Structure and Function of Aqua(glycero)porins

A Path to Drug Design

Dipl. Chem. Gerhard Fischer

GÖTEBORGS UNIVERSITET

Meinen Eltern

Abstract

Membrane proteins are a major class of the drug targets. Rational design of drugs requires knowledge about the structure and function of these proteins. Working with membrane proteins, however, is experimentally challenging as the required technology is still emerging.

Water being the medium of life, Nature has evolved specialized water channels called aquaporins. These transporters are found in all kingdoms of life – from simple microorganisms to humans – and function to maintain water homeostasis. These water channels are surprisingly similar in structure and are divided into two major groups: While orthodox aquaporins transport water only, the closely related aquaglyceroporins also transport small solutes like glycerol or other sugar alcohols. Aquaporins have been suggested to be involved in a wide variety of human disorders, but also as a means to combat malicious parasites.

During the course of this work, we discovered the only aquaporin from the yeast *Pichia pastoris*. We determined its atomic structure using X-ray crystallography to 0.90 Å resolution, the highest resolution for any membrane protein to date. Surprisingly, this structure showed the water pore being closed by its elongated N-terminus and provided in detail information on the water transport mechanism through this pore. Further functional characterization by a combination of size-based water conduction assays and temperature-dependent crystallography revealed that the channel can be opened. We suggested a gating mechanism *via* phosphorylation and mechanosensing, which was found to be beneficial to the survival of the organism.

Medically, protein structures like these can also be used for rational drug design. For aqua(glycero)porins, however, the discovery of inhibitors is hampered by the lack of a medium- to high-throughput assay to test potential channel blockers. Using novel technology based on surface plasmon resonance (Biacore™), we established a fully automated drug screening assay with the capability to test up to 50 compounds/day. Application of this method in combination with virtual screening yielded a drug lead with an IC_{50}-value of 1-10 µM against the aquaglyceroporin from the malaria parasite *Plasmodium falciparum*.

Contribution report

Paper I   I analyzed and compared the crystal structures and prepared figures.

Paper II   I was involved in the entire project. I was responsible for cloning the constructs, producing, purifying and crystallizing protein, collecting diffraction data and determining, refining and analyzing the 1.15 Å crystal structure. I also conducted the functional studies based on light scattering and took part in writing the manuscript.

Paper III   I planned the project, optimized protein crystals, collected diffraction data and performed initial data evaluation and structure solution. I wrote a minor part of the manuscript.

Paper IV   I planned the project, optimized protein crystals, collected diffraction data, determined and analyzed the structure and wrote major parts of the manuscript.

Paper V   I produced and purified protein, prepared proteoliposomes, performed Biacore-experiments on proteoliposomes and docking simulations. I wrote a minor part of the manuscript.

Paper VI   I planned and conducted the Biacore experiments, analyzed the data and took a major part in writing the manuscript.

List of Publications

Paper I   Törnroth-Horsefield S, Hedfalk K, **Fischer G**, Lindkvist-Petersson K, Neutze R. "Structural insights into eukaryotic aquaporin regulation." *FEBS Letters.* 582(12), 2580-2588 (2010). Review.

Paper II   **Fischer G**, Kosinska-Eriksson U, Aponte-Santamarìa C, Palmgren M, Geijer C, Hedfalk K, Hohmann S, de Groot B, Neutze R, Lindkvist-Petersson K. "Crystal Structure of a Yeast Aquaporin at 1.15 Å reveals novel gating mechanism." *PLoS Biology.* 7(6); (2009)

Paper III   Kosinska-Eriksson U, **Fischer G**, Friemann R, Neutze R. "Sub-atomic Resolution Structure of a Yeast Aquaporin Reveals a Low-barrier Mechanism for Selective Water Transport.", *manuscript*

Paper IV   **Fischer G**, Kosinska-Eriksson U, Neutze R. "Ambient temperature crystal structure of a yeast aquaporin implies mechanosensitive gating.", *manuscript*

Paper V   Brändén M, Tabaei SR, **Fischer G**, Neutze R, Höök F. "Refractive-Index Based Screening of Membrane-Protein Mediated Transfer across Biological Membranes." *Biophysical Journal.* 99(1), 124-133 (2010)

Paper VI   **Fischer G**, Moberg A, Sjöhamn J, Tabaei SR, Brändén M, Höök F, Hedfalk K, Fishwick C, Johnson P, Neutze R, Simmons K. "Discovery of Inhibitors of the Aquaglyceroporin of *Plasmodium falciparum* using Virtual High-throughput Screening and a Novel *in vitro* Method for Membrane Protein Channels.", *manuscript*

Related Publications

Paper VII   Öberg F, Sjöhamn J, **Fischer G**, Moberg A, Pedersen A, Neutze R, Hedfalk K. "Glycosylation increases the Thermostability of Human Aquaporin 10." *manuscript*

Paper VIII   Saline M, Rödström KEJ, **Fischer G**, Orekhov VY, Karlsson BG, Lindkvist-Petersson K. "The structure of superantigen complexed with TCR and MHC reveals novel insights into superantigenic T cell activation.", *Nature Communications.* 1(8), 119 (2010)

Paper IX   Wöhri AB, Johansson LC, Wadsten-Hindrichsen P, Wahlgren WY, **Fischer G**, Horsefield R, Katona G, Nyblom M, Oberg F, Young G, Cogdell RJ, Fraser NJ, Engström S, Neutze R. "A lipidic-sponge phase screen for membrane protein crystallization." *Structure.* 16(7), 1003-1009 (2008)

Table of Contents

Abbreviations

AOX	alcohol oxidase
AQP	aquaporin
Aqy1	yeast aquaporin 1 from *Pichia pastoris*
ar/R	Constriction region lined by aromatic residues and an arginine
Å	Ångström (= 10^{-10} m = 0.1 nm)
β-OG	beta-octyylglucopyranoside
DDM	dodecylmaltoside
ΔN36	36 N-terminal residues deleted
E. coli	*Escherichia coli*
Fps1	glycerol facilitator from S. cerevisiae
GlpF	glycerol facilitator from *E. coli*
hAQPX	human aquaporin X
IC_{50}	inhibitory concentration, where activity is reduced to 50%
IEX	ion exchange chromatography
IMAC	immobilized metal affinity chromatography
MR	molecular replacement
MD simulations	molecular dynamics simulations
NDI	nephrogenic diabetes insipidus
NPA	aspartate-proline-alanine signature motif of aquaporins
PDB	Protein Data Bank
PEG	Polyethylene glycol
PfAQP	aquaporin from *Plasmodium falciparum*
P. pastoris	*Pichia pastoris*
RU	resonance units (=0.0001°)
S107D	mutation of serine 107 to aspartate
S. cerevisiae	*Saccharomyces cerevisiae*
SEC	size exclusion chromatography
SPR	surface plasmon resonance

1 Introduction

Water is essential for life. This work contributes to the understanding of how water and other small molecules like glycerol are transported on a molecular level by membrane proteins termed "aquaporins". Their structural and functional studies aim to understand how this specific transport is achieved and suggest solutions how this transport can be measured. Ultimately, external control of water homeostasis in the human body and microorganisms can be used to discover new drugs against a multitude of diseases and disorders.

1.1 The Biological Membrane and Membrane Proteins

Differentiating between an "inside" and an "outside" is the basis for all known forms of life. This compartmentalization can be found on both a macroscopic and a microscopic scale: A human body's boundaries are mainly defined by its skin, whereas at a microscopic level this function is fulfilled by cellular membranes composed of lipid bilayers surrounding a cell. Life has evolved intricate membrane structures even inside the cell, leading to sub-compartmentalization into structures like nucleus, endoplasmic reticulum, Golgi-apparatus, mitochondria, chloroplasts and others.

What leads to this highly complex structure? Life is based on chemical reactions that consume energy. To be most efficient, high concentrations of reactants are needed in a "reaction vessel" – like a cell or a sub-compartment. The energy needed to sustain life is supplied from the outside, which is commonly achieved by the consumption of sunlight in autotrophs or via a carbon source in heterotrophs. Membranes take an essential part in storing, transporting and using this energy in a cell.
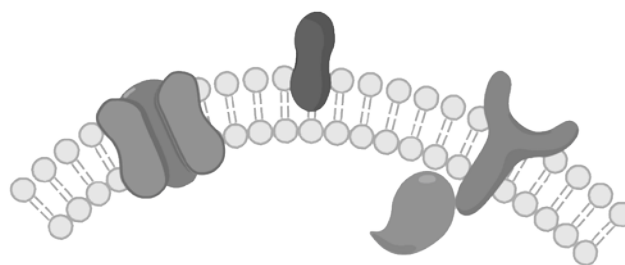
Figure 1. The biological membrane. The lipid bilayer is shown in light grey, with polar headgroups as circles that are connected by fatty acid tails. Integral or peripheral to the membrane are proteins like channels or receptors.

Biological membranes thus separate solutions of different content from each other and create gradients, as the mainly hydrophilic solutes cannot easily cross this barrier. Lipids are composed of a hydrocarbon tail and a polar head group and can form this kind of bilayer by assembling in a way that exposure of the hydrophobic parts to water is minimized (Figure 1). The hydrophobic core of the bilayer is impenetrable to most hydrophilic solutes and thus provides the properties for compartmentalization.

Even though biological membranes are often thought of and depicted as simple lipid bilayers, they can consist of up to 75% membrane proteins [1]. These can be integral or peripheral and their functions range from stabilization to signal transduction and material transport [2]. The type and amount of membrane protein can differ significantly depending on the function of the membrane. In some instances it is crucial to retain molecules, while in others it can be crucial to dispose of them, for example waste products.

1.2 Aquaporins

Water transport is essential for all forms of life. Nature has developed an elaborate system to control water flux across membranes in biological systems. Although water is naturally transported through lipid bilayers to a small extent, water transport at the diffusion limit is necessary more often than not [3].

Aquaporins are integral membrane proteins that are ubiquitous in all organisms. They form pores that facilitate selective diffusion of water across the bilayer. Aquaporins that exclusively transport water are often referred to as "orthodox aquaporins", whereas aquaporins with a wider specificity have a slightly wider channel are called "aquaglyceroporins". They transport also larger solute molecules like glycerol, ammonia or sorbitol [4]. The different terms are not always used consistently in literature – throughout this work the term "aquaporin" will be used generically, "orthodox aquaporin" and "aquaglyceroporins" will be used to specify wherever necessary.

1.2.1 History

The transport of water by biological membranes has been established in the end of the nineteenth century [5]. For a long time, this has been considered only to be caused by passive diffusion. In 1970 however, Farmer and Macy [6] observed that water transport decreases upon exposure to mercury salts. Based on these experiments, they predicted the existence of a protein water channel with a cystein residue at a crucial site. No protein of that function was known at that time, however, and their view was generally not shared within the scientific community [7]. This changed when Agre et al. discovered and isolated a 28 kDa integral membrane protein, which was abundant in erythrocytes and renal tubules [8], then called CHIP28, but today known as "aquaporin 1". Four years later in 1992 they could show that this protein is a water transporter using a burst assay based on *Xenopus laevis* oocytes [9]. The water transport rates of the isolated protein observed by Agre and Verkman in proteoliposomes were surprisingly high – up to 3×10^9 s^{-1} per monomeric unit [10, 11], which is close to the self-diffusion rate in bulk water [12]. It could also be shown that Cys189 is responsible for the inactivation of the channel upon reaction with mercury [13] – thus proving, almost 25-years later, that the prediction of Farmer and Macey was correct.

1.2.2 Structure

Topology predictions and mutational studies [14] predicted that the protein assumes a so-called "hour-glass"-fold [15]: The water pore is formed by six surrounding membrane spanning helices. Two elongated loops form short half helices, which fold back into the membrane from the intra- and extracellular sides respectively. At the tips of these two half helices, the aquaporin signature motif is located: Each of the half helices presents an asparagine-proline-alanine (NPA) sequence, which is characteristic to all aquaporins. Sometimes, this pattern is somewhat modified, in particular in aquaglyceroporins.
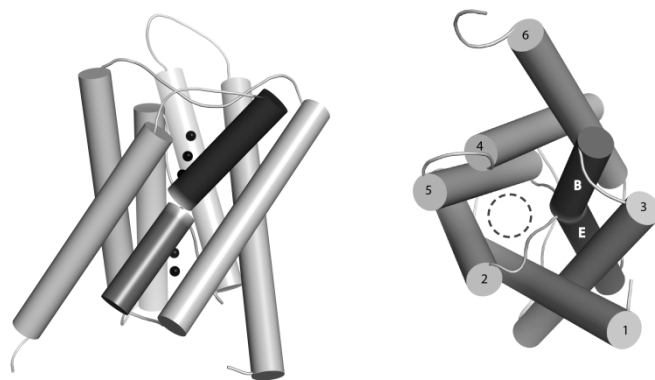
Figure 2. The aquaporin monomer. Six helices 1-6 (light grey) span the membrane. Additionally, the two elongated loops B and E fold into the membrane forming a seventh pseudo helix (dark grey). *Left:* Side view. Water molecules (black) indicate the position of the water pore. R*ight*: Intracellular view. Numbers for the helices and half-helixes are given. The water pore is indicated as a dashed circle.

The first low resolution structures were obtained via two-dimensional cryo-electron microscopy and indicated that aquaporins form tetramers in the membrane [16-19]. In 2000, Stroud et al. [20] described the first high resolution structure of an aquaporin: The glycerol facilitator GlpF from *E. coli*, an aquaglyceroporin. To date, thirteen unique aquaporin structures from all kingdoms of life have been determined to high resolution [21]: AQPZ [22, 23] and GlpF from the prokaryote *E. coli*, AQPM from the archeabacterium *Methanobacter marburgensis*, Aqy1 from the yeast *Pichia pastoris* [24], PfAQP from the protist and Malaria parasite *Plasmodium falciparum* [25], plant SoPIP2;1 from spinach [26, 27], ovine[28, 29] and bovine[30] AQP0, human [31] and bovine [32] AQP1, human [33] and rat [34, 35] AQP4 and human AQP5 [36].

All these proteins exhibit a fold common to all aquaporins and very similar to what had been predicted: six transmembrane helices and a seventh pseudo-helix, formed by two half-helices inserting into the membrane from opposite sides (Figure 2). Conserved in all aquaporins are one or two aromatic residues that in combination with an arginine encompass the pore entrance. This narrowest part of the channel is located at the extracellular side at the pore mouth. The two aquaporin signature NPA-motifs – sometimes in slightly modified form (see below) – are located at the end of each half helix and meet in the middle of the membrane channel, where its proline- and asparagine sidechains stack onto each other.

1.2.3 Transport Mechanism and Proton Exclusion

In orthodox aquaporins, the channel is just wide enough at its narrowest point to allow for a single water molecule to pass. The pore mainly consists of hydrophobic residues, but one side of the pore is lined with hydrophilic residues that can form hydrogen bonds to the water molecules. In all crystal structures where water molecules could be observed, these were located in single file through this pore [3, 37, 38].

The two filter mechanisms along the pore – the aromatic/arginine (ar/R) constriction region and the NPA-motif (Figure 3) – are responsible to exclude the transport of other substances than water, in particular ions and protons. Avoiding the latter is crucial, as decoupling the proton motive force would render the energy metabolism in any organism useless. However, none of the aquaporins studied to date conducts ions or protons at any significant rate.

The mechanism of proton exclusion is still controversial. A suggestion has been made that disruption of the Grotthus-mechanism – conduction along a water chain by proton hopping – is not the main cause for proton exclusion [39]. Instead, most simulations [37, 40-46] indicate that the water column is not entirely static in the channel. The waters rotate and a continuous hydrogen bond chain thus does not exist. The energy barrier against proton transport is highest with approximately 25 kJ/mol [43] at the NPA-region in the centre of the pore – the cause of the barrier however is still not finally determined; suggestions include dehydration effects [41], an involvement of the ar/R-constriction region [47] and the electric field generated by the two

half helices [44]. As experimental evidence for the cause of proton exclusion is difficult to obtain, the true nature of this mechanism still remains elusive.

The ar/R constriction region's main function is defining the permeability towards uncharged solutes [48]. One of the key factors is the pore diameter: In orthodox aquaporins it is usually small; therefore transport of larger solutes like glycerol is not possible resulting in the aquaporin being selective for water. In aquaglyceroporins, however, this constriction region is significantly wider. Apart from the pore width, the hydrogen bonding pattern around the ar/R-region is crucial and leads to surprising effects: most aquaglyceroporins are comparably slow water transporters, whereas the aquaglyceroporin of the malaria parasite *Plasmodium falciparum* PfAQP transports water at high rates, essentially caused by a single additional hydrogen bond [25, 49, 50].
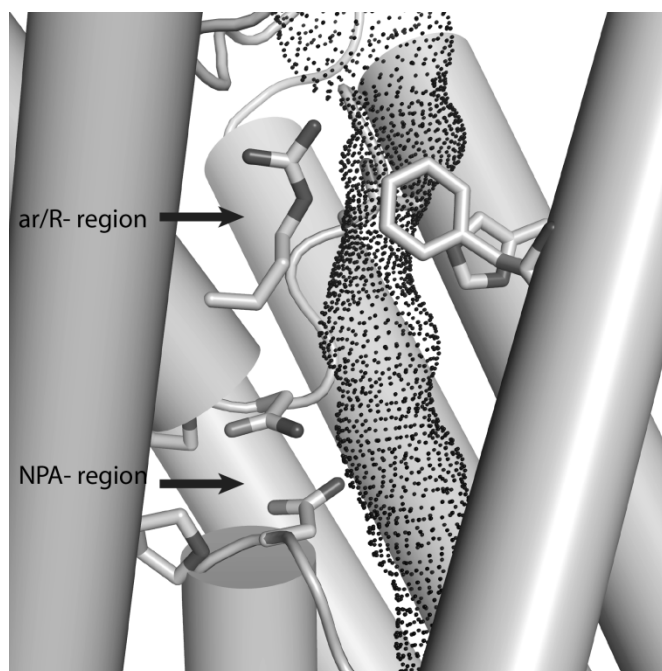
Figure 3. Key regions in the water pore. The ar/R constriction region close to the extracellular entrance of the water pore (dotted surface) creates the narrowest part of the pore (top). Further towards the cytosol, the aquaporin characteristic asparagine-proline-alanine (NPA) motif is found.

Aquaglyceroporins are known to not only transport water, glycerol and longer sugar alcohols, but also other uncharged compounds like urea, arsenate or ammonia [51]. The transport of gases like oxygen and carbondioxid has also been suggested – either through the monomeric or the central pore of the tetramer. Even though this may be possible, it presumably does not have any biological significance, as the permeation barriers of lipid bilayers to these gases are similar or even lower [52].

1.2.4 Regulation and Gating (Paper I)

Although aquaporins are beneficial for all forms of life, there are instances where water transport is undesirable. This makes it necessary to regulate the water flow across the membrane. Like most proteins, the amount of aquaporin produced can be regulated on a transcriptional gene-level [36]. However, this process is comparably slow and sudden changes in the exterior of a cell require a faster response. Two primary mechanisms achieving this goal have been suggested, namely trafficking and gating by different means.

In mammals, a common mechanism for reducing specific membrane transport activity is to remove the protein from the plasma membrane by internalizing it into storage vesicles. When higher membrane permeability is needed, the vesicles can be trafficked back to the plasma membrane quickly. The best studied example for this mechanism is AQP2. Its main function is to reabsorb water from pre-urine in the kidney and is thus essential to concentrate urine before it is excreted. Docking of vasopressin – which is a hormone regulating diuretic activity – to a membrane bound receptor [53] triggers phosphorylation of AQP2 and subsequent membrane fusion of AQP2-containing vesicles. A failure in this mechanism results in the disease nephrogenic diabetes insipidus, which is phenotypically – the excretion of large amounts of urine – identical to "conventional" diabetes, but otherwise unrelated. Analogous mechanisms

have been suggested also for other mammalian aquaporins, e.g. AQP1 [54, 55], AQP2 [56, 57], AQP5 [58-60] and AQP8 [61].

Gating is a more direct way to regulate permeability of a membrane. In contrast to the regulation at a gene level or via trafficking, it is not the amount of channels in the membrane that is altered, but their functionality. This requires the molecular structure of the protein to close the pore, either by a so-called "pinching" or "capping" mechanism [27], i.e. either by narrowing down the pore diameter by changes in the transmembrane region of the protein or by using an external part of the protein as a plug.

The gating mechanism of the spinach aquaporin SoPIP2;1 [62] has been extensively characterized. For plants, maintaining water homeostasis is particularly important, as water does not only affect their cellular function, it also has an impact on their structural integrity *via* the turgor pressure, which keeps the plant cells under tension. SoPIP2;1 has been shown to be gated by its intracellular D-loop, that in a closed confirmation inserts a leucine residue into the pore. This "capping" is triggered both by phosphorylation, the binding of divalent cations like calcium or changes in pH. All of these are known responses of a plant cell to either drought or flooding in its surrounding.

Evidence for the gating of mammalian aquaporins is still scarce. A gating mechanisms has for instance been suggested AQP0. This aquaporin is prevalent in the eye lense and is thought to be involved in mediating lens fiber cell junctions. Despite a low intrinsic water transport rate,it has been suggested to be gated [63]. A gating mechanism has also been suggested for aquaporin 4, which is found in astrocytes in the brain [64], where pH and phosphorylation have been suggested as trigger elements.

In microbial systems, the requirement for aquaporins – and in particular orthodox aquaporins – is still puzzling [65]. Single-celled organisms have a large surface to volume ratio, which should allow them to transport sufficient amounts of water through across the membrane by passive diffusion, without the help of aquaporins. Despite this reasoning, most microbes possess at least one aquaporin, which presumably is important for the organism to react rapidly to external changes, such as osmotic shock. The glycerol transport by the yeast aquaglyceroporin Fps1 has been suggested to provide a rapid response to external osmotic stress. Under normal conditions, however, the channel has to be closed to avoid the loss of solutes. Functional studies showed that Fps1 can be opened and closed by phosphorylation [66]. The orthodox aquaporin Aqy1 shows gating by its N-terminus by phosphorylation and mechanosensing and is described in more detail in Chapter 4.

1.2.5 Aquaporins in Human Health

Aquaporins are involved in a variety of disorders and diseases (Table 1) [51, 67, 68]. In principle two different scenarios arise: When exploiting aquaporins in the treatment of those conditions, the presence of the respective aquaporin can either be beneficial or not beneficial for the recovery of a patient.

As described above, hAQP2 malfunction leads to nephrogenic diabetes insipidus. This can be caused by different effects, for instance direct mutations in the aquaporin gene that render the protein dysfunctional. But also secondary effects like failure to traffic the protein to its destined location would result in the same symptoms. Depending on the cause of the defect, treatment might require advanced techniques, like gene therapy.

Conditions that originate from an undesired water transport should be easier to treat. In humans, hAQP4 is known to play a role in brain edema, i.e. the swelling of the brain after traumatic head injuries [69, 70]. Parasites like *P. falciparum* (see Chapter 1.4) also depend on aqua(glycero)porins to be able to survive and reproduce in their hosts. Blockage of water channels can be achieved by the use of conventional drugs that selectively inhibit the aquaporin in question. Currently however, most inhibitors available are not very specific, if not toxic, like mercury and silver salts [71].

Disease	Aquaporin
Congenital cataracts	AQP0
Glaucoma	AQP1, AQP4
Hereditary nephrogenic diabetes insipidus	AQP2
Chemotherapy-induced polyuric acute renal failure	
Colton-null blood antigen transfusion incompatibility	
Water retention associated with liver cirrhosis	
Water retention during pregnancy	
Brain edema (from head injuries)	AQP4
Seizures	
Brain tumors	
Sjögren's syndrome (dry eyes and mouth)	AQP5
Hyperinsulinemia	AQP7
Malaria	PfAQP

Table 1. Selection of conditions potentially treatable using aquaporins [67]. Aquaporin numbering refers to mammalian aquaporins, except for PfAQP, which is from *Plasmodium falciparum.*

1.3 The Yeast *Pichia pastoris*

Budding yeasts are a single-celled eukaryotic organism. *Saccharomyces cerevisiae* – more commonly known as baker's or brewer's yeast – has been cultured by humankind for millennia. It is probably the best studied eukaryotic organism and serves as a model for higher eukaryotes including animals and humans, in particular with respect to genetics and cell biology. Its 12 Mbp genome was the first to be sequenced in 1996 [72, 73].

While most yeast species are benign, some can also act as pathogens, often caused by excessive growth on immuno-compromised patients. *Candida albicans*, for example, lives on the skin of most people, but can also be the cause candidiasis, which symptoms range from rash to the infection of organs which in rare cases can even lead to death[74, 75].

Pichia pastoris is closely related to *S. cerevisiae*. Discovered in 1914 near Lyon [76], it has the unusual ability to consume methanol as a carbon source using a specialized alcohol oxidase (AOX) enzyme [77]. Initially patented by the Phillips Petroleum Company [78] (Bartesville, OK, USA) during the 1970's as a way of converting methanol – which could cheaply be produced from methane at that time – to biomass for high-protein animal feed, the patent was abandoned due to the heavily increasing prices during the oil crisis.

Later on, P. pastoris was developed into a host for recombinant protein production. Even though *E. coli* is still the most widespread organism for recombinant protein expression, *P. pastoris* has become increasingly popular, in particular for structural studies of eukaryotic proteins, where comparably large amounts of protein are needed. *E. coli* as a prokaryotic organism is unable to perform posttranslational modification like folding, glycosylation and protein secretion. Compared to cell culture systems from higher eukaryotes however, *P. pastoris* is comparably cheap and easy to handle.

Even for *P. pastoris*, the metabolization of methanol is a rare event – which might explain, why its alcohol oxidase [79] has a comparably low turnover number for converting methanol to formaldehyde. To compensate for this, expression of the enzyme is regulated on the gene level by a very strong promoter. Inserting the gene of interest in between this promoter and AOX results in exceptional overproduction upon induction with methanol. In a commercially available expression kit (Invitrogen), the gene of interest is inserted using homologous recombination, which makes the clone significantly more stable and the culture does not have to be kept under constant selective pressure, as opposed to plasmid based techniques.

This system involving the AOX1 promotor for overproduction of proteins and polypeptides has been patented in 1989 [80]. The system has been continually improved by the development of novel Pichia strains [81-83] and the discovery of new promoters [84]. Today, the system is particularly popular in academic settings, but is also used for commercial drug design and the industrial large scale production of proteins [85].

1.4 Malaria

Even today, Malaria is a widespread disease which is prevalent mainly in the tropical and subtropical regions of the earth. 300-500 million people carry the disease and 1-3 million die from it every year – mainly in the lesser developed world [25].

Malaria is caused by single-celled protozoan organisms from the genus *Plasmodium*, discovered by Alphonse Leveran in 1881 [86]. There are over 200 different *Plasmodium* species which often are specific to a particular host. The most common species to infect humans are known to be *Plasmodium falciparum* – predominant in Africa – and *Plasmodium vivax*, which is most common in Asia. The former is the most deadly of the parasites, whereas *P. vivax* can lie dormant for years [87]. Symptoms of the disease can vary from simple fever in uncomplicated cases to severe breathing difficulties, anemia, coma and ultimately death by depletion of the host's hemoglobin supply [88].
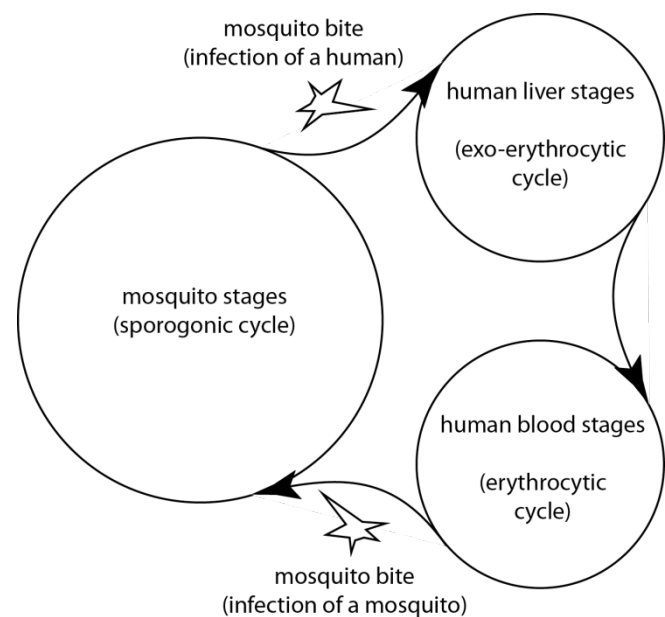
Figure 4. Life cycle of the malaria parasites of the genus *Plasmodium*. The parasite infects humans by the bite of the *Anopheles* mosquito. There, it first infects the liver, where it can lie dormant. An outbreak of the disease is characterized by the parasite proliferating at increased rates in red blood cells (human blood stages). Here, gametocytes are formed, that can infect a new mosquito.

Infection with Malaria occurs via a sting from a female *Anopheles* mosquito. The life cycle of the parasite is complex and involves a so-called "blood stage" in the vertebrate host and a

"mosquito stage" (Figure 4). During the blood stage, the parasite first infects the host's liver cells (exo-erythrotic cycle) to multiply and then enters the erythrocytic cycle, where parasite reproduction takes place in red blood cells. From here, the parasite can be taken up anew by the *Anopheles* mosquito during a blood meal [86].

Currently, there is no approved vaccine against malaria [89]. The two most widely used anti-malarial drugs are chloroquine and sulphadoxinepyrimethamine (Fansidar, Roche) [90]. These are affordable in developing countries, but are losing their efficiency due to emerging resistance in *Plasmodium* strains. As an alternative approach, the *Anopheles* mosquito has been targeted to prevent transmission of the disease using insecticides like DDT [91], with all its negative consequences on the environment. Thus, new cheap and effective drugs against malaria are urgently needed.

1.5   Scope of the Thesis

The aim of this thesis has been to elucidate the transport and gating mechanism of eukaryotic aquaporins (Paper I) by structural and functional studies. Also, the toolbox for functional analysis of aquaporins has been improved by the development and application of novel functional assays.

The project started with the discovery of the novel aquaporin "Aqy1" from the overproduction host *Pichia pastoris* (Paper II). We could determine its three-dimensional structure to a resolution of 1.15Å, to our knowledge the highest resolution achieved for a membrane protein at that time. Based on these structural findings, we conducted further functional investigations and mutational studies, where the functional studies have mainly been performed by a newly developed method based on the shrinkage of *Pichia pastoris* spheroplasts (Chapter 3.7.3).

These crystals were later optimized in size and diffraction quality, diffracting as low as 0.90 Å and thereby advancing membrane protein crystallography into the realm of ultra-high resolution (Paper III). We also addressed a common criticism of crystallographic studies – namely that data collection typically is performed at cryogenic temperature, which might cause unnatural artefacts. Exploiting the size and stability of our crystals, we also collected data at room temperature, yielding new insights into the influence of temperature on the structure and the mechanism of gating (Paper IV).

Protein structures are a great aid for forming hypotheses for biochemical mechanisms, but also have to be confirmed by functional studies. These structures can also be used in a structure based drug design approach to determine candidate inhibitors by virtual high through-put screening. The inhibitory properties of these compounds have to be tested experimentally, however, and no techniques for in vitro medium- or high-throughput screening had been available.

We designed a novel system (Chapter 5) to measure solute transport across membranes in collaboration with Höök et al. (Chalmers Technical University, Sweden). Unlike other methods, this approach is not based on measuring size change of vesicles and takes advantage of surface plasmon resonance biosensor technology. Transport was observed through aquaglyceroporins on the example of the clinically relevant *P. falciparum* aquaporin PfAQP (Paper V).

This method has been developed further into a medium through-put inhibitor screen. Fishwick et al. (Univ. of Leeds, UK) provided us with two sets of potential PfAQP inhibitors that were determined *in silico*, using a rational drug design approach based on the recently determined crystal structure of PfAQP [25]. Screening these substances yielded a compound that successfully inhibited PfAQP with an IC_{50}-value of 1-5 μM (Paper VI).

2   General Methodological Considerations

2.1   Production and Purification of Membrane Proteins in *Pichia pastoris*

2.1.1   Overproduction and Cloning

Producing functional membrane protein in sufficient amounts for structural studies is one of the bottlenecks in the determination of membrane protein structures. Working with eukaryotic proteins in particular puts special requirements on the expression system. Unlike prokaryotic proteins, they undergo various post-translational modifications in the ER and major parts of the machinery of the otherwise commonly used *E. Coli* system are often not compatible.

Since one of the proteins – Aqy1 (Chapter 4) – is endogenous to *P. pastoris*, it was the obvious choice as overexpression host. The other protein used in this study – PfAQP – had been demonstrated earlier to also express well in this system [92].

P. pastoris can either be cultured in shaker flasks or in bioreactors. The latter provide accurate control over crucial parameters like pH, temperature, feed rate and aeration. When grown in bioreactors for five days, a typical culture of *P. pastoris* yields around 250 g of wet cells per litre culture, more than six-fold of what can be achieved by cultivation in traditional shaker flasks. Also, the expression levels of the recombinantly produced protein tend to be higher. The limited aeration has been identified as one of the main factors to influence these large differences [85].

Compared to *E. coli*, the growth process is more time consuming which is caused by the longer generation time of *P. pastoris* (approx. 20 min vs. 2 hours). Cloning and modifications – like adding tags or point mutations – of the gene of interest are usually first performed in *E. coli* using the shuttle vector PICZB (Invitrogen). The vector is then linearized and transformed into *P. pastoris*, where it integrates into the genome by homologous recombination. Finally, expression screens are performed to select a clone which produces the protein to good levels. The fact that this procedure is comparably lengthy is outweighed by a resulting clone that is easy to handle, does not require growth on selective media other than methanol and shows good and constant protein expression levels.

2.1.2   Purification

Highly pure membrane protein is considered paramount for successful crystallisation trials [93]. Separating the cell membrane – which is the matrix containing membrane proteins – removes most of the soluble proteins. It requires a way to disrupt the cells, for instance using a French Press or an X-Press [94] system. Although the protein is comparably stable while still in the membrane, heating of the solution at this energy intense step is best avoided to optimize the yield of functional protein. The cell debris is removed by differential centrifugation and subsequent ultracentrifugation to pellet the cell membrane. This results in isolated membranes still containing all membrane proteins. Optionally, this fraction can be washed with a dilute NaOH- or a concentrated urea solution to remove peripherally bound proteins.
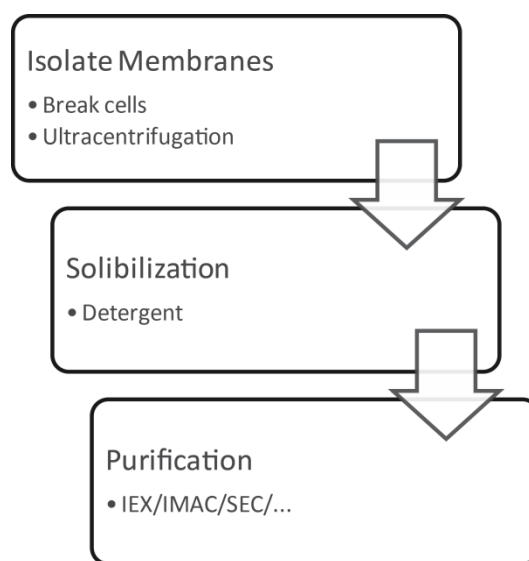

Figure 5. Flowchart for the purification of membrane proteins.

To extract the protein from the membrane, amphiphiles are added (Figure 5). They dissolve most of the membrane and its proteins and form mixed detergent/lipid/protein micelles. Depending on the choice of detergent, not all proteins are solubilized, which makes solubilization also an (often underestimated) purification step.

Solubilized membrane protein can generally be handled as if they were soluble proteins. All buffers have to contain the respective detergent above their CMC – typically a concentration twice the CMC is used – though, which can be a significant cost factor. Alkylglucosides with comparably high CMCs, e.g. OG or NG, and alkylmaltosides like DDM work reasonably well for many membrane proteins and in particular aquaporins [93].

Most contaminants, lipids and undesired protein, can be removed by a subsequent purification step like affinity- or ion exchange chromatography. Sometimes, the protein of interest interacts strongly with one of the contaminants – namely lipids have been shown to remain bound to proteins throughout the purification and crystallization process and are thought to play an important role for the proteins function in some cases [36, 95].

Finally, size exclusion chromatography is used as a polishing step. It allows assessment of the size, homogeneity and multimeric state of the protein. Here, buffer exchange is also easily possible without major changes to the purification protocol, so adjustments can be made to the final buffer and detergent used for crystallization. The aim is to obtain protein as pure as possible after this step (>99%); in many cases this cannot be achieved, however, because of protein degradation.

2.2 Crystallography

Structural Biology is based on observing biological macromolecules – typically proteins or nucleic acids. Traditional methods using visible light in combination with a light microscope have even in an optimal case a physical resolution limit of $\lambda/2$, half of the wavelength used. Thus, the smallest structural features that can be observed using visible light with wavelengths of 400-700 nm are around 200 nm. Atom diameters and atomic bond lengths are typically in the order of magnitude of 1-3 Å, i.e. 0.1 – 0.5 nm.

To obtain this kind of resolution, the use of electromagnetic radiation with a very short wavelength called X-rays is necessary. This radiation with wavelengths between $10^{-8} – 10^{-12}$ m was discovered by Carl Gustav Röntgen in 1895 and was shortly thereafter used by Max von Laue to prove the regular atomic structure of salt crystals [96]. Today crystallography is an indispensable analytic technique in most natural sciences: ranging from the classical field of mineralogy to material sciences, physics, chemistry and structural biology.

X-rays, however, cannot be used to magnify a sample in the same way as visible light in a microscope: this requires lenses that are able to focus the electromagnetic radiation onto a focal point. Visible light can be bent by lenses made of glass or of the transparent tissue in the eye. Even though there are mirrors and lenses available [97] for X-rays, these are not sufficiently perfect to allow direct imaging at atomic resolution. Moreover, as X-rays are high energy radiation, they have a destructive effect on most materials, and in particular biological samples. This makes the use of crystals indispensable, as a large number of molecules arranged in regular fashion produce an interference pattern upon irradiation and counteract radiation damage.

2.2.1 Theory of X-ray Diffraction

When three dimensional crystals are irradiated with X-rays, diffraction is observed. The radiation is scattered by the electrons in the crystal and interference occurs due to the regular assembly of the crystal lattice. Thus, most waves cancel out each other due to destructive interference and reflections can only be observed if the reflection condition, also known as Bragg's law, is fulfilled:

$$n\lambda = 2d\sin\theta \qquad \text{(Eq. 1)}$$

where n is the order of diffraction, λ is the wavelength of the radiation used, θ is the angle between the incoming beam and the lattice plane and d is the distance between the lattice planes (Figure 6, *left*). This in turn means that the reflection pattern is independent of the content of a unit cell but only depends on the lattice, that is to say the spacegroup and cell dimensions. During the diffraction experiment (Chapter 3.5), the crystal is rotated to cover most parts of the reciprocal space (Figure 6, *right*).

The measured intensities contain information about the content of the unit cell, namely the electron density and thus atom positions. The information that can be obtained by a diffraction experiment is incomplete, however. The electron density $\rho(x,y,z)$ at a position x,y,z in the unit cell is fully described by the Fourier transformation of the structure factor

$$\rho(x,y,z) = \frac{1}{V}\sum_h \sum_k \sum_l |F_{hkl}|\, e^{-2\pi i(hx+ky+lz)+i\alpha(hkl)} \qquad \text{(Eq. 2)}$$

with V being the volume of the unit cell, h, k and l being the Miller indices, $|F_{hkl}|$ the absolute value of structure factor for a respective reflection hkl. Experimentally, only intensities – which are proportional to $|F_{hkl}|^2$ – can be measured, i.e. no phase information of this complex number can be obtained. This is generally referred to as the "phase problem" of crystallography.
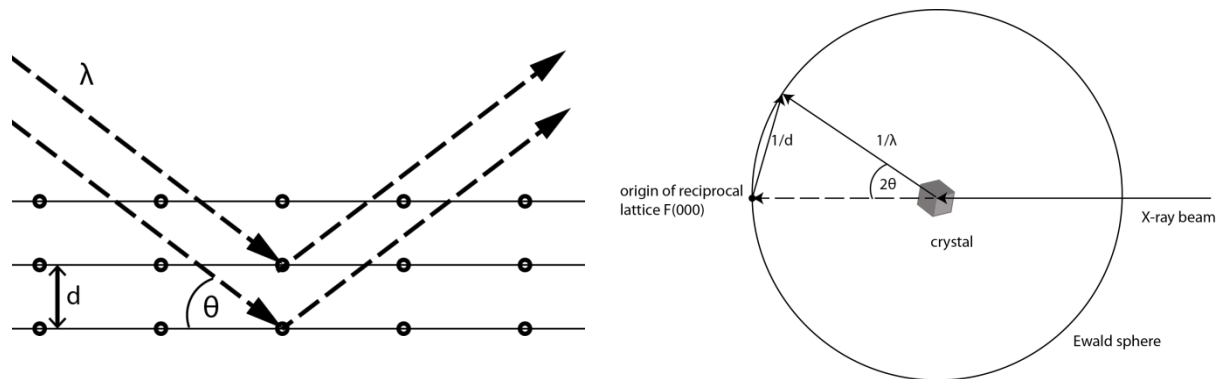

Figure 6. *Left:* Bragg diffraction. The X-ray beam of the wavelength λ is reflected at the lattice plane. *Right:* The Ewald sphere, indicating which reflections of the reciprocal lattice fulfil the reflection condition. The origin of the reciprocal lattice is on the Ewald-sphere in line with the incident beam and can thus never be observed. Reflections occur, when points of the reciprocal lattice meet the sphere as the crystal in the center of the sphere is rotated.

2.2.2 Protein Crystal Growth

By many, growing protein crystals is considered more of an art than an exact science. Set aside the problem of protein production (Chapter 3.1), determining the growth conditions for a crystal in a rational way is virtually impossible.
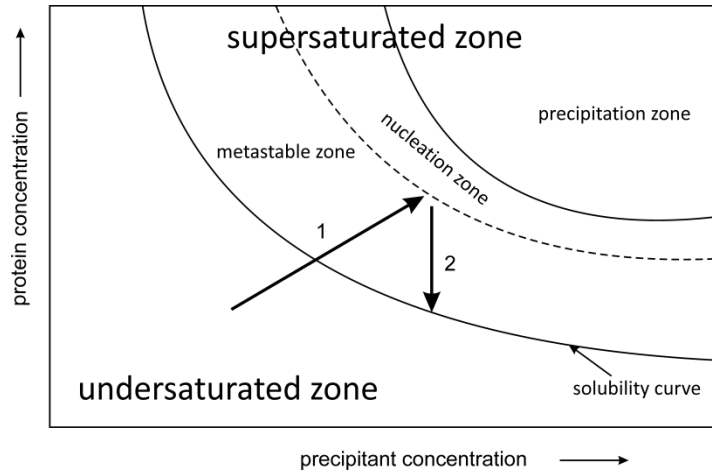

Figure 7. The phase diagram. For successful crystal growth in a vapour diffusion experiment,, the mixture of precipitant and protein solution should fall into the undersaturated zone of the phase diagram. As the drop slowly loses water to the reservoir solution, the concentrations of protein and precipitant increase (1). Once the nucleation zone is reached, crystal growth occurs, decreasing the concentration of protein in the solution (2). After N. Asherie [98].

The basic principle is built on the idea to first have the protein dissolved and then alter the composition of this solution in a way that it precipitates. If precipitation occurs fast, one usually observes amorphous precipitate, often only referred to as unwanted "precipitate", which does not contain any internal order and is thus unsuitable for crystallographic purposes. As described in the phase diagram (Figure 7), the protein can either be completely dissolved ("undersaturation"), in a metastable zone or supersaturated. To obtain crystals, initial nuclei have to be formed in the supersaturated nucleation zone which then can gain volume (Figure 8) in the metastable zone, where no new nucleation occurs.

As straightforward as this basic concept may sound – the number of different factors influencing crystal growth is vast, and even minor deviations may lead to entirely different outcomes of an experiment. Typically, pH, temperature, concentrations and volumes of all solutions used are tightly controlled and yet reproducibility is often not easy or sometimes not possible at all.
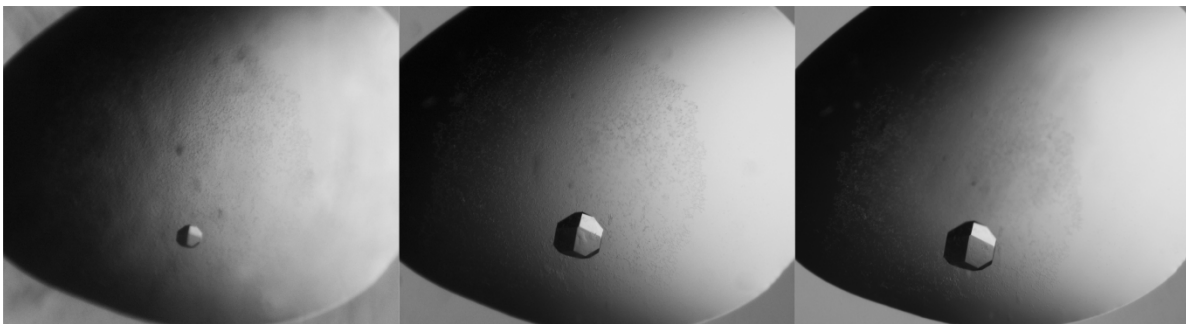

Figure 8. Growth of Aqy1 crystals over time. Shown is a hanging drop experiment where 1 μL protein solution (10 mM HEPES pH=7.5, 100 mM NaCl) was mixed with 1 uL of precipitant solution (26% PEG600, 100 mM Tris (pH = 8.0), 100 mM CaCl$_2$). Pictures were taken 5 (*left*), 7 (*center*) and 10 (*right*) days after experiment start. Final crystal size was approximately. 0.15 x 0.15 x 0.15 mm^3.

There are a number of techniques to grow protein crystals, most commonly based on vapour diffusion. A highly pure protein solution is mixed with a precipitant solution, so the protein just remains soluble. This drop is then brought into an air tight container – typically as a hanging or

sitting drop (Figure 9) – which also contains a reservoir solution separated from the drop. For the ease of handling, the precipitant solution is often identical to the reservoir solution. As the osmolarity in the precipitant solution is higher than in the protein drop, water evaporates over time from the protein drop and moves into the precipitant solution. Crystal growth – if at all – usually occurs within days to weeks, but can be significantly faster or slower as well.

Commercial screens are available which in combination with robotic devices often are used during the initial stages of screening chemical space. Once a suitable condition – e.g. one containing crystal leads – has been found it is refined by changing pH, salt concentrations or adding other substances, aiming for crystals of at least 0.1 x 0.1 x 0.1 mm^3 in size and of sufficient diffraction quality.
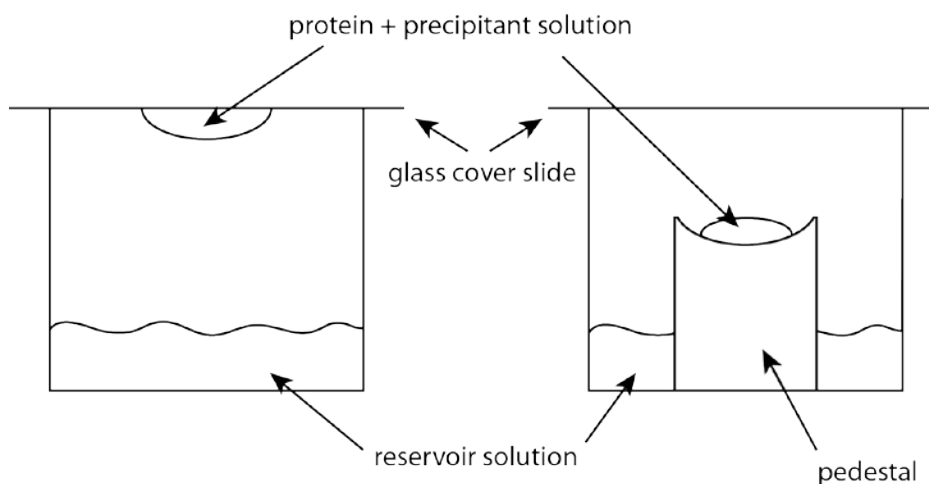

Figure 9. The vapour diffusion experiment. A mixture of protein and a precipitant solution is sealed in a container (typically the well of a plastic microplate) together with a reservoir solution. In a hanging drop arrangement (*left*), the protein drop is turned upside-down and hold in place by the surface tension to a glass cover slide. The sitting drop setup (*right*) uses a pedestal to separate the protein from the reservoir solution.

2.2.3   Crystal Growth of Membrane Proteins

Membrane proteins that are solubilized by a detergent would be expected to crystallize as soluble proteins do. And indeed, most of the standard techniques described above can be and are used frequently. The quality of the crystals obtained – if obtained at all – is usually significantly worse than of comparable soluble proteins. While the first structure of a soluble protein (myoglobin) was determined in 1958 [99], it took until 1984 [100] before the first membrane protein structure could be elucidated. This head start is still visible: To date, more than 60,000 structures of soluble proteins have been determined, while membrane protein crystallography celebrated its 1000's structure in the year 2010 (Figure 10).

The reason for this lies in the amphiphilic nature of membrane proteins. While the intra- and extra-cellular domains are hydrophilic, the transmembrane part, which interacts with the lipid bilayer *in vivo*, is hydrophobic on its surface. Detergent molecules act as a replacement for the bilayer in the crystal by forming a micelle-like shape around the hydrophobic parts.

The size of this micelle is crucial for the crystallization of membrane proteins: too small a micelle and the protein will not be soluble enough to conduct a crystallization experiment. On the other hand, if the micelle is too large, crystal contacts between the single molecules that build up the stable crystal lattice are not possible, as they are shielded by detergent molecules (Figure 11). Thus finding the optimal detergent – the most common ones being LDAO, DDM and OG – is often the key to obtain membrane protein crystals.
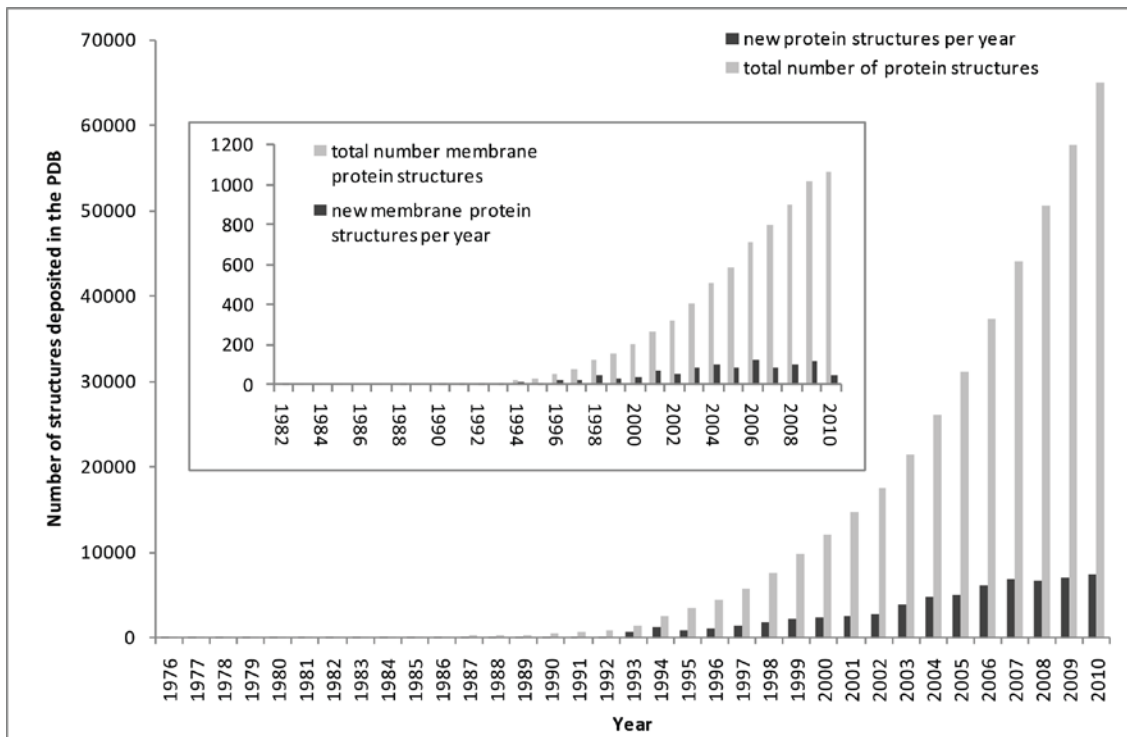
Figure 10. Protein structures deposited the Protein Data Bank. Total number of protein structures determined is shown in light grey, the number of new structures added in the respective year in black. The inset shows the development for membrane proteins. Data for all proteins was obtained from the Protein Data Bank (www.pdb.org), the data for membrane proteins from the Membrane Protein Data Bank (www.mpdb.ul.ie).
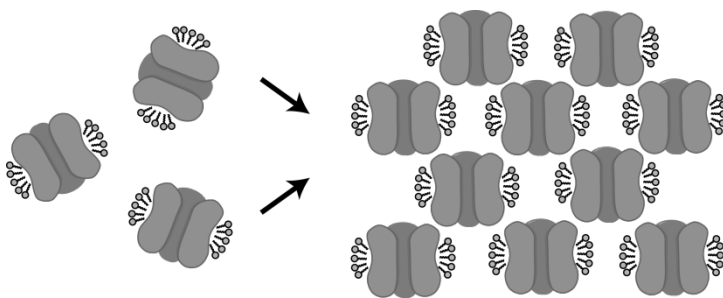

Figure 11.   Crystal formation of membrane proteins. Hydrophilic parts of the protein – here depicted as a channel protein – are covered by detergent molecules, both in solution (*left*) and in the crystalline form (*right*). Crystal contacts are mainly formed between the exposed hydrophilic parts of the protein.

There have been efforts to find new ways towards the crystallization of membrane proteins. Two-dimensional crystals which can be used for electron microscopy can be formed by adding lipids into the solution and removing the detergent by dialysis. In these cases frequently the two-dimensional arrangement of the molecules resembles the surrounding of the protein found in Nature. Other techniques to yield three-dimensional crystals have been developed around similar assumptions. In lipidic cubic phase [101], sponge phase [102] and bicelle [103] crystallization techniques, the protein is first reconstituted into a two dimensional bilayer that allows the protein to remain stable while still exposing its hydrophilic domains for crystal contacts.

All of these techniques still have their limitations, and it is still a challenge to crystallize a membrane protein. If none of the above techniques succeed, it is common practice to either only crystallize parts – such as the soluble domains – of a protein or to modify the protein genetically to maximize the crystal contacts and/or protein stability. Using any of these approaches, great care has to be taken not to affect the proteins properties that are to be studied.

2.2.4 The Diffraction Experiment

The diffraction experiments require a strong source of X-rays. If large and well diffracting crystals can be obtained, sealed-tube or rotating anode sources can be used to collect diffraction data; this kind of experiment can be performed in-house. For more difficult cases, for instance large complexes, membrane proteins or small crystals, synchrotrons are necessary as a stronger source of X-rays.

Synchrotrons are large, often multi-national facilities, where a relativistic electron beam is accelerated and kept in a storage ring, several hundred meters in circumference. Strong electromagnetic radiation is emitted, while the beam transverses "bending magnets" which keep the beam on a circular path. Even stronger, more intense, brilliant and monochromatic radiation can be produced using "undulators" or "insertion devices" as can be found in third generation synchrotrons, which bend the electron beam multiple times within a short distance.

For data collection, the crystal is mounted on a diffractometer, where it is aligned to be hit by an X-ray beam, typically around 100 μm in diameter (Figure 12). Specialized microfocus beamlines allow beamsizes below 5 μm (Figure 13) [104, 105], whereas older sources often operate with larger beams [106]. The crystal is rotated while being exposed to the radiation. The diffraction pattern is collected using a detector, most commonly a CCD-detector, which allows integration of spot intensities. To avoid overlapping of the diffraction spots from the three dimensional lattice on the two dimensional detector, an image is collected after only a small – 0.1 – 2 degrees – rotation of the crystal. The latest generation of detectors are so-called single-pixel detectors (PILATUS, [107]), which can read out up to several hundred frames per second, thus increasing the speed of data collection significantly.
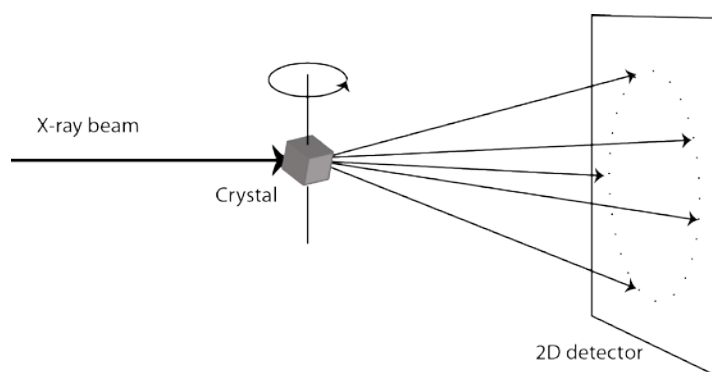

Figure 12. Schematics of a diffraction experiment. X-ray diffraction is produced by a crystal in the path of an incident beam. The crystal is rotated and the diffraction images are collected in regular intervals.
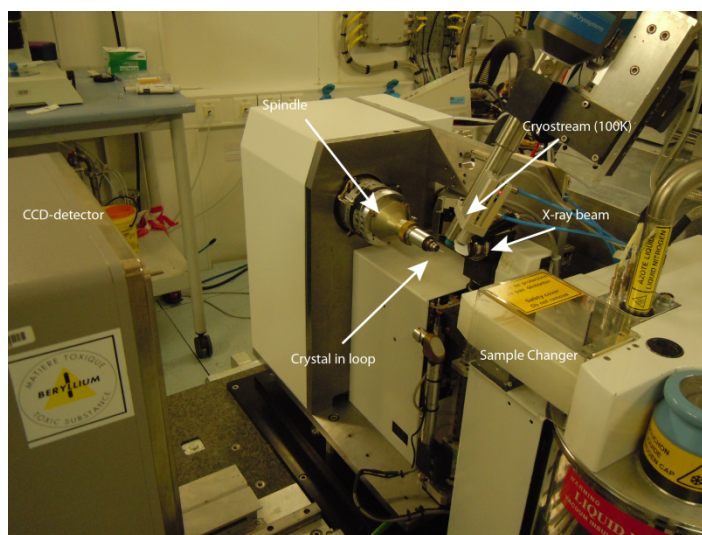

Figure 13. Beamline setup at the microfocus beamline ID23-2, ESRF, France.

2.2.5 Data Collection and Radiation Damage

Collecting good data from a protein crystal is a trade-off between multiple parameters. A good dataset should have a good data/parameter ratio – i.e. extend to high resolution, have a good signal/noise-ratio, high redundancy, low mosaicity and high completeness. On the other hand, there is damage introduced by the radiation dose the crystal is exposed to. This leads to the creation of free radicals in the crystal which leads to its decay and crystal disorder.

Cryocrystallography

This problem can be reduced by cooling the crystal, e.g. in liquid nitrogen (77K), as the mobility of the radicals is drastically reduced. However, when water freezes in its crystalline form, it expands and ruptures the protein crystal. This problem can be avoided by freezing the crystal quickly, so that the water is caught in an amorphous, glass-like state. To achieve the transition to amorphous ice, the crystal would have to be frozen below 155 K at cooling times below 10^{-5} s [108, 109], which is not possible in practice. By introducing so-called cryo-protectants, e.g. short polyethylene glycols or glycerol, into the crystal, the cooling speed required for vitrification can be reduced to 1-2 seconds.

In practice, these substances are either directly present in the crystallization condition or are introduced later on by soaking. Optimizing cryo-conditions can be a time-consuming part of crystal optimization; if possible, moving to crystallization conditions that are cryo-compatible is preferred over soaking attempts, as the osmotic shock introduced can often be more damaging than helpful. As their heat capacity is larger, it is particularly important to find good cryoprotection conditions for large crystals [110].

Despite the best attempts, freezing a crystal increases its mosaicity. Even though adding cryo-protectants to the crystal water increases the density to better match the density at room temperature, some freezing defects will always appear, reducing the order in the crystal.

Data Collection at Room Temperature

Originally, diffraction data were collected at room temperature by sealing a crystal in a glass capillary. As biological crystals consist of a significant fraction of water – which is essential to satisfy hydrogen bonds and to fill gaps in the crystal lattice – a small amount of mother liquor has to be added to the capillary before sealing it to keep the crystal from drying out. When exposed to an X-ray beam – in particular from a second or third generation synchrotron source – the crystal suffers significant radiation damage as some of the radiation dose is absorbed and causes the creation of free radicals. This limits the amount of data that can be collected from a single crystal, and in the past it was common practice to collect data from multiple crystals that had to be merged afterwards – introducing an additional error, as two crystals are never exactly the same. There are other approaches to minimize the damage at room temperature, for example the use of scavenger molecules that neutralize the radicals created by the radiation, but their use is limited [111]. Even today – collecting a dataset at ambient temperature typically requires the use of large crystals.

Nevertheless, the technique is still sometimes necessary. To evaluate the quality of initial crystals, diffraction at room temperature is a convenient way to quickly obtain information on the native crystal, as flash-cooling requires optimization. To minimize physical damage introduced by crystal handling, this can even be performed directly in crystallization plates, both at synchrotrons and in-house diffractometers.

Another area where room temperature measurements can be useful is when the dynamics of a protein shall be studied. Wöhri et al. demonstrated in 2010 that light-induced structural changes of light sensitive membrane proteins can be followed using Laue-diffraction [112]. Here, the

non-frozen state of the crystal is necessary to permit this motion. Moreover, low crystal mosaicity is crucial for Laue data to minimize overlaps between diffraction spots. An alternative approach comprises the comparison of structures obtained at different temperatures, which can contain valuable information on protein dynamics and stability (Chapter 4.5).

A new development is that X-ray data can be combined with neutron diffraction data to increase accuracy [113]. Neutron diffraction is essentially a Laue technique and gives more insight into the position of hydrogen atoms in a structure and therefore often unique insight into a protein's mechanism. As cell parameters and atom location can vary slightly in crystals at different temperatures, all data have to be obtained at the same, i.e. ambient, temperature.

2.2.6 Structure Determination

In practice, X-ray diffraction data is first indexed and the space group determined. Then the spot intensities are integrated and saved into a file where they are associated to their Miller indices. Software packages referred to as "integration" or "data processing" software like XDS [114] and MOSFLM [115] are used to do this in an automated way.

The data is then scaled with programs like SCALA [116] or XDS, i.e. consideration is taken for factors like detector sensitivity, crystal absorption and decay of the crystal over time. These programs also take care of data merging, which is performed depending on the symmetry of the Laue-group to provide more accurate data.

2.2.6.1 *Molecular replacement*

As described in Chapter 3.3, the data obtained by X-ray diffraction is lacking phase information and can thus not be used to calculate the electron density directly. Several ways have been developed to obtain initial phases, which can later on be refined iteratively using the molecular model. For *de novo* phasing, experimental phasing methods like MAD and SAD (multiple and single anomalous diffraction) can be used, together with the more traditional methods of MIR and SIR (multiple/single isomorphous replacement) [117]. If data could be obtained to high resolution (<1.2Å) [118], direct methods can be employed to solve the phase problem; latest developments allow *ab initio* phasing with data up to 2 Å (ARCIMBOLDO [119]).

Most protein structures determined however are similar to a structure that has been solved earlier. The Patterson function

$$P_{uvw} = \frac{1}{V^2} \sum_{hkl} F_{hkl}^2 \cos\left[2\pi(hu + kv + lw)\right] \qquad \text{(Eq. 3)}$$

is an alternative way to express Eq. 1, where only the intensities ($I \sim F^2$), which are obtained during native data collection, are used. Not using the phase information leads to a set of interatomic vectors, i.e. the loss of the point of origin and the orientation of the molecule in the unit cell. If a structural fragment – usually a structure similar to the one to be solved – is known, there is a correlation between the Patterson-maps of the model and the target structure. By performing rotation and translation, the new molecule can be placed in the unit cell, from which initial phases can be obtained.

The required degree of similarity between the proteins varies a lot, but generally 25-40 % homology is required. Programs commonly solving structures by molecular replacement (MR) comprise e.g. PHASER [120] and Molrep [121]. Newer developments like BALBES [122] also provide an integrated MR pipeline with a comprehensive database based on the Protein Data Bank to facilitate model search.

2.2.6.2 Structure Validation and Refinement

The initial phases obtained by the various methods described above have to be extended. An electron density map can be calculated to from the preliminary model and the diffraction data by Fourier transformation. Into this density, a model can be built, either by hand using programs like COOT [123] or by automated methods as provided by Buccaneer [124] or ARP/wARP [125] if the resolution is sufficient, typically better than 2.5 Å.

The obtained model will then be refined by iteration between manual real space refinement and automated refinement against the reflection intensities (e.g. SHELXL [126], REFMAC [127], Phenix [128]) until convergence. The agreement of the model with the data is assessed with help of the crystallographic R-factor:

$$R = \frac{\sum_{hkl}||F_{obs}(hkl)| - |F_{cal}(hkl)||}{\sum_{hkl}|F_{obs}(hkl)|} \quad \text{(Eq. 4)}$$

As refinement progresses, the experimental structure factor F_{obs} and the model structure factor F_{cal} should ideally become identical, and the R-value thus approach 0. In practice, a value "maximum resolution divided by 10" is generally considered to be reasonable, that is to say 20% for a structure at 2.0 Å resolution. If parts of the structure cannot be modeled appropriately, for instance as they are flexible or disordered, the R-factor will increase. Additionally, cross-validation with an R_{free}-factor is commonly performed. Usually 5% of the data is excluded from the refinement and used to calculate R_{free}. These two values should be in good accordance with each other, large discrepancies can implicate a poor structural model, often caused by over-interpretation ("overfitting") of the data.

To exclude that the phases, particularly if the initial phases were obtained from MR, are biased towards the used model, parts of the model can omitted and the map recalculated. From the resulting omit map – which should also show the electron density of the parts of the model not used for calculation – the quality of the MR solution can be estimated. This can also be performed in an automated way (e.g. by using the CNS program suite [129]) by calculating a "composite omit map", which omits a part of the structure at any one time and combines the resulting maps to give an unbiased picture of the electron density.

Proteins are built up by the same building blocks, that is to say amino acids. Their individual structures vary very little with regards to bond lengths, and also their connection angles Φ and Ψ have been observed to be restricted to a range represented in the Ramachandran plot [130]. These restraints can both be used during the refinement as well as the final validation of the model.

2.3 Functional Assays

2.3.1 Liposomes

The lipid bilayer of cells is the natural environment of all membrane protein, including aquaporins. It is an important part of all those proteins, as it provides stabilization from the lateral, hydrophobic sides, without which the protein would aggregate.

Most functional assays for channels and transporters make use of artificial liposomes to mimic these surroundings. They often consist of either artificial phospholipids, like POPC, or of lipid extracts from natural sources like e.g. *E. coli*, which can make up lipid bilayers. When resuspended in aqueous buffers, multilamellar vesicles – onion-like shapes with several concentric lipid bilayers – are quickly and spontaneously formed. These provide sufficient

stabilization for the protein, but are of no use for their functional characterization, as their size distribution is usually not homogeneous and their interior is already filled with lipid.

Unilamellar vesicles are thermodynamically stable , and can thus be produced by disrupting multilamellar vesicles, e.g. by sonication [131], extrusion through filters [132], freeze-thawing [133] or the evaporation of organic solvents [134]. This kind of vesicles is what is needed for successful functional transport assays; their quality and size can be characterized by dynamic light scattering. The methods described above are often not compatible with proteins, which are fragile and instable towards these treatments. Thus, unilamellar vesicles are generally produced first and the protein is reconstituted into the vesicle in a separate subsequent step.

Although also the reconstitution process is thermodynamically favourable, it is kinetically hindered, as the lipid bilayer is tightly packed and thus does not allow the protein to enter the membrane. By adding detergent – β-octylglucopyranoside has been shown to be effective and is compatible with most aquaporins – the bilayer can be destabilized and added protein can integrate into the membrane. It is crucial that the detergent is removed after the reconstitution process, as destabilized vesicles are leaky and thus not suited for transport studies due to an increased background signal. Detergents with high CMCs can be quickly removed by dilution of the sample; the proteoliposomes can then be collected via ultracentrifugation. Other means of detergent removal, like dialysis [135], are possible, but more time consuming and are hence often used with low CMC detergents.

2.3.2   Stopped-Flow Light Scattering

Water transport through aquaporins is a fast event. Its detection is particularly difficult, as all biological systems to a large extent consist of water and observing water transport directly with the sufficient time resolution is limited to very few recently developed and still experimental techniques [136]. Determining water transport across a vesicle bilayer containing the protein in question can be achieved by exposing the vesicle to an osmotic gradient, created by adding a high-osmolarity sucrose or sorbitol solution to the outside of the vesicle. The osmotic pressure will cause efflux of water from the vesicle, resulting in shrinkage (Figure 14). This technique can be used with all kinds of vesicles, both and artificial liposomes (Chapter 3.7.1) and natural cells (Chapter 3.7.3).

The amount of light that is scattered by particles depends on their size. It can be described by the Mie theory in the case the particle size is similar to the wavelength used. The most commonly known example of this phenomenon is probably the scattering of light by small lipid droplets in milk, which give it its white colour. This behaviour can be
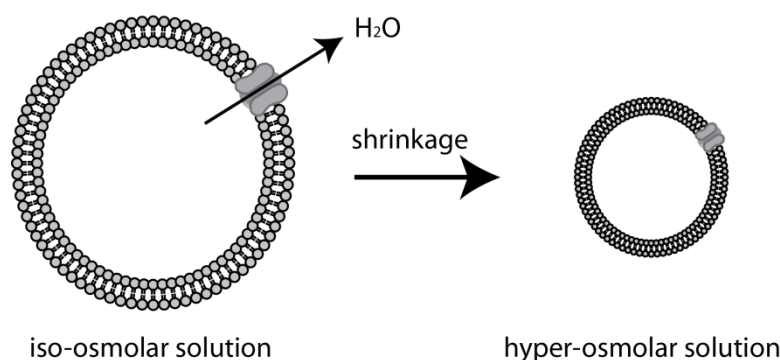

Figure 14. Shrinkage of vesicles upon application of a hyper-osmolar gradient. As the cells are subjected to a concentrated sugar- or sugaralcohol solution, water flows out of the cell due to osmotic pressure. Aquaporins facilitate this transport, resulting in faster shrinkage.

used to follow the size change of the vesicles upon application of an external gradient as described above. The scattering signal is typically measured at a 90° angle to the light source to minimize the background signal.

The half-life for unfacilitated water transport through the bilayer of a vesicle with a diameter in the order of 100 nm is around 100 ms, whereas aquaporin mediated transport is up to a factor of 5 faster. This excludes the use of conventional spectrophotometer devices, instead stopped-flow devices are necessary with a dead time <5 ms and high data acquisition rates to obtain data for the initial kinetics of water transport.

A single exponential function (Eq. 5) usually fits well to the measured data. In the case of mixed populations of vesicles – where the reconstitution of protein into vesicles has not succeeded in all cases and thus a fraction of the vesicles do not contain aquaporins – the fit to a double exponential function (Eq. 6) may be more appropriate. From these curve fits, rate constants k in dependence of the time t and signal amplitude y can be obtained that allow quantitative comparisons.

$$y = A_1 e^{-kt} + A_2 \qquad \text{(Eq. 5)}$$

$$y = A_1 e^{-k_1 t} + A_2 e^{-2t} + A_3 \quad \text{(Eq. 6)}$$

2.3.3   Spheroplast Assay

Spheroplasts are cells, whose cell wall has been fully or partially removed. Their shrinkage can be studied in a similar manner as the shrinkage of liposomes upon application of an osmotic gradient [137]. Yeast spheroplasts, as used in this study, are significantly larger (>1 μm) than the unilamellar vesicles described above and thus benefit from an increased signal in light scattering or light absorption measurements.
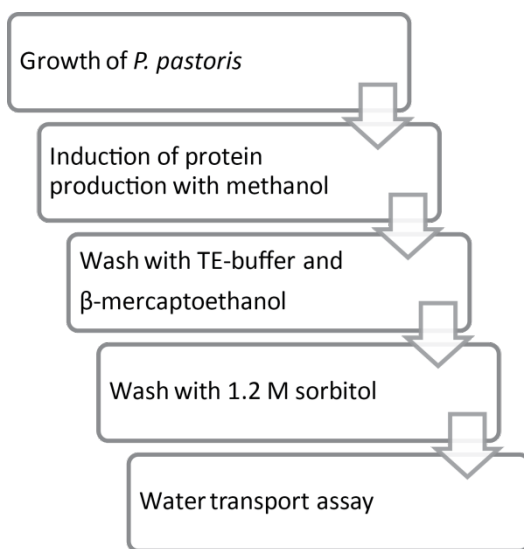

Figure 15. Preparation of spheroplasts. A small culture of *P. pastoris* is grown over night, after which the production of an aquaporin is induced using methanol. The cells are washed and treated with β-mercaptoethanol to destabilize the cell wall, before the water transport assay is conducted in a stopped flow apparatus.

The yeast's cell wall provides integrity to the cell and can prevent it from shrinkage. It can be removed or destabilized by chemical (e.g. β-mercaptoethanol) or enzymatic (e.g. by Zymolyase) means. Complete removal of the cell wall results in protoplasts that are extremely fragile and have a tendency to aggregate in solution, which makes spectroscopy based measurements unreliable. Treating the cells with β-mercaptoethanol only destabilizes the cell wall sufficiently to allow shrinkage, while avoiding the afore mentioned effects. Nevertheless, these spheroplasts have to be treated with care. Osmotic shock can easily result in the rupture of the cells, so that washes and other handling have to be performed in a high-osmolar sorbitol solution (1.2 M) that matches the cell's internal osmolarity.

Unlike the reconstitution of protein into liposomes, the membrane protein is produced directly by the yeast cell – here *P. pastoris* – and processed by its internal machinery to be integrated into the cell wall. This has the benefit of a larger fraction of protein in the membrane being functional and inserted in the correct orientation, but the water transport rates obtained heavily rely on the expression levels

of the protein in the host. As these levels cannot be controlled easily, the assay is best suited for comparative mutational studies of the same protein.

2.3.4 Surface Plasmon Resonance

Surface Plasmon Resonance spectroscopy (SPR) is one of the dominating optical biosensor technologies [138]. It is mainly used to study the interaction between proteins and other substances, e.g. drugs, pesticides, antibodies or other proteins [139]. One of its main advantages is that it does not require the use of chemical labels, and hence neither modification of the receptor or the ligand is necessary. The real time measurement allows the observation of binding kinetics, which can give valuable information, for instance during drug design projects.

The phenomenon of SPR is characterized by a drop in total reflection intensity at a certain angle. This effect is caused by surface plasmons, an electromagnetic field found at a metal-dielectric interface. At the correct incident angle, these surface plasmons can be excited by a light wave, where this resonance angle depends on the difference in refractive index between the metal and the dielectric medium (Figure 16).
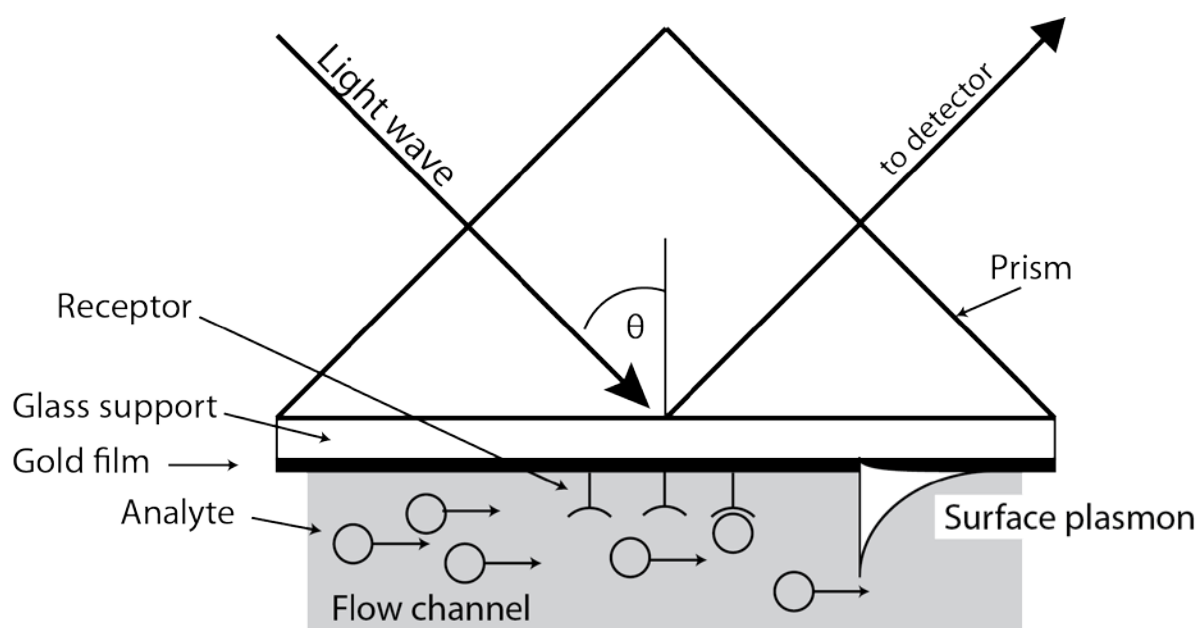
Figure 16. Experiment setup using SPR as a biosensor. An analyte is flowing above a gold surface, to which a receptor is covalently attached. As the analyte binds, the refractive index above the surface increases, changing the SPR angle θ. This signal decreases exponentially with the distance from the surface.

The concept for biosensing applications is usually executed as follows: The metal surface, often a thin layer of gold, is functionalized and the receptor is immobilized onto it using appropriate chemical techniques. This metal surface of the chip is connected to a microfluidic device, while its back side – consisting of a glass support – is coupled to a prism, which is part of the detection system. Running buffer is constantly passed through the microfluidic system (Figure 16) throughout the experiment. To test binding between a ligand and the receptor molecule, the ligand is added to the running buffer and the binding of the ligand to the receptor increases the refractive index on the surface. This alters the surface plasmon, which in turn results in a change in the resonance angle [140]. This signal depends on the mass and amount of the molecule bound and on the distance to the surface [141]. The latter is often also referred to as the "evanescent field" with a decay length of ca. 150 nm above the surface.

In this study, a Biacore2000 machine developed by Biacore AB, Sweden has been used. It is equipped with a robotic autosampler that can be used with 96-well microplates. The microfluidic system creates 4 separate flow channels on a gold chip (Figure 17), which can be used to run controls or study several receptors in parallel. Injection sequences can be programmed for the device using the Biacore2000 Control Software (Biacore AB, Sweden). The software also converts the difference in resonance angle into so-called Resonance Units (RU) to represent the amount of mass bound to the surface. 1 RU unit equals a shift of 0.0001° in resonance angle, which corresponds to approx. 1 pg/mm^2.
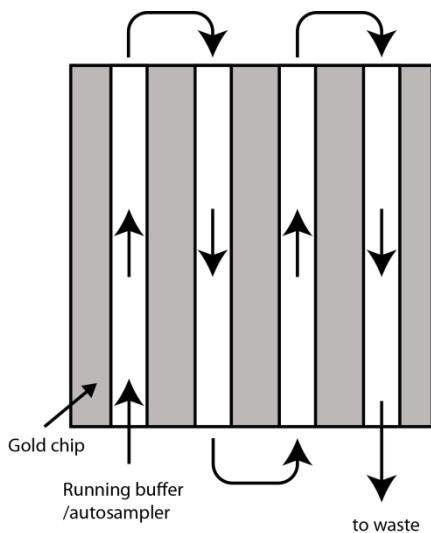

Figure 17. Serial flow of running buffer or sample on a Biacore gold chip. Four flow cells are created on the chip that can be functionalized in differently.

3 Structural and Functional Studies on *Pichia pastoris* Aquaporin

Even though *P. pastoris* is a well established and characterized host for protein production, aquaporins from this yeast have hitherto been unknown. This can be attributed in parts to that although its genome had been sequenced, it had not been publicly available until 2009 [142, 143]. The very similar and well studied baker's yeast *Saccharomyces cerevisiae* is known encode for two orthodox aquaporins, Aqy1 and Aqy2, and two aquaglyceroporins, Fps1 and Yfl054 in its genome [144].

In 2007, we discovered that *P. pastoris* contains an orthodox aquaporin, which we characterized both structurally and functionally. As many other aqua(glycerol)porins from various yeast species, it has an extended N-terminus of more than 30 amino acid residues, whose function has been unclear until now. Its high similarity (67%) to the pathogenic yeast *Candida albicans* makes it also medically interesting.

In Paper I, we determined the initial X-ray structure of Aqy1 to 1.15 Å (pH=3.5) and 1.4 Å (pH=8.0) respectively – which is still the highest resolution published for a membrane protein to date. We also characterized the function of the protein extensively by a mutational study and newly developed functional assays. From this and molecular dynamics simulations we were able to suggest two complementary gating mechanisms for its regulation.

Based on these results, crystal quality and size was improved significantly. These optimized crystals also showed improved diffraction; for the first time, a membrane protein structure could be determined to below 1 Å resolution (Paper II). Also, the structure was re-determined at ambient temperature to 1.3 Å that yielded more information on the protein's dynamics (Paper III).

3.1 Discovery and Structure Determination

During our work with human aquaporin 1 (hAQP1), we obtained crystals that diffracted to the exceptional resolution of 1.15 Å. This is still the highest resolution published for a membrane protein to date. Solving the structure using molecular replacement (PHASER [120]) with hAQP1 as search model yielded a solution. However, it became quickly evident from the electron density that the protein was not hAQP1 (Figure 18). At that exceptional resolution, it was possible to manually determine the amino acid sequence for most parts of the protein from the density – including the elongated N-terminus that is not present in the hAQP1. The mystery of what protein we were dealing with was approached by bioinformatics methods (BLAST [145]), which showed that our sequence had high similarity to a protein found in *Pichia stipidis*. This in turn led to the conclusion that
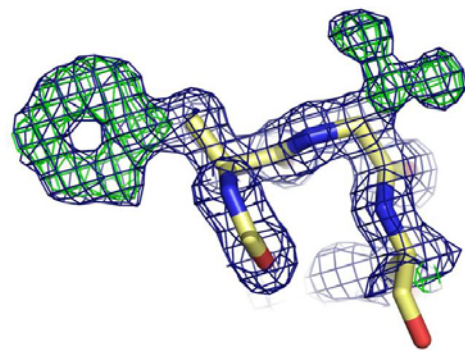

Figure 18. Electron density obtained after molecular replacement and initial stages of refinement. The electron density (2FoFc-map, blue) and difference density (FoFc, green) clearly shows that the docked sequence of hAQP1 (yellow backbone) was incorrect.

the protein had to stem from our production host, *Pichia pastoris*. As the genomic sequence of *P. pastoris* was not publicly available at that time, we contacted James M. Cregg (Keck Graduate Institute, USA), who provided us with the DNA sequence of this aquaporin. This information was used to clone the gene from genomic DNA by PCR and to reintroduce it into *P. pastoris* to optimize protein production for further studies. In analogy to the orthodox aquaporin from *S. cerevisiae* [144], we named this protein from *P. pastoris* Aqy1.

3.2   Structure and Function of Aqy1 (Paper II)

3.2.1   Structural Overview

The monomeric unit of *Pichia pastoris* aquaporin 1 (Aqy1) consists of six transmembrane helices. As for other aquaporins, loops B and E are elongated and fold back into the membrane from the intracellular and extracellular side respectively to form an additional seventh pseudo-helix (Figure 19A). Where the two helices meet at the centre of the membrane, the two asparagines from the aquaporin NPA-signature motifs (Asn 112 from half-helix B and Asn224 from half-helix E) stack onto each other (Figure 19E). Together, all these helices surround the water conducting pore (3.4.2).

Because of the exceptionally high resolution of 1.15 Å (Table 2), the water molecules in the pore can be refined anisotropically (Figure 19E). At the aromatic/arginine (ar/R)-constriction region – less than 2 Å wide – on the extracellular pore entrance, the water ellipses are heavily elongated along the pore axis. Looking at the electron density of the water molecules makes this phenomenon even clearer: The density that is observed around the ar/R-region consists of a continuous, cigar-like shape of ca. 7 Å in length, where waters are disordered in the crystal structure. The remaining water molecules inside the pore are clearly defined, with anisotropic displacements being smallest close to the central of the pore, i.e. the NPA motif. This – together with the hydrogen bond between the water and Asn112 – is indicative of a possible rotation by anchoring of the water to the NPA-motif, as predicted in MD simulations (Chapter 1.2.3).

Surprisingly, the water pore is blocked further towards the cytosolic side. Tyr31, located in the elongated N-terminus acts as a plug, which does not allow passage of any water molecules (Figure 19A). The water molecule closest to Tyr31 in the channel can be observed clearly and forms a very strong and well located hydrogen bond network to Tyr31, Gly108 and Gly109. Individual hydrogen atoms forming the bond can be observed as difference density (Figure 19G), which emphasizes the strength of this bond. Tyr31 is also part of a wider network of interactions, where Tyr27 anchors the N-terminus to the aquaporin core and Pro29 breaks a short helix that would otherwise be formed by the N-terminus to allow Tyr31 access to the pore entrance (Figure 19B).

Both in the crystal structure as well as *in vivo*, aquaporins are arranged as homotetramers (Figure 19D). Crystallizing in the space group I4 with the four-fold axis penetrating the central pore, it is difficult to obtain information on this part of the structure from crystallographic data. It seems however, that the central pore is mainly empty as no significant electron density has been observed. This pore is lined with hydrophobic residues and has several constriction regions, which makes it unlikely that any solutes or water might be able to pass this pore.

The tetrameric arrangement seems to be crucial for the closure of the monomeric pore: The elongated N-termini of the neighbouring monomers interact with each other via countless hydrogen bonds to form a stable bundle. This bundle does not appear to be caused by crystal packing effects: it is placed in a large cavity of the crystal grid and is surrounded by solute (Figure 19F). Molecular dynamics (MD) simulations confirm that the N-terminal bundle is stable.
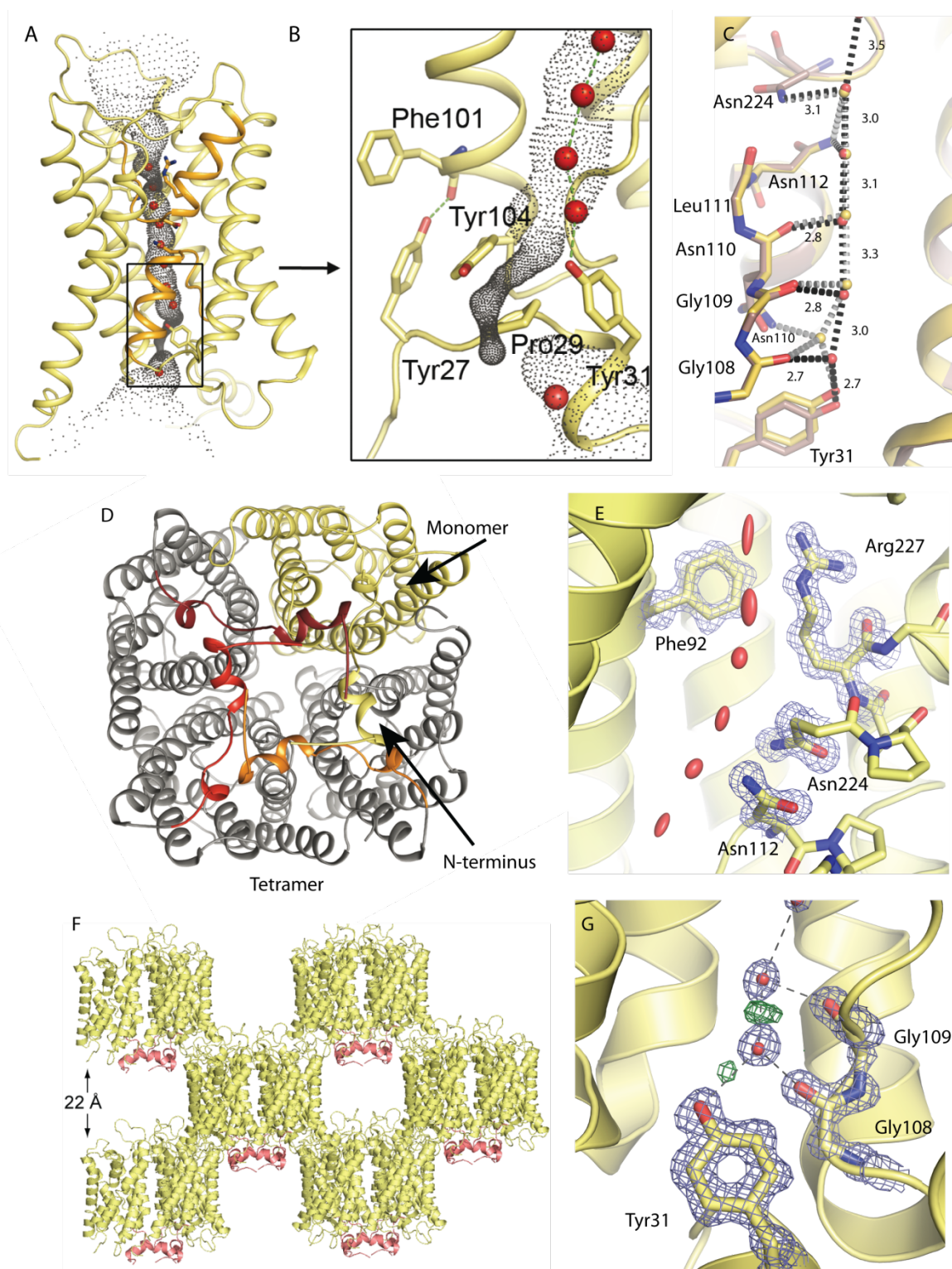
Figure 19. Overview of the structure of Aqy1. (A) Aqy1-monomer with water pore (calculated by HOLE [146]) shown in grey. (B) Zoom onto the intracellular channel part, which is blocked by Tyr31, Tyr104 and Pro29. Tyr27 anchors the N-terminus onto the protein core via a hydrogen bond. (C) The water pore at room temperature (red) and at 100 K. Water molecules form hydrogen bonds to the hydrophilic face of the pore. Distances in Å. (D) Aqy1 tetramer. One monomer is highlighted in yellow. The N-termini (yellow to red) are placed above the neighbouring monomers and are intertwined to a bundle. (E) Anisotropic refinement. The water molecules show anisotropic motion along the channel. In the ar/R region (Phe92, His212 (not shown), Arg227) anisotropy is high due to unclear density at 1.15 Å resolution. $2F_oF_c$-map contour level = 2.0 σ (blue). (F) Crystal packing, view along the a-axis. The N-terminus is located in a large solvent gap. (G) Hydrogen bonds stabilizing Tyr31 at the intracellular pore entrance. $2F_oF_c$-map contour level= 2.0 σ (blue), F_oF_c-map contour level = 2.2 σ (green).

3.2.2 Functionality and Gating

The blockage of the channel in the crystal structure raises the question: Why would any organism invest the energy to produce something that apparently does not have any function, like a closed water channel?

This question was further examined by testing whether the channel can be opened. To this end, a mutational study in combination with a stopped-flow light scattering functional assay based on the shrinkage of spheroplasts (Chapter 2.3.3) was performed. To reduce the background of the water transport assay, *P. pastoris'* endogenous aquaporin was deleted from the yeast strain (aqy1Δ) [147].

Removal of the side chain of Tyr31 (Y31A) that blocks the pore in the crystal structure increases water permeability 6-fold compared to wildtype Aqy1 (Figure 20). Likewise, N-terminal truncation of the protein by 36 residues (ΔN36) results in a 6-fold increase in transport rate, confirming that the channel can transport water when it is not blocked by Tyr31.

The most likely hypothesis for a closed channel is that it is advantageous for the organism if it transports water only under certain conditions. While the general permeability of a cell's membrane towards water is often regulated on a transcriptional level, this type of regulation known as gating is thought to provide the cell with a means to respond rapidly to an extracellular stimulus, such as a sudden change in osmolarity or rapid freezing [65]. The spinach aquaporin SoPIP2;1, for example, is gated by pH and phosphorylation [26]. The former can be excluded for the case of Aqy1 as both crystal structures – obtained at pH 3.5 and 8.0 respectively – show an RMSD of only 0.08 Å for 248 C_α-atoms, indicating that their structures are identical within the experimental error.

However, bioinformatics tools (NetPhos 2.0 Server [148]) suggest that Ser107 is located in a consensus phosphorylation sequence. The hypothesis of Ser107 being involved in a gating mechanism was confirmed *in silico* using MD simulations: mimicking phosphorylation by the mutation of Ser107 to aspartate (S107D) lead to a widening of the pore from 0.8 Å to more than

Figure 20. Light-scattering traces from spheroplasts subjected to a hyperosmolar solution. aqy1Δ refers to a strain of P. pastoris, where the endogenous Aqy1 has been genetically deleted. *Top*: Transport rates k (inset) for the deletion strain (aqy1Δ, red), wildtype Aqy1 (aqy1Δ+Aqy1, orange), and an N-terminally truncated construct (aqy1Δ+Aqy1ΔN36, gray). For comparison, transport rates for human aquaporin 1 are shown (aqy1Δ+hAQP1, light blue). *Bottom*: Putatively pore opening mutations of Aqy1. Wildtype (aqy1Δ+Aqy1) and the mutant S107A (aqy1Δ+Aqy1S107A, dark blue) show low permeability, whereas S107D (aqy1Δ+Aqy1S107D, green) and Y31A (aqy1Δ+Aqy1Y31A, purple) show a significant increase.

2 Å, enough for a water molecule to pass through. A transport assay on *P. pastoris* mutant spheroplasts tests corroborated these calculations: The S107D- mutant experimentally showed an increase of water transport similar to the mutationally opened channels Y31A and ΔN36.
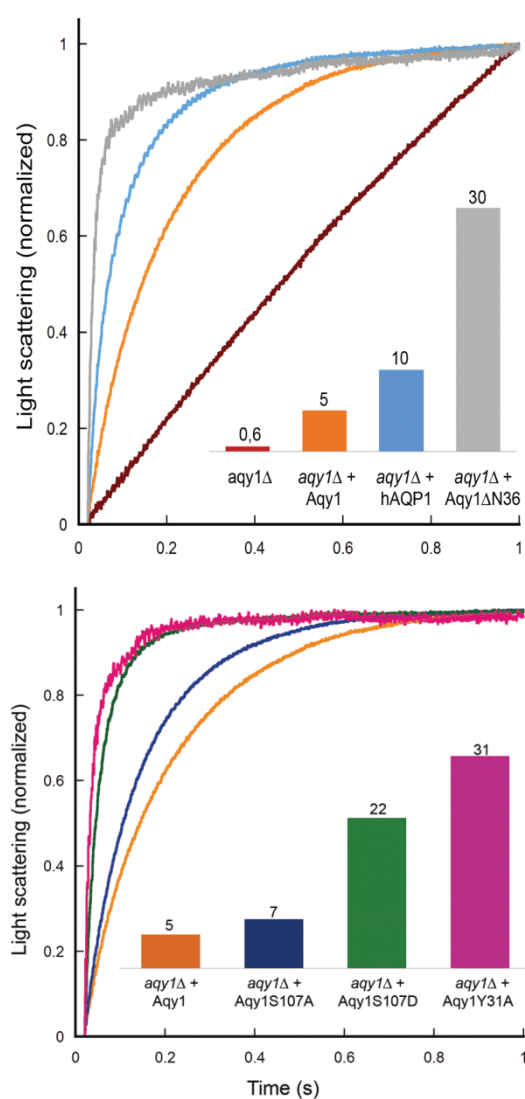
The control mutant's S107A permeability was substantially lower, indicating that indeed a mutation mimicking phosphorylation is needed to open the channel (Figure 20). Phosphorylation would provide the yeast with a way to rapidly change the water permeability of its membrane, where regulation by other means might be too slow.

An even faster mode of regulating water flux across the membrane is gating by mechanosensing, which provides an instant response to external stress that affects membrane tension or curvature. Confirming this idea experimentally is difficult, as most transport assays are based on measuring size change, where the lateral pressure and curvature change during the experiment. MD simulations, however, supported the idea that the channel can be opened by mechanosensing. Our observation that wildtype Aqy1 reconstituted into proteoliposomes – that have higher curvature and a different lipid composition than *P. pastoris* – exhibits the same water transport rate as the truncated construct ΔN36 further strengthens this hypothesis.

	2W2E	2W1P	Sub-atomic Resolution	Ambient Temperature
pH	3.5	8.0	8.0	8.0
Temperature	100 K	100 K	100 K	293 K
Data collection				
Space group	I4	I4	I4	I4
Cell dimensions				
a, b, c (Å)	91.4, 91.4, 80.8	90.9, 90.9, 80.6	90.7, 90.7, 80.3	92.5, 92.5, 81.2
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å) [1]	1.15 (1.20-1.15)	1.40 (1.48-1.40)	0.90 (0.95-0.90)	1.30 (1.37-1.30)
R_{sym} (%)	8.0 (73.3)	8.8 (70.6)	4.2 (68.6)	7.0 (59.2)
$I / \sigma I$	9.3 (2.2)	12.2 (1.8)	22.6 (3.8)	11.7 (2.4)
Completeness (%)	94.7 (99.1)	100 (100)	99.7 (100)	99.6 (100)
Redundancy	3.9 (3.2)	5.3 (3.4)	10.7 (10.9)	3.4 (3.3)
Refinement				
Resolution (Å)	20.0 – 1.15	64.5-1.4	28.7-0.90	64.5-1.3
No. Reflections used	111,455	61,170	238202	79250
R_{work} / R_{free}	14.0/16.5	16.2/18.1	11.4/11.5	13.2/15.7
No. atoms				
Protein [2]	2,073	2,080	2,103	2056
β-OG	120	80	100	40
Cl⁻	3	2	1	2
Ca²⁺				1
Water	143	133	221	75
B-factors				
Protein	15.9	10.4	16.9	15.6
β-OG	48.7	24.3	27.7	32.1
Cl⁻	18.9	17.1	15.6	19.8
Ca²⁺				24.3
Water	28.6	24.1	29.6	34.0
R.m.s. deviations				
Bond lengths (Å)	0.014	0.008	0.009	0.026
Bond angles (°)	1.9	1.07	1.71	2.15

Table 2. Crystallography statistics for data collection and refinement of Aqy1, crystallized or obtained in different conditions. The sub-atomic and room temperature structure were obtained from identical crystallization conditions. 1) Values for the highest resolution shell are shown in parentheses. 2) Not including riding hydrogen atoms.

3.3   Crystal Optimization

The discovery of such well diffracting crystals opened up new possibilities that hitherto only had been accessible to soluble proteins. In particular neutron diffraction studies, which show the exact location of hydrogen atoms, are suddenly feasible even for a membrane protein. A nuclear density map would help to provide structural evidence for the many, sometimes contradictory, suggestions that have been made by theoretical methods and mutational studies over time.

Neutron diffraction, however, is still in its infancy. The low fluxes that can be produced by neutron sources today, require large crystals ($>1.0 \times 1.0 \times 1.0$ mm^3) to keep the time for data collection in the order of weeks or months. Essentially being a Laue-technique, that is to say using a polychromatic neutron beam, a low crystal mosaicity is crucial to avoid overlapping reflections. Hence, neutron diffraction experiments are performed at room temperature, as crystal freezing increases this mosaicity drastically. The latest developments, high energy neutron spallation sources (Spallation Neutron Source, Oak Ridge National Laboratory, TN, USA) have improved the obtainable flux of neutrons by a factor of 10, in consequence reducing the required crystal size and data collection times significantly. Yet, to the standards of membrane protein crystallography, comparably large and stable crystals are necessary.

To this end, we optimized the well diffracting crystals of Aqy1 in size and diffraction quality. Recombinant overproduction of Aqy1 in *P. pastoris* – which at the same time is its natural host – yielded excellent amounts of protein for this task. Despite a purification protocol that was tailored towards crystallization and not protein yield, 70 – 100 mg of pure protein could be obtained per litre of cell culture. Trials with his$_6$-tagged protein were performed, but the crystals obtained showed a higher tendency for twinning and hence these experiments were not pursued further. Even though this phenomenon was also observed in crystals from non-tagged protein, it appeared to be present to a lesser extent.
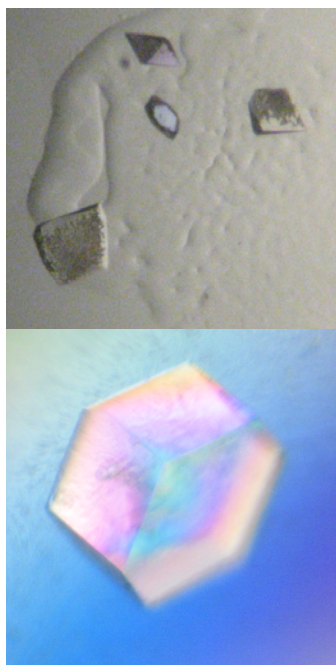

Figure 21. Crystals of Aqy1. *Top*: initial crystals obtained, the largest being approx. $0.1 \times 0.1 \times 0.1$ mm^3. *Bottom*: The optimized crystal is approx. $0.5 \times 0.5 \times 0.4$ mm^3 in size.

Crystallization condition were optimized by screening to 0.1 M Tris (pH=8.0), 26% PEG600, 0.2M CaCl$_2$. Also, temperature during crystallization was increased to 20 (initially 4 °C), which resulted in less amorphous precipitate in the crystallization drops and was necessary for room temperature experiments, as the crystals were observed to crack upon temperature change.

Interestingly, increasing the crystallization drop size required a different ratio of protein solution: precipitant solution: While for 2 µL protein solution, 1 µL of precipitant solution yielded the best crystals, an increased drop size of 20 µL of protein required only 4-5 µL of precipitant to obtain single crystals. The protein solution had a concentration of around 5 mg/mL for setup of crystallization experiments, but seemed not to be the main factor for the growth of single crystals. An explanation to this phenomenon could be that the protein peak observed from the size exclusion chromatography profile indicated a tendency of the protein to aggregate reversibly at high concentrations in the given buffer (0.1 M HEPES (pH=8.0), 100 mM NaCl, 1% OG). Hence, the effective concentration of free Aqy1-tetramers in the drop remains constant, even when higher protein concentrations are used.

Using this protocol, crystals up to 0.8 x 0.5 x 0.4 mm³ (Figure 21) were obtained within three to four weeks. These crystals were stable for extended amounts of time (>5 months) in their drops.

3.4   Sub-atomic Resolution Structure (Paper III)

3.4.1   Data Collection

Testing the diffraction quality of these new crystals revealed their true potential. Initial studies indicated that diffraction beyond 1.0 Å – hitherto unheard of in membrane protein crystallography – was possible, which we have mainly attributed to the increased volume of the crystals, which leads to the observation even of weak reflections and higher radiation tolerance. The comparably long time necessary for crystal growth to occur may, however, also result in a more perfectly packed crystal and hence also play a role.

Further improvement was made by freezing the crystals directly in the cryo-stream at the synchrotron beamline (ID-29, ESRF, France). In contrast to cooling the sample in liquid nitrogen, which is considered the standard technique, this approach has the advantage of avoiding a temperature change from 77 K (liquid nitrogen) to 100 K when transferred to the cryo-stream. As a result, crystals up to a size of 0.4 x 0.4 x 0.4 mm³ could be cooled without any significant damage. Larger crystals on the other hand suffered from freeze damage, presumably due to an insufficient cooling rate.

Modern synchrotrons are optimized for data collection from small crystals and the most intense beamlines often only provide small and highly focused beam sizes. The discrepancy between the available beam size (max. 75 μm in diameter at ID-29) and the crystal size required an advanced data collection strategy.
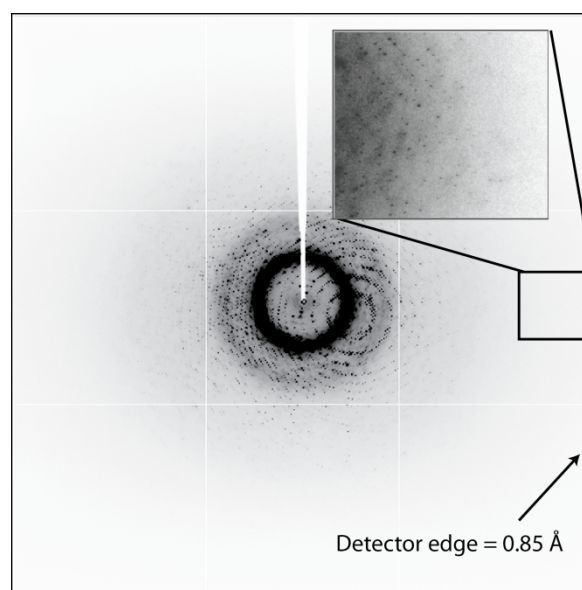

Figure 22. Diffraction image of Aqy1-crystals collected at ID-29, ESRF, France to a resolution of 0.85 Å at a wavelength of 0.65 Å.

To take advantage of the full crystal volume, the beam was translated during data collection ("helical script"). A decrease in wavelength to 0.65 Å allowed for the collection of data without repositioning the detector ("swing-out") with the additional benefit of reduced radiation damage. The data was collected in several passes with an offset of 30° between individual data sets; this resulted in a total of 2 data sets to low resolution (2.9 Å), 2 datasets to medium resolution (1.3 Å) and 8 datasets to high resolution (0.85 Å, Figure 22), of which 6 datasets were finally combined and scaled to 0.90 Å (Table 2).

3.4.2   The Water Channel

The primary function of aquaporins is the transport of water. As outlined in Chapter 1.2.3, it is crucial to a cells survival that this transport occurs in a very specific manner. In particular, the transport of protons would uncouple the proton gradient, which is used to drive the creation of ATP. Undesired transport of ions like sodium or potassium would on the other hand have a negative impact on, for instance, neurons, which rely on an ion gradient to conduct an electric signal.

A structure determined to sub-atomic resolution (Table 2) as for Aqy1 opens up the possibility to examine the chemical surroundings, which a water molecule has to pass as it is traversing the pore, in great detail. In particular, the orientation of side chains like asparagines can be determined unambiguously as the delocalized *p*-electron pair of the C=O double bond is observed, while the orientation of a histidine side chain can be determined by the intensity difference in the electron density. This is valuable for establishing hydrogen bonding patterns that are not only based on chemical experience, but on experimental evidence.

As mentioned in Chapter 1.2.3, the NPA region is located at the centre of the pore and has been suggested to be the main cause for proton exclusion. The orientation of the asparagines in this region can be clearly observed: Asn112 and Asn224 both direct their nitrogen atom towards the pore and act as a hydrogen bond donor (Figure 23A). The opposite face of the pore is very hydrophobic, and hence thought to force the water molecule to undergo a while defined rotation, while traversing that region in the pore. Unfortunately, even at this resolution, we were not able to directly establish the hydrogen bonding pattern from the electron density through the observation of protons. Whether this might be an indication for a high mobility of the water molecules in the channel or not is a matter of speculation.

At the ar/R constriction region, which is lined by Arg227, His212 and Phe92 (Figure 23B), the channel diameter narrows down to below 2 Å, hence allowing only water molecules to pass. As shown in Paper V, the most favourable place for the binding of sugaralcohols (which are chemically similar to water) is this ar/R region. This is not surprising in that sense that this part of the channel being narrow requires a way to actively recruit water molecules. Binding the water molecules too tightly however, would result in a decrease in transport speed, which is contrary to the purpose of a water channel.

The finding, that the outstanding electron density of all things was blurry around the ar/R region gives reason to take a closer look: Modelling the correct number of water molecules is difficult – the best solution was found to either model water molecules with extreme anisotropicity (Figure 19E) or to name several defined positions for water molecules. We found, that two pairs of water molecules with an occupancy of 70 and 30 % respectively fitted the density best (Figure 23B). This "complementary" modelling, i.e. only every second position along the ar/R is occupied at any one time (Figure 23C, D), is in agreement with the small distances of approximately 1.5 Å between the individual atom positions. While all waters in the constriction region can form up to 4 very strong H-bonds of less than 2.8 Å, their geometry is less than ideal for any water molecule along the ar/R-region. This lowers the energy barrier for water transport significantly as the binding affinity is reduced and results in a highly selective, but yet fast filter region.

Figure 23. *A*) Stereo view of the NPA-region. The conformations of Asn112 and Asn224 can be determined unambiguously ($2F_oF_c$-map, contour level = 5.3 σ, blue). Leu208, which is part of the hydrophobic face of the channel, forces two water molecules towards the NPA-motif. The delocalized electrons in the C=O-bonds can be observed as continuous density, while the density of the $C-NH_2$-bond is disconnected. *B*) Stereo representation of the aromatic arginine constriction region of Aqy1 at 0.90 Å resolution. Two sets of waters can be observed (red and orange). When electron density for the first set of waters (red) is modelled, difference density appears (green, F_oF_c at contour level = 4.0 σ), into which the second set of waters (orange) can be modelled with lower occupancy. The electron density for the high occupancy water molecules is displayed at a contour level of 2.0 σ (2FoFc-map). Possible hydrogen bonding patterns are shown for the high (black) and low (orange) occupancy sets as dashed lines. Electron density (2FoFc) for Arg227 and His212 is contoured at 5.3 σ in dark blue.
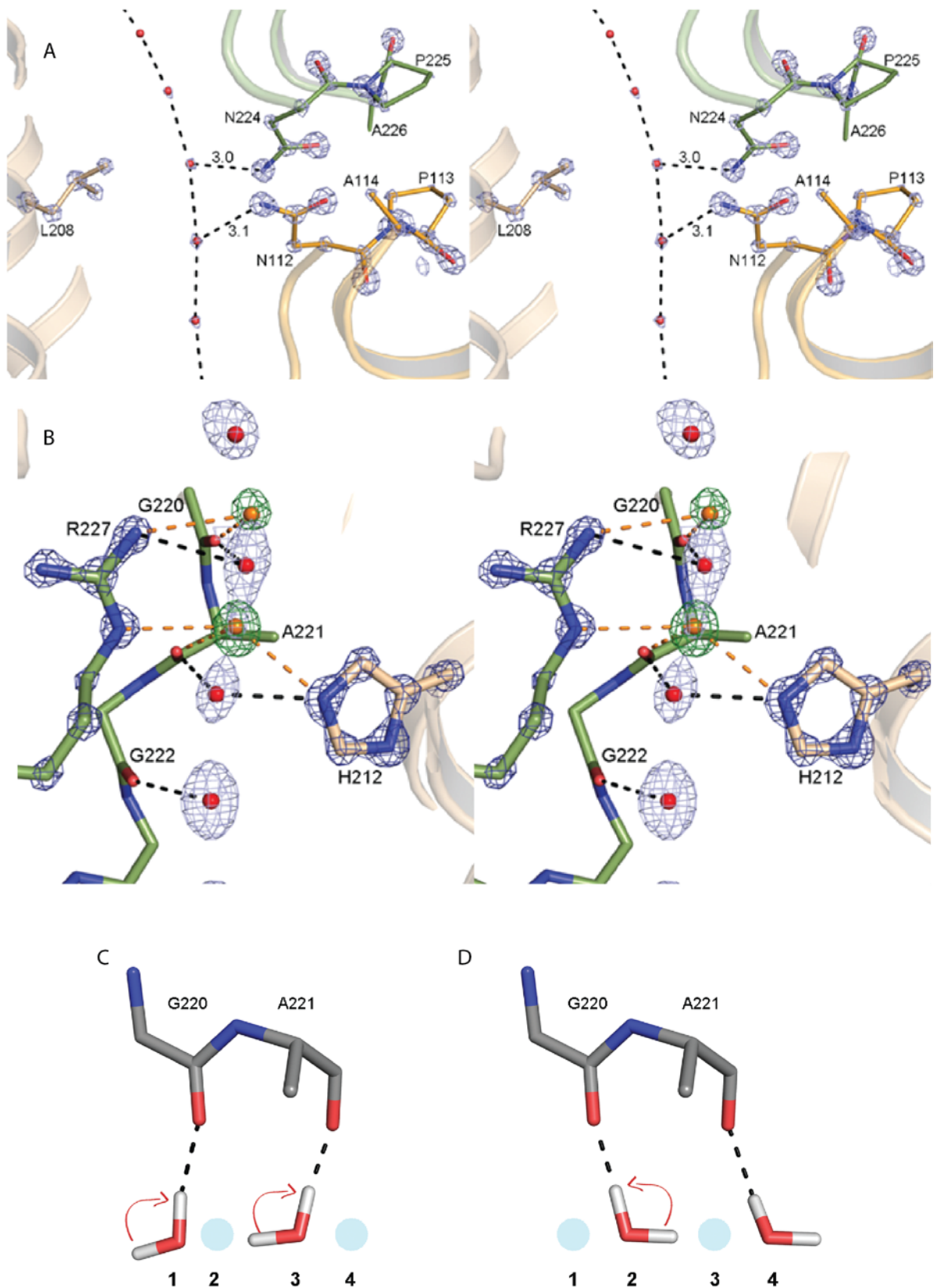
C, D) Schematic illustrating the movement of two water molecules through the ar/R water selectivity filter. *C*, As water molecules at positions 1 and 3 move to positions 2 and 4 they must rotate clockwise, since their movement is constrained by H-bond donor interactions with Arg227 and His212. *D*, Conversely, as the water molecule at position 2 moves to position 3 it must rotate counter-clockwise. This synchronized pair-wise hopping suppresses proton transport via a Grotthuss mechanism.

3.5   Structure at Ambient Temperature (Paper IV)

3.5.1   Data Collection

Exploiting the extraordinary size, stability and diffraction quality of Aqy1, we also determined its structure at ambient temperature. While initially considered as a pre-experiment to neutron diffraction studies, where today often a combination room temperature X-ray and neutron data are processed simultaneously [113] to maximize accuracy, the experiment is interesting in its own right: So far, (to my knowledge) only two membrane protein structures – photosynthetic reaction center [100] and the outer membrane porin ompF [149], have been determined at room temperature. These date back to the time before data collection was routinely carried out at cryogenic temperatures and is certainly owed to the high susceptibility of this class of proteins to radiation damage. Their high solvent content (typically $>> 50\%$) allows for quick and long range diffusion of radicals formed by high-intensity X-ray radiation, which damage the crystal. In addition, weak crystal contacts between proteins exacerbate this problem.

Hence, it was not surprising that X-ray data collection on Aqy1-crystals was challenging. Initial trials on an in-house diffractometer successfully showed diffraction. This attempt was constricted by the hardware, however. Keeping in mind that data was collected at room temperature, the processed data showed unexpectedly high mosaicity, most likely due to a high beam divergence of the anode. In addition, long exposure and slow read-out times of the CCD-detector lead to long experiments, with significantly reduced diffraction quality over time. Collecting data to high resolution – a resolution better than 1.7 Å was observed – requires repositioning of the detector ("swing-out") – the time of exposure to radiation is further increased. This phenomenon can be explained by the fact, that at room temperature, the radiation damage induced is not only proportional to the radiation dose the crystal has been exposed to. It is also a function of the time of exposure, as free radicals can freely diffuse through the crystal, causing secondary radiation damage [150].

This prompted us to attempt data collection at a synchrotron source, where the hardware is not limiting. Performing the experiment at the third-generation synchrotron ESRF, France and Diamond, UK proved to be problematic, both logistically and experimentally. Transporting crystal trays (which contain liquid) on an airplane is not very well looked upon by the authorities. Since the crystals were also very sensitive to the pressure changes during the flight, they often cracked upon opening the well at the destination – crystals that survived showed decreased diffraction quality. For this reason, crystallization trials were directly set up at ESRF, France. These crystals showed good diffraction at room temperature (up to 1.2 Å), and the collection of a data set was possible. However, all data suffered from severe radiation damage over time. It is likely, that this was caused by the high-intensity X-ray beam (even at maximum attenuation) and the small beam size of 0.1×0.1 mm^2, which does not exploit the full crystal volume.

The best results were obtained from crystals transported by car to the second generation synchrotron Maxlab, Sweden at beamline Cassiopeia 911-5. To maximize the amount of data not influenced by radiation damage, crystals similar in size as the final crystal ($0.5 \times 0.5 \times 0.35$ mm^3) were sacrificed to estimate the dose the crystal would tolerate, which was found to be approximately 120 seconds of exposure with a beam 0.2×0.3 mm^2 in size. Overall, the significantly weaker and larger beam allowed the rapid collection of a dataset to 1.3 Å, which showed little to no radiation damage (Table 2).

3.5.2 Structural Comparison
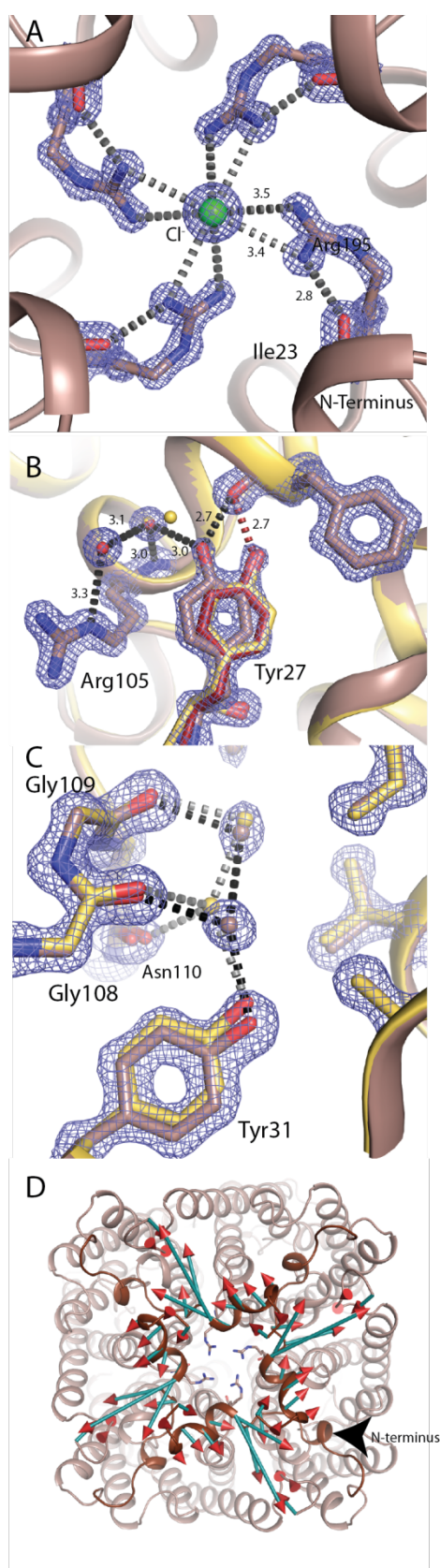
Structural analysis showed that the overall fold of the protein is conserved between the structures determined at ambient and cryogenic temperature. The high resolution of the structures permitted an in-detail and quantitative comparison. Cruickshank's DPI model [151] provided an estimation for the error in the atomic coordinates, which were found to be as small as 0.033 Å (ambient temperature) and 0.010 Å (100 K) respectively. With this unusually accurate data, we could address a number of crystallographic as well as biological questions.

The B-factor is commonly referred to as a "temperature-factor", describing thermal motions of individual atoms. In the determined structures, the average B-factors for both protein chains are similar (16.9 at 100K vs. 15.6 at RT, Table 2), but closer investigation revealed that this average is misleading. B-factors at RT are surprisingly smaller in well ordered regions like the protein core, while they are identical or larger in loop and terminal regions of the protein (Figure 25). This indicates that B-factors observed at low temperatures to a large extent represent the disorder that has been induced through freezing and not thermal motions, at least in the inflexible parts of the protein.

From a structural point of view, the most significant change becomes clear when looking at the tetrameric assembly. The N-terminal bundle, which is involved in the aquaporin's gating mechanism, moves outwards by up to 1.5 Å during freezing (Figure 24D). This movement is caused by Arg195, which is located at the central pore and forms a new hydrogen bond to Ile23 in the N-terminus at RT, while being disordered at low temperature (Figure 24A).

Figure 24. Structural changes upon freezing. Crystal structure at RT (red) and at 100 K (yellow). All distances given in Å. A) View along the central pore. Arg195 is clearly defined at RT, while disordered at 100K. B) Tyr 27 showing a double conformation at RT, creating a new hydrogen bond network to Arg105. C) Tyr31 blocking the cytosolic side of the pore. At RT, Tyr31 moves out of the pore by ca. 0.7 Å. D) Outwards movement of the N-terminus upon freezing. Arrows are 7 times as long as the actual movement, no movements shown smaller than 0.5 Å.
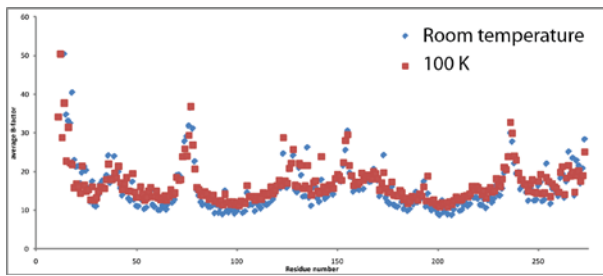
Figure 25. Average B-factors vs. residue number for structures determined at 100 K (red) and at room temperature (blue). In the transmembrane regions, the B-factors are significantly smaller at room temperature compared to the structure determined at 100 K, while they are similar or larger in the loop- and terminal regions.

This tightening of the N-terminal bundle propagates conformational changes in other residues along the N-terminal chain. Tyr27, which anchors the N-terminus to the protein core, undergoes a rotation by 60° and changes its hydrogen bonding pattern (Figure 24B). Further along the chain, the pore blocking residue Tyr31 is pulled out of the water channel by 0.7 Å (Figure 24C). Although the channel is still not wide enough to transport water, this can be regarded as the initial conformational change towards an opening of the pore.

3.5.3   Gating by Mechanosensing

Apart from changes in the flexibility and the possibility of trapping the protein in a certain conformation, the main differences between data obtained at room- and cryogenic temperature are the crystallographic cell dimensions. These are typically larger at room temperature, as the crystallization solution shrinks upon cooling. Water and protein have different expansion coefficients: While water, or more precisely the crystallization solution, has been seen to reduce its volume by 2-7 % upon freezing to 100 K, the equivalent change for the protein corresponds to only 1-3 % [152]. In Aqy1-crystals, the total unit-cell volume is reduced by 5% and thus well in accordance with literature. Interestingly, this contraction is not uniform in Aqy1-crystals: While the cell dimensions in the *ab*-plane are reduced by 2.0 % (1.8 Å), the *c*-axis only contracts by 1.2 % (0.9 Å). This has implications on the protein. Since the protein crystallizes in the space group I4, the aquaporin tetramers are coplanar with the *ab*-plane. The non-isometric shrinkage hence causes an increase in lateral pressure on the protein upon freezing and is likely to be the basis for the observed rearrangements between the two structures.

This finding corroborates the idea of an aquaporin, whose water flow is regulated by mechanosensitive gating. In contrast to the
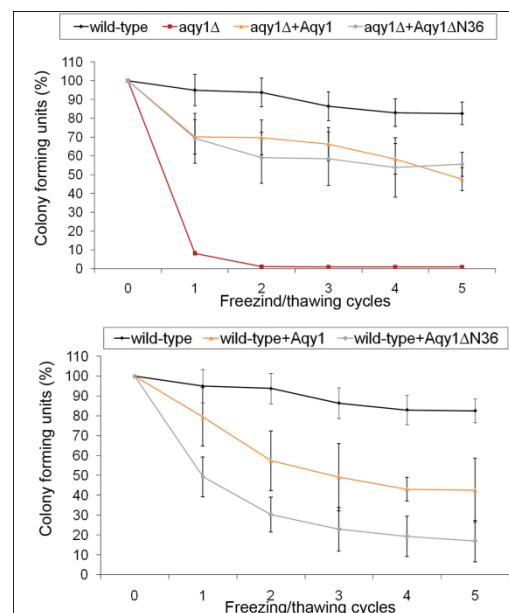

Figure 26.   *Top*: Relative number of *P. pastoris* colony forming units (CFU) following multiple freeze/thaw cycles (x-axis) of the wild-type strain (black), *P. pastoris* with the *AQY1* gene disrupted (aqy1Δ) (red); aqy1Δ overproducing Aqy1 (orange) and Aqy1ΔN36 (grey). *Bottom*: Similar to above, but using the wild-type *P. pastoris* strain (black) and overproducing recombinant Aqy1 (orange) and Aqy1ΔN36 (grey). The recombinantly overproduced proteins negatively impact survival, with Aqy1ΔN36 having a significantly larger effect.

suggestions from molecular dynamics simulations presented in Paper I, this new experimental evidence suggests that the channel closes upon exertion of lateral pressure and opens in the relaxed state. This result also makes the biological interpretation of yeasts' freeze-tolerance (Figure 26) appear in a new light:

The fact that microorganisms, such as yeast, possess aquaporins at all is unexpected. The intrinsic water permeability of their membranes should be sufficient to perform the required water transport, as the surface-to-volume ratio of single cellular-organisms is large [65]. And indeed, many strains cultured for a long time in the non-stressing conditions of a laboratory do not produce aquaporins, whereas their wildtype relatives do [153]. Mainly two areas where aquaporins can be beneficial have been suggested, namely osmotic stress tolerance and the survival of rapid freezing. Even though the latter is somewhat counterintuitive, microorganisms are frequently

exposed to the rapid change in temperature, for instance when they are exhaled with an animal's breath during a cold day. In order to survive freezing, yeasts concentrate their intracellular content by the efflux of water, which results in a decrease in harmful water crystal formation upon freezing. At slow cooling rates, the intrinsic water permeability of the membrane seems to be sufficient, whereas rapid freezing benefits from the presence of aquaporins.

Our finding of an aquaporin that is gated by mechanosensing upon the increase of lateral pressure further enhances this model. While the channel is at least partially opened under ordinary conditions and thus helps the cell to export water quickly, it closes as the cell shrinks. In this respect, the advantage of having a gated aquaporin does not differ from having a non-

Figure 27. Schematic view over the advantage of a gated aquaporin. See text for details.

gated aquaporin. Once a frozen cell is thawed, however, a closed aquaporin could prove valuable: Fast influx of water through a non-gated aquaporin, which is constitutively open, into an incompletely thawed cell might cause damage to the cytoskeleton and organelles. A gated aquaporin could hence buy the cell sufficient time to first thaw entirely, before expanding its volume (Figure 27). Consequently, one can hypothesize that while any aquaporin is beneficial with regards towards freeze-tolerance, a gated aquaporin additionally increases the survival of thawing.
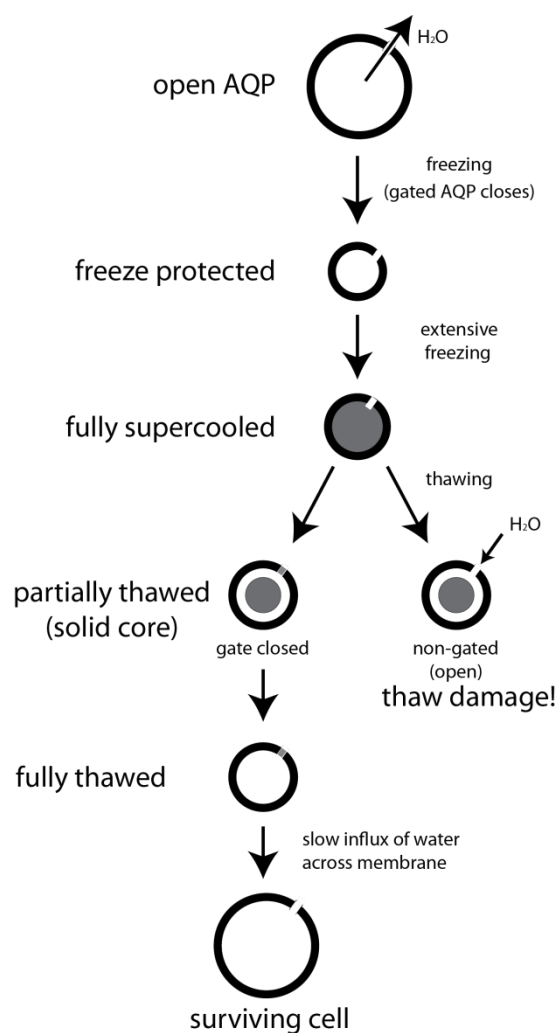
3.6 Summary

For the first time, a membrane protein structure has been determined to below 1 Å resolution. This milestone shows that although crystallizing membrane protein is difficult due to their amphiphilic nature, atomic resolution structures are possible. The extraordinary stability of these crystals allowed us to perform advanced structural studies also at ambient temperature.

These studies provided us with novel insights into the proton exclusion mechanism and into the significance of the elongated N-terminus of *P. pastoris*. The in fungal aquaporins often observed appendices are required for a gating mechanism designed to rapidly open and close the channels upon external stress like changes in osmolarity or rapid freezing to aid the yeast's survival. Phosphorylation and mechanosensing are the two most likely ways that this gating mechanism is regulated *in vivo*.

4   Drug Screening Assay Based on Surface Plasmon Resonance

Structures like the one presented in the previous chapter do not only provide basic biochemical information, but can also be used to rationally design drugs. Both these application require functional assay to confirm the proposed mechanism or inhibitor respectively. In membrane protein biochemistry there is frequently a lack of reliable and potentially high-throughput transport assays – in particular for the transport of water and uncharged solutes.

The transport of solutes like glycerol or shorter sugar alcohols across cell membranes is for most life forms just as crucial as the transport of water. *In vivo*, the transport of sugar alcohols is often facilitated in a similar fashion as the transport of water. In fact, the families of the channel proteins are closely related: orthodox aquaporins and aquaglyceroporins (Chapter 1.2).

Studying the transport of solutes on a molecular level is often achieved by similar means as the transport of water, i.e. by observation of size change of vesicles upon application of a gradient. Vesicles are first prepared and loaded with solute. A suspension of these vesicles is mixed with a solution of the same osmolarity, but a different solute concentration, which thus creates a solute gradient. This results in an efflux of the solute from the vesicle, which also triggers efflux of water and leads to the shrinkage of the vesicle. This method requires meticulous equilibration of the osmolarities of the buffers used on the inside and the outside of the vesicle. Otherwise, the rapid in- or efflux of water would shrink or burst the vesicle without being based on the solute gradient and lead to a significant decrease in signal amplitude, if the size change is followed by e.g. light scattering. Analysis of this secondary effect can be complicated, as not only the size, but also factors like aggregation and change in refractive index due to changes in the liposome structure contribute to the light scattering signal [154].

In the first part of this chapter, a novel method for measuring transport directly using SPR is described and demonstrated using PfAQP (Paper V). In the second part, the method has been developed further into a medium-throughput drug screening system (Paper VI).

4.1   *Plasmodium falciparum* Aquaporin

4.1.1   Medical Relevance

The malaria parasite *Plasmodium falciparum* (Chapter 1.4) produces a single aquaglyceroporin, PfAQP, that transports water and glycerol equally well [25, 49, 155, 156]. This is surprising, as most organisms have two distinct types of channels for transporting water and glycerol: Orthodox aquaporins and aquaglyceroporins. It has been reasoned, that by limiting the amount of membrane proteins, a parasite like *P. falciparum* reduces the number of cell surface exposed structures, which can be recognized by the host and lead to an immune response [157]. This makes PfAQP – and membrane proteins in general – a valuable target for drugs against those parasites, which otherwise are known to rapidly develop resistance [49].

PfAQP has been found to be important for the parasite during its blood stage, i.e. when it enters the erythrocytes of its human host. Here, the parasite proliferates massively, which requires – amongst other things – the generation of large amounts of lipids to form new cell membranes. *P. falciparum* imports glycerol from the blood serum to accommodate its increased need for lipid biosynthesis [158]. Studies showed that growth of *Plasmodium berghei* – the rodent ortholog to *P. falciparum* – was impeded when knocking out PbAQP. Mice infected with an aquaporin knock-out strain of *P. berghei* survived significantly longer than mice that were injected the wild-type strain, confirming that glycerol transport by PbAQP is vital for the parasite [159].

4.1.2 Production and Purification

PfAQP was produced recombinantly in *Pichia pastoris* in bioreactors and shaker flasks. For optimal production yield of PfAQP, an optimised gene construct was used [92]. Compared to the production of orthodox aquaporins, the growth of *P. pastoris* expressing PfAQP has been observed to struggle during the induction phase with methanol, a phenomenon which has also been observed for all other aquaglyceroporins worked with in our laboratory. One can only speculate on the origin of this phenomenon, but that methanol – as the smallest of all "sugar alcohols" – can be transported through the aquaglyceroporin too quickly for the microorganism to process it, and thus becomes toxic, seems a reasonable hypothesis.

The stability of the protein – which has a great tendency to aggregate in solution – was significantly improved from several hours to several days by reconstituting it into lipid vesicles directly after Ni-NTA affinity chromatography. During the purification (Chapter 3.1.2) of the protein it became evident, that its tendency to aggregate is high, since it could not be concentrated to high concentrations or stored over extended periods of time. For the functional studies performed here, purity was found to be sufficient after this purification step (Figure 28). The protein was then reconstituted into sonicated vesicles made from *E. coli* polar lipid extract (Avanti Lipids, USA) at a lipid to protein ratio of 50 (w/w) using the dilution method (Chapter 3.7.1). All samples were filtered through a 0.22 μm filters before injecting them into the Biacore device to avoid clogging.
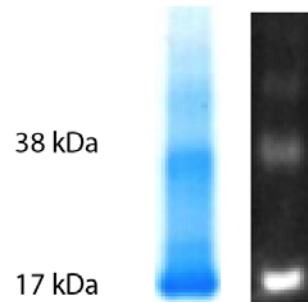
Figure 28. SDS-PAGE (left) and Western Blot (Anti-His) of purified PfAQP. The protein shows the distinct banding pattern s, as well as a high inaccuracy in weight estimation (MW (PfAQP=29 kDa), as it can be observed for many membrane protein

4.2 SPR-based Transport Measurement (Paper V)

4.2.1 Immobilization of Vesicles to the SPR Surface

For all experiments, a pure gold chip (SIA-Au-Kit, GE Healthcare) was used. The unilamellar vesicles used are around 120 nm, as determined by dynamic light scattering. As the decay length of the SPR-signal is approximately 150 nm, it is important that the liposomes are located as close to the surface as possible. Other chips available (like CM5) are usually coated with a dextrane layer. This increases the amount of ligand that can be bound; however this is only true for small ligands up to the size of proteins. Large vesicles do not benefit from this effect, but suffer from reduced signal due to the longer distance from the surface.
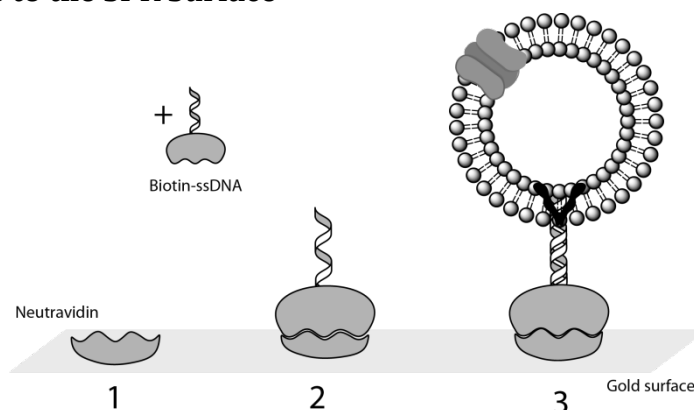

Figure 29. Immobilization of lipsomes. The gold surface is first functionalized with Neutravidin using amino-coupling (1). Adding biotinylated-ssDNA results in strong binding (2). Finally, proteoliposomes are added to bind to the ssDNA-fragment (3). Prior to this, the proteoliposomes are incubated with with a cholesterol-DNA construct that inserts irreversibly into the membrane (not shown). The free DNA-end has a single-stranded overhand ("sticky end") which is complementary to the biotin-ssDNA construct so that binding can occurr via hybridization.

The chip was first cleaned in a boiling mixture of 25% ammonia/30% hydrogenperoxide/water (1:1:5 v/v/v). The pure gold surface was then functionalized by incubation in an 11-mercapto undecanoic acid/ethanol solution (10 mg/mL) over night, which yields a carboxy-terminated chip. The biotin-binding protein NeutrAvidin (50 µg/mL, Invitrogen, USA) is then linked to the carboxy group using amine-coupling (Amine-Coupling Kit, GE Healthcare Life Sciences, Sweden). The high affinity of biotin to NeutrAvidin was then used to stack an additional layer of a single-stranded, 15- base long DNA fragment that was covalently modified with biotin ("DNA-B"). The resulting surface thus exposes single stranded DNA.

Complementary ssDNA is present on the surface of the vesicles. This is achieved by incubating the proteoliposomes with a DNA-cholesterol construct for 30-60 minutes. The DNA construct consists of a 15-base pair long stem, which is attached to two cholesterol molecules that can integrate into the proteoliposome bilayer. From that stem, a 15 base long sticky end protrudes, that can hybridize with the surface-ssDNA upon contact (Figure 29, Figure 30). This bond is strong enough to keep the vesicles on the surface at a slightly basic pH, but can easily be used to remove the vesicles by e.g. flushing the system with water.
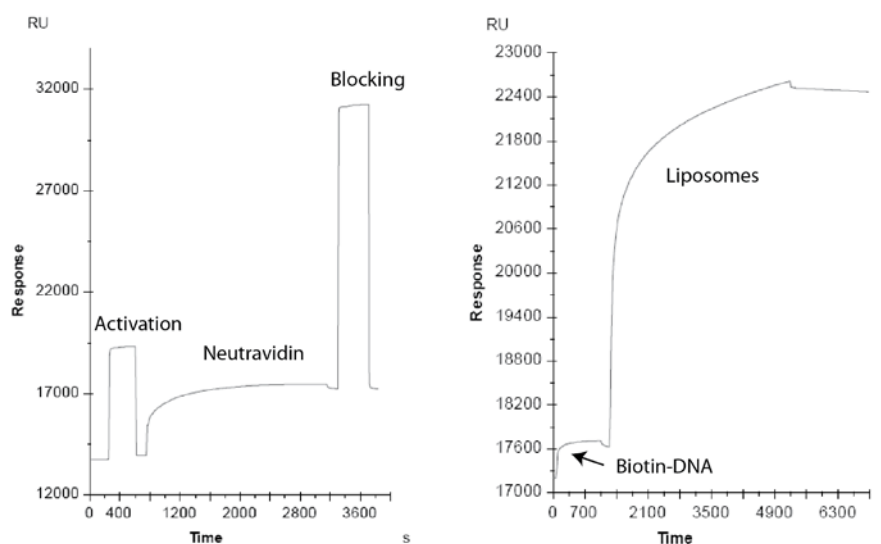
Figure 30. Surface functionalization. *Left*: The carboxylated surface is first activated with a mixture of EDS/NHC, functionalized with Neutravidin using amine coupling. Unreacted carboxy sites are finally inactivated with ethanolamine. *Right:* First, biotinylated single stranded DNA is bound to Neutravidin, before liposomes containing complementary ssDNA are injected.

4.2.2 Transport of Solutes across Lipid Bilayers

The content of the vesicles can be modified by incubation in a different buffer. Running buffer (typically 50 mM Tris pH=8.0, 150 mM NaCl) is constantly flowing over the surface. The proteoliposomes can be filled with solute, e.g. glycerol, by switching to a buffer containing that solute. In practice, this is done by injecting buffer containing the solute onto the chip and waiting for the vesicles to be filled, i.e. for equilibration between the volume outside and inside the vesicle. The sugar alcohols glycerol, erythritol, xylitol and sorbitol were used for that purpose (Figure 31). As they have increasing molecular weights, but a very similar chemical structure, they show increasing refractive indices and are expected to be transported at decreasing rates through the membrane or the protein channel.

The refractive index and thus the SPR signal increases significantly – between 3000 and 6000 RU – by adding a concentrated sugar alcohol solution onto it (Figure 32). Reverting the flow back to running buffer results in a removal of solute from outside the vesicle, followed by a drop in RU. However, solute contained inside the vesicles is not simply washed away, but is released slowly instead. This efflux of the solute from the vesicle that is caused by the generated concentration gradient can be traced using SPR (Figure 33).

Figure 31. Chemical formulae of the sugar alcohols glycerol, erythritol, xylitol and sorbitol. Where applicable, the D-enantiomer is shown.

The rate constant of transport can be obtained by fitting a curve to the data (Figure 34). Protein reconstitution into liposomes, however, seldom yields a defined number of protein molecules per vesicle, but rather a distribution. The reconstitution protocol used aims to incorporate one functional unit – i.e. a single aquaporin tetramer – into the majority of the liposomes. Even though it is possible to describe this using a Poisson distribution, it is often sufficient to use two terms in a double exponential function. One term accounts for the slow component caused by the non-facilitated efflux from empty liposomes and one being fast from liposomes providing facilitated transport through an aquaporin.
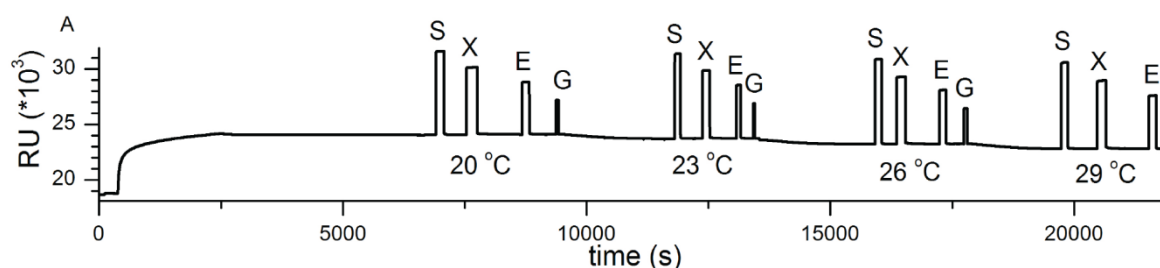

Figure 32.   Representative trace (overview) obtained from an efflux assay using PfAQP liposomes. Initially, liposomes are immobilized onto the surface. When running buffer is exchanged to sorbitol (S), xylitol (X), erythritol (E) or glycerol (G), the refractive index above the surface changes drastically due to the change in refactive index in the buffer according to the molecular weight of the substances.
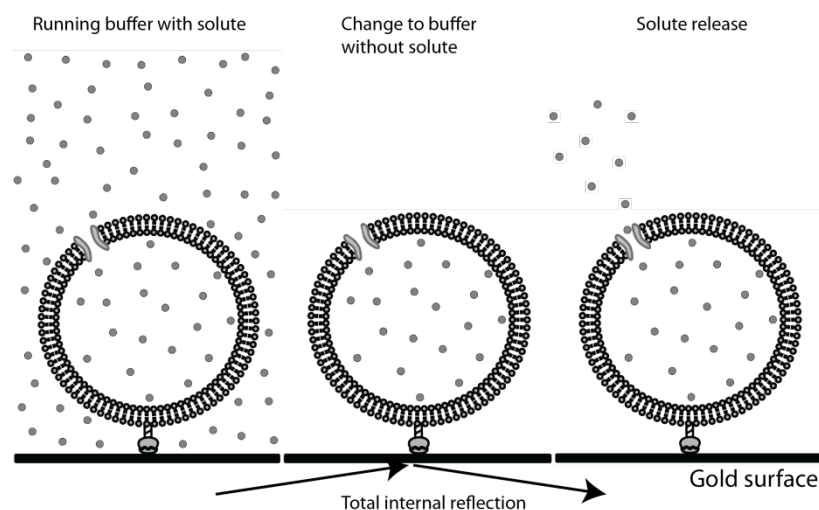
Figure 33. Release of a solute from a liposome. Liposomes are loaded with solute from the surrounding buffer. Upon exchange to a buffer without solute, a concentration gradient arises, causing the slow efflux of the solute. This process is traced by surface plasmon resonance.
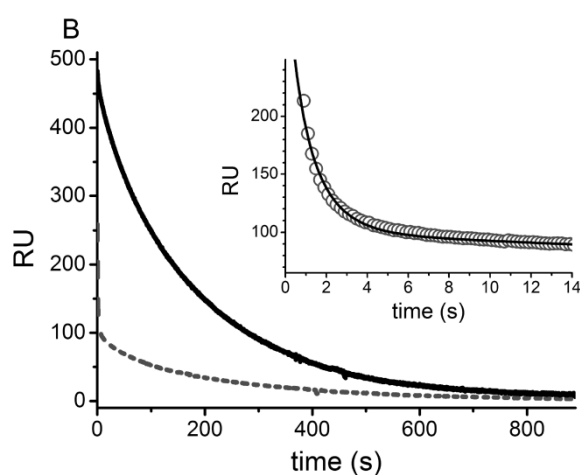

Figure 34. Release of 400 mM xylitol from liposomes prepared from *E. coli* polar lipid extract. Liposomes without aquaporin (black) release the sugar alcohol significantly slower than proteoliposomes with reconstituted PfAQP (grey, dashed. The inset shows the first 14 seconds of xylitol release from PfAQP-proteoliposomes.

In the case of longer sugar alcohols like xylitol or sorbitol, which are transported sufficiently slow across the membrane even in the case of facilitated diffusion, rate constants for both components can be determined. When using a double-exponential function to fit a curve to the data, the rate constant for the slower component has been found to be very similar to the rate constant of the empty liposome sample, confirming the assumption that a fraction of empty liposomes is present.

For study of small molecules like glycerol and erythritol, their initial release is fast and not compatible with the time resolution – 100 ms – of the Biacore2000 machine used in the case of facilitated diffusion. Thus the vesicles have already been largely emptied when the first data point can be acquired. Since it it slower, however, the non-facilitated transport – both in a control sample containing non-proteoliposomes and in the non-proteoliposome fraction of the protein sample – can be observed and quantified. The difference in initial signal amplitudes between a control sample without protein and the proteoliposome sample is indicative of the fraction of empty liposomes in the proteoliposome sample. When reconstituting at LPR 50 with the protocol described in Chapter 3.7.1, a typical fraction of empty liposomes has been found to be between 0.3 and 0.5, while the remaining liposomes contain one or more functional aquaporins.

4.2.3 Functional Studies on PfAQP

PfAQP has been shown to have been reconstituted functionally into the liposomes both by stopped-flow light scattering and surface plasmon resonance experiments. The novel SPR method allowed repeated functional studies on the same set of liposomes, making accurate comparisons significantly easier. In particular, this overcomes the issue that different protein preparations and reconstitution attempts are not always equally efficient.

The aquaglyceroporin has been found to transport glycerol, erythritol, xylitol and sorbitol. Rate constants k (s^{-1}) could be determined accurately for the longer sugar alcohols sorbitol and xylitol and estimated for the transport of erythritol. From these, permeability coefficients P can be determined:

$$P = kVS^{-1} = \tfrac{1}{3}kr \quad \text{(Eq. 7)}$$

with V being the volume, S the inner surface area and r the inner radius of the liposome. Together with Fick's law

$$J = P\Delta C \quad \text{(Eq. 8)}$$

the flow J (mol x cm^{-2} x s^{-1}) can be estimated when the concentration gradient ΔC (mol x cm^{-2}) is known. The radius of the liposomes used has been determined to be between 45-60 nm, which thus yields the permeability coefficients given in Table 3. The order of permeability coefficients is glycerol < erythritol < xylitol < sorbitol, i.e. mainly dependent on their size.

	Activation Energy E_A [kcal/mol]	Permeability coefficient [cm/s]
Glycerol	n.d.	>1 x 10^{-5}
Erythritol	n.d.	~3 x 10^{-6}
Xylitol	5.5 ± 0.5	1.0 x 10^{-6}
Sorbitol	5.6 ± 0.3	1.1 x 10^{-7}

Table 3. Permeability coefficients and activation energies for various sugar alcohols through PfAQP. Values were obtained from PfAQP reconstituted into liposomes from E. coli polar lipid extract.

The system can also be employed to measure activation energies E_a (Table 3) via the measurement of transport rates at different temperatures. Using the Arrhenius equation

$$\ln(k) = \ln(A) - \frac{E_a}{RT} \quad \text{(Eq. 9)}$$

with k being the rate constant, A the preexponential factor and R the universal gas constant. Values can be extracted from the slope of a ln(k) versus 1/T-plot (Figure 35). E_a for passive diffusion through a cell membrane (12.8±1.8 kcal/mol) is significantly slower than facilitated transport. Surprisingly however, the activation energies E_a for the transport for sorbitol and xylitol through PfAQP are identical with 5.6±0.3 and 5.5±0.5 kcal/mol respectively – despite their different size and transport rates. This leads to the conclusion that it is the preexponential factor A (Eq. 9) is decisive for the speed of transport. This factor represents the frequency of "attempts" of a solute molecule to enter the pore for transport, which is lower for larger molecules due to a slower rate of diffusion and steric considerations.
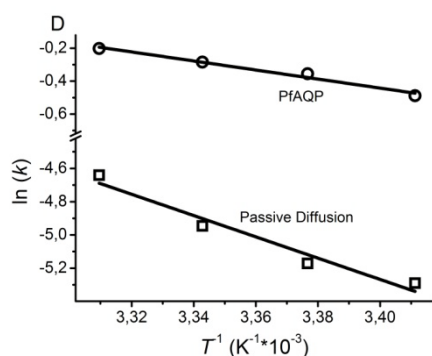
Figure 35. Arrhenius plot. Rate constants were determined at 20, 23, 26 and 29 °C for the efflux of xylitol from liposomes without aquaporin ("passive diffusion") and proteoliposomes containing PfAQP ("PfAQP").

Docking simulations with Autodock [160] indicate that the most stable site for the binding of any of the sugar alcohols inside the channel is at the ar/R constriction region (Figure 36). Binding energies were calculated to be -3.7 kcal/mol and -4.0 kcal/mol for xylitol and sorbitol respectively. This is in good agreement with the values obtained experimentally (see above) and confirms earlier theoretical studies on GlpF from *E. coli* [48], that it is indeed the ability of a solute to displace waters from this site, which is the key determinant for the selectivity of a channel.
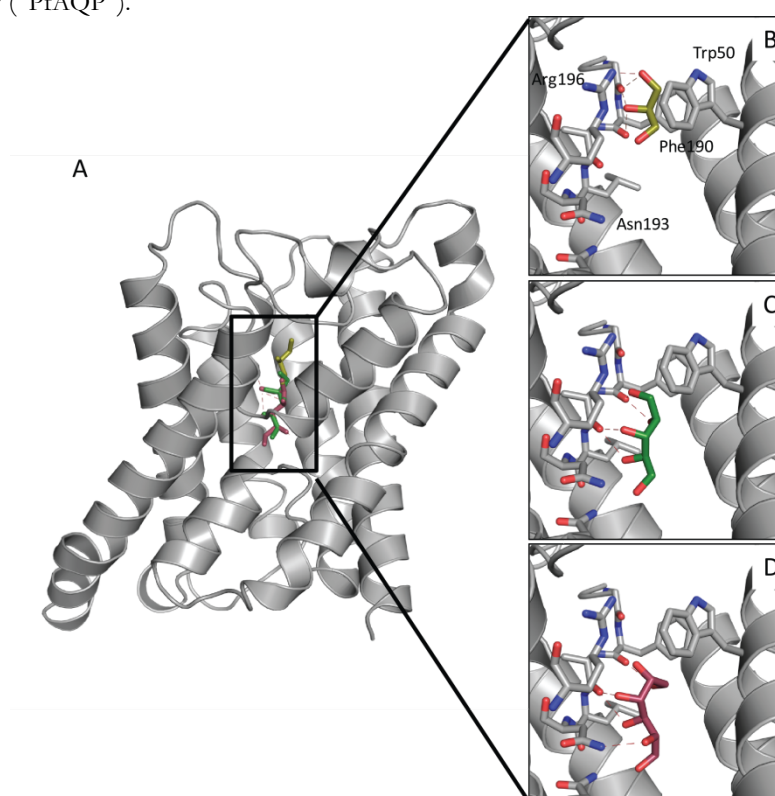
Figure 36. Docking of sugar alcohols the water channel of PfAQP using AutoDock 4.2. All sugar alcohols (B: glycerol, C: xyltiol, D: sorbitol) showed the highest binding affinity to the ar/R constriction region when allowed to dock freely on both the intracellular and extracellular vestibule as well as the channel.

4.3   Medium-throughput Drug Screening Assay (Paper VI)

Being based on a set of immobilized and reusable proteoliposomes, the functional assay provides the ideal basis for the development of an inhibitor screening assay. Using SPR, which is a well established technique for measuring e.g. protein-protein binding, provides the advantage that robotics for automatization via autosamplers are readily available.

4.3.1   Screening Setup

Inhibition assays were developed for use with 96-well plates, which allowed convenient pipetting using multichannel pipettes or pipetting robotics respectively. All changes in buffer on the chip ("injections") were programmed using Biacore Control Software Version 3.1.1 (Biacore AB, Sweden) for a completely unattended run of the machine.
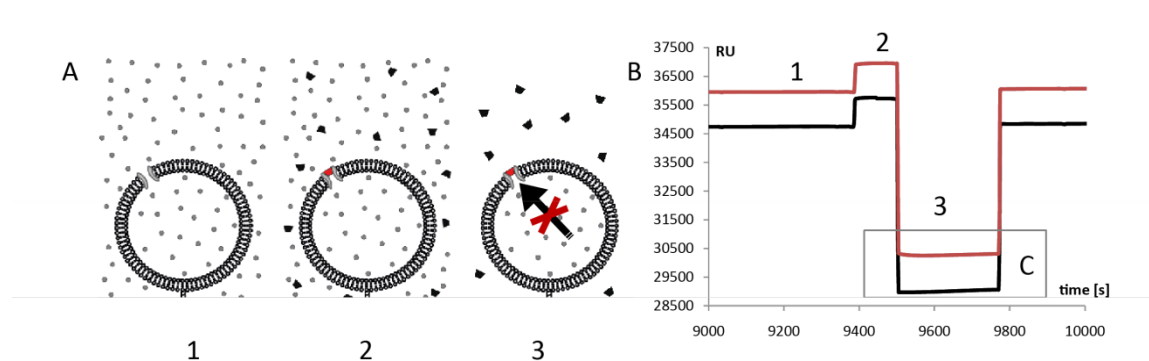
Figure 37. Assay sequence. A) Schematic of loading of vesicles with sorbitol (1), incubation with inhibitor and with sorbitol present (2) unloading of liposomes after removing sorbitol from the running buffer with inhibitor present (3), here depicted with a successfully bound inhibitor, blocking the efflux of sorbitol otherwise caused by the osmotic gradient. B) Sensogram from a Biacore experiment, showing the sequence of events. Liposomes containing protein (red) are incubated with the respective buffers in parallel to a liposome blank without protein (black).

The setup has been optimized for use with PfAQP, but measuring transport for other proteins should be possible without major changes in the concept. As we have seen in Paper V, PfAQP transports glycerol which is – apart from water – its substrate in nature at very high rates. Sorbitol on the other hand is transported significantly slower, while still giving an indication of
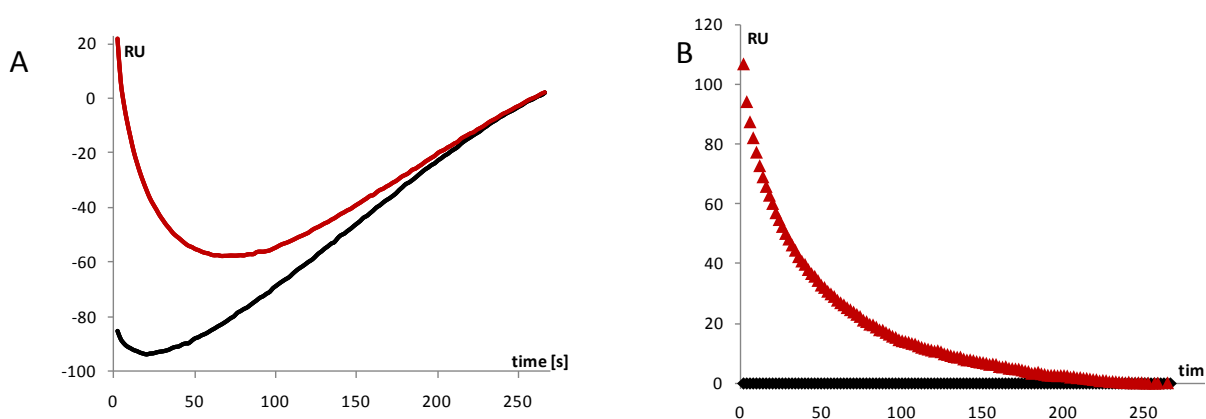
Figure 38. Release of sorbitol from the liposomes. A) Raw data from the PfAQP-proteoliposome trace (red) and the liposome blank (black). B) Subtraction of the blank from the proteoliposome trace yields a single exponential decay function

the degree of functionality of the channel. Used in combination with a bound inhibitor, this transport rate can be used to quantify the efficiency of the inhibitor. Another advantage of using sorbitol is the absence of passive diffusion across the lipid bilayer at any significant rate, when no protein is present. This eradicates the problem of having a mixture of

proteoliposomes and empty liposomes present on the surface, as only the former contribute to the change in SPR signal.

Sorbitol is present in the running buffer at all times (50 mM Tris pH=8.0, 100 mM NaCl, 5% DMSO (v/v), 300 mM sorbitol) to minimize the number of buffer changes. To test an inhibitor, it is dissolved in running buffer and then injected for 2 minutes onto the surface to allow for binding. Directly thereafter, a buffer containing the inhibitor, but not sorbitol is injected onto the surface for 4 minutes. This triggers unloading of the sorbitol from the vesicle, which can be followed by SPR. Each of these injection sequences is followed by a control injection, where the same procedure is performed, but without the presence of any inhibitor. This injection yields information on the integrity of the surface and the protein and is used for normalization.

To account for any unspecific binding to the surface and liposomes as well as mixing effects, we took advantage of the possibility to run the PfAQP sample and a blank measurement (empty liposomes) in parallel on the same gold chip. Even though using the multichannel mode in the Biacore2000 device reduces the maximum data collection rate from 10Hz to 2 Hz, this has been found to be sufficient for the time scale of sorbitol release. After aligning these two traces, subtraction yields the signal only caused by sorbitol release (Figure 38).

Solubility of the compounds to be screened in the buffer used is a common issue in drug design. DMSO – being known to dissolve most substances due to its chemical nature – is the solvent most commonly used for this purpose. Most samples can be dissolved in 100% DMSO, but upon dilution precipitation can easily occur. Including a relatively high concentration of DMSO – 5% (v/v) – into the running buffer aims to minimize this problem. The surface and the liposomes are resistant against DMSO and remain stable up to 8% DMSO (Figure 39). Moreover, a variation in DMSO concentration between different buffers, which can arise experimentally, does not show any significant effect on the sorbitol transport rates. Most likely, DMSO diffuses quickly through the membrane when compared to sorbitol transport.
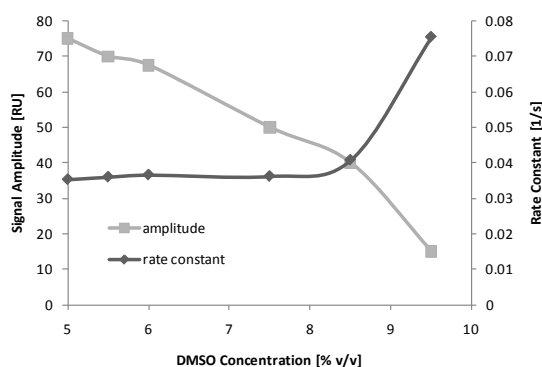

Figure 39. Stability of the surface towards different concentrations of DMSO. Sorbitol release transport rates (black) after four 115 second treatments on the same surface with the respective DMSO-concentration are shown. In grey, the decrease in amplitude is shown.

For an efficient assay, it is also important that the surface is reusable for an extended amount of time. Comparing the rate constants of the control injections allows an estimation of the surface quality. The liposomes on the surface remained stable for ca. 10 hours. By re-immobilizing new liposome sample on top of the existing sample, the lifetime could be extended, although it never regains the initial quality.

The surface also remained stable when inhibitors were injected. Depending on the inhibitor however, unspecific adsorption to the liposomes were observed. The amount of adsorption was found to be dependent on the type of inhibitor (Table 3). Earlier studies by Myszka et. al [161-163] suggested that the magnitude of this effect might provide a better value than the conventional octanol-water partition coefficients logP [164] for the ability of the intestinal tract or other cells to resorb drugs.

Finally, data processing was performed using BiaEvaluation 3.1 (Biacore AB, Sweden). The trace obtained from the blank-subtracted proteoliposome trace fitted well to a single

exponential function (Eq. 5, Chapter 2.3.2) from which the rate constants could be obtained. To obtain a relative value that shows the potency of an inhibitor, k(inhibitor)/k(control), which is the relative permeability through the pore, was calculated.

4.3.2   Inhibitor Screening of PfAQP

We used the assay described above to screen for inhibitors against PfAQP. A pre-selection of inhibitors for the medium-throughput assay was obtained using *in silico* screening of a compound library (Maybridge). The virtual inhibitor screening of membrane channels like aquaporins is difficult however, as they do not possess a defined active site where an inhibitor could strongly bind. Thus, the experimental confirmation of a virtual "hit" is essential. The 10 compounds with the best binding scores from virtual screening were selected in the first round.

Of these inhibitors, 3 were sufficiently soluble in the buffer used. The other compounds either did not dissolve in 100% DMSO or precipitated when diluted to 5% DMSO. As compound solubility is a crucial for its potential use later on, it is rational to remove them from further screening at this stage. The remaining 3 compounds were tested in triplicates at 0.1, 1 and 10 μM concentrations to estimate IC_{50}-values. Compound **1** successfully inhibited PfAQP of up to 80% with an IC_{50} value between 1 and 10 μM (Figure 40), while no inhibition was observed for Compounds **2** and **3**.

For verification, compound **1** has also been tested in an *in vivo* assay against *P. falciparum* and was found to affect the parasite at EC_{50} = 0.884 mM, which is in very good agreement with the *in vitro* result, given the entirely different nature of the assays. This outcome is also a confirmation that PfAQP is indeed a drug target, a fact that has been cause for discussion in the past [159].

Figure 40. *Top*: Transport inhibition at different concentrations of Compound **1**, showing an IC_{50} value between 1 and 10 μM. The observed maximal inhibition was 80 %. *Bottom*: Chemical structure of Compound **1**

Based on the effective Compound **1**, a set of further 14 inhibitors was obtained via virtual screening. These inhibitors did not show any problems with solubility, presumably because they are similar in structure to Compound **1**. All compounds were screened at a concentration of 10 μM (Figure 41) and showed inhibition between 0 and 80%. IC_{50} values for the best three compounds **9**, **5** and **7** were determined by screening at 0.1, 1, 5 and 10 μM to be between 5 and 10 μM, i.e. in the same range as the initial hit Compound **1**. Thus these second generation inhibitors confirms the validity of the approach using virtual high-throughput screening with a medium-throughput *in vitro* assay, but also shows that improvement of inhibitor efficacy against channel proteins is not easy and most likely requires additional structural information.
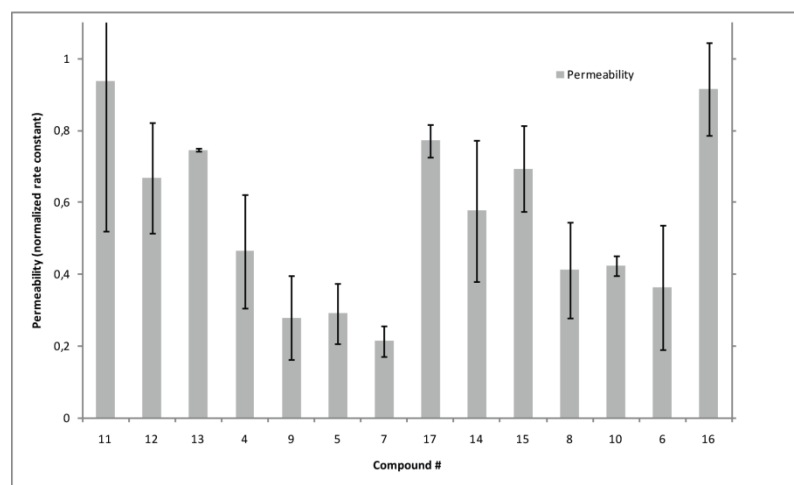
Figure 41.   Screening of inhibitors at 10 μM concentration. Relative transport rates through PfAQP-proteoliposomes indicate blocking of the channel. Rate constants were obtained from triplicates on the same set of liposomes. The compounds are displayed in the order screened. Screening order was reversed during replicate 2.

4.4 Summary

The presented method is capable of studying aquaglyceroporin function and medium-throughput inhibitor studies. In contrast to conventional methods that observe a secondary effect like size change caused by solute flux, here mass change is detected directly. The protein in question is functionally reconstituted into small lipid vesicles and the rate of solute unloading upon buffer exchange is traced via SPR. Currently limiting are the buffer exchange rates and the time resolution of the detector.

The solute transport capabilities of PfAQP have been studied, and glycerol, erythritol, xylitol and sorbitol were found to be transported through the channel at increasing rates. Activation energies measured for xylitol and sorbitol were identical within the experimental error, which – in combination with docking studies – lead to the conclusion that the transport rate is mainly determined by the stability of the sugar alcohol inside the pore.

Further development lead to a medium-throughput assay for membrane channel and transporter proteins, currently capable of measuring ca 50 compounds/day. The assay is fully automated in a 96-well format. It is important to note is that it is protein function and not inhibitor binding that is measured. As it is the large amount of sorbitol, which is released from the vesicle that is traces instead of the inhibitor itself, this leads to significant signal amplification when compared to traditional binding studies of membrane proteins. This benefit is particularly important for the studying the effects of small molecules or fragments, which otherwise would only lead to a small change in SPR signal upon binding. Applying the screening method to PfAQP in combination with high-throughput pre-screening *in silico* yielded successful inhibitors with an IC_{50} value between 1-10 μM, which have been confirmed been confirmed by an *in vivo* assay.

5   Concluding Remarks and Future Perspectives

In recent years, the field of membrane protein science has grown exponentially. With the arrival of new technologies, the scientific community has gained a better understanding of this medically relevant class of proteins. However, working with membrane protein is still a challenge, in particular when compared to soluble proteins. The reasons are obvious: Most biochemical and biophysical experiments are carried out in solution, which is the natural environment of soluble proteins. Membrane proteins on the other hand, are usually have to be removed from their natural surroundings, the biological membrane.

Pichia pastoris aquaporin

Solving the crystal structure of *P. pastoris* Aqy1 to atomic resolution was a milestone in membrane protein crystallography. At 1.15 and 0.90 Å resolution, hitherto unknown structural details of this otherwise well studied class of proteins were observed, providing experimental evidence for the otherwise elusive mechanism of proton exclusion. Through mutational studies and one of the few X-ray structures at ambient temperature of a membrane protein, we could address the mystery of elongated N-termini in yeast. These are involved in a gating mechanism, which protects the yeast from freezing and thawing, by elegant regulation *via* phosphorylation and in response to mechanical stress.

The extraordinary size and quality of the crystals obtained will allow for further experiments. Even neutron diffraction experiments, a type of experiment even rare for soluble proteins, are feasible. They will provide more exact locations of hydrogen atoms and the orientation of water molecules inside the channel. Also, X-ray crystallography experiments at different temperatures are thinkable, taking advantage of the variation in crystal water density to further corroborate the mechanism of mechanosensing.

SPR based transport and inhibitor screening assay

For assessing the potency of aquaporin inhibitors, a medium to high-throughput assay was urgently needed. To ensure protein functionality, these experiments were performed with the protein embedded in the lipid bilayer of artificial liposomes, which resemble its natural surroundings. Surface plasmon resonance technology provided a neat way to probe transport through the channel directly. The future challenges for this assay are mainly of technical nature: probing faster transport events will require the optimization of detector- and microfluidics technology. Initial experiments also showed that not only solute transport, but also water transport may be possible to observe.

Finally, transformation of this functional assay into a medium-throughput platform for the screening of aquaporin inhibitors resulted in a new drug lead against the aquaporin from the malaria parasite *Plasmodium falciparum*. The combination of virtual high-throughput- and *in vitro*-screening provides a cost-effective way for drug discovery. Further efforts in this direction has the potential to yield improved and novel drug leads against Malaria and many other diseases that involve membrane proteins.

6   Acknowledgments

7   References

1.   Guidotti, G., *Membrane proteins.* Annu Rev Biochem, 1972. **41**: p. 731-52.
2.   von Heijne, G., *The membrane protein universe: what's out there and why bother?* J Intern Med, 2007. **261**(6): p. 543-57.
3.   de Groot, B.L. and H. Grubmuller, *Water permeation across biological membranes: mechanism and dynamics of aquaporin-1 and GlpF.* Science, 2001. **294**(5550): p. 2353-7.
4.   Gomes, D., et al., *Aquaporins are multifunctional water and solute transporters highly divergent in living organisms.* Biochim Biophys Acta, 2009. **1788**(6): p. 1213-28.
5.   Pfeffer, W., *Osmotische Untersuchungen – Studien zur Zellmechanik*1877, Leipzig: Verlag von Wilhelm Engelmann.
6.   Farmer, R.E. and R.I. Macey, *Perturbation of red cell volume: rectification of osmotic flow.* Biochim Biophys Acta, 1970. **196**(1): p. 53-65.
7.   Agre, P., *The aquaporin water channels.* Proc Am Thorac Soc, 2006. **3**(1): p. 5-13.
8.   Denker, B.M., et al., *Identification, purification, and partial characterization of a novel Mr 28,000 integral membrane protein from erythrocytes and renal tubules.* J Biol Chem, 1988. **263**(30): p. 15634-42.
9.   Preston, G.M., et al., *Appearance of water channels in Xenopus oocytes expressing red cell CHIP28 protein.* Science, 1992. **256**(5055): p. 385-7.
10.   Zeidel, M.L., et al., *Reconstitution of functional water channels in liposomes containing purified red cell CHIP28 protein.* Biochemistry, 1992. **31**(33): p. 7436-40.
11.   van Hoek, A.N. and A.S. Verkman, *Functional reconstitution of the isolated erythrocyte water channel CHIP28.* J Biol Chem, 1992. **267**(26): p. 18267-9.
12.   Engel, A. and H. Stahlberg, *Aquaglyceroporins: channel proteins with a conserved core, multiple functions, and variable surfaces.* Int Rev Cytol, 2002. **215**: p. 75-104.
13.   Preston, G.M., et al., *The mercury-sensitive residue at cysteine 189 in the CHIP28 water channel.* J Biol Chem, 1993. **268**(1): p. 17-20.
14.   Preston, G.M., et al., *Membrane topology of aquaporin CHIP. Analysis of functional epitope-scanning mutants by vectorial proteolysis.* J Biol Chem, 1994. **269**(3): p. 1668-73.
15.   Jung, J.S., et al., *Molecular structure of the water channel through aquaporin CHIP. The hourglass model.* J Biol Chem, 1994. **269**(20): p. 14648-54.
16.   Walz, T., et al., *The three-dimensional structure of human erythrocyte aquaporin CHIP.* EMBO J, 1994. **13**(13): p. 2985-93.
17.   Walz, T., et al., *Biologically active two-dimensional crystals of aquaporin CHIP.* J Biol Chem, 1994. **269**(3): p. 1583-6.
18.   Mitsuoka, K., et al., *The structure of aquaporin-1 at 4.5-A resolution reveals short alpha-helices in the center of the monomer.* J Struct Biol, 1999. **128**(1): p. 34-43.
19.   Cheng, A., et al., *Three-dimensional organization of a human water channel.* Nature, 1997. **387**(6633): p. 627-30.
20.   Fu, D., et al., *Structure of a glycerol-conducting channel and the basis for its selectivity.* Science, 2000. **290**(5491): p. 481-6.
21.   Tornroth-Horsefield, S., et al., *Structural insights into eukaryotic aquaporin regulation.* Febs Journal, 2010. **277**: p. 21-21.
22.   Savage, D.F., et al., *Architecture and selectivity in aquaporins: 2.5 a X-ray structure of aquaporin Z.* PLoS Biol, 2003. **1**(3): p. E72.
23.   Jiang, J., B.V. Daniels, and D. Fu, *Crystal structure of AqpZ tetramer reveals two distinct Arg-189 conformations associated with water permeation through the narrowest constriction of the water-conducting channel.* J Biol Chem, 2006. **281**(1): p. 454-60.
24.   Fischer, G., et al., *Crystal structure of a yeast aquaporin at 1.15 angstrom reveals a novel gating mechanism.* PLoS Biol, 2009. **7**(6): p. e1000130.

25.   Newby, Z.E., et al., *Crystal structure of the aquaglyceroporin PfAQP from the malarial parasite Plasmodium falciparum.* Nat Struct Mol Biol, 2008. **15**(6): p. 619-25.

26.   Tornroth-Horsefield, S., et al., *Structural mechanism of plant aquaporin gating.* Nature, 2006. **439**(7077): p. 688-94.

27.   Hedfalk, K., et al., *Aquaporin gating.* Curr Opin Struct Biol, 2006. **16**(4): p. 447-56.

28.   Gonen, T., et al., *Aquaporin-0 membrane junctions reveal the structure of a closed water pore.* Nature, 2004. **429**(6988): p. 193-7.

29.   Gonen, T., et al., *Lipid-protein interactions in double-layered two-dimensional AQP0 crystals.* Nature, 2005. **438**(7068): p. 633-8.

30.   Harries, W.E., et al., *The channel architecture of aquaporin 0 at a 2.2-A resolution.* Proc Natl Acad Sci U S A, 2004. **101**(39): p. 14045-50.

31.   Murata, K., et al., *Structural determinants of water permeation through aquaporin-1.* Nature, 2000. **407**(6804): p. 599-605.

32.   Sui, H.X., et al., *Structural basis of water-specific transport through the AQP1 water channel.* Nature, 2001. **414**(6866): p. 872-878.

33.   Ho, J.D., et al., *Crystal structure of human aquaporin 4 at 1.8 A and its mechanism of conductance.* Proc Natl Acad Sci U S A, 2009. **106**(18): p. 7437-42.

34.   Hiroaki, Y., et al., *Implications of the aquaporin-4 structure on array formation and cell adhesion.* J Mol Biol, 2006. **355**(4): p. 628-39.

35.   Tani, K., et al., *Mechanism of aquaporin-4's fast and highly selective water conduction and proton exclusion.* J Mol Biol, 2009. **389**(4): p. 694-706.

36.   Horsefield, R., et al., *High-resolution x-ray structure of human aquaporin 5.* Proc Natl Acad Sci U S A, 2008.

37.   Chakrabarti, N., et al., *Molecular basis of proton blockage in aquaporins.* Structure, 2004. **12**(1): p. 65-74.

38.   de Groot, B.L. and H. Grubmuller, *The dynamics and energetics of water permeation and proton exclusion in aquaporins.* Curr Opin Struct Biol, 2005. **15**(2): p. 176-83.

39.   Hub, J.S., H. Grubmüller, and B.L. deGroot, *Aquaporins*, in *Handb Exp Pharmacol*, E. Beitz, Editor 2009, Springer Verlag: Berlin Heidelberg. p. 58-76.

40.   Chakrabarti, N., B. Roux, and R. Pomes, *Structural determinants of proton blockage in aquaporins.* J Mol Biol, 2004. **343**(2): p. 493-510.

41.   Burykin, A. and A. Warshel, *What really prevents proton transport through aquaporin? Charge self-energy versus proton wire proposals.* Biophys J, 2003. **85**(6): p. 3696-706.

42.   Burykin, A. and A. Warshel, *On the origin of the electrostatic barrier for proton transport in aquaporin.* FEBS Lett, 2004. **570**(1-3): p. 41-6.

43.   Chen, H., Y. Wu, and G.A. Voth, *Origins of proton transport behavior from selectivity domain mutations of the aquaporin-1 channel.* Biophys J, 2006. **90**(10): p. L73-5.

44.   de Groot, B.L., et al., *The mechanism of proton exclusion in the aquaporin-1 water channel.* J Mol Biol, 2003. **333**(2): p. 279-93.

45.   Ilan, B., et al., *The mechanism of proton exclusion in aquaporin channels.* Proteins, 2004. **55**(2): p. 223-8.

46.   Kato, M., A.V. Pisliakov, and A. Warshel, *The barrier for proton transport in aquaporins as a challenge for electrostatic models: the role of protein relaxation in mutational calculations.* Proteins, 2006. **64**(4): p. 829-44.

47.   Li, H., et al., *Enhancement of Proton Conductance by Mutations of the Selectivity Filter of Aquaporin-1.* J Mol Biol, 2011.

48.   Hub, J.S. and B.L. de Groot, *Mechanism of selectivity in aquaporins and aquaglyceroporins.* Proc Natl Acad Sci U S A, 2008.

49.   Beitz, E., *Aquaporins from pathogenic protozoan parasites: structure, function and potential for chemotherapy.* Biology of the Cell, 2005. **97**(6): p. 373-383.

50.   Beitz, E., *Aquaporin water and solute channels from malaria parasites and other pathogenic protozoa.* ChemMedChem, 2006. **1**(6): p. 587-92.

51. King, L.S., D. Kozono, and P. Agre, *From structure to disease: the evolving tale of aquaporin biology.* Nat Rev Mol Cell Biol, 2004. **5**(9): p. 687-98.

52. Hub, J.S. and B.L. de Groot, *Does CO2 permeate through aquaporin-1?* Biophys J, 2006. **91**(3): p. 842-8.

53. Robben, J.H., et al., *Intracellular activation of vasopressin V2 receptor mutants in nephrogenic diabetes insipidus by nonpeptide agonists.* Proc Natl Acad Sci U S A, 2009. **106**(29): p. 12195-200.

54. Conner, M.T., et al., *Membrane trafficking of aquaporin 1 is mediated by protein kinase C via microtubules and regulated by tonicity.* Biochemistry, 2010. **49**(5): p. 821-3.

55. Han, Z. and R.V. Patil, *Protein kinase A-dependent phosphorylation of aquaporin-1.* Biochem Biophys Res Commun, 2000. **273**(1): p. 328-32.

56. de Mattia, F., et al., *Lack of arginine vasopressin-induced phosphorylation of aquaporin-2 mutant AQP2-R254L explains dominant nephrogenic diabetes insipidus.* J Am Soc Nephrol, 2005. **16**(10): p. 2872-80.

57. McDill, B.W., et al., *Congenital progressive hydronephrosis (cph) is caused by an S256L mutation in aquaporin-2 that affects its phosphorylation and apical membrane accumulation.* Proc Natl Acad Sci U S A, 2006. **103**(18): p. 6952-7.

58. Yang, F., J.D. Kawedia, and A.G. Menon, *Cyclic AMP regulates aquaporin 5 expression at both transcriptional and post-transcriptional levels through a protein kinase A pathway.* J Biol Chem, 2003. **278**(34): p. 32173-80.

59. Ishikawa, Y., et al., *Molecular mechanisms and drug development in aquaporin water channel diseases: The translocation of aquaporin-5 from lipid rafts to the apical plasma membranes of parotid glands of normal rats and the impairment of it in diabetic or aged rats.* Journal of Pharmacological Sciences, 2004. **96**(3): p. 271-275.

60. Kosugi-Tanaka, C., et al., *Protein kinase A-regulated membrane trafficking of a green fluorescent protein-aquaporin 5 chimera in MDCK cells.* Biochim Biophys Acta, 2006. **1763**(4): p. 337-44.

61. Garcia, F., et al., *The water channel aquaporin-8 is mainly intracellular in rat hepatocytes, and its plasma membrane insertion is stimulated by cyclic AMP.* J Biol Chem, 2001. **276**(15): p. 12147-52.

62. Nyblom, M., et al., *Structural and functional analysis of SoPIP2;1 mutants adds insight into plant aquaporin gating.* J Mol Biol, 2009. **387**(3): p. 653-68.

63. Nemeth-Cahalan, K.L. and J.E. Hall, *pH and calcium regulate the water permeability of aquaporin 0.* The Journal of biological chemistry, 2000. **275**(10): p. 6777-82.

64. Zelenina, M., et al., *Water permeability of aquaporin-4 is decreased by protein kinase C and dopamine.* American journal of physiology. Renal physiology, 2002. **283**(2): p. F309-18.

65. Tanghe, A., P. Van Dijck, and J.M. Thevelein, *Why do microorganisms have aquaporins?* Trends in Microbiology, 2006. **14**(2): p. 78-85.

66. Thorsen, M., et al., *The MAPK Hog1p modulates Fps1p-dependent arsenite uptake and tolerance in yeast.* Mol Biol Cell, 2006. **17**(10): p. 4400-10.

67. Castle, N.A., *Aquaporins as targets for drug discovery.* Drug Discovery Today, 2005. **10**(7): p. 485-493.

68. Morishita, Y., et al., *Molecular mechanisms and drug development in aquaporin water channel diseases: aquaporin superfamily (superaquaporins): expansion of aquaporins restricted to multicellular organisms.* J Pharmacol Sci, 2004. **96**(3): p. 276-9.

69. Amiry-Moghaddam, M. and O.P. Ottersen, *The molecular basis of water transport in the brain.* Nat Rev Neurosci, 2003. **4**(12): p. 991-1001.

70. Zelenina, M., *Regulation of brain aquaporins.* Neurochem Int, 2010. **57**(4): p. 468-88.

71. Niemietz, C.M. and S.D. Tyerman, *New potent inhibitors of aquaporins: silver and gold compounds inhibit aquaporins of plant and human origin.* FEBS Lett, 2002. **531**(3): p. 443-7.

72. Goffeau, A., *Four years of post-genomic life with 6,000 yeast genes.* FEBS Lett, 2000. **480**(1): p. 37-41.

73. Goffeau, A., et al., *Life with 6000 genes.* Science, 1996. **274**(5287): p. 546, 563-7.

74. Villar, C.C. and A. Dongari-Bagtzoglou, *Immune defence mechanisms and immunoenhancement strategies in oropharyngeal candidiasis.* Expert Rev Mol Med, 2008. **10**: p. e29.

75. Rubin, R.H., *Fungal and bacterial infections in the immunocompromised host.* Eur J Clin Microbiol Infect Dis, 1993. **12 Suppl 1**: p. S42-8.

76. Guilliermond, C.H., Seances, Mem. Soc. Biol., 1919. **82**: p. 466-470.

77. Higgins, D.R. and J.M. Cregg, *Introduction to Pichia pastoris.* Methods Mol Biol, 1998. **103**: p. 1-15.

78. Cregg, J.M., et al., *Recombinant protein expression in Pichia pastoris.* Mol Biotechnol, 2000. **16**(1): p. 23-52.

79. Cregg, J.M., et al., *Functional characterization of the two alcohol oxidase genes from the yeast Pichia pastoris.* Mol Cell Biol, 1989. **9**(3): p. 1316-23.

80. Cregg, J., *Sire Selective Genomix Modification of Yeast of the Genus Pichia*, 1989, Phillips Petroleum Company, Bartesville, Okla.: USA.

81. Daly, R. and M.T. Hearn, *Expression of heterologous proteins in Pichia pastoris: a useful experimental tool in protein engineering and production.* J Mol Recognit, 2005. **18**(2): p. 119-38.

82. Jacobs, P.P., et al., *Engineering complex-type N-glycosylation in Pichia pastoris using GlycoSwitch technology.* Nat Protoc, 2009. **4**(1): p. 58-70.

83. Cereghino, J.L. and J.M. Cregg, *Heterologous protein expression in the methylotrophic yeast Pichia pastoris.* FEMS Microbiol Rev, 2000. **24**(1): p. 45-66.

84. Waterham, H.R., et al., *Isolation of the Pichia pastoris glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter.* Gene, 1997. **186**(1): p. 37-44.

85. Bushell, M.E., et al., *Cyclic fed-batch culture for production of human serum albumin in Pichia pastoris.* Biotechnol Bioeng, 2003. **82**(6): p. 678-83.

86. Bruce-Chwatt, L.J., *Alphonse Laveran's discovery 100 years ago and today's global fight against malaria.* J R Soc Med, 1981. **74**(7): p. 531-6.

87. Mueller, I., et al., *Key gaps in the knowledge of Plasmodium vivax, a neglected human malaria parasite.* Lancet Infect Dis, 2009. **9**(9): p. 555-66.

88. Wells, T.N.C., P.L. Alonso, and W.E. Gutteridge, *New medicines to improve control and contribute to the eradication of malaria.* Nature Reviews Drug Discovery, 2009. **8**(11): p. 879-891.

89. Feachem, R.G., et al., *Shrinking the malaria map: progress and prospects.* Lancet, 2010. **376**(9752): p. 1566-78.

90. Fidock, D.A., et al., *Antimalarial drug discovery: Efficacy models for compound screening.* Nature Reviews Drug Discovery, 2004. **3**(6): p. 509-520.

91. Shah, N.K., et al., *Antimalarial drug resistance of Plasmodium falciparum in India: changes over time and space.* Lancet Infect Dis, 2011. **11**(1): p. 57-64.

92. Hedfalk, K., et al., *Production, characterization and crystallization of the Plasmodium falciparum aquaporin.* Protein Expr Purif, 2008. **59**(1): p. 69-78.

93. Newby, Z.E., et al., *A general protocol for the crystallization of membrane proteins for X-ray structural investigation.* Nat Protoc, 2009. **4**(5): p. 619-37.

94. Edebo, L., *A New Press for the Disruption of Micro-Organisms and Other Cells.* Journal of Biochemical and Microbiological Technology and Engineering, 1960. **2**(4): p. 453-479.

95. Schobert, B., et al., *Crystallographic structure of the K intermediate of bacteriorhodopsin: conservation of free energy after photoisomerization of the retinal.* J Mol Biol, 2002. **321**(4): p. 715-26.

96. Massa, W., *Kristallstrukturbestimmung.* 4th ed2005: B.G. Teubner.

97. Mimura, H., et al., *Image quality improvement in a hard X-ray projection microscope using total reflection mirror optics.* Journal of synchrotron radiation, 2004. **11**(Pt 4): p. 343-6.

98. Asherie, N., *Protein crystallization and phase diagrams.* Methods, 2004. **34**(3): p. 266-72.

99.   Kendrew, J.C., et al., *A three-dimensional model of the myoglobin molecule obtained by x-ray analysis.* Nature, 1958. **181**(4610): p. 662-6.

100.   Deisenhofer, J., et al., *X-ray structure analysis of a membrane protein complex. Electron density map at 3 A resolution and a model of the chromophores of the photosynthetic reaction center from Rhodopseudomonas viridis.* J Mol Biol, 1984. **180**(2): p. 385-98.

101.   Landau, E.M. and J.P. Rosenbusch, *Lipidic cubic phases: a novel concept for the crystallization of membrane proteins.* Proc Natl Acad Sci U S A, 1996. **93**(25): p. 14532-5.

102.   Wohri, A.B., et al., *A lipidic-sponge phase screen for membrane protein crystallization.* Structure, 2008. **16**(7): p. 1003-9.

103.   Faham, S. and J.U. Bowie, *Bicelle crystallization: a new method for crystallizing membrane proteins yields a monomeric bacteriorhodopsin structure.* J Mol Biol, 2002. **316**(1): p. 1-6.

104.   Flot, D., et al., *The ID23-2 structural biology microfocus beamline at the ESRF.* J Synchrotron Radiat, 2010. **17**(1): p. 107-18.

105.   Mosselmans, J.F., et al., *I18--the microfocus spectroscopy beamline at the Diamond Light Source.* J Synchrotron Radiat, 2009. **16**(Pt 6): p. 818-24.

106.   Mammen, C.B., et al., *Design of a 5-station macromolecular crystallography beamline at MAX-Lab.* Acta Physica Polonica A, 2002. **101**(5): p. 595-602.

107.   Broennimann, C., et al., *The PILATUS 1M detector.* J Synchrotron Radiat, 2006. **13**(Pt 2): p. 120-30.

108.   Johari, G.P., A. Hallbrucker, and E. Mayer, *The Glass Liquid Transition of Hyperquenched Water.* Nature, 1987. **330**(6148): p. 552-553.

109.   Garman, E.F. and R.L. Owen, *Cryocooling and radiation damage in macromolecular crystallography.* Acta Crystallogr D Biol Crystallogr, 2006. **62**(Pt 1): p. 32-47.

110.   Garman, E., *'Cool' crystals: macromolecular cryocrystallography and radiation damage.* Curr Opin Struct Biol, 2003. **13**(5): p. 545-51.

111.   Barker, A.I., et al., *Room-temperature scavengers for macromolecular crystallography: increased lifetimes and modified dose dependence of the intensity decay.* J Synchrotron Radiat, 2009. **16**(Pt 2): p. 205-16.

112.   Wohri, A.B., et al., *Light-induced structural changes in a photosynthetic reaction center caught by Laue diffraction.* Science, 2010. **328**(5978): p. 630-3.

113.   Afonine, P.V., et al., *Joint X-ray and neutron refinement with phenix.refine.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 11): p. 1153-63.

114.   Kabsch, W., *Automatic-Indexing of Rotation Diffraction Patterns.* Journal of Applied Crystallography, 1988. **21**: p. 67-71.

115.   Leslie, A.G.W., *Recent changes to the MOSFLM package for processing film and image plate data.* Joint CCP4 and ESF-EAMCB Newsletter on Protein Crystallography, 1992(26).

116.   Evans, P., *Scaling and assessment of data quality.* Acta Crystallogr D Biol Crystallogr, 2006. **62**(Pt 1): p. 72-82.

117.   Dauter, Z., *New approaches to high-throughput phasing.* Curr Opin Struct Biol, 2002. **12**(5): p. 674-8.

118.   Hauptman, H., *Phasing methods for protein crystallography.* Curr Opin Struct Biol, 1997. **7**(5): p. 672-80.

119.   Rodriguez, D.D., et al., *Crystallographic ab initio protein structure solution below atomic resolution.* Nat Methods, 2009. **6**(9): p. 651-3.

120.   McCoy, A.J., *Solving structures of protein complexes by molecular replacement with Phaser.* Acta Crystallogr D Biol Crystallogr, 2007. **63**(Pt 1): p. 32-41.

121.   Vagin, A. and A. Teplyakov, *MOLREP: an automated program for molecular replacement.* Journal of Applied Crystallography, 1997. **30**: p. 1022-1025.

122.   Long, F., et al., *BALBES: a molecular-replacement pipeline.* Acta Crystallogr D Biol Crystallogr, 2008. **64**(Pt 1): p. 125-32.

123.   Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 12 Pt 1): p. 2126-32.

124.   Cowtan, K., *The Buccaneer software for automated model building. 1. Tracing protein chains.* Acta Crystallogr D Biol Crystallogr, 2006. **62**(Pt 9): p. 1002-11.

125.   Perrakis, A., et al., *ARP/wARP and molecular replacement.* Acta Crystallogr D Biol Crystallogr, 2001. **57**(Pt 10): p. 1445-50.

126.   Sheldrick, G.M., *A short history of SHELX.* Acta Crystallogr A, 2008. **64**(Pt 1): p. 112-22.

127.   Murshudov, G.N., A.A. Vagin, and E.J. Dodson, *Refinement of macromolecular structures by the maximum-likelihood method.* Acta Crystallogr D Biol Crystallogr, 1997. **53**(Pt 3): p. 240-55.

128.   Zwart, P.H., et al., *Automated structure solution with the PHENIX suite.* Methods Mol Biol, 2008. **426**: p. 419-35.

129.   Brunger, A.T., et al., *Crystallography & NMR system: A new software suite for macromolecular structure determination.* Acta Crystallogr D Biol Crystallogr, 1998. **54**(Pt 5): p. 905-21.

130.   Ramachandran, G.N. and V. Sasisekharan, *Conformation of polypeptides and proteins.* Adv Protein Chem, 1968. **23**: p. 283-438.

131.   Kozono, D., et al., *Functional expression and characterization of an archaeal aquaporin. AqpM from methanothermobacter marburgensis.* J Biol Chem, 2003. **278**(12): p. 10649-56.

132.   Mayer, L.D., M.J. Hope, and P.R. Cullis, *Vesicles of variable sizes produced by a rapid extrusion procedure.* Biochim Biophys Acta, 1986. **858**(1): p. 161-8.

133.   Geertsma, E.R., et al., *Membrane reconstitution of ABC transporters and assays of translocator function.* Nat Protoc, 2008. **3**(2): p. 256-66.

134.   Levy, D., et al., *Reconstitution of the sarcoplasmic reticulum Ca(2+)-ATPase: mechanisms of membrane protein insertion into liposomes during reconstitution procedures involving the use of detergents.* Biochim Biophys Acta, 1992. **1107**(2): p. 283-98.

135.   Rigaud, J.L., B. Pitard, and D. Levy, *Reconstitution of Membrane-Proteins into Liposomes - Application to Energy-Transducing Membrane-Proteins.* Biochimica Et Biophysica Acta-Bioenergetics, 1995. **1231**(3): p. 223-246.

136.   Aslund, I., et al., *Filter-exchange PGSE NMR determination of cell membrane permeability.* J Magn Reson, 2009. **200**(2): p. 291-5.

137.   Daniels, M.J., M.R. Wood, and M. Yeager, *In vivo functional assay of a recombinant aquaporin in Pichia pastoris.* Applied and Environmental Microbiology, 2006. **72**(2): p. 1507-1514.

138.   Rich, R.L. and D.G. Myszka, *Survey of the year 2007 commercial optical biosensor literature.* J Mol Recognit, 2008. **21**(6): p. 355-400.

139.   Homola, J., *Surface plasmon resonance sensors for detection of chemical and biological species.* Chem Rev, 2008. **108**(2): p. 462-93.

140.   Liedberg, B., C. Nylander, and I. Lundstrom, *Biosensing with surface plasmon resonance--how it all started.* Biosens Bioelectron, 1995. **10**(8): p. i-ix.

141.   Stenberg, E., et al., *Quantitative-Determination of Surface Concentration of Protein with Surface-Plasmon Resonance Using Radiolabeled Proteins.* Journal of Colloid and Interface Science, 1991. **143**(2): p. 513-526.

142.   De Schutter, K., et al., *Genome sequence of the recombinant protein production host Pichia pastoris.* Nat Biotechnol, 2009. **27**(6): p. 561-6.

143.   Mattanovich, D., et al., *Open access to sequence: browsing the Pichia pastoris genome.* Microb Cell Fact, 2009. **8**: p. 53.

144.   Pettersson, N., et al., *Aquaporins in yeasts and filamentous fungi.* Biol Cell, 2005. **97**(7): p. 487-500.

145.   Bhagwat, M. and L. Aravind, *PSI-BLAST Tutorial*, in *Comparative Genomics: Volumes 1 and 2*, N.H. Bergman, Editor 2007: Totowa (NJ).

146.   Smart, O.S., J.M. Goodfellow, and B.A. Wallace, *The pore dimensions of gramicidin A.* Biophys J, 1993. **65**(6): p. 2455-60.

147.   Lindkvist, K., et al., *Modified Pichia Strain*, 2008: Sweden. Patent.

148. Blom, N., S. Gammeltoft, and S. Brunak, *Sequence and structure-based prediction of eukaryotic protein phosphorylation sites.* J Mol Biol, 1999. **294**(5): p. 1351-62.

149. Jeanteur, D., et al., *Structural and functional alterations of a colicin-resistant mutant of OmpF porin from Escherichia coli.* Proc Natl Acad Sci U S A, 1994. **91**(22): p. 10675-9.

150. Garman, E.F. and C. Nave, *Radiation damage in protein crystals examined under various conditions by different methods.* J Synchrotron Radiat, 2009. **16**(Pt 2): p. 129-32.

151. Cruickshank, D.W., *Remarks about protein structure precision.* Acta Crystallogr D Biol Crystallogr, 1999. **55**(Pt 3): p. 583-601.

152. Juers, D.H. and B.W. Matthews, *Reversible lattice repacking illustrates the temperature dependence of macromolecular interactions.* J Mol Biol, 2001. **311**(4): p. 851-62.

153. Laize, V., et al., *Polymorphism of Saccharomyces cerevisiae aquaporins.* Yeast, 2000. **16**(10): p. 897-903.

154. Levitt, D.G. and H.J. Mlekoday, *Reflection coefficient and permeability of urea and ethylene glycol in the human red cell membrane.* J Gen Physiol, 1983. **81**(2): p. 239-53.

155. Beitz, E., et al., *Molecular dissection of water and glycerol permeability of the aquaglyceroporin from Plasmodium falciparum by mutational analysis.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(5): p. 1153-1158.

156. Hansen, M., et al., *A single, bi-functional aquaglyceroporin in blood-stage Plasmodium falciparum malaria parasites.* J Biol Chem, 2002. **277**(7): p. 4874-82.

157. Pavlovic-Djuranovic, S., J.E. Schultz, and E. Beitz, *A single aquaporin gene encodes a water/glycerol/urea facilitator in Toxoplasma gondii with similarity to plant tonoplast intrinsic proteins.* Febs Letters, 2003. **555**(3): p. 500-504.

158. Vial, H.J., et al., *Phospholipid metabolism in Plasmodium-infected erythrocytes: guidelines for further studies using radioactive precursor incorporation.* Parasitology, 1989. **98 Pt 3**: p. 351-7.

159. Promeneur, D., et al., *Aquaglyceroporin PbAQP during intraerythrocytic development of the malaria parasite Plasmodium berghei.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(7): p. 2211-2216.

160. Cosconati, S., et al., *Virtual screening with AutoDock: theory and practice.* Expert Opinion on Drug Discovery. **5**(6): p. 597-607.

161. Abdiche, Y.N. and D.G. Myszka, *Probing the mechanism of drug/lipid membrane interactions using Biacore.* Anal Biochem, 2004. **328**(2): p. 233-43.

162. Baird, C.L., E.S. Courtenay, and D.G. Myszka, *Surface plasmon resonance characterization of drug/liposome interactions.* Anal Biochem, 2002. **310**(1): p. 93-9.

163. Frostell-Karlsson, A., et al., *Biosensor analysis of the interaction between drug compounds and liposomes of different properties; a two-dimensional characterization tool for estimation of membrane absorption.* J Pharm Sci, 2005. **94**(1): p. 25-37.

164. Buchwald, P. and N. Bodor, *Octanol-water partition of nonzwitterionic peptides: predictive power of a molecular size-based model.* Proteins, 1998. **30**(1): p. 86-99.