



LUND UNIVERSITY

From Sequence to Structure- Characterizing Human and Plant Aquaporins

Nordén, Kristina

2012

[Link to publication](#)

Citation for published version (APA):

Nordén, K. (2012). *From Sequence to Structure- Characterizing Human and Plant Aquaporins*. [Doctoral Thesis (compilation), Biochemistry and Structural Biology]. Department of Chemistry, Lund University.

Total number of authors:

1

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

From Sequence to Structure

Characterizing Human and Plant Aquaporins

Kristina Nordén



LUND
UNIVERSITY

Department of Biochemistry and Structural Biology

2012



From Sequence to Structure

Characterizing Human and Plant Aquaporins

Copyright © Kristina Nordén 2012

All rights reserved

Department of Biochemistry and Structural Biology

Center for Molecular Protein Science

P.O. Box 124

SE-221 00 Lund

ISBN 978-91-7422-304-0

Printed in Sweden by Media-Tryck, Lund University
Lund 2012

Abstract

Aquaporins constitute a class of membrane proteins that create pores through the lipid bilayer of biological membranes, facilitating the passage of water and other small solutes. Nearly all living organisms possess a more or less intricate set up of aquaporin proteins and the importance of these channels has been implicated in many aspects of health and disease. The aim of the projects summarized in this thesis was to increase the knowledge about the structure and function of aquaporins. A method for obtaining high levels of recombinantly expressed membrane proteins in the yeast *Pichia pastoris* by screening for clones with multiple recombinant gene inserts was developed. Subsequent to overexpression, a number of different purification schemes for obtaining preparations for molecular characterization were established and three aquaporin isoforms; human *HsAQP5* and *HsAQP8* and a plant aquaporin, *NbXIP1;1*, were successfully purified. The two human isoforms were reconstituted into artificial liposomes and by functionality assessments their water transporting capacity was confirmed. The verification of retained functionality after recombinant expression and subsequent purification trials is crucial in order to validate the quality of the protein that goes into crystallization and structure determination trials. By using two different crystallization techniques, diffracting crystals of both *HsAQP5* and *HsAQP8* were obtained and subsequently, a high resolution structure of *HsAQP5* was solved.

A different approach was used to identify hypothetical aquaporins in the chloroplast, which is the organelle responsible for the conversion of sunlight into chemical energy in plants. By immunoblotting procedures and mass spectrometry analyses, the presence of several aquaporin isoforms was confirmed, suggesting a possible dependence of the photosynthetic reactions on an assisted trans-membrane water flux.

List of publications

This thesis is based on the following papers, which will be referred to by their Roman numerals in the text. The papers are appended at the end of the thesis.

- I. Increasing gene dosage greatly enhances recombinant expression of aquaporins in *Pichia pastoris***
Nordén K, Agemark M, Danielson JÅH, Alexandersson E, Kjellbom P, Johanson U
BMC Biotechnology, 11: 47, 2011
- II. Overexpression and purification of the *Nicotiana benthamiana* XIPI1;1 aquaporin**
Ampah-Korsah H, Nordén K, Kjellbom P, Johanson U
Manuscript
- III. Reconstitution of water channel function and 2D-crystallization of human AQP8**
Agemark M, Kowal J, Kukulski W, Nordén K, Gustavsson N, Johanson U, Engel A, Kjellbom P
Biochimica et Biophysica Acta, 1818: 839-850, 2012
- IV. High resolution x-ray structure of human aquaporin 5**
Horsefield R*, Nordén K*, Fellert M, Backmark A, Törnroth-Horsefield S, Terwisscha van Scheltinga AC, Kvassman J, Kjellbom P, Johanson U, Neutze R
Proc Natl Acad Sci USA, 105 (36): 13327-32, 2008
- V. PIP water channels of the chloroplast inner envelope and thylakoid membranes**
Nordén K, Agemark M, Lambert W, Kjellström S, Andersson A, Danielsson R, Cem Emek S, Emanuelsson C, Albertsson PÅ, Johanson U, Kjellbom P
Manuscript

**These authors contributed equally to this paper*

Contribution to the papers

- I.** KN participated in the design of the study, constructed the *P. pastoris* clones, performed cell culturing and subsequent preparations, carried out immunoblotting assays, qPCR experiments and qPCR data analyses. KN drafted the manuscript, took part in the subsequent revision and was the corresponding author.
- II.** KN participated in the design of the study, supervised parts of the work and performed qPCR experiments. KN drafted the manuscript.
- III.** KN took part in the protein production and reconstitution experiments. KN took part in the revision of the manuscript.
- IV.** KN took part in the design of the study, constructed the *P. pastoris* clones, and optimized the purification protocol. KN performed functional studies and analyzed the data. KN took part in crystallization screens and in the drafting and revision of the manuscript.
- V.** KN took part in the design of the study, in membrane preparations and in immunoblotting and mass spectrometry analyses. KN took part in analysis of the data and in the drafting of the manuscript.

Additional publications

A novel plant major intrinsic protein in *Physcomitrella patens* most similar to bacterial glycerol channels

Gustavsson S, Lebrun AS, Nordén K, Chaumont F, Johanson U
Plant physiol, 139 (1); 287-95, 2005

Affinity tags can reduce merohedral twinning of membrane protein crystals

Backmark A, Nyblom M, Törnroth-Horsefield S, Kosinska-Eriksson U, Nordén K, Fellert M, Kjellbom P, Johanson U, Hedfalk K, Lindkvist-Petersson K, Neutze R, Horsefield R.
Acta Crystallogr D Biol Crystallogr, 64(Pt 11):1183-6, 2008

Contents

Preface	9
Aquaporins	11
2.1. Structural Features	11
2.2 Proton Exclusion and Substrate Specificity	13
2.3 Regulation	15
2.4 Plant Aquaporins	16
2.5 Aquaporins in Man	23
Methods	29
3.1 Recombinant Protein Expression	29
3.2 Membrane Protein Purification	36
3.3 Functionality Assessments	39
3.4 Crystallization and Structure Determination	42
3.5 Identification of Aquaporins at Novel Sites	47
Summary of the Present Investigation	49
4.1. Overexpression	49
4.2. Purification	52
4.3. <i>In Vitro</i> Functionality Assessments	54
4.4. Crystallization and Structure	55
4.5. Aquaporins in Chloroplasts	57
Populärvetenskaplig sammanfattning	59
Acknowledgements	61
References	63

Preface

The diversity of life forms on planet earth is mind-boggling, ranging from tiny unicellular microbes to giant whales and huge sequoia trees. However, one feature that unites all living species is their requirement to distinguish all biochemical processes essential for survival from the surrounding environment. This is achieved by the presence of the lipid membrane that surrounds the smallest unit of life, the cell. Such membranes are also responsible for the confinement of subcellular compartments. Besides constituting a boundary that defines the cell and its organelles, the membrane is the residence for 20-30% of all protein species that exists (Wallin and von Heijne 1998). These membrane proteins are involved in numerous biological processes such as enzymatic reactions, intercellular communication and trans-membrane transport of substances.

Many membrane proteins are positioned at the cell boundary and are commonly involved in signaling pathways in response to extracellular cues. This feature together with the fact that they, again due to their localization, are readily available for modulation by pharmaceutical molecules in the extracellular fluid makes this class of proteins attractive as drug targets. Approximately 60% of the marketed medical substances are directed towards membrane proteins (Yildirim *et al.* 2007).

Even though almost a third of all existing protein species are membrane localized and despite the biochemical and pharmacological interest in these proteins, only a small fraction of the high resolution protein structures deposited in the protein data bank derive from crystalized membrane proteins. This humble quantity is explained by the commonly sparse expression levels of membrane proteins in nature and by their embedment in the lipid bilayer, which are both factors that complicates their purification. Since protein crystallization, which is a prerequisite for structure determination, depends on large quantities of pure protein, success is seldom acquired using membrane proteins from native sources. Hence, optimized protocols for heterologous expression in recombinant cells as well as elaborate purification schemes are imperative to increase the knowledge about membrane protein structure and function.

In this thesis the results of the projects that I have engaged myself in the last few years are summarized. The work has been focused on the characterization of a class of integral membrane channels, referred to as aquaporins, which facilitate the passive diffusion of water and small solutes across the membranes of the cell. The identification of aquaporins at hitherto unknown subcellular sites is presented and

the results are put into a biological context. Moreover results from *in vitro* aquaporin characterization experiments are presented. Recombinant aquaporin gene sequences were successfully incorporated into the genome of the yeast *Pichia pastoris* and efficient protein production was obtained by optimizing gene dosage and culturing conditions. Three aquaporin isoforms were successfully purified from *P. pastoris* and two of these were subsequently functionally characterized and crystallized. In one case a high resolution membrane protein structure was obtained, and the journey *from sequence to structure* was accomplished.

Aquaporins

In order to retain the osmotic balance in the cell, water needs to pass over the membranes enclosing it and its intracellular compartments. Due to the diverse water permeation properties of different biological membranes, an aided water transport was suggested already in the 1960s. However, it was not until Professor Peter Agre and his coworkers purified and characterized the first water channel aquaporin 1 (AQP1), from red blood cells in the mid-1980s through the early 1990s, that the assisted water transport theory was fully acknowledged (Denker et al. 1988; Preston and Agre 1991; Preston et al. 1992). The discovery of the water transporting aquaporin protein earned Professor Peter Agre the Nobel Prize in Chemistry in 2003.

Even though AQP1 was the first aquaporin to be purified and proven to create a water channel in the membrane, other members of the same channel family had previously been recognized. For instance, the highly abundant MIP (later renamed AQP0), expressed in the eye lens fiber had been reported to have a predicted channel-like structure (Gorin et al. 1984) and the *Escherichia coli* glycerol facilitator GlpF was identified already in the 1970s (Richey and Lin 1972).

In the last decades, immense efforts have been made to characterize the aquaporins. Besides AQP1, and 12 additional aquaporin isoforms subsequently identified in mammals, a considerable number of members of this protein family have been discovered in other animals, plants, bacteria and fungi. Moreover an unexpected complexity regarding substrate specificities and biological functions has emerged.

2.1. Structural Features

The basic aquaporin topology, with six transmembrane helices (H1-H6) connected by five loops (A-E) was early on predicted by hydrophathy plotting (Gorin *et al.* 1984; Preston and Agre 1991). At the same time, the two tandem repeats that constitute the aquaporin polypeptide, that arose from an ancient gene duplication event, were recognized (Pao *et al.* 1991). When sequence information of different members of the aquaporin family started to sum up, it soon became clear that two amino acid stretches containing the asparagine-proline-alanine (NPA) motif,

situated in the B-loop and E-loop respectively, were highly conserved among isoforms (Reizer *et al.* 1993). By introducing a series of point mutations in these conserved elements it was revealed that these amino acid residues are involved in the water transporting mechanism of AQP1, and the “hourglass model” with loop B and E dipping into the membrane, positioning the NPA motifs at the center of the membrane and at this position creating a constriction, was presented for the first time (Jung *et al.* 1994).

While attempts were made to biochemically characterize the aquaporins, structural studies were also initiated. Soon, the first low resolution structural data of AQP1 was presented and a tetrameric organization of the protein in the membrane, as had previously been predicted (Smith and Agre 1991) was confirmed (Walz *et al.* 1994). When resolution was improved, additional structural features such as the right-handed helical bundle structure of the monomer (Walz *et al.* 1997) and the presence of two half α -helices in the center of the monomer (Mitsuoka *et al.* 1999) were revealed.

Since the 1990s a number of high resolution structures of aquaporins have been solved and they all share the same basic structural features: They assemble as tetramers, with each monomer creating a pore through the membrane. Each monomer is built up of six membrane spanning helices with a seventh transmembrane domain being formed by the two half helices of loop B and E, that inserts into the membrane from opposite sides. These two short helices (HB and HE) harbor the conserved aquaporin NPA signature sequences, important for function and selectivity (discussed below). The basic aquaporin topology is presented in Figure 2.1.

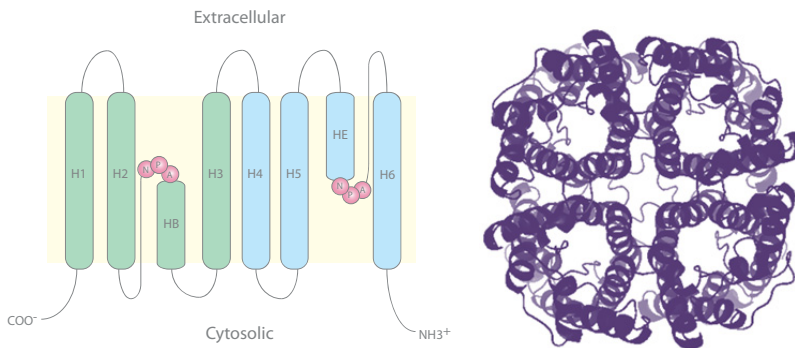


Figure 2.1. Left; The basic aquaporin topology with six transmembrane helices (H1-H6) and two membrane inserted half helices (HB and HE). The two tandem repeats are indicated in green and blue respectively and the conserved NPA motifs are illustrated as pink spheres. Right; Extracellular view of *HsAQP5* showing four monomers assembled as a tetramer.

2.2 Proton Exclusion and Substrate Specificity

Since the characterization of AQP1, which is a strict water transporter, several other solutes have been reported to pass through aquaporins. An alternative to the somewhat misleading designation “aquaporin” has hence been suggested, and the synonymous appellation “major intrinsic protein” (MIP) is often encountered in the literature.

The aqueous pore that penetrates the aquaporin monomer is lined with hydrophobic residues, with only a few water-binding sites. This construction, with a carefully balanced proportion between hydrophobicity and hydrogen bonding possibilities, facilitate the exceptionally high water permeability seen in e.g. AQP1, since the availability of water binding nodes reduces the energy barrier to water transport while the relatively low number of such sites minimizes water-pore interactions (Murata *et al.* 2000; Sui *et al.* 2001).

Close to the extracellular surface of the aquaporin monomer, a constriction region, known as the aromatic arginine (ar/R) filter is positioned. At this site, side chains of four amino acids, situated in helix 2, helix 5 and loop E, protrudes into the pore, creating its narrowest passage (Figure 2.2). The ar/R filter constitutes an important size exclusion barrier and amino acid diversity at this position, together with a pore diameter variation connected to the overall protein structure, render the aquaporins miscellaneous regarding their substrate specificities (Savage *et al.* 2010). In GlpF, which is highly permeable to glycerol but less to water, the diameter of the ar/R region is wider than in AQP1, allowing the passage of small and linear sugar molecules (Fu *et al.* 2000). This feature together with an increased hydrophobicity in the GlpF ar/R filter allows energetically favorable interactions between the glycerol and the protein (Hub and de Groot 2008). The importance of this site in size exclusion has been further established by introducing relevant point mutations. By substituting bulky residues such as a histidine (H180), a phenylalanine (F56) and an arginine (R195) with smaller residues, the minimum pore diameter of AQP1 was increased, rendering this natively highly water specific channel permeable to larger molecules, such as glycerol, urea and ammonia (Beitz *et al.* 2006).

In order to retain the electrochemical potential across the plasma membrane, a prevention of ion fluxes through the aquaporin channel is required. Hydrated ions are believed to be efficiently excluded due to the narrow diameter of the pore and the limited number of exposed carbonyl oxygen atoms that could potentially substitute for hydrating waters (Sui *et al.* 2001). However, allowing a rapid transport of water whilst, simultaneously, excluding protons from permeating the pore poses a particularly difficult challenge, since the positive charge of the proton can move along a column of water by hydrogen bond exchange, a phenomenon referred to as the Grotthuss mechanism (Agmon 1995; Pomes and Roux 1996).

There are some main features of the aqueous pore structure that confer proton exclusion, which occurs at the center of the pore where the NPA motifs, localized in the two half helices (HB and HE), are situated. Firstly, the dipole moment of these helices generates a net positive charge which causes repulsion of positively charged molecules. Secondly, the limited hydrogen bonding possibilities at this site forces the water molecules to reorient and the single file water column is broken, preventing the transfer of positive charge by hydrogen bond exchange (Chakrabarti *et al.* 2004). Moreover, the arginine residue of the ar/R filter has also been proven to be involved in proton exclusion, supposedly by adding to the electrostatic repulsion with its positive charge (Beitz *et al.* 2006).

The main structural features determining aquaporin selectivity and specificity are schematically visualized in Figure 2.2.

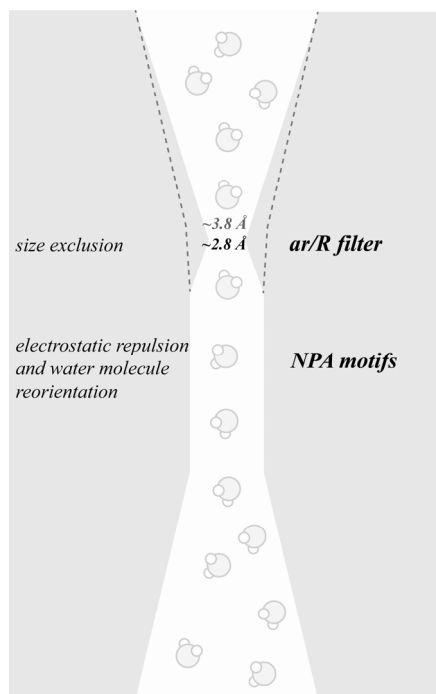


Figure 2.2. Schematic illustrating the main structural determinants of aquaporin substrate specificity. Approximately 20 Å from the extracellular side (top of figure), the channel narrows to ~2.8Å at the ar/R constriction region. This region is slightly wider (~3.8 Å) for the glycerol facilitating aquaporins as illustrated by the dashed line. Another 8Å down the pore, the NPA motifs are situated. At this site, the dipolar moment of helix HB and HE creates a net positive charge that, together with the water reorientation that occurs at this site, prevents the passage of protons.

2.3 Regulation

Aquaporins do not only generate an increased membrane permeability of water and other small solutes. They also offer a possibility for the organism to regulate the trans-membrane flux of these molecules. Three main modes of aquaporin regulation have been described in the literature and are presented below.

By regulating the promoters connected to the aquaporin genes, a complex expression pattern can be achieved, with organ- end even cell specific localization of particular isoforms. The site of expression is often reflected by the substrate specificity of the isoform. In the kidney for instance, where water reabsorption occurs during the concentration of the urine, efficient water transporters such as AQP1, AQP2 and AQP4 are expressed whereas glycerol transporters such as AQP7 and AQP9 are localized in adipose tissue and liver cells respectively, both involved in the glycerol metabolism (King *et al.* 2004). By transcriptional control, expression of aquaporins may also be up- or down regulated in response to stress, as seen when exposing plants to water deprivation (Alexandersson *et al.* 2005; Alexandersson *et al.* 2010) or rats to starvation (Carbrey *et al.* 2003).

Subcellular re-localization, or trafficking, of aquaporins offers a second regulation strategy, with the most well studied example probably being the vasopressin induced trafficking of AQP2 containing vesicles in cells lining the collecting duct of the kidney (Nedvetsky *et al.* 2009). Upon water loss, vasopressin hormone is released from the pituitary gland. In response to the elevated plasma levels of vasopressin, cyclic AMP levels are increased in the renal cells, which activate protein kinase A. The kinase phosphorylates AQP2 at a site in the C-terminus (Fushimi *et al.* 1997) which triggers the trafficking of the aquaporin to the apical plasma membrane. When the AQP2 density of the membrane is elevated, water reabsorption is increased, restoring the water balance. Aquaporin 5, a close homologue to AQP2 that is expressed in secretory glands, is also regulated in a similar fashion (Yang *et al.* 2003; Kosugi-Tanaka *et al.* 2006). In plants, a subcellular re-localization of aquaporins in root cells, as a response to salt stress, has been established and is believed to be connected to an observed decrease in water permeability of the roots at these conditions (Boursiac *et al.* 2005; Luu *et al.* 2012).

Direct stimuli of aquaporins, already positioned in the membrane can induce a third regulation mechanism by structural gating. Factors such as pH shifts, (Zeuthen and Klaerke 1999; Tournaire-Roux *et al.* 2003; Zelenina *et al.* 2003) phosphorylation (Johansson *et al.* 1998) and changes in osmolarity (Johansson *et al.* 1996) have been reported to influence the activity of certain aquaporin isoforms. Structural mechanisms responsible for aquaporin gating, including the capping of the spinach PIP2;1 channel by the D-loop (Tornroth-Horsefield *et al.*

2006) and the obstruction of the yeast AQY1 channel by its extended N-terminus (Fischer *et al.* 2009) have been presented.

2.4 Plant Aquaporins

Due to its inability of spatial movement the plant needs a complex and fine-tuned system for regulating its water balance. Hence it is not surprising that the aquaporin protein family is extensive in plants with 31 and 35 isoforms identified in maize and *Arabidopsis thaliana*, respectively (Chaumont *et al.* 2001; Johanson *et al.* 2001). Recent evidence of diversity in substrate specificity of plant aquaporins further explains the large number of isoforms reported. In higher plants, five distinct aquaporin subfamilies have been established based on amino acid sequence homology; the Plasma membrane Intrinsic Proteins (PIPs), the Tonoplast Intrinsic Proteins (TIPs), the Nod26-like Intrinsic Proteins (NIPs), the Small basic Intrinsic Proteins (SIPs) and the X-Intrinsic Proteins (XIPs), the latter hitherto only identified in dicots such as tomato and tobacco. The presence of yet two subfamilies, the GlpF-like Intrinsic Proteins (GIPs) and the Hybrid Intrinsic Proteins (HIPs) in more primitive plants such as mosses suggests that there has been a loss of aquaporin subfamilies during plant evolution (Gustavsson *et al.* 2005; Danielson and Johanson 2008). However the five remaining subfamilies seem to have expanded and diversified further and consequently it is sometimes difficult to assign true orthologs between species. The phylogeny of the plant aquaporin subfamilies is presented in Figure 2.3 and in the following sections the major characteristics of the aquaporin subgroups of higher plants will be described.

2.4.1 The Plasma membrane Intrinsic Proteins

The plasma membrane intrinsic proteins are, as the name implies, predominantly localized in the plasma membrane of plant cells, where they can constitute up to 20% of the total membrane protein fraction (Johansson *et al.* 1996). The PIPs can be further subdivided into two groups, PIP1s and PIP2s, which in their primary structure differ mainly in the terminal regions (Kammerloher *et al.* 1994; Chaumont *et al.* 2001). The members of the PIP2 group typically demonstrate high water permeability when reconstituted into liposomes or expressed in oocytes (Johansson *et al.* 1998; Schaffner 1998; Chaumont *et al.* 2000; Karlsson *et al.* 2003), while PIP1 isoforms generally show significantly lower or even nonexistent water transporting activity (Kammerloher *et al.* 1994; Biela *et al.* 1999; Chaumont *et al.* 2000).

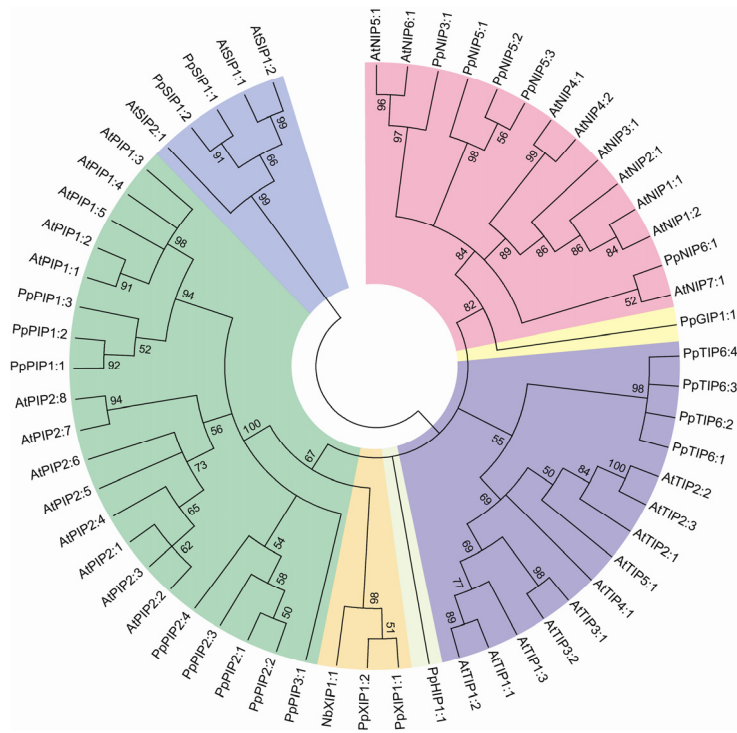


Figure 2.3. Bootstrap consensus tree obtained by the maximum likelihood method showing the division of *A. thaliana* (At) and *Physcomitrella patens* (Pp) aquaporins into seven subfamilies. The *Nicotiana benthamiana* (Nb) XIP1;1 was also included in the analysis. Percentage bootstrap values for the nodes are included and nodes with less than 50% support are collapsed. (The tree was kindly provided by Hanna Anderberg).

The PIPs are known to be regulated at several levels. Besides a transcriptional regulation, their density at the plasma membrane can be regulated by phosphorylation dependent trafficking events in response to stress (Boursiac *et al.* 2005; Boursiac *et al.* 2008; Luu *et al.* 2012). Moreover PIPs are subject to structural gating in a pH dependent manner. In response to a drop in pH, protonation of a conserved histidine in the D-loop of PIP isoforms allows the interaction between this residue and an acidic residue of the N-terminus, leading to the capping of the pore by the D-loop. This closed conformation could also be stabilized by the binding of divalent cations in close vicinity to the acidic residue at the N-terminus. On the other hand phosphorylation events, involving residues in the B-loop, would induce conformational changes unlocking the D-loop from its capping position, destabilizing the closed state and favoring the open conformation (Tornroth-Horsefield *et al.* 2006). The structural features of the PIP gating mechanism is illustrated in Figure 2.4. Yet another regulatory mechanism has been

recognized and it involves the interaction between PIP monomers of different isoforms. While *ZmPIP1;2* did not increase water permeability when expressed in *Xenopus laevis* oocytes (section 3.3.1), co-expression with *ZmPIP2* isoforms lead to variations in water permeability in a *ZmPIP1;2* dose dependent manner. Subsequent fluorescent protein fusion experiments have shown that co-expression increases the density of PIPs at the plasma membrane of both oocytes and plant protoplasts. This finding suggests that PIP2 isoforms assist PIP1 trafficking, possibly by contributing with a ticket for export out from the ER, where PIP1 isoforms are otherwise retained (Fetter *et al.* 2004; Zelazny *et al.* 2007; Zelazny *et al.* 2009). Results from another study, where artificial heterotetramers with different ratios of *NtAQP1* (a PIP1 homolog) and *NtPIP2;1* were expressed in yeast cells, suggest a shift in substrate specificity of *NtAQP1* upon heteromerization.

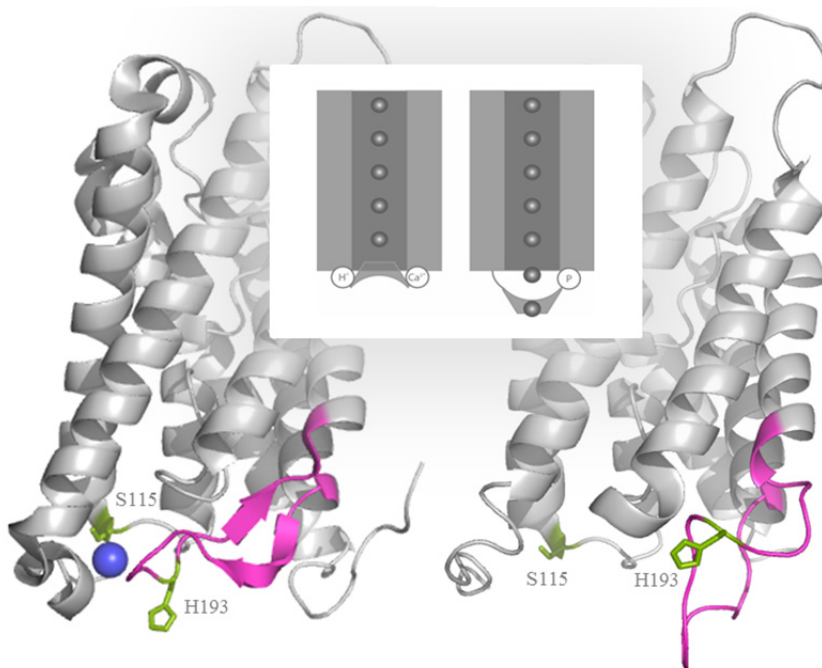


Figure 2.4. The SoPIP2;1 channel in its closed (left) and open (right) conformation. In the closed conformation, the D-loop (pink) is anchored to the N-terminal helix by a hydrogen bonding network involving the Cd^{2+} ion (blue) that is supposedly replaced by a Ca^{2+} ion *in vivo*. In the open conformation, the D-loop is released from the N-terminal region and is shifted towards the cytosol which uncaps the pore. The transition between the different states depends on protonation of H193 (indicated) and phosphorylation of S115 (indicated) and S274 and is also believed to be influenced by the Ca^{2+} concentration, as summarized in the inset schematic.

The *NtAQP1* homotetramer showed no water permeability in yeast protoplasts, while a CO₂-induced intracellular acidification was recorded. However at a 1/1 *NtAQP1/NtPIP2* ratio, the water transport rate was indistinguishable from that of the *NtPIP2*;1 homotetramer, suggesting a conformational change upon heteromerization transforming the *NtAQP1* pore into a water transporter (Otto *et al.* 2010).

Besides water, other PIP substrates have been suggested. PIP2 isoforms have shown to transport hydrogen peroxide, when recombinantly expressed in yeast cells (Dynowski *et al.* 2008; Hooijmaijers *et al.* 2012), a feature that is interesting since, besides being a toxic metabolite, this reactive oxygen species also has a suggested role as messenger molecule. As mentioned above, another suggested PIP substrate is carbon dioxide which has proven to permeate through PIP1 channels (Uehlein *et al.* 2003; Uehlein *et al.* 2008; Otto *et al.* 2010; Uehlein *et al.* 2011). The identification of *NtAQP1*, a tobacco PIP1 homolog, in the inner envelope of the chloroplast, the organelle in which CO₂ is transformed into biomolecules, further supports the hypothesis of a PIP1 mediated CO₂ transport. However, the biological role of such an assisted transport has been debated since the lipid bilayer does not constitute a significant barrier to CO₂ flux (Hub and de Groot 2006). In **Paper V**, both PIP1 and PIP2 isoforms were identified in thylakoid membranes of *A. thaliana* while only the PIP2;1 isoform, which is a proven good water transporter, could be identified in the inner envelope. Since CO₂ is consumed in the stroma, while water is split in the thylakoid lumen during the photosynthetic reactions, we propose a role of aquaporins as water, rather than gas, transporters in chloroplasts. Proposed biological functions of PIP water channels in chloroplasts are summarized in Figure 2.5.

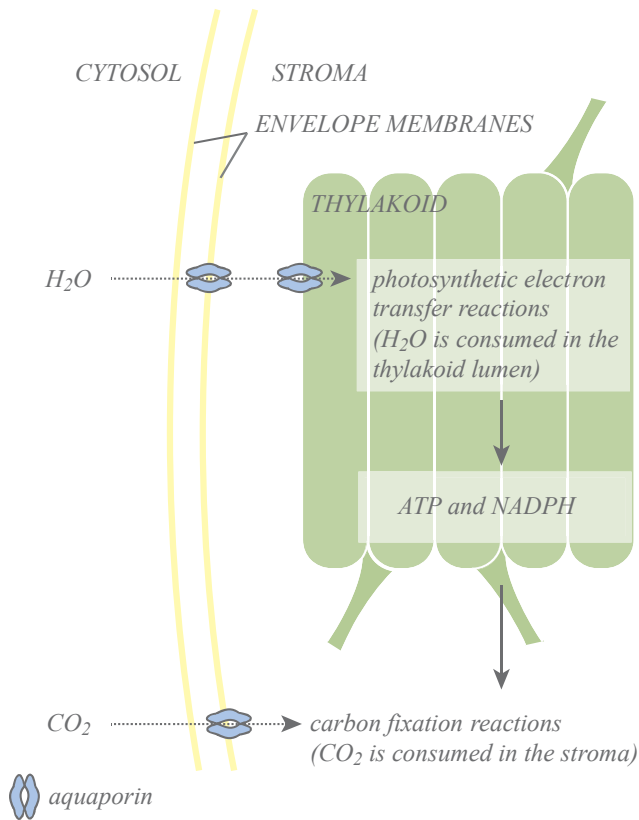


Figure 2.5. Proposed biological functions of PIP aquaporins in chloroplasts. In **Paper V**, *AtPIP1* and *AtPIP2* isoforms were identified in thylakoid membrane fractions while only the efficiently water transporting *AtPIP2;1* channel could be detected in the inner envelope. Since water is consumed in the thylakoid lumen, we propose a biological function of chloroplast localized *AtPIP* channels as water transporters. In other studies, the CO_2 transporting capacity of PIP1 channels has been established and since the *NtAQP1* was identified in the inner envelope of tobacco chloroplasts, it has been proposed to have a role in the CO_2 import into the stroma of the chloroplast (Uehlein *et al* 2003; Uehlein *et al* 2008; Uehlein *et al* 2011).

2.4.2 The Tonoplast Intrinsic Proteins

The vacuolar membrane, also referred to as the tonoplast, is highly permeable to water which is explained by the large number of Tonoplast Intrinsic Protein (TIP) aquaporins being expressed at this site (Maurel *et al.* 1997; Higuchi *et al.* 1998; Karlsson *et al.* 2000). Besides water, other TIP substrates have been identified, among which nitrogen containing compounds such as urea and ammonia implicate that some TIP isoforms may have a role in the nitrogen metabolism of plants (Liu

et al. 2003; Jahn *et al.* 2004; Loque *et al.* 2005). The finding that the transcript levels of several TIPs were upregulated in root tissues in response to nitrogen deprivation supports this hypothesis (Liu *et al.* 2003). Hydrogen peroxide, which has also been shown to be transported by PIP aquaporin isoforms (see above), is another potential TIP substrate that has been shown to permeate through *Arabidopsis* TIP1 isoforms (Bienert *et al.* 2007). The biological importance of TIP aquaporins remains elusive since a number of knockout studies have not generated any obvious phenotypes (Schussler *et al.* 2008; Gattolin *et al.* 2010). However since the TIPs constitute a subfamily with many members that frequently have similar localization patterns, the lack of effects from knockout experiments may be explained by redundancy (Gattolin *et al.* 2010). Interestingly, transformed tomato plants with constitutive *STIP2;2* expression rendered increased biomass and fruit yield as compared with control plants, both under normal and drought-stress conditions, suggesting an impact of this isoform on abiotic stress tolerance (Sade *et al.* 2009).

Even though the tonoplast is the main localization of TIP aquaporins, some isoforms have been suggested to be targeted to other sites such as the plasma membrane (Alexandersson *et al.* 2004; Gattolin *et al.* 2011) and the chloroplast inner envelope and thylakoid membranes of *A. thaliana* seedlings (**Paper V**; Ferro *et al.* 2010). The role of the TIPs at these alternative locations remains to be elucidated.

2.4.3 The Small basic Intrinsic Proteins

Possessing the main aquaporin primary structure features but at the same time being smaller than the PIPs and the NIPs and having an overall more basic amino acid composition than the TIPs, the members of the SIP subfamily were distinguished as a new aquaporin subfamily (Johanson *et al.* 2001; Johanson and Gustavsson 2002). Little is known about the function of the SIP proteins due to their presumably low native expression levels which is reflected by the scarce number of expressed sequence tags (ESTs) reported (Johanson and Gustavsson 2002). Moreover, the SIPs have proven to be difficult to produce in heterologous expression systems for *in vitro* characterization (**Paper I**, and Dr. Maria Agemark, unpublished data). When analyzing transient expression of the three *A. thaliana* SIP isoforms, linked to green fluorescent protein (GFP) it was shown that they were all localized to the ER, which was further confirmed by immunoblotting analyses of membrane fractions (Ishikawa *et al.* 2005; Maeshima and Ishikawa 2008). If the SIPs harbor an ER retention signal they may be localized to this organelle also in heterologous expression systems, offering an explanation to the low recombinant expression levels observed in *P. pastoris* (**Paper I**). The amino acid composition of the SIPs differs from other aquaporins at several conserved positions, suggesting a deviating substrate specificity of these channels (Johanson

and Gustavsson 2002). Hitherto *AtSIP1;1* and *AtSIP1;2* have shown to be water permeable when expressed in yeast whereas no activity could be observed for *AtSIP2;1* (Ishikawa *et al.* 2005). Further biochemical characterization of the SIPs is needed to elucidate the role of this aquaporin subfamily.

2.4.4 The Nod26-like Intrinsic Proteins

The first member of the NIP subgroup was originally identified in the peribacteroid membrane, a lipid bilayer of plant origin, which envelops nitrogen fixing bacteria in the root nodules of soybean (Sandal and Marcker 1988; Weaver *et al.* 1991). Since then, NIPs have been identified at other subcellular sites such as the plasma membrane, which seems to be its main location (Ma *et al.* 2006; Tanaka *et al.* 2008), and the endoplasmic reticulum (Mizutani *et al.* 2006).

Based on the amino acid composition at the ar/R region the *A. thaliana* NIP isoforms can be divided into two groups, with three of the NIPs (NIP5;1, NIP6;1 and NIP7;1), referred to as group II, having a supposedly wider diameter at this site as compared with the other six NIP isoforms, referred to as group I which suggests the accommodation of larger substrates. Moreover, NIP5;1 and NIP6;1 have an amino acid substitution in the second NPA box, where the alanine is replaced by the slightly larger and more hydrophobic valine, creating a NPV motif suggesting a decreased water permeability (Wallace and Roberts 2004). *In vivo* functional assessments in oocytes have confirmed the predicted nonconformity in substrate specificities between group I and group II NIPs. The *A. thaliana* NIP6;1 has been shown to transport larger substrates, such as urea, in addition to glycerol. Glycerol was also efficiently transported by soybean noduline 26 that served as a model protein representing group I NIPs. Furthermore no *AtNIP6;1* assisted water transport was observed while noduline 26 showed readily detectable water transporting activity (Wallace and Roberts 2005). In rice and several other plant species a third functional subgroup, NIP III, has been recognized with an ar/R filter that differs significantly from that of the NIP I and NIP II aquaporins (Danielson and Johanson 2010). Members of the NIP III group have been reported to be involved in the uptake of silicon (Ma *et al.* 2006; Chiba *et al.* 2009; Mitani *et al.* 2009).

A NIP mediated transport of boron which, just like silicon, serves as a plant nutrient, has also been reported (Takano *et al.* 2006; Tanaka *et al.* 2008). Plants with mutated silicon or boron transporting NIP channels show growth impaired phenotypes (Ma *et al.* 2006; Takano *et al.* 2006) which suggests a critical role for NIPs in plant nutrient uptake. Recent studies have also shown that several NIP isoforms are permeable to arsenite, a toxic mineral that is naturally occurring in many agricultural areas and that accumulates in plants and impairs the growth of crops (Bienert *et al.* 2008; Isayenkova and Maathuis 2008; Kamiya *et al.* 2009).

2.4.5 The X-Intrinsic Proteins

The XIP aquaporin subfamily was initially identified in the moss *Physcomitrella patens* and subsequent database searches identified XIP sequences in other organisms such as *Populus trichocarpa*, *Citrus clementina* and *Nicotiana benthamiana* (Danielson and Johanson 2008). Lately, isoforms belonging to this subgroup have also been identified in other dicots such as tobacco, potato, tomato and cotton (Park *et al.* 2010; Bienert *et al.* 2011). The X designation refers to this subfamily being uncharacterized. However in a recent extensive study, this channel was shown to be expressed in all organs of tobacco plants with the subcellular localization being at the plasma membrane. Furthermore, the substrate specificity was investigated and the different XIP isoforms studied all facilitated the transport of glycerol, urea and boric acid while no water transport could be detected (Bienert *et al.* 2011). Still, more data is needed in order to elucidate the role of this aquaporin subfamily and since a protocol for efficient overexpression and purification of *N. benthamiana* XIP1;1 has been developed (**Paper II**) further structural and functional characterization is within reach.

2.5 Aquaporins in Man

In mammals, 13 aquaporin isoforms (AQP0-AQP12) have been identified and they can, on phylogenetic basis, be subdivided into four distinct subfamilies. Since the amino acid composition, upon which the phylogenetic analyses are built, affects structural architecture, the members of a certain subgroup commonly share biochemical features such as substrate specificity. The largest subgroup consisting of AQP0-2 and AQP4-6 is referred to as “orthodox aquaporins” due to their ability to transport water exclusively. The “aquaglyceroporins” have a capacity of transporting glycerol and other substrates in addition to water and this subgroup comprises AQP3, AQP7 and AQP9-10 (Borgnia *et al.* 1999; King *et al.* 2004). A third subgroup is formed by the relatively recently discovered AQP11 and AQP12 which are labeled “subcellular aquaporins” due to their proposed intracellular localization (Ishibashi *et al.* 2009). Aquaporin 8 does not readily group together with any of the other mammalian aquaporins, but rather this isoform constitute a group on its own (**Paper III**; Borgnia *et al.* 1999). The phylogeny of the 13 mammalian aquaporin isoforms is presented in Figure 2.6 and below some of the major features of the mammalian aquaporin isoforms are summarized.

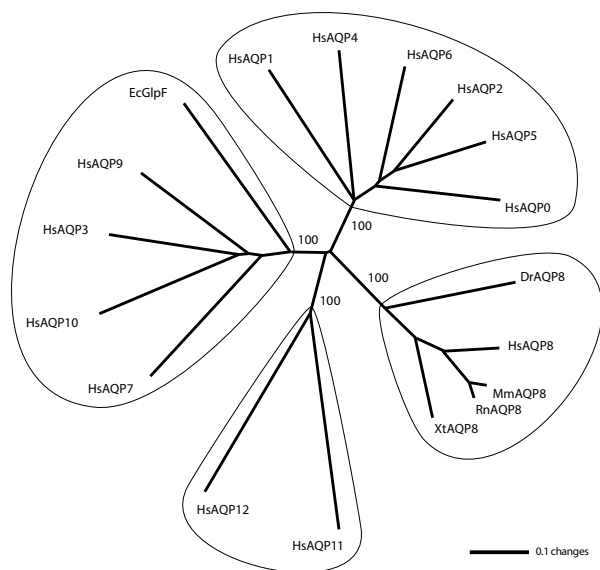


Figure 2.6. Phylogenetic tree showing the division of human (Hs) aquaporin isoforms into four subfamilies. The *EcGlpF* glycerol facilitator was included as a reference and AQP8 isoforms from other animal species (western clawed frog (Xt), rat (Rn), mouse (Mm) and zebrafish (Dr)) were incorporated in the analysis to emphasize the distinction of this subgroup from the others. Bootstrap support for the four subfamilies are indicated. (Modified from **Paper III** with permission from the publisher).

2.5.1 The Orthodox Aquaporins

The strict water transporters belonging to the aquaporin subfamily, has a remarkable capacity of facilitating a rapid trans-membrane flux of water while excluding other molecules (see section 2.2). Being the first aquaporin to be molecularly characterized, AQP1 is probably the mammalian aquaporin isoform that has been most intensively scrutinized. With a well-documented tissue localization pattern, a known structure and characterized biochemical properties, the biological function of AQP1 is to a large extent elucidated (King *et al.* 2004). Besides predictable features such as an importance for the concentration of urine (Ma *et al.* 1998) and for maintenance of the water homeostasis in the eye (Zhang *et al.* 2002), unexpected functions of AQP1, such as an involvement in cell migration and tumor angiogenesis have also emerged (Saadoun *et al.* 2005). In a few reports based on molecular dynamics simulations and *in vitro* experiments, an AQP1 mediated transport of gasses, such as carbon dioxide, with a possible route through the central pore (Prasad *et al.* 1998; Wang *et al.* 2007) was proposed. However, CO₂ permeate most membranes rapidly without assisted transport mechanisms (Hub and de Groot 2006), and no significant reduction in CO₂ permeability was observed in erythrocytes, kidney or lung of AQP1-null mice,

indicating a lack of physiological relevance of these findings (Yang *et al.* 2000; Fang *et al.* 2002).

Another, extensively studied member of the aquaporin subfamily is AQP2, which is predominantly expressed in the collecting duct of the kidney, where it has a crucial function in regulating the water reabsorption. The density of AQP2 in the plasma membrane can be altered in response to disturbed water balance (see section 2.3). Malfunctioning AQP2 trafficking leads to a rare condition; nephrogenic diabetes insipidus, where the ability to concentrate the urine is lost, leading to extensive water loss and severe dehydration (Deen *et al.* 1994; Mulders *et al.* 1998). A close homologue of AQP2 is AQP5, which is localized in secretory glands such as salivary and sweat glands as well as in the respiratory system and the cornea of the eye (Raina *et al.* 1995; Nielsen *et al.* 1997; Hamann *et al.* 1998). Aquaporin 5 is an efficient water transporter as determined by *in vivo* experiments in *X. laevis* oocytes and *in vitro* assessments in proteoliposomes, and its activity was proven to be efficiently and reversibly inhibited by mercury (Raina *et al.* 1995; **Paper IV**). The high resolution structure of *HsAQP5* has been solved and a remarkable feature that had not been observed in previous structures of mammalian AQPs was the presence of a lipid in the central pore. The role of this lipid remains elusive but since it completely occludes the space between the monomers, an alternative transport route through the central pore of AQP5, as suggested for AQP1, is unlikely (**Paper IV**). In resemblance with AQP2, there is a regulation of AQP5 by trafficking, where AQP5 moves from intracellular storages and fuse with the plasma membrane in a phosphorylation dependent manner (Yang *et al.* 2003; Ishikawa *et al.* 2005; Kosugi-Tanaka *et al.* 2006). Impaired AQP5 trafficking has been associated with the autoimmune disorder “Sjögrens syndrome” with clinical manifestations such as dry eyes and dry mouth (Tsubota *et al.* 2001; Fox 2005).

Aquaporin 0 is a bit of an outsider in the aquaporin subgroup, partially due to its relatively poor water transporting capacity but also because of its specific localization in the lens of the eye exclusively (King *et al.* 2004). Another unusual feature of AQP0, is its function in cell-cell adhesion (Engel *et al.* 2008). Aquaporin 0 related junction formation was proven to be due to interactions between the extracellular loops of tetramers situated in adjacent membranes. This, in turn, was shown to be caused by conformational changes at the extracellular protein surface, induced by proteolytic cleavage of the intracellular N- and C-termini (Gonen *et al.* 2004; Gonen *et al.* 2005; Engel *et al.* 2008). As AQP0 transforms into an adhesion molecule, its water transporting capacity is lost due to closure of the pore (Gonen *et al.* 2004). The dual function of AQP0 as a water channel and adhesion molecule is critical for maintaining lens transparency and missense mutations in the *aqp0* gene causes development of cataract (Shiels and Bassnett 1996; Berry *et al.* 2000). Another member of the aquaporin subgroup with unusual properties is AQP6 which resides in intracellular vesicles of the acid

secreting intercalated cells in the collecting duct of the kidney. Since AQP6 is able to facilitate the transport of anions in addition to water and since it is subject to pH regulation it has a presumed role in acid-base homeostasis at its site of expression (Yasui *et al.* 1999; Ikeda *et al.* 2002).

Aquaporin 4 is expressed in a number of tissues including skeletal muscle, kidney, lung and brain (King *et al.* 2004). Besides having a role in maintaining water homeostasis it has a documented function in neural signaling and cell migration (Verkman 2011). Moreover, like AQP0, it is involved in cell adhesion, but AQP4 junctions are rather formed by a certain splice version of AQP4 than resulting from modifications due to proteolytic effects. Unlike the AQP0 tetramers that are exactly stacked on top of each other in the junction, the AQP4 tetramers shows a shifted interaction pattern upon crystallization, with the tetramer in one layer of the crystal interacting with four in the next layer (Engel *et al.* 2008). Recently, a substantial amount of data has been published dealing with the involvement of AQP4 in the neuroinflammatory demyelinating disease neuromyelitis optica, which is a rare condition related to multiple sclerosis. In this disorder, binding of auto antibodies to AQP4 causes initiation of inflammatory responses in the vicinity of the glial cells where AQP4 is expressed and subsequent degradation of the myelin (Lennon *et al.* 2005; Papadopoulos and Verkman 2012).

2.5.2 The Aquaglyceroporins

The four members belonging to the aquaglyceroporin subfamily (AQP3, AQP7, and AQP9-10) are more promiscuous regarding their substrate specificities as compared with the members of the aquaporin subgroup. As the name implies, isoforms belonging to this subgroup transport glycerol in addition to water and urea has also proven to be a substrate of these channels (Ishibashi *et al.* 1994; Ishibashi *et al.* 1997; Ishibashi *et al.* 1998; Tsukaguchi *et al.* 1998; Ishibashi *et al.* 2002). Aquaglyceroporin 3, 7 and 9 shows a wide tissue distribution and have been suggested to be involved in numerous biological processes such as the lubrication of the skin (Hara and Verkman 2003), the release of glycerol from adipose tissue (Kondo *et al.* 2002), and the subsequent import of this carbohydrate into liver cells where it is believed to support gluconeogenesis (Carbrey *et al.* 2003; Jelen *et al.* 2011). Aquaporin 10 is more specifically expressed and has hitherto only been detected in small intestine (Ishibashi *et al.* 2002) with suggested additional weak expression in colon and stomach. It has, together with AQP3, 7 and 11 been suggested to have an involvement in active celiac disease (Laforenza 2012). Another interesting feature of the aquaglyceroporins is their ability to transport metalloids such as arsenate, as reported for AQP7 and AQP9. Uptake of arsenate through AQP9 in leucocytes was suggested to account for the therapeutic effects of this metalloid on certain kinds of leukemia (Liu *et al.* 2002).

2.5.3 The "Subcellular Aquaporins"

Aquaporin 11 and 12, also referred to as the "subcellular aquaporins", were the two last members of the aquaporin protein family to be discovered (Morishita *et al.* 2004; Itoh *et al.* 2005). Since the human genome is now sequenced, it is established that no more aquaporin isoforms are present in man. These two aquaporin isoforms are believed to be expressed in intracellular structures, such as the ER (Morishita *et al.* 2005). They show low sequence identity with the other mammalian aquaporins and one of the otherwise highly conserved NPA-motif differs in these two isoforms (Ishibashi 2006). Despite this difference, AQP11 proved to be a water facilitator when reconstituted into proteoliposomes (Yakata *et al.* 2006). Little is known about the biological function of the "subcellular aquaporins" but AQP11-null mice developed a lethal phenotype with polycystic kidneys (Morishita *et al.* 2005).

2.5.4 Aquaporin 8

Due to low sequence similarity with the other mammalian aquaporins, AQP8 forms a phylogenetic clade on its own (figure 2.6). Rather, this isoform seems to be more closely related to a subgroup of plant aquaporins; the TIPs (Ishibashi *et al.* 1997; Koyama *et al.* 1998). The ability of AQP8 to transport unusual substrates such as ammonia (Saparov *et al.* 2007), ammonium analogues (Liu *et al.* 2006) and hydrogen peroxide (Bienert *et al.* 2007), besides water, further distinguishes AQP8 from the other mammalian aquaporins and unites it with the TIPs (section 2.4.2).

Aquaporin 8 was initially cloned from rat testis (Ishibashi *et al.* 1997), pancreas and liver (Koyama *et al.* 1997) and occurrence of AQP8 messenger ribonucleic acid (mRNA) in rodent heart, kidney, salivary gland and gastrointestinal tract has also been reported (Yang *et al.* 2005). In man, AQP8 mRNA was initially only detected in pancreas and colon (Koyama *et al.* 1998) but its expression has later been identified in other tissues such as placenta (Liu *et al.* 2004) and testis (Yeung *et al.* 2010). The subcellular localization of AQP8 has been debated. Initially AQP8 was, by immunolocalization assessments, identified in endoplasmatic reticulum, subapical vesicles and mitochondria of mouse hepatocytes. (Ferri *et al.* 2003). Subsequently, AQP8 was shown to be localized in the inner membrane of mitochondria from rat liver (Calamita *et al.* 2005). These results were contradicted by a later study, where no AQP8 could be immunodetected in mitochondrial membrane fractions from mouse and where no differences in water permeability could be observed when comparing mitochondria from wild-type *versus* AQP8 knock-out mice (Yang *et al.* 2006). In a recent study, *RnAQP8* that was recombinantly expressed in yeast, proved to be enriched in the mitochondrial fraction. The heterologously expressed protein was shown to increase

mitochondrial formamide (an ammonia analogue) transport by threefold as compared with mitochondria from untransformed yeast cells, while only a limited effect on water permeability was observed. Hence a role of AQP8 as an ammonia transporter rather than a water channel in mitochondria was suggested (Soria *et al.* 2010).

The biological relevance of AQP8 remains elusive since no, or very mild phenotypes were observed when the *aqp8* gene was knocked out in mice (Yang *et al.* 2005). To increase the knowledge about HsAQP8 structure and function, we overexpressed and purified this channel protein for subsequent functional and structural analyses (**paper III**).

Methods

3.1 Recombinant Protein Expression

The proteins that are integrated in or associated with the membranes of the cell contribute to the unique characteristics of different membranes and are often involved in vital biological processes. Moreover, being situated at the surface of the cell, available for modification by molecules in the extracellular fluid, membrane proteins constitute important drug targets. Hence, in order to understand essential biochemical reactions and to gain important information to alleviate drug design, functional and structural experiments aimed at characterizing membrane proteins are required. Such trials depend on the availability of large quantities of protein and since membrane proteins are commonly present at low levels in their native tissues, their characterization has been belated. This is reflected in the relatively small number of unique high resolution membrane protein structures deposited in the protein data bank. In the last decades, however, recombinant expression systems for large scale protein production have been developed leading to a surge in structural information and today more than $\frac{3}{4}$ of the reported membrane protein structures derive from protein produced in heterologous expression hosts (Figure 3.1). In the following sections, some common expression hosts will be described with particular focus on *P. pastoris*, a yeast strain which in our lab has been used for the production of several aquaporin isoforms (**Paper I-IV**).

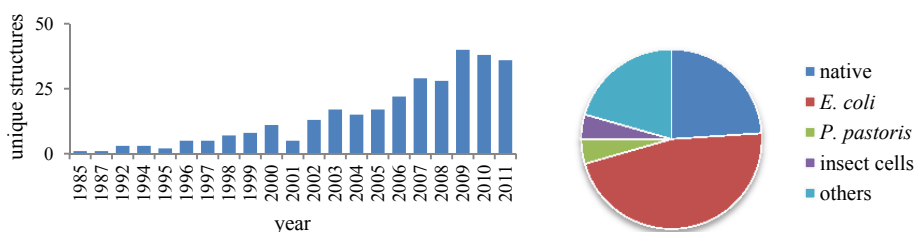


Figure 3.1. Diagrams showing the yearly number of unique membrane protein structures deposited in the protein data bank during the last two decades (left) and the protein source distribution of all membrane protein structures deposited (right). Data was retrieved from http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html and the Membrane Protein Data Bank (Raman *et al.* 2006).

3.1.1 Common Expression Systems

The statistics presented in Figure 3.1 clearly shows that the prokaryote *E. coli* has constituted the dominating recombinant expression system over the years since almost half of the reported membrane protein structures deposited in the PDB derive from proteins produced in this microbe. With a wide range of strains and vectors available, this system offers a possibility to fine tune recombinant protein expression. Moreover sophisticated methods for monitoring protein expression and purification has been developed as exemplified by the GFP fusion system where a fluorescent protein connected to the terminus of the recombinant membrane protein aids in the assessment not only of the yields but also of the folding status of the overexpressed membrane proteins (Drew *et al.* 2005; Drew *et al.* 2006).

The number of prokaryotic membrane proteins purified from recombinant *E. coli* is significant and a number of eukaryotic membrane proteins have also been successfully produced in this host (Freigassner *et al.* 2009). However, since the prokaryotic membrane protein production machinery differs significantly from the eukaryotic one with e.g. a much higher polypeptide chain elongation rate, a deviating translocon complex and an inability to perform post translational modifications, obtaining over expression of eukaryotic membrane proteins in *E. coli* is often not trivial (Junge *et al.* 2008; Freigassner *et al.* 2009).

As an alternative to the prokaryotic expression systems, recombinant expression in insect cells has emerged. By ligating the gene of interest into a viral vector and letting the viruses infect the insect cells, transformation is achieved. In this system, post translational modifications can occur and it has become particularly attractive for G-Protein Coupled Receptor (GPCR) production due to its low endogenous GPCR levels, rendering a low background signal during ligand binding assays. The major drawback with this expression system is the potentially time-consuming vector construction which may take several months, and the complex media required for culturing (Junge *et al.* 2008).

3.1.2 The Methylophilic Yeast *Pichia pastoris*

Possessing the eukaryotic protein production machinery in combination with the ability to be easily manipulated and cultured, different yeast species have become increasingly popular as protein production hosts. One of these is *P. pastoris*, which is a yeast strain with methylophilic properties, meaning that it can utilize methanol as a sole carbon source. Its ability to grow to tremendously high cell densities at aerobic conditions without the accumulation of toxic side products distinguishes *P. pastoris* from *Saccharomyces cerevisiae* which is another yeast strain commonly used for recombinant protein production. Another feature that is characteristic of *P. pastoris* is the presence of the methanol inducible *AOX1*

promoter that regulates the expression of an alcohol oxidase enzyme, which is involved in the methanol catabolism. This promoter is very potent and during growth on methanol, approximately 5% of the total mRNAs derive from the *AOX1* gene (Invitrogen 2005). A replica of this strong endogenous promoter can be used to control the transcription of a recombinant gene.

Specialized *P. pastoris* strains and expression vectors have been developed and by combining these, optimal conditions for secreted and intracellular expression of soluble proteins can be achieved (Daly and Hearn 2005; Invitrogen 2005; Cregg *et al.* 2009). Moreover, this expression system has also appeared to be invaluable in membrane protein production and a number of eukaryotic membrane proteins with different topologies and biochemical characteristics such as enzymes (Martinez Molina *et al.* 2007), transporters (Doring *et al.* 1997; Cai and Gros 2003; Aller *et al.* 2009) and GPCRs (Talmont *et al.* 1996; Sarramegna *et al.* 2003; Hori *et al.* 2010) have been produced in *P. pastoris*. The first aquaporin to be successfully overexpressed in *P. pastoris* was the spinach aquaporin *SoPIP2;1* which, in our lab, was purified in mg amounts enabling the functional characterization of this protein after reconstitution into artificial liposomes (Karlsson *et al.* 2003). Subsequently, the purified protein was crystallized and the high resolution *SoPIP2;1* structure was achieved (Tornroth-Horsefield *et al.* 2006). Since then, a number of aquaporins have been successfully overexpressed in *P. pastoris*, both in our lab (**Paper I-IV**) and elsewhere (Nyblom *et al.* 2007; Hedfalk *et al.* 2008; Fischer *et al.* 2009).

3.1.3 Optimizing the Recombinant Expression

Recent research has revealed a number of bottlenecks that may limit heterologous protein production and in order to fully exploit the capacity of the *P. pastoris* expression system, optimization of several parameters may be required. This section covers some of the most prominent hurdles that can influence recombinant membrane protein expression and that may be avoided by optimization.

3.1.3.1 Nucleic Acid Composition

Let us start with the transcription of the deoxyribonucleic acid (DNA) sequence which, for optimal expression, should contain the *P. pastoris* consensus start sequence A/YAA/TAATGTCT and have a balanced A/T/G/C content. A/T rich nucleotide stretches may induce truncation of the transcripts (Scorer *et al.* 1993; Sreekrishna *et al.* 1997) since they resemble potential yeast polyadenylation sites and should be avoided.

The ribonucleic acid (RNA) sequence which is produced during transcription is composed of four different nucleotide species combined to form coding triplets or codons. Since there are more possible tri-nucleotide combinations than there are

amino acids, some amino acid residues are encoded by several synonymous codons. Different organisms tend to utilize synonymous codons to unequal extents and infrequently occurring RNA triplets in the transcriptome of a particular organism are referred to as rare codons. The overall codon distribution of a certain genome is also reflected by the concentrations of the corresponding transfer RNAs (tRNAs). Since rare codons have been found to sometimes gather in clusters in the genomic sequence they have been speculated to have a function in the regulation of the translational speed, which could depend on ribosomal spacing, and in co-translational protein folding (Norholm *et al.* 2012).

When a protein is to be overexpressed in a foreign organism there might be a mismatch in the codon composition of the recombinant gene and the tRNA profile of the host cell, causing impaired translation efficiency. By modifying the DNA sequence by silent mutations, a codon profile of the mRNA that correlates better with the tRNA set up of the expression host can be obtained. Taken together, the factors described above may dramatically improve expression, as exemplified by the successful overexpression of an aquaporin from the malaria causing microbe *Plasmodium falciparum* in *P. pastoris* after optimization of both A/T content and codon composition (Hedfalk *et al.* 2008).

3.1.3.2 Recombinant Gene Dosage

Recombinant expression vectors can be integrated into the *P. pastoris* genome by homologous recombination, creating stable transformants. Upon gene insertion, the homologous recombination site is reconstructed which enables repeated recombination events to occur, resulting in clones with multiple integrated gene sequences, also referred to as “jack pot clones” (Figure 3.2). In several instances, expression of soluble proteins has shown to be dramatically increased by elevating the recombinant gene dosage (Romanos *et al.* 1998). In **Paper I** we show that also the expression of aquaporins in *P. pastoris* is strongly dependent on the number of inserted gene sequences. Since multi-copy integration is a rare event, selection methods based on acquired antibiotic tolerance levels are utilized to identify “jack pot clones”. The *Sh ble* gene which encodes a protein that confers resistance towards the antibiotic zeocin is incorporated in the expression vector and is recombined into the *P. pastoris* genome together with the gene of interest. Increasing the number of *Sh ble* transcripts induces elevated antibiotic tolerance which thereby indirectly reflects the copy number of the gene of interest.

In order to directly assess the recombinant gene dosage, expensive and time consuming techniques depending on radioactively labeled probes, such as quantitative dot blotting or Southern hybridization have previously been used (Invitrogen 2005). The accurateness of these methods can be questioned since the signal of the dot blot may become saturated and since different probes may bind to the extracted DNA with unequal efficiencies. We developed a more convenient method based on qPCR to quantify inserted gene sequences. The primers used in

the reaction are generic since they are directed towards a sequence in the vector and hence an accurate comparison of gene dosage between clones carrying different recombinant genes can be made. By using this method we were able to compare expression levels of clones harboring varying numbers of the same recombinant gene and of clones with equal copy numbers of varying gene constructs (**Paper I**).

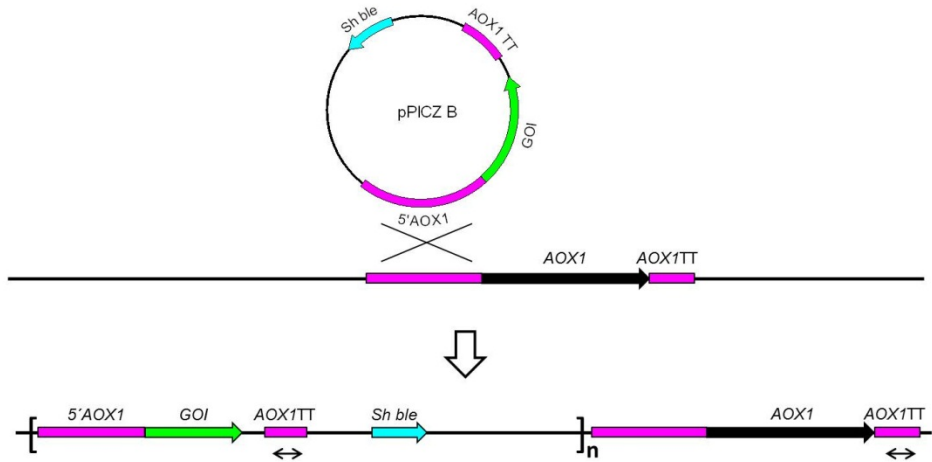


Figure 3.2. Schematic showing the integration of the pPICZB expression vector, carrying the antibiotic resistance *Sh ble* gene and the gene of interest (GOI), into the *P. pastoris* genome at the *AOX1* locus. Upon integration the 5'*AOX1* recombination site is recreated enabling integration of several (n) recombinant gene sequences. The double headed arrows indicate the sequence in the 3'TT region of the *AOX1* gene (*AOX1TT*) toward which the qPCR primers were directed for the gene dosage determination experiments. (Reprinted from **Paper I** with permission from the publisher.)

3.1.3.3 Translation, Folding, Targeting and Degradation

Sometimes, poor expression is observed despite decent transcript levels (Bonander *et al*, 2005; **Paper I**). Several factors downstream of transcription have been reported to influence the yield of recombinant membrane proteins.

The folding and membrane integration of membrane proteins occur in the ER. If the folding requirement exceeds the folding capacity at this site, as happens upon overexpression of recombinant membrane proteins, the cell responds by adjusting the number of chaperones to meet the folding requirements. If accumulation of unfolded proteins in the ER lumen persists, a coordinated cell response, referred to as the unfolded protein response (UPR), is initiated in the eukaryotic cell to avoid detrimental overload of aggregated polypeptides in the ER. In the UPR, the ribosomal activity is reduced at the same time as the transcription of ER-targeted proteins involved in protein folding and degradation of unfolded polypeptides are

up-regulated (Chawla and Niwa 2005). In one study, the expression levels of membrane proteins in *S. cerevisiae* was compared with the degree of induced UPR, showing an inverse relation between these two parameters, suggesting that UPR should be avoided in order to obtain high expression yields (Griffith *et al.* 2003). By adjusting culturing parameters such as pH or temperature it has been shown that the host response can be reduced at the same time as yields are improved (Bonander *et al.* 2005).

Trafficking and subcellular localization of the heterologous protein may also pose a problem in recombinant membrane protein production. Even though a lot is known about protein sorting and targeting signals, the full story about how membrane proteins reaches their final subcellular destinations has not been unraveled. When trying to express the plant *AtSIP1;1* aquaporin that has been reported to be localized to the ER in its native host, we observed tremendously poor yields that were on the verge of detection despite reasonable mRNA levels (**Paper I**). We speculate that this protein may carry an ER-retention signal that is recognized by the *P. pastoris* protein sorting machinery and that causes it to accumulate in the ER, which induces degradation.

If proper plasma membrane targeting is achieved, there are still factors that could potentially limit the final expression yields. For instance, the lipid composition of the target membrane may affect the protein stability and function. Membranes of different organisms have different lipid compositions which, to make matters even more complicated, may alter in response to environmental cues. Specific lipid-protein interactions are common which is reflected by the presence of lipids in many crystal structures of membrane proteins. Possibly, these lipids are needed to support the three dimensional structure of the proteins (Opekarova and Tanner 2003). One study showed that the function of a GPCR overexpressed in *S. cerevisiae* was altered when the lipid composition of spheroplasts was changed, suggesting the importance of lipid composition in structure and function (Lagane *et al.* 2000). Moreover, the membrane itself poses a threshold to recombinant membrane protein yields since the number of protein molecules that can be accommodated is limited due to its restricted surface area.

3.1.3.4 External Factors

By altering a number of external parameters during cultivation, expression levels may be further optimized. In one report, a Design of Experiments (DoE) approach was applied to examine the impact of several input parameters on expression yields of secreted GFP in *P. pastoris*. By studying three parameters at a time and cross-correlating them, optimal conditions in terms of temperature, pH, dissolved oxygen levels and methanol supply were determined, and by using this method the functional yields of certain GPCRs were also improved (Holmes *et al.* 2009; Bawa *et al.* 2011). In our lab, cultivation in shake flasks at fixed pH and temperature but with sporadic (i. e. every 24th hour) rather than continuous methanol supply used

to be the standard protocol and high yields of crystallization quality *HsAQP8* (**Paper III**), *HsAQP5* (**Paper IV**) and *SoPIP2;1* (Karlsson *et al.* 2003; Tornroth-Horsefield *et al.* 2006) were obtained. However, it has been shown that the more controlled growth in bioreactors have a positive effect of aquaporin production in *P. pastoris*, not only as a result of the significantly increased biomass generated, but also due to improved recombinant expression per cell (Nyblom *et al.* 2007). By applying this cultivation strategy, the quality and amount of *NbXIP1;1* aquaporin produced was improved (**Paper II**) and the yield of functional *HsAQP5* was significantly increased, with 145 mg of pure protein obtained from 1L fermenter culture (unpublished results), as compared with the 10 mg/L obtained from shake flasks.

Factors that could influence recombinant membrane protein expression in *P. pastoris* are summarized in Figure 3.3.

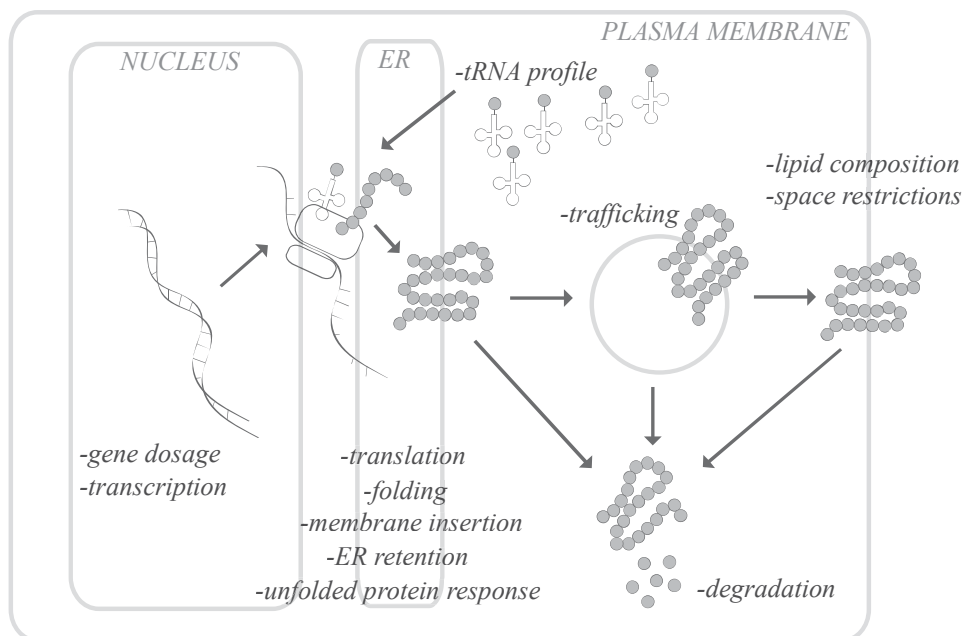


Figure 3.3. Schematic summarizing factors that could potentially influence the recombinant expression yields of a membrane protein.

3.2 Membrane Protein Purification

Once satisfactory expression levels are obtained, an optimized purification scheme needs to be designed in order to efficiently extract the protein out of the membrane and separate it from endogenous proteins. Below, some major bottlenecks for efficient purification are described and possible solutions to these are proposed.

3.2.1 Membrane Preparation

Initially, the membranes of the *P. pastoris* cells, where the protein of interest resides, need to be harvested. Since yeast cells are surrounded by a thick cell wall, a cell lysis step is required. We have compared the effectiveness of two cell breakage methods, the French Press method, where cells are broken due to shear forces caused by pressure variations and the BeadBeater method, where cell lysis is induced by glass bead grinding. The BeadBeater showed to break the cells more efficiently (**Paper II** and Dr. Maria Agemark, unpublished results). Yet, in many instances, cell lysis efficiency is only ~50% (**Paper I-II**) indicating that yields may be drastically increased by optimization of this step. Cultivation time has been proposed to correlate with cell wall thickness of *P. pastoris* cells and may be adjusted to improve the cell disruption efficiency (Bornert *et al.* 2012). Enzymatic digestion of the cell wall may also improve cell lysis.

Once cells are broken, membranes are harvested by ultracentrifugation and according to our experience, aquaporins are usually associated with the membrane fraction, suggesting correct folding and targeting. However this observation does not exclude that mistargeted polypeptides do exist since they may be degraded and hence undetectable.

To decrease the complexity of the membrane protein content in the harvested fractions, a membrane wash can be applied. We have used an urea/alkali treatment which is known to abolish peripherally bound proteins from the membrane and has also been shown to alleviate subsequent membrane solubilization. (Hasler *et al.* 1998; Fotiadis *et al.* 2001). While this method was successfully used in the *HsAQP5* purification scheme, it showed to cause extraction of *HsAQP8* and *NbXIP* from the membrane. Hence, in the latter cases the membrane washing procedure was excluded or modified to reduce losses (**Paper III** and **II** respectively).

3.2.2 Solubilization

In order to separate the protein of interest from endogenous membrane proteins of the expression host, it needs to be extracted from the membrane, a procedure that

is usually accomplished by detergent based solubilization. Detergents are amphipathic molecules with a hydrophilic moiety connected to a hydrophobic tail. They differ from lipids in the sense that they, rather than forming bilayers, assemble into spherical structures referred to as micelles.

Detergents may be grouped according to their composition and chemical properties. The most fundamental categorization is based on charge and divides detergents into the ionic, non-ionic and zwitterionic subgroups. Ionic detergents can be either positively or negatively charged and are generally harsh with denaturing properties. A commonly used ionic detergent is sodium dodecyl sulfate (SDS) that is employed to denature proteins prior to gel electrophoresis. Ionic detergents can be used for efficient membrane protein solubilization in applications such as mass spectrometry analyses, where a retained tertiary structure is not crucial. In functionality assessments and structural determination, however, denaturation should of obvious reasons be avoided and hence the milder non-ionic detergents are most frequently used for these purposes (Prive 2007). The zwitterionic detergents carry both a negative and a positive charge creating a neutral net charge. They are generally less denaturing than the ionic detergents and have been used successfully in crystallization trials (Seddon *et al.* 2004).

The downstream applications should be considered when choosing detergent. If the purified protein is to be reconstituted into an artificial membrane, the critical micelle concentration (CMC), above which monomers will associate and form micelles (Arnold and Linke 2008) should be thought of. For instance, a detergent with low CMC will be prone to micelle formation and will consequently be difficult to remove by dialysis (section 3.3.2 and 3.4.1), since the large diameter of the micelles would prevent the detergent from passing the dialysis tubing. Hence a higher CMC detergent may serve the purpose better. If the protein is to be used for three dimensional crystallization trials (section 3.4.2), additional factors such as detergent sample homogeneity and micelle size may be of greater importance, since particle heterogeneity and extensive embedding of the protein may prevent crystal formation (Prive 2007).

The capacity of a given detergent to solubilize the protein of interest cannot be predicted and hence empirical screens are necessary to identify the detergent that generates maximum recovery. Typically, uncharged (nonionic or zwitterionic) detergents with different properties regarding size of the head group and length of the hydrophobic tail are surveyed for their capacity to extract the protein from the membrane and keep it stable in solution subsequent to solubilization. As a general rule of thumb, detergents with a small head group and a short hydrophobic tail tend to have more denaturing properties and are referred to as being harsh as opposed to mild detergents with larger head groups and longer tails (Prive 2007). Harsh detergents are generally easier to work with in downstream applications due to their small size and high CMC and could preferably be analyzed in the initial screens. However, if aggregation which implies denaturation is observed in the

solubilized sample, a detergent with milder properties should be considered. The effects of harsh versus mild detergent properties was observed when purifying *HsAQP8* (**Paper III**) which aggregated after solubilization with n-Octyl- β -D-glucopyranoside (OG) but remained stable in solution when extracted with milder detergents such as n-decyl- β -D-maltopyranoside (DM) and n-dodecyl- β -D-maltopyranoside (DDM).

To conclude, the choice of detergent is not easy since the detergents most suitable for functional assessments and structural studies are the ones that are most prone to induce denaturation of the solubilized protein. However, since a difficult downstream experiment is still much better than a nonexistent ditto, the primary focus during the solubilization trials should be on getting the protein of interest stable in solution.

3.2.3 Obtaining a Pure Protein Preparation

Finally, to obtain a pure preparation, the protein of interest needs to be separated from the endogenous membrane proteins that still remain in the solubilized fraction. By utilizing differences in charge, size or binding affinity to specific molecules, separation is obtained. We have most commonly used immobilized metal affinity chromatography (IMAC) for aquaporin purification (**Paper II** and **III**). A poly histidine (His) sequence added at either of the termini of the construct is utilized for this purpose. Since the side chain of histidine can be coordinated by divalent metal ions such as nickel, His-tagged proteins can be trapped in a nickel containing resin while other proteins remain in solution. After washing away unbound proteins, the protein of interest can be eluted by introducing an agent, such as imidazole, that competes with the histidine-metal ion interactions.

However, even though a His-tag significantly aids in the purification process, it confers a modification of the protein primary sequence that may influence the folding or function of the protein. Hence other strategies such as a separation based on charge, as accomplished by ion exchange chromatography, can be applied (**Paper IV**). Alternatively a protease cleavage site may be introduced proximally to the tag, enabling its removal subsequently to the affinity chromatography trials.

A second purification step, by the means of size exclusion, is commonly employed after the affinity- or ion exchange chromatography to further improve homogeneity (**Paper IV**). Besides improving the purity of the sample, size exclusion chromatography also serves as a quality control step since any polydispersity caused by aggregation or degradation of the protein will be revealed.

3.3 Functionality Assessments

There are several methods that can be applied in order to reveal new functionality features of an uncharacterized membrane protein or to confirm that functionality is retained after recombinant expression and subsequent purification. Some of the most common methods used to establish aquaporin functionality and substrate specificity are described below.

3.3.1 *In Vivo* Assays

To assess aquaporin functionality *in vivo*, transgenic and knock-out systems have been developed. In the most classical example, eggs from the frog *Xenopus laevis* are injected with cRNA, encoding the gene of interest, which is subsequently expressed and inserted into the plasma membrane of the oocyte. By introducing an osmotic gradient that causes water to move either in to or out of the cell, the permeability of the aquaporin containing membrane can be measured and compared to that of the uninjected control eggs which exhibit an intrinsically low water permeability. This method was used when the water transporting capacity of the aquaporins was initially discovered (Preston *et al.* 1992). By employing radioactively labeled molecules and monitoring their accumulation inside the oocytes, it is also possible to verify aquaporin substrates other than water using this system.

Several-knock out systems in plants and mammals also serve to establish aquaporin function (section 2.4 and 2.5) and perhaps applying such methods gives the best view of the biologically relevant properties of these channels. However, it has often proven to be difficult to observe knock-out related phenotypes, possibly due to redundancy among the aquaporins and many isoforms may only have a crucial function at specific stress conditions which may be hard to predict.

Transgenic yeast strains have also been used to validate aquaporin substrate specificities. In one approach that resembles the oocyte swelling assay, aquaporin expressing yeast cells were stripped of their cell walls to generate spheroplasts which were subsequently subjected to hypo-osmotic shock. By monitoring the decrease in light scattering or the optical density at 600 nm, the spheroplast lysis rate which reflects membrane water permeability was recorded (Laize *et al.* 1999; Daniels *et al.* 2006). Methods for identification of other aquaporin substrates based on yeast complementation assays have also been developed. By deleting endogenous transporters of the molecule under survey and subsequently observing growth phenotype recovery after introduction of the aquaporin of interest, new substrates may be identified (Jahn *et al.* 2004; Bienert *et al.* 2007).

In **Paper II** we have applied a simplified yeast growth assay to monitor the activity of overexpressed *NbXIP1;1*, which has previously been shown to transport boric acid (Bienert *et al.* 2011). By plating dilution series of *P. pastoris* cells with induced *NbXIP1;1* expression on plates containing boric acid we were able to observe an impaired amplification rate of the recombinant cells. This phenotype could be readily distinguished from that of the wild type cells, suggesting that any background boron transport conferred by endogenous transporters was overridden by the large amount of recombinant *NbXIP1;1* present in the membrane.

3.3.2 Stopped Flow Measurements on Reconstituted Protein

One major drawback with the *in vivo* transgenic systems is the presence of endogenous biomolecules and signaling substances that may affect the function of the protein of interest. Moreover, if the direct effects of a given substance on the aquaporin function are to be studied it is preferable to have the protein of interest positioned in a more controllable system, such as an artificial lipid bilayer. By combining solubilized proteins with solubilized lipids and subsequently withdrawing the detergent, lipids and proteins will spontaneously associate and form proteoliposomes, which are lipid bilayers in a spherical conformation with integrated membrane proteins. There are several means by which detergent removal can be achieved (Rigaud and Levy 2003).

In the simplest approach, dilution of the lipid, protein, detergent ternary mixture is applied to lower the detergent concentration below the CMC. Dilution induces a shift in the free versus associated detergent equilibrium which causes detergents to dissociate from the lipids and the proteins. This in turn induces the spontaneous formation of proteoliposomes which subsequently can be harvested by centrifugation. The dialysis method (Figure 3.4) is related to the dilution method and has frequently been used for aquaporin reconstitution in our lab (**Paper III and IV**). By enclosing the lipid-protein-detergent ternary mixture in dialysis tubing through which only the detergent monomers are small enough to penetrate, detergent removal will be obtained when the dialysis bag is positioned in a large volume of dialysis buffer. The benefit of this method over the dilution method is that lipids and proteins remain concentrated since they are trapped in the dialysis bag, and that no additional centrifugation step is required. The major drawback with both methods listed above is that they are only suitable when working with high CMC detergents that do not require huge dilution volumes to dissociate from the solubilized protein and lipids and that are not prone to form micelles that would be trapped inside the dialysis tubing due to their large diameter. To remove low CMC detergents from solution, adsorption methods are applied. By introducing hydrophobic polystyrene beads or cyclodextrin, which is a barrel shaped molecule with a hydrophobic interior (Degrip *et al.* 1998), detergents are adsorbed and proteoliposome formation is accomplished.

Once proteoliposomes are obtained, *in vitro* functional assessments are enabled. To investigate aquaporin mediated water transport, the proteoliposomes are exposed to a hypertonic solution by rapid mixing in a stopped flow device (Mathai and Zeidel 2007). The outwardly directed osmotic gradient will induce water efflux from the proteoliposome that causes it to shrink, and this shrinkage can be monitored as an increase in light scattering perpendicular to the light beam (**Paper III** and **IV**). Alternatively the liposomes may be loaded with a fluorescent dye prior to measurements, and the self-quenching induced by the increased concentration of the dye upon water efflux can be recorded as a reduction in fluorescence. If transport of a solute distinct from water is to be analyzed, a slightly different approach could be applied. By loading proteoliposomes with the

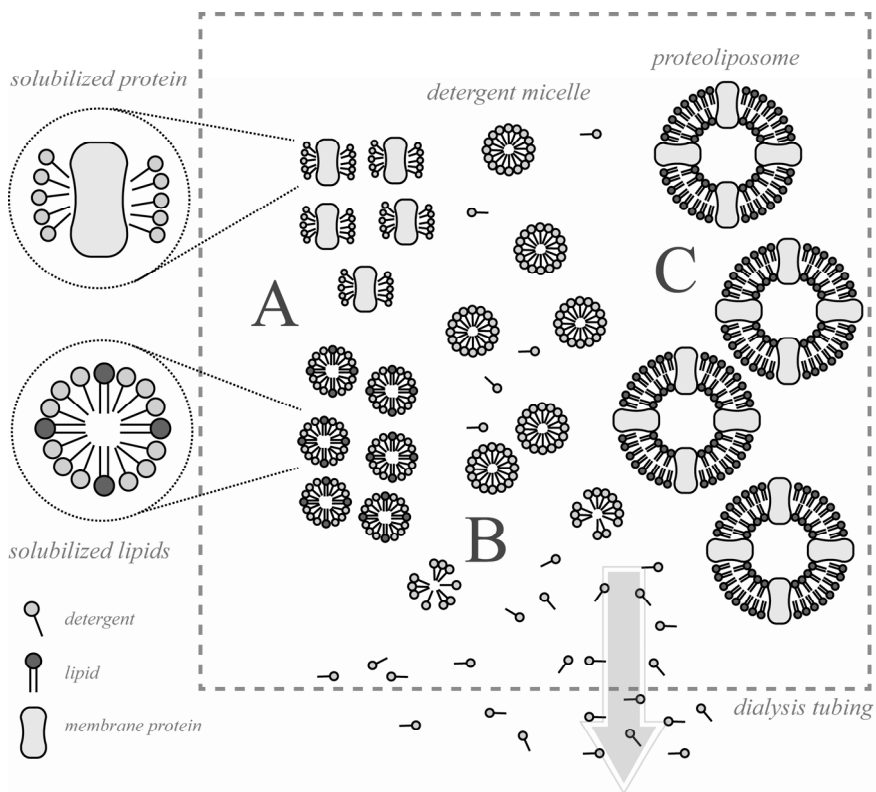


Figure 3.4. Schematic showing the principles of membrane protein reconstitution by dialysis. **A.** Solubilized protein and lipids are mixed to a desired lipid to protein ratio (LPR) in dialysis tubing. **B.** The dialysis bag is positioned in a large volume of dialysis buffer and detergent molecules will be withdrawn from the protein and lipids due to dilution effects since detergent monomers are small enough to penetrate the dialysis tubing. **C.** When detergent is removed, lipids and protein will start to spontaneously form proteoliposomes.

proposed substrate and exposing them to a buffer that lacks the substrate but is otherwise isosmotic, a water efflux that is solely dependent on the outward diffusion of the solute under survey is induced. The water efflux can be recorded by applying the self-quenching assay described above. However, if the aquaporin under survey is a poor water transporter, the water flux may be limited since water will pass mainly through the lipid bilayer. A restricted water efflux would lead to an underestimation of the channel mediated substrate transport rate. Besides offering a means of characterizing aquaporins, the proteoliposome/stopped flow approach also poses a possibility to screen for inhibiting or activating agents.

3.4 Crystallization and Structure Determination

In order for structure determination to be achieved, the protein of interest needs to be crystallized since a single protein molecule would not be able to diffract electrons or X-rays. In a protein crystal, the protein molecules are arranged in a symmetrical fashion where the smallest repetitive element is referred to as a unit cell. Crystals could be either two-dimensional (2D), with the protein embedded in a lipid bilayer, or three-dimensional (3D), where solubilized protein molecules arrange into a well ordered lattice held together by non-covalent interactions between the polar parts of the protein. These two methods with their respective pros and cons will be briefly described in the following sections.

3.4.1 Two-Dimensional Crystallization

By dense packing of membrane proteins situated in a lipid bilayer, two-dimensional (2D) crystalline structures can be obtained. Such 2D crystals can be used for structure determination by electron microscopy. The benefit of 2D crystallization lies within the presence of lipids in the crystalline lattice. Specific lipid-protein interactions and the lateral support that the lipids offer may be crucial for retaining the native conformation of some membrane proteins. Moreover, since loosely attached lipids are retained in the crystal, interactions between the membrane protein and adjacent lipids may be studied (Figure 3.5).

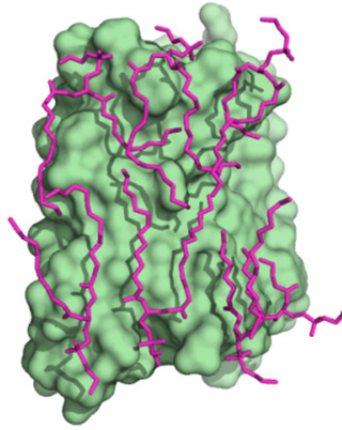


Figure 3.5. The AQP0 monomer (PDB ID: 3M9I) and annular lipids that occupies grooves of the surface that faces the lipid bilayer. The AQP0 protein was reconstituted in *E. coli* polar lipids to form 2D crystals and the structure was solved to 2.5 Å resolution by electron diffraction (Hite *et al.* 2010).

In some rare cases 2D crystallization occurs spontaneously in nature, as is the case with AQP0 in the eye lens and with bacteriorhodopsin which forms highly ordered 2D-arrays in the *Halobacterium salinarum* plasma membrane. The naturally occurring 2D-crystals of bacteriorhodopsin were used to determine the very first, three-dimensional membrane protein structure in the 1970s (Henderson and Unwin 1975). However, in most cases membrane proteins need to be solubilized, purified and reconstituted into an artificial lipid bilayer in order for 2D crystallization to take place. Reconstitution is commonly achieved by detergent removal (section 3.3.2 and figure 3.4), from a mixture of solubilized protein and lipids with a very low lipid to protein ratio, by dialysis (Rigaud *et al.* 2000).

There are several parameters that may affect the crystallization process that need to be optimized in order to obtain highly ordered 2D crystals. The choice of lipids and detergent may be crucial as well as lipid to protein ratio (LPR), speed of detergent removal, temperature, pH and composition of the crystallization solution. To unravel the effects of all these variables, a vast number of screening conditions need to be evaluated. This is a time consuming process and the screening procedure has indeed constituted a major bottleneck in electron crystallography. In the last few years, progress towards automation of the 2D crystal screening has been made by introducing small scale crystallization in microtiter plates and liquid handling robots (Ubarretxena-Belandia and Stokes 2012). By enabling screening of more variables, the hitherto relatively poor success rate of 2D crystallization may be improved, making it a more attractive alternative or complement to 3D crystallization for determination of membrane protein structures at high resolution.

In **Paper III**, the 2D crystallization of *HsAQP8* is described. Numerous parameters were optimized in order to promote crystallization, and well-ordered and diffracting crystals were obtained. However, despite the large efforts put into crystal screening, no high resolution data was achieved from the subsequent electron microscopy trials. This illustrates the major drawback with the electron crystallography method; since the limited expansion of the crystals in the third dimension will entail weak diffraction and since these thin crystal structures will suffer from damage by the electron beam that carries a lot of energy, atomic resolution data is seldom acquired (Rhodes 2006).

3.4.2 Three-Dimensional Crystallization

The most frequently used technique to achieve high resolution membrane protein structures is based on X-ray diffraction data attained from three dimensional protein crystals (Bill *et al.* 2011). 3D crystals are commonly obtained by slow precipitation which is achieved by a gradual increment of the concentration of the protein and precipitants (such as salts or polyethylene glycol (PEG)). This is usually done by applying vapor diffusion methods, such as hanging drop crystallography (Figure 3.6 B). In this method, a small droplet of buffer solution containing the protein and a low concentration of the precipitant is allowed to equilibrate with a larger reservoir of buffer, containing a concentration of precipitant believed to be optimal for crystal growth, in a sealed system. Water will start to move from the droplet and into the reservoir solution which leads to an increase in protein and precipitant concentration in the droplet which will induce nucleation (Figure 3.6 A), which is the first step in crystal formation. When nucleation is initiated, the protein concentration in the droplet will decrease which counteracts further nucleation, and rather growth of the crystal will be promoted. Generally, the slower the precipitation is, the bigger the crystals will grow and the less is the likelihood of amorphous aggregation of the protein (Rhodes 2006). A big crystal is preferred to a smaller one since it is more likely to generate high resolution data.

3D-crystallization of membrane proteins is particularly challenging due to the presence of detergents that surrounds the hydrophobic parts of the protein and that may induce heterogeneity due to non-specific binding and obstruct surfaces important for crystal contact formation. Several measures could be taken to improve membrane protein crystals. For instance crystal contacts may be favoured by the use of a detergent with a smaller micelle radius (Prive 2007) or by extending the hydrophilic surface by the addition of antibody fragments (Hunte and Michel 2002).

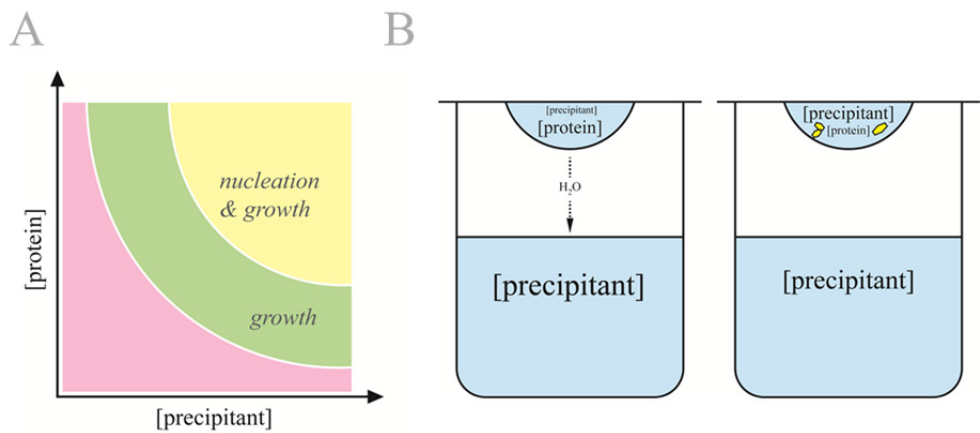


Figure 3.6. **A.** Schematic illustration of a phase diagram showing the correlation between conditions that favor nucleation (yellow) and crystal growth (green). At low protein and precipitant concentrations (pink) the solution is not saturated and no precipitation will occur. **B.** In hanging drop crystallography a small droplet containing the protein of interest and a low concentration of the precipitant is allowed to equilibrate with a larger reservoir of buffer with a higher precipitant concentration. Initially water will move out of the droplet and the conditions will move from the pink to the yellow area of the phase diagram as protein and precipitant concentrations are increased. At these conditions nucleation will be induced and protein will be withdrawn from the solution. Due to the lowered protein concentration in solution, the conditions will be shifted towards the green field in the phase diagram. At these conditions, growth rather than nucleation is promoted and hence few but large crystals will be obtained rather than plentiful small ones.

3.4.3 How to Make Use of the Structure

The step from a well ordered, diffracting protein crystal to a high resolution protein model requires advanced radiation physics and complex mathematical calculations that will not be covered here since it is not within the scope of this thesis.

With a high resolution membrane protein structure at hand, a range of possibilities to characterize the protein emerges. Primarily the atomic model may reveal unknown structure related features of the protein of interest. For instance, novel substrates or regulatory mechanisms may be identified. The model of the structure can also be used for molecular dynamics simulations, which for aquaporins have been frequently used to study movement of solutes through the pore (de Groot and Grubmuller 2005; Hub and de Groot 2008). Moreover, by performing computer based docking experiments with ligands, drug leads may be identified. By screening a library of ligands for binding to *HsAQP5*, potential inhibitors were identified (Figure 3.7) and these are currently surveyed *in vitro* for water permeation inhibitory effects. Since aquaporins are involved in pathological conditions such as tumor angiogenesis and edema, inhibitors may be of pharmaceutical interest.

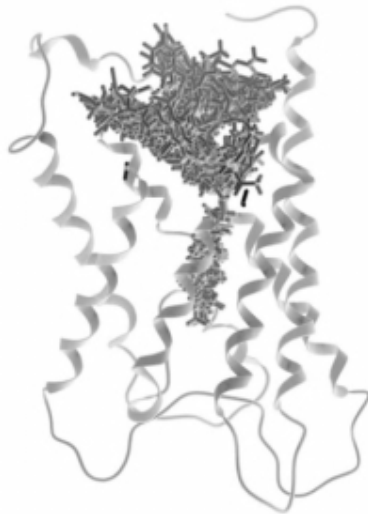


Figure 3.7. Model of a *HsAQP5* monomer and an overlay of potential inhibitors identified by *in silico* docking experiments.

3.5 Identification of Aquaporins at Novel Sites

In order to increase the understanding of the biological significance of aquaporin mediated water and solute transport it is of importance to map the temporal and spatial expression patterns of these channels. There are several methods that can be applied in order to verify aquaporin occurrence, some of which that are used in plant research I will describe in this section.

To assess the tissue specific activation of aquaporin promoters in plants, the β -glucuronidase (GUS) reporter system (Jefferson *et al.* 1987) is commonly employed. In this assay, the promoter of the aquaporin isoform of interest is connected to the GUS encoding *uid* gene and subsequently incorporated into the plant genome by transfection with *Agrobacterium tumefaciens*. The GUS protein catalyzes the cleavage of the artificial X-gluc substrate which, is infiltrated into the tissue, generating a clear blue precipitate that reports in which tissues the promoter is activated (Figure 3.8).

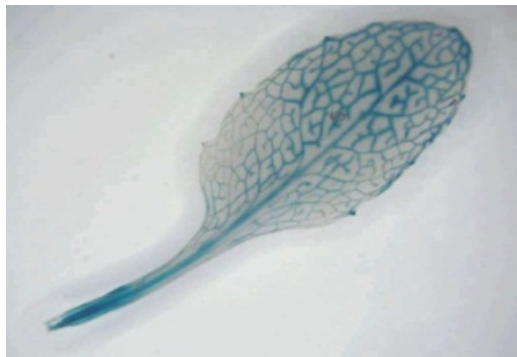


Figure 3.8. Leaf from a plant transformed with an AQP-GUS construct showing aquaporin promoter activation (blue staining) in the vascular tissues. (This picture was kindly provided by Professor Urban Johanson).

Another common approach used to identify tissue specific aquaporin expression patterns is to screen different tissue sections for the presence of the corresponding mRNAs. This can be done either by the use of quantitative polymerase chain reactions (qPCR) or by microarray techniques. Besides conferring information regarding localization these methods are also quantitative and offer a means of assessing transcriptional regulation. These methods were applied when the tissue specific up or down regulation of *A. thaliana* aquaporin mRNAs in response to drought stress were analyzed (Alexandersson *et al.* 2005).

The promoter activity and mRNA detection assays are superior when mapping the tissue specific expression of membrane proteins, but do not give any information regarding subcellular localization patterns. In order to confirm in what parts of the cell the aquaporins reside, methods based on GFP fusion constructs or antibody detection in tissue sections have commonly been used. However the addition of the relatively large GFP protein may alter the localization pattern and antibodies may exhibit non-specific binding, potentially introducing artifacts.

By using an approach based on western blotting for identification of specific proteins in purified and solubilized membrane samples, signals derived from non-specific antibody binding can be distinguished from “true hits” since information about protein size is integrated in this method. Since aquaporins frequently show a very characteristic migration pattern on SDS-PAGE gels with multiple bands corresponding to the different oligomeric states (**Paper V**), additional confidence is added to a positive result. A drawback with this method as compared with *in situ* immunolocalization of proteins at subcellular sites is the requirement of a perfectly pure membrane preparation since contaminating membranes may introduce false positive results. By performing contamination control experiments with antibodies directed towards marker proteins for different types of membranes, the purity of the fractions analyzed may be validated.

If, as in our case, an antibody directed towards an epitope that is common to many isoforms is applied during the immunoblotting trials, further resolution may be obtained by employing mass spectrometry analyses. By combining the time of flight (TOF) and iontrap mass spectrometry methods we were able to establish the identity of several aquaporin isoforms in chloroplast samples (**Paper V**).

Summary of the Present Investigation

In this section, the major findings obtained during the projects summarized in this thesis will be presented and discussed. Four of the pieces (**Paper I-IV**) are focused on overexpression, purification and characterization of specific aquaporin isoforms, and will be presented in parallel. By comparing the schemes developed for successful production of the different aquaporin isoforms, I will try to pinpoint major hurdles encountered during membrane protein production and present strategies to overcome these. An overview of the methods used and the results obtained in **Paper II-IV** is presented in Figure 4.1.

4.1. Overexpression

To circumvent the limitations caused by scarceness of membrane proteins in nature, recombinant expression hosts have been developed for large scale membrane protein production. In our lab we have routinely employed the yeast *P. pastoris* to produce aquaporins. Initially, efforts were put into optimizing the cultivation and purification protocols rather than into the generation of strains with multiple gene inserts and the associated putatively higher recombinant expression levels. Hence protein fractions used for the characterization of *SoPIP2;1* and *HsAQP8* were purified from single copy clones (unpublished results). In **Paper I**, the effect of gene dosage on recombinant expression levels of four aquaporin isoforms (*HsAQP5*, *HsAQP8*, *SoPIP1;2* and *AtSIP1;1*) in *P. pastoris* was evaluated. A qPCR-based method for a straight forward and generic gene copy quantitation was developed to aid in the evaluation process. Expression levels of all four isoforms under survey responded strongly to increases in gene dosage, suggesting a positive effect of multiple gene copies on aquaporin overexpression. The *HsAQP5* protein used for functional assessments and structure determination in **Paper IV** and the *NbXIP1;1* protein used for the yeast growth assay and purification trials in **Paper II** both derive from multicopy clones showing that the increased workload imposed on the yeast cells protein production machinery does not seem to impair neither protein folding, targeting to the membrane, nor function.

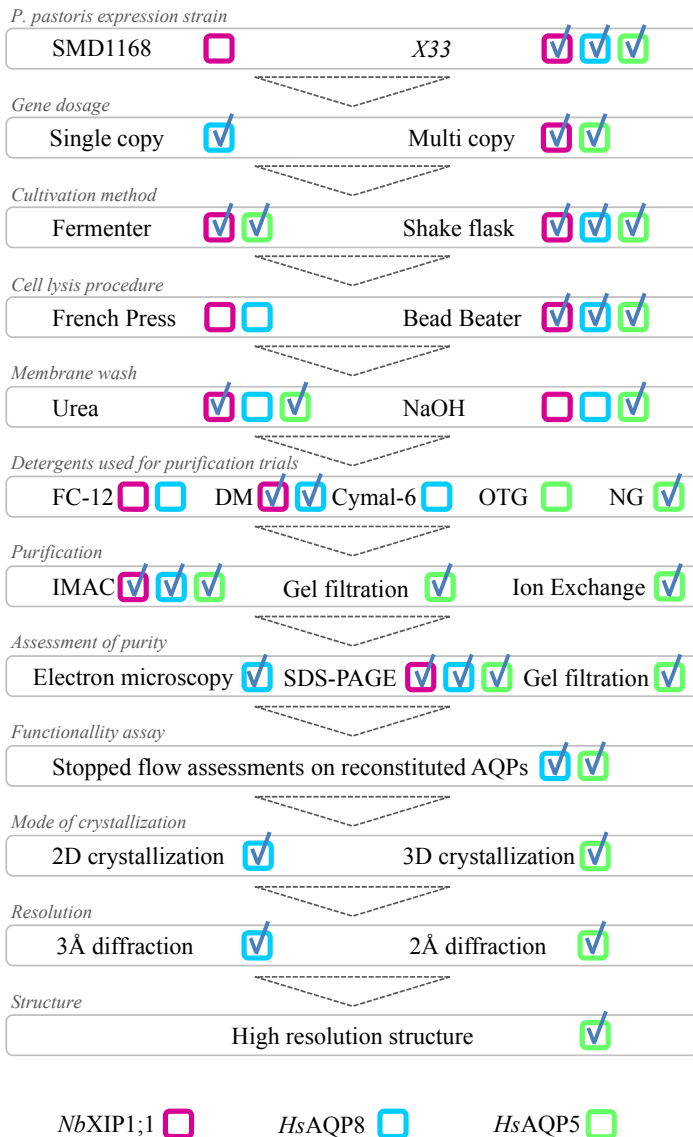


Figure 4.1. Summary of methods used and results obtained in **Paper II, III and IV**. Boxes in pink, blue and green indicate conditions evaluated or results obtained during experiments with *NbXIP1;1*, *HsAQP8* and *HsAQP5* respectively. Checked boxes indicate successful conditions.

The multicopy strategy is now routinely employed in early phase membrane protein production trials in our lab and if satisfactory expression levels are not obtained subsequent to transformation, gene dosage is established by the qPCR method presented in **Paper I** to validate the outcome of the antibiotic selection. If

no high copy clones are identified, efforts are put into redoing the transformation and zeocin selection, since high expression levels may save a lot of time and endeavor in subsequent purification steps. Apart from several aquaporin isoforms, both GPCRs and transient receptor potential (TRP) ion channels have now been successfully overproduced from multicopy *P. pastoris* clones in our lab (unpublished results).

However, there are other factors, besides recombinant gene dosage, that could influence recombinant membrane protein production (section 3.1.3). This became particularly evident when comparing expression levels of clones carrying a single copy of different aquaporin genes (**Paper I**). Expression of *AtSIP1;1* proved to be exceptionally low as compared with that of *HsAQP5*, *HsAQP8* and *SoPIP1;2*. When analyzing the *AtSIP1;1* mRNA levels it appeared that the quantity of recombinant transcripts was low as compared with that of a single copy *HsAQP5* clone. However, *AtSIP1;1* transcripts were still several times as abundant as actin transcripts, suggesting that even though transcription was impaired or that degradation of the transcripts was occurring, additional downstream factors were also contributing to the poor expression levels.

One valuable experience we made during the multicopy evaluation project was that prolonged incubation with zeocin, which is the antibiotic used in the selection for multicopy clones, may increase the occurrence of so called false positive clones i.e. clones with low gene dosage on high antibiotic levels. This could either be a result of a reduced antibiotic effect over time or an acquired antibiotic resistance, not coupled to recombinant gene dosage, of the clones. The conclusion drawn was that the plates used in the antibiotic selection should not be incubated for more than three days, since after that time interval the occurrence of false positives increased dramatically

A problem that, according to our experience, is frequently encountered during membrane protein production is proteolysis. To reduce degradation of the *NbXIP1;1* protein, expression in the protease deficient SMD1168 strain was evaluated (**Paper II**). A reduction in degradation effects was observed, but since overexpression was impaired we choose to employ the wild type X33 strain for large scale production. Subsequently it also appeared that controlled cultivation in a fermenter in combination with careful sample handling could almost eradicate the previously observed degradation. Besides having a positive effect on protein stability, bioreactor cultivation also induced significantly increased yields of purified *HsAQP5* and *NbXIP1;1* as measured per grams of cells. Moreover, since fermenter growth generally renders much higher cell densities than cultivation in flasks, it should be the method of choice for large scale membrane protein production.

4.2. Purification

Aquaporins exhibit relatively high sequence identity, shared topology and overall structure. Yet purification schemes need to be carefully optimized in order to obtain maximal yields of the individual isoform since factors such as differences in protein-lipid interactions, presence of protease cleavage sites and the proteomic set up of the recombinant host cell may differ significantly. In this section I will describe how purification procedures of the three aquaporin isoforms *HsAQP5*, *HsAQP8* and *NbXIP1;1* were optimized.

Achievement of efficient cell lysis constitutes the first challenge in membrane protein purification. Regarding lysis of recombinant *P. pastoris* cells we have routinely used the BeadBeater method by which a more efficient cell breakage has been obtained than when applying a French Press device. However, still a significant fraction of the cells remain intact even after repeated BeadBeater runs and subsequently a substantial loss of recombinant protein is observed in this step (**Paper II**). Efforts should be put into developing the cell lysis procedure since a more efficient cell breakage could significantly increase the final yields.

A urea/alkali based membrane wash, to eradicate peripherally bound membrane proteins, was evaluated (**Paper II, III and IV**). While *HsAQP5* seemed to withstand the washing, the other two isoforms were extracted from the membrane by this treatment. Regarding *NbXIP1;1*, the urea wash was included in the purification scheme since it did not induce as a major loss of the protein as was observed during the alkali treatment. Human AQP8 on the other hand did not withstand either of the washing steps and hence these were omitted. Successful urea/alkali treatment may significantly aid in subsequent purification steps since the protein of interest may be highly enriched during this process (Figure 4.2).

The next major hurdle encountered in membrane protein purification is solubilization by detergents (section 3.2.2.) to release the integral proteins from the membrane. Human AQP5 was successfully solubilized with n-Octyl- β -D-Thioglucopyranoside (OTG) and n-Nonyl- β -D-Glucopyranoside (NG) which are both high CMC detergents that are readily compatible with subsequent reconstitution and crystallization experiments. Initial purification experiments with OTG revealed that this detergent tend to precipitate at low temperatures and to avoid clogging of the tubing in the ÄKTA system, NG was applied in subsequent purification trials. High CMC detergents did not solubilize *NbXIP1;1* efficiently and caused *HsAQP8* to precipitate. Hence low CMC detergents with longer hydrophobic tails were evaluated. FOS-CHOLINE[®]-12 (FC-12) which is a zwitterionic detergent solubilized these proteins well but turned out to interfere with the subsequent nickel affinity chromatography. Possibly the charge of the head group affected the binding of the His-tag to the nickel or caused a dissociation of the metal ions from the agarose matrix. Cymal-6 was also a good

candidate for *HsAQP8* solubilization but the final yield after purification was not satisfactory. Instead DM proved to be the best alternative for *HsAQP8* solubilization and this detergent was also used in the *NbXIP1;1* purification scheme.

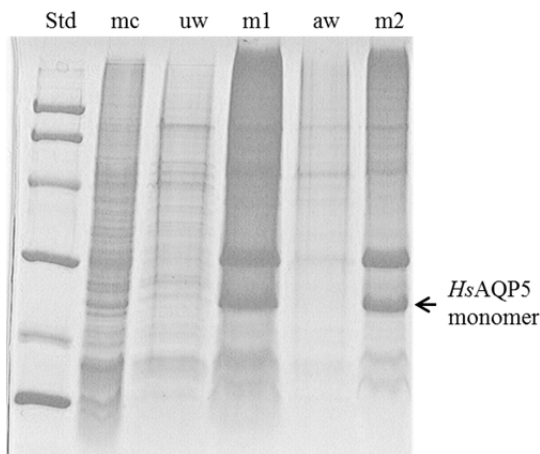


Figure 4.2. SDS-PAGE with the fractions from the urea/alkali wash of *P. pastoris* membranes from cells expressing *HsAQP5*. Protein content in the crude membrane fraction (mc), urea wash (uw), urea treated membranes (m1), alkali wash (aw) and alkali treated membranes (m2) is visualized. The *HsAQP5* monomer is indicated.

Immobilized metal ion affinity chromatography was used for purification of His-tagged *HsAQP5*, *HsAQP8* and *NbXIP1;1* constructs. Since *HsAQP5* was already highly enriched after the urea/alkali wash, satisfactory purity was achieved by a simple batch technique in which protein binding, washes and elution was performed in a test tube by repeated resuspension of the agarose matrix in the different buffers (unpublished data). For the *HsAQP8* and *NbXIP1;1* proteins more controlled conditions with the matrix packed in a column was required to obtain pure fractions. Moving the IMAC purification of the *NbXIP1;1* protein from a column system based on gravity flow to an ÄKTA system with controlled pressure and flow over the column further improved purity of this protein. The non-tagged *HsAQP5* protein was purified by ion exchange chromatography and subsequent gel filtration. This procedure was fairly straight forward, again due to the major enrichment of *HsAQP5* after urea/alkali washes.

To assess purity, several different methods have been applied. All three proteins were analyzed by SDS-PAGE to check for major contaminants. To further validate sample homogeneity the *HsAQP8* protein was analyzed by electron microscopy, while the chromatogram of the gel filtration served the same purpose for *HsAQP5*.

4.3. *In Vitro* Functionality Assessments

The two purified human aquaporin isoforms were reconstituted into proteoliposomes in order to verify that their functionality was retained after overexpression and purification. To induce proteoliposome formation, detergent withdrawal by dialysis (section 3.3.2) was performed. Since *HsAQP5* was solubilized with NG which is a high CMC detergent, proteoliposomes were formed after 48 h of dialysis. Human AQP8 on the other hand was solubilized with DM which has a much lower CMC and hence the time required for proteoliposome formation was extended to 168 h. Another difference in the reconstitution scheme was the lipid composition. A complex lipid mixture extracted from *E. coli* was successfully used in reconstitution of *HsAQP5*. These lipids did however not form liposomes after solubilization with DM even after prolonged dialysis, suggesting that the affinity of this detergent to the *E. coli* lipids was too high and that detergent removal could not be obtained. Hence a defined 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoserine/1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPS/POPC) lipid mixture was applied to obtain *HsAQP8* reconstitution. These lipids created tight control liposomes with exceptionally low water permeability which rendered them inert to the osmotic gradient applied in the stopped flow assay. Both *HsAQP5* and *HsAQP8* proved to be functional since proteoliposomes exhibited much faster shrinkage rates than control liposomes without reconstituted protein. Moreover, water efflux rates could be correlated with LPR values and the lower the LPR the faster the proteoliposome shrinkage.

Mercury is known to inhibit the function of most aquaporins since it binds to cysteine residues in the channel, causing occlusion of the pore (Savage and Stroud 2007). By adding mercurial salts to the *HsAQP5* and *HsAQP8* proteoliposomes it was established that water was moving mainly through the water channel pores since shrinkage rates were significantly reduced. Water flux through *HsAQP5* channels was almost completely inhibited by incubation with mercury whereas a ~40% inhibition was observed for *HsAQP8* proteoliposomes. This could be explained by the expected random insertion of protein into the lipid bilayer. Since only ~50% of the tetramers would have their extracellular domain, harboring the cysteines, facing towards the outside of the proteoliposome, only this half of the protein population would be available for mercury inhibition. Liposomes created by *E. coli* lipids appeared to be leakier than the POPS/POPC liposomes and hence the complete inhibition observed for *HsAQP5* could be explained by permeation of mercury into the proteoliposome interior rendering the entire protein population exposed to this inhibitory compound.

4.4. Crystallization and Structure

Two different approaches were used to crystallize *HsAQP8* and *HsAQP5*. By applying 2D crystallization, highly ordered *HsAQP8* crystals diffracting down to 3 Å were obtained. Optimal conditions for 2D array formation were attained using DM as a detergent and a 3:7 molar ratio of 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE) and 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) lipids. Crystal quality was further improved by proteolytic removal of the C-terminal His-tag and by subjecting the sample to carefully controlled temperature alterations. Despite the major efforts put into the 2D crystallization trials, atomic resolution data was not achieved. The only high resolution aquaporin structures obtained by electron crystallography are those of AQP0 and AQP4 (Gonen *et al.* 2005; Tani *et al.* 2009; Hite *et al.* 2010). Possibly, these proteins' innate capacity to form ordered arrays (section 2.5.1) was critical to attain this success. Due to the intrinsic limitations of the electron crystallization method (section 3.4.1), 3D crystallization techniques may be required in order to achieve a high resolution *HsAQP8* structure.

Human AQP5 was crystallized by hanging drop vapor diffusion to generate 3D crystals (section 3.4.2). Both His-tagged and native protein constructs rendered diffracting crystals, but higher resolution data was obtained from the crystals of the non-tagged protein. Possibly, the increased flexibility inflicted by the extension of the C-terminus impaired crystal packing. However, the crystals of the native protein suffered from twinning effects. In a twinned crystal, two crystalline lattices are superimposed and rotated in relation to each other. Hence, the symmetry of the crystal is distorted which influences the diffraction pattern that becomes difficult to interpret. Since several native aquaporin constructs have rendered crystals with more severe twinning than their His-tagged counterparts it has been speculated that the increased asymmetry induced by the terminal extension that the tag confers may prevent twinning effects (Backmark *et al.* 2008). Luckily the *HsAQP5* twinning issue could be solved by advanced computational refinement methods.

The structure of *HsAQP5* was solved to 2.0 Å resolution, at which atomic details such as water molecules and amino acid side chains can be readily distinguished. No other human aquaporin structure had previously been solved to atomic resolution. According to its primary sequence the *HsAQP5* is classified as a pure water transporter and, as expected, the model of the structure revealed a membrane spanning pore resembling that of other water transporting aquaporins. No surprises regarding tertiary or quaternary structure were uncovered in the structure since *HsAQP5* adapts the classical aquaporin helical bundle and tetrameric conformation (section 2.1.). However, somewhat unexpectedly, a phosphatidylserine lipid was found to be wedged in-between the monomers,

occluding the central pore of the tetramer (Figure 4.3). In all other mammalian AQP structures available the central pore does not appear to be occupied and a similar feature had previously only been observed in the *E. coli* AQPZ structure, where a phosphatidylethanolamine lipid could be modeled into the center of the tetramer. Since in most AQP structures a lipid in the central pore is absent, it is unlikely that the lipid-protein interactions at this site are crucial for tetrameric stability. Neither, the function of the lipid as a blocker of the central pore could be justified since a constriction at the cytoplasmic side of the protein would have rendered the pore impermeable also in the absence of a lipid. Hence, if the lipid in the central pore has a biological function, it remains elusive.

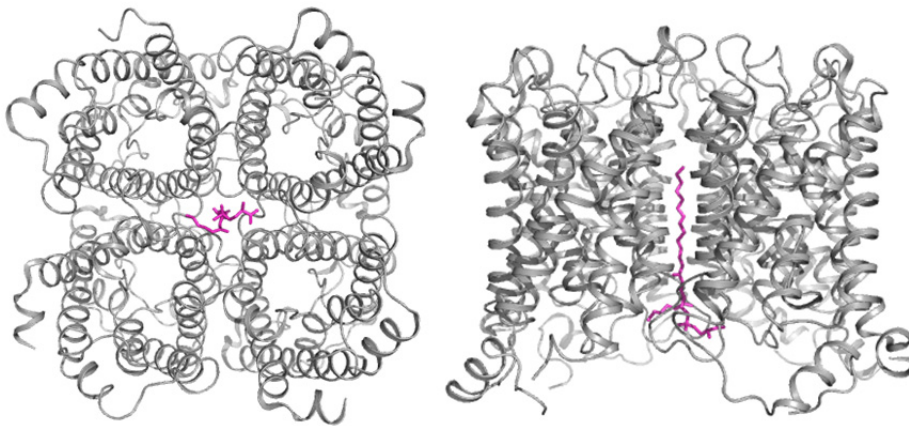


Figure 4.3. Cytosolic- (left) and Side- (right) view of the *HsAQP5* tetramer showing the phosphatidylserine lipid (magenta) that is situated in the central pore.

Aquaporin 5 has been reported to be re-localized from intracellular storages to the plasma membrane in response to phosphorylation. Interestingly, the structure revealed several phosphorylation sites to be exposed to the cytoplasm, accessible for phosphorylation. The D-loop harbors two serines (S152 and S156) that can potentially be phosphorylated and subsequently induce the dissociation of the loop from the C-terminus that it is anchored to by hydrogen-bond interactions. If such a conformational change would induce trafficking or if phosphorylation *per se* would suffice to signal for subcellular re-localization, remains to be elucidated. Interestingly, the structure of the S156E *HsAQP5* mutant, where phosphorylation of the S156 is mimicked by the introduction of a negatively charged amino acid residue, was recently deposited in the PDB. Unfortunately, the coordinates are not publically available yet, but this structure may reveal new information about the impact of phosphorylation on the *HsAQP5* structure.

4.5. Aquaporins in Chloroplasts

The chloroplast is the organelle in the plant cell where the light-driven carbon dioxide fixation takes place. In this process water is consumed in the thylakoid lumen as it donates electrons to the electron transport chain in the thylakoid membranes. Upon electron donation, the water molecule is split into oxygen and protons, the latter which contributes to the electrochemical proton gradient that is required for synthesis of adenosine 5'-triphosphate (ATP). A rapid flux of water from the cytosol and into the thylakoid lumen is a prerequisite for photosynthetic reactions to occur and theoretical calculations have suggested a need for an assisted water transport to meet these requirements (Per Åke Albertsson and Per Kjellbom, unpublished data). We speculated that aquaporins may be present in the chloroplast membranes and hence set out to identify these (**Paper V**).

Chloroplasts obtained either from *A. thaliana* or *Spinacia oleracea* plants, were fractionated into inner envelope-, thylakoid-, stroma lamellae- and grana membrane sub-fractions. In initial immunoblotting experiments, antibodies directed either towards the first of the, among aquaporins, highly conserved NPA motifs (section 2.1) or towards an epitope characteristic for isoforms of the PIP2 subgroup were used to detect the presence of potential AQPs. Strong signals indicated the occurrence of at least one PIP2 group member and additional aquaporin isoforms in several of the membrane fractions. To establish the identity of the indicated aquaporins, mass spectrometric analyses were performed on peptides extracted from bands cut out from SDS-PAGE gels. The immunoblots served as templates to localize the aquaporin bands. In thylakoid membranes or sub-fractions thereof, several PIP isoforms (*AtPIP1;2*, *AtPIP1;3*, *AtPIP2;1* and *AtPIP2;7*) were identified. By contamination control experiments we were able to confirm that the chloroplast membrane fractions were not contaminated by plasma membranes in which PIP isoforms are highly abundant and hence these hits were further validated. Two TIP isoforms (*AtTIP1;2* and *AtTIP2;1*) were also identified, but since contamination by vacuolar membranes was indicated, the presence of these isoforms in chloroplasts needs to be further confirmed. The working scheme employed to identify aquaporins in chloroplast is presented in Figure 4.4.

Whether the aquaporin isoforms identified in this study are crucial for sustaining photosynthesis remains to be elucidated. By up or down regulating the expression of these, phenotypic effects may reveal the biological significance of aquaporins in chloroplast membranes.

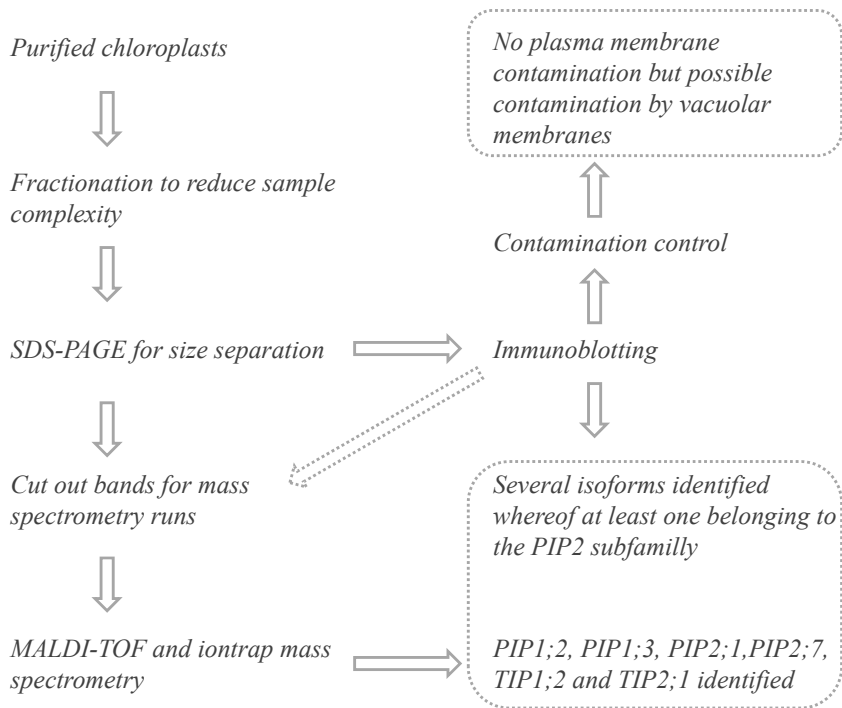


Figure 4.4. Outline of the strategy for identification of aquaporins in chloroplasts. By combining immunoblotting and mass spectrometric analyses several aquaporin isoforms were identified. By using antibodies directed towards markers of subcellular compartments it was also possible to monitor the degree of contamination of the chloroplast membrane fractions by vacuolar- and plasma membranes.

Populärvetenskaplig sammanfattning

Biokemi, livets kemi, är den gren av naturvetenskapen som avhandlar de molekyler och kemiska processer som är nödvändiga för allt levandes existens. Genom att karakterisera livets molekyler och deras funktion vill man öka förståelsen kring vad som krävs för att en organism ska kunna leva, växa, fortplanta sig och anpassa sig till nya miljöer och förhållanden. Som biokemist har man ofta till uppgift att visualisera det som inte med blotta ögat kan uppfattas, en tuff men spännande utmaning.

De biomolekyler vi valt att studera tillhör en klass av proteiner som kallas aquaporiner, och utgör kanaler genom de lipidrika höljen, eller membran, som omger en cell och som dessutom definierar intracellulära strukturer. Aquaporiner underlättar flödet av vatten (därav namnet) eller andra små och oladdade molekyler över cellens membran. Individer från denna proteinfamilj har återfunnits i alla livsformer, såväl hos djur och växter som hos bakterier och svampar. Vissa aquaporiner återfinns i stora kvantiteter i naturen, medan merparten produceras i relativt blygsam skala. Detta innebär ett problem då man vill studera dessa vattentransportörers biokemiska egenskaper. Vi har därför utvecklat ett protokoll för att, på genetisk väg, kunna modifiera jästsvampen *Pichia pastoris*, en släkting till den bagarjäst man kan köpa i livsmedelsbutiken, så att den kan producera aquaporiner och andra membranlokaliserade proteiner i stora mängder (**Paper I**). Fortsättningsvis har vi utarbetat optimerade metoder för att utvinna rena preparationer av dessa, artificiellt producerade proteiner genom att extrahera dem från jästcellerna och vidare separera dem från de proteiner som naturligt finns i *P. pastoris* (**Paper II, III och IV**). Då en stor mängd av det protein man önskar studera erhållits, möjliggörs biokemisk karakterisering av detta.

Två vattentransporterande kanaler, aquaporin 5 (AQP5) och aquaporin 8 (AQP8) från människa, har vi studerat närmare. Aquaporinerna, som producerats i *P. pastoris* och därefter extraherats och renats, sattes in i liposomer. Liposomer är artificiella lipidmembraner i sfärisk konformation, som grovt sett kan liknas vid såpbubblor i mikroformat. Insatt i ett sådant konstgjort membran kan vattenkanalerna studeras utan att andra proteiner och kemiska processer påverkar analyserna. Genom att undersöka AQP5 och AQP8 insatta i liposomer kunde vi se att det protein vi producerat var funktionellt och hade en förmåga att transportera

vatten. Då vi fastställt funktionaliteten, övergick vi till att försöka bestämma den tredimensionella strukturen hos dessa båda aquaporiner (**Paper III** och **IV**).

Eftersom en aquaporin är mycket liten med en ungefärlig diameter på 6 nm (sedd från cellens utsida) krävs raffinerade metoder för att visualisera dess uppbyggnad i detalj. Först måste proteinerna packas i ordnade kristaller, likt kolatomerna i en diamant eller vattenmolekylerna i en isbit. Genom att skjuta partiklar genom kristallerna och detektera hur dessa sprids då de studsar mot proteinets atomer som sitter i ett ordnat gitter, erhåller man ett så kallat diffraktionsmönster. Med hjälp av en rad avancerade matematiska beräkningar kan man sedan omvandla diffraktionsmönstret till en mängd datapunkter som beskriver atomernas position i proteinstrukturen. Vi har lyckats skapa ordnade kristaller av AQP5 och AQP8. Kristallerna av AQP5 visade sig dessutom vara av så god kvalitet att de resulterade i ett diffraktionsmönster som kunde översättas till en högupplöst proteinstruktur, det vill säga en avbildning av proteinet där aminosyrornas sidokedjor och de vattenmolekyler som befinner sig i kanalen, kan urskiljas. Förutom att man, med hjälp av denna molekylära avbildning, bättre kan förstå funktionen hos AQP5, kan proteinstrukturer som denna dessutom användas till exempel vid datasimuleringar av proteinets interaktioner med mindre molekyler. Sådana simuleringar ger viktiga upplysningar om man till exempel vill designa en molekyl som blockerar den aquaporin-assisterade vattentransporten.

I ett annat delprojekt undersökte vi växternas solenergiomvandlingsfabrik; kloroplasten, den del av växtcellen där klorofyllet, som ger växter deras gröna färg, återfinns. Eftersom vatten konsumeras vid fotosyntesen, då koldioxid omvandlas till energirika sockermolekyler och syre, spekulerade vi att en aquaporin assisterad vattentransport över kloroplastens membran kunde vara motiverad. Genom att tillämpa biokemiska metoder kunde vi fastställa förekomsten av vattenkanaler i kloroplastens olika membran (**Paper VI**). Om förekomsten av dessa är kritisk för effektiviteten hos fotosyntesen återstår att utröna, men om så är fallet kan dessa aquaporiner utgöra ett potentiellt mål för genetisk modifiering med syfte att skapa mer snabbväxande grödor.

Acknowledgements

There are a number of people that have been very important to me during the last six years when I have worked to accomplish this thesis and that deserve to be acknowledged.

First of all I would like to express my sincere gratitude to my supervisors, **Urban Johanson** who with great patience and consideration has taught me to work as an independent researcher and **Per Kjellbom** who has contributed with a lot of invaluable scientific advice and insight into how one can commercialize scientific ideas.

All past and present members of the AQP/TRP/RNA group: **Maria** for picking me up in the course lab in the first place and for being my mentor in the lab as well as a very dear friend also outside the lab. The “old guys”: **Erik, Sara, Sofia** and **Mia** for looking after me during my early time in the lab. **Jonas** and **Hanna** my best “next door neighbors”, who without ever complaining have helped me out with statistics as well as computer- and qPCR issues. **Anders**, for providing moral support when best needed. **Andreas**, for interesting discussions about crystallization and stopped flow assessments. **Sabeen** and **Lavanya** for bringing some eastern spirit and new proteins and methods into the lab. **Anna** for keeping me company in the “bås” for a while. **Henry**, for teaching me how to be a better supervisor during your masters project and for all your efforts put into the XIP project. **Angelica**, for being a source of inspiration with all your crazy projects. Former project students; **Zandra, Anna Maija** and **Lars** for all your hard work. **Michael** and **Jan** for providing scientific experience. And of course our invaluable technician **Adine**, for helping out with equipment, preparations and many other things.

At the CMPS we have a very inspirational and creative working environment and I would like to acknowledge everyone within this consortium for contributing to that. Especially I would like to thank **Katja** and **Wietske** for introducing me to the basics of mass spectrometry and for precious friendship. **Sinan, Ravi, Cecilia, Sven** and **Per-Åke** for the chloroplast collaboration. **Thom**, for input on liposomes and **Anna** for the informal stress management courses. **Maivi, Susanna** and **Gert** for keeping track of papers, computers and lab equipment.

For the fruitful collaboration resulting in the *HsAQP5* structure I would like to thank **Rob Horsefield**, **Anna Backmark**, **Susanna Törnroth-Horsefield** and Professor **Richard Neutze** at the University of Gothenburg.

Tack till mina vänner i och utanför Lund för att ni finns där för mej. Ett speciellt tack till **Anna** för att du alltid har kloka ord och uppmuntran till hands när jag som bäst behöver det.

Tack till min stora familj i Östergötland, Halland and Skåne för att ni alltid har en sommarfest, en god middag eller ett kräftfiske på lut när energipåfyllning behövs som mest. Speciellt tack till min fantastiska farmor **Fia**, för att du och **farfar** alltid haft dörren öppen för mej.

Mamma, Pappa, Evelynn, Rickard, Karolina, Roine, Marie och **Anton** - Tack för att ni tror på mig och för att ni alltid ställer upp och hjälper oss med stort och smått.

Ellen och **Love**, tack för att ni hjälper mig att se det stora i det lilla och för att ni påminner mig om att det bara är fantasin som sätter gränserna.

David, tack för att du är den omtänksamma och generösa person du är.

References

- Agmon, N. (1995). "The Grotthuss Mechanism." Chemical Physics Letters **244**(5-6): 456-462.
- Alexandersson, E., *et al.* (2010). "Transcriptional regulation of aquaporins in accessions of *Arabidopsis* in response to drought stress." Plant J **61**(4): 650-660.
- Alexandersson, E., *et al.* (2005). "Whole gene family expression and drought stress regulation of aquaporins." Plant Mol Biol **59**(3): 469-484.
- Alexandersson, E., *et al.* (2004). "Arabidopsis plasma membrane proteomics identifies components of transport, signal transduction and membrane trafficking." Plant Cell Physiol **45**(11): 1543-1556.
- Aller, S. G., *et al.* (2009). "Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding." Science **323**(5922): 1718-1722.
- Arnold, T. and D. Linke (2008). "The Use of Detergents to Purify Membrane Proteins." Current Protocols in Protein Science.
- Backmark, A., *et al.* (2008). "Affinity tags can reduce merohedral twinning of membrane protein crystals." Acta Crystallogr D Biol Crystallogr **64**(Pt 11): 1183-1186.
- Bawa, Z., *et al.* (2011). "Understanding the yeast host cell response to recombinant membrane protein production." Biochem Soc Trans **39**(3): 719-723.
- Beitz, E., *et al.* (2006). "Point mutations in the aromatic/arginine region in aquaporin 1 allow passage of urea, glycerol, ammonia, and protons." Proc Natl Acad Sci U S A **103**(2): 269-274.
- Berry, V., *et al.* (2000). "Missense mutations in MIP underlie autosomal dominant 'polymorphic' and lamellar cataracts linked to 12q." Nat Genet **25**(1): 15-17.
- Biela, A., *et al.* (1999). "The *Nicotiana tabacum* plasma membrane aquaporin NtAQP1 is mercury-insensitive and permeable for glycerol." Plant J **18**(5): 565-570.
- Bienert, G. P., *et al.* (2011). "Solanaceae XIPs are plasma membrane aquaporins that facilitate the transport of many uncharged substrates." Plant J **66**(2): 306-317.
- Bienert, G. P., *et al.* (2007). "Specific aquaporins facilitate the diffusion of hydrogen peroxide across membranes." J Biol Chem **282**(2): 1183-1192.
- Bienert, G. P., *et al.* (2008). "A subgroup of plant aquaporins facilitate the bi-directional diffusion of As(OH)(3) and Sb(OH)(3) across membranes." Bmc Biology **6**.
- Bill, R. M., *et al.* (2011). "Overcoming barriers to membrane protein structure determination." Nat Biotechnol **29**(4): 335-340.
- Bonander, N., *et al.* (2005). "Design of improved membrane protein production experiments: quantitation of the host response." Protein Sci **14**(7): 1729-1740.
- Borgnia, M., *et al.* (1999). "Cellular and molecular biology of the aquaporin water channels." Annu Rev Biochem **68**: 425-458.

- Bornert, O., *et al.* (2012). "Overexpression of membrane proteins using *Pichia pastoris*." Curr Protoc Protein Sci **Chapter 29**: Unit 29 22.
- Boursiac, Y., *et al.* (2005). "Early effects of salinity on water transport in *Arabidopsis* roots. Molecular and cellular features of aquaporin expression." Plant Physiology **139**(2): 790-805.
- Boursiac, Y., *et al.* (2008). "The response of *Arabidopsis* root water transport to a challenging environment implicates reactive oxygen species- and phosphorylation-dependent internalization of aquaporins." Plant Signal Behav **3**(12): 1096-1098.
- Cai, J. and P. Gros (2003). "Overexpression, purification, and functional characterization of ATP-binding cassette transporters in the yeast, *Pichia pastoris*." Biochim Biophys Acta **1610**(1): 63-76.
- Calamita, G., *et al.* (2005). "The inner mitochondrial membrane has aquaporin-8 water channels and is highly permeable to water." J Biol Chem **280**(17): 17149-17153.
- Carbrey, J. M., *et al.* (2003). "Aquaglyceroporin AQP9: solute permeation and metabolic control of expression in liver." Proc Natl Acad Sci U S A **100**(5): 2945-2950.
- Chakrabarti, N., *et al.* (2004). "Molecular basis of proton blockage in aquaporins." Structure **12**(1): 65-74.
- Chaumont, F., *et al.* (2000). "Plasma membrane intrinsic proteins from maize cluster in two sequence subgroups with differential aquaporin activity." Plant Physiol **122**(4): 1025-1034.
- Chaumont, F., *et al.* (2001). "Aquaporins constitute a large and highly divergent protein family in maize." Plant Physiol **125**(3): 1206-1215.
- Chawla, A. and M. Niwa (2005). "The unfolded protein response." Curr Biol **15**(22): R907.
- Chiba, Y., *et al.* (2009). "HvLsi1 is a silicon influx transporter in barley." Plant J **57**(5): 810-818.
- Cregg, J. M., *et al.* (2009). "Expression in the yeast *Pichia pastoris*." Methods Enzymol **463**: 169-189.
- Daly, R. and M. T. Hearn (2005). "Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production." J Mol Recognit **18**(2): 119-138.
- Daniels, M. J., *et al.* (2006). "In vivo functional assay of a recombinant aquaporin in *Pichia pastoris*." Appl Environ Microbiol **72**(2): 1507-1514.
- Danielson, J. A. and U. Johanson (2008). "Unexpected complexity of the Aquaporin gene family in the moss *Physcomitrella patens*." BMC Plant Biol **8**(1): 45.
- Danielson, J. A. and U. Johanson (2010). "Phylogeny of major intrinsic proteins." Adv Exp Med Biol **679**: 19-31.

- de Groot, B. L. and H. Grubmuller (2005). "The dynamics and energetics of water permeation and proton exclusion in aquaporins." Curr Opin Struct Biol **15**(2): 176-183.
- Deen, P. M., *et al.* (1994). "Requirement of human renal water channel aquaporin-2 for vasopressin-dependent concentration of urine." Science **264**(5155): 92-95.
- Degrip, W. J., *et al.* (1998). "Selective detergent-extraction from mixed detergent/lipid/protein micelles, using cyclodextrin inclusion compounds: a novel generic approach for the preparation of proteoliposomes." Biochem J **330** (Pt 2): 667-674.
- Denker, B. M., *et al.* (1988). "Identification, purification, and partial characterization of a novel Mr 28,000 integral membrane protein from erythrocytes and renal tubules." J Biol Chem **263**(30): 15634-15642.
- Doring, F., *et al.* (1997). "Expression and functional characterization of the mammalian intestinal peptide transporter PepT1 in the methylotropic yeast *Pichia pastoris*." Biochem Biophys Res Commun **232**(3): 656-662.
- Drew, D., *et al.* (2006). "Optimization of membrane protein overexpression and purification using GFP fusions." Nat Methods **3**(4): 303-313.
- Drew, D., *et al.* (2005). "A scalable, GFP-based pipeline for membrane protein overexpression screening and purification." Protein Sci **14**(8): 2011-2017.
- Dynowski, M., *et al.* (2008). "Plant plasma membrane water channels conduct the signalling molecule H₂O₂." Biochem J **414**(1): 53-61.
- Engel, A., *et al.* (2008). "Junction-forming aquaporins." Curr Opin Struct Biol **18**(2): 229-235.
- Fang, X., *et al.* (2002). "Evidence against aquaporin-1-dependent CO₂ permeability in lung and kidney." J Physiol **542**(Pt 1): 63-69.
- Ferri, D., *et al.* (2003). "Ontogeny, distribution, and possible functional implications of an unusual aquaporin, AQP8, in mouse liver." Hepatology **38**(4): 947-957.
- Fetter, K., *et al.* (2004). "Interactions between plasma membrane aquaporins modulate their water channel activity." Plant Cell **16**(1): 215-228.
- Fischer, G., *et al.* (2009). "Crystal structure of a yeast aquaporin at 1.15 angstrom reveals a novel gating mechanism." PLoS Biol **7**(6): e1000130.
- Fotiadis, D., *et al.* (2001). "Structural characterization of two aquaporins isolated from native spinach leaf plasma membranes." J Biol Chem **276**(3): 1707-1714.
- Fox, R. I. (2005). "Sjogren's syndrome." Lancet **366**(9482): 321-331.
- Freigassner, M., *et al.* (2009). "Tuning microbial hosts for membrane protein production." Microb Cell Fact **8**: 69.
- Fu, D., *et al.* (2000). "Structure of a glycerol-conducting channel and the basis for its selectivity." Science **290**(5491): 481-486.

- Fushimi, K., *et al.* (1997). "Phosphorylation of serine 256 is required for cAMP-dependent regulatory exocytosis of the aquaporin-2 water channel." J Biol Chem **272**(23): 14800-14804.
- Gattolin, S., *et al.* (2010). "Tonoplast intrinsic proteins and vacuolar identity." Biochem Soc Trans **38**(3): 769-773.
- Gattolin, S., *et al.* (2011). "Mapping of tonoplast intrinsic proteins in maturing and germinating Arabidopsis seeds reveals dual localization of embryonic TIPs to the tonoplast and plasma membrane." Mol Plant **4**(1): 180-189.
- Gonen, T., *et al.* (2004). "Aquaporin-0 membrane junctions form upon proteolytic cleavage." J Mol Biol **342**(4): 1337-1345.
- Gonen, T., *et al.* (2005). "Lipid-protein interactions in double-layered two-dimensional AQP0 crystals." Nature **438**(7068): 633-638.
- Gonen, T., *et al.* (2004). "Aquaporin-0 membrane junctions reveal the structure of a closed water pore." Nature **429**(6988): 193-197.
- Gorin, M. B., *et al.* (1984). "The major intrinsic protein (MIP) of the bovine lens fiber membrane: characterization and structure based on cDNA cloning." Cell **39**(1): 49-59.
- Griffith, D. A., *et al.* (2003). "A novel yeast expression system for the overproduction of quality-controlled membrane proteins." FEBS Lett **553**(1-2): 45-50.
- Gustavsson, S., *et al.* (2005). "A novel plant major intrinsic protein in *Physcomitrella patens* most similar to bacterial glycerol channels." Plant Physiol **139**(1): 287-295.
- Hamann, S., *et al.* (1998). "Aquaporins in complex tissues: distribution of aquaporins 1-5 in human and rat eye." Am J Physiol **274**(5 Pt 1): C1332-1345.
- Hara, M. and A. S. Verkman (2003). "Glycerol replacement corrects defective skin hydration, elasticity, and barrier function in aquaporin-3-deficient mice." Proc Natl Acad Sci U S A **100**(12): 7360-7365.
- Hasler, L., *et al.* (1998). "Purified lens major intrinsic protein (MIP) forms highly ordered tetragonal two-dimensional arrays by reconstitution." J Mol Biol **279**(4): 855-864.
- Hedfalk, K., *et al.* (2008). "Production, characterization and crystallization of the *Plasmodium falciparum* aquaporin." Protein Expr Purif **59**(1): 69-78.
- Henderson, R. and P. N. Unwin (1975). "Three-dimensional model of purple membrane obtained by electron microscopy." Nature **257**(5521): 28-32.
- Higuchi, T., *et al.* (1998). "Molecular cloning, water channel activity and tissue specific expression of two isoforms of radish vacuolar aquaporin." Plant Cell Physiol **39**(9): 905-913.
- Hite, R. K., *et al.* (2010). "Principles of membrane protein interactions with annular lipids deduced from aquaporin-0 2D crystals." EMBO J **29**(10): 1652-1658.

- Holmes, W. J., *et al.* (2009). "Developing a scalable model of recombinant protein yield from *Pichia pastoris*: the influence of culture conditions, biomass and induction regime." Microb Cell Fact **8**: 35.
- Hooijmaijers, C., *et al.* (2012). "Hydrogen peroxide permeability of plasma membrane aquaporins of *Arabidopsis thaliana*." J Plant Res **125**(1): 147-153.
- Hori, T., *et al.* (2010). "Expression, purification and characterization of leukotriene B(4) receptor, BLT1 in *Pichia pastoris*." Protein Expr Purif.
- Hub, J. S. and B. L. de Groot (2006). "Does CO₂ permeate through aquaporin-1?" Biophys J **91**(3): 842-848.
- Hub, J. S. and B. L. de Groot (2008). "Mechanism of selectivity in aquaporins and aquaglyceroporins." Proc Natl Acad Sci U S A.
- Hunte, C. and H. Michel (2002). "Crystallisation of membrane proteins mediated by antibody fragments." Curr Opin Struct Biol **12**(4): 503-508.
- Ikeda, M., *et al.* (2002). "Characterization of aquaporin-6 as a nitrate channel in mammalian cells. Requirement of pore-lining residue threonine 63." J Biol Chem **277**(42): 39873-39879.
- Invitrogen (2005). "EasySelect *Pichia* Expression Kit, version G. Accessed from the website: http://tools.invitrogen.com/content/sfs/manuals/easyselect_man.pdf."
- Isayenkov, S. V. and F. J. Maathuis (2008). "The *Arabidopsis thaliana* aquaglyceroporin AtNIP7;1 is a pathway for arsenite uptake." FEBS Lett **582**(11): 1625-1628.
- Ishibashi, K. (2006). "Aquaporin subfamily with unusual NPA boxes." Biochim Biophys Acta **1758**(8): 989-993.
- Ishibashi, K., *et al.* (2009). "Aquaporin water channels in mammals." Clin Exp Nephrol **13**(2): 107-117.
- Ishibashi, K., *et al.* (1997). "Cloning and functional expression of a new water channel abundantly expressed in the testis permeable to water, glycerol, and urea." J Biol Chem **272**(33): 20782-20786.
- Ishibashi, K., *et al.* (1998). "Cloning and functional expression of a new aquaporin (AQP9) abundantly expressed in the peripheral leukocytes permeable to water and urea, but not to glycerol." Biochem Biophys Res Commun **244**(1): 268-274.
- Ishibashi, K., *et al.* (1997). "Cloning and functional expression of a second new aquaporin abundantly expressed in testis." Biochem Biophys Res Commun **237**(3): 714-718.
- Ishibashi, K., *et al.* (2002). "Cloning and identification of a new member of water channel (AQP10) as an aquaglyceroporin." Biochim Biophys Acta **1576**(3): 335-340.
- Ishibashi, K., *et al.* (1994). "Molecular cloning and expression of a member of the aquaporin family with permeability to glycerol and urea in addition to water expressed at the basolateral membrane of kidney collecting duct cells." Proc Natl Acad Sci U S A **91**(14): 6269-6273.

- Ishikawa, F., *et al.* (2005). "Novel type aquaporin SIPs are mainly localized to the ER membrane and show cell-specific expression in *Arabidopsis thaliana*." FEBS Lett **579**(25): 5814-5820.
- Ishikawa, Y., *et al.* (2005). "Identification of AQP5 in lipid rafts and its translocation to apical membranes by activation of M3 mAChRs in interlobular ducts of rat parotid gland." Am J Physiol Cell Physiol **289**(5): C1303-1311.
- Itoh, T., *et al.* (2005). "Identification of a novel aquaporin, AQP12, expressed in pancreatic acinar cells." Biochem Biophys Res Commun **330**(3): 832-838.
- Jahn, T. P., *et al.* (2004). "Aquaporin homologues in plants and mammals transport ammonia." FEBS Lett **574**(1-3): 31-36.
- Jefferson, R. A., *et al.* (1987). "GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants." EMBO J **6**(13): 3901-3907.
- Jelen, S., *et al.* (2011). "Aquaporin-9 protein is the primary route of hepatocyte glycerol uptake for glycerol gluconeogenesis in mice." J Biol Chem **286**(52): 44319-44325.
- Johanson, U. and S. Gustavsson (2002). "A new subfamily of major intrinsic proteins in plants." Mol Biol Evol **19**(4): 456-461.
- Johanson, U., *et al.* (2001). "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 **126**(4): 1358-1369.
- Johansson, I., *et al.* (1998). "Water transport activity of the plasma membrane aquaporin PM28A is regulated by phosphorylation." Plant Cell **10**(3): 451-459.
- Johansson, I., *et al.* (1996). "The major integral proteins of spinach leaf plasma membranes are putative aquaporins and are phosphorylated in response to Ca²⁺ and apoplastic water potential." Plant Cell **8**(7): 1181-1191.
- Jung, J. S., *et al.* (1994). "Molecular structure of the water channel through aquaporin CHIP. The hourglass model." J Biol Chem **269**(20): 14648-14654.
- Junge, F., *et al.* (2008). "Large-scale production of functional membrane proteins." Cell Mol Life Sci **65**(11): 1729-1755.
- Kamiya, T., *et al.* (2009). "NIP1;1, an aquaporin homolog, determines the arsenite sensitivity of *Arabidopsis thaliana*." J Biol Chem **284**(4): 2114-2120.
- Kammerloher, W., *et al.* (1994). "Water channels in the plant plasma membrane cloned by immunoselection from a mammalian expression system." Plant J **6**(2): 187-199.
- Karlsson, M., *et al.* (2003). "Reconstitution of water channel function of an aquaporin overexpressed and purified from *Pichia pastoris*." FEBS Lett **537**(1-3): 68-72.
- Karlsson, M., *et al.* (2000). "An abundant TIP expressed in mature highly vacuolated cells." Plant J **21**(1): 83-90.
- King, L. S., *et al.* (2004). "From structure to disease: the evolving tale of aquaporin biology." Nat Rev Mol Cell Biol **5**(9): 687-698.

- Kondo, H., *et al.* (2002). "Human aquaporin adipose (AQPap) gene. Genomic structure, promoter analysis and functional mutation." *Eur J Biochem* **269**(7): 1814-1826.
- Kosugi-Tanaka, C., *et al.* (2006). "Protein kinase A-regulated membrane trafficking of a green fluorescent protein-aquaporin 5 chimera in MDCK cells." *Biochim Biophys Acta* **1763**(4): 337-344.
- Koyama, N., *et al.* (1998). "Cloning and functional expression of human aquaporin8 cDNA and analysis of its gene." *Genomics* **54**(1): 169-172.
- Koyama, Y., *et al.* (1997). "Molecular cloning of a new aquaporin from rat pancreas and liver." *J Biol Chem* **272**(48): 30329-30333.
- Laforenza, U. (2012). "Water channel proteins in the gastrointestinal tract." *Mol Aspects Med*.
- Lagane, B., *et al.* (2000). "Role of sterols in modulating the human mu-opioid receptor function in *Saccharomyces cerevisiae*." *J Biol Chem* **275**(43): 33197-33200.
- Laize, V., *et al.* (1999). "Molecular and functional study of AQY1 from *Saccharomyces cerevisiae*: Role of the C-terminal domain." *Biochemical and Biophysical Research Communications* **257**(1): 139-144.
- Lennon, V. A., *et al.* (2005). "IgG marker of optic-spinal multiple sclerosis binds to the aquaporin-4 water channel." *J Exp Med* **202**(4): 473-477.
- Liu, H., *et al.* (2004). "Quantitative comparison of placental expression of three aquaporin genes." *Placenta* **25**(6): 475-478.
- Liu, K., *et al.* (2006). "Purification and functional characterization of aquaporin-8." *Biol Cell* **98**(3): 153-161.
- Liu, L. H., *et al.* (2003). "Urea transport by nitrogen-regulated tonoplast intrinsic proteins in *Arabidopsis*." *Plant Physiol* **133**(3): 1220-1228.
- Liu, Z., *et al.* (2002). "Arsenite transport by mammalian aquaglyceroporins AQP7 and AQP9." *Proc Natl Acad Sci U S A* **99**(9): 6053-6058.
- Loque, D., *et al.* (2005). "Tonoplast intrinsic proteins AtTIP2;1 and AtTIP2;3 facilitate NH₃ transport into the vacuole." *Plant Physiol* **137**(2): 671-680.
- Luu, D. T., *et al.* (2012). "Fluorescence recovery after photobleaching reveals high cycling dynamics of plasma membrane aquaporins in *Arabidopsis* roots under salt stress." *Plant J* **69**(5): 894-905.
- Ma, J. F., *et al.* (2006). "A silicon transporter in rice." *Nature* **440**(7084): 688-691.
- Ma, T., *et al.* (1998). "Severely impaired urinary concentrating ability in transgenic mice lacking aquaporin-1 water channels." *J Biol Chem* **273**(8): 4296-4299.
- Maeshima, M. and F. Ishikawa (2008). "ER membrane aquaporins in plants." *Pflugers Arch* **456**(4): 709-716.
- Martinez Molina, D., *et al.* (2007). "Structural basis for synthesis of inflammatory mediators by human leukotriene C4 synthase." *Nature* **448**(7153): 613-616.

- Mathai, J. C. and M. L. Zeidel (2007). "Measurement of water and solute permeability by stopped-flow fluorimetry." Methods Mol Biol **400**: 323-332.
- Maurel, C., *et al.* (1997). "Purified vesicles of tobacco cell vacuolar and plasma membranes exhibit dramatically different water permeability and water channel activity." Proceedings of the National Academy of Sciences of the United States of America **94**(13): 7103-7108.
- Mitani, N., *et al.* (2009). "Identification of maize silicon influx transporters." Plant Cell Physiol **50**(1): 5-12.
- Mitsuoka, K., *et al.* (1999). "The structure of aquaporin-1 at 4.5-Å resolution reveals short alpha-helices in the center of the monomer." J Struct Biol **128**(1): 34-43.
- Mizutani, M., *et al.* (2006). "Aquaporin NIP2;1 is mainly localized to the ER membrane and shows root-specific accumulation in *Arabidopsis thaliana*." Plant Cell Physiol **47**(10): 1420-1426.
- Morishita, Y., *et al.* (2005). "Disruption of aquaporin-11 produces polycystic kidneys following vacuolization of the proximal tubule." Mol Cell Biol **25**(17): 7770-7779.
- Morishita, Y., *et al.* (2004). "Molecular mechanisms and drug development in aquaporin water channel diseases: aquaporin superfamily (superaquaporins): expansion of aquaporins restricted to multicellular organisms." J Pharmacol Sci **96**(3): 276-279.
- Mulders, S. M., *et al.* (1998). "An aquaporin-2 water channel mutant which causes autosomal dominant nephrogenic diabetes insipidus is retained in the Golgi complex." J Clin Invest **102**(1): 57-66.
- Murata, K., *et al.* (2000). "Structural determinants of water permeation through aquaporin-1." Nature **407**(6804): 599-605.
- Nedvetsky, P. I., *et al.* (2009). "Regulation of aquaporin-2 trafficking." Handb Exp Pharmacol(190): 133-157.
- Nielsen, S., *et al.* (1997). "Aquaporins in complex tissues. II. Subcellular distribution in respiratory and glandular tissues of rat." Am J Physiol **273**(5 Pt 1): C1549-1561.
- Norholm, M. H., *et al.* (2012). "Manipulating the genetic code for membrane protein production: What have we learnt so far?" Biochim Biophys Acta **1818**(4): 1091-1096.
- Nyblom, M., *et al.* (2007). "Exceptional overproduction of a functional human membrane protein." Protein Expr Purif **56**(1): 110-120.
- Opekarova, M. and W. Tanner (2003). "Specific lipid requirements of membrane proteins-a putative bottleneck in heterologous expression." Biochim Biophys Acta **1610**(1): 11-22.
- Otto, B., *et al.* (2010). "Aquaporin tetramer composition modifies the function of tobacco aquaporins." J Biol Chem **285**(41): 31253-31260.
- Pao, G. M., *et al.* (1991). "Evolution of the MIP family of integral membrane transport proteins." Mol Microbiol **5**(1): 33-37.

- Papadopoulos, M. C. and A. Verkman (2012). "Aquaporin 4 and neuromyelitis optica." Lancet Neurol **11**(6): 535-544.
- Park, W., *et al.* (2010). "Identification of the family of aquaporin genes and their expression in upland cotton (*Gossypium hirsutum* L.)." BMC Plant Biol **10**: 142.
- Pomes, R. and B. Roux (1996). "Structure and dynamics of a proton wire: a theoretical study of H⁺ translocation along the single-file water chain in the gramicidin A channel." Biophys J **71**(1): 19-39.
- Prasad, G. V. R., *et al.* (1998). "Reconstituted aquaporin 1 water channels transport CO₂ across membranes." Journal of Biological Chemistry **273**(50): 33123-33126.
- Preston, G. M. and P. Agre (1991). "Isolation of the cDNA for erythrocyte integral membrane protein of 28 kilodaltons: member of an ancient channel family." Proc Natl Acad Sci U S A **88**(24): 11110-11114.
- Preston, G. M., *et al.* (1992). "Appearance of water channels in *Xenopus* oocytes expressing red cell CHIP28 protein." Science **256**(5055): 385-387.
- Prive, G. G. (2007). "Detergents for the stabilization and crystallization of membrane proteins." Methods **41**(4): 388-397.
- Raina, S., *et al.* (1995). "Molecular cloning and characterization of an aquaporin cDNA from salivary, lacrimal, and respiratory tissues." J Biol Chem **270**(4): 1908-1912.
- Raman, P., *et al.* (2006). "The Membrane Protein Data Bank." Cell Mol Life Sci **63**(1): 36-51.
- Reizer, J., *et al.* (1993). "The MIP family of integral membrane channel proteins: sequence comparisons, evolutionary relationships, reconstructed pathway of evolution, and proposed functional differentiation of the two repeated halves of the proteins." Crit Rev Biochem Mol Biol **28**(3): 235-257.
- Rhodes, G. (2006). Crystallography made crystal clear. USA, Academic Press; Elsevier.
- Richey, D. P. and E. C. Lin (1972). "Importance of facilitated diffusion for effective utilization of glycerol by *Escherichia coli*." J Bacteriol **112**(2): 784-790.
- Rigaud, J., *et al.* (2000). "Use of detergents in two-dimensional crystallization of membrane proteins." Biochim Biophys Acta **1508**(1-2): 112-128.
- Rigaud, J. L. and D. Levy (2003). "Reconstitution of membrane proteins into liposomes." Methods Enzymol **372**: 65-86.
- Romanos, M., *et al.* (1998). "The generation of multicopy recombinant strains." Methods Mol Biol **103**: 55-72.
- Saadoun, S., *et al.* (2005). "Impairment of angiogenesis and cell migration by targeted aquaporin-1 gene disruption." Nature **434**(7034): 786-792.
- Sade, N., *et al.* (2009). "Improving plant stress tolerance and yield production: is the tonoplast aquaporin SITIP2;2 a key to isohydric to anisohydric conversion?" New Phytol **181**(3): 651-661.

- Sandal, N. N. and K. A. Marcker (1988). "Soybean nodulin 26 is homologous to the major intrinsic protein of the bovine lens fiber membrane." Nucleic Acids Res **16**(19): 9347.
- Saparov, S. M., *et al.* (2007). "Fast and selective ammonia transport by aquaporin-8." J Biol Chem **282**(8): 5296-5301.
- Sarramegna, V., *et al.* (2003). "Heterologous expression of G-protein-coupled receptors: comparison of expression systems from the standpoint of large-scale production and purification." Cell Mol Life Sci **60**(8): 1529-1546.
- Savage, D. F., *et al.* (2010). "Structural context shapes the aquaporin selectivity filter." Proc Natl Acad Sci U S A **107**(40): 17164-17169.
- Savage, D. F. and R. M. Stroud (2007). "Structural basis of aquaporin inhibition by mercury." J Mol Biol **368**(3): 607-617.
- Schaffner, A. R. (1998). "Aquaporin function, structure, and expression: are there more surprises to surface in water relations?" Planta **204**(2): 131-139.
- Schussler, M. D., *et al.* (2008). "The effects of the loss of TIP1;1 and TIP1;2 aquaporins in *Arabidopsis thaliana*." Plant J **56**(5): 756-767.
- Scorer, C. A., *et al.* (1993). "The intracellular production and secretion of HIV-1 envelope protein in the methylotrophic yeast *Pichia pastoris*." Gene **136**(1-2): 111-119.
- Seddon, A. M., *et al.* (2004). "Membrane proteins, lipids and detergents: not just a soap opera." Biochim Biophys Acta **1666**(1-2): 105-117.
- Shiels, A. and S. Bassnett (1996). "Mutations in the founder of the MIP gene family underlie cataract development in the mouse." Nat Genet **12**(2): 212-215.
- Smith, B. L. and P. Agre (1991). "Erythrocyte Mr 28,000 transmembrane protein exists as a multisubunit oligomer similar to channel proteins." J Biol Chem **266**(10): 6407-6415.
- Soria, L. R., *et al.* (2010). "Aquaporin-8-facilitated mitochondrial ammonia transport." Biochem Biophys Res Commun **393**(2): 217-221.
- Sreekrishna, K., *et al.* (1997). "Strategies for optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast *Pichia pastoris*." Gene **190**(1): 55-62.
- Sui, H., *et al.* (2001). "Structural basis of water-specific transport through the AQP1 water channel." Nature **414**(6866): 872-878.
- Takano, J., *et al.* (2006). "The *Arabidopsis* major intrinsic protein NIP5;1 is essential for efficient boron uptake and plant development under boron limitation." Plant Cell **18**(6): 1498-1509.
- Talmont, F., *et al.* (1996). "Expression and pharmacological characterization of the human mu-opioid receptor in the methylotrophic yeast *Pichia pastoris*." FEBS Lett **394**(3): 268-272.
- Tanaka, M., *et al.* (2008). "NIP6;1 is a boric acid channel for preferential transport of boron to growing shoot tissues in *Arabidopsis*." Plant Cell **20**(10): 2860-2875.

- Tani, K., *et al.* (2009). "Mechanism of aquaporin-4's fast and highly selective water conduction and proton exclusion." J Mol Biol **389**(4): 694-706.
- Tornroth-Horsefield, S., *et al.* (2006). "Structural mechanism of plant aquaporin gating." Nature **439**(7077): 688-694.
- Tournaire-Roux, C., *et al.* (2003). "Cytosolic pH regulates root water transport during anoxic stress through gating of aquaporins." Nature **425**(6956): 393-397.
- Tsubota, K., *et al.* (2001). "Defective cellular trafficking of lacrimal gland aquaporin-5 in Sjogren's syndrome." Lancet **357**(9257): 688-689.
- Tsukaguchi, H., *et al.* (1998). "Molecular characterization of a broad selectivity neutral solute channel." J Biol Chem **273**(38): 24737-24743.
- Ubarretxena-Belandia, I. and D. L. Stokes (2012). "Membrane protein structure determination by electron crystallography." Curr Opin Struct Biol.
- Uehlein, N., *et al.* (2003). "The tobacco aquaporin NtAQP1 is a membrane CO₂ pore with physiological functions." Nature **425**(6959): 734-737.
- Uehlein, N., *et al.* (2008). "Function of Nicotiana tabacum aquaporins as chloroplast gas pores challenges the concept of membrane CO₂ permeability." Plant Cell **20**(3): 648-657.
- Uehlein, N., *et al.* (2011). "The Arabidopsis aquaporin PIP1;2 rules cellular CO₂ uptake." Plant Cell Environ.
- Wallace, I. S. and D. M. Roberts (2004). "Homology modeling of representative subfamilies of Arabidopsis major intrinsic proteins. Classification based on the aromatic/arginine selectivity filter." Plant Physiol **135**(2): 1059-1068.
- Wallace, I. S. and D. M. Roberts (2005). "Distinct transport selectivity of two structural subclasses of the nodulin-like intrinsic protein family of plant aquaglyceroporin channels." Biochemistry **44**(51): 16826-16834.
- Wallin, E. and G. von Heijne (1998). "Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms." Protein Sci **7**(4): 1029-1038.
- Walz, T., *et al.* (1997). "The three-dimensional structure of aquaporin-1." Nature **387**(6633): 624-627.
- Walz, T., *et al.* (1994). "The three-dimensional structure of human erythrocyte aquaporin CHIP." Embo J **13**(13): 2985-2993.
- Wang, Y., *et al.* (2007). "Exploring gas permeability of cellular membranes and membrane channels with molecular dynamics." J Struct Biol **157**(3): 534-544.
- Weaver, C. D., *et al.* (1991). "Calcium-dependent phosphorylation of symbiosome membrane proteins from nitrogen-fixing soybean nodules : evidence for phosphorylation of nodulin-26." Plant Physiol **95**(1): 222-227.
- Verkman, A. S. (2011). "Aquaporins at a glance." J Cell Sci **124**(Pt 13): 2107-2112.
- Yakata, K., *et al.* (2006). "Aquaporin-11 containing a divergent NPA motif has normal water channel activity." Biochim Biophys Acta.

- Yang, B., *et al.* (2000). "Carbon dioxide permeability of aquaporin-1 measured in erythrocytes and lung of aquaporin-1 null mice and in reconstituted proteoliposomes." J Biol Chem **275**(4): 2686-2692.
- Yang, B., *et al.* (2005). "Phenotype analysis of aquaporin-8 null mice." Am J Physiol Cell Physiol **288**(5): C1161-1170.
- Yang, B., *et al.* (2006). "Evidence against functionally significant aquaporin expression in mitochondria." J Biol Chem **281**(24): 16202-16206.
- Yang, F., *et al.* (2003). "Cyclic AMP regulates aquaporin 5 expression at both transcriptional and post-transcriptional levels through a protein kinase A pathway." J Biol Chem **278**(34): 32173-32180.
- Yasui, M., *et al.* (1999). "Rapid gating and anion permeability of an intracellular aquaporin." Nature **402**(6758): 184-187.
- Yeung, C. H., *et al.* (2010). "Aquaporins in the human testis and spermatozoa - identification, involvement in sperm volume regulation and clinical relevance." Int J Androl **33**(4): 629-641.
- Yildirim, M. A., *et al.* (2007). "Drug-target network." Nat Biotechnol **25**(10): 1119-1126.
- Zelazny, E., *et al.* (2007). "FRET imaging in living maize cells reveals that plasma membrane aquaporins interact to regulate their subcellular localization." Proc Natl Acad Sci U S A **104**(30): 12359-12364.
- Zelazny, E., *et al.* (2009). "An N-terminal diacidic motif is required for the trafficking of maize aquaporins ZmPIP2;4 and ZmPIP2;5 to the plasma membrane." Plant J **57**(2): 346-355.
- Zelenina, M., *et al.* (2003). "Nickel and extracellular acidification inhibit the water permeability of human aquaporin-3 in lung epithelial cells." J Biol Chem **278**(32): 30037-30043.
- Zeuthen, T. and D. A. Klaerke (1999). "Transport of water and glycerol in aquaporin 3 is gated by H(+)." J Biol Chem **274**(31): 21631-21636.
- Zhang, D., *et al.* (2002). "Aquaporin deletion in mice reduces intraocular pressure and aqueous fluid production." J Gen Physiol **119**(6): 561-569.