1545-614.
6. Klauda, J.B., et al., *Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types.* J. Phys. Chem. B, 2010. **114**(23): p. 7830-43.
7. Guvench, O., et al., *Additive Empirical Force Field for Hexopyranose Monosaccharides.* J. Comput. Chem., 2008. **29**(15): p. 2543-64.
8. Guvench, O., et al., *CHARMM Additive All-Atom Force Field for Glycosidic Linkages between Hexopyranoses.* J. Chem. Theory. Comput., 2009. **5**(9): p. 2353-2370.
9. Ley, K., *The Role of Selectins in Inflammation and Disease.* Trends Mol. Med., 2003. **9**(6): p. 263-8.
10. Kneuer, C., et al., *Selectins--Potential Pharmacological Targets?* Drug. Discov. Today, 2006. **11**(21-22): p. 1034-40.
11. McEver, R.P. and C. Zhu, *Rolling Cell Adhesion.* Annu. Rev. Cell Dev. Biol., 2010. **26**: p. 363-96.
12. Croce, K., et al., *Interaction between Soluble P-Selectin and Soluble P-Selectin Glycoprotein Ligand 1: Equilibrium Binding Analysis.* Biochemistry, 1998. **37**(47): p. 16472-80.
13. Klopocki, A.G., et al., *Replacing a Lectin Domain Residue in L-Selectin Enhances Binding to P-Selectin Glycoprotein Ligand-1 But Not to 6-Sulfo-Sialyl Lewis x.* J. Biol. Chem., 2008. **283**(17): p. 11493-500.
14. Mehta, P., R.D. Cummings, and R.P. McEver, *Affinity and Kinetic Analysis of P-Selectin Binding to P-Selectin Glycoprotein Ligand-1.* J. Biol. Chem., 1998. **273**(49): p. 32506-13.
15. Poppe, L., et al., *Conformation of sLex Tetrasaccharide, Free in Solution and Bound to E-, P-, and L-Selectin.* J. A. C. S., 1997. **119**(7): p. 1727-1736.
16. Somers, W.S., et al., *Insights into the Molecular Basis of Leukocyte Tethering and Rolling Revealed by Structures of P- and E-Selectin Bound to SLe(X) and PSGL-1.* Cell, 2000. **103**(3): p. 467-79.
17. Ley, K. and G.S. Kansas, *Selectins in T-Cell Recruitment to Non-Lymphoid Tissues and Sites of Inflammation.* Nat. Rev. Immunol., 2004. **4**(5): p. 325-35.
18. Sperandio, M., C.A. Gleissner, and K. Ley, *Glycosylation in Immune Cell Trafficking.* Immunol. Rev., 2009. **230**(1): p. 97-113.
19. Brzezinski, P. and R.B. Gennis, *Cytochrome c Oxidase: Exciting Progress and Remaining Mysteries.* J. Bioenerg. Biomembr., 2008. **40**(5): p. 521-31.
20. Kaila, V.R., M.I. Verkhovsky, and M. Wikstrom, *Proton-Coupled Electron Transfer in Cytochrome Oxidase.* Chem. Rev., 2010. **110**(12): p. 7062-81.

21. von Ballmoos, C., et al., *Proton Transfer in ba(3) Cytochrome c Oxidase from Thermus thermophilus.* Biochim. Biophys. Acta, 2012. **1817**(4): p. 650-7.
22. Gilderson, G., et al., *Subunit III of Cytochrome c Oxidase of Rhodobacter sphaeroides is Required to Maintain Rapid Proton Uptake Through the D Pathway at Physiologic pH.* Biochemistry, 2003. **42**(24): p. 7400-9.
23. Rauhamaki, V. and M. Wikstrom, *The Causes of Reduced Proton-Pumping Efficiency in Type B and C Respiratory Heme-Copper Oxidases, and in some Mutated Variants of Type A.* Biochim. Biophys. Acta, 2014. **1837**(7): p. 999-1003.
24. Svensson-Ek, M., et al., *The X-ray Crystal Structures of Wild-Type and EQ(I-286) Mutant Cytochrome c Oxidases from Rhodobacter sphaeroides.* J. Mol. Biol., 2002. **321**(2): p. 329-39.
25. Tiefenbrunn, T., et al., *High Resolution Structure of the ba3 Cytochrome c Oxidase from Thermus thermophilus in a Lipidic Environment.* PLoS One, 2011. **6**(7): p. e22348.
26. Qin, L., et al., *Identification of Conserved Lipid/Detergent-Binding Sites in a High-Resolution Structure of the Membrane Protein Cytochrome c Oxidase.* Proc. Natl. Acad. Sci. USA, 2006. **103**(44): p. 16117-22.
27. Ostermeier, C., et al., *Structure at 2.7 A Resolution of the Paracoccus denitrificans Two-Subunit Cytochrome c Oxidase Complexed with an Antibody FV Fragment.* Proc. Natl. Acad. Sci. USA, 1997. **94**(20): p. 10547-53.
28. Yoshikawa, S., et al., *Redox-Coupled Crystal Structural Changes in Bovine Heart Cytochrome c Oxidase.* Science, 1998. **280**(5370): p. 1723-9.
29. Branden, M., et al., *The Entry Point of the K-Proton-Transfer Pathway in Cytochrome c Oxidase.* Biochemistry, 2002. **41**(35): p. 10794-8.
30. Chang, H.Y., et al., *The Cytochrome ba3 Oxygen Reductase from Thermus thermophilus Uses a Single Input Channel for Proton Delivery to the Active Site and for Proton Pumping.* Proc. Natl. Acad. Sci. USA, 2009. **106**(38): p. 16169-73.
31. Kaila, V.R., et al., *Glutamic Acid 242 is a Valve in the Proton Pump of Cytochrome c Oxidase.* Proc. Natl. Acad. Sci. USA, 2008. **105**(17): p. 6255-9.
32. Yamashita, T. and G.A. Voth, *Insights into the Mechanism of Proton Transport in Cytochrome c Oxidase.* J. Am. Chem. Soc., 2012. **134**(2): p. 1147-52.
33. Yang, S. and Q. Cui, *Glu-286 Rotation and Water Wire Reorientation are Unlikely the Gating Elements for Proton Pumping in Cytochrome C Oxidase.* Biophys. J., 2011. **101**(1): p. 61-9.
34. Mehta, P. and R.P. McEver, *Structure of Lectin and EGF Domains of L-Selectin.* To be published.
35. Namslauer, A., et al., *Intramolecular Proton-Transfer Reactions in a Membrane-Bound Proton Pump: the Effect of pH on the Peroxy to Ferryl Transition in Cytochrome c Oxidase.* Biochemistry, 2003. **42**(6): p. 1488-98.
36. Adelroth, P., R.B. Gennis, and P. Brzezinski, *Role of the Pathway Through K(I-362) in Proton Transfer in Cytochrome c Oxidase from R. sphaeroides.* Biochemistry, 1998. **37**(8): p. 2470-6.
37. Konstantinov, A.A., et al., *The Roles of the Two Proton Input Channels in Cytochrome c Oxidase from Rhodobacter sphaeroides Probed by the Effects of Site-Directed Mutations on Time-Resolved Electrogenic Intraprotein Proton Transfer.* Proc. Natl. Acad. Sci. USA, 1997. **94**(17): p. 9085-90.
38. Branden, M., et al., *On the Role of the K-proton Transfer Pathway in Cytochrome c Oxidase.* Proc. Natl. Acad. Sci. USA, 2001. **98**(9): p. 5013-8.

39. Lepp, H., et al., *Charge Transfer in the K Proton Pathway Linked to Electron Transfer to the Catalytic Site in Cytochrome c Oxidase.* Biochemistry, 2008. **47**(17): p. 4929-35.
40. Smirnova, I., et al., *Functional Role of Thr-312 and Thr-315 in the Proton-Transfer Pathway in ba3 Cytochrome c Oxidase from Thermus thermophilus.* Biochemistry, 2010. **49**(33): p. 7033-9.