

# **Regulation of Permeation in Aquaporins**

Dissertation

for the award of the degree

*Doctor rerum naturalium*

of the Georg-August-Universität Göttingen

within the doctoral program Physics of Biological and Complex Systems (PBCS)  
of the Georg-August-University School of Sciences

submitted by

Shreyas Sanjay Kaptan

Göttingen, 2015

## DECLARATION

I hereby declare that this thesis is composed independently by me at Max Planck Institute for Biophysical Chemistry, Goettingen, Germany, under the supervision of Prof. Dr. Bert de Groot. All notions, taken directly or indirectly from elsewhere have been identified as such. The subject matter presented in this thesis has not previously formed the basis of the award of any degree, diploma, associateship, fellowship or any other similar title in any other University.

Goettingen, 31/01/2015

Shreyas Sanjay Kaptan

*To my family*



## ACKNOWLEDGEMENTS

I would like to thank my thesis supervisor Prof. Bert de Groot for guiding me through the thesis work. He has been a source of constant encouragement to me, acting not only as my guide, but also as a mentor, a teacher and a parental figure. He always allowed me to be independent and supported my endeavors with genuine concern and affection, for which I am truly grateful. Similarly, I would thank Prof. Helmut Grubmueller, who also shared in the role of a mentor and inculcated within me the need for good scientific practices. I would like to extend my gratitude towards Jan Peters and David Koepfer who were there to cheer me up in gloom and share in my times of happiness. They have always been patient with me throughout my time here and for that I am grateful. I sincerely thank my thesis advisory committee members, Prof. Marina Benatti and Prof. Marcus Mueller, for participating in this journey and seeing to it that it was successful. Their guidance and advice was crucial for the thesis. I would like to express gratitude towards Dr. Petra Kellers, Dr. Hadas Leonov and Dr. Rudolfo Briones for reading my thesis work and making excellent suggestions. I would like to thank Martin Fechner and Ansgar Ezsterman for looking after the technical side of the projects and helping me with all the troubles concerning the hardware. Also, I cannot thank Antje Erdmann, Frauke Bergmann and Tina Trost for being patient with me concerning the exchanges with the administration while being a part of the GGNB program. I would like to thank all the group members for providing a wonderful environment where doing science was fun. I would like to thank my collaborators Prof. Nanna Macaulay and Dr. Mette Assentoft who carried out the experimental work concerning the AQP4 project. I would like to also thank Prof. Urban Johansson and Andreas Kirscht who collaborated with us in the work related to ammonia permeation aquaporins and provided the crystal structure of the TIP2;1. I am thankful to Dr. Manish Kumar and Prof. Tom Walz for the experimental collaboration in the AQP0 project.

Lastly and most importantly I would like to express my gratitude towards my parents and thank my wife, who supported me emotionally through the PhD and always were there for me when I needed them.

There are many others in my life, teachers, friends and family, who directly or indirectly helped make this thesis a success. I would like to apologize that I can not mention them all and thank them from the bottom of heart for their help.

Shreyas Sanjay Kaptan

Goettingen

January, 2015



## SYNOPSIS

Regulation of water permeation is an essential aspect of homeostasis in living organisms. Proteins of the ubiquitous aquaporin family act as trans-membrane tetrameric channels through which water may permeate under osmotic gradients in a bidirectional manner. These channels are highly specific in the choice of the conducting molecules, but at the same time several members of this family allow permeation of alternative solutes. In the last two decades, the structures of several aquaporin proteins have been unveiled through x-ray and electron crystallography and have helped us understand the features of this protein family that make solute permeation across bio-membranes possible. Molecular Dynamics (MD) simulations have contributed substantially to the understanding of the mechanisms that govern the efficiency and the specificity of the aquaporin family of protein. These simulations have helped shed light on the finer mechanical details of the process of water permeation at the atomistic level. MD simulations add a new dimension, that of the dynamics, to the ensemble of available knowledge. This added information is crucial to understand the inherently dynamic nature of the permeation process.

In this study we explore the molecular mechanisms that regulate the permeability in three aquaporin proteins, namely, AQP4, AQP0 and TIP2;1. We find that the permeability of AQP4 protein can be modulated with an extrinsic parameter such as pH. We provide an explanation for the native low permeability of the protein AQP0 and suggest means to manipulate this protein through mutations so it can be made more water permeable. Finally, we study the permeability of ammonia through the plant aquaporin TIP2;1 and explore the origin of the modified specificity of this protein. To achieve the results stated above, we use a variety of techniques related to MD simulations and highlight several aspects of regulation of permeation that could be general features of the protein family as a whole.





# Contents

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Introduction</b>   | <b>3</b>  |
| 1.1      | Background . . . . .  | 3         |
| 1.2      | The aquaporin protein family . . . . .  | 6         |
| 1.3      | The structure of the aquaporin channel . . . . .                                | 8         |
| 1.4      | The biophysics of water permeation in aquaporins . . . . .                      | 11        |
| 1.4.1    | Water permeability of aquaporins . . . . .                                      | 11        |
| 1.4.2    | Selectivity of aquaporins . . . . .   | 13        |
| 1.5      | Why molecular dynamics? . . . . .   | 16        |
| 1.6      | Goals of this project . . . . .   | 17        |
| 1.7      | Organization of the thesis . . . . .  | 18        |
| <b>2</b> | <b>Methodology</b>  | <b>23</b> |
| 2.1      | Molecular Dynamics Simulations . . . . .  | 23        |
| 2.1.1    | Physical basis of MD simulations . . . . .                                      | 23        |
| 2.1.2    | Numerical integration of equations of motion . . . . .                          | 26        |
| 2.1.3    | Thermostats and Barostats . . . . .   | 27        |
| 2.1.4    | Water models . . . . .  | 28        |
| 2.1.5    | Handling the electrostatics and the van der Waals interactions                  | 28        |
| 2.1.6    | Periodic boundary conditions . . . . .  | 29        |
| 2.2      | Functional Mode Analysis (FMA) . . . . .  | 29        |
| 2.2.1    | Dimensionality reduction . . . . .  | 29        |
| 2.2.2    | Relating to the function . . . . .  | 30        |
| 2.2.3    | Partial Least Squares based Functional Mode Analysis . . . . .                  | 31        |
| 2.3      | Enhanced sampling . . . . .   | 32        |
| 2.3.1    | Umbrella sampling . . . . .   | 32        |
| <b>3</b> | <b>Regulation of permeability in Aquaporin-4</b>                                | <b>37</b> |
| 3.1      | Overview . . . . .  | 37        |
| 3.1.1    | Biological role of Aquaporin-4 . . . . .  | 37        |
| 3.1.2    | Regulation of AQP4 function . . . . .   | 39        |
| 3.2      | Computational details . . . . .   | 40        |
| 3.3      | Can phosphorylation of S111 gate the AQP4 channel? . . . . .                    | 41        |
| 3.3.1    | Background . . . . .  | 41        |
| 3.3.2    | Comparison of the phosphorylated and non-phosphorylated forms of AQP4 . . . . . | 41        |

---

|          |   |            |
|----------|---|------------|
| 3.4      | Is aquaporin-4 gated by pH? . . . . .                         | 45         |
| 3.4.1    | Exploring the effect of H95 protonation state on AQP4 gating  | 46         |
| 3.4.2    | Experimental verification of pH regulation in AQP4 . . . . .  | 53         |
| 3.5      | Conclusions . . . . .   | 54         |
| <b>4</b> | <b>Modulating the permeability of Aquaporin-0</b>             | <b>61</b>  |
| 4.1      | Introduction . . . . .  | 61         |
| 4.1.1    | Role of aquaporin-0 . . . . .                                 | 61         |
| 4.1.2    | Structure of aquaporin-0 . . . . .                            | 63         |
| 4.2      | Motivation . . . . .  | 65         |
| 4.3      | Computational details . . . . .                               | 66         |
| 4.4      | Can we mutate AQP0 to have high water permeability? . . . . . | 66         |
| 4.5      | What governs the low permeability of AQP0? . . . . .          | 70         |
| 4.6      | Conclusions . . . . .   | 86         |
| <b>5</b> | <b>Ammonia permeation in plant aquaporin TIP2;1</b>           | <b>91</b>  |
| 5.1      | Overview . . . . .  | 91         |
| 5.2      | Motivation . . . . .  | 92         |
| 5.3      | Is the TIP2;1 structure a functional water channel? . . . . . | 95         |
| 5.4      | Novel features of the selection filter of TIP2;1 . . . . .    | 96         |
| 5.5      | Unusual side-pore in the TIP2;1 . . . . .                     | 98         |
| 5.6      | Does TIP2;1 permeate ammonia? . . . . .                       | 103        |
| 5.7      | Conclusions . . . . .   | 106        |
| <b>6</b> | <b>Conclusions and perspectives</b>                           | <b>109</b> |
| 6.1      | Conclusions . . . . .   | 109        |
| 6.2      | Future . . . . .  | 113        |
| 6.2.1    | Substrate specificity of AQP4 . . . . .                       | 113        |
| 6.2.2    | 'Rescuing' the ammonia permeation in Human AQP1 . . . . .     | 113        |
| 6.2.3    | A physiological model for the biological membranes . . . . .  | 114        |
| 6.2.4    | Generality of the 'Arginine mode' . . . . .                   | 114        |









# Chapter 1

## Introduction

### 1.1 Background

One of the characteristic features of living organisms is the isolation of their cellular chemistry through compartmentalization. A living cell requires a continuous function of its metabolic pathways, such as anaerobic glycolysis or the aerobic Tri-Carboxylic Acid (TCA) cycle for efficient function and survival. It is observed that the chemical cycles that drive the cell are typically shielded from the 'bulk' environment by some form of semi-permeable insulation. In higher organisms this shielding extends to the genetic material which is additionally isolated from the cytoplasm into the nucleus. Further segregation can occur in internal organelles, such as lysosomes, mitochondria, chloroplasts where the machinery required to perform specific chemical function is sequestered from the rest of the cellular mass. A constant of the chemical nature of all these physical barriers is that they are composed of lipid membranes in the form of a bilayer. Across this bilayer a permeation of nutrients must occur so that a homeostatic chemical milieu can thrive with minimum interruption inside the cell. This particular function is generally handled by a variety of proteins that are embedded within the membrane. Acting as channels and transporters, they use either passive or active means to shuttle the required substrates in and out of the compartments [1] .

Water is the most abundant component of the chemical mix present on either side of the bilayer where it functions chiefly as a solvent. Therefore, managing the

permeation of water is an important function associated with the proteins embedded in the bilayer. The historical development concerning the understanding of water permeability is covered in detail in the review by Parisi *et al.* [2]. It was initially unknown if permeability of water was attributed particularly to proteins. The bilayer itself was deemed sufficiently permeable to allow for the osmotic movement of water [3]. However, studies conducted from the 1950s to the 1960s indicated that water permeability of the biological cell could not be attributed to diffusive or osmotic entry of water through the cell membrane alone.

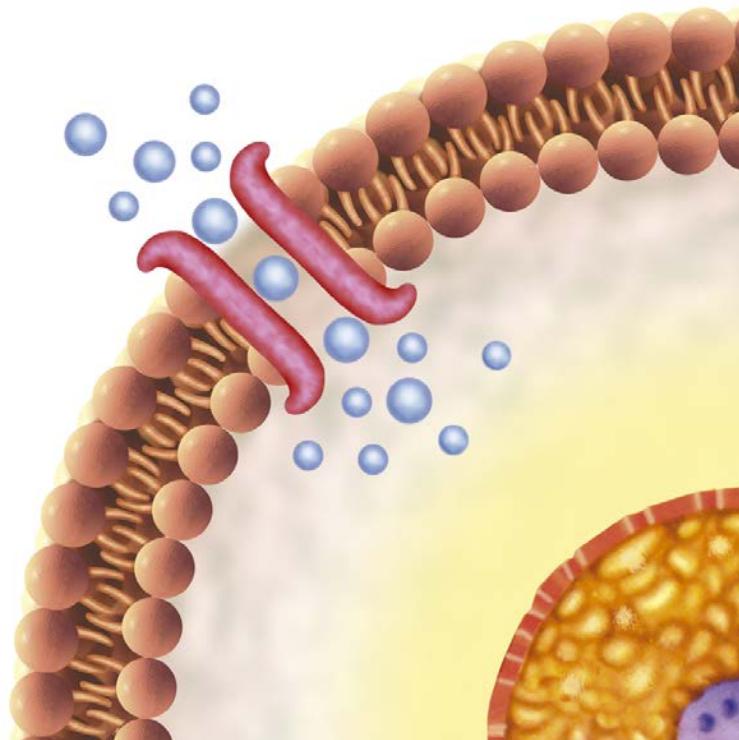


FIGURE 1.1: Separation of the cytoplasm from the cell exterior via the plasma membrane. Molecular channels allow transmission of nutrients across this barrier. (Source: <http://www.medicalscienzenavigator.com>)

This gave rise to the 'membrane pore' hypothesis that postulated the presence of dedicated poriform structures which could facilitate the transfer of water across the cell boundary. Initial evidence arrived in the form of observation of the permeation rate of tritiated water across the Red Blood Cell (RBC) membrane, which exceeded the rate that could be expected from a purely diffusive transport

across the cell membrane [4]. Although this hypothesis was initially challenged by contemporary investigators [5, 6], it regained attention upon the discovery of dedicated water channels which acted as antibiotics, such as the Gramicidin popypeptide [7]. In the 1980’s evidence of a specific water channels began to accumulate. The observation that the anti-diuretic hormone (ADH) could stimulate the formation of protein aggregates which increased water permeability [8] lead to a clear evidence that biological membranes had proteinaceous members capable of acting as water channels. Furthermore, the water permeation affected by ADH was found to be modulated via pH [9] in a manner that pH could not lead to either formation or disassembly of the aggregate itself. This indicated a regulatory mechanism placed on the aggregate itself thus pinning the function on the protein/s. It was also shown that the permeability in RBCs could be inhibited by application of mercury [10] in a reversible manner. This implied that the protein potentially responsible had free sulfhydryl groups, that could be chemically bound with mercury. This protein was finally identified by Preston *et al.* in 1992, [11, 12]. Initially called CHIP28, due to its mass of 28kDa and the assay used to detect this protein, it was later named to aquaporin-1 (AQP1).

Now, after more than two decades of the discovery of AQP1, the world of aquaporins has expanded from putative models and innovative function to a broad understanding of their structure and well established empirical methods to assess their functionality. Computational models have become available to explain their behavior at an atomistic level, shedding light on the relation between their structure and their function. Despite this familiarity, a multitude of questions regarding the regulation of their behavior remain unanswered. In this thesis we will try to address some of these questions and hopefully advance our understanding of the working of aquaporins protein channels. In this chapter, a brief overview of important historical and scientific background is provided to bring the reader in tune with the current state of understanding regarding the topic. Several important ideas and facts related to aquaporin function and structure are also introduced so as to facilitate familiarity with details later discussed in the thesis.

## 1.2 The aquaporin protein family

Aquaporins are tetrameric transmembrane channel proteins, in which each of the monomers is an independently functional channel pore. After the discovery of AQP1, a large number of proteins were identified across the entire biome, sharing with AQP1 sequence, structural and functional similarity [13]. These proteins constitute what is now called the aquaporin family. There have been more than 450 proteins identified to date. Most of these are isoforms present in multicellular organisms, involved in varied physiological roles associated with the function of the tissue where they are expressed. Plants typically possess a larger number of isoforms per species compared to animals, with most angiosperms having more than 30 aquaporin proteins. *Arabidopsis thaliana* has 38 isoforms of the protein compared to 13 for humans. The human isoforms are comparatively well studied due to the clinical significance associated with their function. They are designated with the names AQP attached to a number as AQP0 to AQP12.

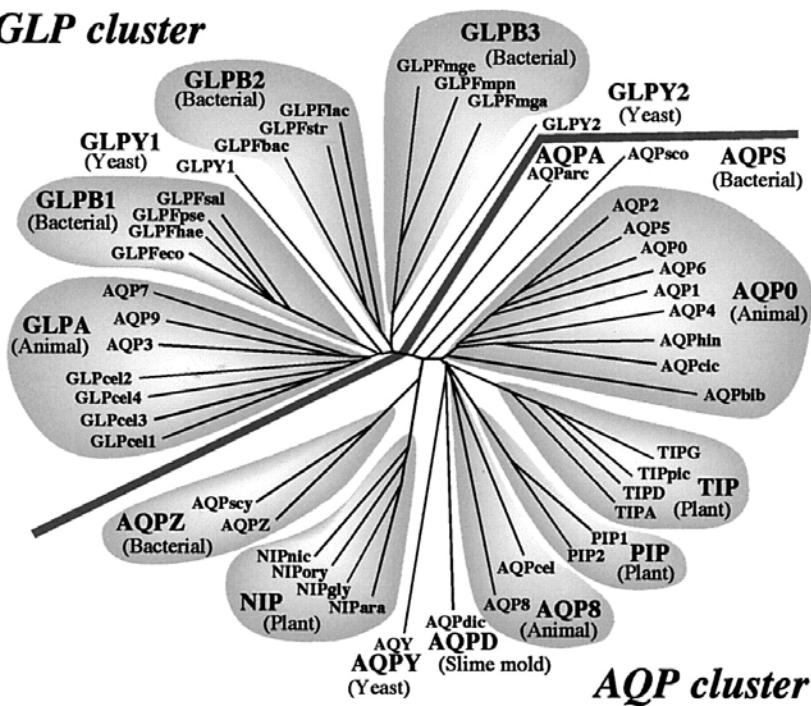


FIGURE 1.2: Division of the aquaporin protein family into clusters. The GLP cluster represents the glycerol permeable aquaporins. The AQP cluster is the water specific set of the family.

Figure taken from [14]

Aquaporins, although predominantly water channels, are not restricted to only one permeating solute. There are two major aquaporin sub-families (fig.1.2). The first is the Glycerol Facilitator Proteins (GLPs) family which allows permeation of GLycerol and water through the pore. The second subfamily is the water specific Aquaporins (AQPs). In humans, AQP0,1,2,4,5,6 and 8 act as water channels and AQP3,7,9 and 10 are permeable to glycerol and water. In plants, the AQPs are further sub-divided into three categories. The aquaporins present on the plasma membrane are termed Plama-membrane Integral Proteins (PIPs) and are sub-divided into two types PIP1 and PIP2. PIP1 proteins exhibit significantly smaller water conductance compared to PIP2 proteins. The second category is present intra-cellularly on the Tonoplast membrane and its members are called Tonoplast Integral Proteins (TIPs). The last identified plant aquaporin subgroup is called Nodule Integral Protien (NIPs) as they were originally found on the Nodules of leguminous plants, which are involved in fixing nitrogen from the soil. The aquaporin family phylogeny is reviewed in detail by Heymann and Engel *et al.* [14].

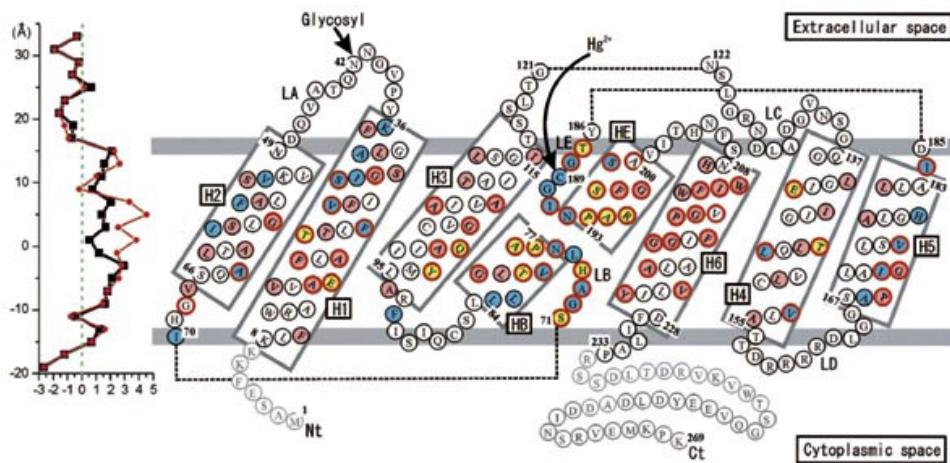


FIGURE 1.3: Aquaporin family members are integral membrane proteins with 6 transmembrane helices and two semi-transmemebrane helices. The latter end into conserved NPA motifs, which form part of the so called 'filter' at the center of the monomer channel. Picture adapted from [15]

The aquaporin gene and consequently the protein shows two-fold pseudo-symmetry. Similarity between these two halves of the aquaporin gene sequence suggests a duplication event in the early genetic history of the family. The two halves constitute hemi-pores that associate to form the water channel. Conserved sites are not uncommon among closely related members. Globally, the sequence of aquaporin proteins has two strongly conserved motifs present in the channel lumen; the NPA-duplex motif and an arginine residue (R195 for AQP1) placed in a highly hydrophobic region of the channel interior. The latter forms a narrow funnel in the pore which gives rise to the hourglass shaped structure of aquaporin channels. This funnel portion of the protein is termed aromatic/arginine (ar/R) region. The combination of the NPA motifs and the ar/R region give rise to the 'filter' in the protein. F56, H180 and C189 form the aromatic environment in AQP1. The arginine residue if mutated from the protein results in a highly permeable but otherwise less selective channel. The sulphydryl group of the cystine residue inserts into the pore and forms a binding site for mercury. This binding results in the channel being blocked from permeation and is now known to be the origin of the mercury based inhibition of a large subset of aquaporins [15].

### 1.3 The structure of the aquaporin channel

Several members of the aquaporin family proteins have been studied with X-ray and electron crystallographic methods, which have been used to elucidate their structural and topological features (AQP1 [15, 16]); AQPZ ([17]); AQP0 ([18]) ; AQYM ([19]); SOPIP2;1 ([20]); PfAQP ([21]); AQP5 ([22]); AQP4 ([23]),GlpF ([24]) ). These structures have provided us with insights into the inner mechanical working of the channel.

The aquaporin protein is a homo-tetramer (fig.1.4), where all four monomers are fully functional water channels. Each of the AQP monomers is made up of 6 transmembrane helices arranged as a helical bundle. The helices are labeled with an index H1 to H6 (fig.1.3). Alongside these transmembrane helices, there are two semi-trans-membrane helices that insert themselves halfway into the channel

pore. These are labeled HB and HE. The two asparagine residues of the NPA-duplex present at the termini of these helices and act as hydrogen donors to the permeating substrates. The channel pore of the monomer is 25 Å in length. It has a sharp constriction near the conserved arginine site, surrounded by hydrophobic residues. The channel can be as narrow as 1.4 Å (approximate radius of a water molecule) or even lower in this region for water specific aquaporin proteins. In some aquaporins that also conduct other substrates such as Glycerol, this constriction can be wider, and up to 3.4 Å. Another slightly larger constriction is formed by the NPA-duplex placed on the cytosolic side of the ar/R region. Near the extracellular and the cytoplasmic regions the pore is relatively wide, reaching a diameter of 6 to 7 Å. The junction of the four monomers gives rise to a single central pore. This pore does not allow permeation of water. However it has been shown that it can potentially participate in permeation of smaller hydrophobic residues [25] and possibly ions [26].

The monomeric pores, called the 'water pores' are usually constitutively in an open conformation and are typically co-crystallized with water molecules trapped in the channel. The lumen of the water pore is lined mostly with hydrophobic residues that present the carbonyl groups of the mainchain as hydrogen bonding partners to permeating water molecules. Near the ar/R region and the NPA motifs, the side-chains also contribute as hydrogen bonding partners to permeating water molecules. These interactions have been suggested to be important for compensating the desolvation of water as it enters the aquaporin channel [27]. The actual permeation through the channel was thought to occur via single file motion of water molecules across the channel [9]. This proved to be partially true. For aquaporins with narrow pores, such as AQP-Z, the number of binding sites calculated experimentally matched the actual number of water molecules observed in the crystal structures. On the other hand, aquaporins with wider pores, such as the Glycerol permeable GlpF, the number of binding sites was found to be much smaller in permeation experiments as compared to the X-ray structure, thus demonstrating that these pores showed a 'multi-lane' permeation. In general, it appears that a 'single file' motion of water across the channel lumen is a useful approximation, that can yield qualitative results in agreement with experiments.

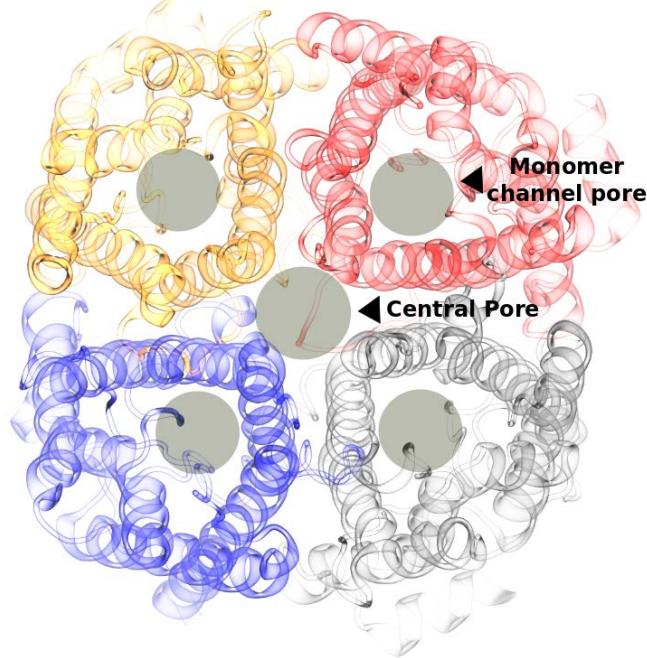


FIGURE 1.4: A typical aquaporin protein (in this case AQP1). Viewed from the top, the tetramer has four independent fully functional water pores. The pore formed at the center does not allow water permeation.

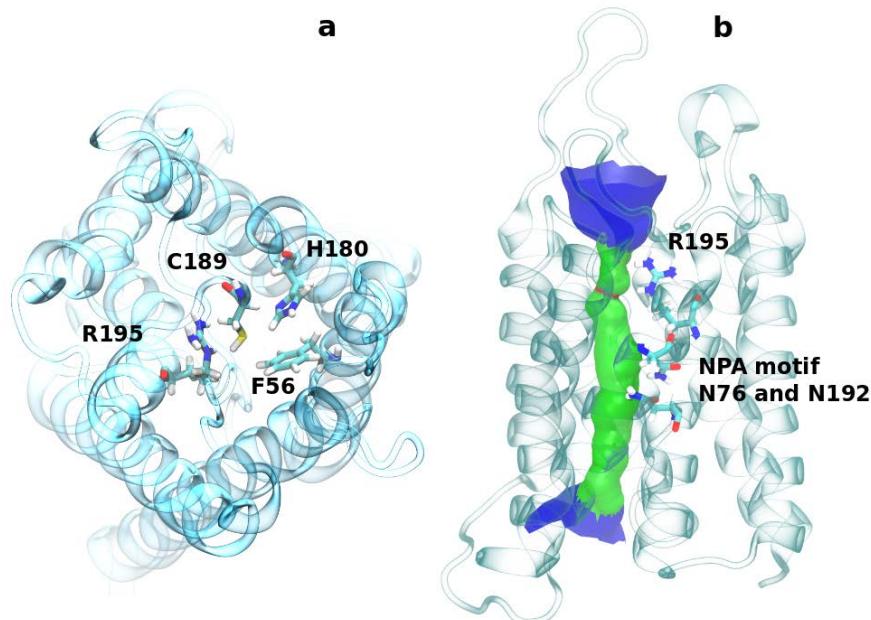


FIGURE 1.5: Structure of the water pore. a. The top view of a monomer is shown in cartoon representation. The four residues that make up the ar/R part of the filter for the AQP1 protein, are shown in licorice representation. b. The tunnel across the protein monomer channel is shown. The arginine of the ar/R region and the asparagines of the NPA motifs that form the selectivity filter are highlighted. The tunnel itself is shown as a space-filling model, with the narrowest constriction shown in red.

## 1.4 The biophysics of water permeation in aquaporins

### 1.4.1 Water permeability of aquaporins

Water permeates through the channel lumen in an ordered and choreographed fashion. This picture is substantiated by the crystal structures, which depict conserved locations in the pore where water molecules are co-crystallized with the protein. Physiologically, the aquaporins are bidirectional channels. The permeation process across the channel is diffusive or osmotically driven [28]. Experimentally, the measurement of aquaporin permeability is done via the CHIP assay. This assay was first used by Agre *et al.* to identify the CHIP28 protein as a water channel [11]. It involves measurement of the swelling in *Xenopus laevis* oocytes. These oocytes are transfected with aquaporin cDNA. Upon expression the protein is transported to the cell surface. The relative swelling rate is then contrasted with a control oocyte that does not express the protein, thus allowing measurement of the permeability. The inability to accurately determine the copy number of the proteins functionally present on the membrane is chiefly responsible for the lack of accuracy in the measurement. The typical aquaporin permeability is on the order of  $10^{-14}$  cm<sup>3</sup>/second. It measures the rate of volume transfer of water across the pore. This rate results in approximately a billion permeation events per channel per second on an average. It can fluctuate across the protein family ranging from a minuscule  $0.2 \times 10^{-14}$  cm<sup>3</sup>/second in AQP0 to  $20 \times 10^{-14}$  cm<sup>3</sup>/second in TIP2;1. Measurements carried in computational studies qualitatively agree with empirical measurements.

It appears that the control of permeability of the aquaporin channel is strongly related to the constriction region, as it posits the largest Arrhenius barrier to the permeation of solutes through the pore (fig.1.6). This consideration rationalizes the impermeability of the channel to substrates of a size larger than water molecules. The constriction region is formed from a two stage filter as mentioned earlier, the first of which is the ar/R region. The second half is formed out of the NPA-duplex motif.

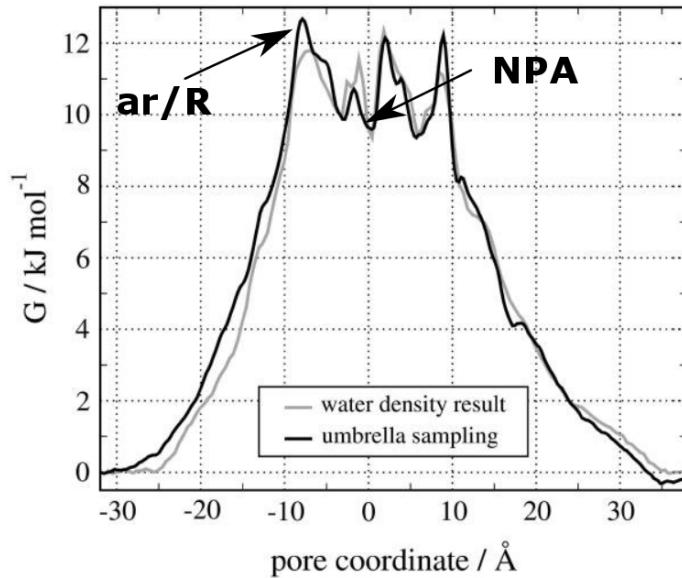


FIGURE 1.6: Potential of mean force for water across the monomer channel. The ar/R region is close to the  $-10 \text{ \AA}$  value on the pore coordinate and poses the largest free energy barrier.

[29]

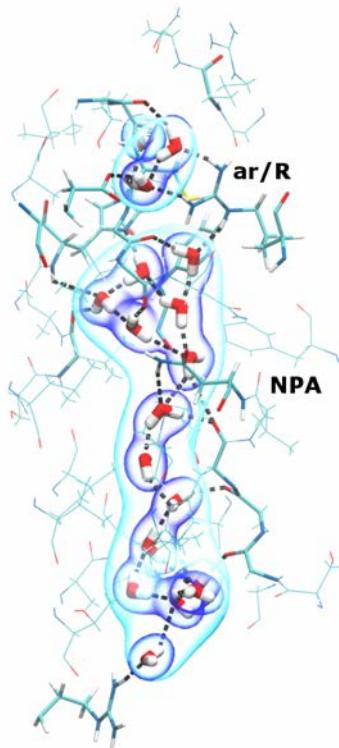


FIGURE 1.7: A snapshot of water permeating through the channel pore. The water molecules at all points during the permeation are coordinated by other water molecules or though the polar atoms of the sidechain or the mainchain of the protein. The permeation is only approximately single file. The hydrogen bonds are shown with dashed black lines. The possible hydrogen bond partners and water molecules are shown with the licorice models. The continuity of the inter-water hydrogen bonds is broken near the ar/R region.

Why is the permeability of the protein to water larger than the direct osmotic transfer rates across the membrane? This question can be answered by studying the channel pore interior. In the bulk, water has an intricate hydrogen bond network with neighboring water molecules. The entry of water into the channel lumen is entropically disfavored. This is because the water molecules entering the narrow pore lose both translational and rotational degrees of freedom that can be accessed in the bulk. In order to 'accommodate' these water molecules into the channel, the lumen offers complementary hydrogen bonding network formed by polar sidechain atoms and mainchain carboxyl groups. This complementarity explains the reduced activation barrier compared to the path through the hydrophobic core of the membrane for the entry of water molecules in the constricted channel lumen. However, the presence of a large number of strong hydrogen bonding residues in the pore is also detrimental to the permeation. Moreover, excessive hydrogen bonding in the pore can kinetically trap water molecules in the lumen by creating deep free energy minima along the permeation pathway. It appears that the protein has achieved a trade-off in these effects through evolutionary selection. Most of the pore lining residues in the channel are in fact hydrophobic and they provide a 'well oiled' and 'non-sticky' pathway for the water molecules to traverse the lumen.

#### 1.4.2 Selectivity of aquaporins

Recent *in vivo* and *in silico* experiments highlight an important feature of the water channels: the ability to distinguish between permeating entities. There are already several examples where the members of the aquaporin family can permeate alternative substrates. Aquaporins such as GlpF are specialized in the transport of Glycerol alongside that of water. Aquaporin-6 [30] acts as an anion channel, transporting nitrate molecules. But beyond these roles, it has been suggested that the protein family can serve in permeation of many other species of molecules. There has been evidence indicating that molecules such as urea, gaseous substrates such as carbon dioxide, ammonia and oxygen may permeate via aquaporins protein. Generally, gaseous neutral molecules are typical candidates for the application of Meyer-Overton's rule. This rule states that small and neutral molecules can readily pass through lipid membranes, while charged or polar residues may not [31]. The

rationalization behind this principle comes from the knowledge that hydrophobic solvation of small, apolar molecules in aqueous environment into non-aqueous aliphatic liquids is thermodynamically favored due to the so called hydrophobic effect. This rule, although quite simple, forms the backbone of modern rational drug design, where small drug-like molecules are synthesized to be sufficiently apolar so they can effectively diffuse across the cell membrane. Thus, intuitively, it can be claimed that gas permeation in living organisms should also occur directly through the cell membrane, bypassing the need for any dedicated channel.

However, in many cases aquaporins have been suggested to present an alternative to this route. Knocking out the gene or disrupting the function of the protein can lead to a severe loss in gas permeation. This dependence of permeability of gases on the presence of aquaporins is surprising in view of the Meyer-Overton rule. To reconcile this apparent paradox, hypothetical conditions can be proposed where Meyer-Overton's rule may not apply. The permeability of biological membranes is known to depend on the cholesterol content of the membrane. Higher eukaryotes contain a significant portion of sterols ( 20-40 %) [32] in the lipid composition of their cell membranes. This may have the effect of reducing the contribution of permeation of neutral gas molecules directly through the lipids.

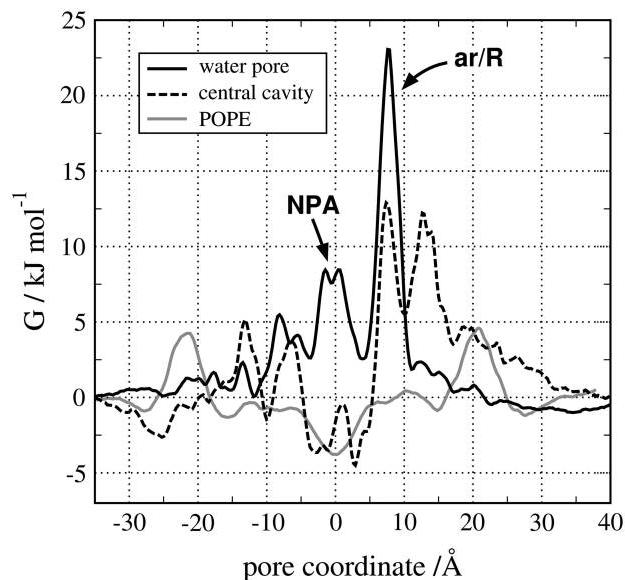


FIGURE 1.8: The free energy barrier for a CO<sub>2</sub> molecule across the AQP1 water pore and central pore compared with the barrier across a pure lipid bilayer. The AQP1 has a larger barrier for the permeation of CO<sub>2</sub> in accordance with the Meyer-Overton rule. Picture taken from [29]

Although most aquaporins are highly permeable to water, a purely structural consideration of the filter cannot explain why the channel would not permeate smaller solutes. The pore constriction, as mentioned earlier in the chapter, is approximately of the same size threshold as the diameter of a water molecule. This would imply that molecules smaller than the size of the constriction would have little trouble navigating the channel. As aquaporins are passive channels, an osmotically driven proton permeation might, for example, be expected to result in a dissipation of the proton gradient across biological membranes. This, however would be catastrophic as free permeation of protons would distort the proton motive force due to a loss of function of biological motors which involve proton pumps.

Proton translocation in bulk water is thought to occur via the Grotthuss mechanism [33], which involves shuttling of a proton between hydronium ion clusters via chains of unbroken hydrogen bonds. These clusters are highly coordinated entities solvated by hydrogen bonds. Three possible reasons for the exclusion of protons from permeation through aquaporins were proposed to explain this phenomenon. The first one was discerned from the crystallographic data. When the water molecules are co-crystallized in the channel pore, they appear to have a strict orientation along the channel axis, which involves the discontinuation of the unbroken hydrogen bonded chain between water molecules in the lumen [34], which was proposed to occur at the NPA site. As stated earlier, a continuous chain may be necessary for the function of a Grotthuss mechanism. Thus its absence could lead to loss of proton transfer. Also, Tajikhorshid *et al.* [35] suggested that a high constriction of the rotation of the water molecules lumen would additionally hinder permeation as it would inhibit the Grotthuss mechanism. The second proposed mechanism involved a purely electrostatic barrier to the transfer of protons due to the presence of a strong electric dipole in the protein [36]. The pseudo-symmetry of the protein structure leads to the formation of two opposing dipoles in the channel interior which create a large free energy barrier for the proton. This barrier would prohibit the transfer of proton in either direction. The third mechanism is concerned with desolvation effects that occur while transferring the proton from the 'bulk' to the channel [37]. When a proton enters the channel interior,

its solvation is drastically reduced, as the donors in the lumen can not completely compensate the high dielectric shielding of the solvent bulk ( $\epsilon = 80$ ) with the low screening inside the protein (approximately  $\epsilon = 8$ ). This leads to a thermodynamically unfavorable situation for the proton. Of these three effects, the first is now considered the least important. This is so because when the proton is forced through the channel in computational simulations, it can efficiently transfer via the Grotthuss mechanism, implying that an intact continuous proton chain exists in the protein, as the water molecules observed in the crystal structure are not static in the channel [36].

## 1.5 Why molecular dynamics?

The function of aquaporins is invariably coupled to their structure. In fact, most differences in the functional aspects of the members of the aquaporin family can be reconciled by observing their structural dissimilarities. The crystal structures successfully explain experimental observations that are dependent on purely static properties of the protein. However, permeation is a dynamic phenomenon and it can not be completely studied without observing the protein and the permeating solutes in motion. Furthermore, permeation occurs at a timescale of nanoseconds and the important interactions involved in the process take place at an atomistic level of detail. These conditions, although accessible to modern experimental tools, are still difficult to capture efficiently in an empirical setup and demand a high resolution and sensitivity to be studied in detail. This is where molecular dynamics simulation can assist in the search for answers. Modern computational resources can access the requisite nanosecond timescales at fully atomistic detail. The prime requirement for these simulations is the availability of high quality coordinates, which are now made available for a large selection of the aquaporin family via x-ray and electron crystallography. Using these as the starting points and combining them with semi-empirical potentials, it is now possible to study working models of aquaporins. This offers various conveniences, such as the ability to modulate the molecular features of the protein structure via in silico mutations and alchemical transformations. Biophysical processes and measurements that

are otherwise not feasible can be suitably made in such models. Thermodynamic quantities such as entropies, enthalpies and free energies can be readily obtained by utilizing both equilibrium and non-equilibrium methods. It is also possible to study microsecond timescale physics using either coarse grained methods, which trade the atomic resolution for long-term simulations or with Markov state models which can be used to build kinetics data that explains behavior temporally out of reach of atomistic simulations. Molecular dynamics has thus been used extensively to understand the mechanism of water permeation through aquaporins and to explain molecular mechanisms of regulation and modulation of permeability([20, 22, 27, 29, 34, 36, 38–42]).

The functional parameter associated with simulations is typically the permeability. In molecular dynamics the problem of measuring the permeability is complicated. The osmotic gradients required to generate a unidirectional flux are hard to set up in the *in silico* environment as this process requires maintaining a large concentration difference. It is possible to drive unidirectional flux with a purely hydrostatic pressure. But again this requires the introduction of large mechanical forces in the system. In the work presented within the thesis the collective diffusion method is used for the calculations of the Pf [43]. This method measures the diffusion of a collective coordinate formed out of the single file region within the channel pore. Diffusive hops of this coordinate at equilibrium are used to calculate the permeability of the protein, without a need for application of either osmotic or hydrostatic pressure. This method is much more robust compared to experimental approaches to assess permeability quantitatively of the protein. This advantage derives from the knowledge of the permeability associated with the concentration of the protein in the membrane, details of which are lacking in experiments.

## 1.6 Goals of this project

The goal of this thesis is to study specific instances of aquaporins water channels in order to understand the regulation of their function and the selectivity associated with it, with the hope that a light can be shed upon the generalities in the

structure-dynamics-function relationships in the aquaporin protein family. The various sub-projects in this thesis have been carried out, in collaboration with experimentalists from the fields such as crystallography, biochemistry and molecular biology. Three members of the aquaporin protein family have been studied in the thesis; AQP4, TIP2;1 and AQP0. The common principle behind the work conducted here is the focus on molecular mechanisms that are involved the regulation of the permeation and the specificity of the permeating molecule: 'Are there gating mechanisms inherently encoded in the aquaporin structure?' 'What conditions determine the activation of such gates?' 'Do aquaporins allow permeation of substrates other than water? If yes, then are these broad specificities relevant in physiological situations?' - are some of the questions addressed. Whenever possible, we have attempted to answer these questions with quantitative parameters derived from various statistical tools.

## 1.7 Organization of the thesis

The thesis is organized so as to transit from a theoretical and historical background of the work conducted to the actual research work and its analysis. The first two chapters deal chiefly with the former and the next three focus on the latter. In this first chapter of the thesis, a quick qualitative perspective is provided towards the history of research concerning the aquaporin family. The important structural and biophysical details of the protein relevant to understanding the following chapters are presented. In chapter 2, we describe briefly the important methods utilized in this project. In addition, a detailed summary is provided for the various parameters utilized in the design of the *in silico* experiments, so as to facilitate the reproduction of these results. Chapter 3 deals with the regulation of permeability in the AQP4 protein. A covalent modification of aquaporins has been suggested to be one of the important mechanisms through which permeability is regulated. In most cases, this is achieved via control of expression of the protein on the bio-membrane. However, recent evidence indicates that phosphorylation of conserved residues can activate gating behavior in aquaporins. Here, we investigate such a

possibility in the case of AQP4. Additionally, we study the effect of pH on possible gating mechanisms identified within the protein using molecular dynamics. In chapter 4 we investigate the possibility of gas permeation in an aquaporin. Here, a novel aquaporin crystallographic structure is modeled using molecular dynamics simulations and then studied for its substrate specificity. This structure demonstrates certain novel features which we reproduce in the computational models. Using equilibrium simulations and enhanced sampling methods, we explore the possibility if ammonia can permeate the protein water channel. Within the fifth chapter, we study an unusual member of the aquaporin family, AQP0. We try to decipher the molecular mechanisms behind the low permeability of this protein using machine learning methods. In a collaboration with our experimental partners, we propose a variety of mutational forms AQP0 which can explain the functional aspects of the protein and enhance its conduction. Finally, in the sixth and the last chapter, the conclusions for the overall thesis are presented. As science is a never-ending pursuit of facts and explanations, an outlook towards possible future avenues that can develop into interesting projects is also provided.







# **Chapter 2**

## **Methodology**

In this chapter we provide a brief introduction to a variety of methods used in the thesis. The aim here is to convey the general purpose and the physical basis of the methodology involved in the most important techniques used to facilitate a clear reading of the research work performed.

### **2.1 Molecular Dynamics Simulations**

Molecular Dynamics (MD) is a well established computational technique used to simulate macromolecules at a near-atomistic level. The chief advantage of MD is the ability to access properties at time and length scales which are either typically difficult to discern or outright unavailable at the experimental level. The general idea behind the methodology is to simulate the Newtonian mechanics of molecules using model inter-atomic potentials and generating time series data which can be analyzed for statistical behavior of the simulated entity. MD has been used extensively to understand thermodynamics and kinetics of molecular processes [44].

#### **2.1.1 Physical basis of MD simulations**

To our best understanding the dynamics of molecules is governed by the laws of quantum mechanics. To enable simulating bio-molecular structures, which is

the goal of the present study, certain approximations must be sequentially made to simplify the complex and computationally costly description of the quantum paradigm to simplify it to the Newtonian level. The goal of MD simulations is to study the time-dependence of the *positions* of the atoms. Other properties associated with them are of less interest in this scenario. The wave-function that describes the molecule typically contains information content that is higher than this requirement. Thus we have to extract only the requisite piece of functional interest from it.

The first approximation required to achieve this is that we can neglect any relativistic effects involved in the dynamics for the time and length scales under consideration. This allows to start with the time dependent Schrödinger's formulation [45] as the basis of our considerations. For a wave function  $\psi$  and a Hamiltonian  $\hat{H}$ , this equation is given as:

$$i\hbar \frac{\partial \psi}{\partial t} = \hat{H}\psi \quad (2.1)$$

$\psi$  is the wave-function that describes both the electronic and nuclear contributions. The Born-Oppenheimer approximation allows us to separate these two contributions so that the  $\psi$  can be written as a product of the nuclear wave-function  $\psi_n$  and the electronic wave-function  $\psi_e$ ,

$$\psi = \psi_n * \psi_e \quad (2.2)$$

Such a separation can be justified by observing that the nuclear mass far exceeds the electronic mass. This allows electronic modes in the total wave-function to relax quickly and independently of the nuclear motions. While solving the dynamics of large biomolecules we are typically not interested in the exact solution for the electronic wave-functions. Rather the nuclear wave-functions are the subject of interest as they broadly describe the motions of the atoms in space. Hence, we focus only on the nuclear part of the equation 2.1. This approximation is known as the Born-Oppenheimer approximation [46]. This is written as:

$$(\hat{T}_n + \hat{V}(r))\psi_n = i\hbar \frac{\partial \psi}{\partial t} \quad (2.3)$$

where,  $\hat{T}_n$  is the kinetic energy operator corresponding to the nuclear motion and  $\hat{V}(r)$  is the potential energy operator that represents the nuclear and electronic potential. At this point, equation 2.3 describes the time dependence of the entire wave-function. We are interested only in the time-dependence of the *expectation value* of the position. This information can be acquired using the Ehrenfest theorem [47], which states that:

$$\frac{d\langle \vec{r} \rangle}{dt} = \langle \vec{v} \rangle \quad (2.4)$$

and

$$\frac{\langle \vec{v} \rangle}{dx} = -\nabla V(r)/m \quad (2.5)$$

Equations 2.4 and 2.5 are essentially Newton's equations of motions, where  $\vec{v}$  is the expectation value of the velocity of the particle involved and 'm' is the mass of the particle. The velocity is required to calculate the future position of the same.  $V(r)$  is the potential experienced by the particle and it is typically broken down into multiple components that describe the nuclear motions in molecules. A typical breakdown is shown in table 2.1 [44]:

Forcefield terms are typically divided into *bonded* and *non-bonded* parameters describing various kinds of interactions between atoms. These forcefields thus allow us to approximate molecular interaction on a *classical* regime and simulate dynamics using Newtonian laws of motion. Forcefields can be specialized by parameterizing the individual terms in the potential to represent better certain situations. Thus the choice of the forcefield depends on the property of the macromolecule that the researcher is trying to explore. In our simulations, we have worked with either the CHARMM36 [48, 49] or the Amber99sb-ILDN\* [50] forcefields. CHARMM36 has the advantage of a large library of all-atom lipids for

| Interaction type               | Mathematical model  |
|--------------------------------|---|
| <i>Bonded interactions</i>     |   |
| Bond stretching                | $\Sigma_{bonded} k_b/2 (r_{ij} - r_0)^2$                              |
| Bond angle Bending             | $\Sigma_{bending} k_a/2 (\theta_{ij} - \theta_0)^2$                   |
| Dihedral twisting              | $\Sigma_{dih} k_\phi/2 (1 + \cos(P_n\phi - \gamma))$                  |
| Out of Plane twisting          | $\Sigma_{improper} k_\eta (\eta_{ij} - \eta_0)^2$                     |
| <i>Non-bonded interactions</i> |   |
| Electrostatic                  | $\Sigma_{i,j} \frac{q_1 q_2}{4\pi\epsilon_0 r_{ij}}$                  |
| van der Waals interactions     | $\Sigma_{i,j} - \frac{A_{ij}}{r_{ij}^6} + \frac{B_{ij}}{r_{ij}^{12}}$ |

TABLE 2.1: The terms involved in a typical potential used for MD simulations. The force constant terms are shown as  $k_i$ . The term for the phase angle is shown as  $\gamma$ . The distances between two particle indices, i and j, are shown as  $r_{ij}$ .  $P_n$  is the periodicity in the dihedral rotation potential.

which high quality parameters have been tested to simulate experimental observables. The latter, Amber99sb-ILDN\* has the advantage of having tested parameters for modified amino acids such as phospho-serine. The two major requirements for MD simulations are the forcefield parameters and a set of coordinates to start the simulation from. The latter is typically obtained from either x-ray diffraction studies, cryo-electron microscopy or NMR spectroscopy. Generally, high fidelity structures with resolution in the Å scale are needed. Usually, these are obtained from the Protein Data Bank (PDB) website. When such a structure is available, the initial velocities are generated using a Maxwell-Boltzmann distribution of velocities, assigned randomly to atoms.

### 2.1.2 Numerical integration of equations of motion

An important facet of MD simulations is the integrator. Once the initial structure and the velocities are generated and the forcefield parameters are known, the equations of motions can be used to predict the subsequent set of coordinates after a given amount of time, as well as the velocities that are to be assigned to these atoms. The function of the integrator is to determine the velocities and

positions at a later point in time by ‘integrating’ equations of motion. The timestep used for this process is an important determinant of the dynamics. The value of this parameter must be smaller than the frequency of the fastest possible mode defined by the potential used for the forcefield. Typically, this corresponds to the bond stretching mode. To account for it, a timestep in the femtosecond (fs) regime is used. The exact number depends on the kind of constraints applied to the simulation system. Sometimes it is possible to eliminate certain fast modes, such as those resulting from the light hydrogen atoms bound to carbon atoms, using a technique called virtual-sites [51], which replaces these flexible bonds with virtual particles that reduce their stretching degree of freedom. In our simulations, we have used this technique alongside the AMBER99SB-ILDN\* [50] forcefield. In these simulations, we used a 4fs timestep as it was found to be appropriately faster by an order of magnitude than the fastest mode in the structure. For the simulations with the CHARMM36 forcefield we used a 2 fs timestep, as here we dealt with an all-atom description of the simulation system. We used the leap-frog integrator for all our simulations. This integrator generates a stable, energy conserving ensemble and was found to be well suited for our needs. The Taylor series based expansion for this algorithm is given below in equations 2.6 and 2.7.

$$r(t + \Delta t) = r(t) + v\Delta t * (t + \Delta t/2) \quad (2.6)$$

$$v(t + \Delta t/2) = v(t - \Delta t/2) + \Delta t * \nabla V(r)/m \quad (2.7)$$

### 2.1.3 Thermostats and Barostats

The NPT thermodynamical ensemble is desired to be used for our simulations as it corresponds best to the experimental physiological conditions. To obtain such a condition, the simulation needs to be regularly adjusted for maintaining an average constant temperature and pressure with statistical deviation consistent with the Boltzmann ensemble. This is achieved with the use of barostat and thermostat algorithms. The ideal setup for every force field can differ in terms of which barostat or thermostat is in effect. For our Amber99sb-ILDN\* simulations

we used the Berendsen barostat [52] and the velocity-rescale thermostat [53]. For the CHARMM36 simulations the same thermostat was used and the Parrinello-Rahman barostat was used to maintain a constant pressure. The temperatures maintained by the thermostat depended on the lipid composition in the simulation boxes; they were always adjusted to be 10 degrees higher than the phase transition temperature, which would be necessary to maintain the system in the liquid-disordered phase. The pressure was always restricted to 1 atmosphere.

#### 2.1.4 Water models

Water typically is the biggest fraction of the system contributing the largest number of particles in the simulation setup. Thus, it is no wonder that a large amount of effort has been spent in developing computational models for water molecules. These models are developed with a particular forcefield in mind and are best suited when used alongside them. In our simulations for the Amber99sb-ILDN\* we used the SPC/E [54] water model while the CHARMM36 setup we employed the TIP3P water model [55], based on their efficacy in reproducing experimental membrane properties.

#### 2.1.5 Handling the electrostatics and the van der Waals interactions

The biggest contribution to the computational cost of creating a simulation trajectory is calculations involving the force terms. The bulk of these consist of the van der Waals and the coloumbic potential terms. If all the possible pairwise interactions of this form are considered for the simulation, the number can grow quadratically and make any practical application of the MD technique to large biomolecules in solution impossible. In order to facilitate a compromise, several cut-off schemes are utilized in these simulations so that interactions, that ignore effects considered insignificant to measure the property of interest. This works particularly well for the van der Waals interactions as these decay very fast due to the large negative exponents involved. However, it is observed that terminating the Coloumbic interactions with a cut off can lead to distorted simulations and thus alternative approaches are needed to account for them. In the simulation

setup used for the thesis, we used the Particle Mesh Ewald (PME) method [56] which uses Ewald summation to calculate the Coloumb interactions beyond the cut-off in the reciprocal space. For our Amber99sb-ILDN\* simulations we used a cutoff of 1nm for both the interaction types. For the CHARMM36 forcefield we used a slightly modified setup with a switch function with the switch at a distance of 0.8 nm and a cut-off at 1 nm for the van der Waals interactions. The cut-off for the Coloumbic terms was changed to 1.2. Beyond these cutoffs, the PME grid was used to estimate the coloumbic potential. These conditions were found to influence the Area per Lipid (APL) for the bilayer and provided a good correlation with experiments.

### 2.1.6 Periodic boundary conditions

To ensure that finite area/size effects do not affect simulation results, a simple yet effective method called Periodic Boundary Condition (PBC) is used to account for long range interactions. This is achieved by implementing a algorithm which reintroduces a molecule diffusing out of the simulation box back into the box by the linear translation given by the box vectors. This creates an effect similar to that of having an infinite number of identical boxes adjacent to each other through which the simulation particles can diffuse freely. Thus, the number of particles in a box is kept a constant and the concentrations of all the elements in the mix can be maintained.

## 2.2 Functional Mode Analysis (FMA)

### 2.2.1 Dimensionality reduction

One of the problems with simulation of large biomolecules is that the dimensionality of the dynamics involved creates practical barriers to the analysis of the relation between the collective motions of the structure and a function of interest. Thus, dimensionality reduction of the coordinate space of these molecules has been attempted to make the problem of analysis more tractable. Among the most popular methods is the principal component analysis (PCA). PCA is a well

established statistical technique where the 'principal components' are the first few eigenvectors of the covariance matrix of the trajectory, ordered according to their eigenvalue. The covariance matrix in a  $3N$  dimensional Cartesian space is defined as:

$$\mathbf{C} = \left\langle (\mathbf{r}_n - \langle \mathbf{r} \rangle)(\mathbf{r}_n - \langle \mathbf{r} \rangle)^N \right\rangle \quad (2.8)$$

Here, the  $\mathbf{r}_1, \dots, \mathbf{r}_N$  are the  $N$  points in the trajectory. The diagonalization of this matrix provides an orthonormal transformation which creates an eigenvector basis along which the variance in the dynamics of the molecule is maximized. The ordering of the eigenvalues can be arbitrary, but a natural descending arrangement allows us to identify the eigenvectors which contribute largely to the variance of the molecule. Typically when applied to proteins, it is found that for the eigenvalue spectrum obtained, the magnitudes decrease very rapidly along the ordering dimension and quickly identifies a sufficiently small complement of vectors that can be used for further analysis.

### 2.2.2 Relating to the function

The reduced dimensionality subspace obtained from the PCA may or may not be functionally relevant. The particular 'mode' that is described by the eigenvector could be statistically significant in terms of the collective motions of the protein, but it is not guaranteed (and observed) that the largest modes are evolutionarily associated with the characteristic we are interested in. In order to understand if these modes are related to the property of interest, the method Functional Mode Analysis (FMA) is used to create a model which best correlates with the function being explored. This is a linear model given as:

$$\mathbf{f} = \mathbf{X}\beta + \epsilon \quad (2.9)$$

Here,  $\mathbf{f}$  is a vector array of dimensionality equal to that of  $\mathbf{X}$  which itself in an array of the Cartesian coordinates obtained from the an ensemble or a

time-series. The important idea is that every member of  $\mathbf{f}$  should be directly computed from the corresponding member of  $\mathbf{X}$  and directly quantifies a property of functional interest. Equation 2.9 is a regression problem with  $\epsilon$  as the residuals and  $\beta$  is a single dimensional set of coefficients of the PCA eigenvectors used as regressors for the problem. This method provides us with a linear combination of the eigenvectors from the PCA subspace that have maximum correlation with the functional property  $\mathbf{f}$  used to train the model. Although useful, this method has certain associated issues. The regressors used to create the model describing the function are pre-determined from the diagonalization of the covariance matrix. Thus, they present a potentially very large space from which the coefficients need to be extracted. If the dimensionally reduced subspace of the 'principal' components is used, then the correlation obtained between the model and the function is jeopardized as we observed that there is no guarantee that PCA subspaces are functionally relevant.

### 2.2.3 Partial Least Squares based Functional Mode Analysis

The Partial Least Squares (PLS) algorithm presents an alternative method to generate a model and solve the regression problem presented in equation 2.9. The idea behind the PLS method is to find the *smallest* possible basis that has the maximum correlation with the functional property of interest. PLS defines  $k$  regressors,  $\mathbf{T}_k$ , successively via iteration. The generation of these new coordinates takes into account two properties; first, that each is a linear combination of the original coordinate  $\mathbf{X}$  in Cartesian space; second, that each regressor is uncorrelated to the previous one. The transformation required to get  $\mathbf{T}_k$  is given by  $\mathbf{T}_k = \mathbf{XW}_k$  involving the maximization of the covariance between the function vector  $\mathbf{f}$  and  $\mathbf{T}_k$ . This has two significant advantages over the PCA based FMA. The first is that the new basis generated from PLS algorithm is de facto associated with maximization of the correlation between the function and the structure. The second advantage is that substantial dimensionality reduction obtained from the successive lack of correlation between the individual elements of the basis.

In the thesis the implementation provided by Kribokova *et al.* [57] has been

used. This implementation itself is based on the earlier implementation of the Helland's algorithm [58] by Denham [59].

## 2.3 Enhanced sampling

Unconstrained simulations can explore local minima in the Potential Energy Surface (PES) of a biomolecule quite well. However, as the most interesting behavior is typically characterized by a transition between minima, it can be difficult to explore pertinent regions of the energy landscape due to large free energy barriers involved. Sampling and equilibration is hard to achieve in the regions lying between the minima, as the system tends to have a very short lifespan in regions of the transition where the gradient of the PES is very high. The critical information contained in the transitions is concerned with stable or metastable intermediate states, the height of the barriers included in the transition and the free energy difference between the states of interest. This information can be used to deduce the thermodynamics and kinetics of biomolecular behavior and can be employed to modulate or even alter function of the molecule. Many a time the approximate reaction coordinate or an order parameter that changes along a transition can be guessed by studying the molecular structure. This information can be used to enhance the sampling in the region of the phase space between the states of interest thus improving the overall knowledge of the process. There are quite a few techniques available today which can make use of a variety of constraints to improve the statistics in the simulation. The one we have employed here most extensively is the Umbrella Sampling methodology. A brief discussion is provided below to describe what this technique entails.

### 2.3.1 Umbrella sampling

The umbrella sampling method [60] is named so due to the shape of the normal distribution function, which it uses to infer the free energy profile along a given reaction coordinate, also known as the Potential of Mean Force (PMF). The PMF

(W) introduced by J.G. Kirkwood in 1935, along a generalized coordinate  $\zeta^i$  in a set of N coordinates  $\zeta$  is defined with:

$$W(\zeta) = -k_b T \ln \left( \frac{\rho(\zeta)}{\rho(\zeta_0)} \right) \quad (2.10)$$

where, the  $\rho$  is the distribution function for the system of interest for a given value of  $\zeta$ . The mean force acting along the coordinate of interest  $\zeta$  is given by the gradient of this potential leading to the term PMF. The PMF is a relative scale, normalized with the value computed for a particular point in the coordinate ( $\zeta_0$ ). This quantity is directly available by simple counting of individual instances of the reaction coordinate assuming that the simulation is 'converged' i.e. has explored the free energy surface 'sufficiently' well to reproduce a Boltzmann distribution. This is exactly the catch, since such a convergence is very time consuming and there is no real guarantee that it will be achieved in a finite simulation time. Umbrella sampling overcomes the barrier to exploring the regions in where the gradient of the PES is very high or where the relative free energy itself is very large, by introducing a restraining harmonic potential of the form:

$$w_i(\zeta) = K_i / 2(\zeta - \zeta_i^c)^2 \quad (2.11)$$

which adds an additional potential well at the positions  $\zeta_i^c (i = 1....N_w)$ . These restraints are governed by the force constant  $K_i$ . The advantage of this construction is that the underlying potential can be recovered after simulating 'sufficiently' in every position of interest along the coordinate. Here the 'sufficiency' depends on the degrees of freedom available to the biomolecule for the given value of the coordinate. The reconstruction itself is performed via the unbiasing procedure called 'Weighted Histogram Analysis Method' (WHAM) [61]. This method has a predetermined inefficiency that is associated with the integrated autocorrelation time of the umbrella window,  $\tau_i$ . This quantity is the characteristic of the underlying local free energy profile; the larger the local barrier, the large is the integrated autocorrelation time to reach near equilibrium situation locally. The statistical inefficiency attached to the window is given by  $g_i = 1 + 2\tau_i$ . The

reconstruction of the local unbiased distribution from the histogram of the local biased distribution is done as follows:

$$P(\zeta) = \frac{\sum_{i=1}^{N_w} g_i^{-1} h_i(\zeta)}{\sum_{j=1}^{N_w} n_j g_j^{-1} \exp(-\beta(w_j(\zeta) - f_i))} \quad (2.12)$$

where,  $n_j$  is the total number of individual points used to construct the local histogram

$h_i$ ,  $\beta$  is inverse of  $kT$  and the term  $f_i$  is the free energy constant, a relative free energy term computed for the local umbrella window. It is defined as:

$$\exp(-\beta f_i) = \int d\zeta \exp(-\beta w_j(\zeta)) P(\zeta) \quad (2.13)$$

Coupling all these constants gives us the underlying PMF for the reaction coordinate  $\zeta$ .





# Chapter 3

## Regulation of permeability in Aquaporin-4

### 3.1 Overview

#### 3.1.1 Biological role of Aquaporin-4

Aquaporin-4 (AQP4) is the main water channel in the brain and is heavily expressed at the perivascular glial endfeet (fig. 3.1) [62, 63]. Due to the location at the blood brain barrier and to the altered survival rate of AQP4 knock-out mice following experimentally inflicted brain edema formation, AQP4 has been proposed to be involved in ischemic brain edema (reviewed by Zador *et al.* [64]). Short-term regulation of AQP4 under pathophysiological conditions promoting brain edema has therefore attracted scientific interest. During cerebral ischemia, the pH decreases in the extracellular space and in the astrocytic cytoplasm [65–68]. These changes in the brain environment are likely to be involved in the pathogenesis of brain ischemia and the subsequent brain edema [65]. However, the effect of these pH changes on AQP4-mediated water permeability remains to be determined. Earlier work has suggested that AQP4 is subject to regulation through modification via covalent as well as non-covalent modifications. Phosphorylation has been suggested to be responsible for regulation of AQP4 trafficking to the membrane. Two conserved serine residues S180 and S111 have been proposed to be the targets of phosphorylation [69, 70], although evidence in favor of this

regulatory modulation was absent in subsequent experimental and computational work [42].

AQP4 water permeability has been studied in detail with computational molecular dynamics techniques [23, 42, 71]. It has been shown that an external electric field can also modulate the permeability by gating a histidine residue (H201) in the protein [72, 73], although the water permeability of AQP4-expressing oocytes was independent of membrane potentials in the physiological range [74]. There has been a computational study to show that the solute preference of AQP4 extends beyond water to other substrates such as NO, CO<sub>2</sub> and O<sub>2</sub> by Wang *et al.* [75]. Only recently, the histidine residue, H95 has been shown to participate in a gating mechanism which can regulate the permeability of the AQP4 protein [76]. However, the detailed mechanism of modulation of the permeability via this gate remains unresolved.

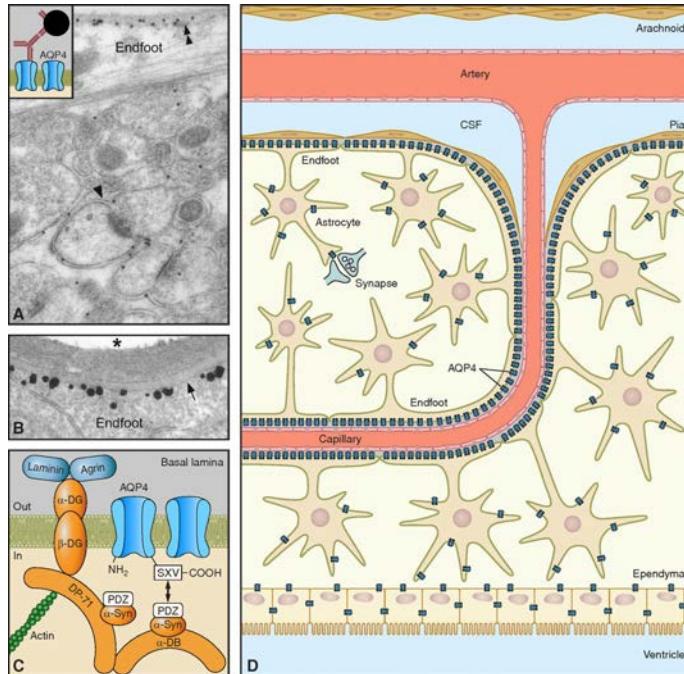


FIGURE 3.1: Distribution of AQP4 in brain. A: electron micrograph showing the distribution of AQP4 immunogold reactivity in the cerebellar cortex. B: The highest density of AQP4 is found along the perivascular glial endfoot membrane (double arrow). [A and B from Nielsen *et al.*] C: AQP4 is anchored to the perivascular basal lamina. D: Anatomy of the perivascular glial environment. Figure adapted from [77]

### 3.1.2 Regulation of AQP4 function

AQP4, as a member of the aquaporin channel family, bears close resemblance in structure and sequence to other aquaporin channels. This has led to the speculation that the regulation mechanism for this protein might be identical to those present in its relatives. Phosphorylation of two conserved serine residues S180 and S111 as mentioned earlier has been suggested to be responsible for regulation of AQP4 trafficking to the membrane [78]. The amino acid sequence of AQP4 contains consensus sequences for a range of different protein kinases [69]. Activation of PKC has been shown to lead to the down-regulation of AQP4 functionality [79, 80]. These mechanisms affect the AQP4 function by controlling its expression and localization. More relevant to the subject matter of this work, other protein kinases such as PKG and PKA have been proposed to phosphorylate S111 in AQP4, thereby inducing a gating mechanism that leads to an increase in the permeability [69, 81]. This mechanism was instigated by comparison with the crystallographic study of spinach aquaporin SOPIP2;1, where the phosphorylated structure was present in the open state [20] (fig. 3.2).

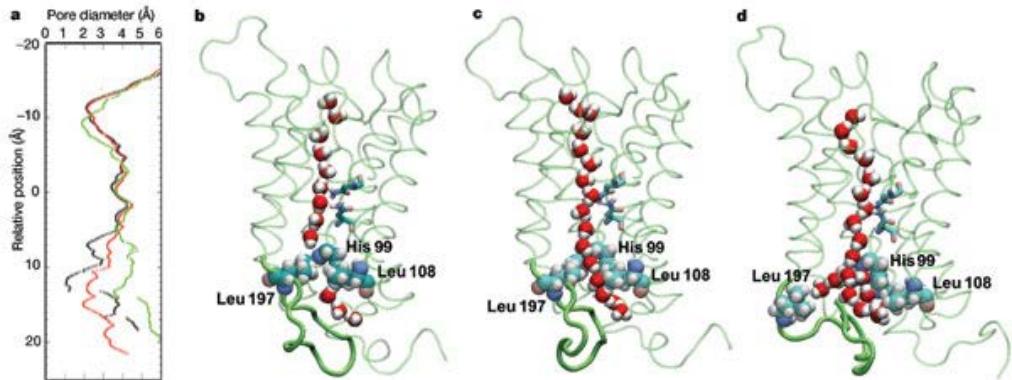


FIGURE 3.2: a. The pore diameter of the non-phosphorylated (black), phosphorylated (red), and induced open (green). b to d. Snapshots of simulations of the non-phosphorylated (b), phosphorylated (c) and induced open (d) structures. Loop D is drawn in a thicker tube representation to highlight its conformational coupling to the cytoplasmic gate. Residues forming the cytoplasmic gate are drawn in van der Waals representation. Two asparagine residues of the NPA motifs are shown in stick representation.. Adapted from [20]

In this work we explore the possibility if phosphorylation of the S111 residue

could gate the AQP4 channel. This project was collaboratively carried out with Dr. Mette Assentoft and Prof. Nanna McCaulay from the Department of Cellular and Molecular Medicine, University of Copenhagen, Denmark. We observed the predicted effects of change in the protonation state of H95 on the permeability of the channel leading to identification of a pH regulated gating mechanism. We demonstrate that such a mechanism can affect the permeability of the protein and further contribute to the understanding of the complex regulation mechanism in AQP4.

## 3.2 Computational details

A molecular structure obtained from x-ray crystallographic methods, with a resolution of 0.18nm, was used for the simulation of AQP4. It was obtained from the Protein Data Bank with the entry 3GD8. We used the package WHATIF [82] to curate the structure and predicted the protonation states of the residues using hydrogen bond networks. The tetramer generated from this structure was inserted into a lipid bilayer consisting of 294 dimyristoylphosphatidylcholine (DMPC) lipids. The standard Amber99sb-ILDN\* forcefield parameters as described in the methods chapter were used for the simulations.

AQP4 was simulated under two conditions: the unmodified amino acid sequence and with phosphorylation of S111. The parameters for the phosphorylated serine residue were obtained from Homeyer *et al.* [50]. The simulations were carried out for 500 ns each and the osmotic permeability was calculated using the collective diffusion method [43]. The first 100 ns of the simulations were discarded to account for equilibration. The osmotic water permeability was then calculated for a 50 ns window, dividing the simulations into eight slices for each monomer. The average over eight windows and four monomers was used to compare the water permeability of AQP4 with its S111 phosphorylated form.

### 3.3 Can phosphorylation of S111 gate the AQP4 channel?

#### 3.3.1 Background

Gunnarson *et al.* [69] proposed a mechanism for AQP4 gating analogous to the one described by Tornroth-Horsefield *et al.* [60] for spinach aquaporin SOPIP2;1. This analogy was based on the ideas of conservation of the S111 residue equivalent to the S115 in SOPIP2;1. However, this comparison raised concerns as crystallization of the human AQP4 yielded an open conformation, despite the lack of a phosphate group at S111. Also, the D-loop and the B-loop in AQP4 were observed to be too short to act as the gate [23]. This could principally prohibit the S111 from forming contacts similar to those described by Tornroth-Horsefield *et al.*. In vivo phosphorylation of S115 on SoPIP2;1 was not detected despite experimental activation of a range of protein kinases [83]. Furthermore, the open structure of the spinach aquaporin was obtained at a pH promoting a closed structure [20].

Oocyte swelling assays studied in the lab of our collaborators indicated that the permeability of AQP4 expressing Xenopus oocytes was unchanged due to the activation of a variety of protein kinases. But this did not rule out the possibility that the phosphorylation itself was present in the protein and that it could gate the channel. To verify if a potential phosphorylation would indeed be able to alter the permeability of the protein, we decided to model AQP4 in a fully atomistic setup. By comparing osmotic permeability ( $p_f$ ) values of the phosphorylated and non-phosphorylated forms of AQP4 we could then shed light on existence of a mechanism similar to the one proposed for SOPIP2;1.

#### 3.3.2 Comparison of the phosphorylated and non-phosphorylated forms of AQP4

MD simulations were performed to contrast AQP4 without a phosphorylation and with phosphorylated S111 (fig. 3.3). All monomers were subjected to the phosphorylation, which improved our statistics four-fold. These simulations were compared with those without such a modification. The phosphorylation makes the S111 doubly negatively charged. This modification was proposed to lead to

the closing of a gate formed by the D-loop, by forming hydrogen bonding and electrostatic contacts with the residues of B-loop in SOPIP2;1 and blocking the channel pore on the cytoplasmic side of the monomer.

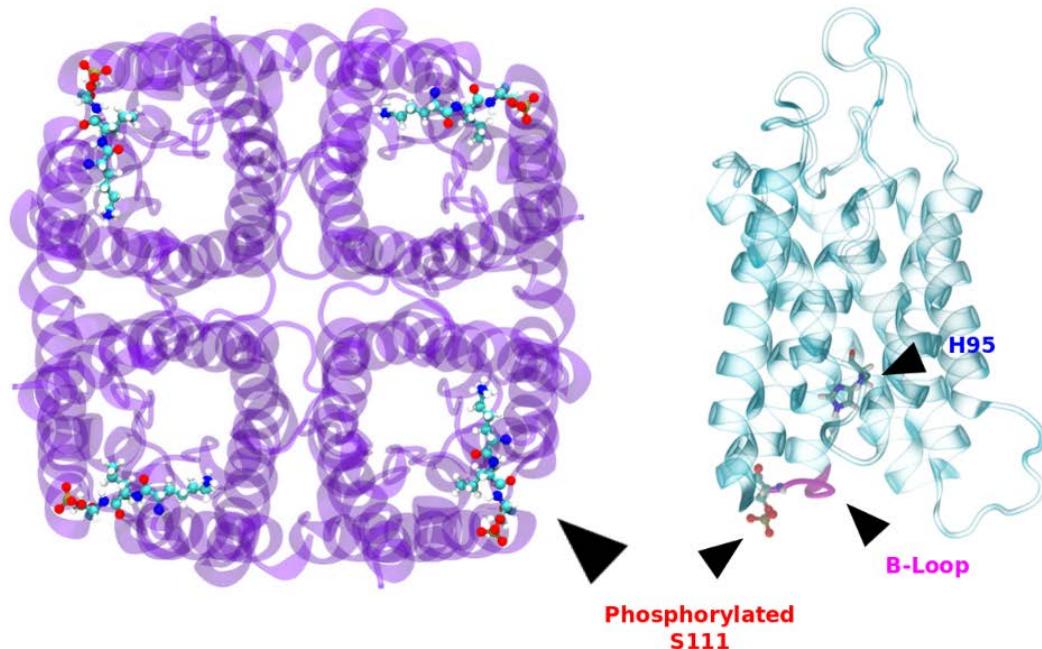


FIGURE 3.3: Top-view of the AQP4 tetramer and a Side-view of an AQP4 monomer. The B loop is shown in magenta and with a phosphate-group attached to S111 shown in ball and stick model. H95 is emphasized in licorice representation.

In the simulation time window of 500ns, we observed no significant effect of the phosphorylation of S111 (pS111-AQP4) on the conformation of the S111-containing loop B or of the loop D (fig. 3.4). To determine the effect of a phosphorylation of S111 on the flexibility of the B-loop, we compared the Root Mean Square Fluctuations (RMSF) of the backbone atoms of the loop residues over 400ns of the simulation. There was no significant change in the motion of the loop between the phosphorylated and the non-phosphorylated forms. The phosphorylation of S111 did not significantly affect the water permeability of AQP4 either , (in  $10^{-14} \text{ cm}^3/\text{sec}$ ); 1.93  $\pm$  0.41 for AQP4 and 1.94  $\pm$  0.39 for pS111-AQP4, n = 8 (fig. 3.5b). The experimental data corroborating this result is also shown (fig. 3.5a). There activation of several protein kinases via cyclic GMP

analogues expected to modify AQP4 S111 residue fail to show a significant change in permeability [83].

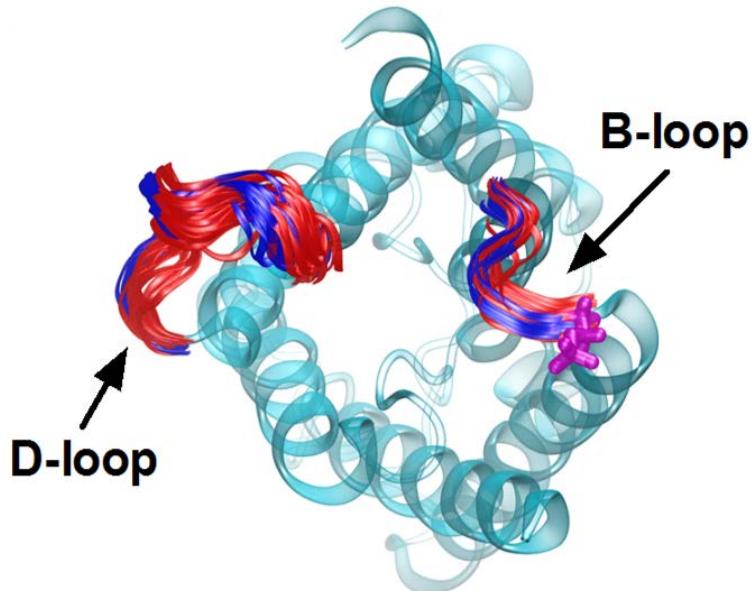


FIGURE 3.4: Phosphorylation does not lead to a gating from the loop D. The motion of the loop D and B in the phosphorylated (Red) S111 simulations is similar to that in the simulation where they are not phosphorylated (Blue) within the simulation timescale.

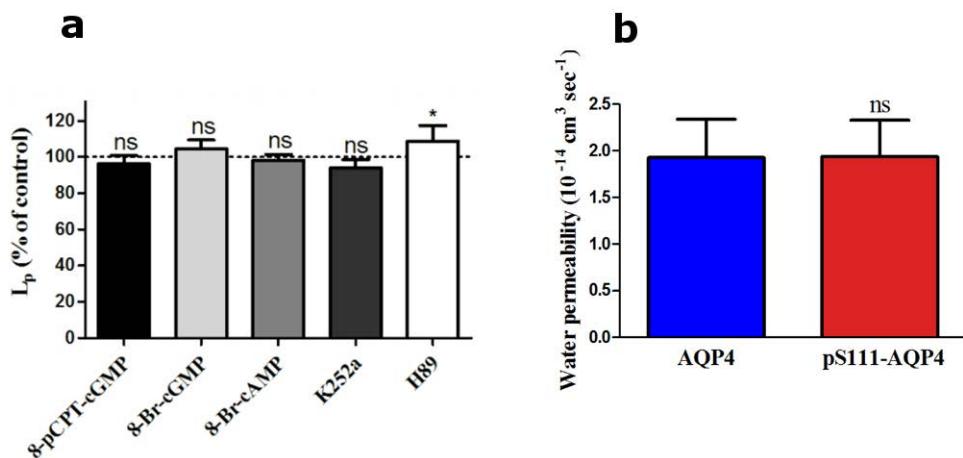


FIGURE 3.5: Molecular dynamics simulation on AQP4 revealed no functional effect of phosphorylation of S111 on the water permeability of AQP4. (a) The experimentally measured relative permeabilities of AQP4 on activation of several protein kinases expected to target S111. The activators used are shown on the x-axis [83] (b) The summarized recorded water permeability of AQP4 whether non-phosphorylated (blue) or phosphorylated (red) was not statistically different ( $n = 8$  blocks, Student's t-test).

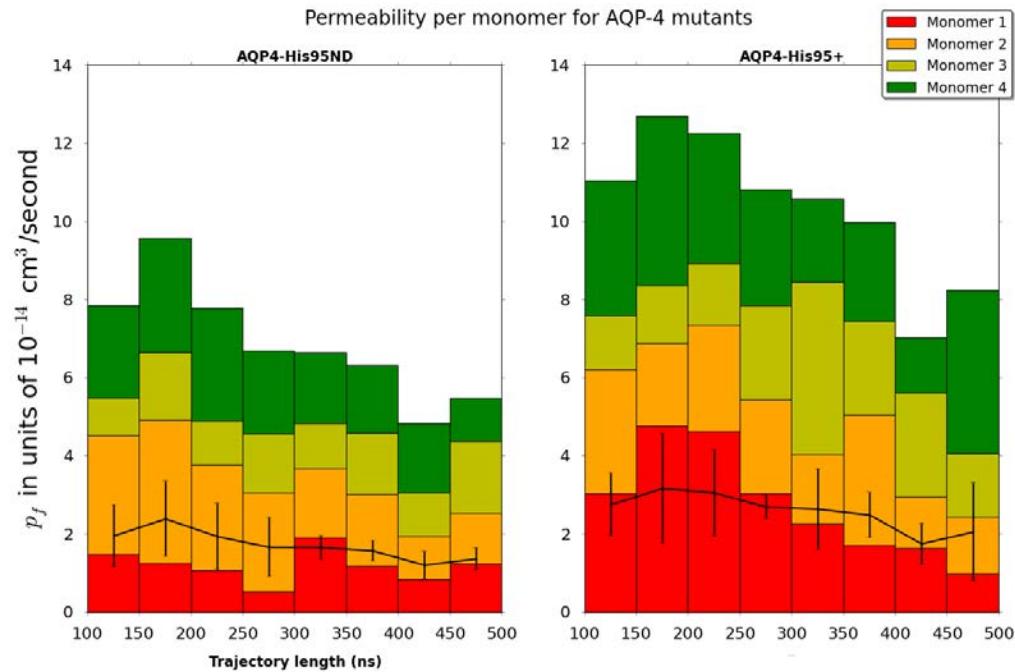


FIGURE 3.6: Osmotic permeability for the non-phosphorylated channel plotted across the simulation window. Individual permeabilities of the monomeric channels are shown to demonstrate correlation in the permeabilities. The black line is the average permeability over the monomer.

Interestingly, we observed that the permeability of the channel pore fluctuated across monomers during the simulation window. There appeared to be a consistent correlation between the monomer permeabilities, some retaining a larger value compared to others across the simulation window (fig. 3.6). This indicated that there could exist a potential mechanism that could trap the protein partially open or closed conformations. During the simulation, the residue H95 was discovered to transiently move in and out of the channel pore of the AQP4. In the initial simulations, we had predicted this particular histidine residue to have a single protonation and thus be neutral. Since this residue was placed within 1 nm distance of the S111, a protonation of this H95 could possibly interact with S111. The neutral H95 could transiently block the water pore without the interference from loop-D, while the retracted doubly protonated H95 interacting with S111 would not. To test this hypothesis, we used both the doubly protonated and the neutral (singly protonated at epsilon nitrogen) forms of this histidine in simulations. In the simulations with a doubly protonated H95 we observed no significant difference in the water permeability between the two forms of S111 in

AQP4. However, the change in the protonation state lead to increase in the observed  $p_f$  in both simulations by 30%. This result implied that the change in the protonation state of H95 and not the S111 phosphorylation was responsible for the increase in the observed permeability.

### 3.4 Is aquaporin-4 gated by pH?

In the previous section, we observed that the individual monomers of AQP4 showed distinct osmotic water permeabilities which were consistent over a long time scale. We identified H95 moving in and out of the channel pore as a possible reason that could lead to these variations. H95 is located in the pore of AQP4 and is conserved within plant aquaporins as well as in AQP4 of human, rats and mice (figs. 3.7a and b). This suggests an important role for this specific residue.

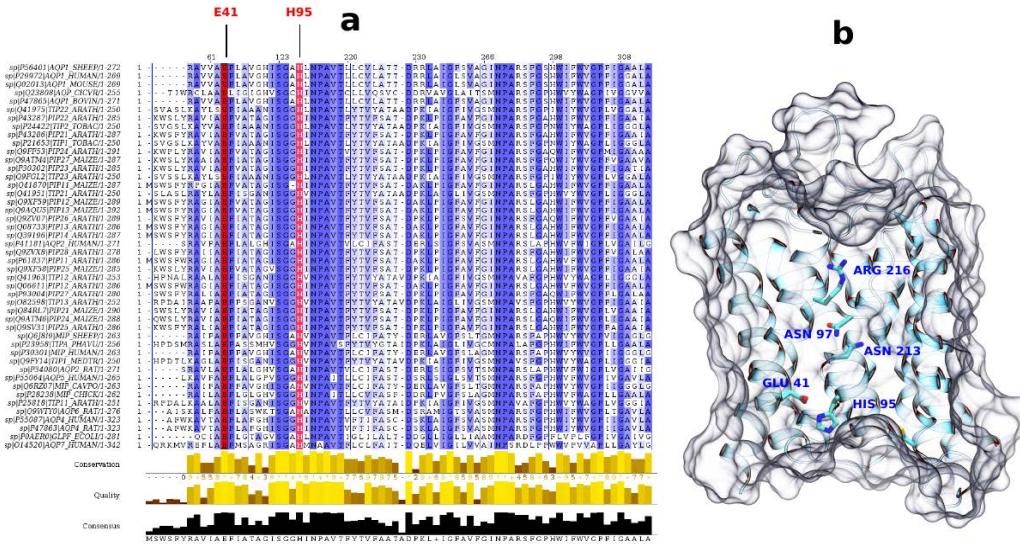


FIGURE 3.7: a. The sequence comparison is shown for the protein sequences of aquaporin family proteins. The evolutionary conservation of the H95 residue as well as the E41 residue in the human AQP-4 is highlighted in red. b. The conserved and important residues in the AQP4 are shown in the licorice representation.

### 3.4.1 Exploring the effect of H95 protonation state on AQP4 gating

To estimate the effect of the change in protonation state of H95 on AQP4 permeability with molecular dynamics simulations, we carried out two 500 ns equilibrium runs with one having H95 in all four monomers in a singly protonated, neutral state and the other with H95 in its doubly protonated, positively charged state. For the neutral state, the H95 residue was protonated in the delta nitrogen of the imidazole ring as predicted by the WHAT-IF package. The simulations showed an increased water permeability with the doubly protonated H95 residue (H95<sup>+</sup>) (compare  $1.72 +/ - 0.26 \times 10^{-14} \text{ cm}^3/\text{second}$  with  $2.62 +/ - 0.37 \times 10^{-14} \text{ cm}^3/\text{second}$ , n = 8, P < 0.038) (fig. 3.8a). To assess if this change was brought about through a purely steric alteration in the pore radius of the monomer, we calculated the radius profile for the channel (fig. 3.8b). The profile for the AQP4 channel with the doubly protonated H95 residue was observed to offer a slightly wider pore in the region where the H95 is located. However, this difference was within the standard deviation in the radius as estimated from the trajectory (fig. 3.8b). This finding indicated that the change in the protonation of the residue did not significantly alter the channel profile in a static manner. Interestingly, the H95 region in the channel profile demonstrated the highest variability in the pore radius. This observation furthered the possibility of a gating-like behavior in the general vicinity of the H95 residue, which could potentially modulate permeability of the monomer channel. As the permeability is affected by the protonation state, a pH-dependent gating mechanism seems likely.

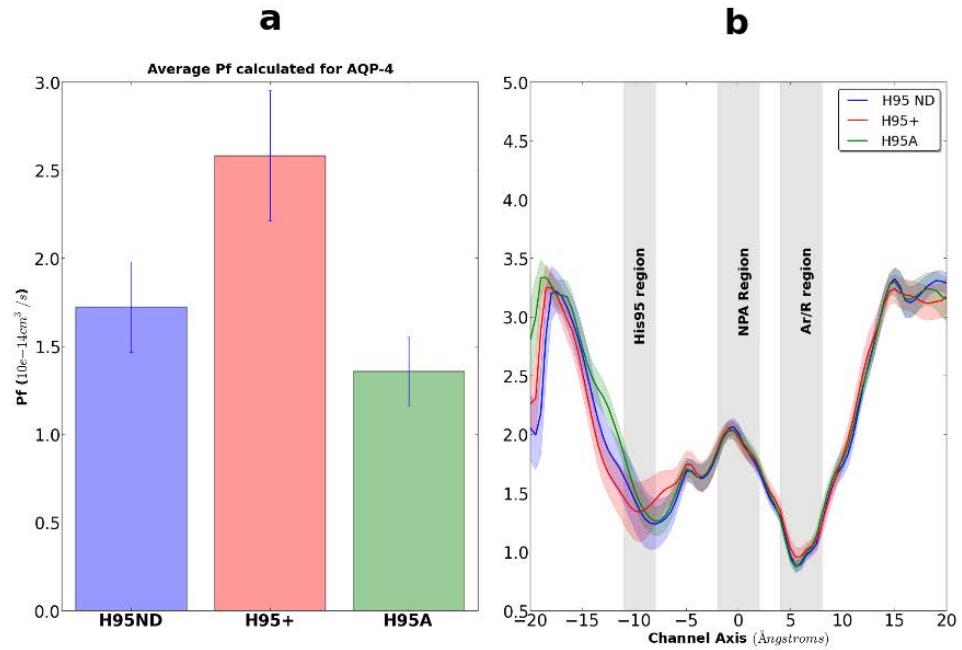


FIGURE 3.8: a. The permeabilities calculated in the the H95 Singly protonated simulation are compared with those calculated from the H95+ and H95A simulations. b. The profile of the radius along the channel lumen is plotted for comparison between the three simulations. The uncertainty is estimated with the standard deviation calculated in the trajectory. Important regions along the channel profile are highlighted.

This local radius change would not be effectively detected by techniques such as Principal Component Analysis (PCA) as such methods concentrate on identifying collective motions that maximize the covariance in the entire simulation system. This method has been proven useful in earlier studies to identify collective modes that are relevant for the functioning of the proteins of interest [84, 85]. But, such a treatment can potentially miss the collective mode of interest, if it does not contribute largely to the covariance in the simulation system. Hence, we decided to use the Partial Least Squares based Functional Mode Analysis (PLS-FMA) methodology as described in the methods section, which allows us to capture the collective motions in the protein that correlates most with the local change in the pore radius. This method attempts to detect the collective mode in a simulation trajectory that has the highest correlation with a function of interest. Ideally, the training function for the PLS-FMA model in this case would be the osmotic

permeability. However, as the permeability of the AQP4 channel appears to have a long auto-correlation time on the order of nanoseconds, we could not obtain sufficient data to generate a mode that reliably describes the change in the water permeability. Hence, we chose to use the radius of the pore as the training function as it was expected to contribute to the overall mobility of water molecules in the channel, acting as a proxy for permeability.

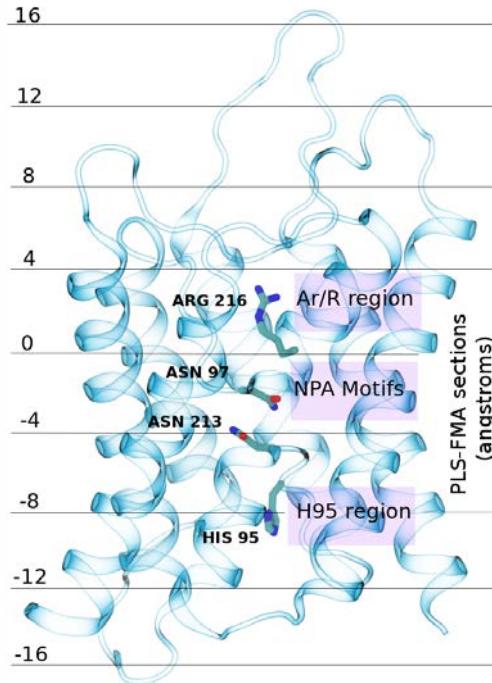


FIGURE 3.9: The AQP4 monomer is divided into 8 non-overlapping regions. From each region the lowest value of the radius is drawn to represent each frame and train the PLS-FMA model.

In this case, we used the set of local radii in eight non-overlapping regions (fig. 3.9) in the channel as the function to train the algorithm. Using the training set with the PLS-FMA method we detected eight modes in the simulation system that could potentially explain the opening and closing of the channel pore in a local region for the simulation window. PLS-FMA modes with a high cross-validation correlation coefficient ( $R^2 > 0.75$ ) were considered for further investigation. Out of the eight modes obtained through this method, three modes, in the -16 to -8 Å region, the -12 to -8 Å region and the -8 to -4 Å region, demonstrated the ability to open and close the AQP4 channel significantly by allowing the radius to fluctuate from a value of 2 Å to 0.5 Å (fig. 3.10).

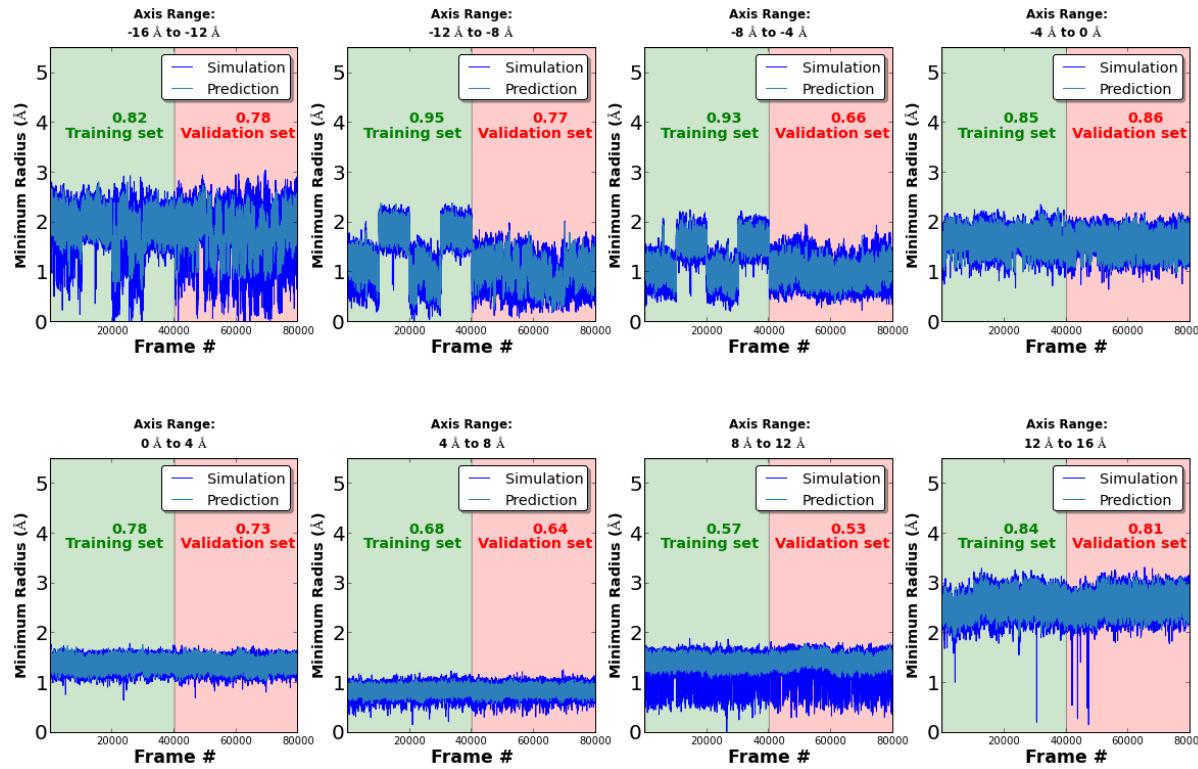


FIGURE 3.10: The AQP4 monomer is divided into 8 non-overlapping regions. From each region the lowest value of the radius is drawn to represent each frame and train the PLS-FMA model. b. The simulation data is compared with the model generated from the PLS-FMA method. Half the data show on green background is used for training the model. The rest half on the red background is used for cross-validation. The correlation coefficient for the two data are shown for each of the sets.

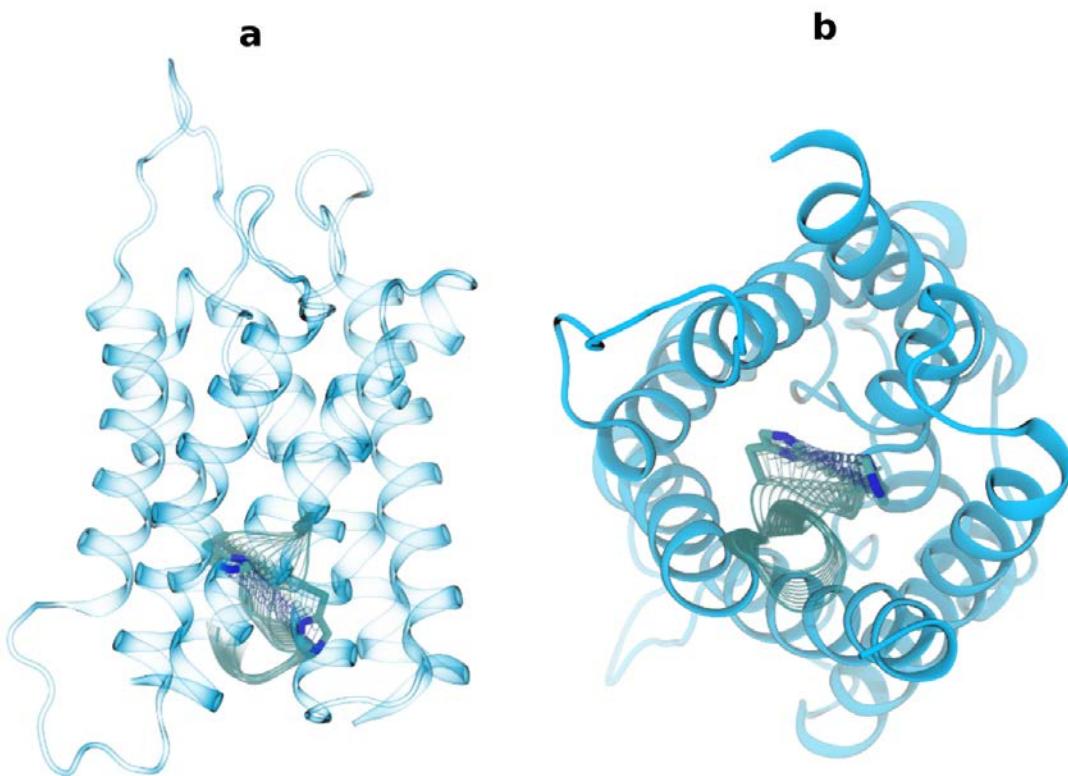


FIGURE 3.11: The mode identified from the FMA is shown with the protein in the cartoon and the H95 residue in the licorice representation. The part of the mode for rest of the protein are not shown for convenience. a. The PLS-FMA mode in the side view of the monomer. b. The PLS-FMA mode in the top-down view.

The modes obtained from other regions did not show similarly large fluctuations implying that a gate-like behavior was absent in these regions. The three 'interesting' regions are located in the vicinity of the H95 residue. The PLS-FMA mode obtained from the -12 to -8 Å region is numerically and functionally identical to the one obtained from the -8 to -4 Å region. This was expected as the H95 residue was on the border of these regions (fig. 3.9). Consequently, it had the potential to affect the channel pore radius in either region of the protein, which is what we observed in the simulation. Also, these modes have a high correlation coefficient in both the training set as well as the validation set. Therefore, the algorithm used here identified the mode that affects the pore radius in the vicinity of the H95 residue. This mode depicts the motion of the H95 residue and the associated loop to move in and out of the channel pore and is described in figure

3.11. A population histogram of the simulations with neutral H95 showed a predisposition of H95 to remain in the 'closed' state of the channel where the pore radius in the vicinity of the H95 residue is narrow ( $>1.4 \text{ \AA}$ , about the radius of a water molecule). In the simulations with the doubly protonated H95, the protein mostly occupies the 'open' state of the channel (fig. 3.12a).

Although the mode identified with the PLS-FMA method correlates with the change in the radius of the pore in the vicinity of H95, it does not guarantee that the actual function of interest i.e. permeability, is regulated by this collective motion. To test for a causal relation between radius modulation and water permeability we carried out Essential Dynamics (ED) simulations. In these simulations we trapped the AQP4 monomers into an 'open' or a 'closed' state. To accomplish that we chose values of the projection of the FMA vector representing the open-close transition in the H95 region. ED simulations were restricted to these values (+2.1 for the open structure and -2.0 for the closed structure as shown in fig. 3.12a) to obtain what we termed 'open' and 'closed' states which corresponded to a wide pore and a narrow pore at the location of H95 respectively (fig. 3.12b). In the 'open' state we observed a significant increase in the permeability ( $3.02 +/ - 0.55 \times 10^{-14} \text{ cm}^3/\text{second}$ ,  $n=8$ , a 76% increase compared to equilibrium simulations with H95 in the neutral state). On the other hand the 'closed' state showed a reduced permeability similar to that of the channel with the neutral H95 residue ( $1.78 +/ - 0.25 \times 10^{-14} \text{ cm}^3/\text{second}$ ,  $n=8$ , unchanged from equilibrium H95ND simulations) (fig. 3.12c).

This result supported the initial assumption that the PLS-FMA mode based on change in the radius profile of the channel could be effectively used to predict the permeability of the AQP4 monomer. In the simulations of the H95A mutant we observed a slightly reduced water permeability compared to the wildtype singly protonated AQP4 channel, although the difference was not statistically significant (compare  $1.72 +/ - 0.26 \times 10^{-14} \text{ cm}^3/\text{second}$  with  $1.35 +/ - 0.19 \times 10^{-14} \text{ cm}^3/\text{second}$ ,  $n = 8$ ,  $p < 0.23$ ).

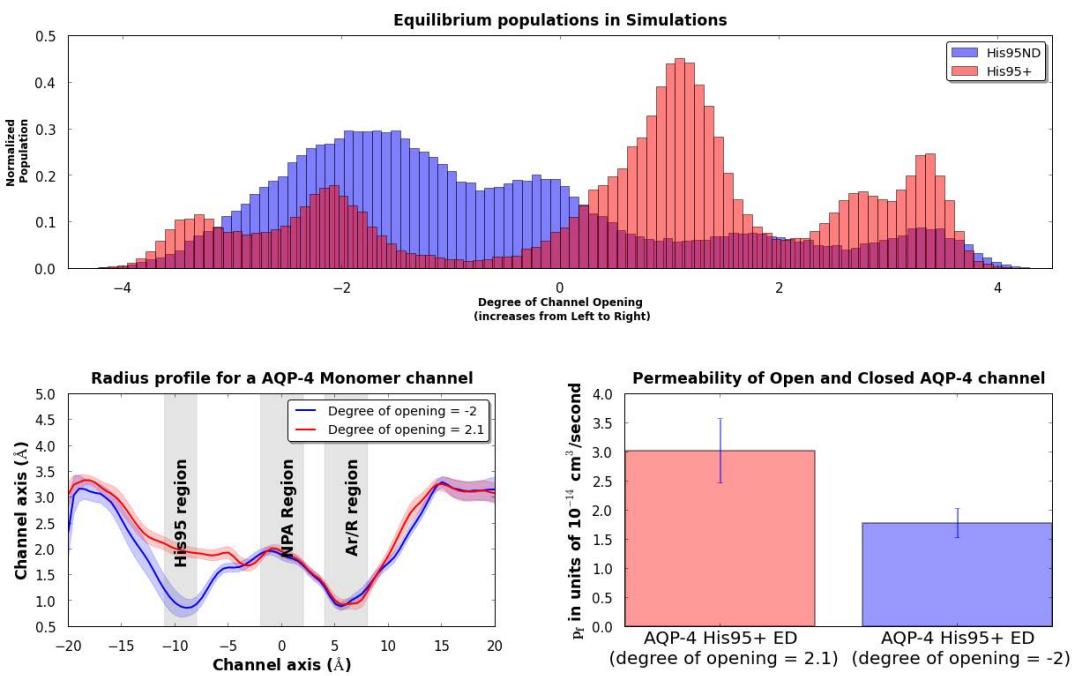


FIGURE 3.12: a. The populations of the structures encountered in the equilibrium simulations are plotted as a histogram. The structures are binned against the PLS-FMA vector describing the degree of channel opening in the H95 region. The zero on the x-axis approximately equals 1.3 Å in radius within the H95 region. b. The channel radius profile at two points along the PLS-FMA vector is shown. At a value of -2.0 on the PLS-FMA vector, the channel is constricted at the H95 region, while at the value +2.1 the local pore near the H95 region is expanded. These values were then used as constraints in Essential Dynamics (ED) simulations. c. The osmotic permeability values obtained from the ED simulations are plotted as bar graphs. The 'open' ED simulation has a comparatively larger osmotic permeability than the doubly protonated H95 simulation, while the 'closed' simulation is reduced to the level of the singly protonated H95 simulation..

### 3.4.2 Experimental verification of pH regulation in AQP4

As a logical consequence of our results, our collaborators experimentally determined the effect of pH on the permeability of the channel using the *Xenopus laevis* oocyte swelling assay.

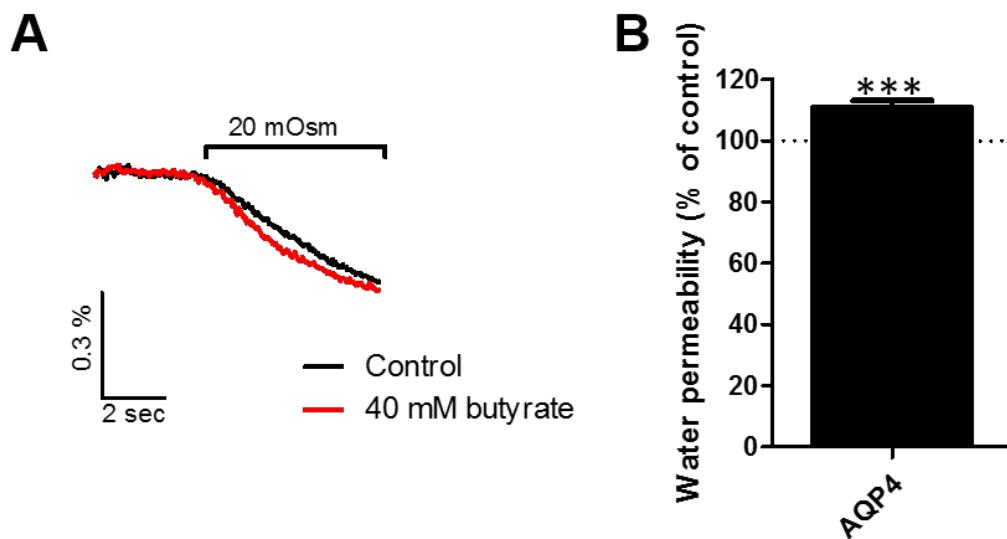


FIGURE 3.13: Intracellular acidification increases the water permeability of AQP4 expressed in oocytes. A. Volume traces from an AQP4-expressing oocyte challenged with a hyperosmotic gradient before (black) and after exposure to intracellular acidification with 40 mM butyrate (red) for 10 min. B. The bar graph shows a summary of the water permeability of AQP4-expressing oocytes after 10 min exposure of 40 mM butyrate relative to two control measurements. In % of control;  $111.0 \pm 2.2$ ,  $n = 17$ . Paired T-test was used as statistical test with  $P < 0.0005$ . Figure adapted from, [83]

In their studies, they found that the acidification of the cellular exterior had little effect on the permeability of the test system. However, the acidification of the intracellular compartment resulted in a significant increase in permeability. This indicated that the pH regulation mechanism was present in the intracellular side of the protein (fig. 3.13). Experimentally, the H95A mutant eliminated the gain in the permeability on acidification certifying that the H95 was indeed the crucial residue responsible for the gating mechanism (fig. 3.14).

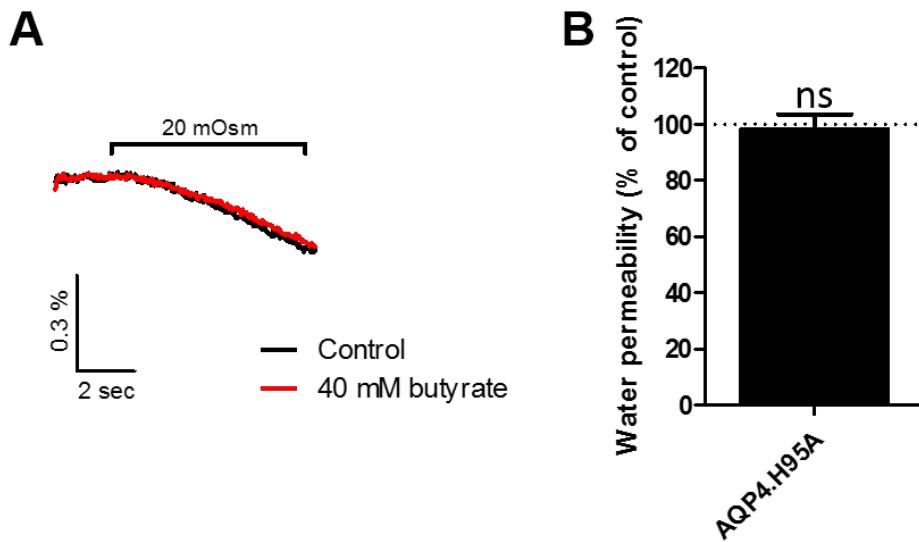


FIGURE 3.14: Intracellular acidification does not significantly change the water permeability of AQP4.H95A expressed in oocytes. A. Volume traces from an AQP4.H95A-expressing oocyte challenged with a hyperosmotic gradient before (black) and after exposure to intracellular acidification with 40 mM butyrate (red) for 10 min. B. The bar graph shows a summary of the water permeability of AQP4-expressing oocytes after 10 min exposure of 40 mM butyrate relative to two control measurements. In % of control; 98.4 +/- 5.2, n = 11. Paired t-test was used as statistical test. [83]

### 3.5 Conclusions

We show that AQP4 is not regulated via phosphorylation of the conserved S111 residue in a manner analogous to that of SOPIP2;1. The reason for the lack of this function appears to be the limited mobility of the loop D in AQP4 compared to the loop D of the SOPIP2;1. The short D-loop of AQP4 does not participate in an interaction with the B-loop in the simulation window failing to close the channel. The phosphorylated S111 does not demonstrate additional mechanisms that can affect the permeability. In fact, the average permeability of the channel in either condition appears to be equivalent.

We also describe in this work a novel pH sensitive mechanism to regulate the permeability of AQP4. Several aquaporins have previously been demonstrated to be regulated by extracellular and/or intracellular pH. Plant AQPs generally appear to decrease their water permeability with decreasing intracellular pH whereas

extracellular pH changes influence mammalian aquaporins in different manners. AQP0 has its highest water permeability at pH 6.5 [86], AQP3 closes at pH <6 [70, 87], and AQP6 opens at pH <5.5 [26]. AQP4 was previously reported not to be regulated by acidic pH down to pH 6.5 [86] but increases its water permeability at pH >8.5 [88]. These reports focus on the role of Histidine residues that are able to influence the water traffic through the channel, although not always directly.

Here, we show that pH changes, mediated by H95, can effectively gate the channel. The H95 residue has a strategic location in the channel as it is placed at the lumen-bulk interface at the cytosolic side of the AQP4 monomer. This location can expose the H95 to the osmotic environment of the cell and allow for pH sensitivity. The conservation of this residue across mammalian and plant species also points to the possibility that it might serve a functional purpose. In an earlier work [89] the role of an H95 analogue, H67 for AQP5, has been postulated to act in a similar role to modulate the permeability in that protein. The work of Alberga *et al.* presents a mechanism very similar to what we observe in our simulation. This work also identified H95 as the key residue driving the gate opening and closing. Our investigation reveals that in fact it forms the basis for pH regulated gating.

We employed a combined experimental and computational investigation strategy to identify the mechanism by which pH can regulate the permeability of the AQP4 channel. We expected to see the pH sensitive elements to be present on one of the lumen openings of the monomer channel, so that they would effectively detect a pH change in the environment. To assess which side of the protein responds to the pH change, our collaborators acidified the two compartments independently and measured the change in permeability across the cell membrane. The acidification of the extracellular compartment did not induce a significant shift in the permeability. The acidification of the intracellular compartment, however, led to an increase in the measured osmotic permeability, indicating the presence of the mechanism of pH sensitivity located on the cytosol facing lumen.

As histidine residues have a pKa close to the physiological pH we expected to see them have a significant role in responding to changes in the bulk hydrogen

ion concentration. To check if this residue could actually affects the osmotic permeability of AQP4, we built computational models analogous to the experimental system. Using MD simulations performed for these models, we compared the effects of the differential protonation of the H95 residue on the permeability of the AQP4 channel. We found that the simulation findings qualitatively agreed with the experiments with the doubly protonated H95 showing a larger permeability when compared to the simulations with the singly protonated H95.

To understand if the molecular basis of the mechanism behind this pH sensitivity was due to a physical change in the channel lumen we used the simulations to identify the effect of the double protonation on the radius profile of the channel. We found that the profile for the H95 protonation variants was statistically similar in our simulation window. But the specific radius of the pore region in the vicinity of the H95 region fluctuated strongly compared to the variation in the radius observed in rest of the monomer. This result indicated, possibly, a direct role of the H95 residue in regulating the pore opening.

To examine this possibility further without introducing a bias, we used the machine learning algorithm, Partial Least Squares based Functional Mode Analysis (PLS-FMA) to generate models that could explain the change of the channel radius in the protein. Using this methodology, we identified a collective mode that could potentially regulate the opening and closing of the monomer channel. This mode had a strong direct contribution from the H95 residue, which was found to be physically responsible for controlling the local pore radius by moving in and out of the channel lumen. Upon further analysis of the simulation trajectory with this information, we found that the doubly protonated H95 simulation largely preferred to occupy the 'open' state of this mode, while the singly protonated H95 simulation mostly occupied the closed state. To investigate whether this mode was indeed the cause of the observed osmotic permeability variation in the simulation, we carried out further ED studies which tested the ability of the this motion to modulate the permeability of the channel. By locking the AQP4 in successively 'open' and 'closed' states we could indeed influence the osmotic permeability of the channel. To validate the results from the simulations with empirical observations, we carried out mutational experiment where we replaced H95 to alanine. This

mutation was expected to abolish the pH sensitivity in the wildtype AQP4. Indeed the acidification of the cytosolic compartment was unable to affect the osmotic permeability of the mutant within any statistical margin. We speculate that the reason for the change in the preference of the H95 orientation is in the electrostatic interactions within the protein. The doubly protonated H95 residue is placed in the proximity of a highly conserved glutamic acid residue (E41) which could potentially trap the positively charged H95 in an open position. This mechanism differs from the one proposed by Alberga *et al.*, as the latter requires the formation of a hydrogen bond with a C178 residue. Although C178 can act as a potential hydrogen donor, we do not observe it to engage in a hydrogen bonded interaction in our simulation window of 500 ns. Instead, we find a significant change in the sampling of the 'open' and 'closed' states of the H95 mode dependent on the protonation state of the H95 residue.

Overall, we report a pH modulated gate for AQP4 that regulates the permeability by modulating the local channel radius. The pore modification calculated in the simulation is significant and can affect the traffic through the channel. The pH sensitivity offered by the H95 residue combined with its evolutionary conservation points towards a physiological role where the channel potentially responds to shifts in the osmolarity in the cytosolic compartment. Thus it can act in a manner that would allow the drop in the cytosolic pH to affect the channel to open wider and through osmotic pressure increase the flow of solvent from bulk into the cell. It must be noticed that the constriction of the channel at the ar/R region is the narrowest region of the monomer lumen and poses the largest barrier to permeants. This barrier appears to be directly maintained by the physical presence of the R216 in the channel pore. This observation would explain why the radius modulation from the H95 residue cannot completely switch off the channel but regulates the permeability more moderately.







# Chapter 4

## Modulating the permeability of Aquaporin-0

### 4.1 Introduction

#### 4.1.1 Role of aquaporin-0

Aquaporin-0 (AQP0) (fig. 4.1) is the water channel specifically located in the ocular lens [90–92]. The lens plays an important role in vision, as it focuses the incoming light for the eye into the retina. The lens must also be flexible enough so that it can adjust to accommodate light from differing distances for focusing by changing its focal length. AQP0 is predominantly localized in the fiber cells of the lens. There it plays a dual role as a facilitator of water permeation and as the protein that forms membrane junctions between neighboring fiber cells of the lens. The lens is an avascular structure, where the need for nutrients is fulfilled via a circular flow facilitated by an uniform and concentric formation made up of lens fiber cells. These cells are 'empty' in the sense that during their maturation they extrude all the organelles. These cells needs to be in tight conjunction with each other for the proper function of the lens, as the small distance between them (smaller than the wavelength of visible light) reduces the scattering loss and increases transparency. Forming these "tight junctions" has been predicted to be an essential role of the AQP0 protein, as it has been demonstrated that mutations in the protein or its deletion leads to congenital cataract of the eye [93, 94].

In the role of a water channel the protein AQP0 helps regulate the shape change of the lens by altering the size of the fiber cells [93]. Surprisingly, AQP0 is a poor water conductor when compared to aquaporin-1 (AQP1) [95]. Its permeability has been shown to be at least an order of magnitude less than that of AQP1, for which no physiological reason has yet been deciphered. More importantly, the mechanism of this low water permeability is also mostly unexplored, although several mutations have been suggested that lead to a drastic change in the water conductance.

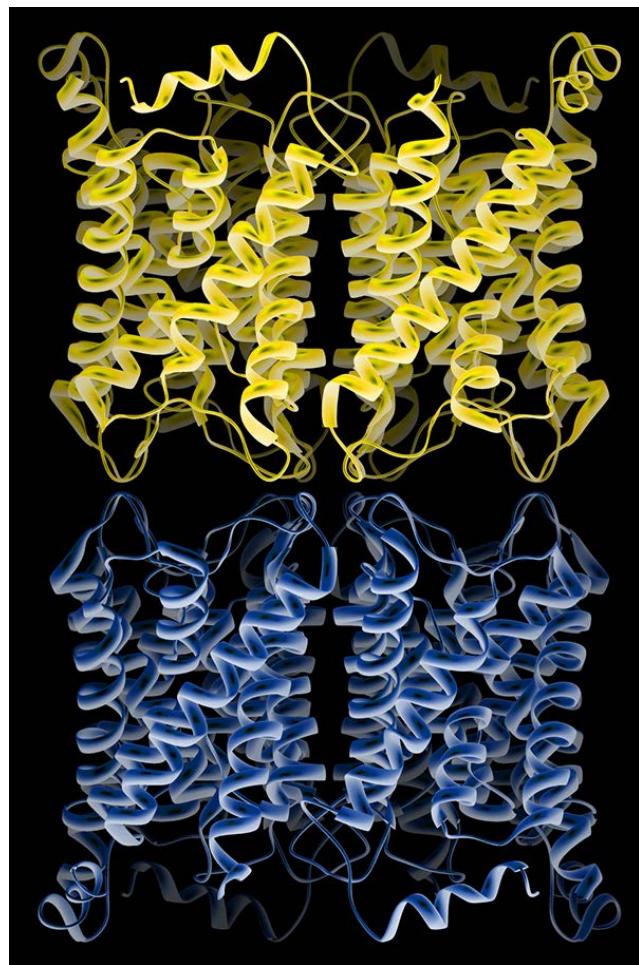


FIGURE 4.1: Structure of the aquaporin-0-mediated membrane junction [96]

#### 4.1.2 Structure of aquaporin-0

The comparison of the sequence of AQP0 with other aquaporins has resulted in the observation that there are two unusual tyrosine residues present in its sequence. Y23 and Y149 occupy locations that can potentially obstruct the passage of water through the channel pore (fig. 4.2) [79, 97, 98]. In AQP1 these residues are replaced with phenylalanine and threonine respectively. This results in a very narrow pore radius in AQP0 close to the ar/R region. Furthermore, it has been shown that the pH can in principle affect the permeability of the protein with as much as 3.4 times by Nemeth-Calahan *et al.* [86]. This pH sensitivity has been attributed to the H40 residue present on the cytoplasmic side of the protein. There also have been speculations regarding the H66 residue which is present in the cytoplasmic lumen of the protein, acting as a pH based gate. This residue is an analogue of the H95 of the AQP4, and is closely spaced along side the Y149 residue that can act as a potential channel gate as well. Overall, the pH sensitivity of AQP4 could result from both the H95 and the H40 residues.

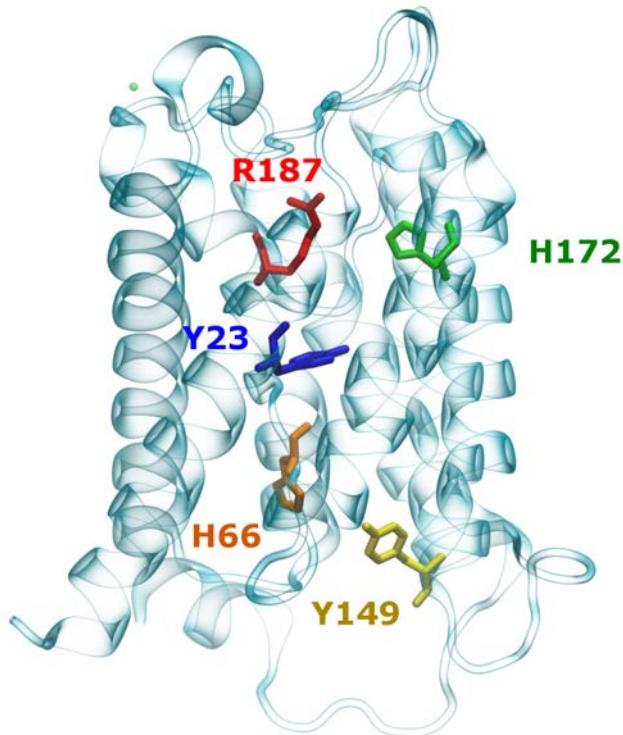


FIGURE 4.2: Cartoon representation of the AQP0 monomer with the important residues in the lumen highlighted in the licorice representation.

The structure of AQP0 was first explored with electron crystallography by Gonen *et al.* in 2004 [99] (PDBID 1SOR). This structure was identified as closed as it was crystallized without any trapped water molecules in the channel. As the protein is expected to be in the open state at the crystallization pH of 6 units, it was proposed that its closed state was due to the formation of membrane junctions, where it was present in a low-permeation state. In the same year another crystal structure of the protein was published [PDBID 1YMG] [100] that utilized X-ray crystallography. The resolution of this structure was comparable to the previous study (2.2 Å). This structure however was termed open due to identifiable water molecules in the channel pore. This new structure was crystallized at pH 10 and it was observed that it remained in the open state despite the latter pH condition being known to prohibit permeability. This indicated that the pH gating may not introduce a static blockade on the water conductance. In the following year, Gonen *et al.* published a crystal structure of AQP0 [PDBID 2B6O] using electron-ray crystallography at resolution of 1.9 Å [96]. Interestingly, this structure was crystallized in an open state in contact with the neighboring lipid molecules in close association. All the three crystal structures agreed in the observation that the channel offered a large barrier to conduction of water molecules through the lumen as it had regions of diameter smaller than twice the radius of a water molecular sphere. This static view, however, can be explored in further detail using computational efforts. There have been efforts to model the AQP0 water conduction using molecular dynamics simulations. Hashido *et al.* [98] have studied the simulations of a variety of Aquaporins to conclude that the permeability can not be completely explained purely based on comparison with the channel pore radii. Han *et al.* [79] showed that the low water permeability can be observed in the simulations thus qualifying somewhat the static picture as presented by crystallographic structures. Qiu *et al.* [97] mutated the Y23 to a phenylalanine residue to observe a 4 fold increase in the permeability of the channel. Aponte-Santamaria *et al.* [101] explored the interactions of the protein with the lipid environment co-crystallized with AQP0 to reveal a strong preference for the arrangement of the lipid tails neighboring certain sites in the protein.

In this work, we perform several computational experiments where we test

the effect of mutations on the residues predicted to be involved in the low permeability of AQP0. The experimental work associated with this project was carried out in the lab of our collaborators Dr. Manish Kumar and Prof. Tom Walz.

## 4.2 Motivation

AQP0 has a low native permeability in the experimental setup. Our collaborators, Dr. Manish Kumar and Prof. Tom Walz, observed that certain mutations in the pore lining residues Y23 and Y149 to phenylalanine and threonine, respectively, resulted in an increased permeability of the channel to the level of aquaporin-1 (fig. 4.3). The goal of this project was to identify if we could reproduce this effect in molecular dynamics and if we could try to understand the molecular mechanism that governs the change in the permeability of the protein as a result of these mutations.

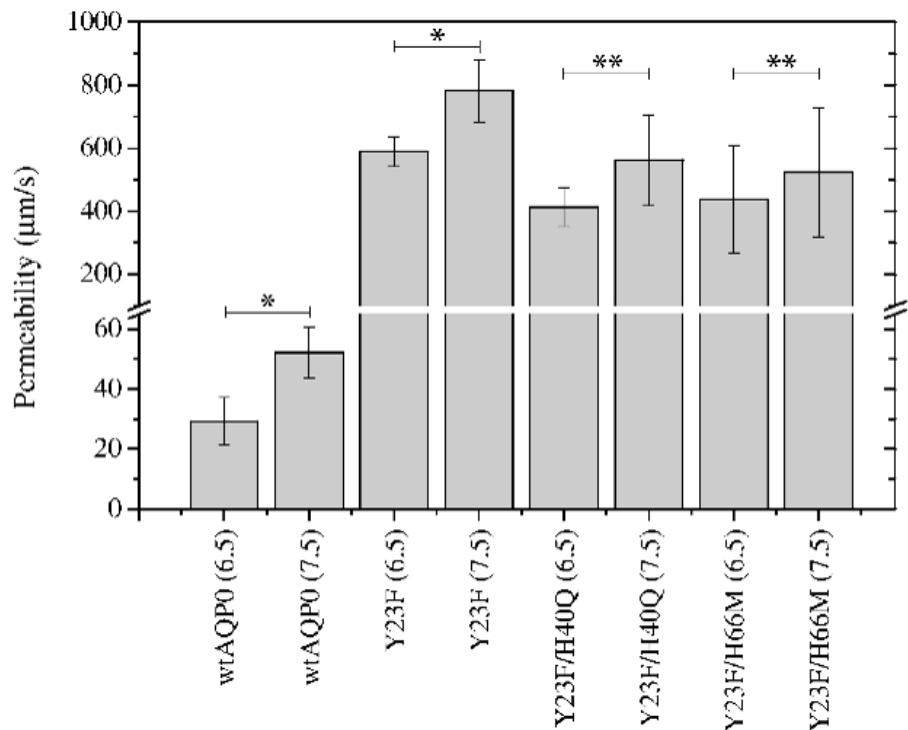


FIGURE 4.3: Osmotic permeability experimentally measured for various AQP0 mutants by Dr. Manish Kumar and Prof. Tom Walz.

### 4.3 Computational details

AQP0 and its mutants were modeled with a CHARMM36 setup adapted for protein-lipid simulations in GROMACS as described in the methods chapter. We used DMPC lipids for simulating the protein environment as the original crystal conditions have these lipids in the environment. Five mutants were generated for the simulations: Y23F, Y149T,H40Q and two double mutants Y23F/Y149T Y23F/H40Q. Of these, only the H40Q mutation did not affect residues in the channel pore. Each mutant was simulated for 500ns without restraints and the Pf was calculated by the collective diffusion method over 50 ns windows after dropping the first 100 ns for equilibration. Essential Dynamics simulations were carried out for 100 ns with a restraint of  $1000 \text{ kJ.mol}^{-1}.\text{nm}^{-1}$  to keep the protein in an open and closed states, respectively.

### 4.4 Can we mutate AQP0 to have high water permeability?

Our first hypothesis was that the structure of the protein is linked closely to the permeability of the channel. Within this conjecture, it can be directly regulated by modulating the width of the channel pore or by altering the thermodynamics of the permeation itself by affecting the protein interactions with the water molecules. To understand what environment the water molecules experience and to identify the chief players in the permeation pathway, we plotted the radius profile of the native protein and the residue population that interacts most strongly with the waters in the lumen by presenting their side-chains and polar atoms of the main chain to the lumen surface (fig. 4.4).

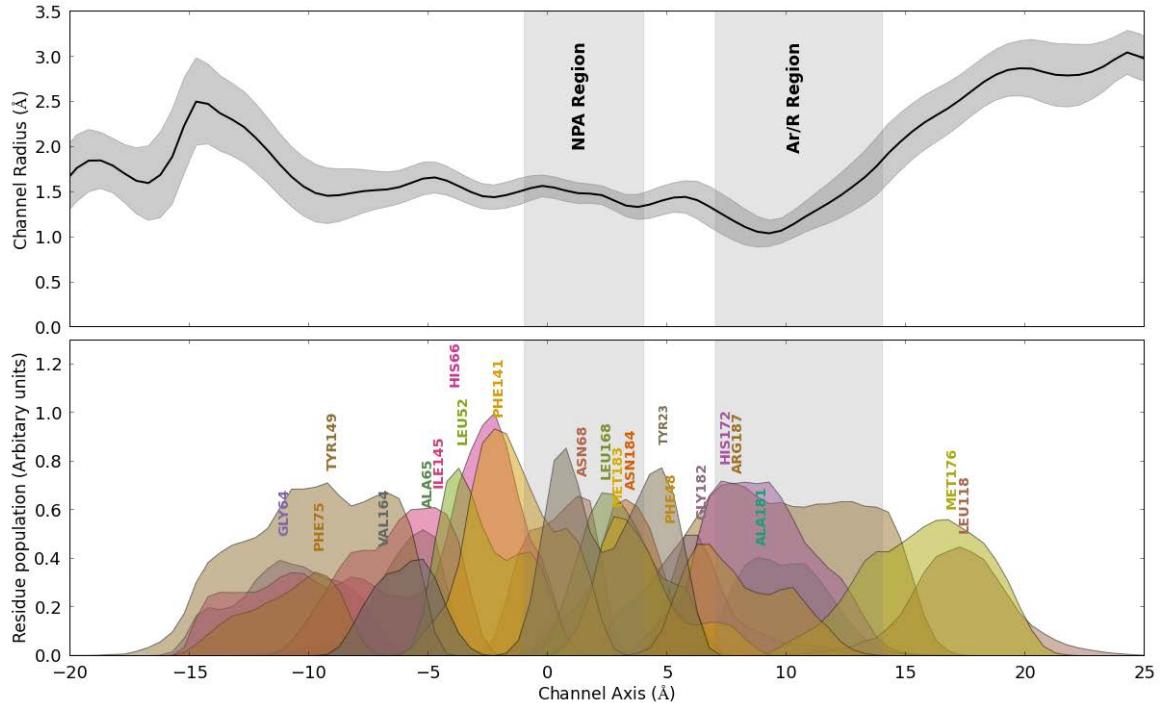


FIGURE 4.4: Radius profile for the AQP0 channel obtained over the simulation trajectory. The various important region of the protein where mutated residues are present are also shown. The shaded region is the standard deviation in the radius calculated over latter 400ns in the 500ns simulation window.

As expected, the narrowest region of the protein lies in the conserved ar/R region (here, near the R187 residue). Incidentally, the Y23 and the Y149 residues strongly participate in the formation of the channel interior as well and seem to present multiple peaks in their distribution along the channel, possibly hinting towards multiple possible stable states in the dynamics of the protein. In the previous chapter concerning the AQP4 protein, we established that the channel pore radius could be related to the permeability of the protein. So as a first guess, we compared the pore radii of all the simulated mutants with each other as well as with AQP1 (fig. 4.5). This, we expected, would give us a preliminary understanding of the channel flexibility as well as any static effect of the mutations on the channel width.

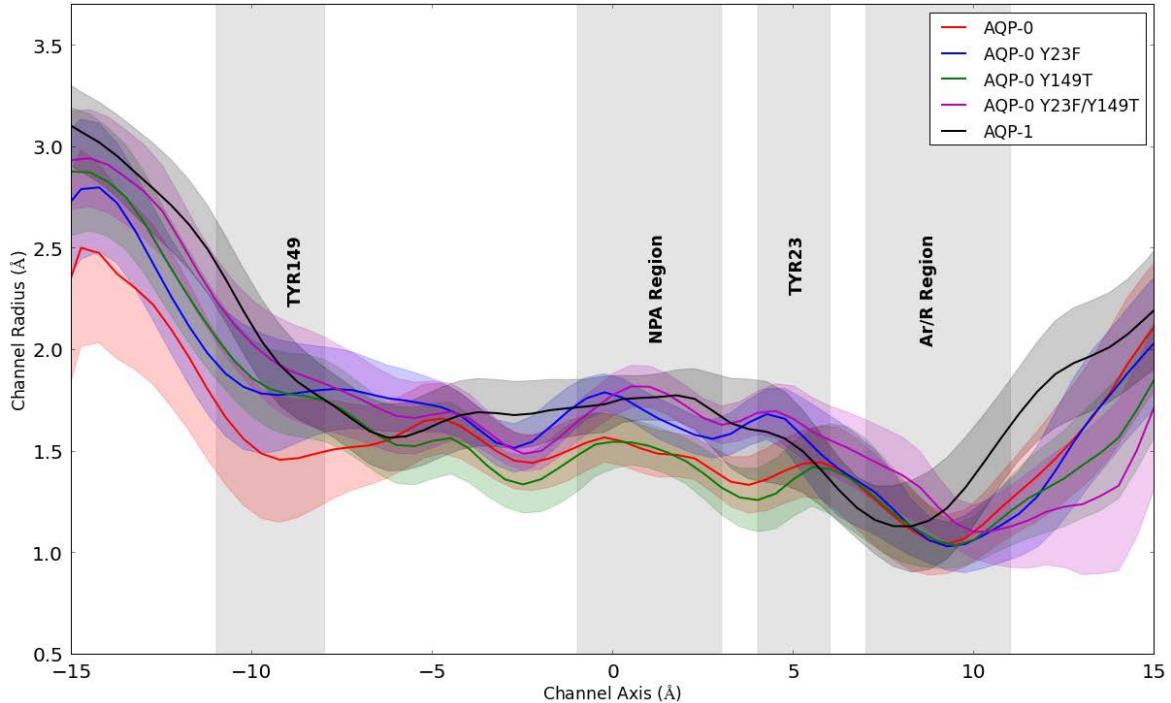


FIGURE 4.5: Radius profiles for the AQP0 protein and its mutants. The comparison is shown with respect to the AQP1 protein. The various important region of the protein where mutated residues are present are also shown.

The comparison shows that the simulated channel profile of the AQP0 protein in the native form is almost uniformly narrower compared to the AQP1 protein. The mutants created by our collaborators were constructed to bring the native protein increasingly closer to the AQP1 sequence. The channel profile for these show expected departures towards the target profile in the specific regions of AQP1. For example the Y23F and the Y23F-Y149T double mutant, both have an increased channel radius, closer to that of the AQP1 protein in the region where the mutation in the Y23 is present. Similarly, the Y149T mutant and the Y23F-Y149T double mutant (henceforth referred to as the "double mutant") have a larger radius in the region where the Y149 residue is present. However, in this region, the Y23F mutant also presents a larger radius in the profile. These differences in the simulation are highly suggestive as they allow us to anticipate that the mutants could have a larger permeability compared to the native protein based solely on the static changes we see here. However, the standard deviation observed here in the radii overlap and weaken such a direct argument. Nevertheless, they are

indicative of flexible regimes within the protein channel and point towards a combined role of dynamics with the static structure of the protein in the regulation of the permeability.

We calculated the permeability in the computational models for these mutants with the collective diffusion method of Zhu *et al.* [43]. These we compared to the AQP1 permeability so we could understand which mutation led to the largest change in the permeability bringing it towards the target value. The comparison is shown in figure 4.6 below.

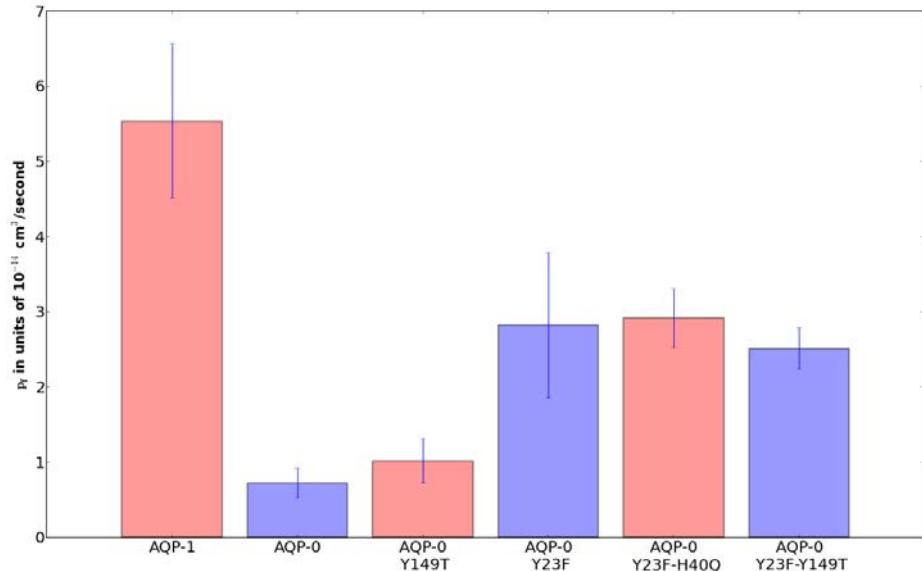


FIGURE 4.6: Osmotic permeability computationally calculated for various AQP0 mutants. The error bars are standard deviations calculated over 8 windows in the simulation trajectory.

We find that AQP0 presents the smallest permeability as expected. However, the measured value is 3 times larger than the one observed in the experiment. This could be due to the limitation of the method used for the calculations as the baseline measurement for a closed AQP channel seems to be in the same range as the one calculated here. This renders any real comparison with the Y149T mutant difficult. But the Y23F mutants always demonstrate a permeability in the same order of magnitude as the AQP1 protein, thus agreeing qualitatively with the experiments. The double mutant Y23F-Y149T does not show an additive effect

in the conductance as demonstrated in the experiments. Hence, for the rest of the chapter we focus on the Y23F mutation and try to decipher the mechanism behind its effect in detail as its effect is clearly discernible in the simulations.

## 4.5 What governs the low permeability of AQP0?

To understand how the Y23F mutant differs from the native protein, we utilized a similar strategy as employed in the AQP4 chapter. This can be rationalized by observing the dependence of the radius profile of the protein on the mutation. Functional mode analysis could provide here an important insight into dynamical aspects of the relation between the profile and the mutation. It would then be possible to test this relation to validate if a more direct causal effect of the mutation to the permeability exists, using essential dynamics simulations. To that effect, we built a Partial Least Squares (PLS) based linear model to connect the time series of the radius data to the structural change in the protein. The same methodology, of selecting the smallest radius in a pore region to detect a functional mode, as from the previous chapter was employed. The only difference within the construction here, was that we used 4 sections instead of the 8 as used previously. This was chosen, as we were interested in the specific regions where the mutations were present and 4 of these covered the residues in question adequately.

Once again, a linear model was trained by using data vectors obtained from the simulation trajectory. Each vector consisted of the smallest radius in each region of the monomer per frame as the training function. The reason for this choice was to identify a concerted motion within the protein that could lead to the closing or opening of the channel pore in that particular region of the channel interior. No initial bias was included for the model, as the entire monomer structure was used for the process of training. Only half the data was used for the training itself while the rest was used for cross-validation of the model. The model was tested by using the Pearson coefficient of correlation ( $R^2$ ) in the cross validation set. Values of this quantity  $>0.75$  were considered adequate for a reasonable model. We expected to see local collective modes emerging from the individual training sets.

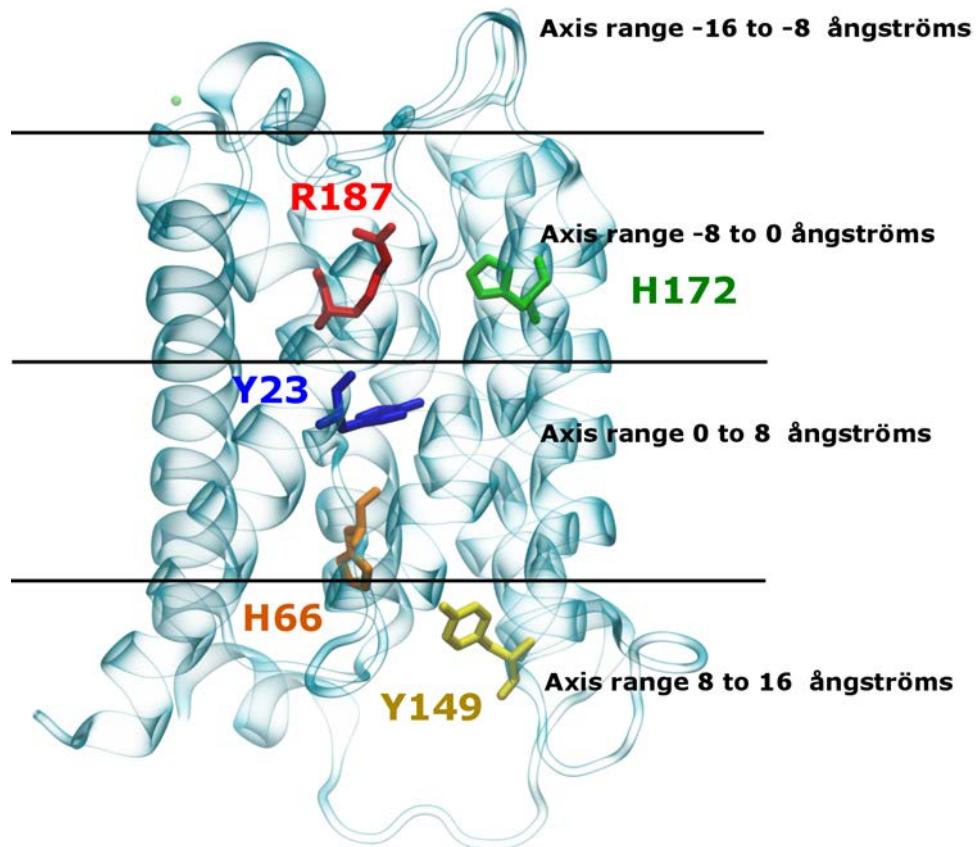


FIGURE 4.7: Subdivisions of the AQP0 monomer as used for the PLS-FMA model. Each region is 8 Å in length. Important residues in each region are highlighted.

To determine the optimal number of components to be used for the PLS modeling, we plotted the correlation coefficient in the validation set against increasing number of PLS components. We expected to see the correlation in the validation set to reach either a maximum or a plateau value when the optimum value was reached (fig. 4.7). With this treatment we identified 15 components as an optimum for our purpose without leading to an over-fitting (fig 4.8). The PLS-FMA methodology could not identify a model with a high correlation coefficient ( $>0.75$ ) in the region of -16 to -8 Å in any of the protein trajectories. Hence the data for that region is not shown. For the other regions however, a maximum correlation coefficient was obtained when 15 components were used.

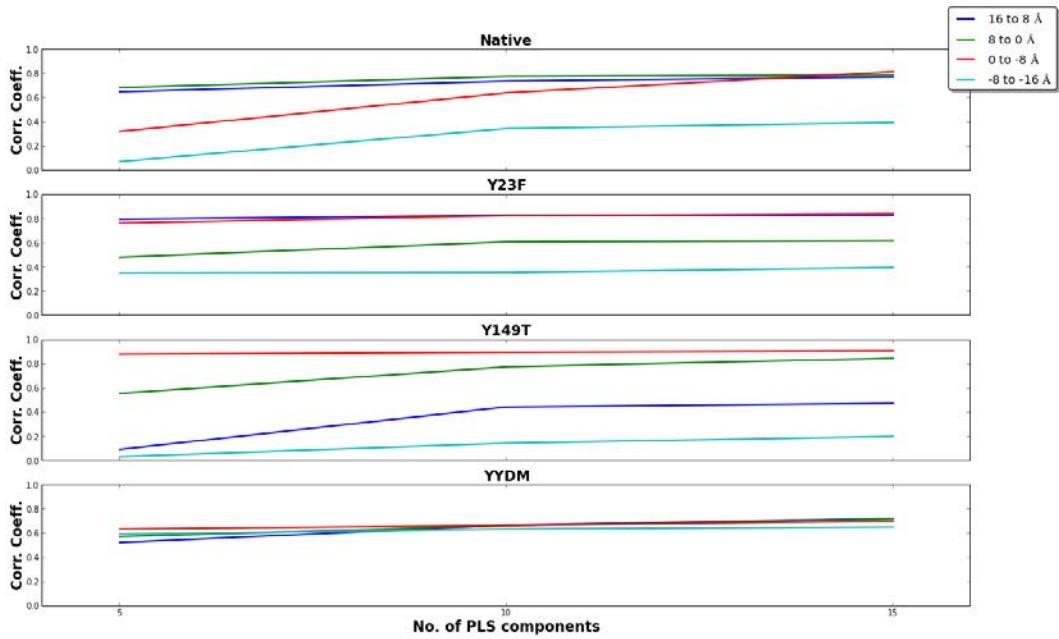


FIGURE 4.8: The coefficients of correlation for the validation sets in different regions of the monomer as a function of PLS components used. We found that this quantity reaches a maximum at 15 components.

The training data and the model is shown for three of the regions for each mutant in figure 4.9. The data was collected from all the four monomers and then concatenated to form a single vector. The assumption here was that the four monomers are equivalent in terms of structure and dynamics and thus could quadruple our statistics. From the figure, it can be seen that we do not obtain a high cross validation correlation for all the sets. Interestingly, the region where the mutation itself is present always leads to a correlation coefficient smaller than the other two windows. This is notable as it indicates that the mutation results in a reduction in the collective motion present in its periphery. Another observation of note is the fact that correlation in the validation set for the region -8 to -0 Å is always high, indicating that there exists a mode that is particularly efficient at regulating the channel radius. This region contains the strongly conserved ar/R motif. The arginine residue R187 of this motif is responsible for constraining the channel to its narrowest point in the profile. Typically, this residue is not highly mobile in other aquaporins.

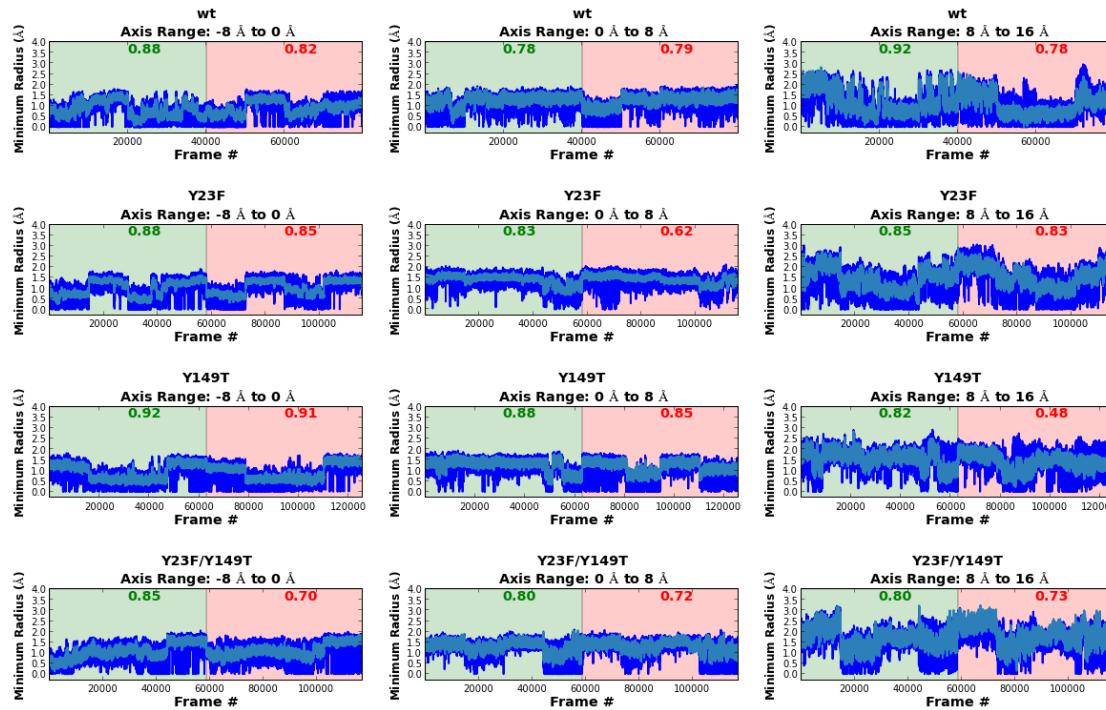


FIGURE 4.9: The training set for each region and the model for all the protein structures (wild type - wt). The sets are divided into two equal regions for the purpose of validation. The training region is highlighted in lighter green color and the validation set is shown in light red. The training data is shown in light blue, while the model is shown in blue. The correlation coefficient is shown for each of the individual sets.



The molecular details of the motion associated with these modes are shown in figure 4.10. All of these were obtained from the wild-type simulations so the role of all the wild-type residues could be shown in detail. The modes highlight the ability of PLS-FMA to capture the effect of relevant residues without any initial bias. In the 0 to 8 Å region, the method identifies the residue Y23 as having the most significant contribution to the collective motion influencing the channel radius in its neighborhood. This mode however does not visually seem to influence the pore radius directly. We henceforth call this mode the *Y23 mode*. The H66 residue also shows significant mobility in this mode, but the effect of that is unclear. Similarly, in the 8 to 16 Å region, we find the highest contributor as Y149. This residue clearly moves in and out of the lumen closing and opening the channel, demonstrating the ability to directly influence permeation. Here, the H66 residue again participates in the collective motion, possibly stabilizing the 'closed' state of the Y149 residue via a hydrogen bond. Henceforth, this mode is referred to as the *Y149 mode*. Most interestingly, the mode obtained from the -8 to 0 Å region captures the motion of the R187 residue. This residue occupies a highly restricted state in terms of channel permeation as the R187 moves into the lumen and a more permeable state formed when the R187 moves away from the lumen. The H172 residue is another chief contributor to this mode. This residue is a part of the aromatic region of the ar/R motif. It also could stabilize the closed state of the ar/R residue through a side-chain mediated hydrogen bond. As this mode involves the R187 of the ar/R region, we call it the *Arginine gate mode*.

PLS-FMA modes allow us to understand the correlation between the change in the protein profile and the dynamics within the protein structure. As the function of physiological interest is permeability, we would like to know how the conductance of the channel changes when the radius of the pore is locally modulated. To establish a causal relation between the PLS-FMA modes and the permeability, we then performed several essential dynamics simulations. The goal of these simulations was to trap the structures of the protein in either 'open' or 'closed' forms of the mode and then calculate the osmotic permeability to validate if indeed we could observe a higher conductance for the former and vice versa.

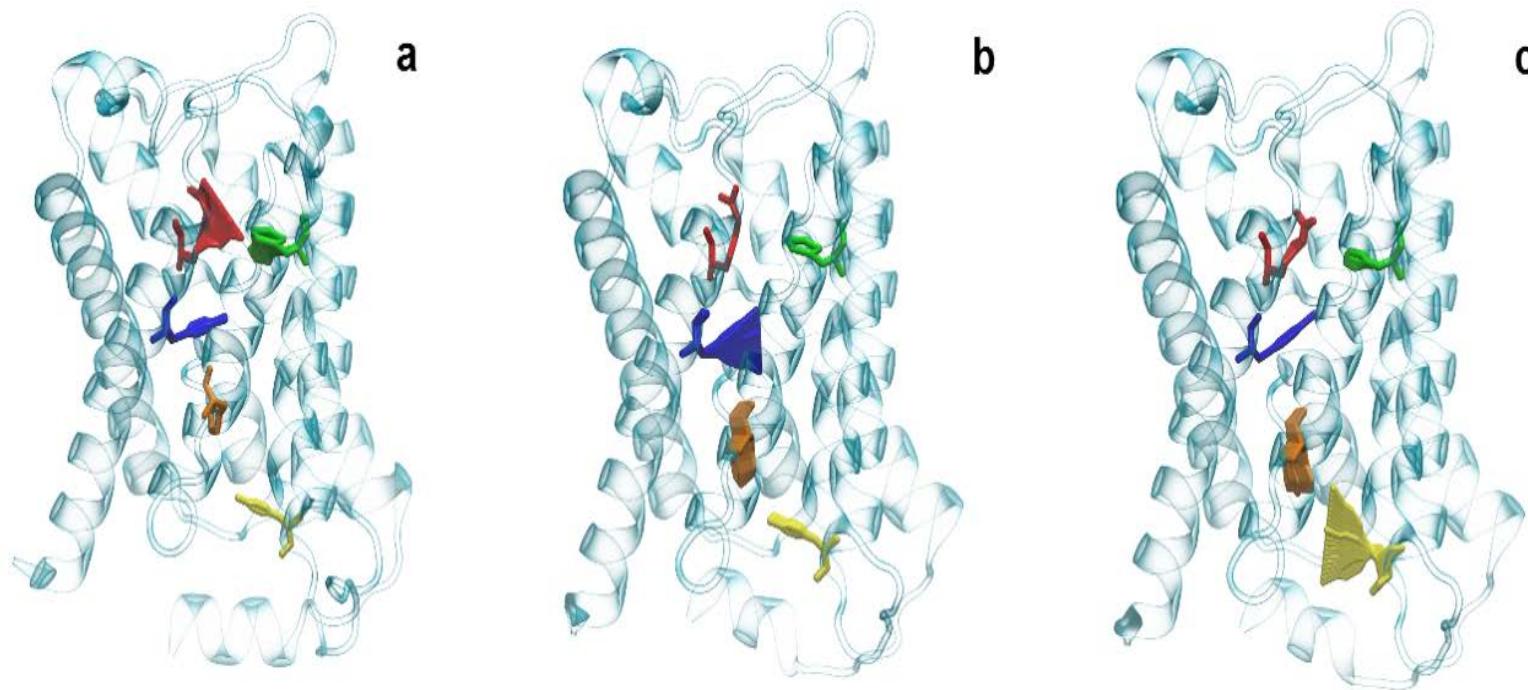


FIGURE 4.10: The collective modes observed in the PLS-FMA model. **a.** The mode in the -8 to 0 Å region (arginine gate mode). This mode captures a dynamic motion of the R187 residue into and away from the AQP0 lumen. **b.** The mode in the 0 to 8 Å region (Y23 mode). The chief contributor to this mode is the Y23 residue. **c.** The collective motion in the 8 to 16 Å region (Y149 mode). The Y149 residue flips in and out of the channel pore.



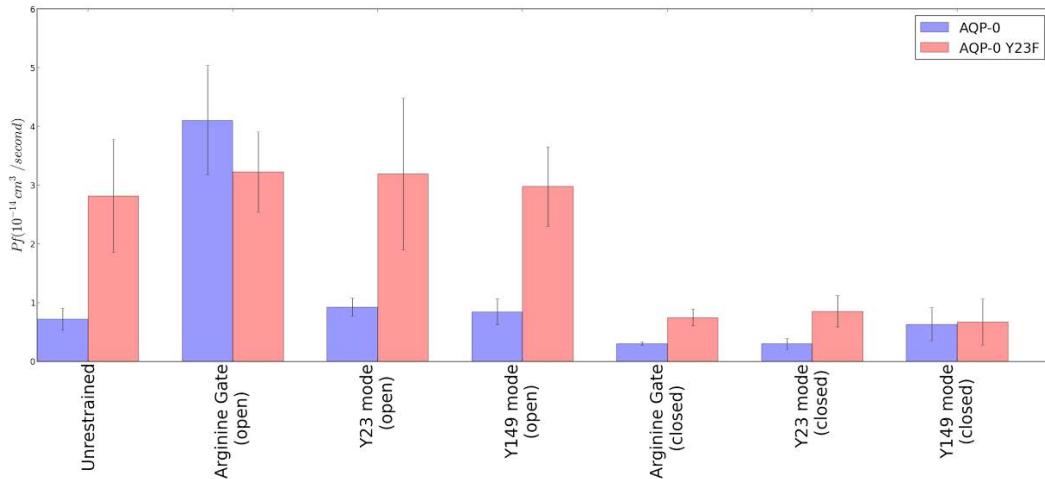


FIGURE 4.11: Osmotic permeabilities as measured by the collective diffusion method for the essential dynamics simulations. The measurement has been done for two forms of AQPO; the wildtype and the Y23F mutant. For each type, the structure has been restricted in an open and a closed state of each of the three PLS-FMS modes shown earlier for the purpose of permeability calculation.

In that direction we chose the wildtype and the Y23F mutant for the comparison. This was so, because in our previous equilibrium measurements we saw the largest change in the permeability for the protein that had this particular mutation. Thus, we expected to see the most significant difference in the Essential dynamics simulations for this pair. We carried out these simulations for all the three modes described earlier. The results for these simulations are plotted in figure 4.11. They are compared with the equilibrium simulations so that the effectiveness of a collective mode can be judged. It is evident that all the 3 modes can 'switch off' the protein's permeation, bringing it down to the level of the wild-type channel. However, only the Arginine gate mode can effectively both 'switch on' and 'switch off' the permeation. Its open state has an osmotic permeability comparable to the Y23F equilibrium simulations indicating that it must have the largest responsibility for the change in the permeation of the protein due to the Y23F mutation. This is surprising as the Y23 residue itself does not directly contribute to the collective motion described by this mode thus begging the question as to how it affects the arginine mode.

A clue might be found in the knowledge regarding how the simulations of

the wildtype and the Y23F mutant are distributed along the arginine gate mode. This distribution will ascertain if the arginine gate mode is at all prominently occupied in the open state of the Y23F mutant. If this would be the case it would strengthen the supposition that there indeed is a relation between the mode and the mutation. In order to check how the protein population is distributed along the arginine gate modes, we made a histogram of the projections of the wild-type protein trajectory and the Y23F mutant trajectory on the first eigenvector associated with the arginine gate mode (fig. 4.12). We found that the Y23F mutant spends half its simulation time in the open state of the arginine mode, which is twice as much compared to the wildtype protein, which spends only a quarter of the simulation time in the open state. The open-ness is calculated based on the projection along the arginine mode. The projection values larger than approx. -0.2, are considered to belong to the closed state as the channel radius is smaller than 1.4 Å from that point onwards in this region, which is the radius of a water molecule. The population distribution implies that the effect of the mode is not *linear* on the permeability, as it can only explain a difference of *two* times through a directly proportional dependence, whereas we observe a difference of 4 times as high conductance in the equilibrium simulations.

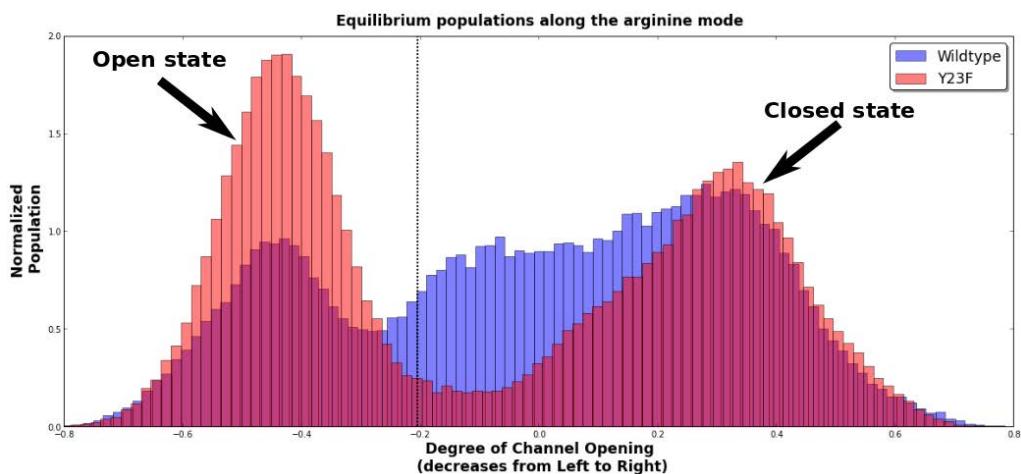


FIGURE 4.12: The population of the wildtype protein and the Y23F mutant structures in the equilibrium simulation is compared against the projection along the arginine mode vector. The population is normalized by its sum. The openness of the channel decreases along the x-axis.

This could be explained by postulating that the actual permeation is not a 'trickle' phenomenon but rather happens in bursts where large number of permeation events take place while the channel is in an open state, not unlike a Poisson process. The essential dynamics simulations and PLS models allow us to confirm that there is a causal link between the pore restricting modes and the permeability. But they leave an important question unanswered: what exactly is the mechanism by which the Y23F mutation closes the channel? We observed in the figure 4.11 that the only mode capable of opening the channel to the same extent as the Y23F mutant is the arginine mode. Does the mutation influence this mode to increase its propensity in the open state? To achieve that the Y23 needs to have a direct or indirect interaction with the R187 residue. This interaction should ideally subside or be eliminated in the presence of the mutation. The chief structural difference in the Phenylalanine residue replaced from the Tyrosine is the lack of the phenolic -OH moiety. This group is polar and can form hydrogen bonds. We checked the simulations of the wildtype protein for possible hydrogen bonding interactions between the R187 and the Y23 residues or even a bridged hydrogen bonding interaction of the form Y23 - X - R187, X being an amino acid in the lumen. But we did not find the direct interaction or an evidence of an indirect linkage within the protein. To assess if the bridge could be formed through the transient water molecules in the permeation pathway, we looked at the densities of water molecules in the channel lumen. The comparisons for the wildtype and the Y23F mutant water densities are shown in the figures 4.13 and 4.14. In these figures we looked at the isodensity surfaces of water molecules in the ED simulations to check what effect the mutations had on the residence position of water molecules near the arginine gate.

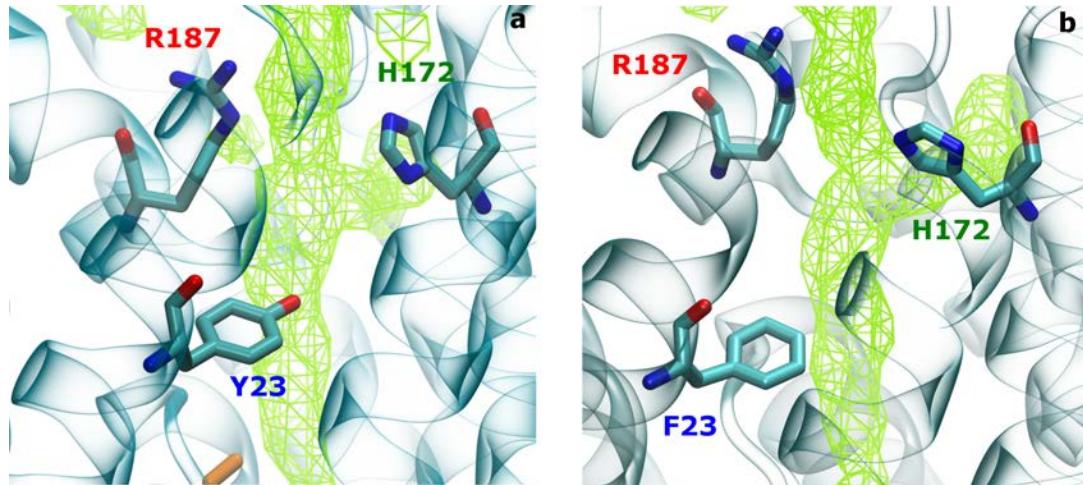


FIGURE 4.13: Comparison of water isodensities near the Y23 and R187 residues with the arginine gate in the open position. The isodensity surface is shown in green in mesh representation. **a.** The ED simulation with the wildtype protein. **b.** The ED simulation with the Y23F mutant. Both simulations show a continuous water density in the Y23/F23 region and near the R187 residue.

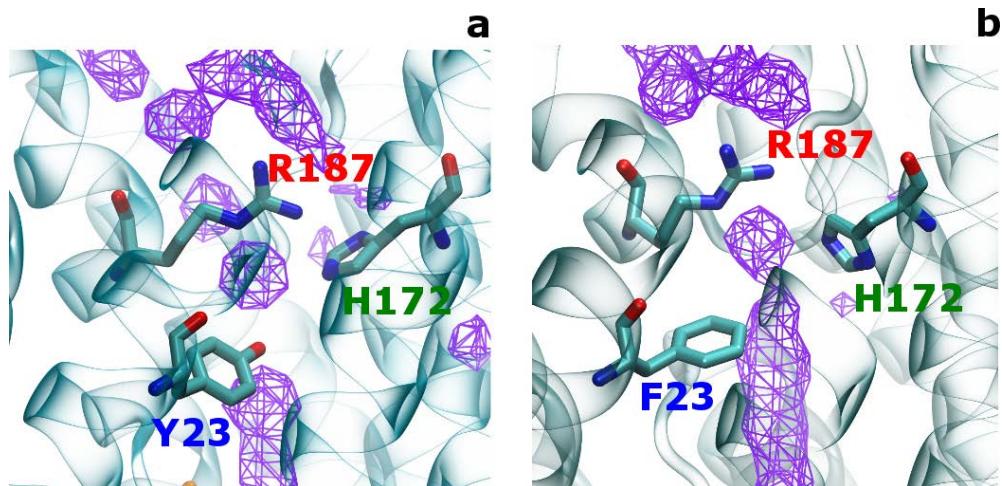


FIGURE 4.14: Comparison of water isodensities near the Y23 and R187 residues with the arginine gate in the closed position.. The isodensity surface is shown in violet in mesh representation. **a.** The ED simulation with the wildtype protein. **b.** The ED simulation with the Y23F mutant. The water density near the Y23 residue is discontinuous and traps a single water molecule between the phenolic group of Y23 and the R187 residue. The F23 residue lacks this ability.

In figure 4.12 we see that the open configuration of the arginine gate presents a continuous water permeation pathway, as the density in that region appears unbroken. In figure 4.14 the closed arginine gate however restricts the pathway in an expected manner by prohibiting water molecules to move unobtrusively across

the lumen. The most interesting feature of this figure is the regime in the middle of the Y/F23 and the R187 residue. Here, the Y23 residue of the wildtype protein clearly traps a single water molecule between the R187 and the Y23 residues via a hydrogen bond bridge through the phenolic group of its sidechain and the N $\delta$  atom of the R187 residue. This isolation of the water molecule is markedly absent in the Y23F mutant as the corresponding hydrogen bond partner is lost. Such an interaction, that couples the R187 residue to the Y23 residue through a hydrogen bond bridge mediated by a water molecule would stabilize the closed state of the arginine gate mode, by holding the R187 residue in the compatible state for restricting the permeation. This seems to suggest that the Y23F mutant regulates the permeation of water through the protein by modulating the arginine gate mode itself.

Finally, we wanted to assess if we could 'rescue' the higher permeability in the AQP0 wildtype protein after mutations that mimic the effect of the Y23 residue. To accomplish this, we searched for possible mutations in the neighborhood of the R187 residue which could induce the R187 in a configuration compatible with the open state of the arginine gate mode. The general strategy in that direction was based on the fact that R187 is a positively charged residue. This could be used to our advantage as we could introduce negatively charged residues in locations surrounding the R187 site so that it could form salt bridges with one of them and be locked in a open state.

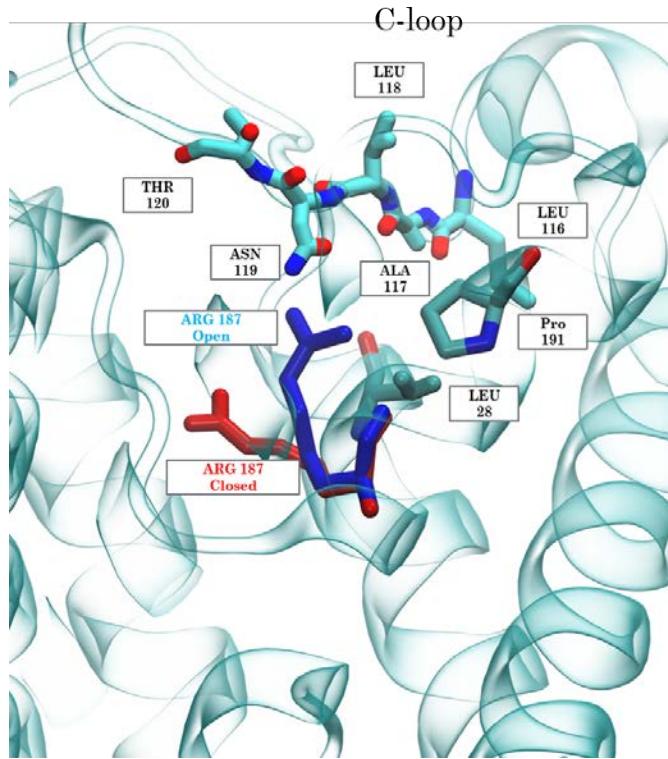


FIGURE 4.15: Seven possible residues that could induce the R187 residue in a closed state (shown in the solid red color), into an open state (solid blue color), of the arginine gate mode.

Figure 4.15 shows the 7 residues chosen to achieve this effect. Of these seven mutations, we chose 5 in a semi-flexible conserved C-loop placed on top of the Y23 residue. The other two (L28 and P191) are present outside of the C-loop but still in strategically suited locations. These seven residues were replaced with either an aspartate (D) or a glutamate (E) residue. Both of these amino acids are negatively charged and differ from each other through a single methylene group. This gives the us the flexibility to provide the mutants with two possible maximum distances from the R187 residue. We conducted equilibrium simulations in a 250 to 300 ns window for all the 14 possible mutants. We tested for their ability to remain in the open state of the arginine gate mode in two ways; by direct measurement of the permeability or by the direct measurement of the projection of the trajectory on the arginine gate mode.

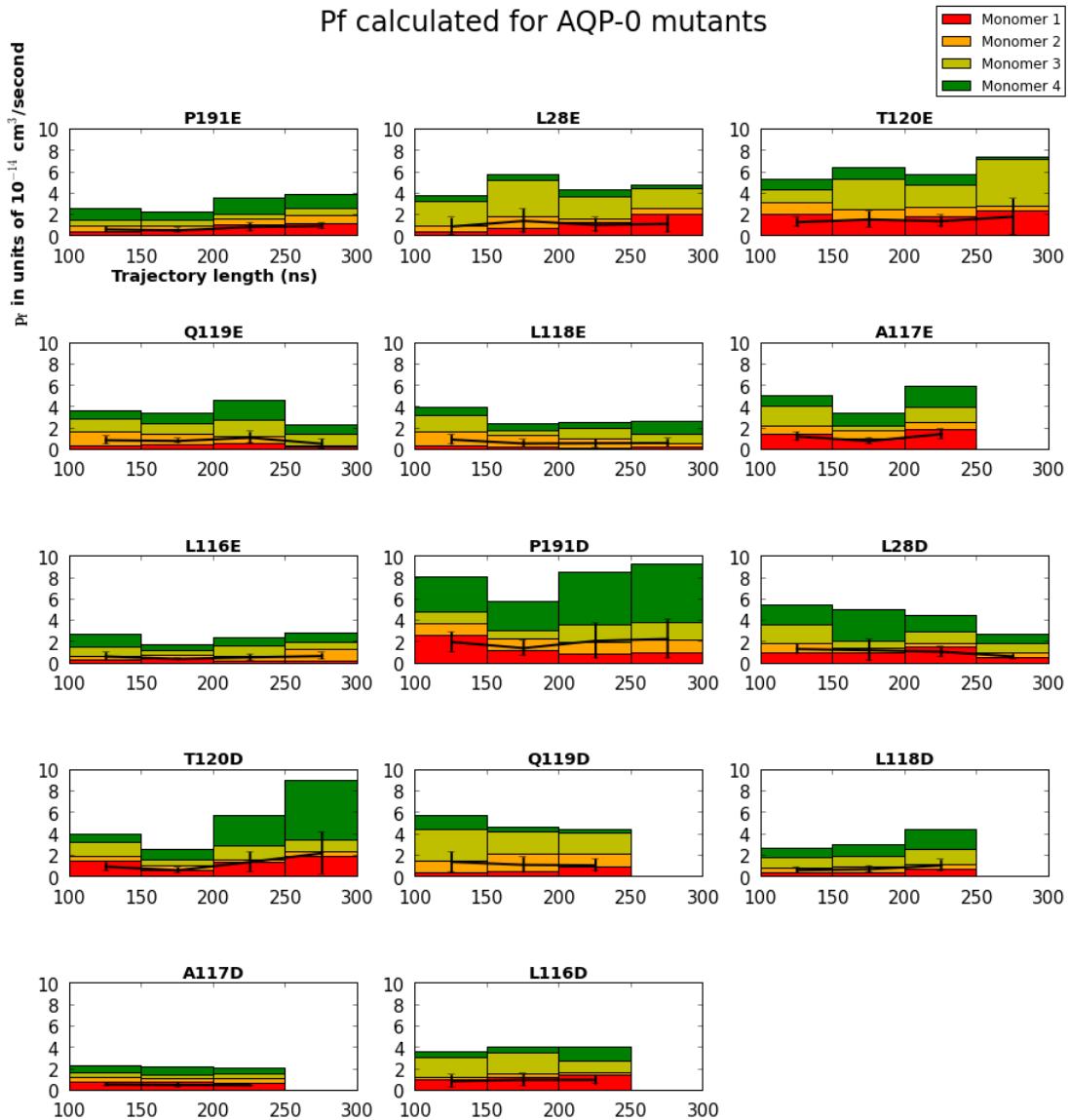


FIGURE 4.16: The osmotic permeability measurement for the 14 mutants of expected to open the arginine gate. The permeabilities are plotted for all four of the monomers in the bar graph for each 50ns time window. The average permeability for the protein is shown by a black line at the appropriate ordinate. The error bar is the standard deviation calculated over the time windows.

The osmotic permeability measurement shown in figure 4.16 is too noisy to allow for a meaningful assessment of the efficacy of the mutants to open the arginine gate mode. This led us to consider the projection of the structures in the trajectory on the arginine gate mode as a more suitable indicator of how well the mutants could pull the R187 residue into an open state. Figure 4.17 shows that the trajectories of most mutants, with the exceptions of P191D, Q119E and P191E,

are trapped in the open state of the arginine mode in the simulation window. This could imply that most of the the mutations selected would be able to increase the permeation of the AQP0 similar to the Y23F mutant.

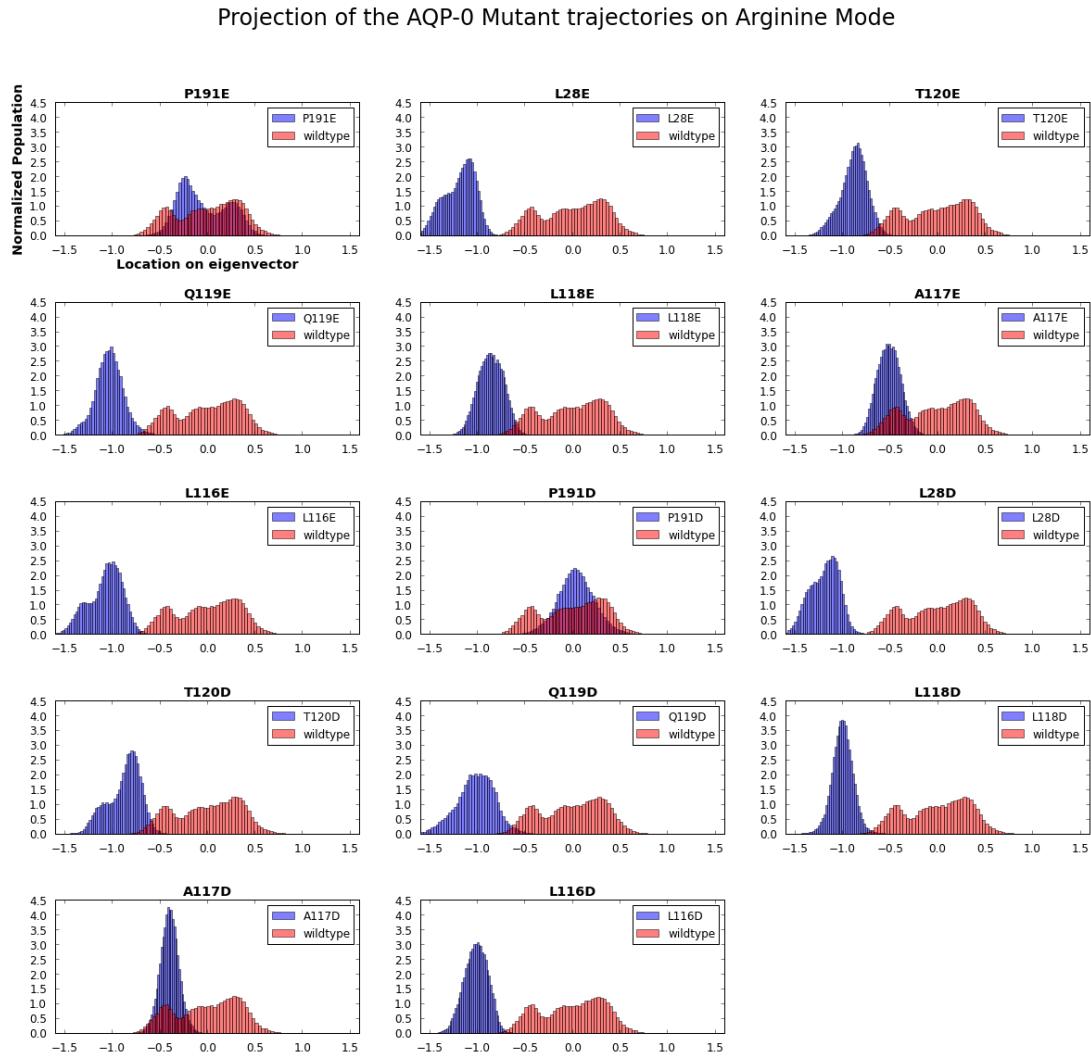


FIGURE 4.17: Histograms of the projections of the mutant trajectories (blue) on the arginine gate mode are compared to the population of the wildtype trajectory projections (red). The x-axis projections measure the ‘open-ness’ of the channel, decreasing from left to right. All the mutations have the projections on the open side (left) of the arginine mode.

## 4.6 Conclusions

In this chapter we have tried to reproduce the experimental observation that the mutations in the unusual tyrosine residues in the lumen of the AQP0 channel increase the permeability of the protein to the level of AQP1. We find that the increase in the permeability observed is not quantitatively in the same range; instead of a 20 fold increase in the experimental results, we see a 4 fold increase. This however is in line with previous molecular dynamics studies [97] Neither do we see an additive effect of the Y149T and the Y23 mutations as present in the experimental setup. However, we do have a qualitative agreement corresponding to the Y23F mutation which indeed brings the permeability of the mutant to the order of magnitude observed in AQP1. We tried to focus on this mutation and used machine learning algorithms to find out what structural or dynamical effects present in the simulations could explain such a drastic change in the conductance of the protein. We identified three collective modes in the simulations that allowed us to switch off the protein conductance even in the high permeability mutant. One of these modes, associated with the conserved ar/R region could induce a high conductance in the wild type protein, indicating that the Y23F mutant must somehow affect this mode to achieve its high permeation.

Upon considering the wild type structure, we found that the possible means of the Y23 residue to promote the closed state of the arginine mode could be via its phenolic sidechain. We found that the -OH group of the Y23 residue was capable of forming a stable and sustained hydrogen bond with a water molecule that could be trapped between R187 of the arginine gate and Y23. This arrangement could potentially stabilize the closed state of the arginine gate in the wild type protein, which is also substantiated in our simulations where we see that the open state is twice as often visited in the mutant compared to the wild type. We expect that this mechanism provides a non-linear increase in the permeability of the Y23F mutant as it can no longer stabilize the closed state. This is feasible, as the ar/R region is the narrowest region in the permeation pathway and thus presents the largest Arrhenius barrier to conductance. The pore in this region is typically smaller than the radius of a water in the wild type simulations. Small changes in the pore radius

thus could allow bursts of permeations that lead to the large osmotic permeability of the mutants. Thus overall we find that the effective mechanism to modulate the conductance of AQP0 can only be described through the dynamic nature of the permeation itself and cannot be explained via changes in the static profile of the channel alone.







# Chapter 5

## Ammonia permeation in plant aquaporin TIP2;1

### 5.1 Overview

How do gas molecules move across biological membranes? In the first chapter of this thesis we discussed the 'Meyer-Overton rule' [31]. This principle describes that permeation of gaseous substrates across lipid bilayers occurs passively and preferably through the membrane itself. In physiological situations, gases such as oxygen, carbon dioxide and ammonia need to be regularly conducted across cell membranes. The exact mechanism of their permeation is still under scrutiny. Several members of the aquaporin protein family have been suggested to mediate the permeation of these gases in a biological context. Carbon dioxide permeation through aquaporins has received special attention with several studies directed towards its role in physiological situations [102–105]. Aquaporins have been proposed to be permeable to other gaseous molecules such as ammonia [106–108] and nitric oxide (NO) [109]. There have been computational studies that have assessed the capacity of aquaporins to conduct gases. Wang *et al.* have studied the permeation of NO through AQP1 [75]. Hub *et al.* have calculated the potential of mean force for the permeation of several gaseous molecules to traverse the AQP1 main pore [29]. These efforts point towards at least potential capacity of this protein family to allow conductance of gas molecules through its channel.

Gas permeation is much more important in plants, where it is involved activities such as photosynthesis and ammonia fixation beyond simple respiration. Subsequently, plants posses specialized organelles to deal with storage and conduction of gas molecules. The most important of these are fluid or gas storage vesicles called vacuoles that regulate the size and the pressure within the plant cell. They are surrounded by a specialized membrane called tonoplast. Tonoplast Intrinsic Proteins (TIPs) are members of the Major Intrinsic Protein (MIP) family [110]. [111–113]. TIPs are ubiquitously present in land plants and there are several isoforms of these required for their survival, labelled TIP1-5. Primitive plants such as mosses typically have only one type of TIP (TIP6) [114]. These proteins can constitute a very large fraction of the protein content of the tonoplast membrane, going as high as 40 % [115]. Their role in ammonia permeation has been suggested to be enhancing nitrogen uptake and detoxifying the ammonium waste created in the vacuoles [112]. TIP has also been suggested to mediate remobilization of the vacuolar ammonia during nitrogen starvation and its reallocation during senescence (aging) [111, 116]. Recently, TIPs were included in a revised model of the 'Futile cycle' under high ammonia conditions [117].

## 5.2 Motivation

Our collaborators Dr.Urban Johansson and Andreas Kirscht have crystallized the TIP2;1 channel for the first time and determined its structure at a resolution of 1.18 angstrom using x-ray diffraction (fig. 5.1).

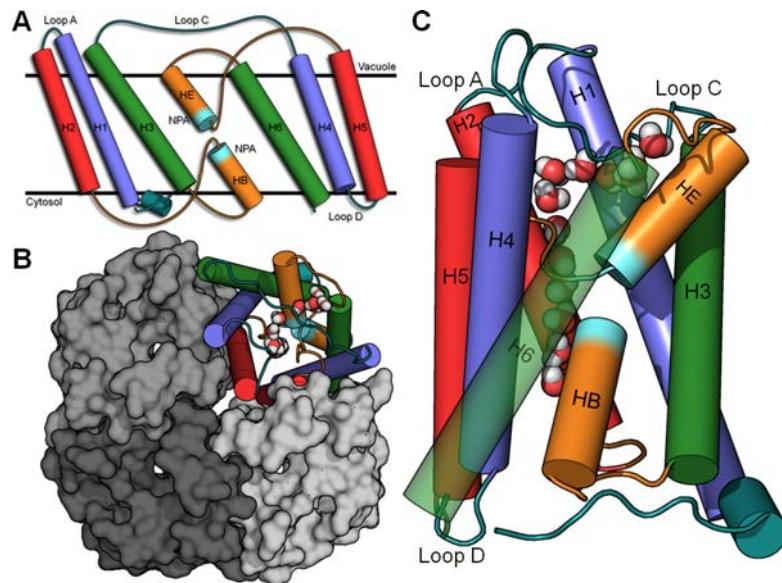


FIGURE 5.1: Structure of AtTIP2;1. (A), Membrane spanning helices (helix 1-helix 6) and two half helices (helix B and helix E), connected via conserved NPA-motifs, form a pore through the vacuolar membrane. Homologous helices in the internal repeat are indicated by colours. (B) AtTIP2;1 tetramer viewed from the vacuolar side. (C) Side view of the monomer with the same orientation as in (A). Eight water molecules form a single file in the main pore and five waters are seen in a side pore underneath loop C. Figure courtesy, Andreas Kirsch and Dr. Urban Johansson (unpublished data)

This structure presents two interesting features which we suggest are relevant to the possible function of gas permeation through the main channel. The first one is the extended selectivity filter (SF). This region formed out of several aromatic residue and a highly conserved arginine residue (ar/R region)(here, R200) is typically the narrowest region of the channel lumen. The ar/R region provides the aquaporin structure its selectivity towards water and its ability to distinguish it from protons. In the TIP2;1 structure this region is comparatively broad and thus is considered 'extended'. The aromatic residues of the ar/R are replaced with two histidine residues (here, H63 and H131) which provide a polar environment around the charged R200 (fig. 5.2).

The second interesting element of the structure is the presence of a water filled side pore, which extends from the loop C near the extracellular side of the protein directly into the main pore right into the selectivity filter (fig. 5.1). This pore provides an unusual second means of entry into the permeation pathway.

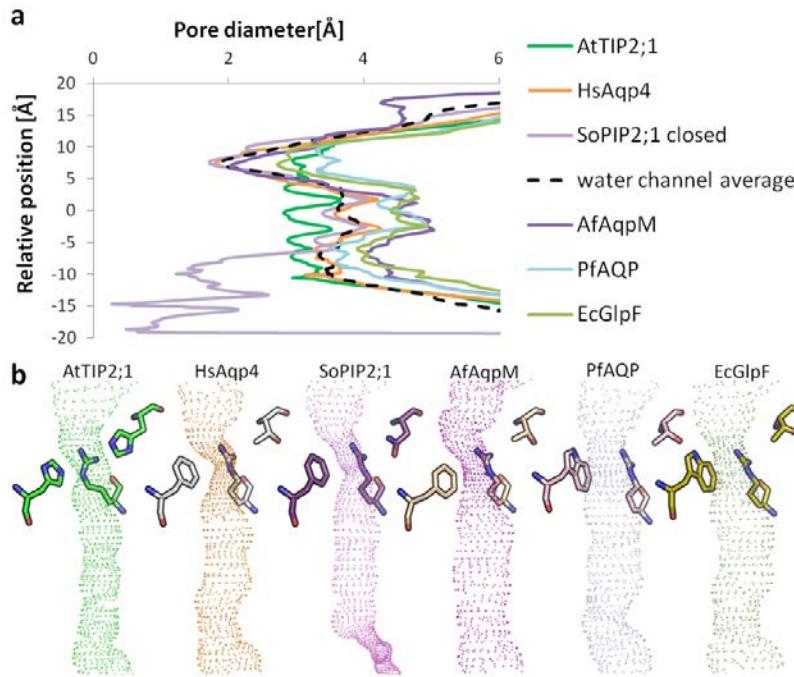


FIGURE 5.2: Comparison of the maximal pore diameter for different MIPs. a, AtTIP2;1 (green) provides a more or less constant/uniform pore diameter at 3 angstrom, illustrating the ability to conduct compounds larger than water. AtTIP2;1 presents the narrowest NPA region (relative position 0 on y-axis), but a much wider SF region; only glycerol containing structures such as PfAQP and EcGlpF have a larger diameter at the SF. b, Graphical representation of MIP pores, indicating the SF region by stick representation of residues in H2 and HE as well as the residue in loop C corresponding to H131 of TIP2;1. Waters and non-proteogenic molecules were removed from available structures and their pores analyzed using the program Hole. Figures courtesy of Andreas Kirsch and Dr. Urban Johansson, unpublished data

As the goal of this project, we wanted to understand what roles were played by the extended selectivity filter and side-pore in a possible mechanism through which ammonia could be conducted through the TIP2;1 channel. Does the dynamics involved in the SF somehow facilitate the entry of ammonia? Does the side pore provide some insight into a mechanism by which ammonia permeation could be enhanced? What free energy barriers does the ammonia molecule experience when it traverses the channel lumen? These are some of the questions we address in this chapter.

### 5.3 Is the TIP2;1 structure a functional water channel?

A TIP protein structure must be a functional water channel in order to qualify as active and perform its designated function in the vacuole. To test if the structure determined by our collaborators is indeed permeable to water, we calculated the osmotic permeability of the protein using the collective diffusion method [43]. The results of these calculations are shown in figure 5.3.

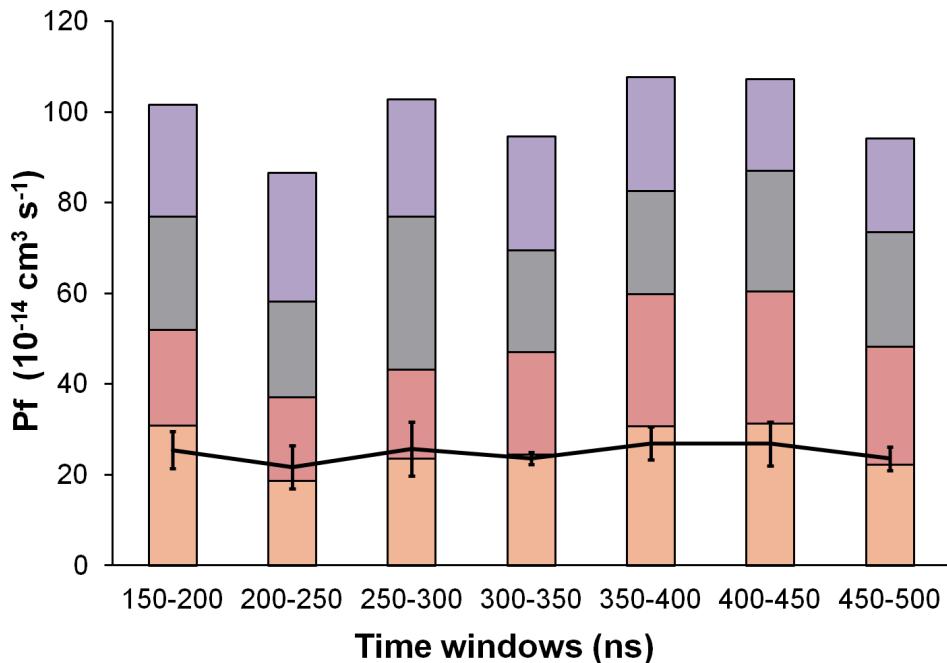


FIGURE 5.3: Permeability values calculated from MD simulations. These values were calculated separately for each monomer in seven 50-nanosecond time windows. The contribution of the individual monomers to the permeability values of the tetramer are indicated by different colours and average values per monomer and SD in each time window are indicated by the black line and error bars.

We find that the TIP2;1 protein has a comparatively high permeability ( 3 times larger than AQP1) which persists over a 500ns simulation window. This is consistent with our understanding that the permeation barrier of the channel has been reduced significantly in the channel due to the enlargement of the SF. This allows a mostly single-file but 'rapid' movement of water across the lumen. This

result substantiates the idea that the protein has been crystallized in a functional form.

## 5.4 Novel features of the selection filter of TIP2;1

The pore diameter of TIP2;1 is rather small at the NPA-region and remains constant at around 3 angstrom throughout the pore (fig.. 5.2). This is unusual since in other structures of open AQPs the SF region constitutes the narrowest part of the pore. Four amino acid residues in helix 2, helix 5, loop E and helix E form the SF and the residues at these positions are believed to be the major determinants for substrate specificity.

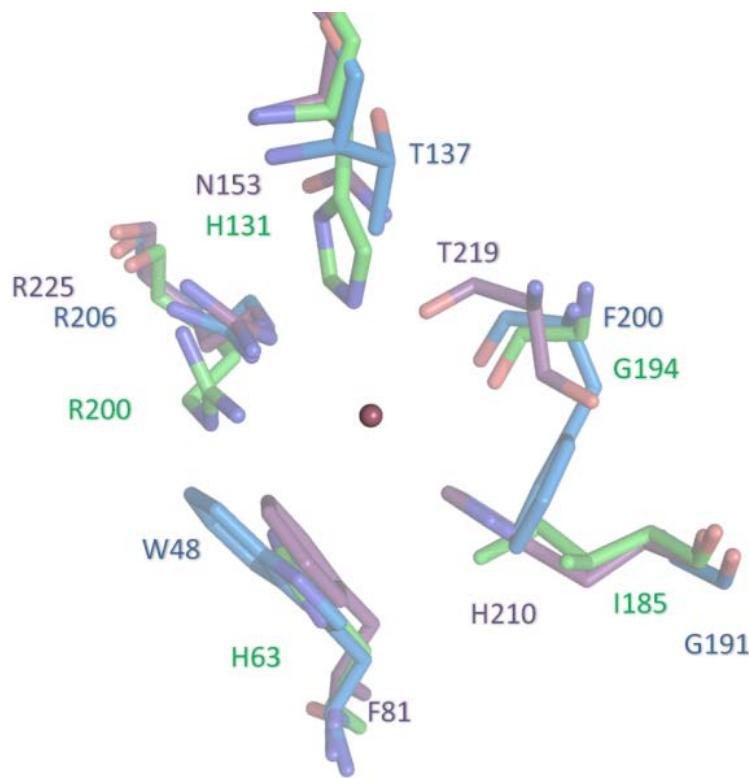


FIGURE 5.4: The histidine (H63) does not directly affect the diameter of the pore compared to known structures of other MIPs. In contrast the isoleucine (I185), replacing the histidine enlarges the pore compared to water specific MIPs. H131 creates an H-bond to water (red sphere) in the pore and sterically forces the arginine away from the center of the lumen. Figure made in collaboration with Andreas Kirsch.

The unusual width of the SF in TIP2;1 appears to be due to the wider pore which is formed by replacing a conserved histidine in water specific AQPs to an isoleucine (I185) (fig. 5.4). In TIP2;1 the residue (R200) is additionally pushed to the side of the pore by a histidine located in loop C (H131). It seems to be further stabilized by a hydrogen bond to the histidine (H63), which occupies essentially the same space as the corresponding aromatic residues of water and glycerol channels without direct effects on the aperture of the pore. The close interaction with R200 can shift the pKa of H63, which is likely to stay singly protonated even in a comparatively acidic environment of the vacuole. The other H131 points to the center of the pore and forms a hydrogen bond to a pore-water in the crystal structure.

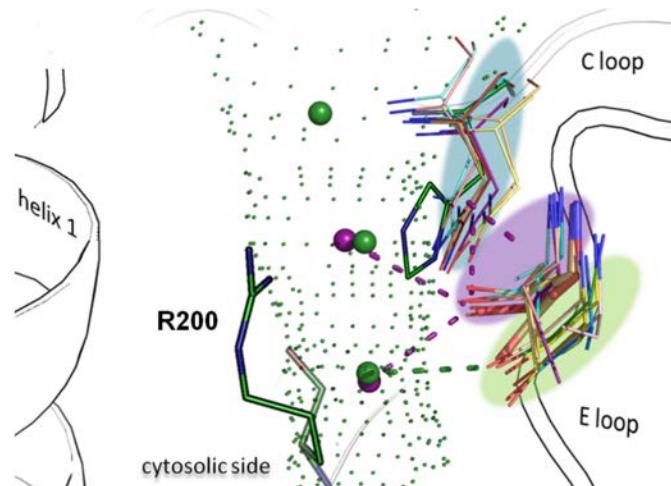


FIGURE 5.5: The positions of corresponding SF carbonyls in loop E cluster in two distinct groups. (A) Carbonyls opposing arginine (R200 from TIP2;1) of water-specific aquaporins form group II (violet shading), and most of them are within hydrogen-bonding distance (indicated by dashed lines) to two water molecules (violet spheres) in their structures as illustrated by SoPIP2;1 (PDB ID 1Z98; violet). Carbonyls of non-water-specific channels gather in a different location (group I; green shading). Among those are TIP2;1 (green) and glycerol transport facilitating and uncharacterized proteins. Like all other members of this group, AQP4 is lacking the asparagine at C-loop that is conserved among the other water-specific proteins (blue shading, only asparagines residues are shown, Group III). A certain flexibility is suggested by the special case of AQP0, where different structures are available (1YMG and 2B6O shown) and the carbonyl highlighted by thick stick representation is seen with both orientations. Figure made in collaboration with Andreas Kirsch.

In the simulations performed to study the dynamics of this region we found that the overall interactions as observed in the crystal structure are preserved. The H131 residue is always seen in strong hydrogen bonding association with the water molecules that are present in the lumen of the main pore. The carbonyl moieties

of the lumen residues present the polar carboxyl oxygen facing the permeation pathway and provide an expected hydrogen bonding environment for compensating the loss of intra-water bulk hydrogen bonding.

## 5.5 Unusual side-pore in the TIP2;1

The unusual side pore seen in the crystal structure is in also reproducibly present in the MD simulations (fig. 5.6 A). The simulations concur that loop C leaves enough space for a continuous side pore reaching all the way to the vacuolar surface. It is always water filled and is open to bi-directional permeation of water molecules to and fro from the main pore to the extracellular surface. Interestingly, this side pore may offer an explanation of how ammonium ions might deprotonate.

Ammonia molecules in the soluble form predominantly exist as positively charged ammonium ions in an aqueous environment. Charged residues cannot traverse the cell-membrane according to the Meyer-Overton's rule. It is also not possible for these ions to move across the aquaporin channel due to the strong repulsion present in the lumen from the conserved arginine of the ar/R region. The opposing dipoles in the channel interior from the hemi-pores also create a large electrostatic barrier to permeation of charged molecules through aquaporins. This would imply that a mechanism to deprotonate the ammonium ions needs to exist to facilitate permeation of ammonia through the channel. Can TIP2;1 protein itself contribute to that process?

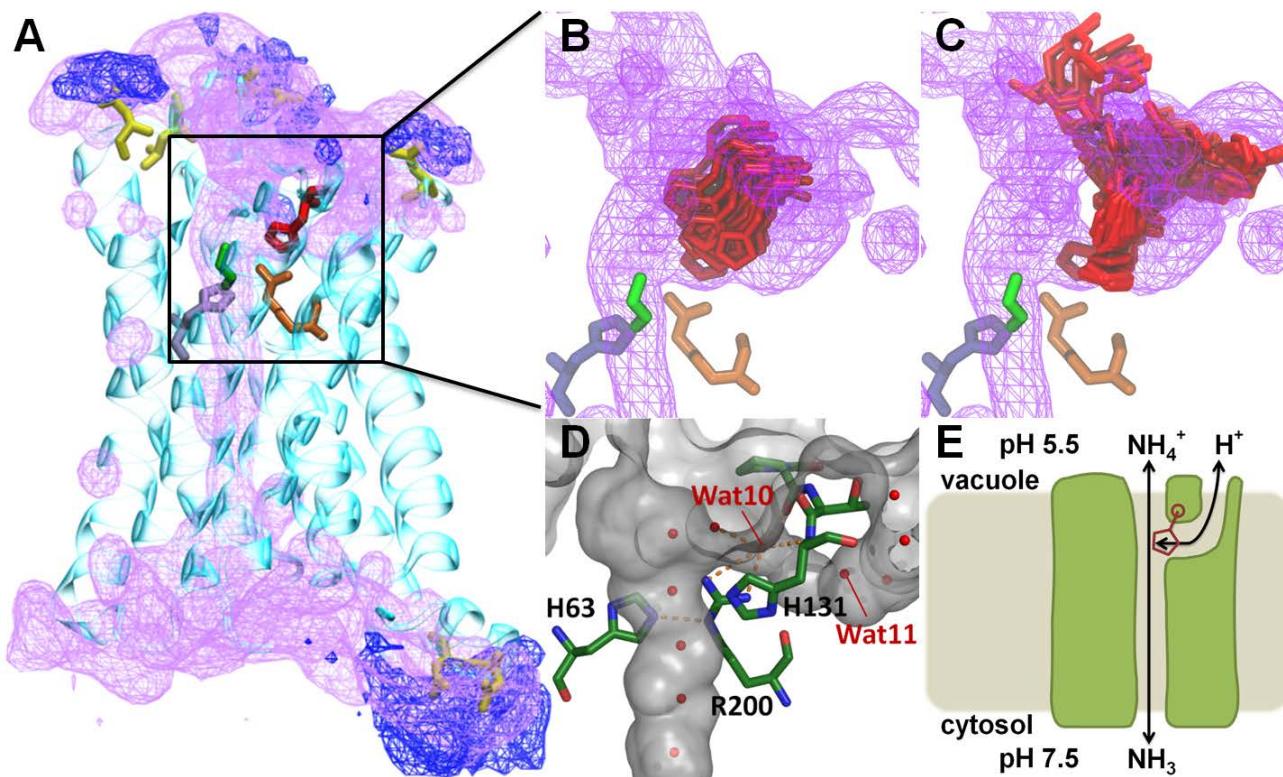


FIGURE 5.6: Ammonium accumulation and proton pathway. (A) MD simulations showing ammonium accumulation (blue mesh) at aspartate residues (yellow sticks). Water density is displayed as purple mesh and residues of extended selectivity filter as sticks (H63 (blue), H131 (red), G194 (green), R200 (brown)). (B) and (C) MD simulations demonstrating flexibility of H131 being neutral and positively charged, respectively. Colour code as in (A). (D) Surface representation of the crystal structure depicting the water-filled side pore beneath loop C. Hydrogen bonds of water 10 as well as between R200 and H63 are indicated by dashed orange lines. (E) Working model of ammonia permeating TIP<sub>2;1</sub>. Ammonium may contribute to permeation by accumulating on vacuolar protein surface and having its protons shuttled back by H131 (red) via a water-filled side pore. D and E were prepared by Andreas Kirsch.



In our simulations where we added ammonium ions to the solvation box, we see regions of the protein surface enriched by this molecule. These regions tend to be in the neighborhood of the negatively charged aspartate residues that are present on the solvent exposed surface on either side of the protein (fig. 5.6 A). Interestingly, one of the regions with such an enrichment lies in the close vicinity of the side pore. Considering the low pH in vacuoles and accumulation of ammonium on the vacuolar surface of TIP2;1, the permeation efficiency would clearly benefit if ammonium contributes to channeling of ammonia. The simulation results could mean that the surface aspartates could be antennas that accumulate the ions for local concentration before deprotonation.

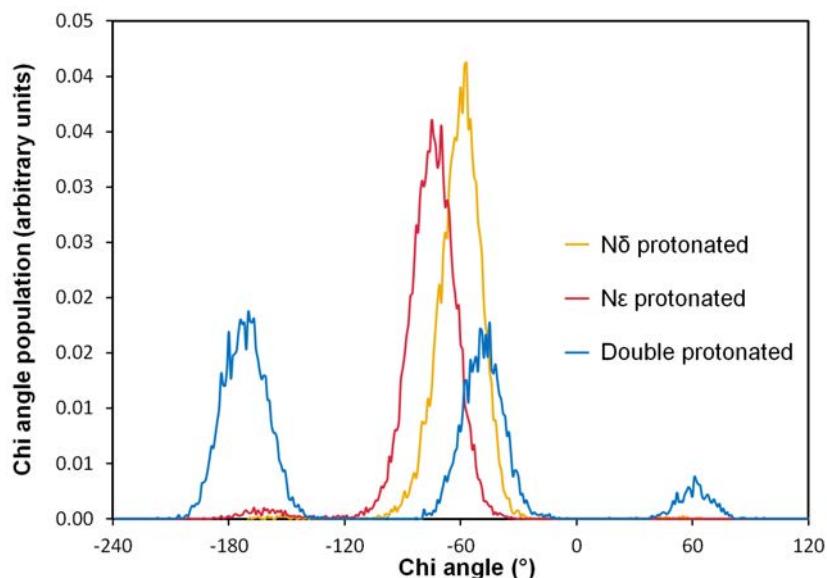


FIGURE 5.7: Dihedral populations of H131 at three different protonated states. Chi1 angles in MD simulations with doubly protonated (positively charged; blue), N $\delta$  protonated (neutral; yellow) and N $\epsilon$  protonated (neutral; red) H131.

The dynamics of the second vacuolar loop (loop C) relative to the SF arginine (R200) hints at a mechanism that could explain how this deprotonation might occur. This loop contains the H131 which can interact directly with the R200 of the SF. As histidines at physiological conditions are titrable, we wanted to explore if H131 and the side pore could play a role in facilitating deprotonation of ammonium. To check that, we plotted the dihedral space occupied by the H131 (fig. 5.7) in different protonation states. The simulations show that the angle of the H131 side-chain (chi 1) remains as in the crystal structure when neutral while

in a doubly protonated, positively charged state we observe indeed an alternative orientation towards main pore. The doubly protonated version is highly flexible and can move in and out of the main pore existing in three differing states in the simulations. For this residue, the peak corresponding to the angle of  $-60^\circ$  is situated within the main pore. It is equally probably occupied at the  $-180^\circ$  position which corresponds to its orientation towards the side pore. This is interesting as such a rotation brings the H131 residue close to the region where we found high density for the ammonium ions accumulating near an aspartate residue. The final state of H131 is the least occupied one and represents the position near the main pore opening (fig. 5.6). This may suggest H131 could potentially serve to abstract a proton from the ammonium and enrich ammonia near the lumen entrance. It might be that the abstraction itself would stabilize the otherwise less likely position of the singly protonated H131.

Furthermore, another possible means of abstracting the proton might be suggested by the arrangement of the side pore. The simulations confirm the side pore as being continuously water-filled. Thus from H131 to the vacuolar exit, it potentially supplies a hydrogen bonded network for transfer of protons. These findings could imply a mechanism where H131 shuttles protons from the main pore to the vacuolar surface via the side pore, using a Grotthuss mechanism, putatively enhancing the permeation rate of ammonia under non-equilibrium flux conditions.

## 5.6 Does TIP2;1 permeate ammonia?

The conditions under which the Meyer-Overton rule might weaken, so as to enhance permeation via membrane embedded channels are physiologically present. These represent the situation where the permeation of gas molecules might occur at highly specialized regions, where the local environment might not offer a large free membrane surface for free diffusion. Another important consideration concerning this process is the lipid constitution of the membrane. Cell membranes of higher organisms contain large proportions (20-40 %) cholesterol. This thickens the membrane and changes its ability to allow permeation of hydrophobic molecules. Thus, it might be necessary to take these factors into consideration when permeation of

ammonia is examined for aquaporins. To get a more detailed view of the substrate specificity in TIP2;1 MD simulations were conducted. We have already shown that water permeation was seen at high frequency corresponding to a Pf with SD of approximately  $25 +/- 4 \times 10^{-14} \text{ cm}^3 \text{ s}^{-1}$ . Spontaneous ammonia permeation events (fig. 5.8 a) were observed in unbiased simulations with a length of 400 ns.

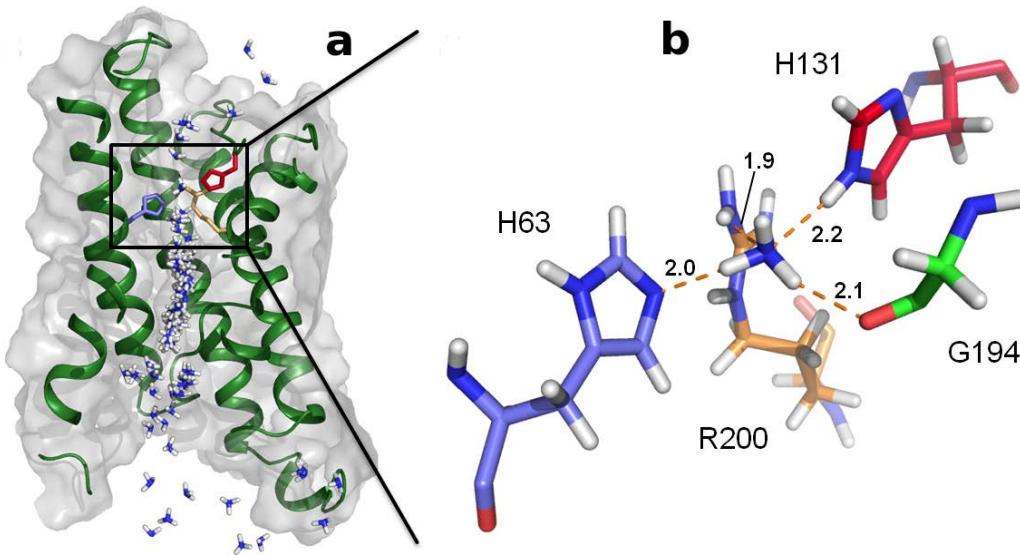


FIGURE 5.8: (a) Snapshots of ammonia permeation. (b) Close-up of ammonia interactions at the SF. Hydrogen bonds are depicted in orange and distances are given in .

We studied the free energy profile of the permeation of ammonia using umbrella sampling simulations. 280 umbrella windows, 0.25 angstrom apart were used to estimate the Potential of Mean Force (PMF) along the channel axis. For each window, we simulated 2ns length of trajectories of which we dropped the first 500 ps to allow for equilibration. Using the WHAM algorithm as described in the second chapter, we calculated the uncertainty in the profile obtained with 100 bootstrap trials. This yielded a free energy barrier of approximately 15 kJ/mol (fig. 5.9), in line with a high ammonia permeability. This desolvation barrier was found to be comparable to a similar barrier to permeation obtained in a purely POPC lipid bilayer. Interestingly, a model asymmetric bilayer constructed with 20% cholesterol in either leaflet yielded an even higher barrier. Therefore, comparing TIP2;1 to a cholesterol containing model membrane with an energy barrier

of 20 kJ/mol (fig. 5.9), the corresponding ammonia permeability of TIP2;1 is an order of magnitude higher. This suggested that physiologically, the permeation across aquaporins might be enhanced due to the presence of cholesterol.

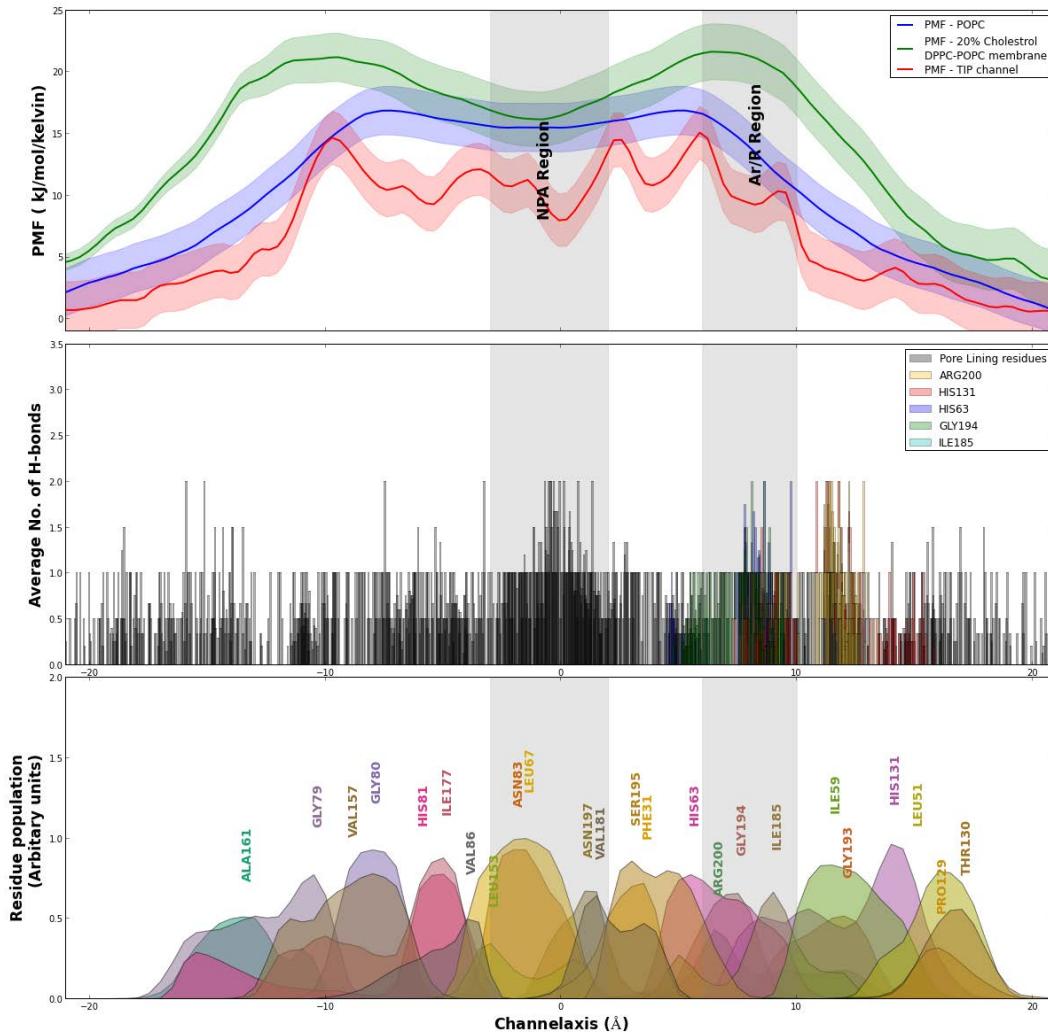


FIGURE 5.9: The potential mean force (PMF) profiles for ammonia through TIP2;1 (red) and through model membranes containing no (blue) or 20% cholesterol (green) (Top panel). Number of hydrogen bonds between ammonia and TIP2;1 as function of position along the pore axis (Middle panel). The residue population at a given point on the channel axis (Lowermost panel)

The umbrella sampling simulations also showed that the SF residues strongly interacted with the ammonia molecule and stabilized it (fig. 5.8 b). This was reflected in the PMF as well, where the ar/R region has a remarkable dip in the free energy where typically a barrier to entry is present for simulations with

the model membranes. MD simulations containing ammonium ions showed no spontaneous permeation events, which is expected due to the electrostatic and desolvation effects in the channel.

## 5.7 Conclusions

The first atomic structure of an aquaammoniaporin provides new insight into the substrate selectivity of AQPs. The structure reveals an extended selectivity filter, including a fifth residue positioned in loop C that also may play a role in defining substrate profiles of the entire superfamily of proteins.

We use MD simulations to analyze the unusual features of the structure to shed a light on what might be their possible functional relevance. The importance of the extended SF is demonstrated within the calculation of the free energy profiles and the role of the SF in stabilizing the entry of the ammonia molecule in the pore. This points towards a gain-of-function of TIP2;1 substrate selectivity compared to the water-specific aquaporins. MD simulations support the notion that ammonia conductance occurs via the main pore and depict a lack of ammonium permeability. Ammonia permeation through the main pore is also comparably a more feasible pathway when set against a lipid bilyare with moderate amount of cholesterol. The observation, that the free energy barrier to ammonia permeation through the model membrane is larger compared to permeation through the channel, combined with the specialized nature of the tonoplast membrane indicates that the role of the aquaporin family is worthy of further study in the context of gas permeation.

Based on structural and MD analyses, we describe a pore that is permeable to ammonia and propose a mechanism in which ammonia permeation may be further increased by ammonium accumulation at the vacuolar protein surface, deprotonation, and proton transfer via a previously unidentified water-filled side pore. The side-pore itself seems to be a functional feature of the structure and its close proximity to the other unusual elements of the TIP2;1, such as the conserved H131 residue and its entry into the main pore at the ar/R region suggest that it is more than a structural curiosity.

Thus, overall we present a study where it is demonstrated that computational efforts can be combined successfully with empirical studies to understand the dynamics of gas permeation in aquaporins.





# Chapter 6

## Conclusions and perspectives

### 6.1 Conclusions

In the previous chapters we have studied three aquaporin proteins using MD simulations. We found that experimentally observed behavior of these proteins can not be explained completely from the structural details alone. Dynamics adds essential information which is needed to understand the full extent of their characteristics. Many a time it is necessary to address the specifics of the mechanism within which an evolutionarily conserved structure mediates permeation to understand the exact role of such a structure. In fact, the structure and the dynamics originating from the former must act in concert for the normal function of aquaporins. To exemplify, we saw that in AQP4, the high conservation H95 might imply that the residue has a physiological role. Studying the dynamics, however, allowed us to find that the collective motion including this residue was affected by protonation and thus modulated by pH. In the AQP0 system it was postulated that the Y23 residue was an unusual replacement for a phenylalanine present in water specific aquaporins. The phenolic group addition to this residue was supposed to be responsible for the reduced permeability of the protein. But, it was with simulations and the functional mode analysis that it became clear that the actual regulation occurred through the conserved arginine residue of the ar/R region and the Y23 residue affected its collective motion to regulate the permeability of the protein. Another interesting facet of regulation is the specificity of the protein towards the permeating molecules. In modeling the TIP2;1 system we saw that the protein

had several unusual features, such as an extended selectivity filter and a highly flexible residue on the extracellular surface. We used MD simulations to verify that these features were not simply artifacts but rather could contribute to a possible gas permeation in a meaningful and plausible model. We believe that these findings expose a more general relation between the protein dynamics and its function. Water permeation is a dynamic phenomenon occurring on a nanosecond timescale. Structures that are functionally static on such a scale, would be too rigid to provide a framework that can be modulated to respond to a variety of stimuli that biological systems face in their lifetime. Mechanisms of regulation and specificity in aquaporins are thus expected to be on the same timescale as well. We discussed how all-atom MD simulations are ideally suited to probe this particular regime in the life of a protein. Our work confirms that such a supposition is not ill founded.

Among the specific details elucidated by this work, the most important is the observation concerning the ar/R region. The arginine residue here has been shown to be *most* important regulator of the rate of permeation. This extends its function beyond its role in the selectivity. From the results shown in the AQP0 project it is clear that in order to make a positive increment in the permeability of an aquaporin, the collective motion within this region *must* be affected. It seems that permeation can be 'switched off' by other means, such as through the H95 gate discussed in chapter three. But, the bottleneck for the channel conductance is strictly contained within the ar/R region. This can be anticipated from the structure itself; but the fact that it can be modulated dynamically was unknown. Additionally, we have shown that the aquaporin protein specificity could also be flexible depending on the exact localization of the protein and the lipid content of the biological membrane in which it is embedded. We confirm the earlier findings [118] that the permeation of alternative gaseous/small hydrophobic substrates might be preferably routed via aquaporins if the membrane is rich in cholesterol. The important determinant of a pore forming protein acting as a channel in the membrane is its capacity to provide an alternative pathway which has a lower free energy barrier compared to the membrane itself. We have observed that the entropic penalty to the entry of a gaseous molecule into a narrow pore of an aquaporin channel might be partially offset by the its stabilization in the selectivity

filter. We saw how the TIP2;1 protein channel had an unusual selectivity filter and an additional side pore to route water molecules. Whether these modifications are necessary or if they only act to further encourage a natural capacity of the protein to assist gas permeation is however still in question. Regardless, new avenues have been opened up as result of this work to identify potential ammonia permeation in aquaporins. This can be accomplished by looking for similar alterations in the lumen of the channel which might act in a manner functionally identical to what can be seen in the TIP2;1 protein.

All in all, MD simulations, have been shown to be effective tools to study the mechanics of regulation. The results of the simulations undertaken as a part of this thesis have independent verification from empirical studies. The collaborations have allowed us to explore a few useful 'rules of thumb' concerning the application of MD simulations to the aquaporin systems. One of our most important observations here was that our results matched *qualitatively* very well with the experiments. We found that exact quantitative match was harder to obtain. We propose a few reasons to explain why such a disparity might exist.

The first is the general problem of equilibration. Mutant proteins constructed from the available X-ray structures are frequently used in this work to verify mechanisms suggested by analyzing the dynamics. We always allow these structures to 'relax' in the simulations prior to analyzing the data from the trajectories. Typically, a hundred nanoseconds worth of data is put aside from the simulation runs to allow for equilibration. This precautionary measure, however, does not guarantee that the end result is exploring a physiological ensemble. Especially mutations that lead to changes in charge or those involved in a lot of structural interactions are expected to be harder to equilibrate within simulation timescales. The problem is further complicated by the high dimensionality of the protein dynamics. Some of the dynamical modes of proteins can relax faster and others might require timescales longer than our simulations windows to equilibrate. This might yield measurement of functional properties associated with these modes either to be very noisy or be trapped in a functional state limited to a local minimum. This was especially true concerning our measurement of the permeability. This function was observed to be very noisy in AQP0, where the baseline was set very low.

This issue could be addressed partially by using the essential dynamics methodology. We used it extensively to trap simulations in functional states of interest and then carry out simulations with restrictions on such a state. This allowed us the benefit of understanding the direct causal relation between the state identified in the analysis and its effect on the characteristic of interest, such as the osmotic permeability. We found the results of these partially constrained simulations to be highly informative and a useful addition to the unconstrained simulation studies. This allowed us to separate purely statistical anomalies from real physical effects within the trajectories. But this can not be guarantee of a *real* effect. The results obtained from such constrained simulations are expected to be used as guides for suggesting experimental verifications and not an end in themselves.

A second possible reason for the observed disparity in the experimental and the simulation osmotic permeability could be the uncertainty in the 'functional' copy number of the aquaporin protein in the membrane within the experiments. Although it is possible to measure the approximate embedded copy number of the molecules in the *Xenopus* oocyte system with fluorescence microscopy, it is harder to determine how many of these are in the 'functional' state with correct localization and permeation conditions. This might make a complete quantitative agreement with the experiment harder. Also, in the computation studies, all the four monomers are assumed to be equivalent in terms of the function. But this may not be true in the experiments, which might have differentially active number of monomers per protein. For example, in the chapter three, the phosphorylation was assumed to be available for all four monomers, while in reality any combination of the four monomers might be subjected to modification. This poses an additional problem in obtaining an exact match between the two regimes of inquiry. But, this particular issue is out of our control and can only be partially addressed in MD simulations by trying out all possible combinations of the modifications although the cost/benefit ratio of this approach is questionable.

The easiest means to compare the simulations and the experiments is thus qualitative, where a comparison can be made between the effectiveness of a modification to the permeability of the protein. If the comparison is off by one or more orders of magnitude, then this approach would be considered inviable.

## 6.2 Future

In all the projects presented here, there are certain avenues which have been left unexplored or under-explored due to the limitations presented by time. Some possible extensions are outlined in the final section so facilitate continuation of the work.

### 6.2.1 Substrate specificity of AQP4

Our collaborator, Dr. Mette Assentoft and Dr. Nanna Macaulay have observed that the protein AQP4 protein can possibly enhance permeation of ammonia across the membranes of cells in nervous tissue. This situation is highly similar to that of the TIP2;1 protein. The membrane of the Glial cells which express AQP4, are highly specialized and a very small portion of that is exposed to the site where ammonia permeation may occur. This allows us to speculate that the lipid composition and the overall very limited surface area for diffusive permeation could work in favor for AQP4 to be the chief pathway for conducting ammonia. In order to verify this, we have calculated the PMF for the ammonia molecule through the channel pore. We are conducting equilibrium simulations with ammonia and ammonium ions in the same lines as the TIP2;1 project to test for the spontaneous permeation of ammonia.

### 6.2.2 'Rescuing' the ammonia permeation in Human AQP1

In the TIP2;1 chapter we saw that the modified SF region of the protein was one of the chief elements speculated to be the gain of ammonia permeation function. Our collaborators, Andreas Kirscht and Prof. Urban Johansson have introduced mutations which modified the SF region of the human APQ1 to resemble that of the TIP2;1 protein. We would like to measure the PMF for ammonia in this modified channel. However, in our previous efforts to achieve this we could not obtain a well equilibrated free energy profile for the mutant. We believe that this was due to the extensive rearrangement needed in the selectivity filter before equilibration of the free protein was achieved. By extending the equilibration

longer we could try and obtain a more relaxed state of the protein which then we could test for the PMF. This would further add credence to the idea that the unusual SF region indeed alters the specificity of permeation in the TIP2;1.

### 6.2.3 A physiological model for the biological membranes

In order to recreate an environment where membrane embedded proteins could 'feel at home' we want to simulate a bilayer with lipid composition similar to that of physiological conditions. Cell membranes of higher organisms tend to be asymmetric and contain a plethora of lipids. Cholesterol is an essential component of the mix and changes the functional properties of the bilayer in many ways including affecting localization of individual proteins and their clustering. We also saw in the fifth chapter that the diffusion barrier to gas permeation also depends on the cholesterol content of the membrane. Recently, new lipid forcefields have become available that provide a significant increase in the number of lipids that can be used for simulations. It would be possible to use these for testing how the asymmetrical nature of the bilayer and the addition of charged lipids such as POPS would affect physical properties of the membrane. Some of the lipids which are very prominent in higher animals, such as Sphingomyelin, have not been parametrized as of yet. Obtaining and validating the corresponding parameters would also be a target of this work. A composition with very high-fidelity with physiological bilayers might be complicated, but simulating the major components would be an ideal starting point.

### 6.2.4 Generality of the 'Arginine mode'

In the fourth chapter we saw that the conserved arginine of the ar/R region constrains the maximum permeation potential in the protein. We would like to know if this collective mode is significantly present in other aquaporins or if it has a unique function in AQP0. The possible generality of this mode might lead to understanding of how the protein might regulate its ability to control permeation of water. The specific mechanism, if it exists, might lead us to understand under

what conditions, such as pH, voltage or mechanical stress does the regulation occur. We already have an ensemble of a variety of members of aquaporins which we could analyze to search for these regulatory switches.







# Bibliography

- [1] Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, and Peter Walter. *Molecular Biology of the Cell*. Garland Science, 4th edition, 2002. ISBN 0-8153-3218-1; 0-8153-4072-9.
- [2] Mario Parisi, Ricardo A. Dorr, Marcelo Ozu, and Roxana Toriano. From membrane pores to aquaporins: 50 years measuring water fluxes. *Journal of Biological Physics*, 33(5-6):331–343, December 2007. ISSN 0092-0606. doi: 10.1007/s10867-008-9064-5. URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2565768/>.
- [3] T Hanai, D A Haydon, and W R Redwood. The water permeability of artificial bimolecular leaflets: a comparison of radio-tracer and osmotic methods. *Annals of the New York Academy of Sciences*, 137(2):731–739, July 1966. ISSN 0077-8923.
- [4] C. V. Paganelli and A. K. Solomon. THE RATE OF EXCHANGE OF TRI-TIATED WATER ACROSS THE HUMAN RED CELL MEMBRANE. *The Journal of General Physiology*, 41(2):259–277, November 1957. ISSN 0022-1295. URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2194835/>.
- [5] J. Dainty and C. R. House. ‘unstirred layers’ in frog skin. *The Journal of Physiology*, 182(1):66–78, January 1966. ISSN 0022-3751. URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1357456/>.
- [6] J. Dainty and C. R. House. An examination of the evidence for membrane pores in frog skin. *The Journal of Physiology*, 185(1):172–184, July 1966. ISSN 0022-3751. URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1395865/>.

- [7] Water permeability of gramicidin a-treated lipid bilayer membranes. *The Journal of General Physiology*, 72(3):341–350, September 1978. ISSN 0022-1295. URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2228543/>.
- [8] Evidence that ADH-stimulated intramembrane particle aggregates are transferred from cytoplasmic to luminal membranes in toad bladder epithelial cells. *The Journal of Cell Biology*, 85(1):83–95, April 1980. ISSN 0021-9525. URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2110591/>.
- [9] M. Parisi and J. Bourguet. The single file hypothesis and the water channels induced by antidiuretic hormone. 71(3):189–193. ISSN 0022-2631.
- [10] R. I. Macey. Transport of water and urea in red blood cells. *The American Journal of Physiology*, 246(3 Pt 1):C195–203, March 1984. ISSN 0002-9513.
- [11] P. Agre, G. M. Preston, B. L. Smith, J. S. Jung, S. Raina, C. Moon, W. B. Guggino, and S. Nielsen. Aquaporin CHIP: the archetypal molecular water channel. *American Journal of Physiology - Renal Physiology*, 265(4):F463–F476, October 1993. URL <http://ajprenal.physiology.org/content/265/4/F463>.
- [12] G. M. Preston, T. P. Carroll, W. B. Guggino, and P. Agre. Appearance of water channels in xenopus oocytes expressing red cell CHIP28 protein. *Science (New York, N.Y.)*, 256(5055):385–387, April 1992. ISSN 0036-8075.
- [13] Jonathan Reizer, Aiala Reizer, and Milton H. Saier. The MIP family of integral membrane channel proteins: Sequence comparisons, evolutionary relationships, reconstructed pathway of evolution, and proposed functional differentiation of the two repeated halves of the proteins. *Critical Reviews in Biochemistry and Molecular Biology*, 28(3):235–257, January 1993. ISSN 1040-9238. doi: 10.3109/10409239309086796. URL <http://informahealthcare.com/doi/abs/10.3109/10409239309086796>.
- [14] J. Bernard Heymann and Andreas Engel. Aquaporins: Phylogeny, structure, and physiology of water channels. *Physiology*, 14(5):187–193, October 1999. ISSN 1548-9213, 1548-9221. URL <http://physiologyonline.physiology.org/content/14/5/187>.

- [15] Kazuyoshi Murata, Kaoru Mitsuoka, Teruhisa Hirai, Thomas Walz, Peter Agre, J. Bernard Heymann, Andreas Engel, and Yoshinori Fujiyoshi. Structural determinants of water permeation through aquaporin-1. *Nature*, 407(6804):599–605, October 2000. ISSN 0028-0836. doi: 10.1038/35036519. URL <http://www.nature.com/nature/journal/v407/n6804/full/407599a0.html>.
- [16] Haixin Sui, Bong-Gyo Han, John K. Lee, Peter Walian, and Bing K. Jap. Structural basis of water-specific transport through the AQP1 water channel. *Nature*, 414(6866):872–878, December 2001. ISSN 0028-0836. doi: 10.1038/414872a. URL <http://www.nature.com/nature/journal/v414/n6866/full/414872a.html>.
- [17] David F Savage, Pascal F Egea, Yaneth Robles-Colmenares, Joseph D. O’Connell III, and Robert M Stroud. Architecture and selectivity in aquaporins: 2.5 x-ray structure of aquaporin z. *PLoS Biol*, 1(3):e72, December 2003. doi: 10.1371/journal.pbio.0000072. URL <http://dx.doi.org/10.1371/journal.pbio.0000072>.
- [18] William E. C. Harries, David Akhavan, Larry J. W. Miercke, Shahram Khademi, and Robert M. Stroud. The channel architecture of aquaporin 0 at a 2.2- resolution. *Proceedings of the National Academy of Sciences of the United States of America*, 101(39):14045–14050, September 2004. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.0405274101. URL <http://www.pnas.org/content/101/39/14045>.
- [19] John K. Lee, David Kozono, Jonathan Remis, Yoshichika Kitagawa, Peter Agre, and Robert M. Stroud. Structural basis for conductance by the archaeal aquaporin AqpM at 1.68 . *Proceedings of the National Academy of Sciences of the United States of America*, 102(52):18932–18937, December 2005. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.0509469102. URL <http://www.pnas.org/content/102/52/18932>.
- [20] Susanna Toernroth-Horsefield, Yi Wang, Kristina Hedfalk, Urban Johansson, Maria Karlsson, Emad Tajkhorshid, Richard Neutze, and Per Kjellbom. Structural mechanism of plant aquaporin gating. *Nature*, 439(7077):688–694,

- February 2006. ISSN 0028-0836. doi: 10.1038/nature04316. URL <http://www.nature.com/nature/journal/v439/n7077/full/nature04316.html>.
- [21] Zachary E. R. Newby, Joseph O'Connell Iii, Yaneth Robles-Colmenares, Shahram Khademi, Larry J. Miercke, and Robert M. Stroud. Crystal structure of the aquaglyceroporin PfAQP from the malarial parasite plasmodium falciparum. *Nature Structural & Molecular Biology*, 15(6):619–625, June 2008. ISSN 1545-9993. doi: 10.1038/nsmb.1431. URL <http://www.nature.com/nsmb/journal/v15/n6/full/nsmb.1431.html>.
- [22] Rob Horsefield, Kristina Nordn, Maria Fellert, Anna Backmark, Susanna Toernroth-Horsefield, Anke C. Terwisscha van Scheltinga, Jan Kvassman, Per Kjellbom, Urban Johanson, and Richard Neutze. High-resolution x-ray structure of human aquaporin 5. *Proceedings of the National Academy of Sciences*, 105(36):13327–13332, September 2008. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.0801466105. URL <http://www.pnas.org/content/105/36/13327>.
- [23] Joseph D. Ho, Ronald Yeh, Andrew Sandstrom, Ilya Chorny, William E. C. Harries, Rebecca A. Robbins, Larry J. W. Miercke, and Robert M. Stroud. Crystal structure of human aquaporin 4 at 1.8 ? and its mechanism of conductance. *Proceedings of the National Academy of Sciences of the United States of America*, 106(18):7437–7442, May 2009. ISSN 0027-8424. doi: 10.1073/pnas.0902725106. URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2678640/>.
- [24] Daxiong Fu, Andrew Libson, Larry J. W. Miercke, Cindy Weitzman, Peter Nollert, Jolanta Krucinski, and Robert M. Stroud. Structure of a glycerol-conducting channel and the basis for its selectivity. *Science*, 290(5491):481–486, October 2000. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.290.5491.481. URL <http://www.sciencemag.org/content/290/5491/481>.
- [25] Gordon J. Cooper, Yuehan Zhou, Patrice Bouyer, Irina I. Grichtchenko, and Walter F. Boron. Transport of volatile solutes through AQP1. *The Journal of Physiology*, 542(1):17–29, July 2002. ISSN 0022-3751, 1469-7793.

- doi: 10.1113/jphysiol.2002.023218. URL <http://jp.physoc.org/content/542/1/17>.
- [26] M. Yasui, A. Hazama, T. H. Kwon, S. Nielsen, W. B. Guggino, and P. Agre. Rapid gating and anion permeability of an intracellular aquaporin. *Nature*, 402(6758):184–187, November 1999. ISSN 0028-0836. doi: 10.1038/46045.
- [27] Anton Burykin and Arieh Warshel. On the origin of the electrostatic barrier for proton transport in aquaporin. *FEBS Letters*, 570(13):41–46, July 2004. ISSN 0014-5793. doi: 10.1016/j.febslet.2004.06.020. URL <http://www.sciencedirect.com/science/article/pii/S0014579304007410>.
- [28] Anne-Kristine Meinild, Dan A. Klaerke, and Thomas Zeuthen. Bidirectional water fluxes and specificity for small hydrophilic molecules in aquaporins 0?5. 273(49):32446–32451. ISSN 0021-9258, 1083-351X. doi: 10.1074/jbc.273.49.32446. URL <http://www.jbc.org/content/273/49/32446>.
- [29] Jochen S. Hub and Bert L. de Groot. Does CO<sub>2</sub> permeate through aquaporin-1? *Biophysical Journal*, 91(3):842–848, August 2006. ISSN 0006-3495. doi: 10.1529/biophysj.106.081406. URL <http://www.sciencedirect.com/science/article/pii/S0006349506717957>.
- [30] Masahiro Ikeda, Eric Beitz, David Kozono, William B. Guggino, Peter Agre, and Masato Yasui. Characterization of aquaporin-6 as a nitrate channel in mammalian cells REQUIREMENT OF PORE-LINING RESIDUE THRE-ONINE 63. *Journal of Biological Chemistry*, 277(42):39873–39879, October 2002. ISSN 0021-9258, 1083-351X. doi: 10.1074/jbc.M207008200. URL <http://www.jbc.org/content/277/42/39873>.
- [31] Kurt H. Meyer. Contributions to the theory of narcosis. 33(0):1062–1064. ISSN 0014-7672. doi: 10.1039/TF9373301062. URL <http://pubs.rsc.org/en/content/articlelanding/1937/tf/tf9373301062>.
- [32] Fr?d?rick de Meyer and Berend Smit. Effect of cholesterol on the structure of a phospholipid bilayer. *Proceedings of the National Academy of Sciences*, 106(10):3654–3658, March 2009. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.0809959106. URL <http://www.pnas.org/content/106/10/3654>.

- [33] de Grotthuss. Sur la d?composition de l'eau et des corps qu'elle tient en dissolution ? l'aide de l'?lectricit? galvanique. *Ann. Chim.*, 58:54?73, 1806.
- [34] B. L. de Groot and H. Grubmueller. Water permeation across biological membranes: mechanism and dynamics of aquaporin-1 and GlpF. *Science (New York, N.Y.)*, 294(5550):2353?2357, December 2001. ISSN 0036-8075. doi: 10.1126/science.1062459.
- [35] Emad Tajkhorshid, Peter Nollert, Morten Jensen, Larry J. W. Miercke, Joseph O'Connell, Robert M. Stroud, and Klaus Schulten. Control of the selectivity of the aquaporin water channel family by global orientational tuning. *Science (New York, N.Y.)*, 296(5567):525?530, April 2002. ISSN 1095-9203. doi: 10.1126/science.1067778.
- [36] Bert L. de Groot, Tomaso Frigato, Volkhard Helms, and Helmut Grubmueller. The mechanism of proton exclusion in the aquaporin-1 water channel. *Journal of Molecular Biology*, 333(2):279?293, October 2003. ISSN 0022-2836.
- [37] Anton Burykin and Arieh Warshel. What really prevents proton transport through aquaporin? charge self-energy versus proton wire proposals. *Biochemical Journal*, 85(6):3696?3706, December 2003. ISSN 0006-3495. URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1303673/>.
- [38] Bert L de Groot and Helmut Grubmueller. The dynamics and energetics of water permeation and proton exclusion in aquaporins. *Current Opinion in Structural Biology*, 15(2):176?183, April 2005. ISSN 0959-440X. doi: 10.1016/j.sbi.2005.02.003. URL <http://www.sciencedirect.com/science/article/pii/S0959440X05000497>.
- [39] Jochen S. Hub, Helmut Grubmueller, and Bert L. de Groot. Dynamics and energetics of permeation through aquaporins. what do we learn from molecular dynamics simulations? *Handbook of Experimental Pharmacology*, (190):57?76, 2009. ISSN 0171-2004. doi: 10.1007/978-3-540-79885-9\_3.
- [40] Yuchun Lin, Zexing Cao, and Yirong Mo. Molecular dynamics simulations on the escherichia coli ammonia channel protein AmtB:? mechanism of ammonia/ammonium transport. *Journal of the American Chemical Society*, 128

- (33):10876–10884, August 2006. ISSN 0002-7863. doi: 10.1021/ja0631549. URL <http://dx.doi.org/10.1021/ja0631549>.
- [41] Jin Yu, Andrea J. Yool, Klaus Schulten, and Emad Tajkhorshid. Mechanism of gating and ion conductivity of a possible tetrameric pore in aquaporin-1. *Structure*, 14(9):1411–1423, September 2006. ISSN 0969-2126. doi: 10.1016/j.str.2006.07.006. URL <http://www.sciencedirect.com/science/article/pii/S0969212606003285>.
- [42] Ruchi Sachdeva and Balvinder Singh. Insights into structural mechanisms of gating induced regulation of aquaporins. *Progress in Biophysics and Molecular Biology*, 114(2):69–79, April 2014. ISSN 0079-6107. doi: 10.1016/j.pbiomolbio.2014.01.002. URL <http://www.sciencedirect.com/science/article/pii/S0079610714000030>.
- [43] Fangqiang Zhu, Emad Tajkhorshid, and Klaus Schulten. Collective diffusion model for water permeation through microscopic channels. *Physical Review Letters*, 93(22):224501, November 2004. doi: 10.1103/PhysRevLett.93.224501. URL <http://link.aps.org/doi/10.1103/PhysRevLett.93.224501>.
- [44] B. Smit D. Frenkel. Understanding molecular simulations: from algorithms to applications. *Physics Today*, 50(7), 1996. ISSN 0031-9228. doi: 10.1063/1.881812.
- [45] E. Schroedinger. An undulatory theory of the mechanics of atoms and molecules. *Physical Review*, 28(6):1049–1070, December 1926. doi: 10.1103/PhysRev.28.1049. URL <http://link.aps.org/doi/10.1103/PhysRev.28.1049>.
- [46] M. Born and R. Oppenheimer. Zur quantentheorie der moleküle. *Annalen der Physik*, 389(20):457–484, 1927. ISSN 1521-3889. doi: 10.1002/andp.19273892002. URL <http://onlinelibrary.wiley.com/doi/10.1002/andp.19273892002/abstract>.
- [47] P. Ehrenfest. Bemerkung ueber die angeneaherte gueltigkeit der klassischen mechanik innerhalb der quantenmechanik. *Zeitschrift fuer Physik*, 45(7-8):

- 455–457, July 1927. ISSN 0044-3328. doi: 10.1007/BF01329203. URL <http://link.springer.com/article/10.1007/BF01329203>.
- [48] Nicolas Foloppe and Alexander D. MacKerell, Jr. All-atom empirical force field for nucleic acids: I. parameter optimization based on small molecule and condensed phase macromolecular target data. *Journal of Computational Chemistry*, 21(2):86–104, 2000. ISSN 1096-987X. doi: 10.1002/(SICI)1096-987X(20000130)21:2<86::AID-JCC2>3.0.CO;2-G. URL [http://onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1096-987X\(20000130\)21:2<86::AID-JCC2>3.0.CO;2-G/abstract](http://onlinelibrary.wiley.com/doi/10.1002/(SICI)1096-987X(20000130)21:2<86::AID-JCC2>3.0.CO;2-G/abstract).
- [49] Jing Huang and Alexander D. MacKerell. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *Journal of Computational Chemistry*, 34(25):2135–2145, September 2013. ISSN 1096-987X. doi: 10.1002/jcc.23354.
- [50] Nadine Homeyer, Anselm H. C. Horn, Harald Lanig, and Heinrich Sticht. AMBER force-field parameters for phosphorylated amino acids in different protonation states: phosphoserine, phosphothreonine, phosphotyrosine, and phosphohistidine. *Journal of Molecular Modeling*, 12(3):281–289, February 2006. ISSN 0948-5023. doi: 10.1007/s00894-005-0028-4.
- [51] K. Anton Feenstra, Berk Hess, and Herman J. C. Berendsen. Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *Journal of Computational Chemistry*, 20(8):786–798, June 1999. ISSN 1096-987X. doi: 10.1002/(SICI)1096-987X(199906)20:8<786::AID-JCC5>3.0.CO;2-B. URL [http://onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1096-987X\(199906\)20:8<786::AID-JCC5>3.0.CO;2-B/abstract](http://onlinelibrary.wiley.com/doi/10.1002/(SICI)1096-987X(199906)20:8<786::AID-JCC5>3.0.CO;2-B/abstract).
- [52] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans. Interaction models for water in relation to protein hydration. In Bernard Pullman, editor, *Intermolecular Forces*, number 14 in The Jerusalem Symposia on Quantum Chemistry and Biochemistry, pages 331–342. Springer Netherlands, January 1981. ISBN 978-90-481-8368-5,

- 978-94-015-7658-1. URL [http://link.springer.com/chapter/10.1007/978-94-015-7658-1\\_21](http://link.springer.com/chapter/10.1007/978-94-015-7658-1_21).
- [53] Giovanni Bussi, Davide Donadio, and Michele Parrinello. Canonical sampling through velocity rescaling. *The Journal of Chemical Physics*, 126(1):014101–014101–7, January 2007. ISSN 00219606. doi: doi:10.1063/1.2408420. URL [http://jcp.aip.org/resource/1/jcpa6/v126/i1/p014101\\_s1](http://jcp.aip.org/resource/1/jcpa6/v126/i1/p014101_s1).
- [54] Christopher J. Cramer and Donald G. Truhlar. Continuum solvation models. In Orlando Tapia and Juan Bertrn, editors, *Solvent Effects and Chemical Reactivity*, number 17 in Understanding Chemical Reactivity, pages 1–80. Springer Netherlands, January 2002. ISBN 978-0-7923-3995-3, 978-0-306-46931-2. URL [http://link.springer.com/chapter/10.1007/0-306-46931-6\\_1](http://link.springer.com/chapter/10.1007/0-306-46931-6_1).
- [55] William L. Jorgensen, Jayaraman Chandrasekhar, Jeffry D. Madura, Roger W. Impey, and Michael L. Klein. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2):926–935, July 1983. ISSN 00219606. doi: doi:10.1063/1.445869. URL [http://jcp.aip.org/resource/1/jcpa6/v79/i2/p926\\_s1](http://jcp.aip.org/resource/1/jcpa6/v79/i2/p926_s1).
- [56] Tom Darden, Darrin York, and Lee Pedersen. Particle mesh ewald: An  $n \log(n)$  method for ewald sums in large systems. *The Journal of Chemical Physics*, 98(12):10089–10092, June 1993. ISSN 0021-9606, 1089-7690. doi: 10.1063/1.464397. URL <http://scitation.aip.org/content/aip/journal/jcp/98/12/10.1063/1.464397>.
- [57] Tatyana Krivobokova, Rodolfo Briones, Jochen S. Hub, Axel Munk, and Bert L. de Groot. Partial least-squares functional mode analysis: application to the membrane proteins AQP1, aquaporin 1, and CLC-ec1. *Biophysical Journal*, 103(4):786–796, August 2012. ISSN 1542-0086. doi: 10.1016/j.bpj.2012.07.022.
- [58] Inge S. Helland. Partial least squares regression and statistical models. *Scandinavian Journal of Statistics*, 17(2):97–114, January 1990. ISSN 0303-6898. URL <http://www.jstor.org/stable/4616159>.

- [59] M. C. Denham. Implementing partial least squares. *Statistics and Computing*, 5(3):191–202, September 1995. ISSN 0960-3174, 1573-1375. doi: 10.1007/BF00142661. URL <http://link.springer.com/article/10.1007/BF00142661>.
- [60] Glenn M. Torrie and John P. Valleau. Monte carlo free energy estimates using non-boltzmann sampling: Application to the sub-critical lennard-jones fluid. *Chemical Physics Letters*, 28(4):578–581, October 1974. ISSN 0009-2614. doi: 10.1016/0009-2614(74)80109-0. URL <http://www.sciencedirect.com/science/article/pii/0009261474801090>.
- [61] Shankar Kumar, John M. Rosenberg, Djamal Bouzida, Robert H. Swendsen, and Peter A. Kollman. THE weighted histogram analysis method for free-energy calculations on biomolecules. i. the method. *Journal of Computational Chemistry*, 13(8):1011–1021, 1992. ISSN 1096-987X. doi: 10.1002/jcc.540130812. URL <http://onlinelibrary.wiley.com/doi/10.1002/jcc.540130812/abstract>.
- [62] Mahmood Amiry-Moghaddam and Ole P. Ottersen. The molecular basis of water transport in the brain. *Nature Reviews Neuroscience*, 4(12):991–1001, December 2003. ISSN 1471-003X. doi: 10.1038/nrn1252. URL <http://www.nature.com/nrn/journal/v4/n12/full/nrn1252.html>.
- [63] H. Nielsen, J. Engelbrecht, S. Brunak, and G. von Heijne. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Engineering*, 10(1):1–6, January 1997. ISSN 1741-0126, 1741-0134. doi: 10.1093/protein/10.1.1. URL <http://peds.oxfordjournals.org/content/10/1/1>.
- [64] Zsolt Zador, Shirley Stiver, Vincent Wang, and Geoffrey T. Manley. Role of aquaporin-4 in cerebral edema and stroke. *Handbook of Experimental Pharmacology*, (190):159–170, 2009. ISSN 0171-2004. doi: 10.1007/978-3-540-79885-9\_7.
- [65] Kaoru Beppu, Takuya Sasaki, Kenji F. Tanaka, Akihiro Yamanaka, Yugo Fukazawa, Ryuichi Shigemoto, and Ko Matsui. Optogenetic countering of glial acidosis suppresses glial glutamate release and ischemic brain damage.

- Neuron*, 81(2):314–320, January 2014. ISSN 1097-4199. doi: 10.1016/j.neuron.2013.11.011.
- [66] R. P. Kraig and M. Chesler. Astrocytic acidosis in hyperglycemic and complete ischemia. *Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism*, 10(1):104–114, January 1990. ISSN 0271-678X. doi: 10.1038/jcbfm.1990.13.
- [67] C. Lascola and R. P. Kraig. Astroglial acid-base dynamics in hyperglycemic and normoglycemic global ischemia. *Neuroscience and Biobehavioral Reviews*, 21(2):143–150, March 1997. ISSN 0149-7634.
- [68] F. Ringel, R. C. Chang, F. Staub, A. Baethmann, and N. Plesnila. Contribution of anion transporters to the acidosis-induced swelling and intracellular acidification of glial cells. *Journal of Neurochemistry*, 75(1):125–132, July 2000. ISSN 0022-3042.
- [69] Eli Gunnarson, Marina Zelenina, Gustav Axehult, Yutong Song, Alexander Bondar, Patrik Krieger, Hjalmar Brismar, Sergey Zelenin, and Anita Apelia. Identification of a molecular target for glutamate regulation of astrocyte water permeability. *Glia*, 56(6):587–596, April 2008. ISSN 1098-1136. doi: 10.1002/glia.20627. URL <http://onlinelibrary.wiley.com/doi/10.1002/glia.20627/abstract>.
- [70] Marina Zelenina, Alexander A. Bondar, Sergey Zelenin, and Anita Apelia. Nickel and extracellular acidification inhibit the water permeability of human aquaporin-3 in lung epithelial cells. *The Journal of Biological Chemistry*, 278(32):30037–30043, August 2003. ISSN 0021-9258. doi: 10.1074/jbc.M302206200.
- [71] Yubao Cui and David A. Bastien. Water transport in human aquaporin-4: Molecular dynamics (MD) simulations. *Biochemical and biophysical research communications*, 412(4), September 2011. ISSN 0006-291X. doi: 10.1016/j.bbrc.2011.08.019. URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3809129/>.
- [72] J.-A. Garate, Niall J. English, and J. M. D. MacElroy. Human aquaporin 4 gating dynamics in dc and ac electric fields: a molecular dynamics study.

- The Journal of Chemical Physics*, 134(5):055110, February 2011. ISSN 1089-7690. doi: 10.1063/1.3529428.
- [73] Riccardo Reale, Niall J. English, Jos-Antonio Garate, Paolo Marracino, Micaela Liberti, and Francesca Apollonio. Human aquaporin 4 gating dynamics under and after nanosecond-scale static and alternating electric-field impulses: A molecular dynamics study of field effects and relaxation. *The Journal of Chemical Physics*, 139(20):205101, November 2013. ISSN 0021-9606, 1089-7690. doi: 10.1063/1.4832383. URL <http://scitation.aip.org/content/aip/journal/jcp/139/20/10.1063/1.4832383>.
- [74] Robert A. Fenton, Hanne B. Moeller, Marina Zelenina, Marteinn T. Snæbjornsson, Torgeir Holen, and Nanna MacAulay. Differential water permeability and regulation of three aquaporin 4 isoforms. *Cellular and Molecular Life Sciences*, 67(5):829–840, March 2010. ISSN 1420-682X, 1420-9071. doi: 10.1007/s00018-009-0218-9. URL <http://link.springer.com/article/10.1007/s00018-009-0218-9>.
- [75] Yi Wang and Emad Tajkhorshid. Nitric oxide conduction by the brain aquaporin AQP4. *Proteins: Structure, Function, and Bioinformatics*, 78(3):661–670, February 2010. ISSN 1097-0134. doi: 10.1002/prot.22595. URL <http://onlinelibrary.wiley.com/doi/10.1002/prot.22595/abstract>.
- [76] Domenico Alberga, Orazio Nicolotti, Gianluca Lattanzi, Grazia Paola Nicchia, Antonio Frigeri, Francesco Pisani, Valentina Benfenati, and Giuseppe Felice Mangiatordi. A new gating site in human aquaporin-4: Insights from molecular dynamics simulations. *Biochimica Et Biophysica Acta*, August 2014. ISSN 0006-3002. doi: 10.1016/j.bbamem.2014.08.015.
- [77] Erlend A. Nagelhus and Ole P. Ottersen. Physiological roles of aquaporin-4 in brain. *Physiological Reviews*, 93(4):1543–1562, October 2013. ISSN 0031-9333, 1522-1210. doi: 10.1152/physrev.00011.2013. URL <http://physrev.physiology.org/content/93/4/1543>.
- [78] Jrme Badaut, Andrew M. Fukuda, Amandine Jullienne, and Klaus G. Petry. Aquaporin and brain diseases. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1840(5):1554–1565, May 2014. ISSN 0304-4165.

- doi: 10.1016/j.bbagen.2013.10.032. URL <http://www.sciencedirect.com/science/article/pii/S0304416513004728>.
- [79] Bong-Gyo Han, Anton B. Guliaev, Peter J. Walian, and Bing K. Jap. Water transport in AQP0 aquaporin: Molecular dynamics studies. 360(2): 285–296. ISSN 0022-2836. doi: 10.1016/j.jmb.2006.04.039. URL <http://www.sciencedirect.com/science/article/pii/S0022283606005134>.
- [80] Marina Zelenina, Sergey Zelenin, Alexander A. Bondar, Hjalmar Brismar, and Anita Aperia. Water permeability of aquaporin 4 is decreased by protein kinase c and dopamine. doi: 10.1152/ajprenal.00260.2001. URL <http://ajprenal.physiology.org/content/early/2002/02/19/ajprenal.00260.2001>.
- [81] Yutong Song and Eli Gunnarson. Potassium dependent regulation of astrocyte water permeability is mediated by cAMP signaling. *PLoS ONE*, 7(4):e34936, April 2012. doi: 10.1371/journal.pone.0034936. URL <http://dx.doi.org/10.1371/journal.pone.0034936>.
- [82] G. Vriend. WHAT IF: A molecular modeling and drug design program. *Journal of Molecular Graphics*, 8(1):52–56, March 1990. ISSN 0263-7855. doi: 10.1016/0263-7855(90)80070-V. URL <http://www.sciencedirect.com/science/article/pii/026378559080070V>.
- [83] Mette Assentoft, Shreyas Kaptan, Robert A. Fenton, Susan Z. Hua, Bert L. de Groot, and Nanna MacAulay. Phosphorylation of rat aquaporin-4 at ser111 is not required for channel gating. *Glia*, 61(7):1101–1112, July 2013. ISSN 1098-1136. doi: 10.1002/glia.22498. URL <http://onlinelibrary.wiley.com/doi/10.1002/glia.22498/abstract>.
- [84] Flavia Autore, Bruno Pagano, Arianna Fornili, Katrin Rittinger, and Franca Fraternali. In silico phosphorylation of the autoinhibited form of p47phox: Insights into the mechanism of activation. *Biophysical Journal*, 99(11):3716–3725, December 2010. ISSN 0006-3495. doi: 10.1016/j.bpj.2010.09.008. URL <http://www.sciencedirect.com/science/article/pii/S0006349510011057>.

- [85] Tobias Linder, Bert L. de Groot, and Anna Stary-Weinzinger. Probing the energy landscape of activation gating of the bacterial potassium channel KcsA. *PLoS Comput Biol*, 9(5):e1003058, May 2013. doi: 10.1371/journal.pcbi.1003058. URL <http://dx.doi.org/10.1371/journal.pcbi.1003058>.
- [86] K. L. Nemeth-Cahalan and J. E. Hall. pH and calcium regulate the water permeability of aquaporin 0. *The Journal of Biological Chemistry*, 275(10): 6777–6782, March 2000. ISSN 0021-9258.
- [87] Thomas Zeuthen and Dan Arne Klaerke. Transport of water and glycerol in aquaporin 3 is gated by h+. *Journal of Biological Chemistry*, 274(31): 21631–21636, July 1999. ISSN 0021-9258, 1083-351X. doi: 10.1074/jbc.274.31.21631. URL <http://www.jbc.org/content/274/31/21631>.
- [88] Karin L. Nemeth-Cahalan, Katalin Kalman, and James E. Hall. Molecular basis of pH and ca<sup>2+</sup> regulation of aquaporin water permeability. *The Journal of General Physiology*, 123(5):573–580, May 2004. ISSN 0022-1295. doi: 10.1085/jgp.200308990. URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2234493/>.
- [89] Lorant Janosi and Matteo Ceccarelli. The gating mechanism of the human aquaporin 5 revealed by molecular dynamics simulations. *PLoS ONE*, 8 (4):e59897, April 2013. doi: 10.1371/journal.pone.0059897. URL <http://dx.doi.org/10.1371/journal.pone.0059897>.
- [90] Michael B. Gorin, S. Barbara Yancey, Janice Cline, Jean-Paul Revel, and Joseph Horwitz. The major intrinsic protein (MIP) of the bovine lens fiber membrane: Characterization and structure based on cDNA cloning. *Cell*, 39 (1):49–59, November 1984. ISSN 0092-8674. doi: 10.1016/0092-8674(84)90190-9. URL <http://www.sciencedirect.com/science/article/pii/0092867484901909>.
- [91] Morten Jensen, Ron O. Dror, Huafeng Xu, David W. Borhani, Isaiah T. Arkin, Michael P. Eastwood, and David E. Shaw. Dynamic control of slow water transport by aquaporin 0: Implications for hydration and junction stability in the eye lens. *Proceedings of the National Academy of Sciences*, 105(38):14430–14435, September 2008. ISSN 0027-8424, 1091-6490. doi:

- 10.1073/pnas.0802401105. URL <http://www.pnas.org/content/105/38/14430>.
- [92] Ana B. Chepelinsky. Structural function of MIP/aquaporin 0 in the eye lens; genetic defects lead to congenital inherited cataracts. In Prof Dr Eric Beitz, editor, *Aquaporins*, number 190 in Handbook of Experimental Pharmacology, pages 265–297. Springer Berlin Heidelberg, January 2009. ISBN 978-3-540-79884-2, 978-3-540-79885-9. URL [http://link.springer.com/chapter/10.1007/978-3-540-79885-9\\_14](http://link.springer.com/chapter/10.1007/978-3-540-79885-9_14).
- [93] K. Varadaraj, C. Kushmerick, G. J. Baldo, S. Bassnett, A. Shiels, and R. T. Mathias. The role of MIP in lens fiber cell membrane transport. *The Journal of Membrane Biology*, 170(3):191–203, August 1999. ISSN 0022-2631, 1432-1424. doi: 10.1007/s002329900549. URL <http://link.springer.com/article/10.1007/s002329900549>.
- [94] M. J. Costello, T. J. McIntosh, and J. D. Robertson. Distribution of gap junctions and square array junctions in the mammalian lens. *Investigative Ophthalmology & Visual Science*, 30(5):975–989, May 1989. ISSN 0146-0404.
- [95] G. Chandy, G. A. Zampighi, M. Kreman, and J. E. Hall. Comparison of the water transporting properties of MIP and AQP1. *The Journal of Membrane Biology*, 159(1):29–39, September 1997. ISSN 0022-2631, 1432-1424. doi: 10.1007/s002329900266. URL <http://link.springer.com/article/10.1007/s002329900266>.
- [96] Tamir Gonen, Yifan Cheng, Piotr Sliz, Yoko Hiroaki, Yoshinori Fujiyoshi, Stephen C. Harrison, and Thomas Walz. Lipidprotein interactions in double-layered two-dimensional AQP0 crystals. *Nature*, 438(7068):633–638, December 2005. ISSN 0028-0836. doi: 10.1038/nature04321. URL <http://www.nature.com/nature/journal/v438/n7068/full/nature04321.html>.
- [97] Hu Qiu, Shaojie Ma, Rong Shen, and Wanlin Guo. Dynamic and energetic mechanisms for the distinct permeation rate in AQP1 and AQP0. *Biochimica Et Biophysica Acta*, 1798(3):318–326, March 2010. ISSN 0006-3002. doi: 10.1016/j.bbamem.2009.11.015.

- [98] Masanori Hashido, Mitsunori Ikeguchi, and Akinori Kidera. Comparative simulations of aquaporin family: AQP1, AQPZ, AQP0 and GlpF. *FEBS Letters*, 579(25):5549–5552, October 2005. ISSN 0014-5793. doi: 10.1016/j.febslet.2005.09.018. URL <http://www.sciencedirect.com/science/article/pii/S0014579305011233>.
- [99] Tamir Gonen, Piotr Sliz, Joerg Kistler, Yifan Cheng, and Thomas Walz. Aquaporin-0 membrane junctions reveal the structure of a closed water pore. *Nature*, 429(6988):193–197, May 2004. ISSN 0028-0836. doi: 10.1038/nature02503. URL <http://www.nature.com/nature/journal/v429/n6988/full/nature02503.html>.
- [100] William E. C. Harries, David Akhavan, Larry J. W. Miercke, Shahram Khademi, and Robert M. Stroud. The channel architecture of aquaporin 0 at a 2.2- resolution. *Proceedings of the National Academy of Sciences of the United States of America*, 101(39):14045–14050, September 2004. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.0405274101. URL <http://www.pnas.org/content/101/39/14045>.
- [101] Camilo Aponte-Santamaría, Rodolfo Briones, Andreas D. Schenk, Thomas Walz, and Bert L. de Groot. Molecular driving forces defining lipid positions around aquaporin-0. *Proceedings of the National Academy of Sciences*, 109(25):9887–9892, June 2012. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1121054109. URL <http://www.pnas.org/content/109/25/9887>.
- [102] G. V. Ramesh Prasad, Larry A. Coury, Frances Finn, and Mark L. Zeidel. Reconstituted aquaporin 1 water channels transport CO<sub>2</sub> across membranes. *Journal of Biological Chemistry*, 273(50):33123–33126, December 1998. ISSN 0021-9258, 1083-351X. doi: 10.1074/jbc.273.50.33123. URL <http://www.jbc.org/content/273/50/33123>.
- [103] V. Endeward, R. Musa-Aziz, G. J. Cooper, L.-M. Chen, M. F. Pelletier, L. V. Virkki, C. T. Supuran, L. S. King, W. F. Boron, and G. Gros. Evidence that aquaporin 1 is a major pathway for CO<sub>2</sub> transport across the human erythrocyte membrane. *The FASEB Journal*, 20(12):1974–1981, October

2006. ISSN 0892-6638, 1530-6860. doi: 10.1096/fj.04-3300com. URL <http://www.fasebj.org/content/20/12/1974>.
- [104] N. L. Nakhoul, B. A. Davis, M. F. Romero, and W. F. Boron. Effect of expressing the water channel aquaporin-1 on the CO<sub>2</sub> permeability of xenopus oocytes. *The American Journal of Physiology*, 274(2 Pt 1):C543–548, February 1998. ISSN 0002-9513.
- [105] Yuko T. Hanba, Mineo Shibasaki, Yasuyuki Hayashi, Takahiko Hayakawa, Kunihiro Kasamo, Ichiro Terashima, and Maki Katsuhara. Overexpression of the barley aquaporin HvPIP2;1 increases internal CO<sub>2</sub> conductance and CO<sub>2</sub> assimilation in the leaves of transgenic rice plants. *Plant and Cell Physiology*, 45(5):521–529, May 2004. ISSN 0032-0781, 1471-9053. doi: 10.1093/pcp/pch070. URL <http://pcp.oxfordjournals.org/content/45/5/521>.
- [106] Lars M. Holm, Thomas P. Jahn, Anders L. B. Mller, Jan K. Schjoerring, Domenico Ferri, Dan A. Klaerke, and Thomas Zeuthen. NH<sub>3</sub> and NH<sub>4</sub><sup>+</sup> permeability in aquaporin-expressing xenopus oocytes. *Pflgers Archiv: European Journal of Physiology*, 450(6):415–428, September 2005. ISSN 0031-6768. doi: 10.1007/s00424-005-1399-1.
- [107] R. Ryan Geyer, Raif Musa-Aziz, Xue Qin, and Walter F. Boron. Relative CO(2)/NH(3) selectivities of mammalian aquaporins 0-9. *American Journal of Physiology. Cell Physiology*, 304(10):C985–994, May 2013. ISSN 1522-1563. doi: 10.1152/ajpcell.00033.2013.
- [108] Raif Musa-Aziz, Li-Ming Chen, Marc F. Pelletier, and Walter F. Boron. Relative CO<sub>2</sub>/NH<sub>3</sub> selectivities of AQP1, AQP4, AQP5, AmtB, and RhAG. *Proceedings of the National Academy of Sciences of the United States of America*, 106(13):5406–5411, March 2009. ISSN 1091-6490. doi: 10.1073/pnas.0813231106.
- [109] Marcela Herrera, Nancy J. Hong, and Jeffrey L. Garvin. Aquaporin-1 transports NO across cell membranes. *Hypertension*, 48(1):157–164, July 2006. ISSN 1524-4563. doi: 10.1161/01.HYP.0000223652.29338.77.
- [110] C Maurel, J Reizer, J I Schroeder, and M J Chrispeels. The vacuolar membrane protein gamma-TIP creates water specific channels in xenopus oocytes.

- The EMBO Journal*, 12(6):2241–2247, June 1993. ISSN 0261-4189. URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC413452/>.
- [111] Thomas P. Jahn, Anders L. B. Mller, Thomas Zeuthen, Lars M. Holm, Dan A. Klrke, Brigitte Mohsin, Werner Khlbrandt, and Jan K. Schjoerring. Aquaporin homologues in plants and mammals transport ammonia. *FEBS Letters*, 574(13):31–36, September 2004. ISSN 0014-5793. doi: 10.1016/j.febslet.2004.08.004. URL <http://www.sciencedirect.com/science/article/pii/S0014579304009743>.
- [112] Dominique Loqu, Uwe Ludewig, Lixing Yuan, and Nicolaus von Wirn. Tono-plast intrinsic proteins AtTIP2;1 and AtTIP2;3 facilitate NH<sub>3</sub> transport into the vacuole. *Plant Physiology*, 137(2):671–680, February 2005. ISSN 0032-0889. doi: 10.1104/pp.104.051268. URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1065367/>.
- [113] Marek Dynowski, Maria Mayer, Oscar Moran, and Uwe Ludewig. Molecular determinants of ammonia and urea conductance in plant aquaporin homologs. *FEBS Letters*, 582(16):2458–2462, July 2008. ISSN 0014-5793. doi: 10.1016/j.febslet.2008.06.012. URL <http://www.sciencedirect.com/science/article/pii/S0014579308005012>.
- [114] Hanna I. Anderberg, Per Kjellbom, and Urban Johanson. Annotation of selaginella moellendorffii major intrinsic proteins and the evolution of the protein family in terrestrial plants. *Frontiers in Plant Science*, 3:33, 2012. ISSN 1664-462X. doi: 10.3389/fpls.2012.00033.
- [115] Masayoshi Maeshima. TONOPLAST TRANSPORTERS: Organization and function. *Annual Review of Plant Physiology and Plant Molecular Biology*, 52(1):469–497, 2001. doi: 10.1146/annurev.arplant.52.1.469. URL <http://dx.doi.org/10.1146/annurev.arplant.52.1.469>.
- [116] Michael M. Wudick, Doan-Trung Luu, and Christophe Maurel. A look inside: localization patterns and functions of intracellular plant aquaporins. *The New Phytologist*, 184(2):289–302, October 2009. ISSN 1469-8137. doi: 10.1111/j.1469-8137.2009.02985.x.

- [117] Devrim Coskun, Dev T. Britto, Mingyuan Li, Alexander Becker, and Herbert J. Kronzucker. Rapid ammonia gas transport accounts for futile transmembrane cycling under NH<sub>3</sub>/NH<sub>4</sub><sup>+</sup> toxicity in plant roots. *Plant Physiology*, 163(4):1859–1867, December 2013. ISSN 0032-0889, 1532-2548. doi: 10.1104/pp.113.225961. URL <http://www.plantphysiol.org/content/163/4/1859>.
- [118] Jochen S. Hub, Fritz K. Winkler, Mike Merrick, and Bert L. de Groot. Potentials of mean force and permeabilities for carbon dioxide, ammonia, and water flux across a rhesus protein channel and lipid membranes. *Journal of the American Chemical Society*, 132(38):13251–13263, September 2010. ISSN 0002-7863. doi: 10.1021/ja102133x. URL <http://dx.doi.org/10.1021/ja102133x>.







# Curriculum Vitae:

## Shreyas Sanjay Kaptan

### Personal Information

**Name:** Shreyas Sanjay Kaptan

**Address:** Stumpfe Eiche 13, 37077 Göttingen

**Date and city of birth:** 24. October, 1985, Nagpur

**Nationaity:** Indian

### Education

- 07.2011– Max Planck Institute for Biophysical Chemistry, Göttingen  
Thesis: *Regulation of Permeation in Aquaporins*  
Supervisor: Prof. Dr. Bert de Groot
- 10.2003–11.2009 University of Pune  
Institute of Bioinformatics and Biotechnology (IBB)  
A six years integrated course  
Thesis: *Developing a QM:MM methodology for studying spectroscopic and structural properties of proteins*  
Supervisor: Dr. Mrinalini Puranik, National Center for Biological Sciences (NCBS)  
Graduated with degree *Master of Technology (M.Tech.)*  
Final grade: "Outstanding" (O)

## Articles in peer-reviewed journals

- Pinjari, R. V., Kaptan, S. S. & Gejji, S. P. *Alkali metals (Li, Na, and K) in methyl phosphodiester hydrolysis.* Phys. Chem. Chem. Phys. 11, 5253-5262 (2009).
- Garces, E. A., Kaptan, S. & Puranik, M. *Mode-Specific Reorganization Energies and Ultrafast Solvation Dynamics of Tryptophan from Raman Line-Shape Analysis.* Biophysical Journal 105, 211-211221 (2013)
- Assentoft, M., Kaptan, S., Fenton, R.A., Hua, S.Z., de Groot, B.L., and MacAulay, N. *Phosphorylation of rat aquaporin-4 at Ser111 is not required for channel gating.* Glia 61, 11011112. (2013).

