THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN NATURAL SCIENCE

Human aquaporins:
Production, Characterization and Interactions

JENNIE SJÖHAMN

University of Gothenburg
Department of Chemistry and Molecular Biology
Göteborg, Sweden, 2015

Thesis for the Degree of Doctor of Philosophy in Natural Science

Human aquaporins: Production, Characterization and Interactions

Jennie Sjöhamn

Cover: Model of the hAQP0-calmodulin complex (PDB ID: 3J41)

ISBN 978-91-628-9573-0 (Print)
ISBN 978-91-628-9574-7 (PDF)
Available online at: http://hdl.handle.net/2077/40589

Department of Chemistry and Molecular Biology
Biochemistry and Biophysics
University of Gothenburg
SE-413 90 Göteborg, Sweden

Printed by Ale tryckteam AB
Göteborg, Sweden 2015

Till min familj

It's not the destination,
It's the journey

Abstract

Membrane proteins are essential components of the cell and responsible for the communication with the outside environment and transport of molecules across the membrane. Water transport is facilitated by aquaporins, which are water selective transmembrane pores that serve to maintain cell homeostasis. Aquaporins have been well characterized in terms of structure and function and a broad variety of cell based assays have given insight into their mechanism of regulation, including membrane localization and conformational changes. High-resolution structural information on aquaporins has emerged mainly using X-ray crystallography, which require large quantities of pure and homogenous protein.

This thesis presents the importance of codon optimization and clone selection in the first step of the pipeline to obtain high yields of recombinant membrane protein. This, in turn, enables biochemical characterization of the protein of interest. One such target, the human aquaporin 10, was found to be glycosylated in *P. pastoris*, increasing the protein stability *in vitro* but without any measurable impact on function.

Aquaporin function is regulated by both physiological signals and interactions with other proteins. The regulation of the plasma membrane abundance of hAQP5 shows that three independent mechanisms – phosphorylation at Ser156, protein kinase A activity and extracellular tonicity – work in synergy to fine-tune the fraction of membrane localized protein. Furthermore, an overview of the literature of AQP protein:protein interactions reveal that the C-terminus is the most diverse sequence between aquaporins and that the majority of the known interactions map there.

Obtaining high-resolution structural information of protein:protein complexes is one of the future challenges in structural biology. We developed a novel method for the characterization and purification of membrane protein complexes using hAQP0 and calmodulin as the proof-of-principle interaction partners. Our approach combined bimolecular-fluorescence complementation to characterize the interaction and fluorescence detection to detect the complex throughout purification. This resulted in a versatile method to purify intact protein complexes in enough yields for crystallization, potentially facilitating future structural determination by X-ray crystallography or electron microscopy.

List of publications

Paper I Öberg F, **Sjöhamn J**, Conner M. T, Bill R. M, Hedfalk K. (2011) Improving recombinant eukaryotic membrane protein yields in *Pichia pastoris*: the importance of codon optimization and clone selection. *Mol. Membr. Biol*. 28(6) 398–411

Paper II Öberg F, **Sjöhamn J**, Fischer G, Moberg A, Pedersen A, Neutze R, Hedfalk K. (2011) Glycosylation increases the thermostability of human aquaporin 10 protein. *J. Biol. Chem.* 286(36), 31915–31923

Paper III Kitchen P, Öberg F, **Sjöhamn J**, Hedfalk K, Bill RM, Conner AC, Conner MT, Törnroth-Horsefield S. (2015) Plasma membrane abundance of human aquaporin 5 is dynamically regulated by multiple pathways. *Accepted for publication in PLOS ONE.*

Paper IV **Sjöhamn J**, Hedfalk K. (2014) Unraveling aquaporin interaction partners. *Biochim Biophys Acta 1840 (5), 1614-1623*

Paper V **Sjöhamn J**, Båth P, Neutze R, Hedfalk K. (2015) A strategy for expressing membrane protein:protein complexes for structural studies - bimolecular fluorescence complementation for membrane protein complex purification. *Submitted*

Other publications

Paper VI Johansson LC, Arnlund D, Katona G, White TA, Barty A, DePonte DP, Shoeman RL, Wickstrand C, Sharma A, Williams GJ, Aquila A, Bogan MJ, Caleman C, Davidsson J, Doak RB, Frank M, Fromme R, Galli L, Grotjohann I, Hunter MS, Kassemeyer S, Kirian RA, Kupitz C, Liang M, Lomb L, Malmerberg E, Martin AV, Messerschmidt M, Nass K, Redecke L, Seibert MM, **Sjöhamn J**, Steinbrener J, Stellato F, Wang D, Wahlgren WY, Weierstall U, Westenhoff S, Zatsepin NA, Boutet S, Spence JC, Schlichting I, Chapman HN, Fromme P, Neutze R. (2013) Structure of a photosynthetic reaction centre determined by serial femtosecond crystallography. *Nature Commun. 4, 2911*

Paper VII Arnlund D, Johansson LC, Wickstrand C, Barty A, Williams GJ, Malmerberg E, Davidsson J, Milathianaki D, DePonte DP, Shoeman RL, Wang D ,James D, Katona G, Westenhoff S, White TA, Aquila A, Bari S, Berntsen P, Bogan M, van Driel TB, Doak RB, Kjær KS, Frank M, Fromme R, Grotjohann I, Henning R, Hunter MS, Kirian RA, Kosheleva I, Kupitz C, Liang M, Martin AV, Nielsen MM, Messerschmidt M, Seibert MM, **Sjöhamn J**, Stellato F, Weierstall U, Zatsepin NA, Spence JC, Fromme P, Schlichting I, Boutet S, Groenhof G, Chapman HN, Neutze R. (2014) Visualizing a protein quake with time-resolved X-ray scattering at a free-electron laser. *Nature methods. 11(9), 923-926*

Contribution report

Paper I I performed the cloning, transformations of the chimeric and fusion constructs and quantitation of protein levels. I was involved in the interpretation of the results and took part in the writing of the manuscript.

Paper II I prepared the mutant version of hAQP10, purified the protein and carried out the functional characterization of the glycosylation. I took part in the interpretation of the results and the writing of the manuscript.

Paper III I cloned the mutant protein constructs and did part of the protein purifications.

Paper IV I did an extensive overview on the literature covering protein-protein interaction methodology and the known interaction partners of the AQPs included in the study. I was involved in drawing the conclusions on common features of aquaporin interactions.

Paper V I planned the project, designed primers and prepared the constructs, transformed cells and performed the microscopy studies and cell counting. I purified the complex, ran native gels, wrote the major part of the manuscript and prepared the figures.

Abbreviations

AOX	Alcohol oxidase
AQP	Aquaporin
ar/R	aromatic/arginine region, AQP restriction site
Å	Ångström (1Å = 0.1nm)
BiFC	Bimolecular fluorescence complementation
CAI	Codon adaptive index
CaM	Calmodulin
CD	Circular dichroism
EM	Electron microscopy
ER	Endoplasmic reticulum
(F)SEC	(Fluorescence) size exclusion chromatography
GFP	Green fluorescent protein
IMAC	Immobilized metal affinity chromatography
MD	Molecular dynamics
MR	Molecular replacement
NPA	Asn-Pro-Ala, signature motif of aquaporins
PCA	Protein complementation assay
PDB	Protein data bank
PKA	Protein kinase A
PPI	Protein:protein interaction
TM	Transmembrane
WT	Wild type
YFP	Yellow fluorescent protein
YTH	Yeast two-hybrid assay

E. coli	*Escherichia coli*
P. pastoris	*Pichia pastoris*
P. falciparum	*Plasmodium falciparum*
S. cerevisiae	*Saccharomyces cerevisiae*

All figures are reprinted with permission from the publishing companies as listed in the beginning of the reference list.

Table of contents

ABSTRACT **V**

1 INTRODUCTION **1**
1.1 Membrane proteins **1**
1.2 Aquaporins **2**
1.2.1 Human aquaporins *5*
1.2.2 Aquaporin 0 *6*
1.2.3 Aquaporin 5 *7*
1.2.4 Aquaporin 10 *8*
1.3 Protein:protein interactions **9**
1.4 Recombinant protein production **12**
1.4.1 Protein production systems *12*
1.4.2 Posttranslational modifications *13*

2 SCOPE OF THE THESIS **15**

3 PROTEIN PRODUCTION OPTIMIZATION **17**
3.1 Generating high yielding clones **18**
3.2 Clone selection **19**
3.3 Protein stability **20**

4 AQUAPORIN CHARACTERIZATION **25**
4.1 Thermostability **25**
4.2 Functional assays **27**
4.3 Trafficking and regulation **29**
4.4 Protein:protein interactions **33**
4.4.1 The AQP0-CaM interaction *35*
4.4.2 Purification of protein complexes *39*
4.5 Future directions for structure determination of protein complexes **41**

5 CONCLUDING REMARKS **43**

6 FUTURE PERSPECTIVES **45**

7 ACKNOWLEDGEMENTS **46**

8 REFERENCES **49**

1 Introduction

Several billion years ago life appeared on planet Earth. How it came about is a fundamental question that has been discussed for thousands of years. Irrespective of whether the process of replication or metabolism was the first step towards life as we know it today, most scientists agree that life first evolved in the oceans of the young Earth, making water an essential molecule for life.

The formation of lipid vesicles that came to enclose the replicating biomolecules created the first plasma membrane and defined an "inside" and an "outside". Text books often visualize the membranes as a lipid sea with membrane proteins floating around like icebergs. Even though the membrane is indeed comparable to a fluid in some respects the complexity of the membrane should not be underestimated. The protein content of membranes can vary between 20% and 80% and the presence of sterols can alter the fluidity and thickness of the membrane significantly [1,2]. In the human body the lipid composition of different tissues and intracellular compartments exhibit large specificity and lipids often influence protein function [3]. As we learn more about the membrane heterogeneity it becomes evident that the studies of membranes and membrane proteins are closely intertwined.

1.1 Membrane proteins

About one third of any given genome is constituted of genes encoding membrane proteins involved in cellular processes ranging from facilitated transports of nutrients and byproducts to cellular structure, environmental sensing, signaling and movement [4,5]. It is estimated that 50% of the drugs available on the market target membrane localized proteins and many disorders are associated with defects in membrane protein functions [6].

The close relationship between structure and function is explored by structural biology. Determining the protein structure is vital in understanding the function of the protein on a molecular level. Despite their biological and physiological relevance, the number of solved membrane protein structures severely lags behind their soluble counterparts constituting less than 1% of the known protein structures [7]. The big discrepancy is due to the difficulty of extracting membrane proteins from their native membrane environment and keeping the

protein functional and stable on the time scales of purification and crystallization. In addition, the natural abundance of membrane proteins is low compared to soluble proteins. Increased yields are obtained by recombinant protein production which has been the most common source of protein for structure determination over the past ten years [8].

1.2 Aquaporins

Water flows in and out of the cell and is facilitated by water channels in the membrane called aquaporins. The existence of aquaporins was postulated in the 1970s and there was much discussion on whether water transporting proteins were actually required for the cell to maintain homeostasis or if the passive diffusion across the plasma membrane was sufficient (reviewed in [9]). Certain cell types show elevated water permeability and treatment with mercurial compounds effectively decreased the water permeability to levels compared to pure lipid bilayers which indicated the presence of proteins with water transporting abilities [10]. In 1991, the Agre lab was investigating the rhesus blood group antigens and a contaminant at 28 kDa kept appearing in the SDS-PAGEs. They injected oocytes with the cRNA of the responsible gene and transferred them to a hypotonic buffer. This resulted in increased inflow of water causing swelling and finally rupturing of the oocyte [11]. This discovery of protein channels with water transporting abilities, the aquaporins, and their impact on physiology was awarded with the Nobel prize in 2003 [12].

Aquaporins adopt a common fold with six transmembrane helices connected by five loops with both the N- and C-termini on the cytosolic side of the membrane. The Asn-Pro-Asp (NPA) aquaporin signature motif is located in loops B and E which fold back into the membrane and create a seventh pseudo helix (**Figure 1A).** Four monomers come together in the physiological unit, the tetramer, where each pore acts as an independent water channel (**Figure 1B).** The central pore formed in the middle of the tetramer has been implicated in transport of both CO_2 and ions, but this is still an ongoing controversy in the aquaporin field [13–16].

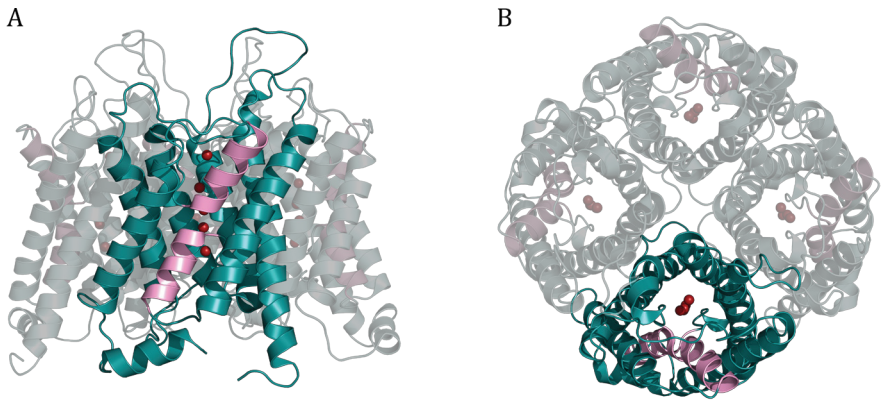


Figure 1. bAQP1 (PMID: 1J4N) show the general fold of aquaporins. A) *Front view of the tetramer showing the half helices (pink).* ***B)*** *Top view of the tetramer showing the water molecules (red) and the central pore.*

Aquaporins can transport water at rates close to the rate of diffusion since the pore is lined with hydrogen bond donating residues. They replace the hydrogen bonds between the water molecules as they pass through the channel, resulting in high water flow capacities and high specificity. A highly selective transport of water is crucial as co-transport of protons in particular would have severe consequences for the cell as it would disrupt the electrochemical proton gradient across the membrane. The proton motive force stored in a transmembrane proton gradient is used to synthesize the majority of the cells' ATP supply.

The available AQP structures together with complementary molecular dynamics (MD) simulations show how water molecules pass through the pore. The aquaporin maintains its specificity by combining channel restriction and electrostatic barriers (reviewed in [17]). The two half helices, where the NPA motif is located, act as two macro dipoles, introducing an electrostatic barrier that repulses positively charged ions and acts as a proton exclusion mechanism. The restriction site with the aromatic/arginine (ar/R) region is located closer to the extracellular side (**Figure 2**). A conserved arginine and aromatic residues narrows the pore to exclude larges solutes [18]. Mutations in the ar/R enable the passage of glycerol, urea and ammonia which suggest that this region is also important in the selectivity of AQPs but exactly how this is achieved is still not known [19].

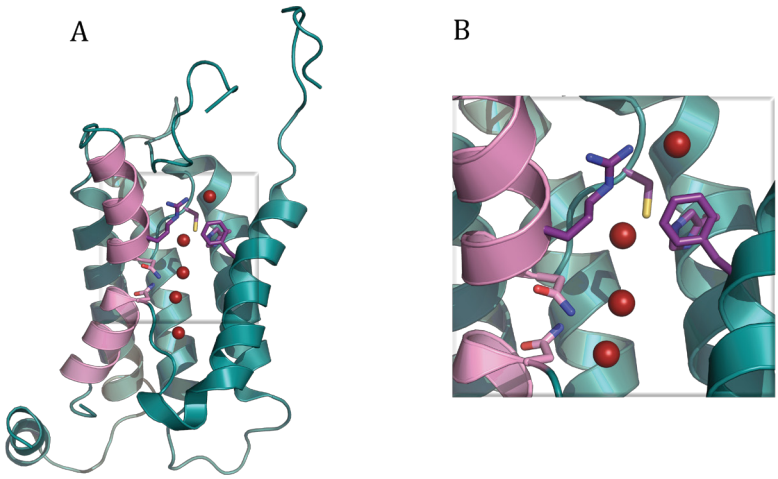


Figure 2. The selective transport of water is created by the NPA motif and the ar/R region. *The asparagines of the NPA motifs on the tips of the half helices acts as an electrostatic barrier. The ar/R region (purple) is part of the constriction site. Two helices have been removed for a better view.*

Proton exclusion requires a mechanism different from other solutes since the hydronium ions are so structurally similar to water. The Grotthuss mechanism predicted the presence of proton conducting networks in water almost 200 years ago [20]. In bulk water protons can "hop" from one water molecule to another, via hydrogen bonds and transient hydronium ions, which substantially increase their diffusion mobility compared to other ions [21]. Disrupting these networks is essential in preventing the protons from passing across the membrane through the water channels, keeping the electrochemical gradient across the membrane intact. Amazing detail of the water flow through the *Pichia pastoris* aquaporin AQY1 was captured at 1.15 Å [22]. Together with the more recent structure at 0.88 Å the movements and hydrogen bonds of the water molecules in the pore can be analyzed, giving insights into how the proton exclusion and selectivity work [23]. From these structures and the complementary MD simulations it seems like the water molecule movements are synchronized and that they move in a pairwise fashion through the pore. A rotation of the water molecules as they move allows the hydrogen bonding to be broken and would effectively prevent proton hopping by the Grotthuss mechanism [23,24].

1.2.1 Human aquaporins

13 aquaporin homologues have been identified in the human body (hAQP0-hAQP12). Historically, they have been divided into three different classes depending on their substrate specificity and sequence homology (**Table 1**). Orthodox aquaporins (hAQP1, hAQP2, hAQP4, hAQP5, hAQP6 and hAQP8) exclusively transport water while the aquaglyceroporins (hAQP3, hAQP7, hAQP9, hAQP10) also transport glycerol, urea and/or other small solutes [25]. The third class, the super aquaporins (hAQP11 and hAQP12), remains relatively unexplored and their specificity is thought to be different from the other aquaporins based on their lower sequence homology [26].

Table 1. Substrate specificity and tissue distribution of the 13 human aquaporins. *The table is prepared based on data from* [25,27–31].

Protein	Substrate specificity	Major tissue distribution
hAQP0	water	eye lens
hAQP1	water	erythrocytes, kidney, lung, brain, eye, vascular endothelium
hAQP2	water	kidney
hAQP3	water, glycerol, urea	skin, kidney, lung, eye
hAQP4	water	brain, kidney, lung, muscle, stomach
hAQP5	water	lungs, salivary-, lacrimal- and sweat glands, lung, eye
hAQP6	water, anions	kidney (intracellular)
hAQP7	water, glycerol, urea, arsenite	adipocytes, kidney, testis
hAQP8	water	kidney, liver, pancreas, intestine, colon, testis
hAQP9	water, glycerol, arsenite	liver, leukocytes, brain, testis
hAQP10	water, glycerol, urea	intestine, skin, adipocytes
hAQP11	water	brain, liver, kidney (intracellular)
hAQP12	water (?)	pancreas (intracellular)

Even though the human aquaporins is a homogenous protein family the recombinant overproduction in the yeast *P. pastoris* give rise to a large variation in production levels (**Figure 3**) [32]. Highly produced hAQPs are found in all three classes as are hAQPs produced at levels below detection. In this host, a higher fraction of the aquaglyceroporins are

inserted into the membrane compared to the orthodox aquaporins, but the determinants of high or low production levels are not obvious.

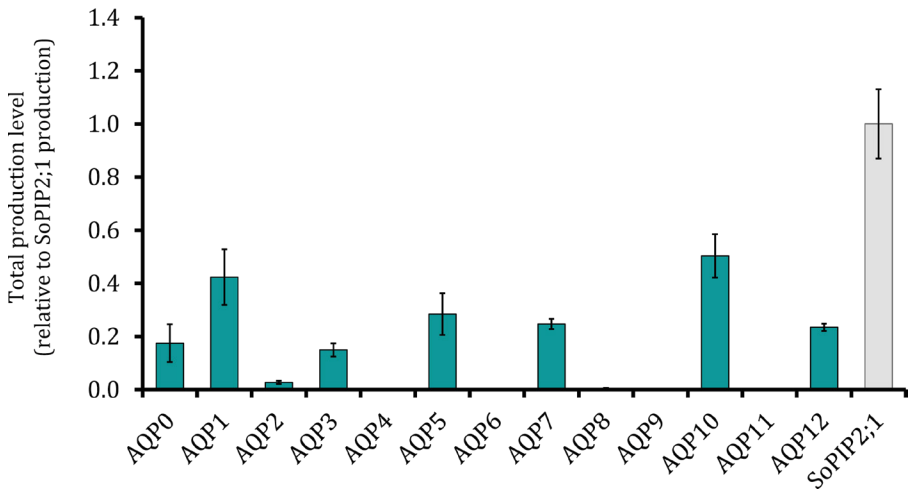


Figure 3. The aquaporin family show large variations in overproduction yields. *The total production levels are normalized to SoPIP2;1, a plant aquaporin used as in-house reference for production levels. The yields range from non-detectable to high production* [32].

This thesis focuses on the characterization of hAQP0, hAQP5 and hAQP10 and these aquaporins are described below.

1.2.2 Aquaporin 0

In the 1980s, AQP0 (also known as Major Intrinsic Protein, MIP26) was the first aquaporin discovered. It is the most abundant membrane protein in lens fiber cells where it regulates the transparency of the lens in the eye [33]. In comparison to other aquaporins, AQP0 transport water at remarkably low rates and instead it is more involved in cell adhesion. It acts as a thin junction protein [34,35], reducing the refractive index in the lens by efficiently decreasing the intercellular distances. AQP0 also regulates the formation of gap junctions, membrane-spanning proteins that form pores that connect the cytoplasm of cells. The gap junctions are made up by complexes of connexins that allow small (<1kDa) molecules to pass between the cells, something that is particularly important in the lens as there are no blood vessels to supply the cells with nutrients and discard the byproducts of the cell metabolism [36].

The function and localization of AQP0 is closely regulated by C-terminal truncation [37], pH and Ca^{2+}/Calmodulin (CaM) [38,39]. The C-terminal is vital for the proper trafficking of AQP0 to the membrane and naturally occurring truncations affect the translocation of the protein to the plasma membrane [40]. Once in the membrane, three histidine residues in the extracellular loops are responsible for the pH dependent reduction in water transport that occur when lowering the pH from 7.5 to 6.5 [38]. Interestingly, CaM turned out to increase the permeability of endogenous AQP0 in lens fiber cells while AQP0 expressed in oocytes showed an CaM-mediated inhibition under similar conditions [39]. It has been reported that both AQP0 and AQP4 have a close relationship between function and the lipid environment with respect to composition, thickness and elasticity [41,42] which could be the reason for the different behavior in the two systems.

1.2.3 Aquaporin 5

hAQP5 is primarily found in tissues such as lungs, airways, tear- and salivary glands. Malfunction of hAQP5 causes Sjögren's syndrome with symptoms such as dry mouth and eyes [25,43,44]. This is an systemic disease with multi-organ consequences and without any treatment [45]. Current evidence suggests that this is an effect of improper trafficking rather than nonfunctional protein. Patients with Sjögren's syndrome show accumulation of hAQP5 at the basal membranes while hAQP5 in healthy individuals is mainly targeted to the apical membranes [46,47].

In 2005, the high resolution structure of hAQP5 was determined to a resolution of 2.0 Å [48]. It is one of the few published aquaporin structures that contain the full tetramer in the asymmetric unit. This gives a chance to see differences between the individual monomers showing two different conformations of the C-terminus, indicating high flexibility and thus a possible role in regulation. A similar flexibility in the C-terminus was observed in the recently solved structure of hAQP2 where its position differed between all four monomers [49]. The space group also allows the central pore to be studied further. In the hAQP5 structure a lipid corresponding to phosphatidylserine was present in the central pore, effectively blocking it. With the implied involvement of the central pore in gas transport of aquaporins [50,51] it is interesting to note that hAQP5 knock-out mice showed a reduction in the water permeability while the CO_2 transport was unaffected [52].

1.2.4 Aquaporin 10

hAQP10 is one of the most recently discovered aquaglyceroporins and it was found in the human small intestine [53]. This first discovery was soon proven to be an incompletely spliced gene product that lacks the sixth transmembrane domain and shows very low water transport rate and no glycerol transport at all. Shortly after, a correctly spliced version was discovered that contained all the six transmembrane domains and displayed the transport of water, glycerol and urea at high rates in a mercury sensitive manner (**Figure 4**) [54]. hAQP10 has also been found in adipocytes [55,56] and in the skin [57] where it could be involved in obesity and eczema, respectively.

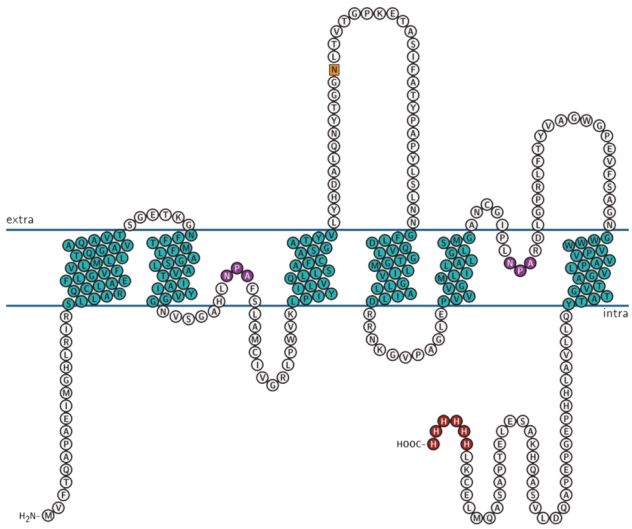


Figure 4. hAQP10 contain a glycosylation site in the extracellular loop C. *The predicted topology of hAQP10 show an extended loop C with the glycosylation site (orange). TM regions (teal), NPA motifs (purple) and his tag (red) are labeled. The figure was made using TexTopo* [58].

The small intestine is the main site for nutrient absorption and as such it has a large surface area of lining cells to increase the uptake. In its relaxed state the small intestine stretches for up to 6.5 meters and the surface area has been estimated to cover ca 300 m^2 [59], although recent studies imply that this is exaggerated and that the more accurate number is 30 m^2 [60]. The large surface is achieved by the intestinal villi that are lined with the absorbing cells, the enterocytes which further increase the surface. hAQP10 was found to localize to the apical membrane of the

enterocytes, thus absorbing water and glycerol from the intestine lumen [61].

1.3 Protein:protein interactions

The famous poem "No man is an island" could easily be related to proteins, as they are connected to the cellular environment in which they exist and come in contact with each other. Protein interactions are the basis for protein function and as such they are vital in our understanding of cellular processes. Protein:protein affinities range from stable protein complexes, which together carry out a function, to transient complexes that mediate signals in cellular pathways. Stable protein complexes can usually be purified using conventional methods. Other approaches are needed to obtain structural information from protein assemblies with low affinities between the proteins within the complex.

In the 1960s it was discovered that the *E. coli* β-galactosidase could be translated as two separate peptide chains that assembled to make up a functional enzyme [62]. Since then, many proteins have been identified where fragments spontaneously associate to form a functional complex including dihydrofolate reductase (DHFR) [63], β-lactamase [64], luciferase [65], TEV protease [66] and fluorescent proteins. This was the basis for the development of protein complementation assays (PCAs). Two target proteins are fused to non-functional fragments and upon complex formation, the activity of the reporter protein is regained and the output can be recorded.

One of the more successful approaches in screening for novel PPIs has been the yeast two-hybrid (YTH) system where multi domain transcription factors are divided into fragments that together enable the transcription of a reporter gene. The first report of YTH in 1989 emphasize the advantage of studying PPIs *in vivo* and today YTH has been used to study interactomes of a wide range of organisms, from bacteria to human [67–69]. However, the high rates of false positives, which has been estimated to be up to 50%, calls for confirmation of the interaction using other methods.

Bimolecular fluorescence complementation (BiFC) is a more recent assay to study PPIs using fluorescence as readout. It is based on the complementation of fluorescent proteins. Green fluorescent protein (GFP) was first discovered in 1962 as an accompanying protein to a protein purified from an *Aequorea* jellyfish [70]. This 27 kDa protein folds

into an 11 strand β-barrel with an α-helix treading through the center. This α-helix contains an amino acid triad – Ser65, Tyr66 and Gly67 – that undergo an autocatalytic post-translational modification, leading to the formation of the chromophore (**Figure 5**). The maturation is a three-step process starting with the cyclization of Ser65 and Gly67 followed by the oxygenation of Tyr66 resulting in a conjugated π system that fluoresce upon excitation.

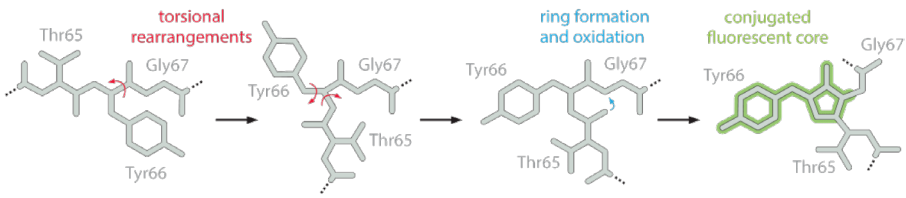


Figure 5. Cyclization and fluorophore formation in GFP is a three-step process*. Figure reprinted with permission from reference* [71].

Through mutagenesis the properties of the wild type GFP has changed, yielding enhanced GFP (GFP-F64L, S65T). These mutations caused a shift in optimal excitation wavelength as well as a 35-fold increase in the fluorescence intensity [72]. From a protein engineering point of view it is interesting to note that mutations that improve the fluorescence properties often destabilize the protein and interfere with folding and maturation, which call for additional mutations to regain protein stability. Mutations in the direct proximity of the chromophore have given rise to a new set of fluorescent proteins emitting light with a variety of wavelengths [73]. Further improvements affecting the oligomeric state [74], fluorescence intensity [75] and maturation efficiency [76] have resulted in protein variants with properties well suited for many different applications. Improving GFP with respect to these factors has created a large number of variants (the variants related to SYFP2, used in this thesis, are summarized in **Table 2**).

Table 2. Yellow fluorescent protein variants derived from GFP.

Fluorescent protein	Mutations
YFP	GFP - T203Y
EYFP [77]	YFP - S56G, S72A
SEYFP [77]	YFP - S56G, F64L, S72A, M153T, V163A, S175G
venusYFP [77]	YFP - F46L, S56G, F64L, S72A, M153T, V163A, S175G
SYFP2 [78]	YFP - F46L, S56G, F64V, S72A, M153T, V163A, S175G

The BiFC assay is based on that YFP can be divided into two non-fluorescent fragments that regain fluorescence as they refold. For protein interaction studies, the target proteins are fused to either of the fragments and upon complex formation YFP is reassembled and fluorescence is obtained [79]. YFP is split at positions in the loop between the β-strands where two main sites are used in BiFC. Most common is splitting between the 8th and 9th strand (generating YFP_N1-173/YFP_C174-239) or alternatively between the 7th and 8th strand (generating the combination used in this thesis, YFP_N1-154/YFP_C155-239) [80]. Several combinations between the YFP fragment and target protein as well as the fusion sites (N-terminal versus C-terminal YFP) should be evaluated as the YFP fragments must be able to associate for the fluorescent complex to form (**Figure 6**).

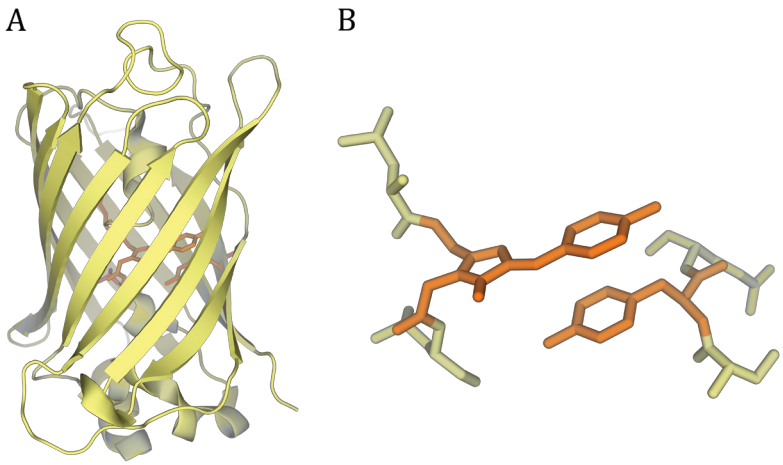


Figure 6. YFP is a β-barrel with an α-helix treading through the center. *All YFP (PDB ID: 1MYW) variants have the GFP-T203Y mutation in the vicinity of the chromophore (orange) which shift the emitted light to longer wavelengths.*

The maturation of the fluorophore is the rate limiting step in the use of fluorescent proteins and for BiFC in particular. The two fragments of the fluorescent protein need to come together in order to form the chemical environment that is necessary for the fluorophore to mature. The two YFP fragments come together in minutes while the formation of the chromophore usually is slower. Important for the BiFC assay is the notion that the assembly to the full fluorescent protein is practically irreversible even though reversibility has been reported in some cases [81–83]. In general, the fluorophore maturation time limits the use of BiFC for monitoring PPIs in real time. However, variants with extremely short maturation times have been developed that has made it possible to follow reactions on the timescale of minutes [84].

1.4 Recombinant protein production

Characterization and structure determination requires large amounts of protein. When working with eukaryotic and human membrane proteins in particular, the conventional approach is to use a recombinant source.

1.4.1 Protein production systems

The recombinant production of eukaryotic integral membrane proteins, with a few exceptions, requires a eukaryotic host [85]. Factors such as lipid composition, chaperones and a post-translational machinery are all necessary to produce properly folded and functional protein located in the membrane [86]. Unicellular organisms such as yeasts combine many of the benefits of a eukaryotic system with the easy and inexpensive means of genetic modification and culturing. In this thesis two different species of yeast, *Saccharomyces cerevisiae* and *Pichia pastoris*, were used in protein production, both with their specific advantages.

S. cerevisiae was the first organism to have its complete genome sequenced [87] and there is a large amount of knowledge concerning the genetics and molecular biology associated with it in the literature [88]. There are a large number of strains and vectors available that can be used to tailor the production system for your needs which allow a multitude of experiments to be carried out.

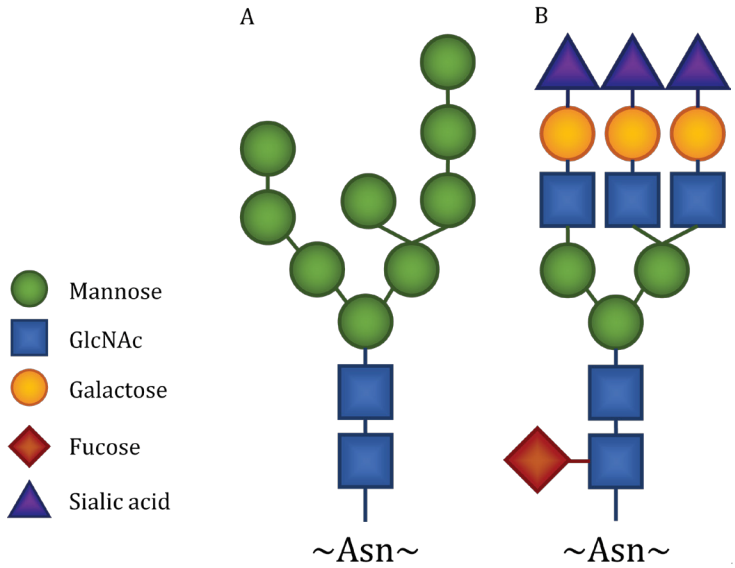
P. pastoris is known for its ability to grow to extremely high cell densities, especially when grown in bioreactors where parameters such as aeration, pH and nutrient supply can be closely regulated [89]. The

P. pastoris genome is now publicly available [90], but the genetic tools are not as elaborate as for *S. cerevisiae*. The presence of a strong, inducible promoter – alcohol oxidase1 (*AOX*1) – is the key to success for this production system. *P. pastoris* is a methylotrophic yeast which can utilize methanol as the sole carbon source. The enzyme is present in the first step of the methanol utilization pathway where it oxidizes methanol. To compensate for its very poor affinity for oxygen, the enzyme is present in large quantities. Replacing the *AOX*1 gene with the recombinant gene generate a potentially high-yielding production system.

1.4.2 Posttranslational modifications

There is a vast diversity in protein modifications that take place during or after the translation of the peptide chain. Amino acids can be chemically altered or have smaller chemical groups attached. More complex molecules such as carbohydrates, lipids or small proteins can be transferred to the peptide chain and proteolytic cleavage is common [91]. The many purposes of these modifications include fine-tuning of protein function, altering protein stability and aiding protein folding and localization [92].

Glycosylations occur in all domains of life [93] and in eukaryotes they can be divided into two classes depending on if the glycan is attached to a nitrogen (N-linked) or oxygen (O-linked) [94]. The assembly of the glycosylations in the two classes is completely different. O-linked glycosylations are characterized by a sequential addition of monosaccharides to hydroxyl groups in sequences rich in serine, threonine and proline without them being part of a particular recognition sequence. O-linked glycosylation in mammalian cells is initiated in the Golgi apparatus, when the protein has already folded and awaits cellular sorting [95,96]. In contrast, the O-linked glycosylation in yeast is exclusively initiated in the endoplasmic reticulum (ER) resulting in a larger potential influence on protein folding [97].



7. The glycosylation pattern in P. pastoris are structurally similar to that of higher eukaryotes. *There are two main types of N-linked glycosylations in humans, mannose-rich oligosaccharides* ***(A)*** *and complex oligosaccharides* ***(B)****. P. pastoris predominantly produce mannose-rich glycosylations with high resemblance to the human high-mannose glycosylations. Figure adapted from reference* [98].

In yeast, the majority of the glycosylations characterized so far have been N-linked [99]. N-linked glycosylations are built on a lipid carrier before transfer *en bloc* to an asparagine of a N-X-S/T sequon in the acceptor protein [100]. Further processing after the transfer allows a wide variation of glycosylation structures, something that is species specific in its appearance. *S. cerevisiae* is known to hyperglycosylate its proteins with up to 150 mannose residues being attached to a single site [99], while *P. pastoris* adds 8-14 mannose residues on average, which is in the same range as observed in higher eukaryotes (**Figure 7**) [101].

The presence of glycosylations introduces micro-heterogeneity into the protein sample which could interfere with crystal growth. Nevertheless, several protein structures have been solved with the glycosylation present and it could be critical in obtaining a functional and properly folded protein [102]. Glycosylation of aquaporins has shown to be required for proper trafficking in plants [103]. hAQP1 and hAQP3 have also been shown to be subjected to glycosylation but without any obvious impact on function [104].

2 Scope of the thesis

This thesis deals with the production of membrane proteins and membrane protein complexes. It is also a thorough examination of protein characterization with respect to aquaporin function and regulation.

Membrane proteins are the gatekeepers of the cell and the targets of the majority of the drugs on the market today. The mechanism of protein function is tightly connected to protein structure and high resolution structures by X-ray crystallography have provided exciting details of an ever increasing number of protein families. Method development is necessary to enable structure determination and this thesis starts at the first step: the protein production which creates the foundation for a successful project in structural biology.

In **Paper I** we studied the production of aquaporins, a membrane protein family that despite high sequence similarity shows large variations when produced recombinantly in *P. pastoris*. The importance of careful clone selection and codon optimization of the targets was evident and something that could be incorporated in all the following work.

Characterization of aquaporin regulation was carried out in the following papers. **Paper II** investigates the effect of glycosylation on the thermostability of hAQP10, without any significant effect on function or selectivity. **Paper III** reports how hAQP5 translocation is regulated by at least three independent mechanisms involving phosphorylation at Ser156, protein kinase A activity and extracellular tonicity.

Human aquaporins are mainly regulated by trafficking and protein:protein interactions are vital for proper function and localization. The extensive literature study of a selection of AQPs presented in **Paper IV** identifies the C-terminus as the major site of aquaporin interaction and also summarizes the diversity of the protein families that are known to directly interact with AQPs. As protein:protein interactions provide the means of regulation, a novel method to produce and purify protein complexes was developed and reported in **Paper V**. The use of BiFC to study protein:protein interactions *in vivo* was extended and used to successfully purify a protein complex of high purity in quantities suitable for crystallization.

3 Protein production optimization

The first membrane protein structures determined by X-ray crystallography were purified from native sources where the protein was present in high amounts [105]. The development of recombinant protein production made it possible to produce targets that are less predominant in their native environment. Indeed, the fraction of membrane proteins purified from recombinant sources has surpassed the protein purified from native sources and is now the method of choice for most structural biology projects (**Figure 8**) [106].

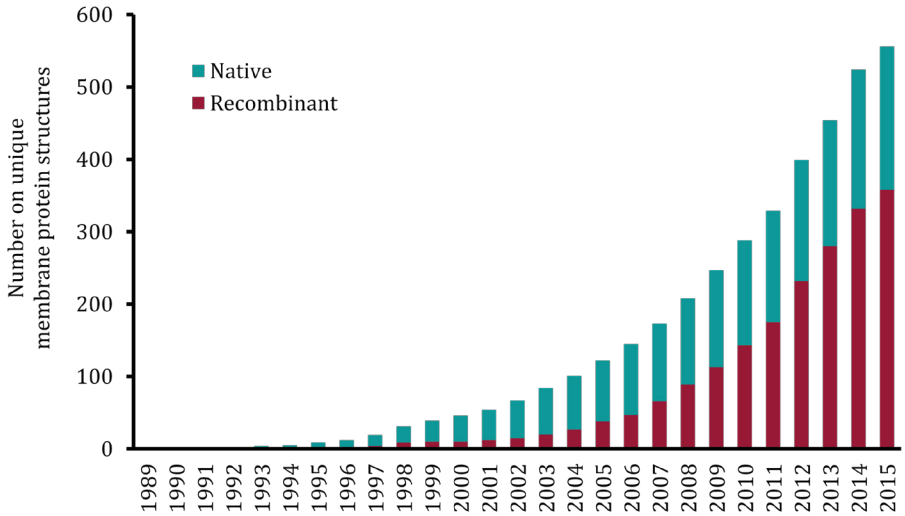


Figure 8. The number of unique membrane proteins increases exponentially. *The majority of the structures are determined on protein from recombinant hosts* [106].

With the increased use of recombinant protein overproduction, the development of strategies to optimize the yields took off. There are two different approaches to optimize the production depending on how well characterized the target is. General methods to increase protein yields can be applied to all targets while specific mechanisms that improve protein stability requires knowledge of the protein to be an option. In this thesis both approaches have been used on different targets.

P. pastoris as a protein factory has been optimized and the workflow is now available from Invitrogen [99]. The gene of interest is cloned into a pPICZ vector and amplified in *E. coli*. Following linearization, the plasmid is transformed to *P. pastoris* where homologous recombination inserts it

into the genome, localizing the gene downstream of the AOX promoter. The Zeocin selection marker accompanies the gene and is used to assist in clone selection. *P. pastoris* was used as the host for protein production in **Paper I** and **Paper II**.

3.1 Generating high yielding clones

Working with membrane proteins the strategies for overproduction can be a bit different when compared to soluble proteins. It may not be beneficial to push the production to a maximum as the protein sorting and translocation systems need to keep up for the protein to be properly processed and translocated. The rate at which the proteins are transcribed and translated is faster in prokaryotes compared to eukaryotes which is one of the reasons why eukaryotic proteins are often more successfully produced in eukaryotic hosts [89]. By slowing down the translation rate in *E. coli*, Siller *et al* showed how the protein folding was improved and resulted in less aggregation [107]. Taken together, strategies that have proven successful for soluble proteins need to be evaluated for membrane proteins. The effects of codon optimization and transformation methods were evaluated for human aquaporins in **Paper I**.

The genetic code is built from four nucleotides that combine into triplets to be translated into amino acids. Since there are more triplets (codons) than unique amino acids most amino acids are encoded by more than one codon, a phenomenon referred to as the degeneracy of the genetic code. Different species use the degenerate codons with a different frequency and in general, highly expressed genes are biased towards codons that are recognized by the most abundant tRNA. The tRNA pool is species specific and the use of different codons can be compared between organisms using the codon adaptive index (CAI). For example *Plasmodium falciparum*, the malaria parasite, has a remarkable high A and T content in its genome [108] which could be a problem when expressing genes originating from it in recombinant hosts (**Table 3**). For the only aquaglyceroporin present in the parasite, PfAQP, codon optimization of the gene for production in *P. pastoris* increased the yields enough to make protein characterization and initial crystallization possible [109].

Table 3*. The codon adaptive index for PfAQP, hAQP1 and hAQP4 of the wildtype sequence compared to the sequence optimized for S. cerevisiae. A higher value indicate a higher proportion of the more abundant codons.*

Protein	CAI native seq	CAI optimized seq
PfAQP	0.55	0.97
hAQP1	0.05	0.97
hAQP4	0.05	0.97

By understanding the host organism and its translation machinery, substantial improvements can be made by optimizing the codon usage to fit the production host [109]. In **Paper I** the yields of hAQP4 could be increased 10-fold by codon optimizing the gene (creating Opt-hAQP4) for production in *P. pastoris*. **Table 3** shows that the large difference in production between hAQP1 and hAQP4 cannot be explained solely by relating the native DNA sequence and the yields in yeast. There must be other factors also affecting the final yields in protein overproduction.

P. pastoris strains were traditionally created by chemical transformation of linearized DNA, but more efficient and convenient methods have now been developed [110]. The linearized DNA is incorporated into the genome by homologous recombination and electroporation has been shown to increase the chance of insertion of several copies of the gene. The multi-copy integrants occur with a frequency ranging from 1-10% and is generally correlated with a higher protein production when evaluated for soluble proteins. In **Paper I** we investigated the influence of transformation method on the production of aquaporins, comparing chemical transformation by standard LiCl protocols and electroporation. The effect on production of high yielding targets was negligible, as exemplified by hAQP1, while hAQP8, normally produced at moderate levels, showed a five-fold increase when using electroporation compared to the chemical transformation. The same effect was seen for low-producing targets where the production of hAQP4 was increased from non-detectable to detectable levels.

3.2 Clone selection

An important consequence of the multi-copy integration events of the transformed gene is the introduced variation between the clones. Evaluating the growth at a higher selection pressure could help in estimating the number of inserts introduced into the *P. pastoris* genome [99]. Since overproduction of functional membrane proteins can affect the viability of the cell and large amounts of protein can choke the

protein production machinery, the localization and protein yields should be confirmed by more thorough analysis. Growth on high Zeocin media should be used as a screening procedure where clones producing large colonies are taken to quantitative analysis. Further evaluation with respect to protein stability and localization are made from medium scale cultures after 6 h of induction followed by western blot on total cell lysate and fractioned membranes as described earlier [32]. There is a strong correlation between colony size and protein signal detected by western blot (**Figure 9**). This correlation is not given for membrane proteins but holds for the aquaporins. To confirm proper membrane localization and stability, a small number of colonies should be selected for quantitation (**Paper I**).

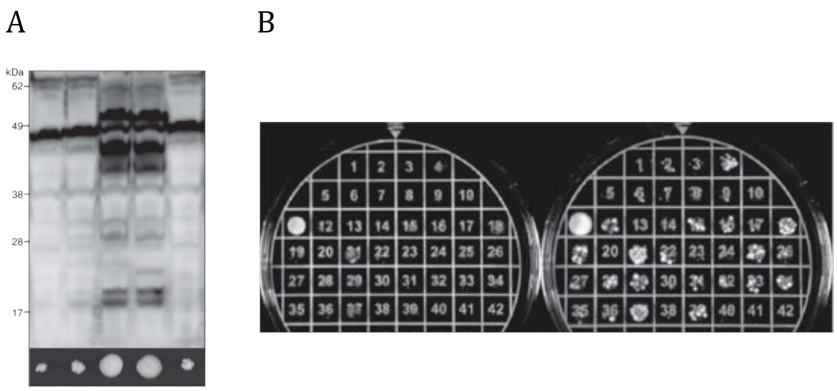


Figure 9. The protein production can be correlated to growth on high concentration Zeocin medium. A) *Quantitation of the protein signal from cell lysate. The protein signal correlate well with the colony size (bottom).* ***B)*** *Growth on high Zeocin plates can be used to select clones for further evaluation. Figure adapted from* ***Paper I****.*

3.3 Protein stability

Protein stability is key to successful overproduction and crystallization of proteins. Fusion proteins are frequently used to improve the solubility, production yields and/or trafficking of the target. For the G-protein coupled receptors (GPCRs), signal transducers that are highly flexible and thus difficult to crystallize, T4 lysozyme (T4L) was critical in obtaining high-diffraction quality crystals [111]. T4L was engineered into one of the intracellular loops, stabilizing it and providing critical crystal contacts via the hydrophilic domain, leading to the first high resolution structure of a GPCR [112].

In membrane protein production the use of GFP fused to the target protein significantly streamlined the screening of optimal conditions for protein production, solubilization and purification. GFP-fusions of eukaryotic membrane proteins have been produced in a high-throughput fashion in *S. cerevisiae* [113]. Rapid cloning, taking advantage of the homologous recombination, allow an efficient evaluation of the production levels and protein behavior during solubilization and purification. In **Paper III** we utilize hAQP5-GFP fusions to study the trafficking of hAQP5 (Section 4.3). Further, **Paper V** utilizes the experiences from work on GFP fusions in a novel purification method for protein complexes (Section 4.4).

Mistic, a soluble bacterial protein that strongly associates with the membrane [114], has shown to be a successful fusion protein for overproducing both human and bacterial membrane proteins in *E. coli* [115,116]. However, as **Paper I** reports, this approach is not transferable to yeast. Mistic was codon optimized for *P. pastoris* and used as a leader sequence for hAQP1 and hAQP8 without any increase in protein yields. As membrane proteins are processed in the ER, the mechanisms for protein sorting in eukaryotes seem to suppress the positive effects of Mistic seen in prokaryotic systems. We also tried fusions of hAQP1 to the N-terminal of hAQP8 which resulted in substantial degradation and aggregation, suggesting that the fusions were unstable.

We turned to investigate why aquaporins display large variation in production levels despite their high sequence homology. We sought to find the sequence(s) within hAQP1 that is responsible for high production. Chimeric constructs where hAQP8, produced at moderate levels, was stepwise replaced by the corresponding sequence of hAQP1 was tested to determine the effect on production levels. The fact that the chimeras resulted in a complete loss of protein production (N-terminal and TM1) or no effect (TM1-2 and TM1-3) indicate that hAQP8 does not fold properly when merged with hAQP1, possibly explained by different folding pathways.

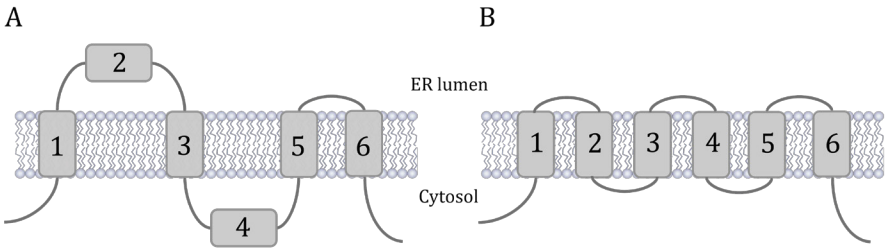


Figure 10. The folding pathway of aquaporins can be determined by two amino acids. *The folding pathway of hAQP4* ***(A)*** *can be changed to mimic the folding pathway of hAQP1* ***(B)*** *by the mutations of two amino acids close to TM2 and TM4. Figure inspired by reference* [117]***.***

Extensive studies of aquaporin stability have been carried out using hAQP1 and hAQP4, two orthodox aquaporins with almost 50% sequence identity that nevertheless show one of the largest differences in production yields [32]. hAQP4 folds sequentially into a six transmembrane (TM) helix structure as the peptide emerges from the ribosome (**Figure 10**). hAQP1 on the other hand, folds into an four helix intermediate state before rearranging to the final topology [117,118]. The signals that stop translocation of the helices have been shown to involve two hydrophilic amino acids in the beginning of TM2 and TM4, and explain this intermediary topology of hAQP1 [119]. The corresponding amino acids in hAQP4 are of hydrophobic nature and point mutations to exchange them for the corresponding amino acids in hAQP1 (hAQP4-M48N, L50K, generating hAQP4*) changes the folding pathway to mimic that of hAQP1. These mutations have a dramatic effect on the hAQP4 yield which is increased more than 10-fold (**Figure 11**).

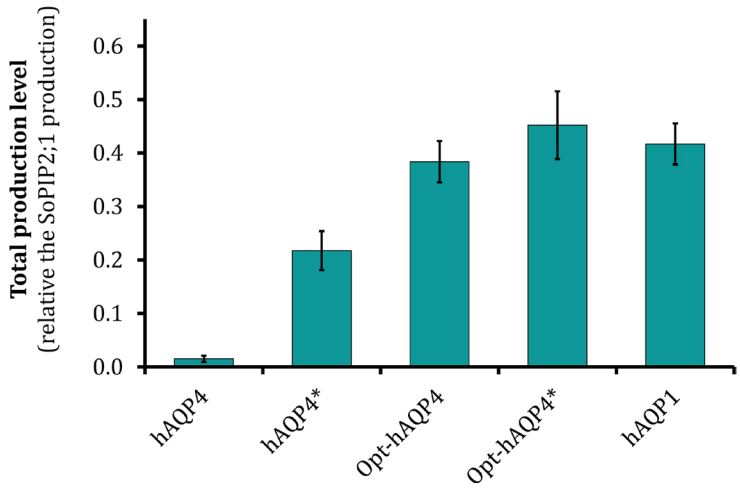


***Figure 11**. **Codon optimization and rational site directed mutagenesis increased the yields of hAQP4 significantly.** Figure adapted from **Paper I**.*

The final yield of hAQP4 is a result of optimization on many levels. A higher gene dosage, improved transcriptional efficiency (sequence analysis performed previously by Öberg *et al* [32]) and more efficient translation result in higher protein levels. Including the improved protein stability due to the modified folding pathway, we have generated a functional hAQP4 construct that can be obtained in sufficient amounts for functional and structural investigation (hAQP4*-Opt).

4 Aquaporin characterization

Protein characterization is a sub discipline that involve scientists from a wide variety of fields including physicians, protein chemists and structural biologists. Together we aim to describe a biological system, approaching it from different angles. This thesis contribute to the characterization of hAQP0, hAQP5 and hAQP10.

hAQP10 is one of the human aquaporins with the highest production levels in *P. pastoris* and the only aquaporin that consistently give rise to distinct double bands on SDS-PAGE [32]. In **Paper II** we showed that this is caused by a glycosylation. Two out of three putative glycosylation sites in hAQP10 are found in the extracellular loops and are thus accessible for glycosylation under physiological conditions. To determine the glycosylation pattern of hAQP10, the two bands discussed were analyzed by mass spectrometry, unambiguously locating the glycosylation to Asn133 in loop C. This glycosylation was shown to be a standard mannose-rich, N-linked glycosylation with nine mannose residues attached to the glycosylation backbone.

Mutagenesis of a glycosylation site is a common strategy to prevent a glycosylation from being transferred to the protein. By mutating the Asn in the N-X-S/T recognition site to a Gln, the site is disrupted while the chemical properties of the amino acid are maintained. In **Paper II** we investigated the effect of the glycosylation on protein function and stability. We created hAQP10-N133Q which specifically abolishes the glycosylation motif. Additional purification steps made it possible to isolate the unglycosylated population of the wild type mixture, hAQP10Δglyc. These hAQP10 samples were compared with respect to function and thermostability.

4.1 Thermostability

Glycosylations play a major role in the protein sorting and even small deviations from the normal glycosylation patterns can have severe effects. Glycosylations can often modulate protein folding, function and stability [120,121], which lead us to investigate the impact this specific glycosylation might have on hAQP10.

Protein stability was measured as the loss of secondary structure with an increase in temperature using circular dichroism (CD) (**Figure 12**). By

monitoring the CD-signal at 222 nm the loss of α-helical content can be observed over time while ramping the temperature. Reproducible curves revealed a 3-5 °C difference in apparent melting temperature between the non-glycosylated hAQP10 and the wild type (**Paper II**). The monomodal appearance of all curves suggest that we have a uniform population in all the samples of hAQP10. Quantitation of the ratio of glycosylated:nonglycosylated protein in the wild type sample lead us to the conclusion that one monomer per tetramer is glycosylated. The relation to protein stability in a physiological context is not entirely obvious, but nevertheless the glycosylation seem to play a role in protein stability.

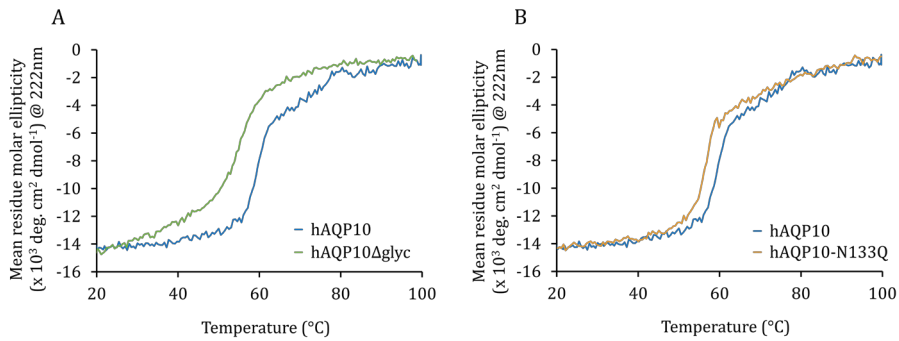


Figure 12. The glycosylation increases the thermostability of hAQP10 by 3-6°C*. The thermostability of hAQP10 was measured as loss of α-helical content with increased temperature using CD at 222 nm. hAQP0Δglyc has an apparent melting temperature approximately 5°C lower than wild type hAQP10. Figure from* ***Paper II.***

Could it be that the asymmetrical glycosylation pattern of hAQP10 is an artefact from a strong overproduction where the post-translational modification machinery cannot keep up with the amounts of protein produced? Our standard protocol of cultivation by fermentation, where the cells are fed 100% methanol for induction, could indeed push the cells into a very non-physiological growth pattern. Fermentation protocols involving a lower induction was investigated and a mixed feed (40% methanol + 60% sorbitol) was employed for induction. The cells are provided with a non-repressing carbon source that allow the cells to continue to grow. This had no impact on the yield or relative amount of glycosylation in hAQP10. On the other hand, some protein degradation products were no longer seen with the mixed feed induction and this approach was used subsequently for hAQP10 cultivation.

For crystallographic purposes the heterogeneity introduced by this partial glycosylation is decreasing the probability of obtaining well-ordered crystals. In order to obtain the first structure of a human aquaglyceroporin, this problem might need to be addressed from a different perspective, maybe by generating more thermostable hAQP10 mutants that could give a stable, homogenous protein suitable for crystallization.

4.2 Functional assays

An important part of protein characterization is to determine or confirm functionality. The difficulties associated with measuring water transport are related to the large abundance of water, the very fast transport rates and the background transport across the membrane. These experiments can be carried out *in vivo* – using either oocytes or spheroplasts (yeast cells with the cell wall removed) – or in artificial membrane mimicking systems such as liposomes. *In vivo* systems will produce and transport the protein to the plasma membrane and this is advantageous as all the protein will have the same, correct localization. As the protein is never removed from its more native, cellular environment the problems associated with protein purification are bypassed. However, targets with low water transporting abilities could be difficult to measure reliably and inherent variations in protein production levels are challenging to correct for.

In **Paper II** the functional studies of hAQP10 were carried out by reconstituting purified protein into liposomes. The compartment made up by the lipid vesicles is smaller than the cell based systems and the lipid membrane will mimic the native lipid bilayer. With this artificial approach it is possible to manipulate the lipid composition to suit the target protein. The lipids are solubilized in a detergent solution with the addition of the protein, and as the detergent is slowly adsorbed by polystyrene beads, proteoliposomes are formed. As the water transport through aquaporins is both bidirectional and passive, a random orientation of the protein in the liposome is often not a problem. Using stopped flow spectroscopy the liposomes are rapidly mixed with a hypertonic sucrose buffer which cause an outward flow of water and a shrinking of the vesicles. This can be followed by measuring the increase in light scattering as a function of time resulting in exponential curves which can be fitted with a one or a two exponential function. The control liposomes of only lipids are fitted with a one exponential equation

estimating the passive water transport across the membrane. This contribution is also present in the proteoliposome measurement, but a faster component is also present which correspond to the transport facilitated by the aquaporins.

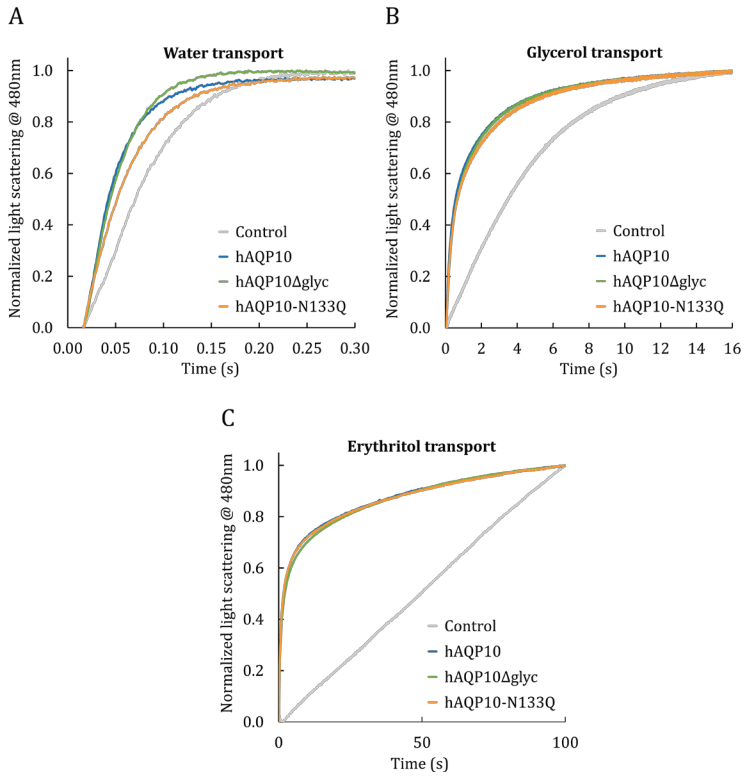


Figure 13. Glycosylated hAQP10 transport water, glycerol and erythritol as the wild type hAQP10. *The graphs show the transport of water* ***(A)****, glycerol* ***(B)*** *and erythritol* ***(C)*** *for wildtype* ***(blue)****, hAQP10Δglyc* ***(green)*** *and hAQP10-N133Q* ***(orange)*** *and control liposomes* ***(grey)****. Figure from* ***Paper II.***

In addition to water, hAQP10 has also shown to transport glycerol [54]. This can be measured in the liposome system by suspending the liposomes in a glycerol containing buffer that is isotonic to the sucrose buffer used to drive the transport. As the liposomes are mixed with the sucrose buffer, glycerol is transported out of the cell, along the gradient. The increase in solutes on the outside of the liposomes cause water to be co-transported and the vesicles to shrink. The same approach was used to measure the transport of erythritol.

Erythritol is a four-carbon sugar alcohol commonly used as a food additive in calorie-reduced beverages (known as E968). Larger sugar alcohols (sorbitol, xylitol etc.) are not absorbed very well by the intestine and irritate the large intestine causing a laxative effect. Interestingly, 90% of the erythritol ingested is absorbed and never reach the large intestine [122]. In **Paper II** we show that hAQP10 transport erythritol (**Figure 13**) and can facilitate the uptake in the small intestine. We've also showed that hAQP10 transports xylitol at very low rates and that the sorbitol transport is non-detectable (Fischer, Öberg, Sjöhamn, unpublished data). As a consequence, xylitol and sorbitol end up in the large intestine where they cause the side-effects commonly associated with sugar alcohols. The rates of glycerol and erythritol is unaltered by the glycosylation while hAQP10-N133Q showed a slower water transport compared to the hAQP10 with wild type sequence (**Figure 13**, **Paper II**).

Reports on glycosylated aquaporins in their native tissue also suggest an asymmetric glycosylation of the tetramer. hAQP1 purified from erythrocytes had N-linked glycosylations attached to 50% of the monomers, but removing it had no impact on functionality or oligomerization state [123]. Whether or not the glycosylation affect hAQP10 trafficking in the small intestine is something that remains to be determined, but in *P. pastoris* there is no difference in membrane localization between the wildtype and the mutated hAQP10 protein.

4.3 Trafficking and regulation

Modulation of aquaporin activity is critical in controlling the water flux across the membrane. Aquaporins in plant and fungi have been shown to undergo conformational changes that close the pore. Structural evidence of gating has been found in the plant aquaporin SoPIP2;1 [124] and in the yeast aquaglyceroporin AQY1 [22]. Conformational changes of loop D and the N-terminus, respectively, cause a single residue to be inserted into the channel resulting in closing of the pore. Gating of human aquaporins seems to be more controversial but has been implied in AQP0, hAQP2 and hAQP4 while the evidence of gating of hAQP5 mainly include molecular dynamics simulations [125].

In mammals, the water permeability of the membrane is instead fine-tuned by regulating the amount of aquaporin present in the membrane. The protein is stored in intracellular vesicles that merge with the plasma membrane in response to external stimuli, a mechanism referred to as trafficking. The most well-understood case of aquaporin trafficking is the

absorption of primary urine by hAQP2 in the collecting ducts of the kidneys [126]. hAQP2-containing vesicles are fused to the plasma membrane in response to the hormone vasopressin and in a matter of minutes the permeability of the plasma membrane is increased [127]. Vasopressin activates protein kinase A (PKA) in a cAMP-dependent manner, triggering phosphorylation of serine residues in the hAQP2 C-terminus where Ser256 is the most prominent one [128]. In the case of hAQP2 phosphorylation, dephosphorylation and ubiquitination together regulate the amount of hAQP2 active in the membrane [129]. Mutations in hAQP2 that interfere with trafficking lead to conditions such as nephrogenic diabetes insipidus (NDI) where an inability to reabsorb water in the kidneys cause severe dehydration [130].

hAQP5, the closest homologue to hAQP2 with 63% sequence identity, is also regulated by trafficking. hAQP5 contains a number of putative PKA consensus sites of which Ser156 in loop D is thought to be the main responsible in trafficking. In **Paper III** we investigate the role of the phosphorylation of residue Ser156 by mutating the residue to either abolish the phosphorylation site or introducing an amino acid that mimics a phosphorylation (S156A and S156E respectively). Using hAQP5-GFP fusions in HEK293 cells the cellular localization of hAQP5 can be observed in real-time. In agreement with previous results, wild type hAQP5 and hAQP5-S156A have the same relative membrane localization. In comparison, the basal membrane localization of hAQP5-S156E is elevated (**Figure 14A**). The results presented in **Paper III** show for the first time how a phosphomimicking mutant is linked to a difference in translocation of hAQP5 and it suggests that the phosphorylation state of Ser156 does indeed affect the membrane targeting of hAQP5. Interestingly, the crystal structure of hAQP5-S156E does not deviate significantly from the wild type structure (**Paper III**). Thus, the phosphorylation of Ser156 does not act through conformational changes of the C-terminus.

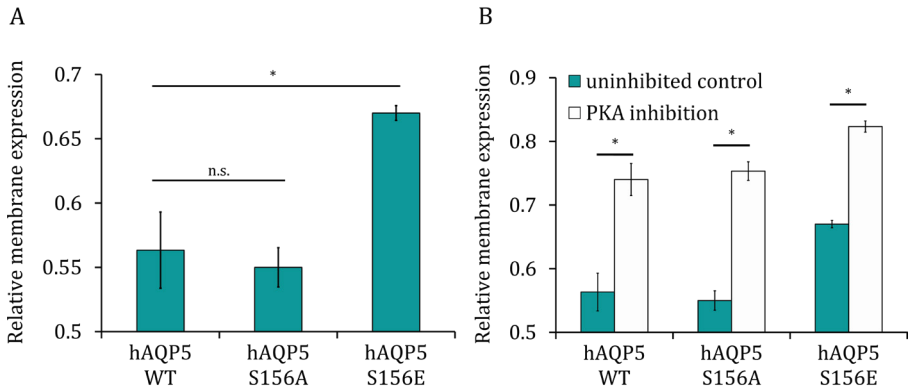


Figure 14. hAQP5 membrane localization depend on the phosphorylation state of Ser156. A) *Introducing the S156E mutation increases the membrane localization in unstimulated cells in comparison to WT hAQP5 and hAQP5-S156A.* ***B)*** *PKA inhibition increases the membrane localization of all hAQP5 variants. Figure adapted from* ***Paper III****.*

AQP5 membrane localization displays a biphasic response to elevated cAMP levels [131]. Within 2 min of cAMP exposure the membrane abundance was reduced. However, after 8 h of incubation, the amount of hAQP5 localized to the membrane was higher than the baseline. These effects were abolished by a PKA inhibitor indicating that PKA activity is important in both responses. In our experiments, wild type hAQP5, hAQP5-S156A and hAQP5-S156E all exhibit increased membrane localization upon PKA inhibition (**Figure 14B**). Whether PKA affects hAQP5 trafficking via the phosphorylation of other proteins or simply other residues in hAQP5 is not possible to determine just from these result. There is evidence of phosphorylation of Thr259 in human cell lines and from tissue specimens of mouse salivary glands, but without an effect on hAQP5 trafficking [132].

Changes in tonicity have been shown to increase the amount of hAQP5 found in mouse lung tissue. Zhou *et al* show that an increased transcription and mRNA stability lead to higher hAQP5 levels when cells are subjected to hypertonic induction [133]. However, the study did not specify the cellular localization of hAQP5 and the membrane localized fraction is not known. From the results presented in **Paper III,** it is clear that the membrane localization of hAQP5 in HEK293 cells is increased as the extracellular tonicity is decreased (**Figure 15**). This mechanism is independent of the phosphorylation of Ser156 and of PKA activity.

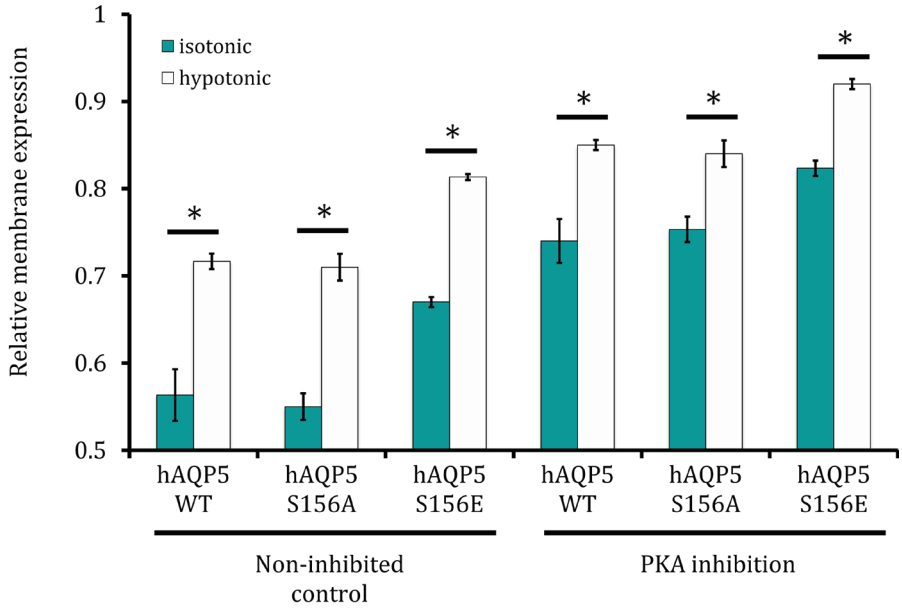


Figure 15. Extracellular tonicity increase the membrane localized hAQP5. *The mechanism by which extracellular hypotonicity change the membrane localization of hAQP5, hAQP5-S156A and hAQP5-S156E is unrelated to Ser156 phosphorylation and PKA activity. Figure adapted from* ***Paper III****.*

This systematic investigation on the effect of phosphorylation of Ser156, PKA activity and the surrounding tonicity lead to the conclusion that three independent mechanisms regulate the trafficking in a coordinated fashion (**Figure 16**). As with most processes in the cell, trafficking is an equilibrium. Both the rate of trafficking to the membrane and the rate of internalization are regulated and affect the final membrane abundance. Changes in membrane permeability are accommodated by a shift in the trafficking equilibrium leading to changes in the amount of membrane localized hAQP5.

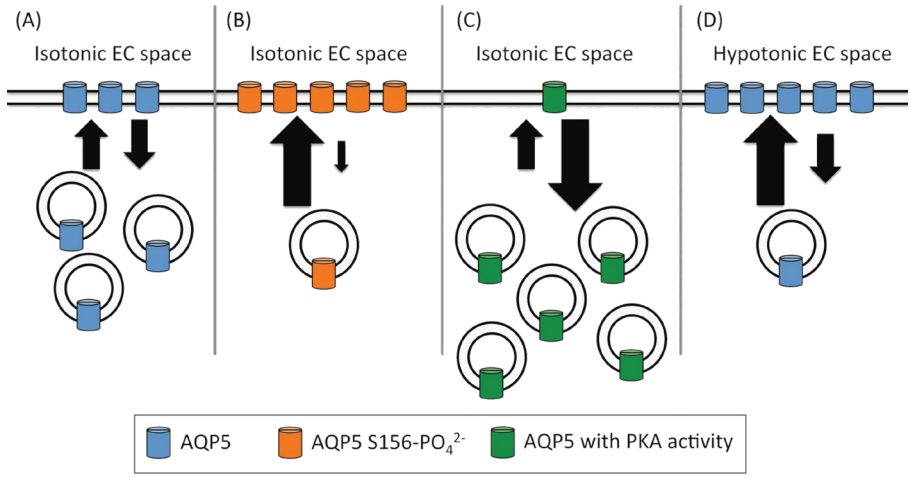


Figure 16. hAQP5 is regulated by trafficking by three independent mechanisms. A) *wild type AQP5 under isotonic conditions.* ***B)*** *Phosphorylation of hAQP5 target the protein to the membrane.* ***C)*** *PKA relocalizes hAQP5 to intracellular vesicles.* ***D)*** *An extracellular decrease in tonicity increase the hAQP5 abundance in the membrane. Figure adapted from* ***Paper III****.*

The complex regulation of hAQP5 trafficking has been shown to also be depending on an interaction with the prolactin-inducable protein (PIP) via a C-terminal interaction. PIP is a 17 kDa glycoprotein with a large number of interaction partners including actin, myosin and CD4 and has been considered a biomarker in breast and prostate cancer due to its overexpression and secretion from those tissues [134,135]. PIP is expressed at lower levels in mouse models of Sjögren's syndrome and has been implicated in the incorrect hAQP5 localization associated with the condition. The evidence of proteins interacting with aquaporins to affect cellular translocation and protein function is increasing and a more complete understanding of aquaporin regulation is now emerging.

4.4 Protein:protein interactions

The extensive literature study of aquaporin protein:protein interactions presented in **Paper IV** summarizes the direct interaction partners known for some medically relevant aquaporins: AQP0, hAQP2, hAQP4 and hAQP5. These aquaporins represent diverse mechanisms of regulation and cellular function. We conclude that almost 70% of the interactions summarized in **Paper IV** involve the C-terminus. This is the

most diverse sequence in the aquaporin family and as such it is often implicated as the main site of regulation.

hAQP2, hAQP4 and hAQP5 are all regulated by trafficking. This is reflected in the nature of the interaction partners which affect the translocation. Even though phosphorylation is an important factor in the trafficking signaling, only three kinases have been identified. As these interactions are likely to be more transient than the other interaction partners, the sensitivity of the assays used may not be high enough to detect them.

The majority of the interaction partners reported for AQP0 and APQ4 are proteins involved in intercellular interactions and cell shape. Both AQP0 and hAQP4 are known to be involved in maintaining the cellular structure in the lens and brain respectively. The arrangement of hAQP4 into crystalline arrays in the membrane is reminiscent of the junction formation of AQP0 in the lens. Both processes are regulated by the termini of protein. There are two splice variants of hAQP4 ($hAQP4_{M1}$ and $hAQP4_{M23}$ respectively) resulting from two translation initiation sites. They differ by 22 amino acids in the N-terminus which has been shown to be important in regulating the formation of the large crystalline ordering of hAQP4 seen in astrocytes [136,137]. In a similar fashion, proteolytic C-terminal truncation of AQP0 regulate the formation of tight junction which in turn affect the water permeability through the pore.

Several AQP0 structures determined by X-ray crystallography and electron diffraction have linked the formation of tight junctions to an open and closed state of the pore [138–142]. A 2D electron crystallography structure solved to 1.9 Å revealed how the two constriction sites, one which is unique to AQP0, elongates to involve a stretch up to 10 Å in junctional AQP0, making the channel too narrow for water to pass [143]. The AQP0 interaction partners identified to date are proteins mainly involved in the formation of gap junctions or otherwise affecting the cell organization (**Paper IV**). AQP0 is the only known example of an aquaporin where a protein interaction directly affects the water permeability by induced gating. The AQP0-Calmodulin interaction is well characterized and used as a proof-of-principle system in the development of a novel strategy for purification of protein complexes presented in **Paper V**.

4.4.1 The AQP0-CaM interaction

In cell signaling Ca^{2+} acts as a second messenger and sudden increases in intracellular Ca^{2+} concentration activates a huge number of downstream targets. Calmodulin (CaM) is one of the major calcium sensor proteins in eukaryotes with the ability to interact with a diverse collection of proteins in the response to Ca^{2+} [144]. This 17 kDa protein has four EF-hand Ca^{2+} binding motifs localized in two globular C- and N-terminal domains separated by an α-helical linker. As Ca^{2+} binds these domains undergo large conformational changes (**Figure 17**), shifting outward and exposing the α-helix linking the EF-hands together [145,146].

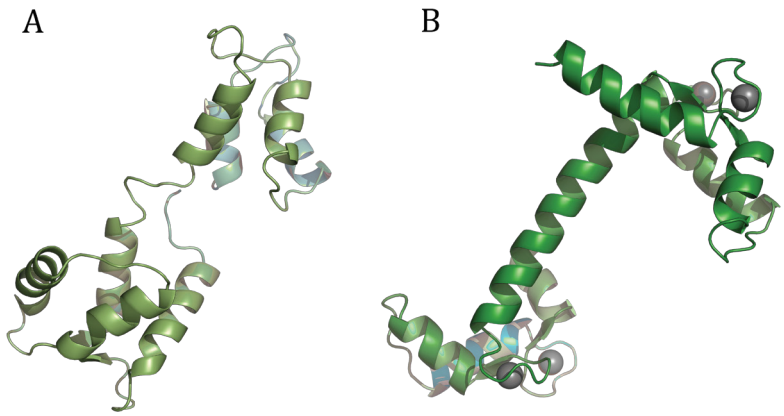


Figure 17. Calmodulin undergo large structural changes upon Ca^{2+} binding. A) *CaM without Ca^{2+} (PDB ID: 1CFD).* ***B)*** *CaM with Ca^{2+} (grey spheres) (PDB ID: 1CLL).*

The hydrophobic stretch of amino acids in the linking α-helix interact with the amphipathic α-helix in the C-terminus of AQP0, an interaction that has been characterized using both biochemical and structural methods [147–149]. NMR titration experiments of labeled CaM showed binding of a peptide corresponding to the AQP0 C-terminus with a 1:2 stoichiometry, confirmed by the EM structure of the AQP0-CaM complex published in 2013 [148,150]. MD simulations show an allosteric mechanism that links the CaM binding to movement of residues in the second constriction site, ultimately closing the channel [148,151]. Altogether, this led to a model where CaM binds the C-termini of two neighboring AQP monomers simultaneously in an antiparallel fashion (**Figure 18**). This is an unusual arrangement represented by only 1% of the structures available of CaM complexes available in the protein data bank [106].

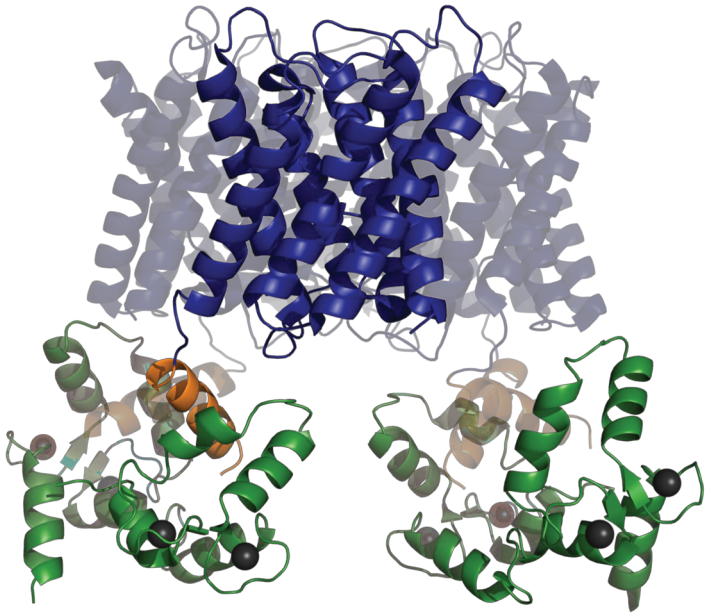


Figure 18*. AQP0-CaM model based on EM (PDB ID: 3J41)* [151]*. Two calmodulins (green) bind to the C-termini (orange) of AQP0 (blue). This interaction is depending on* Ca^{2+} *(grey spheres) binding to calmodulin.*

One of the future challenges in the field of structural biology is to understand the mechanisms of protein interactions. We developed a purification strategy to purify intact protein complexes with the benefit of the characterization and localization confirmation *in vivo*.

In **Paper V** the complex formation of hAQP0 and CaM was observed in *S. cerevisiae* using BiFC. The constructs were based on two halves of YFP (YFP_N and YFP_C) fused to the target proteins in different combinations. Positive controls showing hAQP0 tetramerization confirm correct localization to the membrane and that the BiFC complexes develop fluorescence as a result of the protein interaction with signals similar to an YFP-hAQP0 fusion (**Figure 19**). A BiFC pair of hAQP0 and hAQP2 cause a loss in fluorescence consistent with a decrease in tetramer formation. The combination with YFP_C fused to hAQP0 showed no fluorescence while the other combination, YFP_N-hAQP0 + YFP_C-hAQP2, did show some fluorescence, indicating that there is an interaction between hAQP0 and hAQP2. hAQP2 was selected as the negative control of hAQP0 tetramerization based on their high sequence (60%) and the different tissue distribution in the human body. The fluorescence of the

YFP_N-hAQP0 + YFP_C-hAQP2 complex suggest that there is a hAQP0-hAQP2 complex forming. hAQP2 has previously been shown to interact with hAQP5, indicating a role for aquaporin regulation by each other [152]. In plants, AQPs have been confirmed to interact [153,154] and heterotetramerization has been suggested to regulate the trafficking of some plant aquaporins [155].

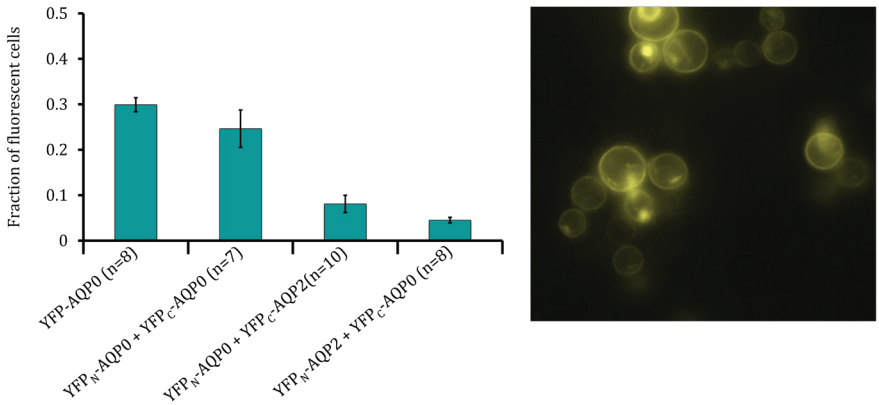


***Figure 19**. **Aquaporins form tetrameric complexes in S. cerevisiae. A)** In vivo fluorescence of YFP-hAQP0 and a tetramerization control (YFP_N-hAQP0+YFP_C-hAQP0) were tested. As negative control we combined hAQP0 and hAQP2 in two ways where YFP_N-AQP0+YFP_C-hAQP2 showed some fluorescence. B) Microscopy image of YFP-hAQP0. Figure adapted from **Paper V**.*

The complex formation of AQP0 and calmodulin was also investigated. For AQP0-CaM, fluorescence was observed when YFP_N was fused to the N-terminus of hAQP0 (**Figure 20**). For all the cases reported, fusing YFP_C to hAQP0 always resulted in a loss of fluorescence, possibly caused by steric hindrance where the YFP fragments consistently end up too far away from each other. Truncation of the hAQP0 C-terminus (resulting in hAQP0Δ) removed the CaM interaction site, leading to an efficient and specific loss of fluorescence as the complex formation was impaired (**Figure 20**). BiFC locks the interacting proteins by linking them via the YFP. The YFP_N-hAQP0Δ + YFP_C-CaM combination demonstrate that the fluorescence is a result of the protein interaction and not the other way around.

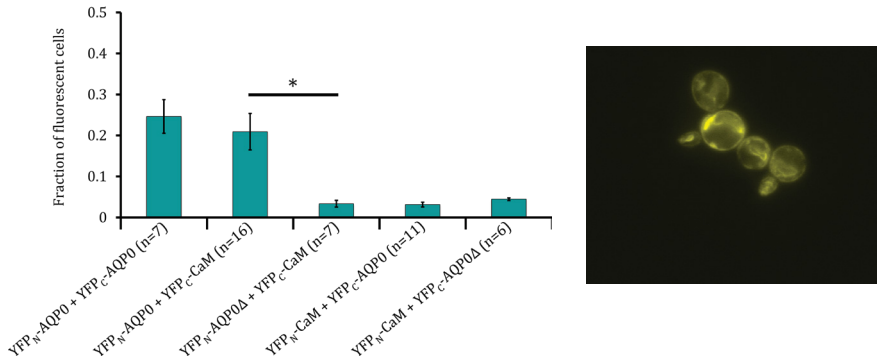


Figure 20. hAQP0 and CaM form a complex in vivo*. A) YFP$_N$-hAQP0 + YFP$_C$-CaM form a complex in vivo. Truncation of hAQP0 at residue 221 removes the CaM binding site and result in a complete loss of fluorescence.* ***B)*** *Microscopy image of YFP$_N$-hAQP0 + YFP$_C$-CaM cells. Figure adapted from* ***Paper V****.*

By the emergence of fluorescence, we can observe the complex formation of hAQP0-CaM using BiFC in *S. cerevisiae* (**Figure 21**). The membrane localization of the fluorescence signal indicate that the YFP fragments do not interfere with the production and translocation of the proteins. To understand the interaction on a molecular level the complex was purified, taking advantage of YFP to hold hAQP0 and CaM together in stoichiometric ratios in addition to acting as a tracing probe during all the purification steps.

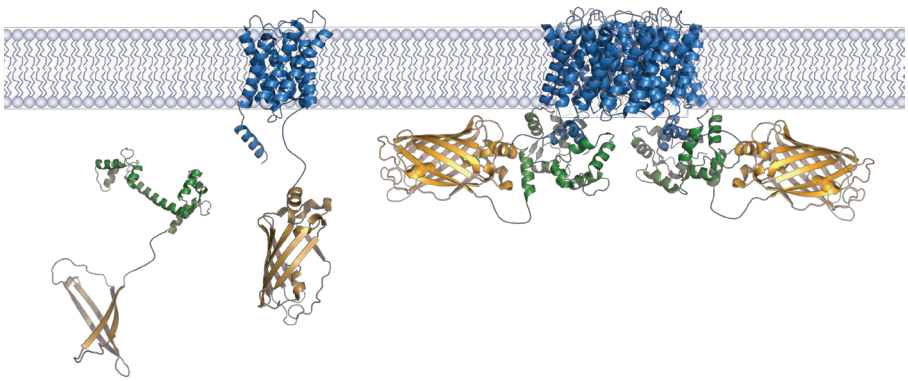


Figure 21. Schematic figure of the hAQP0-YFP-CaM complex assembly. *YFP$_N$-hAQP0 (blue) is located in the membrane while YFP$_C$-CaM (green) is located in the cytoplasm. Together they form the hAQP0-YFP-CaM complex.*

4.4.2 Purification of protein complexes

Using X-ray crystallography we want to understand the interaction between hAQP0 and CaM on a molecular level. The complex was purified using standard protocols including ion exchange and size exclusion chromatography with fluorescence detection (**Figure 22**). Recombinant hAQP0 is known to form octamers during the purification [139], which correlate well with the retention volume seen for the major peaks in the size exclusion chromatography (SEC).

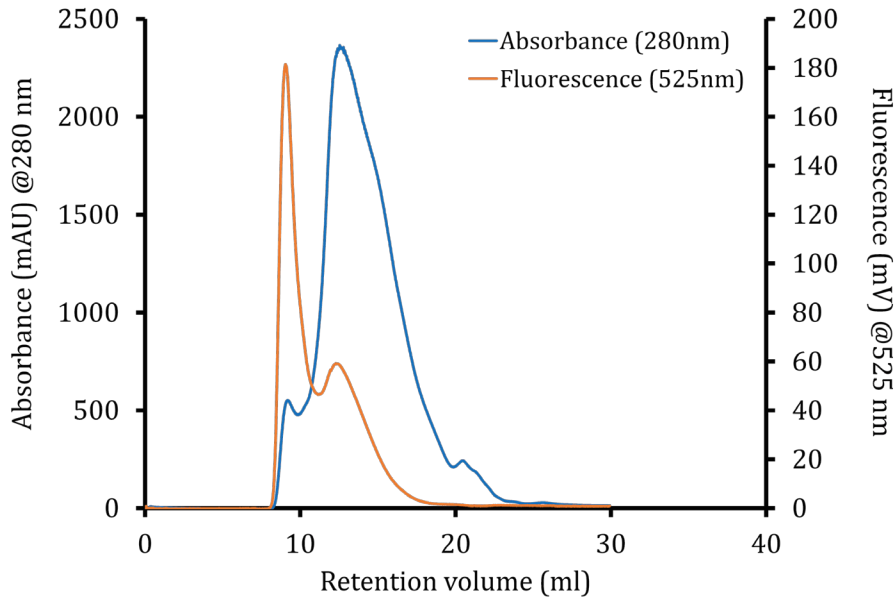


Figure 22. Purification of the YFP_N-hAQP0 + YFP_C-CaM complex show a peak at a retention volume consistent with an octameric complex.

Analysis of the purified complex on native-PAGE using immunodetection with antibodies specific for the his-tag and CaM, respectively, show a co-localization of the two signals (**Figure 23**). The band appears close to 500 kDa, confirming the purification of the full hAQP0-CaM-YFP complex (expected molecular weight of an octameric complex is 495 kDa). There is also a substantial shift in comparison to the positive control, purified hAQP0 (octamer expected at 230 kDa).

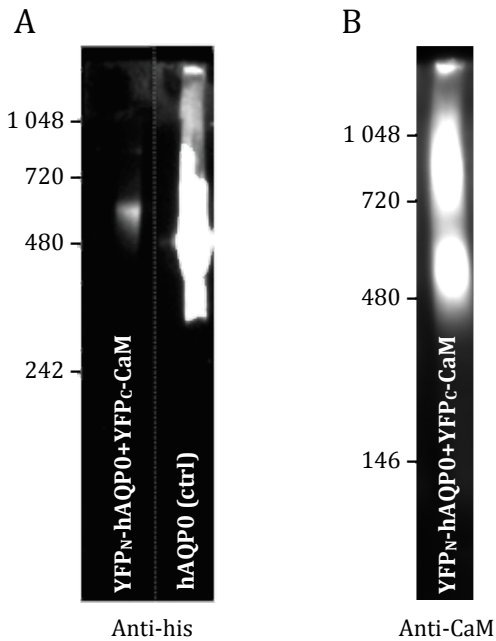


Figure 23. hAQP0 and CaM are both detected in a band around 500kDa, indicating an intact complex. A) *Anti-his (detecting the YFP_N-hAQP0 fusion) show a signal larger than 480 kDa. Purified hAQP0 is loaded as a reference for size.* ***B)*** *Anti-CaM western also show a signal larger than 480 kDa. Together, this is a strong indication of an intact YFP_N-hAQP0+YFP_C-CaM complex. Figure adapted from* ***Paper V.***

Using BiFC in protein complex purification aiming at crystallization has several benefits. First, the complex formation *in vivo* confirms that the protein targets localize to their native cellular compartment. This can be compared to PPI identified by immunoprecipitation where a large number of the false positives is a direct result of the method of preparation. Since immunoprecipitation is based on cell lysate all proteins in the cell, irrespective of localization, are able to interact and can be detected.

Second, the mature fluorescent complex can be traced throughout the solubilization and purification steps. As with the conventional GFP tag, the optimal conditions for each purification step can easily be determined using fluorescence detection.

Third, during crystallization the YFP provide an increased hydrophilic surface beneficial for making crystal contacts. In the case of aquaporins and other membrane proteins with small soluble domains this can prove

important, as shown by the use of GFP previously [156]. By crystallizing the full complex the termini or other flexible regions that are part of the interaction site will be stabilized allowing them to be visible in the structure. Further, depending on the size of the complex, YFP can be used as an initial search model to obtain the phases for molecular replacement. Structural biology will benefit from revealing the molecular mechanisms of protein:protein interactions. The combination of BiFC and fluorescence detection provide a promising tool for producing and purifying protein complexes to crystallization grade quality in sufficient yields.

4.5 Future directions for structure determination of protein complexes

Most methods for protein interaction studies have problems including full length membrane proteins due to their high hydrophobicity. With the aquaporin C-terminus being the main site for protein interactions (**Paper IV**), a solution has been to generate peptides corresponding to the terminal sequences. Using these peptides as bait it is possible to get around the problems associated with membrane proteins in the interaction assays. BiFC allows for full-length proteins to be used and since the complex formation takes place *in vivo* only proteins that localize to the same cellular compartment can interact (**Paper V**). In addition, the fluorescence resulting from the complex formation of the target proteins makes it possible to confirm their proper localization using fluorescence microscopy, an especially important quality control in the case of membrane proteins [157]. The fluorescence of YFP, just like GFP, needs a reducing environment in order to develop, limiting the use to the intracellular environment [158,159]. As a consequence, this particular BiFC system can only be applied to targets where the interaction site is intracellular. However, there are GFP variants that retains their folding when secreted to the extracellular space [160] which could be adapted to BiFC to cover virtually every interaction in the cell.

Another consideration when working with protein complexes is the molecular weight. Although X-ray crystallography is the dominating method for structure determination of membrane proteins, other methods may be worth considering. Knowing that hAQP0 has a high tendency to form octamers, interacting via the extracellular loops, the hAQP0-CaM complex linked by YFP (495 kDa) could be considered for cryo-electron microscopy (cryo-EM). The combination of BiFC and cryo-

EM could be used to determine structures of large protein complexes held together by low affinity interactions.

The ongoing advances in cryo- EM, casually referred to as the "EM-revolution" due to its rapid development, has improved the resolution of the recovered protein complex structures immensely during the past 25 years, and even more so the past 2 years [161,162]. Single particle cryo-EM effectively bypass the process of crystallization and hence the need for large amounts of protein. As the protein is present in a liquid environment upon freezing, the hydration layer around the protein remains intact and the protein environment is more similar to a native state [163]. The initial limitations to study only very large protein complexes is now starting relax with exciting structures also emerging for smaller protein complexes of 200-300 kDa. To date, the highest resolution obtained with the method was the structure determined of a β-galactosidase tetrameric complex (1.8 MDa) at 2.2 Å [164] while human γ-secretase (170 kDa) is the smallest protein complex to be structurally determined by cryo-EM to 3.4 Å [165]. For a successful EM experiment, a structurally homogenous sample is more important than protein purity. Similarly to X-ray crystallography, regions displaying high degrees of random movement are difficult to resolve but, in contrast, ordered conformational changes can sometimes be visualized. The trapping of conformational states by cryo-EM and fitting the density with existing atomic models can be used to define large structural changes [166].

5 Concluding remarks

Aquaporins, the cellular water channels, are more than pores in the membrane: they facilitate water transport at the rate of diffusion with high specificity. The 13 isoforms found in the human body display a wide variation in tissue distribution, cellular localization, regulation mechanism and specificity. Aquaporins have been shown to transport water, glycerol, urea, anions, cations and even CO_2. The complexity of the function and regulation of these proteins increase as we learn more about them.

Bottlenecks associated with structural determination of proteins get narrower when dealing with membrane protein. Robust strategies to improve yields and protein stability are necessary to achieve the first goal: sufficient amounts of protein for functional and structural characterization.

The aquaporin protein family was used as a model system for investigating general and specific means of increasing the obtained protein yields. We clearly show that codon optimization and transformation method can be adapted as general methods to obtain higher yields of this class of membrane proteins. This effect was more pronounced for proteins produced at low to moderate levels. hAQP4, an aquaporin found in the brain, was previously not detectable when overproduced in *P. pastoris*. With knowledge of how rational mutagenesis increased the yield and together with the codon optimization presented here, we now have a hAQP4 mutant with wildtype characteristics and increased stability that can be subject to a large number of experiments.

Post-translational modifications can influence protein behavior and are known to influence trafficking of aquaporins. **Paper III** provided new insights to the trafficking of hAQP5. These experiments, performed in HEK293 cells, led us to the conclusion that a number of mechanisms act together to determine the final membrane abundance of hAQP5. Based on the independent nature of the PKA inhibition, it is possible that some regulatory mechanism is mediated via protein:protein interactions. Also, the behavior of hAQP5-GFP seem to differ between cell lines when comparing HEK293 and MDCK cells. It is always important to keep in mind that while heterologous cell based assays are a convenient way of studying protein regulation, the results must be correlated to information from *in vivo* observations.

Glycosylation has been shown to affect the stability of hAQP10 *in vitro*, but has so far not been seen to have any effect on aquaporins *in vivo*. Since glycosylations are not encoded by the genome but a result of enzymatic reactions, they introduce heterogeneity into the protein population that could be relevant for modulating protein function.

Protein:protein interactions are critical for cellular function and thus important for our understanding of the working mechanisms in the cell. We combined established methods for investigating PPIs (BiFC) and protein purification (fluorescence detection) to achieve a promising tool to purify intact complexes where the constituting proteins have low affinity for each other. The large flexibility of the method will potentially provide the means to study a wide variety of protein targets.

Structural biology has so far given us a number of revelations regarding protein function and mechanisms of action. To fully understand the role of proteins within the cell, structures of protein complexes will be essential. The collective knowledge of crystallizing difficult targets, including membrane proteins, and new developments in technology will all be important in determining high resolution structures of protein complexes. Combining BiFC with protein X-ray crystallography or cryo-EM can potentially cover a wide range of complexes with respect to affinity and size. The recent advances presented here lay technical foundations for the structural characterization of novel targets and the elucidation of protein complexes are expected to be one of the new frontiers of structural biology.

6 Future perspectives

Membrane proteins are key players in many biological processes and crucial targets of drugs and antibiotics. Understanding the protein structure, function and interactions with other proteins are of great importance in developing new drugs for existing and novel targets. The ongoing unraveling of aquaporin regulation and function is far from complete and many exciting discoveries are still waiting to be made.

For the aquaporins, the variations seen in specificity between isoforms are most likely based on structural differences. It would be interesting to understand how these structural changes ultimately affect the fine-tuning of aquaporin function. Some aquaporins have very special, and for the protein, essential variations in the NPA motif and the effect on function is an area that has not received much attention so far.

There is a large number of interaction partners found that affect the aquaporin function and regulation. New methods to study and structurally determine these protein complexes will provide us with a way to answer novel scientific questions.

7 Acknowledgements

You're holding in your hands the result of 5 years of graduate studies (2008 days!). It's a weird feeling to write the acknowledgements and an even more unreal feeling of reaching the finishing line (almost). There's a lot of people to thank for this time:

Thank you, **Richard**, for accepting me as your PhD student and giving me the opportunity to do cool science in a creative environment (not to mention creative meeting environments (!)). Your ideas are never ending. I've learned a lot of science, the fundamental rules on how to handle your supervisor and how to pick the weirdest conferences☺. It's been a great time!

Kristina, ever since the master thesis you have guided me through cloning and overproduction and everything else. Our discussions have always been useful and I've really enjoyed it! I hope we can bring another long-lived project (10:an) to a conclusion soon!

Gisela, thanks for all the encouragement especially now in the end, you'll be a great group leader!

Susanna, for resurrecting that long lost AQP5 manuscript, with a lot of help from the UK crowd:**Ros, Matt**, **Alex**, **Rebecca** and **Phil**, you really saved us in Romania!

Thanks to **Martin** for being my examiner.

Ida, the Goddess of both programming and heavy lifts, for making the lab such a nice place and for being good at being good ☺! It will be so weird not seeing you every day. **Annette**, for all the discussions about cucumbers and cows!

Rebecka, the loudest room member I had, but also the funniest! You are truly unique! Thanks for all the good and crazy times! **Alex,** for practicing pedagogics while sharing your best tips and tricks with Matlab or just going to the gym, keep working that treadmill! **Cecilia**, we really connected during the Japan trip, thanks for bringing me! You are now what remain of the *Pichia* legacy in the corridor, don't let that crush you ☺. **Petra**, the human FACS machine! Call me when you hit the $1{\cdot}10^{10}$ in the cell counting. I can't thank you enough for pushing so hard these last months! Thanks **Rajiv** for being the nicest person ever on the nozzle making week in Hamburg, good luck with the braaainz. **Greger**, I'll call you if I ever have to use a fax machine!

The grey hairs are spreading like wildfire in our group and I think we're all going to be infested before our defenses. **David**, thanks for all the help with the data analysis way back when, which is probably the biggest reason for grey hairs in my case. Unfortunately, the hair color doesn't go back to normal when MatLab is finally cooperating ;) **Amit**, thank you for very pedagogic way of explaining even the basics of WAXS analysis and for your help during the WAXS data collections. **Peter**, for enthusiasm and great discussions on science and life in general, it'll be awesome to join you in Oz!

My bodyguard **Rosie**, thanks for the many late dinner dates and discussions on void peaks. Thanks for being you and bringing so much positive energy to the lab, I'll really miss it!

Old Rob, take care of the apartment! **New Rob**, congratulations to the fancy-colored-protein project!

Proud member of the 3.5 Å club since 2015 (= me and **Rhawnie**)

Many thanks to **Viktor**, **Niada** and **Ehab** for all the effort put into the projects during your master projects!

To all the rest of the people in the lab: **Oskar**, **Ashley**, **Weixiao, Maria**, **Stephan**, **Alex**, **Jiao, Elin C, Elin D, Cecilia, Parveen, Heikki**, **Rafal**, **Sebastian, Gergely** for making the corridor an international and enjoyable place to work and having a lot of input on laser-related issues! **Emil**, keep the grill going! Every day is a BBQ day! **Petra**, I still remember our great cooperation on UbOx back in the day, it was great fun! No fermentor room accidents for us ;). **MaJo**, you must be the most efficient traveler ever. I'm impressed with your pace!

Anna, Gerhard and Mikael for the introduction to PhD life and for all the information I only understand now in the end. Thanks to **Fredrik** for sharing an always tidy lab bench and showing me a whole new level of organization. And for keeping so good track of the data that I can get my hands on it 5 years later!

Ann, **Susanne**, **Helena** and **Desiree**, thanks for all the help with everything related to administration! **Lars**, **Bruno** and **Valida**. The lab simply doesn't run without you!

Ågot, vi har gjort några riktigt grymma insatser på jympagolvet/bryggan, det har varit superkul att dela det och allt doktorandrelaterat med dig!

Tack till marsvinen såklart**: Newton**, **Neo**, **Teodor**, **Dexter**, **Sheldon**, **Paprika** och **Kiwi** som (oftast) fått mig att åka hem i (någorlunda) tid.

Tack **Susanne** för att du alltid är så underbar! Vi borde äta mycket mer glass ihop! Charmigaste **Johan** och **Anna** (och **Maj** såklart, även om vi inte haft så djupa diskussioner än), för mysiga glöggkvällar och träningspass.

Mamma, **Pappa**, **Emma**, **Josefin**, **Novali**, **mormor**, **farmor**, **farfar**: denna är till er, ! Tack för allt stöd, nu och alltid.

Morfar, jag är ett steg närmare!

Mike, du känner mig bättre än vad jag känner mig själv. Du pushar mig att göra det jag egentligen vill men inte alltid vågar. Nya utmaningar väntar och det ska bli spännande att upptäcka nya delar av världen med dig. TO THE ISLAND! Jag älskar dig!

Now… Bring me that horizon

8 References

Figures 9 and 11 were reprinted from **Paper I** with permission from Taylor and Francis publishing group (www.tandfonline.com)
Figures 14-16 was reprinted from **Paper III** with permission from PLOS One.

1 Luckey M (2008) *Membrane structural biology with biochemical and biophysical foundation*, 1st ed. Cambridge University Press.
2 Dufourc EJ (2008) Sterols and membrane dynamics. *J. Chem. Biol.* **1**, 63–77.
3 Koshy C & Ziegler C (2014) Structural insights into functional lipid-protein interactions in secondary transporters. *Biochim. Biophys. Acta - Gen. Subj.* **1850**, 476–487.
4 Wallin E & von Heijne G (1998) Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci.* **7**, 1029–1038.
5 Von Heijne G (2007) The membrane protein universe: What's out there and why bother? *J. Intern. Med.* **261**, 543–557.
6 Russell RB & Eggleston DS (2000) New roles for structure in biology and drug discovery. *Nat. Struct. Biol.* **7 Suppl**, 928–930.
7 http://blanco.biomol.uci.edu/mpstruc/
8 Bill RM, Henderson PJF, Iwata S, Kunji ERS, Michel H, Neutze R, Newstead S, Poolman B, Tate CG & Vogel H (2011) Overcoming barriers to membrane protein structure determination. *Nat. Biotechnol.* **29**, 335–340.
9 Alleva K, Chara O & Amodeo G (2012) Aquaporins: Another piece in the osmotic puzzle. *FEBS Lett.* **586**, 2991–2999.
10 Macey RI, Karan DM & Farmer RE (1972) Properties of water channels in human red cells. *Biomembranes* **3**, 331–340.
11 Preston GM, Carroll TP, Guggino WB & Agre P (1992) Oocytes Expressing. , 26–28.
12 www.nobelprize.org
13 Kaldenhoff R, Kai L & Uehlein N (2014) Aquaporins and membrane diffusion of CO2 in living organisms. *Biochim. Biophys. Acta - Gen. Subj.* **1840**, 1592–1595.
14 Li G, Santoni V & Maurel C (2014) Plant aquaporins: Roles in plant physiology. *Biochim. Biophys. Acta - Gen. Subj.* **1840**, 1574–1582.
15 Herrera M & Garvin JL (2011) Aquaporins as gas channels. *Pflugers Arch. Eur. J. Physiol.* **462**, 623–630.
16 Yu J, Yool AJ, Schulten K & Tajkhorshid E (2006) Mechanism of Gating and Ion Conductivity of a Possible Tetrameric Pore in Aquaporin-1. *Structure* **14**, 1411–1423.
17 Fu D & Lu M (2007) The structural basis of water permeation and proton exclusion in aquaporins. *Mol. Membr. Biol.* **24**, 366–374.
18 Sui H, Han BG, Lee JK, Walian P & Jap BK (2001) Structural basis of water-specific transport through the AQP1 water channel. *Nature* **414**, 872–878.
19 Beitz E, Wu B, Holm LM, Schultz JE & Zeuthen T (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**, 269–274.
20 de Grotthuss CJT (1806) Sur la décomposition de l'eau et des corps qu'elle tient en dissolution à l'aide de l'électricité galvanique. *Ann. Chim.* **58**, 54–73.
21 Choi J-H, Lee H-J & Moon S-H (2001) Effects of Electrolytes on the Transport Phenomena in a Cation-Exchange Membrane. *J. Colloid Interface Sci.* **238**, 188–195.
22 Fischer G, Kosinska-Eriksson U, Aponte-Santamaría C, Palmgren M, Geijer C, Hedfalk K, Hohmann S, de Groot BL, Neutze R & Lindkvist-Petersson K (2009) Crystal structure of a yeast aquaporin at 1.15 angstrom reveals a novel gating mechanism. *PLoS Biol.* **7**, e1000130.
23 Kosinska Eriksson U, Fischer G, Friemann R, Enkavi G, Tajkhorshid E & Neutze R (2013) Subangstrom Resolution X-Ray Structure Details Aquaporin-Water Interactions. *Science (80-.).* **340**, 1346–1349.

24 de Groot BL & Grubmüller H (2001) Water permeation across biological membranes: mechanism and dynamics of aquaporin-1 and GlpF. *Science* **294**, 2353–2357.
25 King LS, Kozono D & Agre P (2004) From structure to disease: the evolving tale of aquaporin biology. *Nat. Rev. Mol. Cell Biol.* **5**, 687–98.
26 Ishibashi K (2009) New members of mammalian aquaporins: AQP10-AQP12. *Handb. Exp. Pharmacol.* **190**, 251–262.
27 Castle N (2005) Aquaporins as targets for drug discovery. *Drug Discov. Today* **10**, 485–93.
28 Carbrey JM & Agre P (2009) Discovery of the aquaporins and development of the field. *Handb. Exp. Pharmacol.* **190**, 3–28.
29 Yasui M, Hazama a, Kwon TH, Nielsen S, Guggino WB & Agre P (1999) Rapid gating and anion permeability of an intracellular aquaporin. *Nature* **402**, 184–187.
30 Morishita Y, Matsuzaki T, Hara-chikuma M, Andoo A, Shimono M, Matsuki A, Kobayashi K, Ikeda M, Yamamoto T, Verkman A, Kusano E, Ookawara S, Takata K, Sasaki S & Ishibashi K (2005) Disruption of aquaporin-11 produces polycystic kidneys following vacuolization of the proximal tubule. *Mol. Cell. Biol.* **25**, 7770–7779.
31 Itoh T, Rai T, Kuwahara M, Ko SBH, Uchida S, Sasaki S & Ishibashi K (2005) Identification of a novel aquaporin, AQP12, expressed in pancreatic acinar cells. *Biochem. Biophys. Res. Commun.* **330**, 832–8.
32 Öberg F, Ekvall M, Nyblom M, Backmark A, Neutze R & Hedfalk K (2009) Insight into factors directing high production of eukaryotic membrane proteins; production of 13 human AQPs in Pichia pastoris. *Mol. Membr. Biol.* **26**, 215–27.
33 Alcalá J, Lieska N & Maisel H (1975) Protein composition of bovine lens cortical fiber cell membranes. *Exp. Eye Res.* **21**, 581–595.
34 Chandy G, Zampighi GA, Kreman M & Hall JE (1997) Comparison of the water transporting properties of MIP and AQP1. *J. Membr. Biol.* **159**, 29–39.
35 Liu J, Xu J, Gu S, Nicholson BJ & Jiang JX (2011) Aquaporin 0 enhances gap junction coupling via its cell adhesion function and interaction with connexin 50. *J. Cell Sci.* **124**, 198–206.
36 Mathias RT, White TW & Gong X (2010) Lens gap junctions in growth, differentiation, and homeostasis. *Physiol. Rev.* **90**, 179–206.
37 Gonen T, Cheng Y, Kistler J & Walz T (2004) Aquaporin-0 membrane junctions form upon proteolytic cleavage. *J. Mol. Biol.* **342**, 1337–45.
38 Németh-Cahalan KL, Kalman K & Hall JE (2004) Molecular basis of pH and Ca2+ regulation of aquaporin water permeability. *J. Gen. Physiol.* **123**, 573–80.
39 Varadaraj K, Kumari S, Shiels A & Mathias RT (2005) Regulation of aquaporin water permeability in the lens. *Invest. Ophthalmol. Vis. Sci.* **46**, 1393–402.
40 Ball LE (2003) Water Permeability of C-Terminally Truncated Aquaporin 0 (AQP0 1-243) Observed in the Aging Human Lens. *Invest. Ophthalmol. Vis. Sci.* **44**, 4820–4828.
41 Tong J, Canty JT, Briggs MM & McIntosh TJ (2013) The water permeability of lens aquaporin-0 depends on its lipid bilayer environment. *Exp. Eye Res.* **113**, 32–40.
42 Tong J, Briggs MM & McIntosh TJ (2012) Water permeability of aquaporin-4 channel depends on bilayer composition, thickness, and elasticity. *Biophys. J.* **103**, 1899–1908.
43 Raina S, Preston GM, Guggino WB & Agre P (1995) Molecular cloning and characterization of an aquaporin cDNA from salivary, lacrimal, and respiratory tissues. *J. Biol. Chem.* **270**, 1908–1912.
44 Delporte C & Steinfeld S (2006) Distribution and roles of aquaporins in salivary glands. *Biochim. Biophys. Acta - Biomembr.* **1758**, 1061–1070.
45 Vitali C, Palombi G & Cataleta P (2010) Treating Sjögren's Syndrome: Insights for the Clinician. *Ther. Adv. Musculoskelet. Dis.* **2**, 155–66.
46 Matsuzaki T, Susa T, Shimizu K, Sawai N, Suzuki T, Aoki T, Yokoo S & Takata K (2012) Function of the Membrane Water Channel Aquaporin-5 in the Salivary Gland. *Acta Histochem. Cytochem.* **45**, 251–259.
47 Steinfeld S, Cogan E, King LS, Agre P, Kiss R & Delporte C (2001) Abnormal distribution of aquaporin-5 water channel protein in salivary glands from Sjögren's syndrome patients. *Lab. Invest.* **81**, 143–148.
48 Horsefield R, Nordén K, Fellert M, Backmark A, Törnroth-Horsefield S, Terwisscha van Scheltinga AC, Kvassman J, Kjellbom P, Johanson U & Neutze R (2008) High-resolution x-ray structure of human aquaporin 5. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 13327–32.

49 Frick A, Eriksson UK, de Mattia F, Oberg F, Hedfalk K, Neutze R, de Grip WJ, Deen PMT & Törnroth-Horsefield S (2014) X-ray structure of human aquaporin 2 and its implications for nephrogenic diabetes insipidus and trafficking. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 1–6.
50 Ramesh Prasad G V., Coury L a., Finn F & Zeidel ML (1998) Reconstituted aquaporin 1 water channels transport CO2 across membranes. *J. Biol. Chem.* **273**, 33123–33126.
51 Wang Y, Cohen J, Boron WF, Schulten K & Tajkhorshid E (2007) Exploring gas permeability of cellular membranes and membrane channels with molecular dynamics. *J. Struct. Biol.* **157**, 534–544.
52 Fang X, Yang B, Matthay M a & Verkman a S (2002) Evidence against aquaporin-1-dependent CO2 permeability in lung and kidney. *J. Physiol.* **542**, 63–69.
53 Hatakeyama S, Yoshida Y, Tani T, Koyama Y, Nihei K, Ohshiro K, Kamiie JI, Yaoita E, Suda T, Hatakeyama K & Yamamoto T (2001) Cloning of a new aquaporin (AQP10) abundantly expressed in duodenum and jejunum. *Biochem. Biophys. Res. Commun.* **287**, 814–9.
54 Ishibashi K, Morinaga T, Kuwahara M, Sasaki S & Imai M (2002) Cloning and identification of a new member of water channel (AQP10) as an aquaglyceroporin. *Biochim. Biophys. Acta* **1576**, 335–40.
55 Laforenza U, Scaffino MF & Gastaldi G (2013) Aquaporin-10 Represents an Alternative Pathway for Glycerol Efflux from Human Adipocytes. *PLoS One* **8**.
56 Madeira A, Moura TF & Soveral G (2014) Aquaglyceroporins: implications in adipose biology and obesity. *Cell. Mol. Life Sci.* **72**, 759–771.
57 Jungersted JM, Bomholt J, Bajraktari N, Hansen JS, Klærke D a., Pedersen PA, Hedfalk K, Nielsen KH, Agner T & Hélix-Nielsen C (2013) In vivo studies of aquaporins 3 and 10 in human stratum corneum. *Arch. Dermatol. Res.* **305**, 699–704.
58 Beitz E (2000) TEXtopo: shaded membrane protein topology plots in LATEX2ε. *Bioinformatics* **16**, 1050–1051.
59 Niess JH & Reinecker HC (2006) Dendritic cells in the recognition of intestinal microbiota. *Cell. Microbiol.* **8**, 558–564.
60 Helander HF & Fändriks L (2014) Surface area of the digestive tract - revisited. *Scand. J. Gastroenterol.* **49**, 681–9.
61 Mobasheri a, Shakibaei M & Marples D (2004) Immunohistochemical localization of aquaporin 10 in the apical membranes of the human ileum: a potential pathway for luminal water and small solute absorption. *Histochem. Cell Biol.* **121**, 463–71.
62 Ullmann A, Jacob F & Monod J (1967) Characterization by in vitro complementation of a peptide corresponding to an operator-proximal segment of the β-galactosidase structural gene of. *J. Mol. Biol.*, 339–343.
63 Pelletier JN, Campbell-Valois FX & Michnick SW (1998) Oligomerization domain-directed reassembly of active dihydrofolate reductase from rationally designed fragments. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 12141–6.
64 Galarneau A, Primeau M, Trudeau L-E & Michnick SW (2002) Beta-lactamase protein fragment complementation assays as in vivo and in vitro sensors of protein protein interactions. *Nat. Biotechnol.* **20**, 619–622.
65 Villalobos V, Naik S & Piwnica-Worms D (2007) Current state of imaging protein-protein interactions in vivo with genetically encoded reporters. *Annu. Rev. Biomed. Eng.* **9**, 321–349.
66 Wehr MC, Reinecke L, Botvinnik A & Rossner MJ (2008) Analysis of transient phosphorylation-dependent protein-protein interactions in living mammalian cells using split-TEV. *BMC Biotechnol.* **8**, 55.
67 Hu JC, Kornacker MG & Hochschild A (2000) Escherichia coli One- and Two-Hybrid Systems for the Analysis and Identification of Protein – Protein Interactions. *Pharm. Res.* **94**, 80–94.
68 Stelzl U, Worm U, Lalowski M, Haenig C, Brembeck FH, Goehler H, Stroedicke M, Zenkner M, Schoenherr A, Koeppen S, Timm J, Mintzlaff S, Abraham C, Bock N, Kietzmann S, Goedde A, Toksöz E, Droege A, Krobitsch S, Korn B, Birchmeier W, Lehrach H & Wanker EE (2005) A human protein-protein interaction network: A resource for annotating the proteome. *Cell* **122**, 957–968.
69 Rual J-F, Venkatesan K, Hao T, Hirozane-Kishikawa T, Dricot A, Li N, Berriz GF, Gibbons FD, Dreze M, Ayivi-Guedehoussou N, Klitgord N, Simon C, Boxem M, Milstein S, Rosenberg J, Goldberg DS, Zhang L V, Wong SL, Franklin G, Li S, Albala JS, Lim J, Fraughton C, Llamosas E, Cevik S, Bex C, Lamesch P, Sikorski RS, Vandenhaute J, Zoghbi HY, Smolyar A, Bosak S,

Sequerra R, Doucette-Stamm L, Cusick ME, Hill DE, Roth FP & Vidal M (2005) Towards a proteome-scale map of the human protein-protein interaction network. *Nature* **437**, 1173–1178.
70 Shimomura O, Johnson FH & Saiga Y (1962) Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. *J. Cell. Comp. Physiol.* **59**, 223–239.
71 http://book.bionumbers.org/what-is-the-maturation-time-for-fluorescent-proteins/
72 Cormack B, Valdivia R & Falkow S (1996) FACS-optimized mutants of the green fluorescent protein (GFP). *Elsevier* **173**, 33–38.
73 Heim R, Prasher DC & Tsien RY (1994) Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc Natl Acad Sci U S A* **91**, 12501–12504.
74 Zacharias D a, Violin JD, Newton AC & Tsien RY (2002) Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science* **296**, 913–916.
75 Ai H, Olenych SG, Wong P, Davidson MW & Campbell RE (2008) Hue-shifted monomeric variants of Clavularia cyan fluorescent protein: identification of the molecular determinants of color and applications in fluorescence imaging. *BMC Biol.* **6**, 13.
76 Subach OM, Gundorov IS, Yoshimura M, Subach F V., Zhang J, Grüenwald D, Souslova E a., Chudakov DM & Verkhusha V V. (2008) Conversion of Red Fluorescent Protein into a Bright Blue Probe. *Chem. Biol.* **15**, 1116–1124.
77 Nagai T, Ibata K, Park ES, Kubota M, Mikoshiba K & Miyawaki A (2002) A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat. Biotechnol.* **20**, 87–90.
78 Kremers G-J, Goedhart J, van Munster EB & Gadella TWJ (2006) Cyan and yellow super fluorescent proteins with improved brightness, protein folding, and FRET Förster radius. *Biochemistry* **45**, 6570–80.
79 Hu C-D & Kerppola TK (2003) Simultaneous visualization of multiple protein interactions in living cells using multicolor fluorescence complementation analysis. *Nat. Biotechnol.* **21**, 539–45.
80 Kodama Y & Hu C-D (2012) Bimolecular fluorescence complementation (BiFC): a 5-year update and future perspectives. *Biotechniques* **53**, 285–98.
81 Hu C-D, Chinenov Y & Kerppola TK (2002) Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. *Mol. Cell* **9**, 789–98.
82 Ghosh I, Hamilton AD, Regan L, V YU & V NH (2000) Antiparallel Leucine Zipper-Directed Protein Reassembly : Application to the Green Fluorescent Protein Department of Molecular Biophysics and Biochemistry The dissection and subsequent reassembly of a protein from peptidic fragments provides an avenue for. , 5658–5659.
83 Guo Y, Rebecchi M & Scarlata S (2005) Phospholipase C 2 Binds to and Inhibits Phospholipase C 1. *J. Biol. Chem.* **280**, 1438–1447.
84 Yu J, Xiao J, Ren X, Lao K & Xie XS (2006) Probing gene expression in live cells, one protein molecule at a time. *Science* **311**, 1600–1603.
85 He Y, Wang K & Yan N (2014) The recombinant expression systems for structure determination of eukaryotic membrane proteins. *Protein Cell* **5**, 658–672.
86 Freigassner M, Pichler H & Glieder A (2009) Tuning microbial hosts for membrane protein production. *Microb. Cell Fact.* **8**, 69.
87 Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H & Oliver SG (1996) Life with 6000 Genes. *Science (80-.).* **274**, 546–567.
88 Barnett JA & Barnett L (2011) *Yeast research: a historical overview*, 1st ed. ASM Press.
89 Öberg F & Hedfalk K (2012) Recombinant production of the human aquaporins in the yeast Pichia pastoris. *Mol. Membr. Biol.* **30**, 1–17.
90 De Schutter K, Lin Y-C, Tiels P, Van Hecke A, Glinka S, Weber-Lehmann J, Rouzé P, Van de Peer Y & Callewaert N (2009) Genome sequence of the recombinant protein production host Pichia pastoris. *Nat. Biotechnol.* **27**, 561–566.
91 Ribet D & Cossart P (2010) Post-translational modifications in host cells during bacterial infection. *FEBS Lett.* **584**, 2748–2758.
92 Lis H & Sharon N (1993) Protein glycosylation. Structural and functional aspects. *Eur. J.*

Biochem. **218**, 1–27.
93 Varki A, Cummings R. D., Esko J. D., Freeze H. H., Stanley P., Bertozzi C. R., Hart G. W. EME (2009) *Essentials of Glycobiology*, 2nd ed. Cold Spring Harbor Laboratory Press.
94 Lodish H, Berk A, Zipursky SL, Matsudaira P, Baltimore D DJ (2000) *Molecular Cell Biology*, 4th ed. W. H Freeman.
95 Davies SJ & Crispin M (2011) *Functional and structural proteomics of glycoproteins* (R. J. Owens and J. E. Nettleship, eds.) Springer New York.
96 Steen P Van Den, Rudd PM, Dwek RA & Opdenakker G (1998) Concepts and Principles of O-Linked Glycosylation. **33**, 151–208.
97 Tanner W & Lehle L (1987) Protein glycosylation in yeast. *Biochim. Biophys. Acta* **906**, 81–99.
98 Lodish H, Berk A, Ziphursky L, Matsadaira P, Baltimore D & Darsell J (2000) *Molecular Cell Biology*, 4th ed. W.H. Freeman.
99 Invitrogen (2010) EasySelect Pichia Expression Kit, version G [Online]. .
100 Schulz BL (2012) *Beyond the Sequon: Sites of N-Glycosylation* (S. Petrescu, ed.).
101 Cregg JM, Vedvick TS & Raschke WC (1993) Recent advances in the expression of foreign genes in Pichia pastoris. *Biotechnology. (N. Y).* **11**, 905–910.
102 Mesters JR & Hilgenfeld R (2007) Protein Glycosylation, Sweet to Crystal Growth? *Cryst. Growth Des.* **7**, 2251–2253.
103 Vera-Estrella R, Barkla BJ, Bohnert HJ & Pantoja O (2004) Novel regulation of aquaporins during osmotic stress. *Plant Physiol.* **135**, 2318–2329.
104 Sato K, Kobayashi K, Aida S & Tamai S (2004) Bronchiolar expression of aquaporin-3 (AQP3) in rat lung and its dynamics in pulmonary oedema. *Pflugers Arch. Eur. J. Physiol.* **449**, 106–114.
105 Vinothkumar KR & Henderson R (2010) Structures of membrane proteins. *Q. Rev. Biophys.* **43**, 65–158.
106 http://www.rcsb.org/pdb/
107 Siller E, DeZwaan DC, Anderson JF, Freeman BC & Barral JM (2010) Slowing Bacterial Translation Speed Enhances Eukaryotic Protein Folding Efficiency. *J. Mol. Biol.* **396**, 1310–1318.
108 Gardner MJ, Hall N, Fung E, White O, Berriman M, Richard W, Carlton JM, Pain A, Nelson KE, Bowman S, Ian T, James K, Eisen J a, Rutherford K, Salzberg SL, Kyes S, Chan M, Nene V, Shallom SJ, Peterson J, Angiuoli S, Pertea M, Allen J, Haft D, Mather MW, Vaidya AB, Martin DM a & Alan H (2013) Europe PMC Funders Group Genome sequence of the human malaria parasite Plasmodium falciparum. **419**, 3–9.
109 Hedfalk K, Pettersson N, Öberg F, Hohmann S & Gordon E (2008) Production, characterization and crystallization of the Plasmodium falciparum aquaporin. *Protein Expr. Purif.* **59**, 69–78.
110 Cregg JM & Russell KA (1998) Transformation. *Methods Mol. Biol.* **103**, 27–39.
111 Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SGF, Thian FS, Kobilka TS, Choi H-J, Kuhn P, Weis WI, Kobilka BK & Stevens RC (2007) High-Resolution Crystal Structure of an Engineered Human 2-Adrenergic G Protein-Coupled Receptor. *Science (80-.).* **318**, 1258–1265.
112 Rosenbaum DM, Cherezov V, Hanson M a, Rasmussen SGF, Thian FS, Kobilka TS, Choi H, Yao X, Weis WI, Stevens RC & Kobilka BK (2007) into b 2 -Adrenergic Receptor Function. **318**, 1266–1273.
113 Newstead S, Kim H, von Heijne G, Iwata S & Drew D (2007) High-throughput fluorescent-based optimization of eukaryotic membrane protein overexpression and purification in Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 13936–13941.
114 Roosild TP, Greenwald J, Vega M, Castronovo S, Riek R & Choe S (2005) NMR structure of Mistic, a membrane-integrating protein for membrane protein expression. *Science* **307**, 1317–1321.
115 Kefala G, Kwiatkowski W, Esquivies L, Maslennikov I & Choe S (2007) Application of Mistic to improving the expression and membrane integration of histidine kinase receptors from Escherichia coli. *J. Struct. Funct. Genomics* **8**, 167–172.
116 Dvir H & Choe S (2009) Bacterial expression of a eukaryotic membrane protein in fusion to various Mistic orthologs. *Protein Expr. Purif.* **68**, 28–33.
117 Foster W, Helm A, Turnbull I, Gulati H, Yang B, Verkman AS & Skach WR (2000) Identification of sequence determinants that direct different intracellular folding pathways for aquaporin-

1 and aquaporin-4. *J. Biol. Chem.* **275**, 34157–34165.
118 Lu Y, Turnbull IR, Bragin a, Carveth K, Verkman a S & Skach WR (2000) Reorientation of aquaporin-1 topology during maturation in the endoplasmic reticulum. *Mol. Biol. Cell* **11**, 2973–2985.
119 Buck TM, Wagner J, Grund S & Skach WR (2007) A novel tripartite motif involved in aquaporin topogenesis, monomer folding and tetramerization. *Nat. Struct. Mol. Biol.* **14**, 762–769.
120 Berg JM, Tymoczko JL & L. S (2007) *Biochemistry*, 6th ed. W. H. Freeman and Company, New York, NY.
121 Wujek P, Kida E, Walus M, Wisniewski KE & Golabek A a. (2004) N-Glycosylation Is Crucial for Folding, Trafficking, and Stability of Human Tripeptidyl-peptidase I. *J. Biol. Chem.* **279**, 12827–12839.
122 Munro IC, Bernt WO, Borzelleca JF, Flamm G, Lynch BS, Kennepohl E, Bär E a. & Modderman J (1998) Erythritol: An interpretive summary of biochemical, metabolic, toxicological and clinical data. *Food Chem. Toxicol.* **36**, 1139–1174.
123 Prestoni GM, Jung JS, Guggino WB & Agre P (1994) Membrane topology of aquaporin CHIP: Analysis of functional epitope-scanning mutants by vectorial proteolysis. *J. Biol. Chem.* **269**, 1668–1673.
124 Törnroth-Horsefield S, Wang Y, Hedfalk K, Johanson U, Karlsson M, Tajkhorshid E, Neutze R & Kjellbom P (2006) Structural mechanism of plant aquaporin gating. *Nature* **439**, 688–694.
125 Janosi L & Ceccarelli M (2013) The Gating Mechanism of the Human Aquaporin 5 Revealed by Molecular Dynamics Simulations. *PLoS One* **8**, e59897.
126 Brown D, Katsura T, Kawashima M, Verkman a S & Sabolic I (1995) Cellular distribution of the aquaporins: a family of water channel proteins. *Histochem. Cell Biol.* **104**, 1–9.
127 Judith Radin M, Yu M-J, Stoedkilde L, Lance Miller R, Hoffert JD, Frokiaer J, Pisitkun T & Knepper MA (2012) Aquaporin-2 regulation in health and disease. *Vet. Clin. Pathol.* **41**, 455–470.
128 Fushimi K, Sasaki S & Marumo F (1997) Phosphorylation of serine 256 is required for cAMP-dependent regulatory exocytosis of the aquaporin-2 water channel. *J. Biol. Chem.* **272**, 14800–14804.
129 Fenton RA, Pedersen CN & Moeller HB (2013) New insights into regulated aquaporin-2 function. *Curr. Opin. Nephrol. Hypertens.* **22**, 551–558.
130 Verkman AS, Anderson MO & Papadopoulos MC (2014) Aquaporins: important but elusive drug targets. *Nat. Rev. Drug Discov.* **13**, 259–77.
131 Sidhaye V, Hoffert JD & King LS (2005) cAMP Has Distinct Acute and Chronic Effects on Aquaporin-5 in Lung Epithelial Cells. *J. Biol. Chem.* **280**, 3590–3596.
132 Hasegawa T, Azlina A, Javkhlan P, Yao C, Akamatsu T & Hosoi K (2011) Novel phosphorylation of aquaporin-5 at its threonine 259 through cAMP signaling in salivary gland cells. *Am. J. Physiol. Cell Physiol.* **301**, C667–C678.
133 Zhou B, Ann DK, Li X, Kim KJ, Lin H, Minoo P, Crandall ED & Borok Z (2007) Hypertonic induction of aquaporin-5: novel role of hypoxia-inducible factor-1alpha. *Am J Physiol Cell Physiol* **292**, C1280–90.
134 Hassan MI, Waheed a., Yadav S, Singh TP & Ahmad F (2009) Prolactin inducible protein in cancer, fertility and immunoregulation: Structure, function and its clinical implications. *Cell. Mol. Life Sci.* **66**, 447–459.
135 Naderi A & Meyer M (2012) Prolactin-induced protein mediates cell invasion and regulates integrin signaling in estrogen receptor-negative breast cancer. *Breast Cancer Res.* **14**, R111.
136 Crane JM, Van Hoek AN, Skach WR & Verkman AS (2008) Aquaporin-4 Dynamics in Orthogonal Arrays in Live Cells Visualized by Quantum Dot Single Particle Tracking. *Mol. Biol. Cell* **19**, 3369–3378.
137 Furman CS, Gorelick-Feldman D a, Davidson KG V, Yasumura T, Neely JD, Agre P & Rash JE (2003) Aquaporin-4 square array assembly: opposing actions of M1 and M23 isoforms. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 13609–14.
138 Gonen T, Sliz P, Kistler J, Cheng Y & Walz T (2004) Aquaporin-0 membrane junctions reveal the structure of a closed water pore. *Nature* **2002**, 193–197.
139 Harries W (2004) The channel architecture of aquaporin 0 at a 2.2-Å resolution. *PNAS* **101**, 14045–14050.
140 Gonen T, Cheng Y, Sliz P, Hiroaki Y, Fujiyoshi Y, Harrison SC & Walz T (2005) Lipid-protein

interactions in double-layered two-dimensional AQP0 crystals. *Nature* **438**, 633–638.
141 Palanivelu D V, Kozono DE, Engel A, Suda K, Lustig A, Agre P & Schirmer T (2006) Co-axial association of recombinant eye lens aquaporin-0 observed in loosely packed 3D crystals. *J. Mol. Biol.* **355**, 605–11.
142 Hite RK, Li Z & Walz T (2010) Principles of membrane protein interactions with annular lipids deduced from aquaporin-0 2D crystals. *EMBO J.* **29**, 1652–8.
143 Gonen T, Cheng Y, Sliz P, Hiroaki Y, Fujiyoshi Y, Harrison SC & Walz T (2005) Lipid-protein interactions in double-layered two-dimensional AQP0 crystals. *Nature* **438**, 633–8.
144 Marshall CB, Nishikawa T, Osawa M, Stathopulos PB & Ikura M (2015) Calmodulin and STIM proteins: Two major calcium sensors in the cytoplasm and endoplasmic reticulum. *Biochem. Biophys. Res. Commun.* **460**, 5–21.
145 Nelson MR & Chazin WJ (1998) Structures of EF-hand Ca2+-binding proteins: Diversity in the organization, packing and response to Ca2+ binding. *BioMetals* **11**, 297–318.
146 Johnson CN, Damo SM & Chazin WJ (2014) EF-Hand Calcium-Binding Proteins. *eLS*.
147 Girsch SJ & Peracchia C (1991) Calmodulin interacts with a C-terminaus peptide from the lens membrane protein MIP26. *Curr. Eye Res.* **10**, 839–849.
148 Reichow SL & Gonen T (2008) Noncanonical binding of calmodulin to aquaporin-0: implications for channel regulation. *Structure* **16**, 1389–98.
149 Hu S, Wang B, Qi Y & Lin H (2012) The Arg233Lys AQP0 mutation disturbs aquaporin0-calmodulin interaction causing polymorphic congenital cataract. *PLoS One* **7**, e37637.
150 Reichow SL, Clemens DM, Freites JA, Németh-Cahalan KL, Heyden M, Tobias DJ, Hall JE & Gonen T (2013) Allosteric mechanism of water-channel gating by Ca(2+)-calmodulin. *Nat. Struct. Mol. Biol.*
151 Reichow SL, Clemens DM, Freites JA, Németh-Cahalan KL, Heyden M, Tobias DJ, Hall JE & Gonen T (2013) Allosteric mechanism of water-channel gating by Ca2+-calmodulin. *Nat. Struct. Mol. Biol.* **20**, 1085–92.
152 Wu H, Chen L, Zhang X, Zhou Q, Li J-M, Berger S, Borok Z, Zhou B, Xiao Z, Yin H, Liu M, Wang Y, Jin J, Blackburn MR, Xia Y & Zhang W (2013) Aqp5 is a new transcriptional target of Dot1a and a regulator of Aqp2. *PLoS One* **8**, e53342.
153 Fetter K, Van Wilder V, Moshelion M & Chaumont F (2004) Interactions between plasma membrane aquaporins modulate their water channel activity. *Plant Cell* **16**, 215–228.
154 Zelazny E, Borst JW, Muylaert M, Batoko H, Hemminga M a & Chaumont F (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**, 12359–64.
155 Murozuka E, Hanisch S, Pomorski TG, Jahn TP & Schjoerring JK (2012) Bimolecular fluorescence complementation and interaction of various Arabidopsis major intrinsic proteins expressed in yeast. *Physiol. Plant.* **2**.
156 Nguyen HB, Hung LW, Yeates TO, Terwilliger TC & Waldo GS (2013) Split green fluorescent protein as a modular binding partner for protein crystallization. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69**, 2513–2523.
157 Kerppola TK (2008) Bimolecular fluorescence complementation (BiFC) analysis as a probe of interactions in living cells. *Annu. Rev. Biophys.*, 465–487.
158 Feilmeier BJ, Iseminger G, Schroeder D, Webber H & Phillips GJ (2000) Green fluorescent protein functions as a reporter for protein localization in Escherichia coli. *J. Bacteriol.* **182**, 4068–4076.
159 Toddo S, Söderström B, Palombo I, Von Heijne G, Nørholm MHH & Daley DO (2012) Application of split-green fluorescent protein for topology mapping membrane proteins in Escherichia coli. *Protein Sci.* **21**, 1571–1576.
160 Ashikawa Y, Ihara M, Matsuura N, Fukunaga Y, Kusakabe Y & Yamashita A (2011) GFP-based evaluation system of recombinant expression through the secretory pathway in insect cells and its application to the extracellular domains of class C GPCRs. *Protein Sci.* **20**, 1720–34.
161 Cheng Y (2015) Single-Particle Cryo-EM at Crystallographic Resolution. *Cell* **161**, 450–457.
162 Bai X, McMullan G & Scheres SH. (2015) How cryo-EM is revolutionizing structural biology. *Trends Biochem. Sci.* **40**, 49–57.
163 Cheng Y, Grigorieff N, Penczek PA & Walz T (2015) A Primer to Single-Particle Cryo-Electron Microscopy. *Cell* **161**, 438–449.
164 Bartesaghi A, Merk A, Banerjee S, Matthies D, Wu X, Milne JLS & Subramaniam S (2015) 2.2 A

resolution cryo-EM structure of -galactosidase in complex with a cell-permeant inhibitor. *Science (80-.).* **348**, 1147–1151.
165 Bai X, Yan C, Yang G, Lu P, Ma D, Sun L, Zhou R, Scheres SHW & Shi Y (2015) An atomic structure of human γ-secretase. *Nature* **525**, 212–217.
166 Saibil HR (2000) Conformational changes studied by cryo-electron microscopy. *Nat. Struct. Biol.* **7**, 711–714.