

THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN NATURAL SCIENCE

**PRODUCTION AND CHARACTERISATION OF
AQUAPORINS AND PROTON-TRANSLOCATING
TRANSHYDROGENASE**

MIKAEL EKVALL



UNIVERSITY OF GOTHENBURG

Department of Chemistry and Molecular Biology
Gothenburg, Sweden
2013

Thesis for the Degree of Doctor of Philosophy in Natural Science

Production and Characterisation of Aquaporins and
Proton-Translocating Transhydrogenase

© Mikael Ekvall, 2013

ISBN: 978-91-628-8756-8

Available online: <http://hdl.handle.net/2077/34322>

University of Gothenburg
Department of Chemistry and Molecular Biology
Lundberg laboratory, Medicinaregatan 9E
SE-413 90 Gothenburg
Sweden

Cover: A montage representing all aspects in the thesis: The structure and channel of open S188E SoPIP₂;1, the ZTH-wt pPICZB construct, a Western blot and a S188E crystal, all lying on top of a nucleotide sequence.

Printed by Ineko AB
Göteborg, Sverige, 2013

*To my beautiful daughters
Victoria, Jasmine & Alice*

“Reality is frequently inaccurate.” – Douglas Adams

Dissertation Abstract

Water transport in eukaryotic cells is a highly regulated and fine-tuned process. Water channel protein known as aquaporins (AQPs), constitute the main cellular water transport system, preserving water homeostasis by maintaining specific selectivity-mechanisms. Dysfunctional AQPs induce a wide variety of diseases in humans thereby enhancing the clinical significance of structural and functional knowledge.

Macromolecular structural research requires large amounts of pure, stable protein to initiate crystallization trials. To achieve this, genetic engineering and overproduction systems such as the methylotrophic yeast *Pichia pastoris* (*P. pastoris*) are employed. The next step, crystallization, involves arrangement of the macromolecule in a repetitive fashion. Once crystals have been obtained, these are exposed by synchrotron radiation (X-rays), producing a diffraction pattern. This reciprocal representation of the arrangement of atoms in the unit cell is converted back to real space by a Fourier transform, which generates a atomic model of the protein.

This thesis is based on a comparative study of the production levels of all human AQPs, an production and purification analysis of eukaryotic transhydrogenases, and a structural and functional investigation of the spinach AQP SoPIP2;1 with associated mutants.

The human AQPs produced in the study displayed a considerable variety in production yield. Although the production yield seemed to depend on multiple factors, a correlation could be drawn between the extent of protein inserted into the membrane and phylogenetic relationship, providing further insight into eukaryotic membrane protein production. Furthermore, zebrafish transhydrogenase was successfully produced in *P. pastoris*, but although the production yield was sufficient, further optimisation of purification conditions is required in order to obtain sample suitable for crystallization. Finally, crystal structures and water transport assays of SoPIP2;1 phosphomimicking mutants as well as of SoPIP2;1 in complex with mercury have given novel insights into the mechanism of plant AQP gating.

Keywords: aquaporins, transhydrogenase, membrane proteins, X-ray crystallography, structure, protein production, overproduction.

List of Publications

This thesis is based on the following publications and are referred to by Roman numerals:

- I Öberg F, Ekvall M, Nyblom M, Backmark A, Neutze R, Hedfalk K. *Insight into factors directing high production of eukaryotic membrane proteins; production of 13 human AQP_s in Pichia pastoris*. Mol Membr Biol, 2009. 26(4): p. 215-27.
- II Nyblom M, Frick A, Wang Y, Ekvall M, Hallgren K, Hedfalk K, Neutze R, Tajkhorshid E, Törnroth-Horsefield S. *Structural and functional analysis of SoPIP₂;1 mutants adds insight into plant aquaporin gating*. J Mol Biol, 2009. 387(3): p. 653-68.
- III Frick A, Järvå M, Ekvall M, Uzdavinys P, Nyblom M, Törnroth-Horsefield S. *Mercury increases water permeability of a plant aquaporin through a non-cysteine-related mechanism*. Biochem J, 2013. 454(3): p. 491-9.
- IV Ekvall M, Frick A, Uzdavinys P, Törnroth-Horsefield S. *Crystal structure of a phosphomimicking spinach aquaporin mutant reveals channel opening*. Manuscript
- V Ekvall M, Sharma A, Törnroth-Horsefield S. *Overproduction and purification of wild-type and cysteine-free Proton-Translocating Transhydrogenase from Zebrafish*. Manuscript

Contribution Report

Referred to publications by Roman numerals:

- I I performed the cloning, production, purification and participated in the analysis of the constructs. I was involved in the quantification and interpretation of the results.
- II I was involved in the protein production and purification. I also took a major part of analyzing and quantifying the functional data.
- III I prepared protein and carried out some of the functional assays.
- IV I performed the production, purification and crystallization of the protein. I was involved in collecting structural data and solving the structure. I took part in writing the manuscript and prepared all the figures.
- V I was involved in planning the project and performed the design, cloning, production and purification of the constructs. I took major part of the interpretation of the results and in preparing the figures and writing the paper.

Table of Contents

1. Introduction	1
1.1 The Biological Membrane.....	1
1.2 Membrane Proteins	3
1.3 Structural Biology	3
1.3.1 Protein Structure and Function	4
1.3.2 Membrane Protein Structural Biology	5
1.3.3 The Problematic Nature of Membrane Protein Purification and Crystallization	5
1.4 Aquaporins	7
1.4.1 Human Aquaporins	7
1.4.2 Plant Aquaporins.....	9
1.4.3 Structural Features of Aquaporins	10
1.4.4 Aquaporin Regulation and Gating	11
1.5 Proton-Translocating Nicotinamide Transhydrogenase	13
1.5.1 Structural Features of TH.....	14
1.6 Scope of the Thesis	16
2. Methodology	17
2.1 From Theory to Protein	17
2.1.1 Cloning Strategy	17
2.1.2 Protein Production Organisms	17
2.1.3 Overproducing Proteins Using <i>Pichia Pastoris</i>	18
2.1.2 Cultivation of <i>Pichia Pastoris</i>	18
2.1.4 Membrane Protein Purification	19
2.2 X-ray Crystallography.....	21
2.2.1 Protein Crystallization	21
2.2.2 X-rays.....	22
2.2.3 X-ray Crystal Diffraction.....	23
2.2.4 The Phase Problem	24

2.2.5	Twining	25
2.2.6	The Data Collection Experiment	25
2.2.7	Data Processing, Refinement and Validation	26
2.3	Functional Assays	28
2.3.1	Liposome Assays	28
2.3.2	Spectroscopic Assays	29
3.	Results and Discussion	30
3.1	Membrane Protein Production in <i>Pichia pastoris</i>	30
3.1.1	Production of human AQPs (Paper I)	30
3.1.2	Production of Zebrafish (<i>Danio rerio</i>) TH (Paper V)	35
3.1.3	Comparative Aspects of hAQPs and TH Production	37
3.1.4	Production of SoPIP ₂ ;1	38
3.2	Structural and Functional Analysis of SoPIP ₂ ;1 (Paper II, III and IV)	38
3.2.1	Gating of SoPIP ₂ ;1 by Phosphorylation	39
3.2.2	Mimicking the Biological On/Off Switch	40
3.2.3	Mimicking Phosphorylation of Ser115 and Ser274 Does Not Cause Channel Opening	41
3.2.4	Significant Increase in Water Flux by the S188E mutant	42
3.2.5	Structure of S188E mutant is Open	44
3.2.6	The Effect of Cations on SoPIP ₂ ;1	46
4.	Conclusions	48
5.	Acknowledgments	50
6.	References	52

Abbreviations and Nomenclature

AcPyAD(H)	3-Acetylpyridine Adenine Dinucleotide*
AOX1	Alcohol Oxidase 1
AQGP	Aqua-glyceroporin
AQP	Aquaporin
β -OG	n-octyl- β -D-glucopyranoside
CMC	Critical Micelle Concentration
DDM	n-Dodecyl- β -D-Maltopyranoside
$\Delta\rho$	Proton Electrochemical Gradient
FT	Fourier Transform function
hAQP	Human Aquaporin
IMAC	Immobilized Metal Affinity Chromatography
kDa	Kilo Dalton
MD	Molecular Dynamic
MIP	Mayor Intrinsic Protein
MW	Molecular Weight
NAD(H)	Nicotinamide Adenine Dinucleotide*
NADP(H)	Nicotinamide Adenine Dinucleotide Phosphate*
NIP	Nodulin-26 like Intrinsic Protein
OD ₆₀₀	Optical Density at Wavelength 600nm
pAQP	Plant Aquaporin
PDB	Protein Data Bank
P_f	Water permeability coefficient
PIP	Plasma membrane Intrinsic Protein
ROS	Reactive Oxygen Species
SDS-PAGE	Sodium-dodesylsulphate polyacrylamid gelelectrophoresis
SIP	Small basic Intrinsic Protein
TH	Proton Translocating Nicotinamide Transhydrogenase
TIP	Tonoplast Intrinsic Protein
TMH	Trans-Membrane Helix
WT	Wild-type
XIP	X Intrinsic Protein
Å	Ångström (10^{-10} m)

* Denoting both oxidized and reduced state of the substrate.

1. Introduction

The immense diversity we all are part of today rose through evolutionary natural selection ^[1]. As much as living organisms themselves diverse, their biochemical processes showcase a vast variation in both complexity and characteristics. Despite the variation, there is a universal analogous regularity which originates back to when life on earth first started to exist. Biogenetic carbon-based life is believed to have started as the first self-replicating molecules became enclosed and thus separated from the surrounding environment ^[2]. Subsequently, the intracellular and extracellular differentiation drove the evolution selectively towards responsive molecular regulation mechanisms, acting both on the inside and across the membrane ^[3]. Accordingly, all living organisms share the same basic fundamental formula in order to store genetic information, metabolize and function.

Numerous assorted biochemical mechanisms constantly oversee and adjust the living organism at a molecular level, to continually adapt to present conditions. These mechanisms are coupled directly or indirectly to each other, collectively constituting a complex molecular biologic system ^[4]. However, despite having mechanisms to correct early biochemical errors, aberrations happen, possibly leading to diseases, some of which modern medicine haven't yet found a cure for. In spite of that, advances in structural biology may contribute to finding the answers to effectively cure and treat these diseases by providing deeper insight in effected biochemical mechanisms and explain how the biomolecular components function at an atomic resolution, ergo attain controllability.

1.1 The Biological Membrane

All eukaryotic cells, including most organelles, consist of a semi-permeable plasma membrane acting as a stable differentiating shell to its outer environment. Nevertheless, cells are far more than apathetic "bubbles" floating around. In fact, cells are highly active in cell-to-cell communication ^[5]. In addition, the defined compartmentalization accommodated by the plasma and organelle membranes, grant specific chemical settings, required for given reactions to take place. Among other things, the compartmentalization grants prokaryotic and eukaryotic

1. INTRODUCTION

biological systems to utilize membrane inserted proton-pumps to transfer protons (H^+) against a gradient and across the membrane to establish an electrochemical proton gradient ($\Delta\rho$) and protonmotive force, ultimately the driving force of adenosine-5'-triphosphate (ATP) synthesis^[6, 7].

The fundamental matrix of the cell membrane is the lipid bilayer which is made-up by lipids organized in two layers with the acyl chains facing each other, forming a hydrophobic centre-core while the lipids hydrophilic head-groups face the surfaces. Along with various types of lipids, proteins and carbohydrates are inserted or anchored through-out the lipid bilayer (Fig.1)^[8].

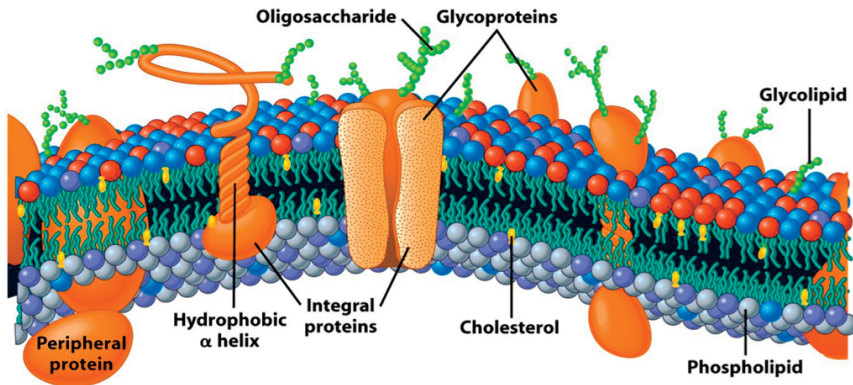


Figure 1. Schematic drawing representing a section of a typical eukaryotic cell membrane bilayer [9].

Assorted ratios of specific lipids, proteins and carbohydrates jointly dictate the characteristics of the cell membrane and hence influence the chemical environment of the cells cytosol, depending on cell type and biological location^[8, 10]. Cell membranes are highly dynamic and asymmetric, opposing the common misleading cartoons of a static plasma membrane found in the literature. The Fluidic Mosaic Model (FMM) describes the organization of cell membranes as such^[11, 12]. Accordingly, the dynamic fluidity of the cell membrane smoothly allows rotational and lateral movements of constituents in the lipid bilayer, explaining how asymmetrical membrane regions such as lipid rafts and assemblies of proteins may emerge^[13]. In addition, the FMM also state the more rare but viable transverse diffusion, commonly known as flip-flop movement, which is the transfer of molecules between the two layers of the lipid bilayer.

1. INTRODUCTION

1.2 Membrane Proteins

Membrane proteins constitute a class of proteins which are associated to the cell membrane either as integral (trans-membrane), peripheral or lipid-anchored proteins ^[14, 15]. As key components to the cell, membrane proteins respond to biochemical changes in the local environment, accounting for cellular signal reception and regulation whilst acting as transporters, channels and receptors ^[16, 17].

In the highly dynamic and quickly responding biological system, molecules need to be passively or actively transported to the inside or outside of the cell. Molecules may in some cases penetrate the cell membrane by a diffusion process but more commonly react directly or indirectly with a substrate specific membrane protein receptor, activating a cascade reaction which results in a biological response ^[10, 14, 15]. Accordingly, the biological system cannot solely rely on concentration gradients as driving force as it will not be effective when a net equilibrium is reached. Moreover, the rate of diffusion is often not quick enough to respond to rapid biochemical changes without having an extremely high concentration on one side. Instead, literal substrate specific membrane protein channels and transporters enable gated openings through the lipid bilayer, allowing quick and precise transfers ^[18]. Often, membrane protein channels and transporters include gating mechanisms, activation sites and/or substrate selectivity features, enabling accurate regulation ^[14, 15, 19].

1.3 Structural Biology

Structural biology incorporates biochemistry, biophysics and molecular biology (Fig.2) toward a better understanding in terms of biological macromolecular three-dimensional structure and function.

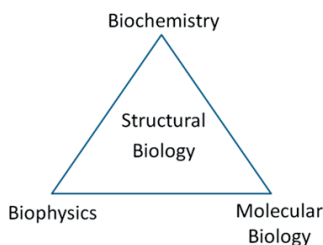


Figure 2. Schematic drawing representing structural biology as an interdisciplinary field constituted by biophysics, biochemistry and molecular biology.

1. INTRODUCTION

The value and usefulness by obtaining macromolecular models at high atomic resolution ($\leq 3\text{\AA}$) stretches beyond the human inquisitive nature as it unveils the proteins function and mechanics [20, 21]. Biological macromolecules such as proteins influence most biochemical processes, consequently, most of our known diseases involves dysfunctional proteins in some way [22-24]. Generally, most drugs used today are used without knowledge of the targets and drugs pharmacodynamics, causing inferior drug utilization [25]. The advantages by knowing a drug targets molecular mechanism and atomic arrangement will not only grant efficient drug design but also provide useful information in the quest to extensively comprehend the whole complex biological machinery.

1.3.1 Protein Structure and Function

"If it looks like a chair, it is most likely that its main purpose is to function as a sitting device" (M. Ekvall). This structure/function relationship is something we interact with on a daily basis but it also stands true on a molecular level. The fact that a proteins biological function has direct correlation to its three-dimensional structure is one of the central dogmas of today's structural biology [26, 27].

However, things may not be that easy as there are exceptions in the protein structure/function-dogma, just as well as our daily life (frequently or less frequently) contain abstractionism (Fig.3).

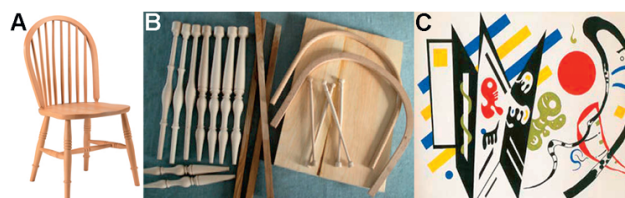


Figure 3. Illustration of proteins structure/function-dogma. (A) A chair. (B) An unfolded chair. (C) The painting "Reciprocal" by Kandinsky (1935).

Disordered proteins that folds only upon binding their target has been investigated and expand the concept of the biological machinery even further [28, 29]. To generalise and perhaps presuppose, the structure may at least gives an indication of the proteins molecular mechanism. The prevision may further be used to theorize and come up with functionality tests and sequential manipulations such as point mutations. Investigating and altering the function and conformation should give more pieces to the puzzle and open up for even further speculation.

1. INTRODUCTION

If practically, another interesting aspect of the structure/function relationship, involves the comparison of protein taken from patients with a medical disorder, confirmed to a specific protein, and compare with the healthy variant. Upon superposition, any deviations may extend the biochemical apprehension and reveal the reason and molecular mechanism to the pathological state, hopefully generating a medical counter.

1.3.2 Membrane Protein Structural Biology

The biological importance and hence the medical significance of membrane proteins cannot be over emphasized. Due to the fact that membrane proteins influence the majority of biochemical processes, they are considered to be critical in pharmaceuticals, accounting for the major part of all known drug targets [22-24, 30]. Additionally, alterations in membrane proteins might result in acquired or inherited diseases such as haemophilia, cystic fibrosis, diabetes, Alzheimer's disease and various forms of cancer [31]. Yet, we know relatively little about them, having few high resolution structures and mechanisms fully determined. One of the main reasons of our lack of structural knowledge lay to a great extent to the challenging task of isolating adequate amounts of pure and stable membrane protein [32, 33]. Subsequently, the pure and stable membrane protein sample should then be used to perhaps an even more challenging task, the generation of well diffracting crystals.

The growth of protein structures deposit in the PDB (Protein Data Bank) increase yearly and are close to a total of 90000. When examining how many of those determined structures are membrane proteins, an unequal distribution of soluble and membrane proteins is revealed, approximately at the ratio of 200/1 [34].

1.3.3 The Problematic Nature of Membrane Protein Purification and Crystallization

To generalise, it's the fundamental duality of the hydrophobic/hydrophilic nature of membrane proteins which creates a challenging task upon purification and stabilisation. Commonly, the targeted membrane protein needs to be extracted from its native membrane bilayer in order to be adequately separated and purified. However, simply disrupting the native bilayer and unleash the membrane protein into solution will destabilise the

1. INTRODUCTION

structural conformation and force a conformational change towards a more energetically favoured structure. Often, the new conformation will not be chemically active and have a tendency to aggregate. Instead, a detergent (surfactant) is used to initially disrupt the native bilayer and subsequently solubilise and stabilise the native conformation ^[35]. Importantly, it is crucial to use a well-suited detergent(s) as well as appropriate critical micelle concentration (CMC) to fit the necessities of the specific membrane protein.

Growing crystals from a protein solution can be a very tough job compared to many inorganic substances which often only require an oversaturated solution to be heated and slowly cooled to generate crystals. However, using such approach on a protein solution will denature to protein into oblivion. So instead of using heat as a concentrator, producing protein crystals generally requires more delicate techniques which allow the protein sample to be supersaturated such as water diffusion, dialysis or gels and capillaries.

Furthermore, there are some basic principles that need to be fulfilled before attempting crystallization. Firstly, the target membrane protein cannot be too shielded by detergent molecules as self-associated ordering into three-dimensional crystals requires some contact area ^[36]. However, the balance between “good” and “bad” contact areas is the balance of stable molecular interactions and nonspecific aggregation and precipitate formation. In addition to stability and self-association ability, the protein sample has to be as homogenous as possible. Paradoxically, the absence of a target specific detergent results in hundreds of other membrane proteins being solubilised, resulting in an extensively inhomogeneous sample. Therefore, it is essential to implement a proficient purification protocol which not only removes unwanted solubilised membrane proteins but also keep the designated membrane protein stable and able to self-assemble. Finally, to figure out optimal crystallization conditions often require several crystallization trials including various additives, detergents and solutions. The fact that protein crystallization has been regarded as an art has only displayed the lack of knowledge of the many subtle yet profound variables protein crystallization involves.

1. INTRODUCTION

1.4 Aquaporins

Water has always played a vital role in the chemistry of life. Animals, plants and bacteria all require intake of water to survive. The evolution of the first living life forms progressed from an abundance of water which became fundamentally integrated into their biochemical systems, acting as solvent, reagent and bulk reaction medium ^[2]. Additionally, the significance of water in the chemistry of life is reflected in the multiple biochemical mechanisms which constantly monitor and regulate water fluctuation ^[19]. It had been suspected since the 1920's that water transportation across the plasma membrane couldn't solely depend on passive diffusion alone ^[37]. But it was not until 1992, when Prof. Peter Agre reported the first aquaporin (AQP), resulting in the 2003 Nobel prize in chemistry ^[38, 39].

AQPs are a group of highly conserved trans-membrane proteins functioning as selective channels, mainly for water but also other small solutes such as glycerol ^[24, 40]. Today, AQPs are well studied potential drug targets as well as molecular schoolbook examples of trans-membrane proteins. A main feature of the approximately 28kDa large AQPs, is the ability of preventing ions and other solutes to pass while still being able to efficiently pass through water, facilitating normal secretory and absorptive functions. Different variations of AQPs are found in just about all living organisms such as unicellular bacteria and yeast to plants and mammals. The highly conserved sequences as well as the occurrence throughout the physiological system manifest the importance of these "plumbing system for cells," as Prof. Peter Agre once called them.

AQPs are allocated to the group of membrane proteins known as major intrinsic proteins (MIPs) but are then further subdivided based on sequence similarity and substrate selectivity into - (i) classic or orthodox AQPs, selectively transporting water, (ii) Aquaglyceroporins (AQGPs), transporting small solutes but mainly water and glycerol and (iii) subcellular AQPs (scAQPs), acting on the inside of the cell ^[41].

1.4.1 Human Aquaporins

13 human AQPs (hAQPs) have been identified, displaying a wide variation in physiological roles ^[42] (Tab.1). Multiple sequence alignment,

1. INTRODUCTION

phylogenetic and functional analysis of the hAQP sequences, has divided these into the subfamilies of classical AQPs (hAQP0, 1, 2, 4, 5, 6 and 8) and AQP3s (hAQP3, 7, 9 and 10).

Perhaps surprisingly, hAQPs are located not only in high fluidic content tissue such as the kidney, but are also encountered in the skin, fat-tissue and brain ^[43]. The largest amount of hAQP-homologues are found in the kidney which reabsorbs approximately 150L water from the blood/day ^[44]. The diversity of different AQPs in the human kidney can in fact compensate a defective urinary concentrating ability cause by hAQP1 deficiency ^[45]. However, many tissues and organs in the human body does not showcase the homologous diversity of AQPs such as the kidney.

Along with substantial significance, defective AQPs arise in substantial malfunction, resulting in a plenitude of medical conditions such as cerebral edema, diabetes insipidus and congenital cataracts ^[24, 31, 46]. Confirming the wide abundance of AQP-related diseases has resulted in an extensive research effort which in addition resulted in deeper insight and understandings of biological macromolecular structure and function in general.

hAQP	Permeability	Major tissue expression
hAQP0	Water	Eye lens fiber cells
hAQP1	Water	Kidney tubules, endothelia, erythrocytes, choroid plexus, ciliary epithelium, intestinal lacteals, corneal endothelium
hAQP2	Water	Kidney collecting duct
hAQP3	Water, Glycerol, Urea	Kidney collecting duct, epidermis, airway epithelium, conjunctiva, large airways, urinary bladder
hAQP4	Water	Astroglia in brain and spinal cord, kidney collecting duct, glandular epithelia, airways, skeletal muscle, stomach, retina
hAQP5	Water	Glandular epithelia, corneal epithelium, alveolar epithelium, gastrointestinal tract
hAQP6	Water and Anions (Cl ⁻ , NO ₃ ⁻)	Kidney collecting duct and intercalated cells
hAQP7	Water, Glycerol, Urea, Arsenite	Adipose tissue, testis, kidney proximal tubule
hAQP8	Water, Urea and NH ₃	Liver, pancreas, intestine, salivary gland, testis, heart
hAQP9	Water, Glycerol, Urea, Arsenite	Liver, white blood cells, testis, brain
hAQP10	Water, Glycerol, Urea	Small intestine
hAQP11	?	Kidney, liver
hAQP12	?	Pancreatic acinar cells

Table 1. The table represents substrate specificity and organ localisation of all human AQP homologues [24, 40, 44].

1. INTRODUCTION

1.4.2 Plant Aquaporins

Osmosis plays a considerable role in plant physiology, creating the plants turgor-pressure which enables the plants rigidity and facilitates nutrition [47, 48]. The substantial role AQPs fulfil in plants is effectively illustrated by the extensive amount of isoforms present in the same species such as 36 genes in maize (*Zea mays*) and 35 in Arabidopsis (*Arabidopsis thaliana*) while as previously mentioned, humans “only” has 13 [49, 50]. However, water intake and flow within the plant cells are mediated through AQPs which additionally to mammalian AQPs, include gating mechanisms to withstand rapid pH changes, drought and flooding [51].

Particularly plants living on land are accustomed of having a high variance in water availability. Accordingly, when available, the water regulation through AQPs responds quickly. Defective or inhibited plant AQPs results in reduced plant growth or even death of the plant. Moreover, AQPs found in plants customarily get subdivided into four main groups, plasma membrane intrinsic proteins (PIPs), tonoplast intrinsic proteins (TIPs), nodulin-26 like intrinsic proteins (NIPs) and small basic intrinsic proteins (SIPs) (Fig.4) [52]. However, recently another isoform-group to include to pAQPs is the x intrinsic proteins (XIPs) [53].

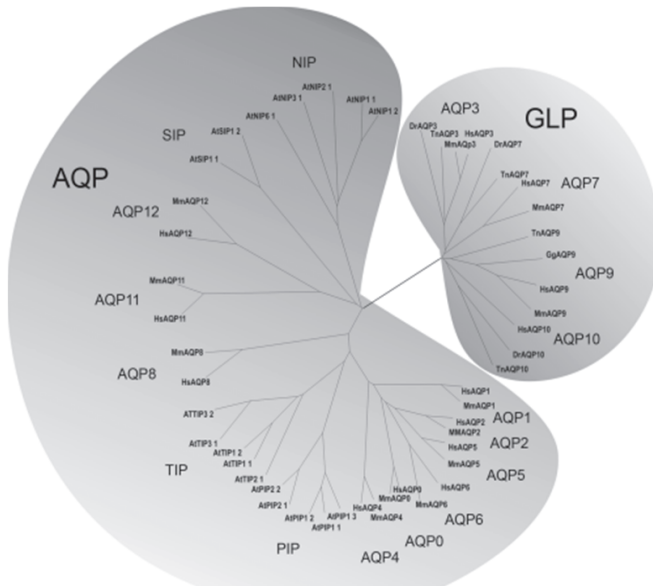


Figure 4. Phylogenetic tree of the evolutionary relationship between mammalian AQPs (AQP0-12) and plant AQPs [plasma membrane intrinsic proteins (PIP), tonoplast intrinsic proteins (TIP), nodulin 26-like intrinsic proteins (NIP) and small intrinsic proteins (SIP)] [52].

1. INTRODUCTION

1.4.3 Structural Features of Aquaporins

Early topological and structural studies of hAQP1, predicted a right-handed “hour-glass”-like structure *i.e.* narrow in the middle and wide at top and bottom ^[54]. The “hour-glass”-conformation has now been confirmed by several high-resolution structures deposit in the PDB ^[51,56, 92-93].

Inspecting the general structural composition of AQPs, reveals that six TM α -helices (1-6) and two half-helices (B and E), forms an approximately 20Å narrow channel, functioning as a selective pore through the membrane (Fig.5A and B).

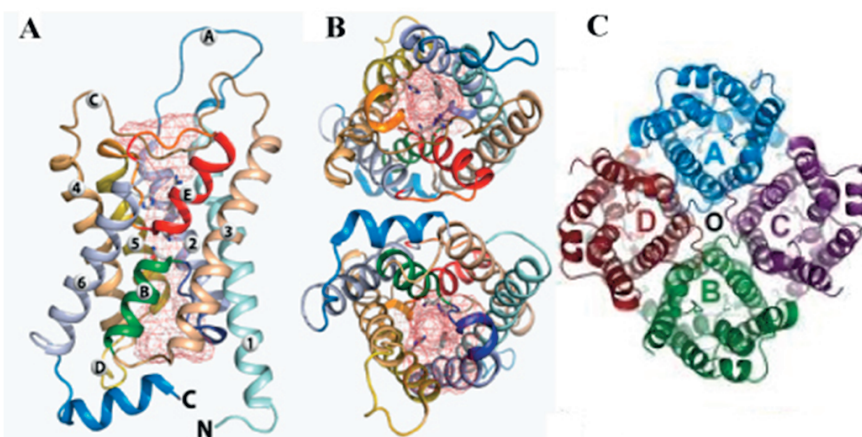


Figure 5. (A) Structure of a hAQP1 monomer in side view, highlighting the NPA-motif containing two half-helices B and E in green and red, respectively. (B) Top and bottom view of hAQP1. (C) Tetrameric assemble of A, B, C and D hAQP1-monomers viewed from the bottom. The O indicates the central pore. The red mesh in (A) and (B) illustrates the pore region.

The diameter of the channel force accessible molecules such as water, to be lined up and transported in a single file (Fig.6A) ^[55]. Furthermore, half-helices B and E, meet in the middle of the bilayer and folds into a seventh α -helix. In addition, the half-helices B and E contain the AQP NPA-signature motif (N/Asparagine-P/Proline-A/Alanine), constituting the narrowest path of the channel. The NPA-motif is characteristic to all AQPs but some variation occur, mainly among the AQPGs. In addition to the NPA-motif, an Arginine (R/Arg) positioned in the proximity to the narrowest part of the channel, functions as a selectivity filter of larger molecules, protons and ions. Residues with mainly hydrophobic nature

1. INTRODUCTION

covers the structures external interface whereas the channels inside contain both hydrophobic and hydrophilic residues *i.e.* an overall amphiphatic character ^[56].

Subsequently, the characteristics of the residues indicate that AQPs are well embedded in the membrane while still being able to maintain a selective channel. Moreover, MD-simulations suggest that the characteristics of the inside of the water channel forces water molecules to flip side. Importantly, rotating water molecules stacked in a single line prevents a chain of continuous hydrogen-bond to occur which potentially could conduct protons by the Grotthuss-mechanism ^[57]. In addition, the construction region is slightly wider in AQGs compared to classic AQPs, 3.4Å and 2.2Å respectively, suggesting that rejection of glycerol by classic AQPs is attributed to mere size ^[58].

However, the decreased potential of water transport by AQGs compared to classic AQPs is caused by increased hydrophobicity inside the channel. Most likely, a dipole moment, created in the interface of the two half-helices B and E, prevents destructive proton passage through the channel, thus preserving the uttermost important electrochemical gradient, $\Delta\mu$ ^[57, 59]. Further on, AQPs assume to cluster in stable tetrameric assemblies, where each monomer functions as a separate independent channel (Fig.5C). As a consequence, the tetrameric assembly creates an additional channel, a central pore, which presumably is permeable to non-polar gas molecules such as O₂ ^[60]. However, it is still debated whether or not this channel is of biological relevance or just an experimental artefact.

1.4.4 Aquaporin Regulation and Gating

A common central regulation strategy used by biological systems is the regulation of accessible molecules to the target(s) *i.e.* trafficking ^[8, 61]. Indeed, an effective regulatory method adopted by most biological systems is to store quantities of inactive molecules which are only transported and activated when needed for. For example, hAQP2s main function is to reabsorb water from the urine and is conveniently situated in the kidneys principal cells in the collecting duct ^[62]. Furthermore, a “reserve pool” of (inactive) hAQP2 is stored inside intracellular vesicles in the plasma membrane. Subsequently, when the body needs to retain water, the peptide hormone Vasopressin is released, which leads to

1. INTRODUCTION

phosphorylation (activation) of hAQP2 (at Ser256) as well as triggering a fusion cascade reaction of intracellular vesicles containing hAQP2 to the membrane. Upon vesicle fusion, activated hAQP2s inserts to the membrane, resulting in an increased potential of water flux and water retention.

In addition, structural and molecular dynamic studies show that some AQPs can assume an open and a closed conformation (Fig.6B) [63]. This physical gating mechanism is commonly triggered by phosphorylation, pH, osmolarity changes and/or binding of a molecule, for example Ca^{2+} .

The plant AQP from spinach (*Spinacia oleracea*) SoPIP2;1, illustrates an ingenious closing mechanism which involves the cytoplasmic loop (loop D) to bend inward, inserting Leu197 as a “plug” in the channel [51, 64]. Contrarily, in the open conformation, loop D is displaced by up to 16Å, allowing a clear passage through the channel. During conditions normal to the plant, the two highly conserved residues Ser115 (conserved in all PIPs) and Ser274 (conserved in PIP2s) of SoPIP2;1, are phosphorylated and the channel is open.

During drought, the channel closes due to a dephosphorylation of the two residues. However, if the plant is flooded, the channel will also be closed but due to the response to a protonation of His193. Both gating by phosphorylation and pH involve interactions with a Cd^{2+} ion (Ca^{2+} ion *in vivo*), bound at the N-terminus which stabilises the loop in the closed state, preventing the loop from open. On the contrary, when Cd^{2+} is absent the stabilizing interactions with the D-loop is not available, causing channel opening.

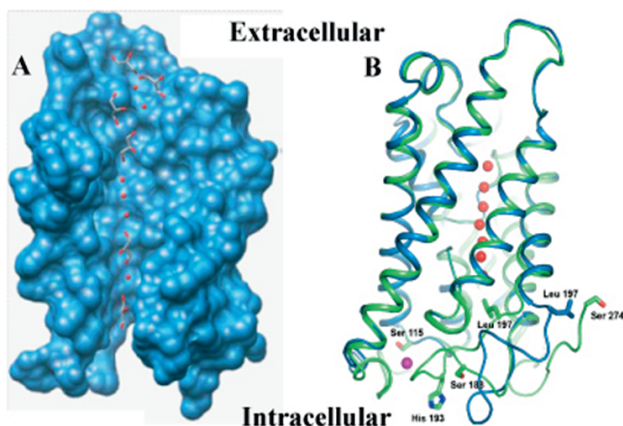
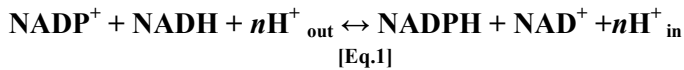


Figure 6. (A) The inside channel region of hAQP1 [55] illustrating the water molecules rotational action while transported on a single file. (B) Structure of monomeric SoPIP2;1 viewed from the side illustrating the closed and open conformation superimposed in green and blue, respectively. Water molecules and a Cd-molecule are shown in red and pink, respectively.

1.5 Proton-Translocating Nicotinamide Transhydrogenase

Located in the inner-membrane of eukaryotic mitochondria and the cytoplasmic membrane in various bacteria, the membrane bound proton-translocating nicotinamide nucleotide transhydrogenase (TH), catalyzes the hydride transfer between NAD(H) (nicotinamide adenine dinucleotide) and NADP(H) (nicotinamide adenine dinucleotide phosphate) [65, 66]. Although TH has many similarities among species concerning gene sequence and functionality, the translation of prokaryotic and eukaryotic TH deviates [67]. That is, eukaryotic TH get translated as a single polypeptide whereas prokaryotic TH produces two or in some cases three polypeptides which get assembled first after translation [68]. Even though the translation processes differ, the enzymatic redox-reaction does not [69, 70].



Where “out” and “in” refers to the matrix and to the intermembrane space of mitochondria, respectively [66, 71]. When addressing prokaryotes, the “out” and “in” refers to the periplasmic space and to the cytosol, respectively. In addition to the reversible redox-reaction between NADP⁺ and NAD⁺, TH features a proton-pump function as part of the mitochondrial respiratory chain [7]. The redox-reaction initiates a conformational change which permits proton transfer across the membrane that either generates or consumes the proton electrochemical gradient ($\Delta\rho$). Commonly, under normal physiological conditions *in vivo*, TH utilises the $\Delta\rho$ to produce NADPH (left to right in Eq.1). In addition, the produced NADPH is used in biosynthesis and in detoxification of free radicals as a defence against cellular oxidative stress *i.e.* reactive oxygen species (ROS) such as superoxide anions (O₂⁻) and hydrogen peroxide (H₂O₂) [72-75]. Moreover, studies indicate that the absence of fully functioning TH effect the potassium (K⁺) and calcium (Ca²⁺) channels in the insulin producing pancreatic β -cells, resulting in increased ROS production. The accumulation of ROS overtime, increases the non-functional pancreatic β -cells, leading to glucose intolerance and reduced insulin secretion *i.e.* *Diabetes Mellitus* (type 2 diabetes) [72, 76].

1. INTRODUCTION

1.5.1 Structural Features of TH

TH (MW ~110kDa) features a general architectural composition of three domains comprised by two hydrophilic domains (dI and dIII) and one hydrophobic domain (dII). The dII is further divided in a dII α and dII β subunit (Fig. 7A) [68, 77].

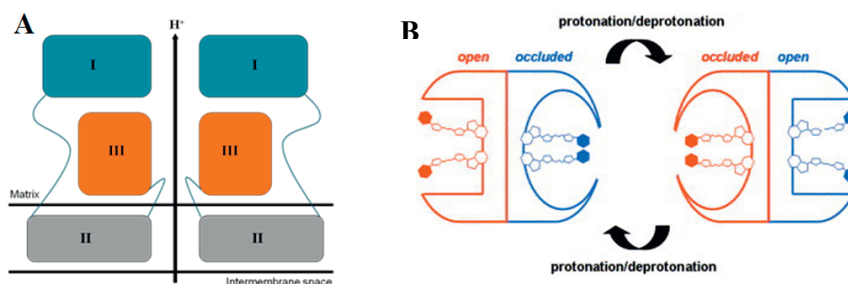


Figure 7. (A) Schematic drawing representing the three domains (I, II and III) of dimeric proton-translocating transhydrogenase. (B) Illustration of the plausible binding-change mechanism of proton-translocating transhydrogenase, showing the monomers in a dimer as alternating between an open and occluded state in red and blue, respectively [78].

However, TH functions as a dimer *in vivo*, consisting of two monomers, each containing a dI+dII+dIII unit [79]. Sequential comparison of nicotinamide nucleotide binding proteins reveals the characteristic NAD⁺/NADH binding-sequence of GXGXXG to be located in dI [80]. Logically, the characteristic binding-sequence of NADP⁺/NADPH binding proteins, GXGXXA, is localized in dIII. Although obscure, dII is mainly localized in the membrane and presumably contains the proton-translocation mechanism [81].

The proton-translocation mechanism remains uncertain as the structural determination of a full-length TH has yet to be done. However, several high resolution structures (2.3-3Å) of the hydrophilic domains (dI and dIII) in various combinations and complexes have been presented in the PDB [77, 82-84]. The binding-change mechanism, originated from combining such structural knowledge to experimental data. Accordingly, the binding-change mechanism of TH, define the redox-reaction as alternating between an occluded and an open state of the two monomers (Fig. 7B) [85, 86].

In the open state, the dihydro- and nicotinamide rings are facing away from each other which prevent bound nucleotides from any redox-reaction but

1. INTRODUCTION

permits dissociation of nucleotide products into the solution or alternatively binding new nucleotide substrates. Contrarily in the occluded state, the dihydro- and nicotinamide rings are facing each other and prevents solution interaction and substrate dissociation but permits redox-reaction of bound nucleotides. It is speculated, that the conformational alternation of the open and occluded state of the two monomers, respectively, are imposed through dII's proton-translocation mechanism.

1.6 Scope of the Thesis

The aim of my research involved the task to ultimately reveal structures and molecular mechanisms. However, the very nature of that goal involves several obstacles to be solved first before even attempting to elucidate the main goal. In most cases, there are two major bottlenecks regularly encountered in membrane protein production; the generation of adequate yield and the implementation of an efficient purification procedure. Thus, the majority of time has been spent on the early stages of molecular biology *i.e.* cloning, mutagenesis, cultivation and purification. Nonetheless, in **Paper V**, the chance to structurally determine the open conformation of a plant AQP mutant appeared, thus completing all elements in the “gene to structure” process.

In **Paper I**, the first bottleneck is addressed by a comparative production study of all 13 highly homologous hAQPs, seeking further insight into factors directing eukaryotic membrane protein overproduction. In **Paper II** and **Paper III**, several plant AQP mutants were produced, purified and tested for functionality. Consequently, further structural insight into plant AQP regulation and mechanism was achieved. Moreover, in **Paper IV**, two constructs of full-length eukaryotic proton-translocating transhydrogenase was designed, produced and purified. Activity could be detected but this was however inconsistently reproducible.

2. Methodology

2.1 From Theory to Protein

Firstly, the experimental strategy needs to be greatly considered in order to practically be able to produce a protein of interest in stable and adequate quantities. Which cloning vectors, amplification and protein production organisms are chosen may vary depending on the projects goal, time plan, economy, and laboratorial capacity.

2.1.1 Cloning Strategy

The overall aim to produce a fully functional recombinant eukaryotic protein in sufficient amount may start with the choice of cloning vector and production host organism. Most eukaryotic proteins require post-translational modifications, such as glycosylation, to assume correct fold and activity, an ability which prokaryotic organisms lack ^[87]. For that reason, prokaryotic hosts are usually only used to express and amplify the DNA-construct (cloning vector + target protein gene) whereas the actual production of the protein of interest is executed by a eukaryotic host. Concurrently follows the choice of a cloning vector. A cloning vector is a circular piece of DNA into which a foreign part of DNA may be inserted, *i.e.* the target proteins gene sequence. The cloning vector favourably contains a multiple cloning site (MCS), an origin of replication (ORI), an inducible promoter region and one or more selectable markers, gaining specific antibiotic resistance. Additional features may be added adjacent to the target protein gene sequence, such as affinity tags to aid purification.

2.1.2 Protein Production Organisms

Different protein production hosts provide different biochemical environments for any protein target and vary according to lipid composition, glycosylation patterns, chaperones and other post-translational modifications ^[9]. When practically possible it may be advantageous to try several different production organisms, however, this is frequently beyond the capability of most research laboratories.

Successfully used eukaryotic protein production organisms include a wide range of choices including; the bacterial hosts *Escherichia coli* (*E. coli*) and *Lactococcus lactis*, the yeast systems *Saccharomyces cerevisiae*

2. METHODOLOGY

and *Pichia pastoris* (*P. pastoris*), the insect cell/baculovirus system, and various mammalian cell lines^[88].

2.1.3 Overproducing Proteins Using *Pichia Pastoris*

Yeast provides a safe, cost effective, and easily manipulated host which is becoming common standard for producing eukaryotic proteins^[89-91] as well as the most successful production host to date used for structural determination of eukaryotic membrane proteins^[51, 92, 93]. In addition, protein to membrane insertion and membrane lipid composition in yeast seems to be a beneficial environment for newly produced eukaryotic membrane proteins^[94, 95].

The methylotrophic yeast *P. pastoris*, has shown to be a particularly powerful host for the production of recombinant proteins mainly due to two facts. Firstly, the capability to grow to high cell densities (~150g L⁻¹ dry weight cells) enables the use of relatively small cultivation volumes, compared to other production organisms such as *E. coli*^[96]. Secondly, *P. pastoris* contains an inherent and very strong, methanol-inducible alcohol oxidase 1 (AOX1)-promoter^[97, 98]. To explicate, *P. pastoris* can utilize methanol (MeOH) as its sole carbon source by producing the enzyme AOX1. However, AOX1 have a very poor affinity for oxygen which is a necessity for survival. Consequently, the *P. pastoris* system compensates the low oxygen affinity by a very high production (overproduction) of the AOX1 enzyme, by that achieving sufficient oxygen quantity.

2.1.2 Cultivation of *Pichia Pastoris*

Using the capability to grow to high cell densities, cultivation using enclosed 3L-bioreactors is advantageous versus conventional shaker flask cultivation (Fig.8A,B)^[99, 100]. Several important parameters such as, pH, temperature, dissolved oxygen, air flow and biomass are easily controlled by bioreactor-cultivation. Furthermore, the requirements for sufficient aeration, agitation and pH adjustment can easily be fulfilled using a bioreactor contrarily conventional shaker flask cultivation.

A typical *P. pastoris* bioreactor-cultivation contains three phases (Fig.8C); (1) a glycerol batch phase, where the yeast cells will utilize glycerol present in the basal salt media, focusing on increasing the biomass. (2) the glycerol fed-batch phase were the yeast cells are fed with limiting glycerol,

2. METHODOLOGY

forcing them to adapt and grow under limiting conditions thus preparing them to the MeOH induction and (3) the MeOH fed-batch phase which starts the production of the target protein^[99].

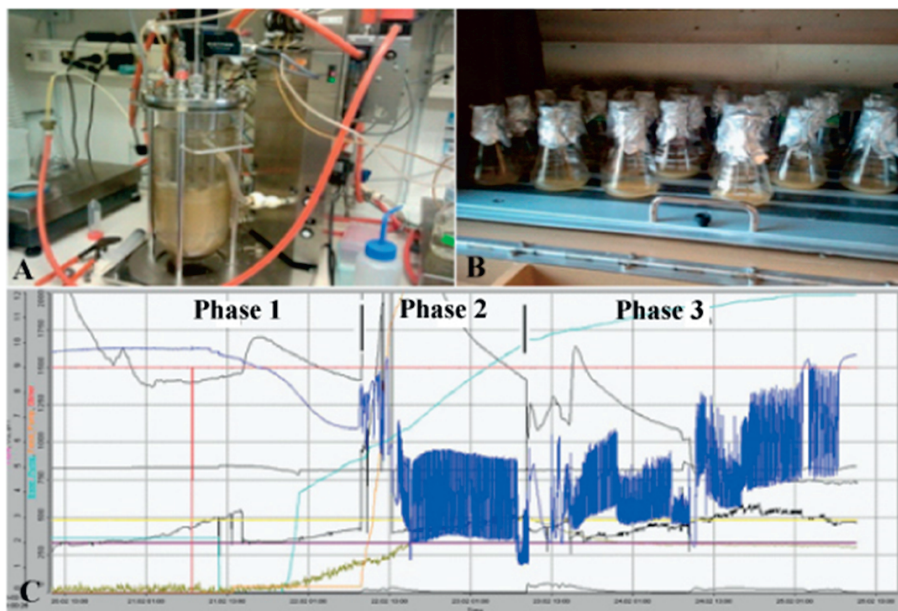


Figure 8. (A) A 3L-bioreactor. (B) 25ml shaker flasks in an incubator. (C) A typical bioreactor diagram showing the oxygen “spiking” in blue over the three cultivation stages denoted phase 1, 2 and 3.

In addition, when cultivating *P. pastoris* it is important not to add more MeOH than is consumed because accumulated MeOH is highly toxic for the yeast cells^[101]. That is, the methanol addition should be precisely matched to the highest consumption rate, ensuring optimal production from the AOX1-promoter.

2.1.4 Membrane Protein Purification

In order to acquire a pure and uniform protein sample, protein purification involves multiple steps specifically target-protein customized. Indeed, establishing a universal membrane protein production protocol might only work regarding as a general approach, not as a definitive protocol (Fig.9). Specific materials such as affinity and gel filtration columns and various buffer solutions will vary depending on the target-proteins nature.

2. METHODOLOGY

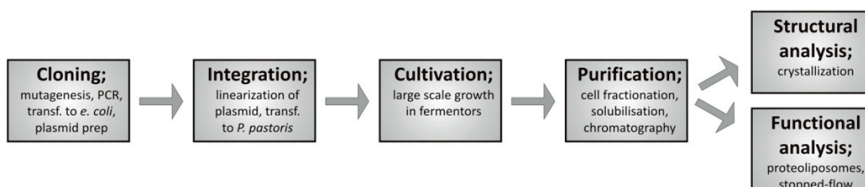


Figure 9. A general schematic illustration of the “gene to structure” process used in my studies.

The following text describes a general protein purification guideline for membrane-bound proteins localized in the inner-membrane of whole cells [102].

Initially, the cells need to be broken enzymatically, chemically or mechanically in order to expose the organelles to the solution. Next step is to remove the cell debris and unbroken cells through centrifugation. The resulting supernatant is subjected to ultracentrifugation resulting in a membrane pellet containing the protein of interest. The membrane pellet generally needs to be washed to remove as many proteins as possible without removing or harm the protein of interest. Choosing what to wash with varies depending of the nature of the protein of interest as well as the degree of its solution exposure. That is, a harsh washing-media, such as Urea and NaOH, may be used if the protein of interest is highly embedded in the inner-membrane, thus being greatly protected and less likely to get disrupted and washed away.

In the next step the membrane-embedded proteins must be removed from the lipid bilayer without causing structural instability and aggregation. Since it is the exposure of the hydrophobic parts to the solution which mostly accounts for the instability, measures to stabilize the hydrophobic parts upon removal are best suited. Generally this is accomplished by adding a water-soluble surfactant (detergent) which will disrupt the native bilayer and stabilize the hydrophobic parts of the protein. As the purification continues en route, the following purification strategy will rely on the target-protein’s physical-chemical parameters such as charge, specific affinity and molecular size. The combination of affinity chromatography and/or ion-exchange followed by gel-filtration is a powerful purification strategy which in theory will separate all proteins since no protein have the same intrinsic values with respect to charge, affinity and molecular size. However, protein isoforms, aggregation,

2. METHODOLOGY

interaction and degradation may prevent proper separation requiring revision of buffers, concentration and perhaps even altering prior purification strategy.

2.2 X-ray Crystallography

An experimental way of obtaining detailed and precise three-dimensional models of proteins, is to use high-intense light (X-rays) whose wavelength (λ) is slightly less than the length of the carbon-carbon bonds within the protein ($\leq 1,5\text{\AA}$). Accordingly, incoming X-rays hitting a protein molecule will be diffracted by the atom's electron cloud in a structure dependent pattern *i.e.* a diffraction pattern. In spite of that, a single protein molecule generates insufficient diffraction intensity inconceivable to differentiate from background signal noise. As a consequence, protein molecules are directed to arrange themselves in a well-ordered manner by forming protein crystals in order to amplify the diffracting X-rays.

2.2.1 Protein Crystallization

Growing crystals from a protein solution can be a very tough job compared to many inorganic substances which often only require an oversaturated solution to be heated and slowly cooled to generate crystals. However, using such approach on a protein solution will denature to protein into oblivion. So instead of using heat to concentrate the sample, producing protein crystals generally requires a more delicate technique such as vapour diffusion ^[103]. The vapour diffusion technique uses a sealed environment and an aqueous protein-solution, kept separated from an aqueous reservoir-solution containing a higher precipitant concentration (Fig.10A).

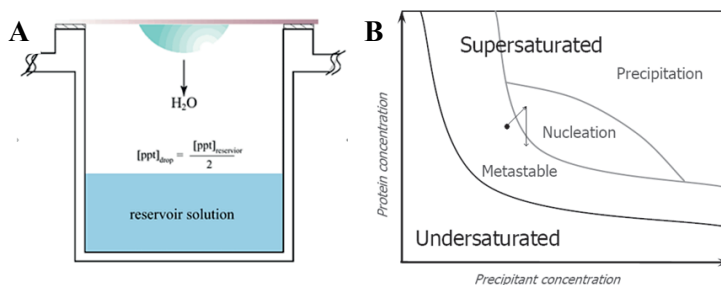


Figure10.

(A) The hanging drop vapour diffusion method [©Michael R. Sawaya]. **(B)** A protein crystallization phase diagram. The intention is to reach the nucleation phase from the offset of the metastable phase via water evaporation. The non-crystallized protein concentration decreases as crystals nucleus form which eventually progress and re-enters the metastable phase where the crystals may continue to grow.

2. METHODOLOGY

The enclosed system will pursue equilibration which is only possible by diluting the reservoir-solution. The only possible mobile substance in the enclosed system is water which by vapour diffusion moves from the protein-solution to the reservoir-solution. Consequently, both the precipitant and the protein concentration in the protein-solution increases as the enclosed system shift towards equilibrium.

The general idea of the crystallization process is to first reach the initial nucleation phase from where additional protein molecules are allowed to be arranged in a systematic fashion, according to a crystallization phase-diagram ^[104]. As the (free)-protein concentration decreases when crystals form and grow, the solution moves towards the metastable phase, the optimal condition for additional crystal growth (Fig.10B). Crystallization phase-diagrams vary somewhat depending on protein, concentrations, buffers and temperature. To figure out optimal crystallization conditions often require several crystallization trials including various additives, detergents and solutions.

2.2.2 X-rays

Discovered 1895 by Professor Wilhelm Röntgen ^[105], X-rays are electromagnetic radiation prepossessing the same physical factors as visible light such as intensity, frequency and polarisation. However, X-rays occur at a wavelength of $\sim 10^{-8}$ to 10^{-12} m ($1\text{\AA} = 10^{-10}$, Ångström) which is much shorter than visible light, occurring at 400 to 700Å. Although applying somewhat different parts of the X-ray wavelength spectra, the main uses of X-rays are for medical and structure determination purposes. For instance in medical radiography, X-ray wavelengths close to 0.3Å is used, as this type of higher energy radiation penetrates most biomaterial without considerable absorption. In crystallography, wavelengths around 1 and up to 10Å are used with the intent to interact and get absorbed by the target sample. Thereby, conducting crystallographic experiments involves many safety regulations as the X-rays may inflict irreversible damage to human tissue and won't distinguish between a crystal and a human body.

Commonly, there are three ways to generate X-rays, by X-ray tubes, rotating anode tubes and particle storage rings (synchrotrons). The intensity of synchrotron X-rays can be at least 10^3 times higher than those obtained with a conventional X-ray tube as well as being better focused

2. METHODOLOGY

[106]. While X-ray and rotating anode tubes are still used as in-house sources, most modern crystallographic experiments takes place at synchrotron radiation facilities such as the ESRF (European Synchrotron Radiation Facility) in Grenoble, France or the NSLS (National Synchrotron Light Source) in Brookhaven NY, USA.

2.2.3 X-ray Crystal Diffraction

X-rays targeting a well-ordered crystal generates a diffraction pattern exclusively related to the crystals atomic arrangement. However, it is only those diffracted X-rays who satisfy Bragg's law which are amplified while the majority of diffraction cancels out (Fig.11).

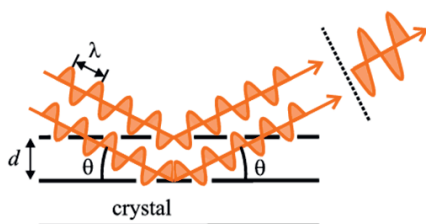


Figure 11. Illustration of two diffracted waves satisfying Bragg's law.

[© Birkbeck, CollegeUniversity, London.
P.Barnes T.Csoka S.Jacques]

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

Bragg's law is given in Eq.2 where n is an integer, λ is the wavelength, d is the distance between the crystals parallel planes and θ is the angle of the planes. The condition for constructive interference implies that the diffracting waves of separate planes are in phase only when Bragg's law is fulfilled which is why a crystals repetitive unit cell arrangement amplifies diffraction signals above signal noise. Intriguingly, X-rays does not directly interfere with the atoms in the crystal but rather with the electrons surrounding the atoms. Thereby the diffraction pattern is interrelated to the electron density of the unit cell.

Furthermore, the diffraction pattern of waves satisfying Bragg's law, does not directly depict an image of the unit cells electron density. Just as a focusing lens in a light microscope reconstructs a samples image, the diffracted X-rays from a crystal also need to be focused. Problematically, there are no lenses available that can focus X-rays directly. Furthermore, each generated reflection contain the unit cell arrangement in an inverted relationship, having the real space unit cell axes a , b and c appear

2. METHODOLOGY

reciprocally as $1/a$, $1/b$ and $1/c$, respectively. To keep on using a light microscope as an example, the light microscope lens performs an operation analogous to the mathematical Fourier transform function (FT), which describes the mathematical relationship of the object and the corresponding diffraction. In this mathematical manner, calculating (the reversible) Fourier transform function on diffraction data makes it possible to move back and forth between reciprocal (diffraction pattern) and real space (electron density) (Eq.3).

$$F_{hkl} \xleftrightarrow{\text{Fourier Transform}} f_{xyz} \quad (\text{Eq.3})$$

In order to do this, every single reflection at a specific coordinate (hkl) is described as a complex wave containing information of the atomic arrangement in the crystals unit cell. Consequently, each appearing reflection is denoted a structure factor equation (F_{hkl}) which summates all the unit cell atoms contributing diffraction. In essence, the main goal of X-ray structure determination is to obtain and sum up the F_{hkl} of all detected reflections and apply a Fourier transform which calculates the electron density ($\rho_{x,y,z}$) in the whole unit cell (Eq.4).

$$\rho_{x,y,z} = \left(\frac{1}{V}\right) \sum_{hkl} F_{hkl} e^{-2\pi i(hx+ky+lz)+(\phi_{hkl})} \quad (\text{Eq. 4})$$

Also, each reflection in a diffraction pattern has a unique phase which is required to solve the electron density equation. However, a typical macromolecular X-ray experiment detects only the reflection intensities (I_{hkl}) and positions, hkl , not the phases. This conundrum is referred to as the phase problem.

2.2.4 The Phase Problem

Several solutions has emerged to solve the phase problem, all having different advantages and disadvantages depending the nature of the diffraction pattern to be solved. Commonly, three main strategies are employed: direct methods, molecular replacement and heavy atom methods. The effort to directly measure the phases becomes increasingly more complex as the molecules size increase and is not used for structural determination of proteins. The direct methods are based on several assumptions, such as atoms being defined as still mathematical points

2. METHODOLOGY

randomly distributed in a black box. In addition, the electron density never becomes negative. Also, the space group of the structure must be known.

In a molecular replacement method, a protein of known structure and high similarity to the protein of interest is used as a template. The phase problem is solved by using rotational and translational movements to superimpose the diffracted data on the solved protein structure.

The heavy atom method is very useful in cases when there are no similar structures available. In principle, a strong scatterer such as the heavy atoms Hg, Pb and Fe is dissolved into a crystal and allowed to interact with the protein. The distinctive signal from the heavy atom makes it possible to determine the position by the application of the Patterson function. Solving the Patterson function will make it possible to retrieve all phases and solve the structure. Different heavy-atom techniques involves MAD (multi-wavelength anomalous diffraction), SAD (single-wavelength anomalous diffraction) and MIR (multiple isomorphous replacement).

2.2.5 Twinning

Twinning is a phenomenon that generates different crystal lattice symmetries within an existing crystal. This is problematic because when a twinned crystal is subjected to X-rays, the diffraction pattern is a mixture of both crystal lattice orientations and one needs to sort out which spots belong to which. Seemingly, AQPs have a tendency to form twinned crystals, most likely due to the fact that the extracellular side is relatively similar to the cytoplasmic-side. During the crystallization process, one AQP molecule might flip which would be the starting point of the new growth direction as other AQP molecules would arrange thereafter, resulting in two growth directions within the same crystal.

2.2.6 The Data Collection Experiment

For X-ray data collection, the crystal sample is collected in a loop which is aligned to the incident X-ray beam (closed shutter). When ready, the X-ray beam shutter is shortly opened generating an X-ray pulse directed at the crystal which (hopefully) generates diffraction. A detector, such as a CCD (charged coupled device), detects the diffracted waves (reflections) and records respectively position and intensity. The crystal sample is rotated (commonly 0.5°) until all reflections have been collected or until the

2. METHODOLOGY

crystals diffraction ceases. Collecting diffraction images often generates a couple of hundred images which further may be processed.

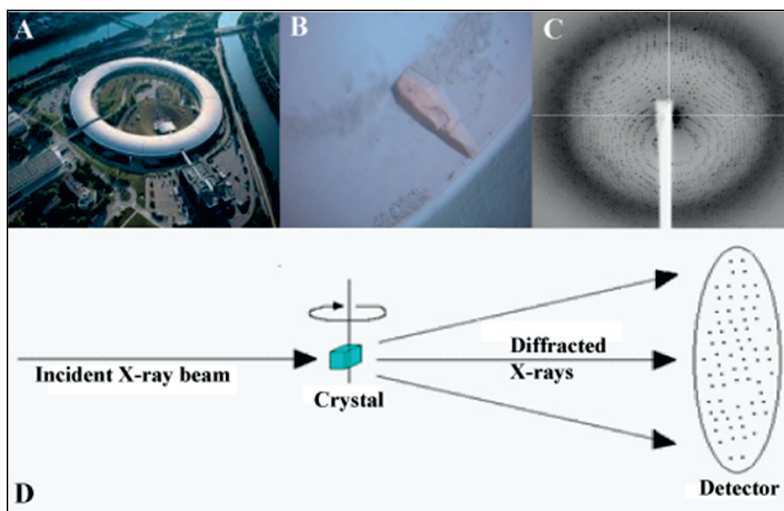


Figure 12. (A) The main storage ring at European Synchrotron Radiation Facility (ESRF) in Grenoble, France. (B) A protein crystal of SoPIP2;1-S188E. (C) An exceptional diffraction pattern (Sso6206, 10.5kDa protein from *Sulfolobus solfataricus*). (D) Simplified setup of a typical macromolecular X-ray experiment.

2.2.7 Data Processing, Refinement and Validation

Having a collected high-resolution data in possession won't do much good until properly processed. The data processing and refinement methods seems to become more and more automatic despite the underlying algorithms and formulas being quite complex (according to me). However, in order to achieve the best results, the processing sequence needs to be generally understood. Various software is available to use, iMosflm, the iCCP4 suite and Phoenix are commonly employed programs used to process, refine and validate crystallographic data and structures.

At first, all reflections need to be indexed. That is, the reflections in the diffraction pattern are located and assigned hkl-indices. The set of indexed reflection are the base of generating a space group and crystal orientation prediction. After choosing a probable space group, the cell parameters are refined. At this stage it is possible to make the program predict the diffraction pattern accordingly to chosen space group and compare the prediction to the reflections. Next is the integration step. That is, the

2. METHODOLOGY

intensity of each reflection in every image is determined. This is done by optimizing the area of the reflection and thereafter determine the intensity peak in that area relatively to the background on that image. However, the different diffraction images backgrounds will most likely vary considerably depending on background noise. Therefore, the intensity peaks from different images cannot be compared without an adjustment. The scale and merge adjustment relates the intensity peaks from different images to each other and in addition merges intensities, such as partial reflections, coming from the same reflection but detected on multiple images. At this point, all the F_{hkl} have been determined which in combination with the phases would solve the electron density map.

If available, MR is a convenient method to use solving of the phase problem. The sequential and thereby structurally similar protein is used as a template as the unknown model is superimposed on using rotational and translational functions. However, the acquired conformation needs further refinement by different modifications so that the structure factors of the modified model (F_{calc}) agree to the structure factors observed (F_{obs}). To assure the quality of the final model conformation, a percentage of the first collected unbiased-data (R_{free}) is used to see if it is a comprehensive agreement to the modelled structure and the experimentally recorded data. Accordingly, the R_{free} -value should decrease during the refinement procedure^[107].

2.2.7.1 Calculating a Composite-Omit-Map

When using Molecular Replacement, it is important to minimize model bias in order to correctly build the new structure. To aid in this process, a composite-omit maps is calculated in which parts of the model is systematically removed from the map calculation process, allowing for a less-biased map to be obtained.

As previously stated, twinning is problematic and requires different algorithms and twin operators that separate and prepare a twinned diffraction data set for further processing. However, it is a great likelihood to include model bias when detwinning perfect twins as the detwinning process itself depend on a model. The SoPIP2;1-S188E open structure (**Paper IV**) is modelled from nearly perfectly twinned diffraction data by

2. METHODOLOGY

using a method which allows calculation of a less biased composite-omit-map adjusted for twinned data ^[108].

At first, an empty three-dimensional space (3D- map) is defined and all atoms in one asymmetric unit are calculated. Thereafter, the asymmetric unit is divided into identical boxes (the number of boxes can be adjusted depending on the diffraction data quality). The atoms comprised in one of the boxes are removed, the data is detwinned (CNS ^[109]) and the resulting asymmetric unit is used to calculate an omit-map. The box contents that were excised are now filled out and the resulting omit-map is complete due to that diffraction data contains structural information from all scatterers. Lastly, the omit-maps structural information corresponding to the box which was excised is pasted in the empty three-dimensional map which is unbiased by the model. The cycle is repeated until every box covering the whole asymmetric unit has been excised, recalculated and pasted into the expanding new map, creating a composite-omit-map.

2.3 Functional Assays

A prerequisite for functionality of any given protein is the ability to assume an active conformation. Most methods to determine proteins functionality generally involves measuring the rate of the proteins specific enzymatic reaction or transport ability. It is also important to choose a suitable method which reflects the proteins natural environment as much as practically possible. Measuring directly *in vivo* would be ideal, however the complexity of a cell is far too great to fully comprehend all including variables. Therefore, it is practical to engineer and simulate a simplified membrane and determine membrane protein functionality *in vitro*.

2.3.1 Liposome Assays

In the study of AQPs, water transport-rate is of high interest as it indicates the functionality of the protein. Creating artificial lipidic vesicles (liposomes) in an aqueous environment is a fairly uncomplicated task. Accordingly, an aqueous solution containing phospholipids is subjected to sonication which organizes the lipids in unilaminar vesicles. Next, the use of a detergent-mediated reconstitution has shown to be successful for the insertion of membrane proteins ^[110]. That is, the protein of choice is added to the liposome mixture which upon detergent removal spontaneously

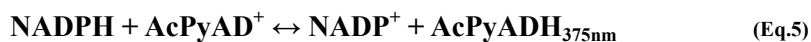
2. METHODOLOGY

reconstitutes into the liposomes. However, finding concrete successful reconstitution conditions include several factors such as protein type, liposomes, the solute as well as detergent type and concentration^[111]. It is advisable to analyse the size distribution of the proteoliposomes by dynamic light scattering (DLS) which determine the size distribution profile by the scattering intensities over time^[112].

2.3.2 Spectroscopic Assays

A stopped-flow experiment is an exquisite technique to test AQP-functionality (**Paper II, III and IV**). In a typical stopped-flow spectroscopy experiment, two separate solution chambers are rapidly mixed and the following absorbance or scattering variation is recorded. Specifically, the solution of proteoliposomes containing AQPs are rapidly mixed with a hyperosmotic non-permeable solute, such as sorbitol, which changes the osmotic gradient. The solute concentration on the outside of the proteoliposomes will thereby be higher than the inside which results in water being transported out of the proteoliposome in order to compensate the imbalance. Accordingly, as water flows out, the proteoliposomes shrinkage is recorded by a decrease in absorbed light at 400nm to 700nm. The resulting curves are plotted to an exponential function, describing the water transport rate and depicting the rate constant.

However, not all membrane proteins require a simulated membrane to arbitrate functionality. Knowledge of the proteins enzymatic reaction can be used to directly measure an absorbance change at a specific wavelength. The TH-activity (**Paper IV**) was determined by adding the NAD⁺ analogue AcPyAD⁺ (3-acetylpyridine adenine dinucleotide) and NADPH to the protein sample (Eq.5).



Typically, protein sample, AcPyAD⁺ and NADPH were mixed in a cuvette (1cm) at room temperature. Then the absorbance per minute at 375nm was monitored in a spectrophotometer. Thus, the enzyme activity becomes indirectly measured by the increase in absorbance at 375nm which is the specific absorbance of AcPyADH.

3. Results and Discussion

3.1 Membrane Protein Production in *Pichia pastoris*

In order to generate a high yield of a recombinant eukaryotic membrane protein it is often important to have a full understanding of the dictating factors of the protein production process. Factors that may be of interest include nucleotide consensus sequence, phylogenetic relationship, folding pathways, polymeric stability and/or charged residues.

In this thesis, the protein production task has been focused on one host, *P. pastoris*. In **Paper I**, the production levels of all 13 human AQPs are compared, given a deeper understanding of what influences membrane protein overproduction in this host. Likewise, *P. pastoris* was used to obtain samples of the SoPIP2;1 and additional mutants for structural and functional studies (**Paper II, III and IV**) as well as developing a production protocol of TH in *P. pastoris* (**Paper V**).

3.1.1 Production of human AQPs (**Paper I**)

All hAQPs were successfully cloned in the pPICZB-vector and transformed into *P. pastoris* in order to compare production levels. Furthermore, the hAQPs production levels were related to low and high production level by Fps1-N-terminus (+) and SoPIP2;1 (+++), respectively. Although, the production level of SoPIP2;1 was not reached, seven hAQPs gave higher yields than Fps1-N-terminus, suggesting that their production levels were sufficient in order to initiate protein purification. In addition, further conclusions concerning production yields and membrane localisation could be correlated to phylogenetic relationship and consensus sequences.

3.1.1.1 Assessment of hAQP Production Level

For quantification purposes and convenience, protein coding genes was inserted in the same vector, the pPICZB-vector (Invitrogen[®]), and transformed into *E. coli* for amplification. Thereafter, all constructs were linearized and transformed by electroporation into *P. pastoris*. The linearized constructs were integrated into the *P. pastoris* genome by homologous recombination. Selection by antibiotics was possible throughout the cloning process due to the pPICZB-vector's inherent

3. RESULTS AND DISCUSSION

Zeocin™-resistance gene. The 13 hAQPs provide an excellent comparative constitution due to the wide range of functionalities amongst and between the classic-hAQPs and the hAQGs. Yet, the hAQPs showcase a high structural and sequential homology indicating that subtle variations generate substantial functional diversity. Judging whether or not a specific construct is well-produced comes down to comparison. As a benchmark reference, we were using the production levels of the well studied plant-AQP SoPIP2;1, knowing that the yield of functional protein has proven adequate for upcoming purification, functional and structural experimental proceedings.

Quantification was performed by comparing Western blot-signals of samples from each protein corresponding to total protein production and membrane localized fraction. Furthermore, the cultivation, purification as well as the Western blot quantification, was performed in sample triplicates, resulting in an average construct-specific signal showing relatively small variations between signal triplets. In addition, a sample of purified hAQP1 of known concentration was used as an internal protein standard, making it possible to compare different Western blots.

Using the software Multi Gauge[®], a precise determination of signal strength (pixel saturation) within a defined area was obtained. The average value of the triplicates of SoPIP2;1 was set to be 1.0. By scaling all other values to this, a relative signal (relatively to SoPIP2;1) was obtained for each hAQP. Additionally, SoPIP2;1 was used as a positive growth control at each cultivation. Later, comparing total production levels from SoPIP2;1 cultivated at different time points provided additional support to the quality of the Western blot-signal quantification.

3.1.1.2 High Variations in Production Level

At first surprisingly, the production levels of the hAQPs showed considerable variation. In summary, seven of the 13 hAQPs (hAQP0, hAQP1, hAQP3, hAQP5, hAQP7, hAQP10 and hAQP12) were successfully produced in *P. pastoris* at sufficient levels for further experimental proceedings. Furthermore, two hAQPs (hAQP2 and hAQP8) generated moderate amounts, just at the verge of useful production level.

3. RESULTS AND DISCUSSION

In addition, four hAQPs (hAQP4, hAQP6, hAQP9 and hAQP11) generated inconsiderable amounts of produced protein. Figure 13 shows the AQPs phylogenetic relationship as well as their production level from our study, where the high yielding AQPs are marked with green boxes, intermediate yield with yellow boxes and the low yielding with red boxes. A swift comparison of the production variation could not be related to phylogenetic relationship *i.e.* classic AQPs and AQGP, indicating a more intricate correlation.

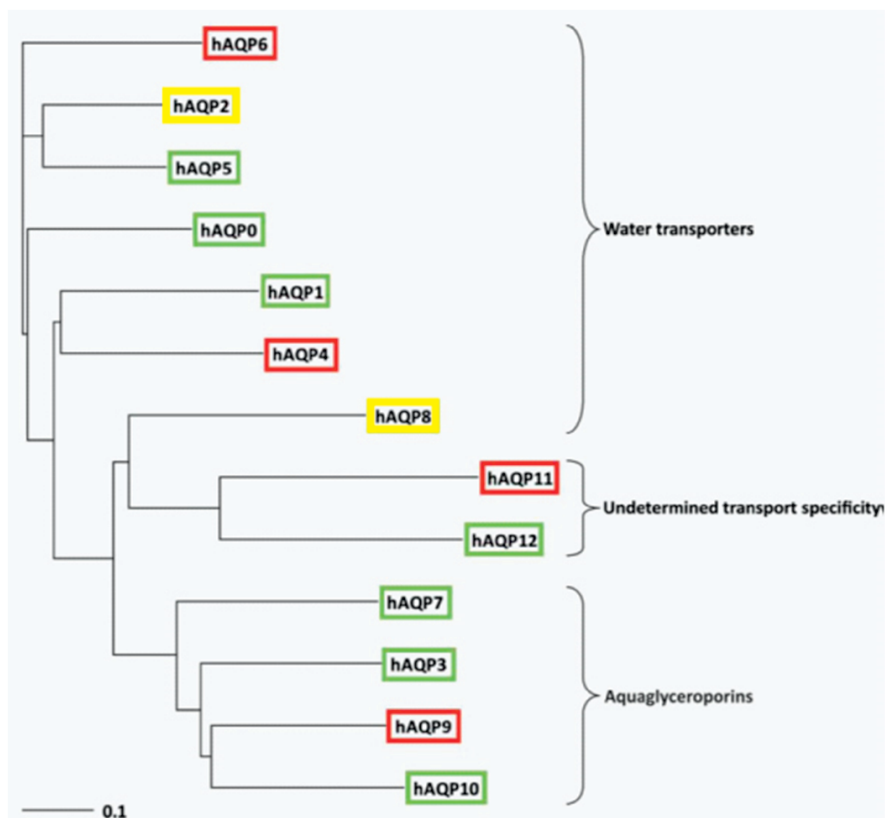


Figure 13. Phylogenetic tree representing all hAQPs. High, intermediate and low yielding protein producers are indicated in green, yellow and red boxes, respectively. Sequences were aligned with ClustalW.

3. RESULTS AND DISCUSSION

3.1.1.3 High Variations in Membrane Localization

Even though the production of a membrane protein works well, the pathway leading to membrane insertion may not. Ideally every produced membrane protein molecule gets correctly folded and transported and into the membrane. However, this might not always be the case. By using cell fractionation and isolating the membranes, we were able to evaluate the total production level as well as the membrane-localized fraction for each protein (Fig.14). Consequently, the ratio gives an indication of how well the produced protein incorporates into the membrane. Once again SoPIP2;1 was used as reference due to previous production and functionality studies indicating a well produced, localized and folded protein. The resulting total protein produced contra membrane-localized ratios of the nine hAQPs exhibiting useful production levels, displayed high individual variation. In contrast, the two hAQPs (hAQP2 and hAQP8) generating just detectable production level, were both inserted into the membrane to a high extent.

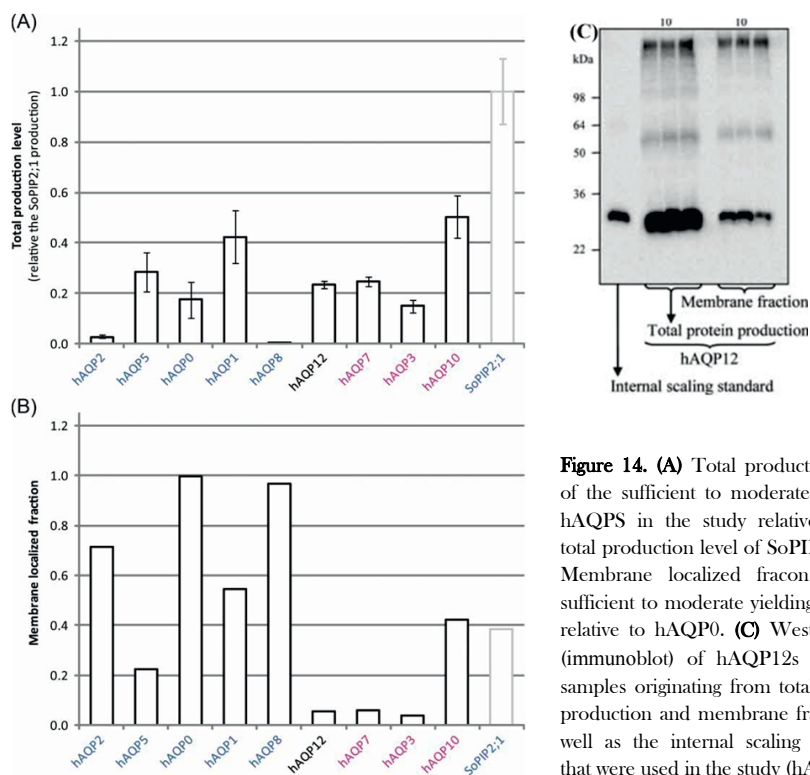


Figure 14. (A) Total production level of the sufficient to moderate yielding hAQPs in the study relative to the total production level of SoPIP2;1. (B) Membrane localized fraction of the sufficient to moderate yielding hAQPs relative to hAQP0. (C) Western blot (immunoblot) of hAQP12s triplicate samples originating from total protein production and membrane fraction as well as the internal scaling standard that were used in the study (hAQP1).

3. RESULTS AND DISCUSSION

In addition, a functionality assay of purified hAQP2 verified functionality by proteoliposome shrinkage. Consequently, a moderate total protein production level may be compensated by a high level of membrane insertion, hence making it worthwhile to proceed further. In contrast, a high production level doesn't necessarily generate a high membrane insertion, exhibited by hAQP5 and hAQP12. Thus, the decisive question will be if either high levels of total production or membrane insertion will compensate inadequate levels of their counterpart?

3.1.1.4 Observed Correlation Factors to High Yield

Interestingly, the extent of membrane insertion (but not total protein production) of the hAQPs could be correlated to the hAQPs subfamilies *i.e.* phylogenetic relationship. Comparing the hAQPs producing high to moderate yields, five of the classic hAQPs was inserted efficiently into the membrane, while among the hAQGPs only hAQP10 was inserted into the membrane efficiently. By speculation, this may depend on factors differentiating the classic hAQPs and hAQGPs apart, such as wider channel and higher monomeric stability of the hAQGPs but also in the lack of native AQGP's in *P. pastoris*. Accordingly, *P. pastoris* might provide insufficient support for the produced hAQGPs, thus influence the membrane insertion negatively. In addition, it was noticed that the nucleotide sequence adjacent to the ATG start codon appears to effect the total protein production.

The hAQPs resume too differ, particularly considering folding pathways. For instance, the hAQP4 folding pathway requires a co-translational translocation process in the endoplasmic reticulum (ER) to achieve final topology. This varies considerably compared to hAQP1 which folds via a four-TM helical spanning topology, stabilizing the tetramer which might impact the final yield. In our study, hAQP4 production level was not noticeable whereas hAQP1 belong to one of the best producing hAQPs. Sequential analysis of hAQP1 reveals 3 polar residues (Asn49, Lys51 in TMH2 and Asp185 in TMH5) responsible for directing and stabilizing the tetramer. In addition, corresponding residues in hAQP4 are hydrophobic (Met48, Leu50 and Asn184), argumentally responsible for the different folding pathway. This hypothesis were tested by constructing two constructs of hAQP4 (hAQP4-M48N-L50K and hAQP4-M48N-L50K-N184D) mimicking hAQP1s folding pathway.

3. RESULTS AND DISCUSSION

Indeed, the production levels were highly increased and high levels of membrane insertion of the triple-mutant were recorded (Fig.15). This verifies the significance of a stable folding pathway in membrane insertion. However, no correlation between tetramer stability and high total production yield could be drawn.

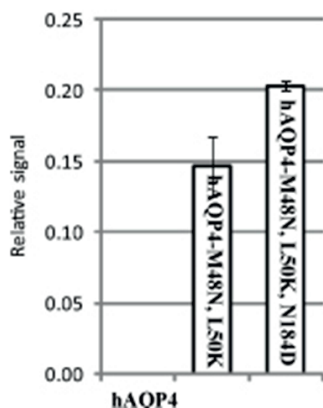


Figure 15. Total produced protein originating from hAQP4 wild-type, the hAQP4-M48N, L50K dubbel-mutant and the hAQP4-M48N,L50K,N184D trippel-mutant, related to hAQP4 yield.

3.1.2 Production of Zebrafish (*Danio rerio*) TH (**Paper V**)

The TH-genes from five different eukaryotic species (human, bovine, zebrafish, mouse and rat) were transformed in *P. pastoris* and tested in a small-scale production screen. One zebrafish TH-transformant (ZTH-wt) was selected to proceed into protein purification based on comparative Western blots. Due to the high instability displayed by the ZT-wt, a TH-construct (ZTH-cys) having all 13 native cysteines replaced by serines, were constructed and cloned in *P. pastoris*.

3.1.2.1 Diversity in Purification, Stability and Activity

The protein purification profiles deviated between the two constructs. ZTH-wt displayed low binding of to the IMAC column. Subsequent gel-filtration resulted in a small peak containing dimeric TH for which activity could be observed, although not in a reproducible manner. Moreover, a large peak at the void volume indicate high levels of aggregation. ZTH-cys displayed a stronger binding affinity which could suggest that the His-tag is more exposed. Surprisingly, a difference in binding affinity was also seen when protease inhibitors were included, resulting in elution at lower imidazole concentration. Alternatively the difference in affinity between

3. RESULTS AND DISCUSSION

the two constructs could be explained by the ZTH-cys construct containing a N-terminal 8xHis-tag whereas the ZTH-wt construct contains an C-terminal 6xHis-tag which might influence the binding capability (Fig.16A and B).

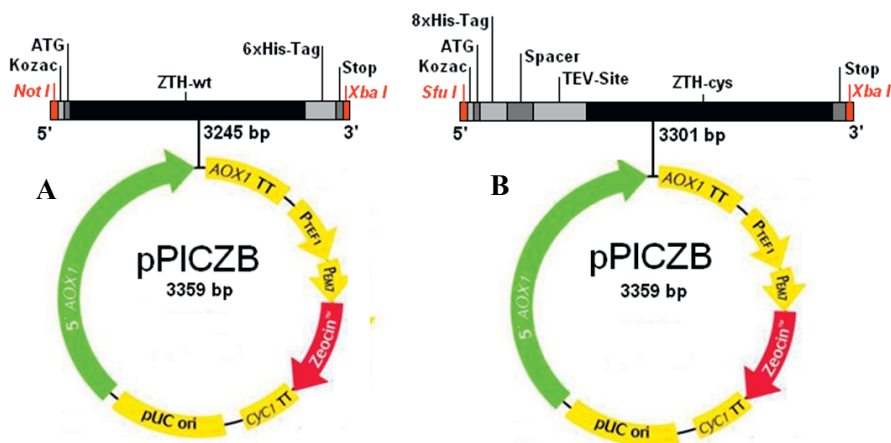


Figure 16. (A) Schematic illustration of the plasmid constructs ZTH-wt and (B) ZTH-cys. The ZTH-wt construct has a C-terminal 6xHis-tag whereas the ZTH-cys construct has N-terminal 8xHis-tag as well as a TEV protease site which enables removal of the His-tag.

The ZTH-cys IMAC chromatogram resulted in two peaks, indicating a polydisperse population (Fig.17A). Furthermore, the two IMAC peaks resulted in different gel-filtration profiles when loaded separately (Fig.17C and D).

Peak 1 generated consistently less aggregate in the gel-filtration void-volume eluting mainly as a dimer (Fig.17B, lane 1) at a similar elution volume as the active form of ZTH-wt. However, the sample was not pure. On the contrary, the ZTH-cys population from peak 2 generated both monomers and dimers (Fig17B, lane 2) for which the monomers were inclined to aggregate. It is therefore possible that cysteines play an important role in intramolecular interactions which keep TH in an dimeric conformation. Moreover, as we failed to observe activity at any point of the ZTH-cys purification, the cysteines seem to be important for functionality as well.

3. RESULTS AND DISCUSSION

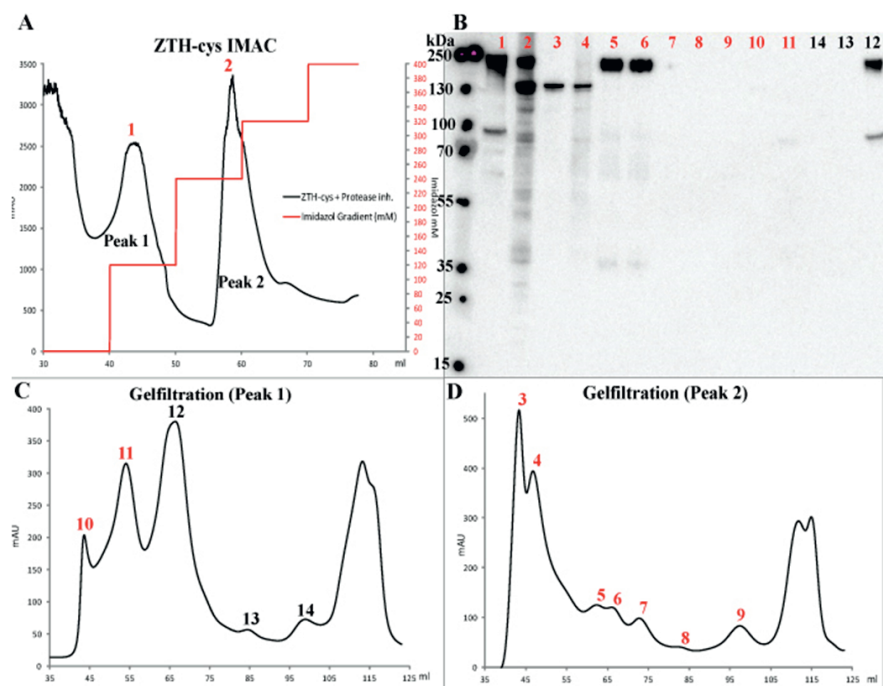


Figure 17. (A) ZTH-cys IMAC-chromatogram indicating two distinctive peaks. (B) Western blot (immunoblot) of IMAC and gelfiltration samples originating both from peak 1 and 2. (C) Gelfiltration chromatogram using fractions originating from IMAC peak 1. (D) Gelfiltration chromatogram using fractions originating from IMAC peak 2.

3.1.3 Comparative Aspects of hAQPs and TH Production

Comparing the purification stability of the hAQPs and ZTH-constructs used in my research, couldn't be more diverse. Although the constructs share the same cloning vector and production host as well as the fact that they are both membrane proteins, illustrates the many considerations that must be taken into account in membrane protein production. However, viewing retrospectively on the results and consider the characteristics of the proteins might provide some insight especially towards improving the ZTH purification protocol.

One of the most noticeable differences during the membrane purification process is that the hAQPs are able to withstand Urea and NaOH-washes whereas the ZTH washes away when using such harsh washes. This is most likely due to the fact that AQPs in general are more embedded into the membrane than THs which are anchored to the membrane only by dII, corresponding approximately to 1/3 of THs molecular weight. Thereby, the

3. RESULTS AND DISCUSSION

hAQP-membrane washes gave the succeeding purification steps a better starting position due to the fact that it was possible to remove loosely attached membrane proteins to a greater extent, having less amount of contaminating proteins in the membrane.

Furthermore, the production host *P. pastoris*, includes native AQP but not native TH. Therefore, it is possible to reason that recombinantly produced hAQPs experience comparable trafficking pathways and folding machinery as in a native human cell, upholding stability for newly produced hAQPs. In addition, the lipidic composition of *P. pastoris* cellular membranes might likewise be in favor for hAQPs.

In vivo, AQPs functions as tetramers which clearly is a preferable conformation as been observed on SDS-PAGE and Western blots. In the case of ZTH protein purification, there is a tendency for the *in vivo*-preferred dimer to break down into monomers and aggregate, implying that the dimeric interactions not sufficiently uphold conformation upon solubilization, as both the dimeric and tetrameric hAQP conformations possess. It has thus become clear that further optimization of the ZTH purification conditions are needed before structural studies can be initialized.

3.1.4 Production of SoPIP2;1

The SoPIP2;1 construct in *P. pastoris* (formerly PM28A), has proven to be an exceptional membrane protein producer, yielding up to 25mg of purified protein/ liter ^[113]. *P. pastoris* native trafficking pathways and folding mechanics seem to be a peculiarly well-matched combination in every production aspect. In addition, it is likely that multiple copies of the construct were inserted during the homologous recombination process, generating a “jackpot-clone” during the transformation. Consequently, the SoPIP2;1 construct and associated mutants produced in *P. pastoris*, has given rise to several high-resolution structures ^[19, 51, 64].

3.2 Structural and Functional Analysis of SoPIP2;1 (Paper II, III and IV)

The plant AQP SoPIP2;1 present a complex channel gating mechanism, indicating that the open and closed states involve an increased or decreased probability of channel opening. Earlier studies as well as MD (molecular

3. RESULTS AND DISCUSSION

dynamics) simulations have suggested that phosphorylation of residues involved in stabilisation of the closed state could cause channel opening. The details of SoPIP2;1s molecular gating mechanism could therefore be investigated by mimicking phosphorylation of these residues and studying the structural effect as well the water transport ability.

3.2.1 Gating of SoPIP2;1 by Phosphorylation

Previously, biochemical as well as structural studies of SoPIP2;1 have indicated two highly interesting residues located at central interaction positions; S115 and S274^[64]. The S115 residue on loop B interacts with an N-terminal cation (Ca^{2+} , Cd^{2+}) binding site which, when occupied, stabilise the capping D-loop by hydrogen bond (H-bond) interactions between multiple residues and water molecules. In addition, Ca^{2+} and Cd^{2+} have shown to have an inhibitory effect in functional assays indicating a decreased channel opening probability. The residue S274 is situated at the C-terminus and mediates D-loop stabilisation through interactions with residues at the end of loop D in the neighbouring monomer in the tetramer. Molecular dynamics (MD) simulations of SoPIP2;1 of phosphorylated at both S115 and S274 revealed a considerable D-loop destabilisation by disrupting D-loop stabilising interactions, thereby increasing the probability of channel opening.

In addition, new potential sites were identified as able to inflict D-loop destabilisation upon phosphorylation. In particular, the central key residue S188 enforced an internal H-bond network which stabilised the D-loop. Although not seen to be phosphorylated *in vivo*, S188 sits in a phosphorylation consensus sequence for the same kinase known to phosphorylate S274.

Furthermore, in the closed structure of wild-type SoPIP2;1, S188 (positioned on the D-loop) interacts with a water molecule (W133) and residues D184 and R187, generating H-bond interactions stabilising the internal structure of loop D. (Fig.18). Phosphorylation of S188 would most likely disrupt these interactions, causing a destabilisation of the closed conformation of loop D. This was supported by MD-simulations of SoPIP2;1 phosphorylated on S188, in which an interaction was seen between this residue and the C-terminus, causing channel opening.

3. RESULTS AND DISCUSSION

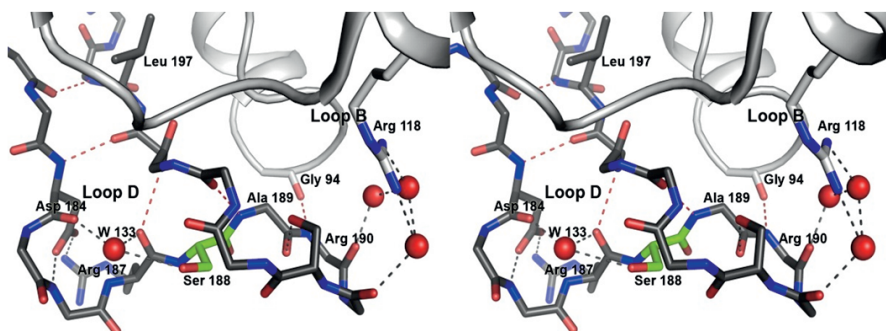


Figure 18. Stereo view of wild-type SoPIP2;1-closed structure. Hydrogen bond network and loop B interactions are shown with dotted lines, water molecules are shown in red and S188 is highlighted in green.

3.2.2 Mimicking the Biological On/Off Switch

Many biological molecules have phosphorylation sites which function as an activator, altering its function. The phosphorylation involves the addition of a phosphate group (PO_4^{3-}) often added to a serine (S), threonine (T), tyrosine (Y) or histidine (H) residue in eukaryotic proteins. The addition of a phosphate group disrupts stable molecular binding networks, resulting in conformational changes.

The residues aspartate (D, Asp) and glutamate (E, Glu), resembles a phosphate group by being relatively bulky and negatively charged which is the main reason to use these in a protein to mimic a phosphorylated state. There are several reported studies where D or E point mutations generated activated proteins^[114-116]. However, even if D and E resemble the addition of a phosphate group, it doesn't necessarily mean they are equivalents. Consequently, a "real" phosphorylated side chain will possess a larger negative charge (-2 at physiological pH) as well as it will be occupying a larger space.

To mimic phosphorylation of SoPIP2;1 and study its effect, we made four constructs; three single mutants (S115E, S188E and S274E) and one double mutant (S115E:S274E). The constructs were successfully transformed into *P. pastoris*, produced, purified, crystallized and assayed for water transport in proteoliposomes. Structural data from X-ray diffraction was collected and structures were solved by molecular replacement using the wild-type structure of SoPIP2;1, PDB code 1Z98 as the search model^[51, 64], (Table 2).

3. RESULTS AND DISCUSSION

Protein structure	So-PIP2;1-wt	So-PIP2;1-wt	S115E	S115E:S274E	S274E	So-PIP2;1-Hg	S188E
PDB code	1Z98	2B5F	3CCL	3CN5	3CN6	4JC6	
Conformation	Closed	Open	Closed	Closed	Closed	Closed	Open
Space group	I4	P2 ₁ 2 ₁ 2	I4	I4	I4	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2
Resolution	2.1Å	3.9Å	2.3Å	2.05Å	2.95Å	2.15Å	3.0Å

Table 2. Comparison of the SoPIP2;1 structures.

3.2.3 Mimicking Phosphorylation of Ser115 and Ser274 Does Not Cause Channel Opening

Surprisingly, even though the effort was to induce an open conformation, the single mutants S115E and S274E as well as the double mutant S115E:S274E crystallized in closed conformations. Superimposing the structures displays close similarities to the closed structure of wild-type SoPIP2;1 but with some structural deviations.

The S115E and S115E:S274E mutants presented the N-terminus protruding into the cytoplasm through an extension of TM helix 1 (Fig.19). This result differ from closed wild-type SoPIP2;1 in which a small helix is formed, harbouring the cation binding site which is involved in stabilising the closed conformation.

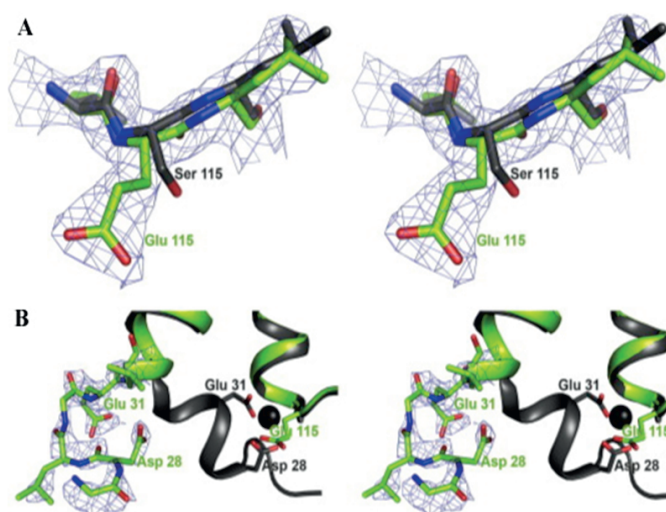


Figure 19. Stereo images of superimposed wild-type SoPIP2;1 (grey) and S115E mutant (green). The electron density belongs to the S115E mutant. **(A)** View of the S115 mutation site showing the prolonged residue of S115E. **(B)** Illustration of the effect of the S115E mutant resulting in a disruption of the divalent cation binding site and extension of TM helix 1. The cation is illustrated as a black ball.

3. RESULTS AND DISCUSSION

The extended N-terminus of S115E and S115E:S274E completely disrupts the cation binding site. Specifically, the added Glu115 interrupts the former Ser115-Glu31 H-bond interaction by repelling the negatively charged Glu31. Consequently, Glu31 C_α is replaced 10Å, compared to wild-type closed SoPIP2;1. Nevertheless, the D-loop remains closed creating other connections to the B-loop by H-bond interaction via two water molecules. The structures of single mutant S274E and S115E:S274E double mutant, shows a disordered C-terminus, suggesting that when S274 is phosphorylated, the stabilising interaction between this residue and loop D in the neighbouring monomer seen in the closed wild-type structure of SoPIP2;1 is lost.

It is possible that the crystallization progress itself induces D-loop stabilisation of the phosphomimicking mutants. Protein molecules who rearrange in a well-ordered pattern may give rise to new interaction possibilities and decreased freedom which could in our case stabilise the closed conformation. To reject crystallization effects, the phosphomimicking mutants were further investigated by a functionality assay (stopped-flow) measuring water transport rate (water permeability) in proteoliposomes. Indeed, the proteoliposome stopped-flow assay showed that S115E, S274E and S115E:S274E were in a closed state also in liposomes as their water permeability coefficient (P_f) was comparable to that of wild-type SoPIP2;1 (Fig.20A).

The inability of the S115E and S274E mutations to generate channel opening could be due to these mutations not fully mimicking the phosphorylated state. Analysis of electrostatic potential maps showed that the phosphorylated serine created significantly more negative potential around the serine than did the phosphomimicking mutation. This additional negative potential could provide the driving force needed to cause a full rearrangement of loop D.

3.2.4 Significant Increase in Water Flux by the S188E mutant

In contrast to the other phosphomimicking mutants of SoPIP2;1, S188E showed a significant increase in water transport indicating that the channel could be in an open conformation (Fig.20A). This could explain why this construct was reluctant to rearrange in an ordered manner (crystallize). It is

3. RESULTS AND DISCUSSION

likely that a flexible D-loop hinders formation of well-ordered crystals as only low resolution data has been obtained for SoPIP2;1 in an open conformation.

Quantitative Western blots indicated similar liposome insertions (Fig.20B). In addition, the relative P_f to empty liposomes (Fig.21C, gray bar) as well as the specific P_f , relating water permeability to the liposome incorporation (Fig.20C, purple bar) was calculated, indicating that the P_f of the S188E mutant does not depend on excessively liposome incorporation relative to the other constructs.

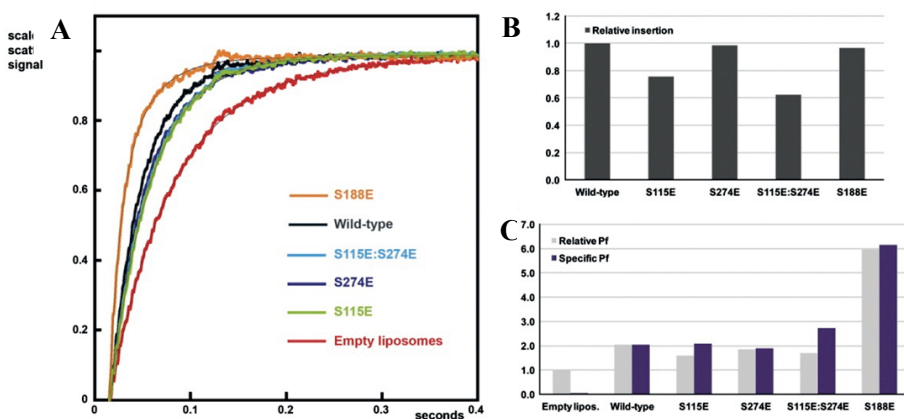


Figure 20. (A) Results of the water transport assay (stopped-flow spectroscopy) showing that the water transport rate of S188E (orange) clearly was higher than S115E (green), S274E (dark blue) and S115E:S274E (light blue) which was similar to wild-type SoPIP2;1 (black). (B) Quantitative Western blot-signals (immunoblot) of incorporated protein in the proteoliposomes relative to SoPIP2;1 wt. (C) Bars indicating water permeability (light gray) which was calculated by relating measured permeability to empty liposomes. The bars indicating specific water permeability (purple) was calculated by relate water permeability to the degree of proteoliposomes incorporation.

The increased P_f of the S188E mutant argue that this single point mutation in fact increase the probability of channel opening. Apparently the H-bond interactions made by the residues S188, D184 and R187 along with the water molecule W133, is essential in stabilising the D-loop (Fig.18). Consequently, upon phosphorylation the residue S188 gets displaced away from the W133 resulting in disconnection of the former H-bond interaction network. In addition, this destabilisation would also effect the interactions between the D-loop and B-loop (G94, R118 and R190) and disconnect the capping D-loop, increasing the probability of channel opening.

3.2.5 Structure of S188E mutant is Open

In contrast to the other phosphomimicking mutants, the S188E mutant turned out to be quite difficult to crystallize. However, after several crystallization trials (producing >200 crystals at 4°C), one crystal condition (**Paper IV**) generated a well defined diffraction pattern and data to 2.7Å was collected. The data was processed and scaled using iMosflm^[117] and Scala^[118]. Unfortunately, the data was highly twinned (perfect pseudomerohedrally twinned), making structural interpretations quite difficult. Despite of that, the structure could be solved at 3Å by MR, using the closed SoPIP2;1 tetramer, PDB code 1Z98 as the model, (section 3.2.2, table 2) and by the use of special composite-omit-maps (Methodology-section: 2.2.7.1). New appearing electron density arose from consecutive composite-omit-map calculations, each with added residues, resulting in gradually improved electron density along a growing D-loop (Fig.21).

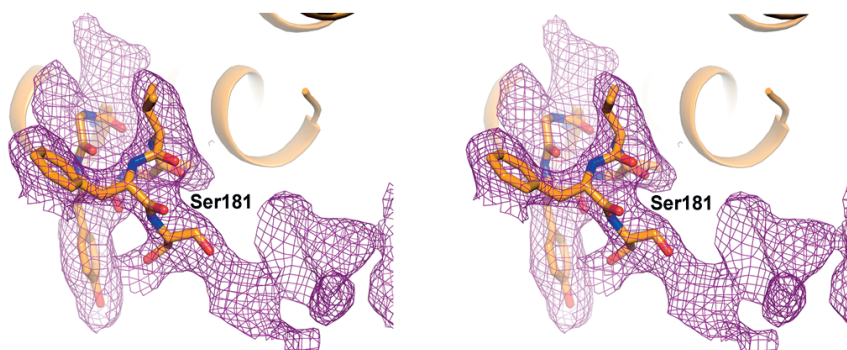


Figure 21. Stereo view of the electron density generated from the initial SoPIP2;1 S188E composite-omit-map (at 0.6σ), indicating where along the D-loop (monomer A) the next amino acids will be added.

Residues continued to be added to the point where no new electron density was appearing from the latest composite-omit-map. Noticeable, one of the monomers (monomer A) seemed to have a more ordered D-loop by displaying clearer electron density after each successive composite-omit-map. Accordingly, the A monomers D-loop was the only one that could be successfully completed whereas only partial electron density could be observed by the other monomers, hinting that the great disorder could result from open D-loops.

3. RESULTS AND DISCUSSION

Structural comparison between the SoPIP2;1 closed and open structure and the S188E mutant, confirms the D-loops flexibility. In the SoPIP2;1 closed structure (Fig.22A, green), the D-loop caps the channel and prevents water transport which on the other hand is accessible in the SoPIP2;1 open conformation (Fig.22A, purple).

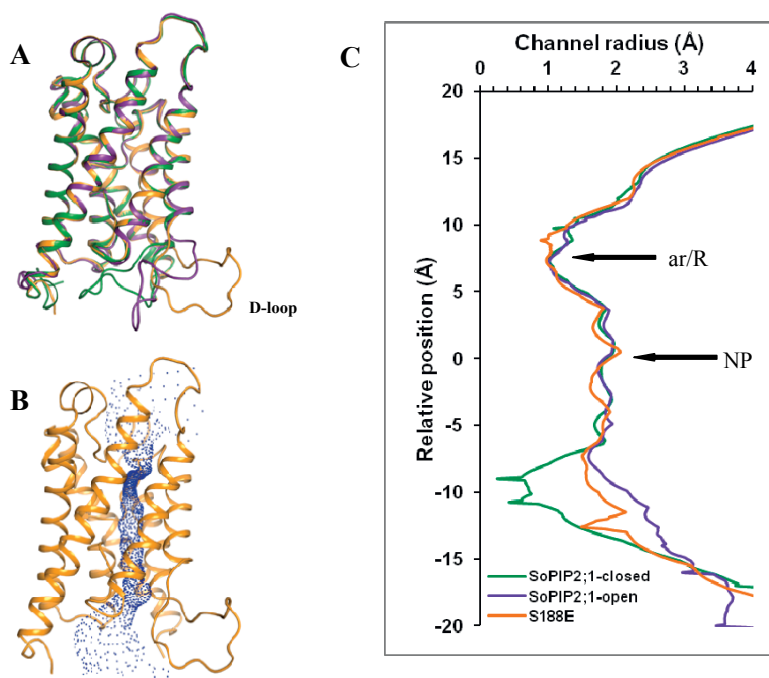


Figure 22. (A) Superimposed monomeric structures of wild-type SoPIP2;1 in closed conformation (green), wild-type SoPIP2;1 in open conformation (purple) and the SoPIP2;1 S188E mutant in open conformation (gold) showing three apparent positions of the D-loop. (B) Structure of S188E and calculated channel (blue) (HOLE [120]). (C) Channel radius of wild-type SoPIP2;1 in closed conformation (green), wild-type SoPIP2;1 in open conformation (purple) and the S188E mutant in open conformation (orange). The y-axis = 0, at the centre of the NPA motif. Location of the NPA motif and ar/R selectivity filter is pointed out by arrows.

However, in the S188E mutant, the D-loop extends further away from the channel opening (Fig.22A, gold), reaching towards the tetrameric centre.

Furthermore, the S188E channel dimensions was calculated (HOLE ^[119]) and compared to wild-type SoPIP2;1 open and closed conformation (Fig.22C). Channel positions (Fig.22C, y-axis) were related to the centre of the NPA-motif (Asn222) of the SoPIP2;1 closed structure. Reasonably, both the SoPIP2;1 open and the S188E conformation showcase comparable

3. RESULTS AND DISCUSSION

channel dimensions whereas the SoPIP2;1 closed conformation showed significant constriction at the D-loop region (Fig.22C, at the relative position -10). These results fit well with the stooped-flow assay (Fig.20A), indicating that mimicking phosphorylation of S188 causes the D-loop interactions to destabilise and open the channel. However, due to quality restrictions, no side-chains could be modelled to the S188E structure which could have given further insight into conceivable D-loop stabilizing interactions.

Further on, comparing the crystal packing of closed SoPIP2;1 in and S188E display similar arrangements having the cytoplasmic interfaces of two tetramers directed to each other (Fig.23).

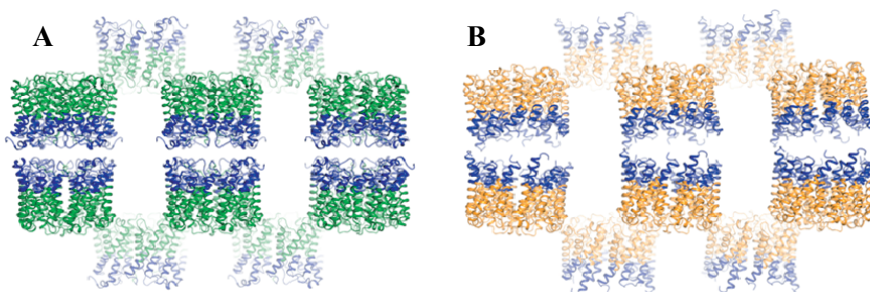


Figure 23. (A) Crystal packing of SoPIP2;1 in closed conformation. (B) Crystal packing of S188E in open conformation. The tetramers cytoplasmic interfaces are highlighted in blue.

However, in the S188E crystal packing (Fig.23.B), the D-loop causes the interface of the two tetramers to tilt towards one side when in fact the interface is straight in the closed SoPIP2;1 crystal packing (Fig.23A). It is highly probable that the D-loop causes the tilt in the S188E crystal packing, indicating that the D-loop is unchained and in an open conformation.

3.2.6 The Effect of Cations on SoPIP2;1

Divalent cations have been shown to have an effect on water transport through plant AQPs in several studies. In particular Cd^{2+} and Ca^{2+} are known to inhibit water transport through plant plasma membrane AQPs such as SoPIP2;1. For AQPs in general, mercury (Hg^{2+}) is often seen to have an inhibitory effect ^[120-123]. In fact, the ability of Hg^{2+} to inhibit water transport across the membrane of red blood cells was a key experiment in

3. RESULTS AND DISCUSSION

identifying the very first AQP, AQP1 [18, 124]. The residue responsible for Hg^{2+} inhibition of AQP1 was later identified as Cys189 in the water-conducting pore [125]. While many AQPs are indeed inhibited by Hg^{2+} , others are in fact insensitive [126, 127].

Surprisingly, adding Hg^{2+} to SoPIP2;1 increased the P_f (+99.5%) compared to when no Hg^{2+} was present. Increased water permeability after Hg^{2+} -treatment has only previously reported for AQP6 [128]. Furthermore, adding Hg^{2+} to liposomes containing S188E, results only in a mediocre increase in P_f (+21.2%). Presumably, adding Hg^{2+} to S188E effects P_f relatively little since this is already situated in an open conformation, supporting that Hg^{2+} affects the conformational state of SoPIP2;1. To study this structural effect, we solved the structure of SoPIP2;1 in complex with Hg^{2+} at 2.15Å (table 2). Contrary to what was expected the structure showed significant similarities to closed SoPIP2;1 also having the D-loop capping the channel.

In the structure of the SoPIP2;1- Hg^{2+} complex, Hg^{2+} was seen to bind to 3 out of 4 cysteines. To investigate which one of these were responsible for the mercury activation, we constructed mutants in which the cysteines were replaced with serines and investigated their water transport ability using proteoliposomes. None of the mutants displayed a water permeability significantly different from wild-type, suggesting Hg^{2+} activation of SoPIP2;1 is not dependent on cysteines. However, an alternative could be that Hg^{2+} affects the liposome membrane which induce a conformational change of SoPIP2;1. In fact, Hg^{2+} is known to be able to react with the polar head groups of the liposome lipids which thereby change the membrane properties, resulting in increased membrane fluidity [129].

Changes effecting membrane properties such as lipidic composition and outer mechanical forces has been suggested to have a physiological role *in vivo* in plant AQP gating [130]. The indications that plant AQPs may be regulated mechanosensitively is quite reasonable because plants generally needs a direct responsive reaction upon mechanical stimuli.

4. Conclusions

This thesis is based on the results which have emerged from various protein production experiments intended for X-ray crystallography structural studies. A general cloning and production methodology has been utilized, allowing for comparison of the data throughout the process. Although not every project developed into a new structure, important aspects and considerations have been made regarding high-yield production of pure and stable membrane protein.

To kinds of eukaryotic membrane protein have been in focus, the AQPs and THs. The hAQPs displayed a considerable variance concerning production yield and membrane localisation in spite of the sequential similarity (**Paper I**). However, the outcome from the hAQP production resulted in the conclusion that an effective trafficking pathway, leading to proper membrane insertion, might surpass a higher production yield, thus still providing adequate yields for structural studies. Importantly, an ineffective trafficking pathway may be improved by mimicking a more effective transporting and membrane insertion process.

However, producing well membrane-inserted proteins is far from enough. Solubilized TH (**Paper V**) turned out to be highly unstable. Removing all cysteines resulted in higher stability but on the other hand the cysteine-free TH became inactive. Further protocol optimization would most likely uncover a better suited purification environment. Nevertheless, it was concluded that full-length eukaryotic TH can be successfully produced in *P. pastoris* at sufficient yields. In addition, the importance of cysteines in eukaryotic TH dimerization and activity was established.

Further on, the very nature of plants and their habitat suggests differences in their AQP molecular mechanisms compared to hAQPs. For instance, a human can just walk away if realizing its standing in a lake while a plant has to adapt to its surroundings, one method of which is AQP gating. Phosphomimicking mutants of the gated spinach AQP SoPIP2;1, were tested in functionality assays as well as structurally determined (**Papers II, IV**). The proposed gating mechanism in which phosphorylation of the residues S115 and S274 result in channel opening appeared reasonable from examining MD simulations. Yet the crystal structures and functional data of the S115E and S274E single and double mutants, respectively,

4. CONCLUSIONS

indicate the channel being closed, most likely due not these not fully mimicking the phosphorylated state.

Although phosphorylation of S188 *in vivo* has yet to be shown, the model of S188E is the first structural evidence of a phosphomimicking mutant causing channel opening (**Paper IV**). However, the 3Å diffraction data from S188E proved to be highly twinned which as first was a big disappointment. But, the use of a special composite-omit-map designed for twinning, provided modelling without inducing additional model bias into the emerging D-loop. In the end, only one of the tetramers D-loop could be completely modelled, revealing an novel conformation of the D-loop where it extends away from the channel opening towards the tetrameric centre. The fact that S188E is in an open conformation is further reinforced by functionality assays, crystal packing arrangement, the crystallization unwillingness and the D-loop disorder seen in the monomers B,C and D's electron density.

Finally, the S188E structure constituted the last component in the “gene to structure” process, which gave me the possibility to have a highly diverse and fulfilling graduate period.

5. Acknowledgments

My time as a PhD-student has been a time which I will look back at with great joy. I truly have enjoyed waking up in the morning, well at about 8AM, and start produce great science which has contributed to many articles, well 5 articles, under a short time period, well 5 years. I have also had the privilege to get to know many nice and fascinating people, ~~well-except~~, and got the opportunity to travel aboard for interesting and rewarding conferences, typically associated to tropical beaches and/or an appealing attraction. However, there are supporting colleagues, friends and family that needs to be acknowledged:

Susanna, thank you for accepting me as your PhD student! You have shown great patience, understanding and desire to share your knowledge. It has been extremely valuable to watch how you solve problem, rationalize and generally relate to research. I have been very fortunate to have you as my supervisor and role model as a researcher. After all, why not try and mimic Sweden's top crystallographer.

Richard, I have always enjoyed your rapid and thoughtful seminar questions (when you haven't asked them to me ;) making you think that little extra. In addition, I still like making Western-blot. However nowadays, I refer to them as: "Immunoblots, Western-style".

Kristina, I appreciated all the cloning discussions and advices that you given. You have always had time to help answer any questions. I will miss the lunch discussions talking about training, kindergarten and life in general.

Former and current PhD students in the group: **Anna**, thanks for all the help with the 188-structure/manuscript, couldn't have made it without you! **Etienne**, thanks for helping out with the thermo-stability assay and at the ESRF! **Mike**, it has been fun to have you in the Susanna-group as well as talking MMORPG's with you. **Amit**, thanks for helping out with the TH-project and for working late and on weekends doing ÄKTA-purification etc, really appreciate it. **Povilas**, thanks for a great diploma work and for all the time you spent in the cold-room. Good luck in your own PhD!

Fredrik, most of the techniques/methods used in the lab, I learnt from you, even the construct-name system (The FO-system). Thanks again for all the help and the great times we had in the lab, teaching and on conferences ("Oh my gah!"-Mykonos). **Gerhard**, förutom roliga konferenser och det du hjälpt mig i labbet/ESRF med så är det den tyska Ohrenreiner'n som kanske varit mest uppskattad, wax on/off! Lycka till med din post-doc. **Linda**, från x-jobbare till doktorand, alltid varit kul att undervisa, luncha och att åka på konferenser tillsammans, fast Pepsi-light kommer jag avstå.

WeiXiao, alltid varit skoj och prata med dig vid lunch och fikastunder, protein engeneering festen blev bra. Hoppas hummern var god! **David**, kul med nån på labbet som man kunnat prata sport och öl (kaffé) med, igen, ingen mer Pepsi-light.

5. ACKNOWLEDGMENT

Karin R, kakan jag ska inte hålla någon längre MONOLOG, uppskattat luncher, kurser och konferenserna vi varit på. BTW, jag är skyldig dig en säng! **Madde**, snacka MMA, barnuppföstran och proteinuttryck på rasterna har varit super! Du har kommit med många bra råd, tack. **Jennie**, det har varit kul och dela kontor med dig, rådfråga och prata om allt möjligt, lycka till i framtiden! **Annette**, tack för bra samarbete i Prot.eng. undervisningen och festen, det har varit kul! **Pierro**, nice having you in the lab and thanks for starting the fermentors when I was sick!

Ida H, Rhawnie, Petra E, Cecilia, Alexander, Linnéa, Ænders, Maria S, Karl, Emelie, Niada, Leif, David, Rajiv, Elin, Oskar, Rob, Emil, Alex, Rebecka, Petra, Rosie, Gergely, Sebastian W, Peter, Gisela, Kiran, Susmita, Maria, Rob H, Sebastian P, Stephan, Veronica, Gabriella, Ingrid, it has been great knowing you all, having sushi, lunches and meetings together. I wish you all the best!

Daniel, Billy, Vlad, er x-jobbstid gick snabbt! Det var kul och spela badminton, tjöta StarCraft och träning med er. Måste snart bli dags för Kina-buffé, Mange bjuder?

Oldies: **Maria, Pontus, Karin H, Magnus, Rob, Johan, Urzula, Erik, Anne-Marie, Nicklas, Daniel F, Jonas & Jonas**, tack för råden, umgänget och välkommandet!

Lars, Bruno, Valida, ni fixar och underhåller så verksamheten flyter på, bra jobbat!

Martin, Örjan, J.R Rydström, tack för biokemikurserna, undervisning och TH hjälp.

Daniel, Fredrik, Save, Alex, Peter, basku, Linné, TNG, N64, grill, WoW, Risk. Ni har till mycket format mig till den jag är. **Look what you have done!**

Henrik och Oscar, bröder som jag ser upp till. Ert stora intresse för naturvetenskap, brittisk humor och pizza i kombination med styrketräning har varit inspirerande. Det går att pressa 130 bänken som akademiker! **Ann-Sofie, Malin**, ni blir också del av vår familj, alltid kul med er! Snart starta X-boxen med Lips?

Lena, Nils-Arvid, Ulla, Leif, tack för norrlandsäventyr och Slapperhäng. Vår familj uppskattar alltid ert sällskap! Ser fram emot fira jular/påskar/födelsedagar etc. med er.

Lisa, Bernt, ni är som vår familj och uppskattar er mycket. Hoppas på fler trevliga möten i framtiden!

Bodil, Susanna, Johan, Tobias, Rasmus, vi har upplevt mycket roligt ihop och alltid känt oss varmt välkomna till Eksjö av er, tack!

Dad, Marlene, Tina, Marc, our vent in the Swedish winter. Hope to see you soon!

Mamma, alltid uppmuntrande och starkt stöd till allt jag gjort. Glad att ha dig, tack!

Annie, du tagit hand om barnen och hemmet och verkligen ställt upp under alla sena kvällar och helger. Tack! Jag älskar dig!

6. References

1. Darwin, C., *On the Origin of Species*. 1859, London: John Murray.
2. Miller, S.L., *A production of amino acids under possible primitive earth conditions*. *Science*, 1953. **117**(3046): p. 528-9.
3. Turing, A.M., *The chemical basis of morphogenesis*. 1953. *Bull Math Biol*, 1990. **52**(1-2): p. 153-97; discussion 119-52.
4. Jänig, W., *The integrative action of the autonomic nervous system : neurobiology of homeostasis*. 2006, Cambridge, UK ; New York: Cambridge University Press. xxi, 610 p.
5. Amos, H., *Introduction to Cell to Cell Communication*. In *Vitro-Journal of the Tissue Culture Association*, 1970. **6**(1): p. 1-&.
6. Padan, E. and H. Rottenberg, *Respiratory control and the proton electrochemical gradient in mitochondria*. *Eur J Biochem*, 1973. **40**(2): p.431-7.
7. Danielson, L. and L. Ernster, *Energy-Dependent Reduction of Triphosphopyridine Nucleotide by Reduced Diphosphopyridine Nucleotide, Coupled to the Energy-Transfer System of the Respiratory Chain*. *Biochem Z*, 1963. **338**: p. 188-205.
8. Zinn, K.M. and I.H. Leopold, *Molecular biology of the cell*. *Int Ophthalmol Clin*, 1973. **13**(3): p. 53-81.
9. Carp, G., *Cell and Molecular Biology: Concepts and Experiments*. John Wiley & Sons, 2008. **5**.
10. Brachet, J., *Cell biology and molecular biology*. *Adv Cytopharmacol*, 1971. **1**: p. 1-12.
11. Singer, S.J. and G.L. Nicolson, *The fluid mosaic model of the structure of cell membranes*. *Science*, 1972. **175**(4023): p. 720-31.
12. Morange, M., *What history tells us XXX. The emergence of the fluid mosaic model of membranes*. *J Biosci*, 2013. **38**(1): p. 3-7.
13. Brown, D.A. and E. London, *Functions of lipid rafts in biological membranes*. *Annu Rev Cell Dev Biol*, 1998. **14**: p. 111-36.
14. Guidotti, G., *Membrane proteins*. *Annu Rev Biochem*, 1972. **41**: p. 731-52.
15. von Heijne, G., *The membrane protein universe: what's out there and why bother?* *J Intern Med*, 2007. **261**(6): p. 543-57.
16. von Heijne, G., *Principles of membrane protein assembly and structure*. *Prog Biophys Mol Biol*, 1996. **66**(2): p. 113-39.
17. Wallin, E. and G. von Heijne, *Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms*. *Protein Sci*, 1998. **7**(4): p. 1029-38.
18. Preston, G.M., et al., *Appearance of water channels in Xenopus oocytes expressing red cell CHIP28 protein*. *Science*, 1992. **256**(5055): p. 385-7.
19. Hedfalk, K., et al., *Aquaporin gating*. *Curr Opin Struct Biol*, 2006. **16**(4): p. 447-56.
20. Guzzo, A.V., *The influence of amino-acid sequence on protein structure*. *Biophys J*, 1965. **5**(6): p. 809-22.
21. Anfinsen, C.B., *Principles that govern the folding of protein chains*. *Science*, 1973. **181**(4096): p. 223-30.
22. Spiegel, A.M., *Defects in G protein-coupled signal transduction in human disease*. *Annu Rev Physiol*, 1996. **58**: p. 143-70.

6. REFERENCES

23. Hopkins, A.L. and C.R. Groom, *The druggable genome*. Nat Rev Drug Discov, 2002. **1**(9): p. 727-30.
24. 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.
25. Imming, P., C. Sinning, and A. Meyer, *Drugs, their targets and the nature and number of drug targets (vol 5, pg 821, 2006)*. Nature Reviews Drug Discovery, 2007. **6**(2): p. 126-126.
26. Martin, A.C., et al., *Protein folds and functions*. Structure, 1998. **6**(7): p. 875-84.
27. Thornton, J.M., et al., *From structure to function: approaches and limitations*. Nat Struct Biol, 2000. **7 Suppl**: p. 991-4.
28. Wright, P.E. and H.J. Dyson, *Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm*. J Mol Biol, 1999. **293**(2): p. 321-31.
29. Nayeem, M.S. and R.H. Khan, *Misfolded proteins and human diseases*. Protein Pept Lett, 2004. **11**(6): p. 593-600.
30. Lundstrom, K., *Structural genomics: the ultimate approach for rational drug design*. Mol Biotechnol, 2006. **34**(2): p. 205-12.
31. King, L.S., M. Yasui, and P. Agre, *Aquaporins in health and disease*. Mol Med Today, 2000. **6**(2): p. 60-5.
32. Grisshammer, R. and C.G. Tate, *Overexpression of integral membrane proteins for structural studies*. Q Rev Biophys, 1995. **28**(3): p. 315-422.
33. Forstner, M., L. Leder, and L.M. Mayr, *Optimization of protein expression systems for modern drug discovery*. Expert Rev Proteomics, 2007. **4**(1): p. 67-78.
34. White, S., *Membrane Proteins of Known Structure*. 2013.
35. Prive, G.G., *Detergents for the stabilization and crystallization of membrane proteins*. Methods, 2007. **41**(4): p. 388-97.
36. Bergfors, T., *Screening and optimization methods for nonautomated crystallization laboratories*. Methods Mol Biol, 2007. **363**: p. 131-51.
37. Sidel, V.W. and A.K. Solomon, *Entrance of Water into Human Red Cells under an Osmotic Pressure Gradient*. Journal of General Physiology, 1957. **41**(2): p. 243-257.
38. 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.
39. Agre, P., S. Sasaki, and M.J. Chrispeels, *Aquaporins: a family of water channel proteins*. Am J Physiol, 1993. **265**(3 Pt 2): p. F461.
40. Carbrey, J.M. and P. Agre, *Discovery of the aquaporins and development of the field*. Handb Exp Pharmacol, 2009(190): p. 3-28.
41. Engel, A., Y. Fujiyoshi, and P. Agre, *The importance of aquaporin water channel protein structures*. EMBO J, 2000. **19**(5): p. 800-6.
42. Krane, C.M. and D.L. Goldstein, *Comparative functional analysis of aquaporins/glyceroporins in mammals and anurans*. Mamm Genome, 2007. **18**(6-7): p. 452-62.
43. Frigeri, A., G.P. Nicchia, and M. Svelto, *Aquaporins as targets for drug discovery*. Current Pharmaceutical Design, 2007. **13**(23): p. 2421-2427.
44. Castle, N.A., *Aquaporins as targets for drug discovery*. Drug Discov Today, 2005. **10**(7): p. 485-93.
45. King, L.S., et al., *Defective urinary-concentrating ability due to a complete deficiency of aquaporin-1*. N Engl J Med, 2001. **345**(3): p. 175-9.

6. REFERENCES

46. Agre, P., et al., *Aquaporin water channels--from atomic structure to clinical medicine*. J Physiol, 2002. **542**(Pt 1): p. 3-16.
47. Javot, H. and C. Maurel, *The role of aquaporins in root water uptake*. Ann Bot, 2002. **90**(3): p. 301-13.
48. Maurel, C., et al., *Molecular physiology of aquaporins in plants*. Int Rev Cytol, 2002. **215**: p. 105-48.
49. Kaldenhoff, R., et al., *Characterization of plant aquaporins*. Methods Enzymol, 2007. **428**: p. 505-31.
50. Chaumont, F., et al., *Aquaporins constitute a large and highly divergent protein family in maize*. Plant Physiol, 2001. **125**(3): p. 1206-15.
51. Tornroth-Horsefield, S., et al., *Structural mechanism of plant aquaporin gating*. Nature, 2006. **439**(7077): p. 688-94.
52. Kaldenhoff, R. and M. Fischer, *Aquaporins in plants*. Acta Physiol (Oxf), 2006. **187**(1-2): p. 169-76.
53. Johanson, U., et al., *The complete set of genes encoding major intrinsic proteins in Arabidopsis provides a framework for a new nomenclature for major intrinsic proteins in plants*. Plant Physiol, 2001. **126**(4): p. 1358-69.
54. Jung, J.S., et al., *Molecular-Structure of the Water Channel through Aquaporin Chip - the Hourglass Model*. Journal of Biological Chemistry, 1994. **269**(20): p. 14648-14654.
55. Gonen, T. and T. Walz, *The structure of aquaporins*. Q Rev Biophys, 2006. **39**(4): p. 361-96.
56. Murata, K., et al., *Structural determinants of water permeation through aquaporin-1*. Nature, 2000. **407**(6804): p. 599-605.
57. 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.
58. Savage, D.F., et al., *Structural context shapes the aquaporin selectivity filter*. Proc Natl Acad Sci U S A, 2010. **107**(40): p. 17164-9.
59. 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.
60. Zhang, Y.B. and L.Y. Chen, *In silico study of Aquaporin V: Effects and affinity of the central pore-occluding lipid*. Biophys Chem, 2013. **171**: p. 24-30.
61. Huang, S., *And now there are five: a new player in intracellular trafficking pathways*. PLoS Biol, 2011. **9**(10): p. e1001173.
62. 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.
63. Tornroth-Horsefield, S., et al., *Structural insights into eukaryotic aquaporin regulation*. Febs Journal, 2010. **277**: p. 21-21.
64. 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.
65. Danielson, L. and L. Ernster, *Demonstration of a mitochondrial energy-dependent pyridine nucleotide transhydrogenase reaction*. Biochem Biophys Res Commun, 1963. **10**: p. 91-6.
66. Eytan, G.D., et al., *Energy-linked nicotinamide-nucleotide transhydrogenase. Characterization of reconstituted ATP-driven transhydrogenase from beef heart mitochondria*. J Biol Chem, 1987. **262**(11): p. 5008-14.
67. Studley, W.K., et al., *Phylogenetic analyses of proton-translocating transhydrogenases*. Microb Comp Genomics, 1999. **4**(3): p. 173-86.

6. REFERENCES

68. Hou, C., M. Potier, and P.D. Bragg, *Cross-Linking and Radiation Inactivation Analysis of the Subunit Structure of the Pyridine-Nucleotide Transhydrogenase of Escherichia-Coli*. *Biochimica Et Biophysica Acta*, 1990. **1018**(1): p. 61-66.
69. Kaplan, N.O., S.P. Colowick, and E.F. Neufeld, *Pyridine nucleotide transhydrogenase. II. Direct evidence for and mechanism of the transhydrogenase reaction*. *J Biol Chem*, 1952. **195**(1): p. 107-19.
70. Colowick, S.P., et al., *Pyridine nucleotide transhydrogenase. I. Indirect evidence for the reaction and purification of the enzyme*. *J Biol Chem*, 1952. **195**(1): p. 95-105.
71. Kaplan, N.O., S.P. Colowick, and E.F. Neufeld, *Pyridine nucleotide transhydrogenase. III. Animal tissue transhydrogenases*. *J Biol Chem*, 1953. **205**(1): p. 1-15.
72. Arkblad, E.L., et al., *A Caenorhabditis elegans mutant lacking functional nicotinamide nucleotide transhydrogenase displays increased sensitivity to oxidative stress*. *Free Radic Biol Med*, 2005. **38**(11): p. 1518-25.
73. Hoek, J.B. and J. Rydstrom, *Physiological roles of nicotinamide nucleotide transhydrogenase*. *Biochem J*, 1988. **254**(1): p. 1-10.
74. Pandolfi, P.P., et al., *Targeted disruption of the housekeeping gene encoding glucose 6-phosphate dehydrogenase (G6PD): G6PD is dispensable for pentose synthesis but essential for defense against oxidative stress*. *EMBO J*, 1995. **14**(21): p. 5209-15.
75. Raha, S. and B.H. Robinson, *Mitochondria, oxygen free radicals, disease and ageing*. *Trends Biochem Sci*, 2000. **25**(10): p. 502-8.
76. Freeman, H., et al., *Nicotinamide nucleotide transhydrogenase: a key role in insulin secretion*. *Cell Metab*, 2006. **3**(1): p. 35-45.
77. Cotton, N.P., et al., *The crystal structure of an asymmetric complex of the two nucleotide binding components of proton-translocating transhydrogenase*. *Structure*, 2001. **9**(2): p. 165-76.
78. Jackson, J.B., *Proton translocation by transhydrogenase*. *FEBS Lett*, 2003. **555**(1): p. 176-7.
79. Rydstrom, J., et al., *Domains, specific residues and conformational states involved in hydride ion transfer and proton pumping by nicotinamide nucleotide transhydrogenase from Escherichia coli*. *Biochim Biophys Acta*, 1998. **1365**(1-2): p. 10-6.
80. Wierenga, R.K., M.C.H. Demaeyer, and W.G.J. Hol, *Interaction of Pyrophosphate Moieties with Alpha-Helices in Dinucleotide Binding-Proteins*. *Biochemistry*, 1985. **24**(6): p. 1346-1357.
81. Yamaguchi, M. and Y. Hatefi, *Energy-transducing nicotinamide nucleotide transhydrogenase. Nucleotide binding properties of the purified enzyme and proteolytic fragments*. *J Biol Chem*, 1993. **268**(24): p. 17871-7.
82. Bhakta, T., et al., *Structures of the dI2dIII1 complex of proton-translocating transhydrogenase with bound, inactive analogues of NADH and NADPH reveal active site geometries*. *Biochemistry*, 2007. **46**(11): p. 3304-18.
83. Johansson, T., et al., *X-ray structure of domain I of the proton-pumping membrane protein transhydrogenase from Escherichia coli*. *J Mol Biol*, 2005. **352**(2): p. 299-312.
84. Prasad, G.S., et al., *Crystal structure of transhydrogenase domain III at 1.2 Å resolution*. *Nat Struct Biol*, 1999. **6**(12): p. 1126-31.
85. Jackson, J.B., et al., *The alternating site, binding change mechanism for proton translocation by transhydrogenase*. *Biochemistry*, 2002. **41**(13): p. 4173-85.

6. REFERENCES

86. Jackson, J.B., et al., *Proton-translocating transhydrogenase in bacteria*. Biochem Soc Trans, 1993. **21**(4): p. 1010-3.
87. Lis, H. and N. Sharon, *Protein glycosylation. Structural and functional aspects*. Eur J Biochem, 1993. **218**(1): p. 1-27.
88. Yokoyama, S., *Protein expression systems for structural genomics and proteomics*. Curr Opin Chem Biol, 2003. **7**(1): p. 39-43.
89. Midgett, C.R. and D.R. Madden, *Breaking the bottleneck: eukaryotic membrane protein expression for high-resolution structural studies*. J Struct Biol, 2007. **160**(3): p. 265-74.
90. 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.
91. Higgins, D.R., *Overview of protein expression in Pichia pastoris*. Curr Protoc Protein Sci, 2001. **Chapter 5**: p. Unit5 7.
92. 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.
93. Horsefield, R., et al., *High-resolution x-ray structure of human aquaporin 5*. Proc Natl Acad Sci U S A, 2008. **105**(36): p. 13327-32.
94. Tate, C.G., *Overexpression of mammalian integral membrane proteins for structural studies*. FEBS Lett, 2001. **504**(3): p. 94-8.
95. Tate, C.G., et al., *Comparison of seven different heterologous protein expression systems for the production of the serotonin transporter*. Biochim Biophys Acta, 2003. **1610**(1): p. 141-53.
96. Wegner, E.H., *Biochemical conversions by yeast fermentation at high cell densities*. Application number US 06/316,164, 1983.
97. Veenhuis, M., J.P. Van Dijken, and W. Harder, *The significance of peroxisomes in the metabolism of one-carbon compounds in yeasts*. Adv Microb Physiol, 1983. **24**: p. 1-82.
98. Lee, J.D. and K. Komagata, *Taxonomic Study of Methanol-Assimilating Yeasts*. Journal of General and Applied Microbiology, 1980. **26**(2): p. 133-158.
99. Invitrogen, *Pichia Fermentation Process Guidelines*. 2002.
100. Cregg, J.M., et al., *Recombinant protein expression in Pichia pastoris*. Mol Biotechnol, 2000. **16**(1): p. 23-52.
101. Li, P.Z., et al., *Expression of recombinant proteins in Pichia pastoris*. Applied Biochemistry and Biotechnology, 2007. **142**(2): p. 105-124.
102. Edwards, A.M., et al., *Protein production: feeding the crystallographers and NMR spectroscopists*. Nat Struct Biol, 2000. **7 Suppl**: p. 970-2.
103. Chayen, N.E., *Comparative studies of protein crystallization by vapour-diffusion and microbatch techniques*. Acta Crystallographica Section D-Biological Crystallography, 1998. **54**: p. 8-15.
104. Asherie, N., *Protein crystallization and phase diagrams*. Methods, 2004. **34**(3): p. 266-72.
105. Rontgen, W.C., *On a New Kind of Rays*. Science, 1896. **3**(59): p. 227-31.
106. Evans, G., D. Axford, and R.L. Owen, *The design of macromolecular crystallography diffraction experiments*. Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 261-70.
107. Brunger, A.T., *Assessment of phase accuracy by cross validation: the free R value. Methods and applications*. Acta Crystallogr D Biol Crystallogr, 1993. **49**(Pt 1): p. 24-36.

6. REFERENCES

108. Artymiuk, P.J. and C.C. Blake, *Refinement of human lysozyme at 1.5 Å resolution analysis of non-bonded and hydrogen-bond interactions*. J Mol Biol, 1981. **152**(4): p. 737-62.
109. 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.
110. Cladera, J., et al., *Liposome solubilization and membrane protein reconstitution using Chaps and Chapso*. Eur J Biochem, 1997. **243**(3): p. 798-804.
111. Rigaud, J.L., *Membrane proteins: functional and structural studies using reconstituted proteoliposomes and 2-D crystals*. Braz J Med Biol Res, 2002. **35**(7): p. 753-66.
112. Hupfeld, S., et al., *Liposome size analysis by dynamic/static light scattering upon size exclusion-/field flow-fractionation*. J Nanosci Nanotechnol, 2006. **6**(9-10): p. 3025-31.
113. Karlsson, M., et al., *Reconstitution of water channel function of an aquaporin overexpressed and purified from Pichia pastoris*. FEBS Lett, 2003. **537**(1-3): p. 68-72.
114. Huang, W. and R.L. Erikson, *Constitutive activation of Mek1 by mutation of serine phosphorylation sites*. Proc Natl Acad Sci U S A, 1994. **91**(19): p.8960-3.
115. McSorley, T., et al., *Mimicking phosphorylation of Ser-74 on human deoxycytidine kinase selectively increases catalytic activity for dC and dC analogues*. FEBS Lett, 2008. **582**(5): p. 720-4.
116. Ouyang, Y., et al., *Structural and kinetic effects of PAK3 phosphorylation mimic of cTnI(S151E) on the cTnC-cTnI interaction in the cardiac thin filament*. J Mol Biol, 2010. **400**(5): p. 1036-45.
117. Batty, T.G., et al., *iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM*. Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 271-81.
118. Kabsch, W., *Automatic Processing of Rotation Diffraction Data from Crystals of Initially Unknown Symmetry and Cell Constants*. Journal of Applied Crystallography, 1993. **26**: p. 795-800.
119. Smart, O.S., et al., *HOLE: a program for the analysis of the pore dimensions of ion channel structural models*. J Mol Graph, 1996. **14**(6): p. 354-60, 376.
120. Daniels, M.J., et al., *Characterization of a new vacuolar membrane aquaporin sensitive to mercury at a unique site*. Plant Cell, 1996. **8**(4): p. 587-99.
121. Roberts, S.K., et al., *Cholangiocytes express the aquaporin CHIP and transport water via a channel-mediated mechanism*. Proc Natl Acad Sci U S A, 1994. **91**(26): p. 13009-13.
122. Lazowski, K.W., et al., *Evidence for the presence of a Hg-inhibitable water-permeability pathway and aquaporin 1 in A5 salivary epithelial cells*. J Cell Physiol, 1995. **164**(3): p. 613-9.
123. Fushimi, K., et al., *Functional characterization and cell immunolocalization of AQP-CD water channel in kidney collecting duct*. Am J Physiol, 1994. **267**(4 Pt 2): p. F573-82.
124. Macey, R.I., *Transport of water and urea in red blood cells*. Am J Physiol, 1984. **246**(3 Pt 1): p. C195-203.
125. Preston, G.M., et al., *The Mercury-Sensitive Residue at Cysteine-189 in the Chip28 Water Channel*. Journal of Biological Chemistry, 1993. **268**(1): p. 17-20.

6. REFERENCES

126. Hasegawa, H., et al., *Molecular cloning of a mercurial-insensitive water channel expressed in selected water-transporting tissues*. J Biol Chem, 1994. **269**(8): p. 5497-500.
127. Daniels, M.J., T.E. Mirkov, and M.J. Chrispeels, *The plasma membrane of Arabidopsis thaliana contains a mercury-insensitive aquaporin that is a homolog of the tonoplast water channel protein TIP*. Plant Physiol, 1994. **106**(4): p. 1325-33.
128. Yasui, M., et al., *Rapid gating and anion permeability of an intracellular aquaporin*. Nature, 1999. **402**(6758): p. 184-7.
129. Delnomdedieu, M., et al., *Interaction of Mercury-Chloride with the Primary Amine Group of Model Membranes Containing Phosphatidylserine and Phosphatidylethanolamine*. Biochimica Et Biophysica Acta, 1989. **986**(2): p. 191-199.
130. Wan, X., E. Steudle, and W. Hartung, *Gating of water channels (aquaporins) in cortical cells of young corn roots by mechanical stimuli (pressure pulses): effects of ABA and of HgCl₂*. J Exp Bot, 2004. **55**(396): p. 411-22.