Structural Studies of Human Aquaporin-1 in Polymer Nanodiscs and an Investigation into a Conserved Hydrogen-bond Network Crucial for Stability

By

Philip Drewniak

A Thesis

presented to

The University of Guelph

In partial fulfillment of the requirements

for the degree of

Master of Science

in

Biophysics

Guelph, Ontario, Canada

ABSTRACT

STRUCTURAL STUDIES OF HUMAN AQUAPORIN-1 IN POLYMER NANODISCS, AND AN INVESTIGATION INTO A CONSERVED HYDROGEN-BOND NETWORK CRUCIAL FOR STABILITY

Philip Drewniak
University of Guelph, 2023

Advisor:
Leonid Brown

The human Aquaporin-1 (AQP1) membrane protein is integral towards many cellular processes involved with water permeability. This thesis explores two separate projects towards the further structural characterization of AQP1. Firstly, the novel reconstitution of AQP1 into SMALPs (styrene-maleic acid lipid nanoparticles) was developed and further assayed for stability and amenability towards solution nuclear magnetic resonance (NMR) experiments. A considerable amount of the most mobile amino acids in AQP1 were observed and resolved using this membrane mimetic system. Secondly, this thesis reports an additional focus into a highly conserved hydrogen-bond network existing within AQP1. This network was explored using site-directed mutagenesis and Fourier-transform infrared (FTIR) spectroscopy coupled with hydrogen-deuterium exchange upon increasing temperatures. The mutants studied affected the stability of AQP1 to varying degrees, revealing the importance of these residues towards the overall stability of the native protein fold.

ACKNOWLEDGEMENTS

I would like to express my sincerest gratitude first and foremost for my advisor, Dr. Leonid Brown, for the tremendous amount of guidance, patience, kindness, and enthusiasm towards the work in this thesis and in my constant development as both a student and an aspiring scientist.

I would also like to extend my appreciation to my committee members; Dr. Vladimir Ladizhansky and Dr. John Dawson, as well as my examination committee and chair members Dr. Rui Huang and Dr. John Dutcher. A special thanks to Dr. John Dutcher as well for being so kind in lending us use of his ATR-FTIR spectrometer, which was crucial towards some of the results in this work. In addition, I would also like to thank Dr. Elyse Roach and Dr. Sameer Al-Abdul-Wahid from the Advanced Analysis Centre here at the University, both for their expertise and in their teaching and mentorship on the instruments used.

Much gratitude goes to my fellow lab members; Xiao, Raoul, Marie, Maryam, Suelen, Azam, Sajjad and Mahdi for all their help both inside and outside the lab. I am also grateful for all the staff members here at the Department of Physics and the Biophysics Interdepartmental group; Ciaran Henry, Reggi Vallillee, Rachel Baker, Cynthia Cheeseman, Janice Ilic, and Dr. Hermann Eberl for all their help over the course of these past 2 years.

Last but certainly not least a special thanks to my friends and family for offering their constant support and much-needed distractions. I am especially grateful for my partner, Liv, for providing the laughs and love needed to get through the writing process the past four months. Finally, and most importantly, I thank Luna and Orion for their attentiveness, care, and curiosity towards my work; that last part being especially pertinent given they are cats, but no cats were harmed in the making of this thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS, ABBREVIATIONS OR NOMENCLATURE

ATR-FTIR Attenuated-total-reflectance FTIR spectroscopy

ar/R Aromatic/Arginine filter

AQP Aquaporin

AQP1 Aquaporin-1

AQGP Aquaglyceroporin

AQPZ *E. coli* aquaporin

AQPY Yeast aquaporin

CAII Carbonic anhydrase II

CD Circular-dichroism spectroscopy

CMC Critical micellar concentration

cryo-EM Cryogenic electron microscopy

CSP Chemical shift perturbation

DDM n-Dodecyl-β-D-maltopyranoside

DLS Dynamic light scattering

DMPC 1,2-dimyristoyl-sn-glycero-phosphocholine

DNA Deoxyribonucleic acid

DTGS Deuterium triglycine sulfate

DSC Differential scanning calorimetry

EDTA Ethylenediaminetetraacetic acid

E. coli *Escherichia coli*

FTIR Fourier-transform infrared spectroscopy

FRET Forster energy resonance transfer

G3P Glycerol-3-phosphate

GLPF Glycerol uptake facilitator protein

GPCR G-protein coupled receptor

GSK3 Glycogen synthase kinase 3

hAQP Human Aquaporin

H-bond Hydrogen-bond

H/DX Hydrogen/deuterium exchange

HSQC Heteronuclear single-quantum coherence spectroscopy

IR Infrared spectroscopy

IRE Internal reflection element

IPTG Isopropyl β-D-1-thiogalactopyranoside

kDa Kilodaltons

MCT Mercury-cadmium-telluride

MD Molecular dynamics simulations

MP Membrane protein

mRNA messenger RNA

MSP Membrane scaffold protein

MW Molecular weight

NADH Nicotinamide adenine dinucleotide

NMR Nuclear magnetic resonance spectroscopy

NPA Asn-Pro-Ala motif

OG **n-**Octyl-β-D-glucopyranoside

OmpA Outer membrane protein A

P. pastoris *Pichia pastoris*

PC Phosphatidylcholine

PDB ID Protein data bank identification

pK_a Acid-dissociation constant

PKA Protein Kinase A

PMSF Phenylmethanesulfonyl fluoride

POPE palmitoyloleoylphosphatidy-ethanolamine

POPC 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine

ppm parts-per-million

PS Phosphatidylserine

RF Radiofrequency

RNA Ribonucleic acid

rpm Revolutions per minute

SDS-PAGE Sodium dodecyl sulfate Polyacrylamide gel electrophoresis

SMA Styrene-maleic acid

SMALPs Styrene maleic acid lipid nanoparticles

ssNMR solid-state NMR

TEM Transmission electron microscopy

TEV Tobacco-mosaic virus

TROSY Transverse relaxation optimized spectroscopy

UV-VIS Ultraviolet-Visible spectroscopy

WT Wild-type

YNB Yeast nitrogen base

YPD Yeast peptone dextrose

YPDS Yeast peptone dextrose sorbitol

Chapter 1

An introduction to the Aquaporins and membrane mimetics

Chapter 1: Overview

In this chapter, the current literature regarding membrane proteins will be discussed, as well as some examples of medically significant proteins, in which then the focus will shift towards the topic of human aquaporins, and a comprehensive introduction into this family of proteins will be presented. Structurally significant conserved motifs and residues will be discussed, and the focus will then narrow towards the topic of human aquaporin-1, its significant structural features and its context within human physiology and disease. Finally, the current state of the use of membrane mimetics towards studying membrane proteins in their native environments will be explored.

1.1 Membrane Proteins

Cells are the singular unit of life, and they are enclosed and compartmentalized entirely via biological membranes [1]. Biological membranes form a necessary barrier between the intracellular cytoplasm and the outside environment, protecting the cell and establishing certain gradients, which power processes such as aerobic respiration and cellular locomotion [2,3]. These membranes are composed, apart from proteins, primarily of phospholipids, sphingolipids, and sterols, each with their own structural diversity [4]. The overall lipid composition of a membrane affects its shape and rigidity, processes that have proven essential towards certain cellular pathways [5,6]. The membrane itself is a bilayer composed of these phospholipids, with the hydrophilic head groups facing out towards the aqueous environment and the hydrophobic acyl chains buried within the bilayer [7].

Within this membrane bilayer there are proteins that function as pumps, channels, transporters, and receptors [8,9]. Proteins of this type are divided into two classes – integral and peripheral membrane proteins; while the former are insoluble proteins which are difficult to separate from the membrane without the use of detergents, the latter class are soluble and more easily extracted from the membrane environment [10–12]. Peripheral proteins usually contain just a few anchoring residues that interact with the membrane and bind very weakly[13]. Within the integral membrane proteins, there are two further sub-categories: transmembrane and integral monotopic proteins [14]. Integral monotopic proteins are a unique instance where the protein only associates with one face of the membrane, most commonly through an amphipathic α-helix that lies parallel to the membrane [15,16]. These proteins are purported to catalyze reactions that involve both hydrophobic and hydrophilic substrates, and so the enzyme lies at the interface between both membrane and soluble environments [17].

Transmembrane proteins span the entire length of the membrane and consist largely of ion transporters, channels, and receptors [18–21]. These integral proteins require the use of mild to harsh detergents to separate them from the membrane for further downstream purification and biophysical analyses. Transmembrane integral proteins have typically either an α-helical or beta-barrel structure, however most of these proteins contain α-helical transmembrane domains [22]. A significant example of these α-helical proteins are the family of rhodopsins – a class of photoreceptor proteins that are light-responsive and trigger a variety of cellular signalling [23,24]. These rhodopsins are found throughout all kingdoms of life, present in eukaryotes, prokaryotes, and archaea [25]. While bacteria and archaea use rhodopsins for energy transduction and signalling, animals use a special class of rhodopsins that belong to the much larger superfamily of G-protein coupled receptors (GPCRs) [26]. The GPCRs are a large focus of

the pharmaceutical drug industry, as the over 900 member protein family is involved in many physiological processes and disease states, thus highlighting the need for research into rhodopsins [27].

Another α-helical transmembrane integral protein of considerable basic science and pharmaceutical interest, is human aquaporin-1, which will be the major focus of this thesis.

1.2 The Aquaporins

Aquaporins (AQP) are a family of water-transporting channel proteins that are also found ubiquitously throughout several kingdoms of life [28–30]. AQPs have evolved mainly to transport water, however some members of the family have also been found to transport some small solutes and even gasses. Through the main function of water transport, AQPs are thus also directly involved in cellular proliferation, migration, adhesion, and angiogenesis[31–33]. Of particular note as well, is the emerging evidence towards certain AQPs involvement in some integral cancer processes, highlighting the need for further work on these proteins as they can be valuable future drug targets [34–37]. Since the discovery of the first aquaporin isolated from human red blood cells in 1993, then-named CHIP-28 and later renamed to AQP1, many more aquaporins have been discovered [38,39]. From microbes to plants and mammals, aquaporins are abundant throughout cellular life and are indeed an old family of proteins. Plant aquaporins are incredibly complex, where a single plant species could contain more than 120 isoforms of AQP [40,41]. As well, similar to some of the mammalian counterparts, some of these plant AQPs' water transport is regulated through gating mechanisms. In *Arabidopsis thaliana,* the closure of AQPs has been reported to be through the protonation of a highly conserved histidine residue in loop D [42,43]. While the overall family of AQPs and related glycerol-facilitator like protein

(GLPF) is large, the focus of this chapter will be more constrained to the mammalian subfamily of AQPs [44].

There are thirteen mammalian aquaporins, designated AQP0 to AQP12, and they are found throughout all human tissues, with some aquaporins expressed in certain cell types while others, like AQP1, more ubiquitous [45]. Among these AQP's further sub-classifications exist based on whether their pores strictly transport water or if they facilitate transport of other small solutes such as glycerol or urea – these aquaporins being aquaglyceroporins [**Table 1-1**][46,47]. AQP0-2, 4-6 and 8 are considered classical or orthodox aquaporins, while AQP3, 7, 9 and 10 are aquaglyceroporins [48]. AQP11-12 functions are largely unknown, with AQP11 and 12 sharing low overall sequence homology in the main water specificity region of AQPs, resulting in their classification as unorthodox aquaporins, or superaquaporins [30].

Among the classical aquaporins, AQP0 is the major lens aquaporin of the eye, and it contains several strongly conserved residues among the water transporting aquaporins, but missing in the aquaglyceroporins [49,50]. Several studies have indicated that AQP0 enhances membrane water permeability under mildly acidic conditions, proposing that AQP0 is gated via pH, thus providing evidence that some mammalian AQPs can also be gated by a change in environment [51,52]. AQP1 is found in many kinds of cell types and will be given a full descriptive treatment in later sections.

AQP2 is a vasopressin-regulated water channel, primarily found in the kidneys in the renal collecting duct, and has been the subject of many studies regarding how AQP2 is trafficked to the plasma membrane, as this dysregulation of trafficking is the cause of a few diseases impacting water homeostasis [53–55]. Several mutations in AQP2 itself are known to cause both dominant and recessive forms of nephrogenic diabetes insipidus, which is characterized by an

inability to concentrate urine in response to vasopressin regulation [56,57]. Phosphorylation of the several serine residues, located on the long and flexible C-terminal tail of AQP2, is both necessary and sufficient for AQP2 localization to the apical plasma membrane, thus highlighting how regulation through trafficking is also an important feature of AQPs [58]. Many mammalian AQPs also possess a longer and flexible C-terminal tail, which could imply the existence of similar phosphorylation or other post-translational modification mechanisms needed for the translocation of mammalian AQPs in general.

AQP4 is the primary water channel in the brain; it is abundantly expressed and involved in several processes that are critical towards cerebral water balance [59]. AQP4 is found highly expressed in human brain tumours [60]. mRNA splicing of the gene encoding AQP4 results in two isoforms of differing molecular weights (MW), M1 and M23 [61]. These isoforms have implications towards cell to cell adhesion as the M23 isoform especially forms orthogonal crystalline arrays, where the interactions between tetramers are mediated by helical contacts involving the short 3_{10} helix in loop C of AQP4 [62]. In these double-layered crystals, each tetramer interacts with four other tetramers, resulting in the lattices slightly shifted from each other. These crystalline arrays that AQP4 tetramers form in the end feet of glial cells, may suggest that AQP4 has a role in weak cell-cell adhesions which are found in the junctions between glial cells [62]. The size and structure of these orthogonal arrays are determined by the ratio of the M1 and M23 isoforms present in the crystals, with a higher M1/M23 ratio disrupting the formation of these arrays [63]. Thus, AQP4's involvement in the transport of water across the blood-brain barrier, its role in astrocyte migration, and neural signal transduction suggests AQP4 is physiologically important and diverse in its functions, and is a valid target for drug development for many brain ailments [63].

AQP5 is another orthodox water channel that only transports water, and is increasingly found to be involved and highly expressed in different tumours and cancerous cells [64]. AQP5 is also widely expressed throughout various human cells [65]. Interestingly, although the core structure is very similar to that of the other classical aquaporins, the crystal structure was solved with a phosphatidylserine (PS) lipid occluding the central gas-conducting pore of the tetramer [66]. This lipid would seem to inhibit the gas transporting function of AQP5. AQP5 is also regulated via phosphorylation of a serine residue, and molecular dynamics simulations suggested that it can also be gated by another histidine residue located on the cytoplasmic side of the channel [67].

AQP6 is a classical water channel, one with no atomic structure available, however it is unique in that it is also an anion-selective channel with a specificity for nitrate ions [68]. AQP6, when compared to other mammalian AQPs, differs in one amino acid at a position near the water pore, where an asparagine resides instead of a glycine – this one amino acid change is entirely responsible for its anion channel conductance [69,70]. A single point-mutation of Asn to a Gly removes the anion channel conductance of AQP6 and relegates it back to being strictly permeable to water. AQP6's water permeation is also known to be activated by mercury ions, confirmed through functional assays and MD simulations [71,72]. The Hg^{2+} ion binds to Cys190, resulting in a broken H-bond between Arg196 and His181, thus creating an open pore conformation near the aromatic/Arginine (ar/R) region [73]. This open conformation increases water permeation tenfold compared to the base rate.

AQP8 is a channel that is permeable to both water and ammonia, whereby its ammonia permeability exceeds that of water by over twofold [74,75]. AQP8 has also been found to localize within the inner mitochondrial membrane, and hypothesized to have involvement in

mitochondrial ammonium detoxification through the process of ureagenesis [75]. Functional assays have determined that AQP8 is a classical water transporting aquaporin, however sequence alignments with other mammalian AQPs suggest that AQP8 is closer in identity to the unorthodox AQPs, AQP11 and 12 [76]. Evidence suggests that AQP8 is also gated via persulfidation of a cysteine and the latter is modulated through cell stress [77]. However no further structural or biochemical mutagenesis information exists for this protein, and the exact physiological role for the ammonium transport is to be fully deciphered.

AQPs 3, 7, 9 and 10 are the mammalian aquaglyceroporins. These AQGPs are known to transport glycerol in addition to water. The most well-known aquaglyceroporin is the *E. coli* glycerol-conducting channel (GlpF). The structure of GLPF is well-resolved at 2.2 Å, and it provides structural clues as to how this class of AQP also transports glycerol [78]. For starters, the pore diameter is larger than the classical AQP's at 3.5 Å wide, and the glycerol molecules diffuse through the pore single file, with the hydroxyl groups forming hydrogen-bonds with exposed carbonyl oxygen atoms and hydrogen bond donors from residues lining the pore [79–81]. Glycerol thus is aligned with its alkyl chain within a two-sided hydrophobic segment in the pore, and hydrogen-bonds to several amino acids near the water pore [79]. Broadly, these features constitute a channel that transports glycerol in addition to water, and there are suggestions towards what physiological relevance they have for human metabolism.

AQP 3, 7, 9, 10 have all been found to be expressed in adipose tissue, as the glycerol transport function may have implications towards energy metabolism, specifically in lipolysis and lipogenesis [82–84]. Glycerol itself is a precursor for the cellular synthesis towards triacyl glycerides and phospholipids. As well, an increase in intracellular glycerol concentration leads to the activation of glycerol kinase, which catalyzes the synthesis of glycerol-3-phosphate (G3P),

an important molecule in the regeneration of intracellular NAD^+ from NADH [85]. G3P is also vital in the glycolysis and gluconeogenesis pathways. Regulation of intracellular glycerol is dominated by the aquaglyceroporins, and so they must play a large role in controlling energy homeostasis [86].

Interestingly, several of these mammalian AQGPs are involved in the transport of arsenic through cells. Transport studies in *Xenopus laevis* oocytes and *Saccharomyces cerevisiae* cells showed both AQP7 and AQP9 involved in the increased transport of arsenite (^{73}Ar[III]) [87,88]. These results may hint at a possible mode of action of arsenic that may require the participation of AQP7 and/or AQP9 in providing a method of transport within the cells, given that arsenite is both a known human toxin and carcinogen.

Table 1.1: The mammalian aquaporins sorted by family and their respective solute transport specificity. Information compiled from Kitchen et al., 2015 [89], and Verkman, 2008 [46].

Family	***Aquaporin***	***Transport***
Orthodox or classical	AQP0	Water
	AQP1	Water, CO_2, NO
	AQP2	Water
	AQP4	Water
	AQP5	Water
	AQP6	Nitrate, water, urea
	AQP8	Ammonia, water
Aquaglyceroporins	AQP3	Water, glycerol, ammonia
	AQP7	Water, glycerol, urea, arsenic
	AQP9	Water, glycerol, urea, arsenic
	AQP10	Water, glycerol, urea
Superaquaporins	AQP11	Water, glycerol
	AQP12	Water

1.3 Aquaporin structure

X-ray crystallography and electron microscopy studies have revealed the three-dimensional structure of several aquaporins, and this overall structure is quite uniform among the different AQPs. AQPs exist as 24-30 kDa monomers that tetramerize inside the internal membranes of cells, before being trafficking to the plasma membrane where water transport occurs through a pore in each of these monomers [47,90]. Aquaporins have six tilted transmembrane helices resulting in an hourglass shape. The structure consists of helices 1-6 (H1-H6), and five loop regions (LA-LE), with both N- and C-termini facing the cytoplasm [**Figure 1-1**]. The main water pore residues reside in the two half-helices HB and HE, which host the functionally important and incredibly conserved NPA (Asn-Pro-Ala) motif, in both re-entrant loops LB and LE [90]. The constrictions of the pore are formed by the NPA motif and an aromatic/arginine (ar/R) selectivity filter. The ar/R region is comprised of two residues from H2 and H6, and two residues from LE. In human aquaporin 1, Phe56, His180, and Arg195 are the major residues involved in this ar/R filter, with the backbone carbonyls of Gly188 and Cys189 contributing as well [**Figure 1-2**] [90]. The two NPA motifs consist of the residues Asn76, Pro77, Ala78, and Asn 192, Pro193, and Ala194 [**Figure 1-2**]. Through these residues, water molecules are transported through the pore single file, with the asparagine from the dual NPA motif on either side of the pore acting as highly oriented hydrogen bond donors from the side chain NH_2 groups [91]. The arginine from the ar/R filter is also well-placed at the outset of the pore to exclude larger positively charged ions. The minimal diameter of the pore is also only 2.8 Å, which is approximately the mean Van der Waals radius of a single water molecule [92]. AQP's structural conformation allows the exclusion of protons within the channel while only

allowing water molecules through. This extreme substrate selectivity is possible through the orientation of the dipoles on the two short pore helices, HB and HE [93]. These dipoles create a local electrostatic field, which forces the water molecule's oxygen atom to orient itself towards the two amide groups on the Asn from the NPA filter. This creates a transient hydrogen bond, one which orients the water molecule in perpendicular fashion to the water channel, excluding it from hydrogen-bonding with any other water molecules in the channel, thus preventing any proton transfer [91,93]. Through these constriction motifs and the size exclusion diameter, AQPs fulfill the function of water transport by significantly reducing the activation energy for water permeation through the membrane. This allows for such a large influx of water that the rate of water inflow is almost that of bulk water diffusion, with $3.9x10^9$ water molecules per second flowing through each monomer subunit [38].

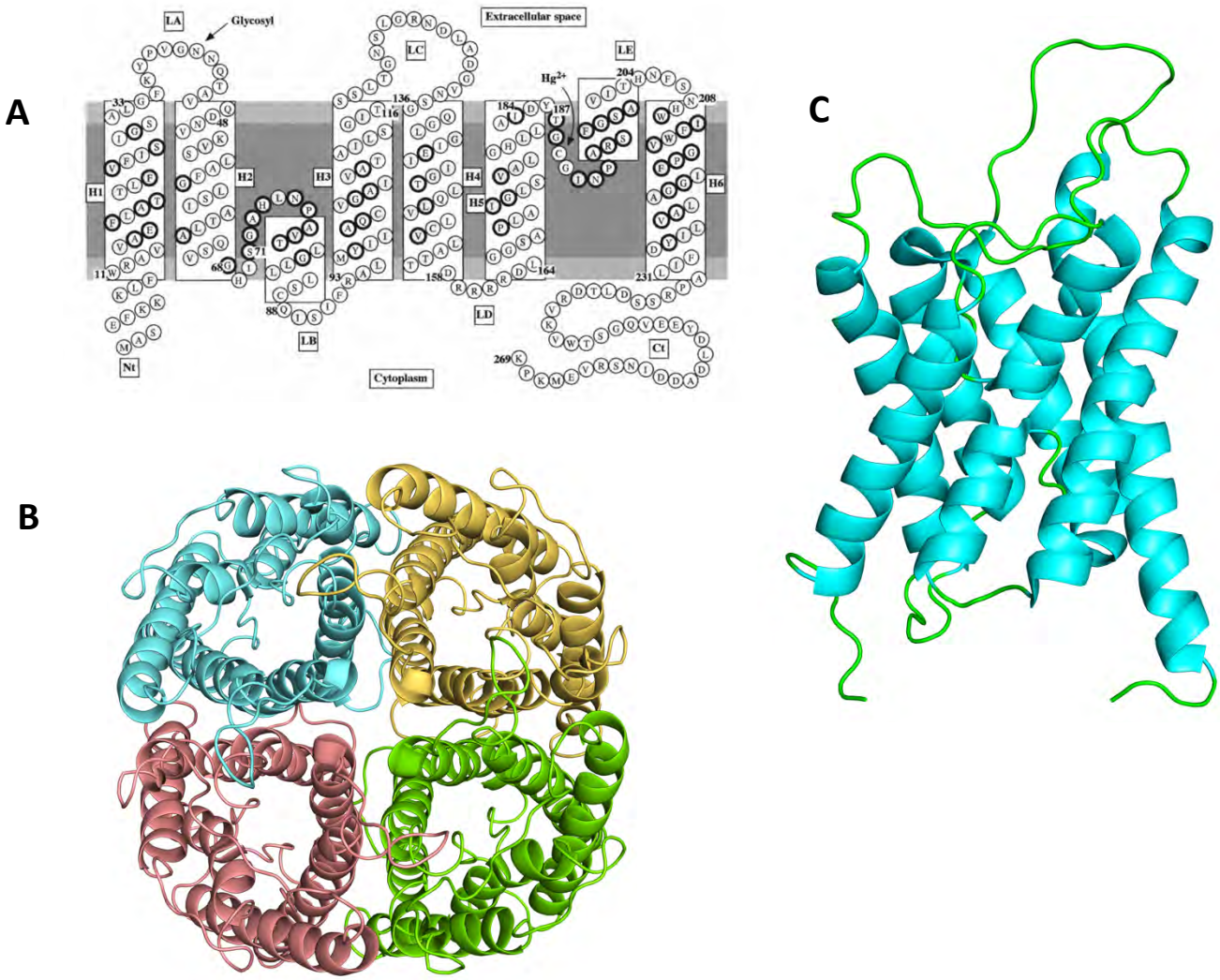

Figure 1-1: *(**A**) Topological model of human AQP1 showing positions of transmembrane helices, external connecting loops, and important residues/motifs. Residues circled in bold represent highly conserved positions. The direction of the N- and C-termini are also shown facing towards the cytoplasmic side of the cell. Helices numbered H1-H6 and the loops numbered LA-LE, with the dual NPA motifs located on LB and LE. Figure modified from Heymann and Engel, 2000, with permission from the Journal of Molecular Biology. (**B**) Tetrameric assembly of AQP1, with the water transporting pore located in the middle of each monomer. The middle cavity of the tetramer is believed to transport smaller ions like* CO_2 NH_3,

*NO and O_2. Bovine AQP1 (PDB ID: 1J4N). (**C**) Monomeric unit of AQP1, highlighting secondary structure features (PDB ID: 1J4N). Figures generated in PyMOL.*

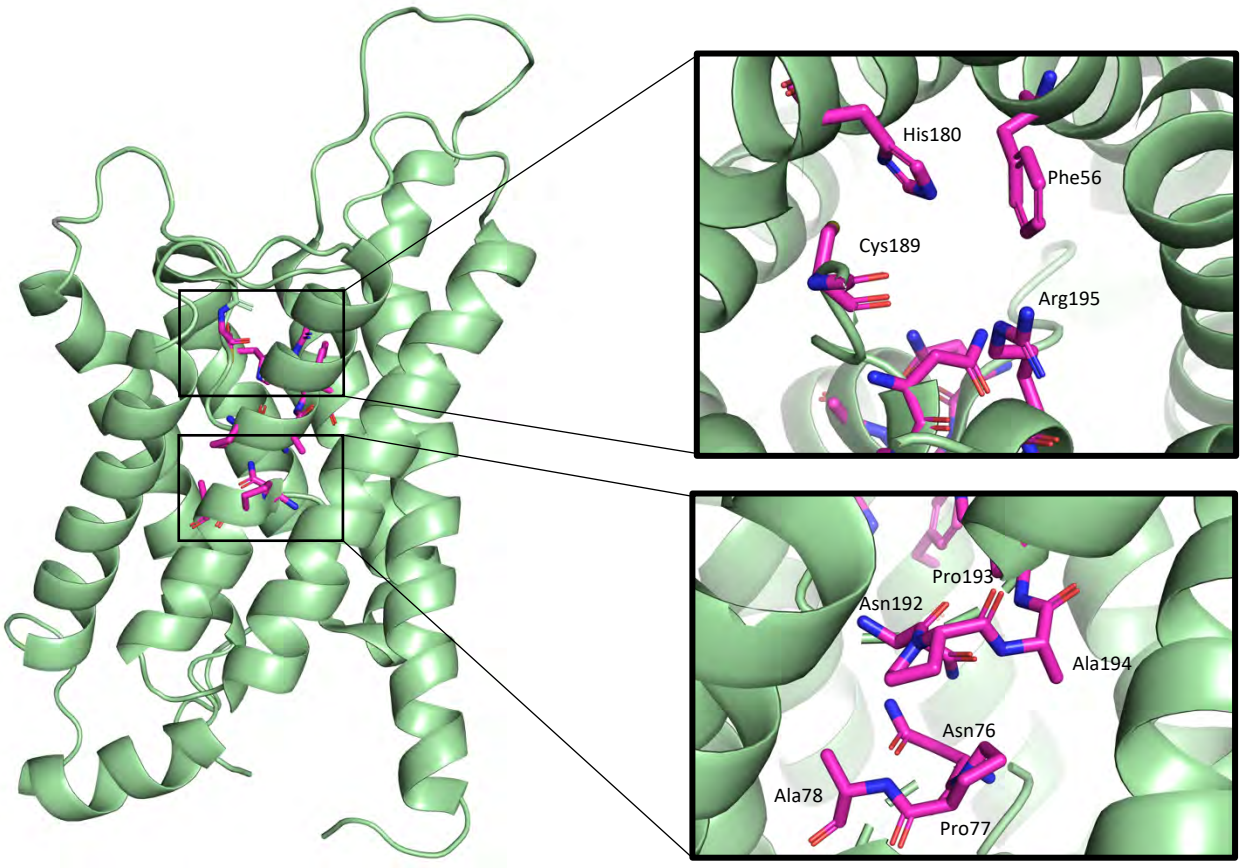

Figure 1-2: *PDB: 1J4N. High resolution X-ray structure of bovine aquaporin-1 with human AQP1 sequence numbering. The ar/R region is shown from the top down, and the NPA filters shown from the side. Residues involved in ar/R filter shown in top window, with residues from NPA motifs shown in bottom window. Figure generated in PyMOL.*

Aside from the main water channel, aquaporins have a few other important structural features that have implications towards human physiology and certain disease states. The middle

pore of the tetramer of AQP1 is believed to be able to transport small gases, such as CO_2, NH_3, NO, and O_2 [94]. CO_2 transport through the middle tetrameric pore is especially physiologically relevant, and it has been found experimentally that AQP1 reconstituted into proteoliposomes is permeable to CO_2 [95]. Whether this CO_2 permeability occurs through the middle tetrameric cavity or through the monomeric water channel is uncertain. MD simulations concerning AQP1 CO_2 permeability suggest that although CO_2 flux can occur through the water channel, it is unlikely as the free energy barrier for this movement is ~23 kJ/mol in the ar/R region, whereas a barrier of just 4 kJ/mol was observed for CO_2 flux in pure palmitoyloleoylphosphatidyethanolamine (POPE) lipid bilayer membrane [96]. When observing the free energy barrier for the central cavity, it was significantly smaller than the barrier for the monomeric channel, however this was done under the assumption that a small ion or organic molecule was not obstructing the pore. These results would suggest that for CO_2 flux in AQP1 to be physiologically significant, it would have to occur in cell membranes with a lower intrinsic CO_2 permeability [96].

Apart from the middle cavity and the central water channel, several AQPs have a long C-terminal tail that is crucial in both subcellular membrane trafficking and regulation [**Figure 1-3]** [97,98]. Phosphorylation at Ser235 of AQP0, located on its C-terminus, has been reported to be involved in the regulation of its water permeability and trafficking [99]. Phosphorylation of this C-terminal tail has also shown that it lowers the binding affinity of calmodulin to AQP0 C-terminal peptides, thus suggesting that phosphorylation has some role in regulating the AQP0-calmodulin binding interaction [100]. This phosphorylation regulating either a protein-protein interaction or membrane trafficking has also been reported for plant aquaporins such as PIP2, and for mammalian aquaporins AQP2 and AQP4 [101–103].

The C-terminal tail of AQP1 also contains putative binding sites for protein-protein interaction events. Of these events, it is proposed that carbonic anhydrase II (CAII), an enzyme which catalyzes the reaction between CO_2 and H_2O, physically interacts with AQP1 [104]. Through immunoblotting and cellular assays, it has been found that CAII interacts with the C-terminus of AQP1, and in so doing increases the water flux across red blood cell membranes [105]. AQP1's tail contains two acidic and negatively charged motifs that are conserved with known carbonic anhydrase I and II binding sequences [105,106]. A CA binding motif is distinguished by a hydrophobic amino acid followed by four amino acids, of which at least two must be acidic. In AQP1's C-terminal tail, two motifs follow this rule: V^{250}EEYD and L^{255}DADD. These regions are the CAI and CAII putative binding regions, and they are also strongly conserved among orthologs of AQP1 from several different species [105]. Sequence alignment of other AQP homologs did not show these binding motifs. A study involving the well-known interaction between CAII and anion-exchanger-1 (AE1) showed that the binding involved the N-terminal region of CAII, which contains a cluster of basic histidines and positively charged lysines [107]. AE1 also contains the CAII acidic binding motif of LDADD, which would further suggest that the interaction event between CAII and AQP1 is dominated by electrostatic interactions between the positively and negatively charged residues [108].

Although the C-terminal tail is implicated in a few cellular processes governing AQPs, it has not yet been structurally resolved. The tail is a very long, flexible, and disordered region of the protein, and so conventional structural approaches like X-ray crystallography have not been able to achieve characterization of the entire tail. Flexible regions, while disordered *in vitro*, can become ordered under crystallization conditions, resulting in inaccurate structural conformation

[109]. Methods such as solution NMR can detect and help characterize these regions more accurately.

Regarding the flexible regions of AQPs, the structure also contains several loop regions connecting the helical bundle. Interestingly, some of these loop regions are not as flexible as previously thought, with discrepancies existing between NMR spectral and crystallographic data for these regions in AQP1 [110,111]. NMR data of AQP1 would suggest that Loops A and C are partially structured, with loop C especially being largely immobile. Loop C is a ~20 residue long extracellular loop, and the conformations of this loop vary greatly between crystal structures, due to this disordered region likely adapting different conformations in a crystalline state [**Figure 1-4].** Hydrogen/deuterium (H/D) exchange experiments were employed to investigate which areas in AQP1 were more exposed to solvent than others, based on the ability of the backbone amide protons to exchange [110]. These H/D exchange experiments highlighted several residues and regions that were unexplored before, and may be crucial towards the overall stability and fold in AQP1. Indirectly, these residues may impact the function of water transport as well, not only just in AQP1 but in other AQPs that share these extremely conserved structural features [110].

In loop C, there exist several possibly important structural features which may serve to stabilize the ar/R constriction. Dihedral backbone angle predictions from NMR analysis and MD simulations indicate the possibility of a type II β-turn between residues Ala130 and Val133, which is further supported by the low exchange of the backbone amide of V133 [111]. This β-turn is stabilized by a few local potential hydrogen bonds and may help to anchor loop C along the top of the structure [**Figure 1-4**]. Of significant note is the lack of exchangeability of N127's sidechain amide protons, which has been confirmed in two separate NMR H/DX experiments and would suggest that N127 is quite buried and involved in hydrogen bonding. Many of the

lowest energy computed structures place N127's sidechain in a downward geometry, less than 4.0 Å away from potential H-bond partners [111].

Another significant defining feature of some aquaporins is that their water transport is inhibited by mercurial compounds, apart from a few exceptions. The precise mechanism of mercurial inhibition is still to be determined, but several structural and molecular dynamics studies have identified the exact binding site of Hg^{2+} and potentially a few residues involved in the water transport inhibition mechanism [112,113]. Mercury binds to cysteine residues on AQPs, specifically to Cys189 in human AQP1 and Cys191 in bovine AQP1 [114]. While AQP1 contains a few cysteines in its structure, it was found that only Cys189 contributed to the complete water transport inhibition of the protein [115]. Through molecular dynamics simulations (MD), it was proposed that mercury binding to Cys189 imparted changes in the local conformations of residues near the pore at the ar/R region, which induced a constriction of the space between Arg195 and His180. These changes led to a pore that was impermeable to water due to size alone [113]. Similarly, using solid-state nuclear magnetic resonance (ssNMR), the bacterial aquaporin homolog AQPZ was studied upon addition of Hg^{2+} ions [73]. It was determined that when the Hg^{2+} ion was bound to Cys190 (Cys189 in human AQP1), this forced the nearby Arg189 residue to be repelled electrostatically, thus pushing the side chain of arginine to point downward. and creating a possible new hydrogen-bond with His174 [73]. This conformational change in Arg189 led to a full constriction of the pore [73]. The NMR spectroscopy was able to discern that while the overall structure of AQPZ was unperturbed, mercury inhibition was guided by residue-specific conformational changes in the pore.

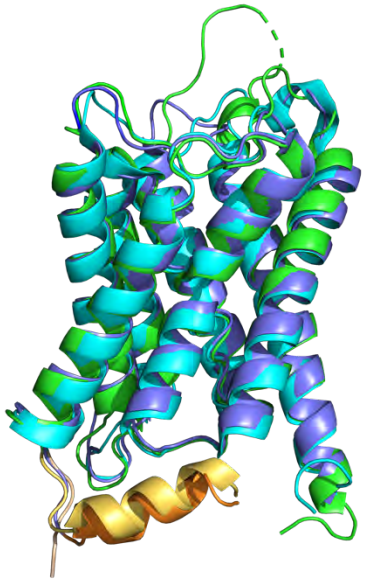

Figure 1-3: *Structural alignment of bovine AQP0 (purple), bovine AQP1 (green), and human AQP2 (cyan), with the C-terminal tails coloured in wheat, yellow, and orange, respectively. The ~40 residue long tail in each homolog is either cleaved or not fully resolved. The adoption of partial helical conformation can be observed in the X-ray structure of AQP2 and the EM structure of AQP0. The C-terminal tail of AQP1 remains unstructured.*

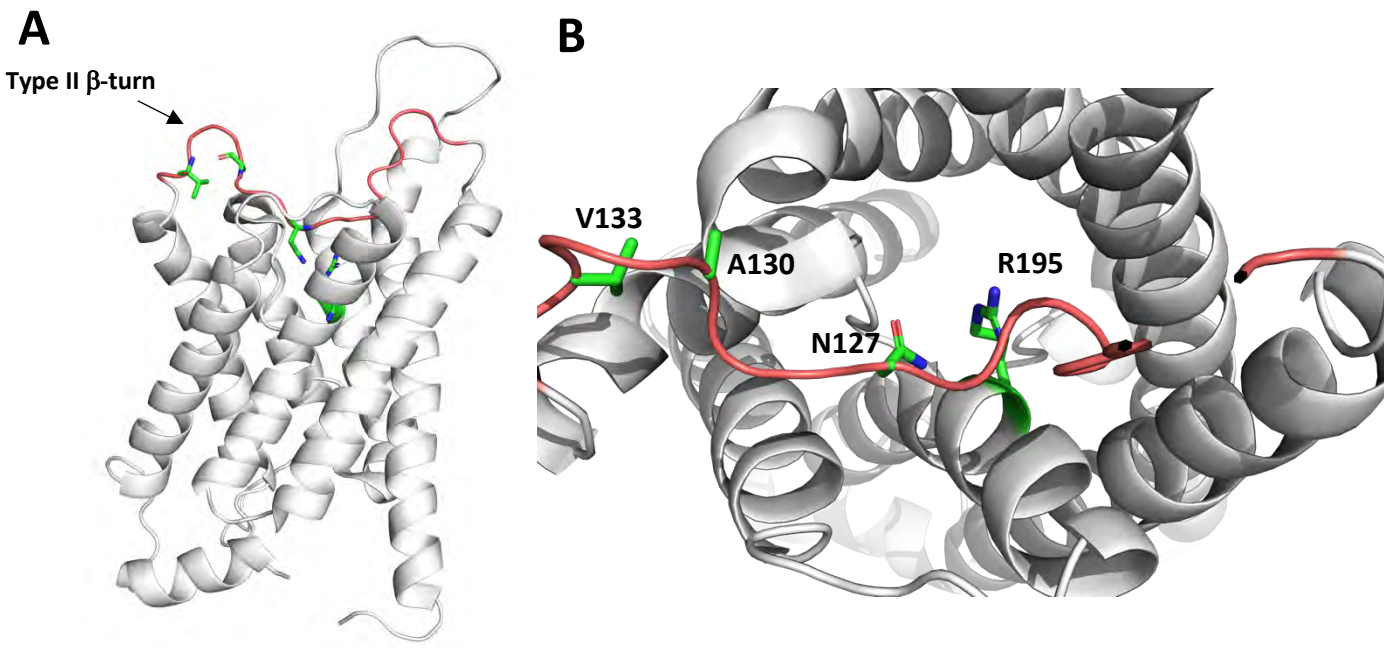

Figure 1-4: *Highlighting the Loop C region of AQP1 and the type II β-turn. Loop C coloured in red, with notable residues coloured by element. Bovine AQP1 shown with human numbering. Figure made in PyMOL. PDB ID: 1J4N.*

1.4 hAQP1 physiology

Human aquaporin 1 was the first aquaporin discovered, and many years later structural elucidation has yielded many atomistic insights into its water transporting functions. However much less is known about its involvement in processes that are only secondary to its main water channel function. To date, new literature suggests hAQP1 is involved in several processes integral to many cancers [116–118]. Tumour angiogenesis, migration, and adhesion are among a few of the processes that AQP1 is involved in [119–122]. Although AQP1 has been found to be highly expressed in several differing cancer cell types, information regarding its trafficking or regulation is sparse. Even sparser is the information surrounding AQP1's accompanying protein-

protein interactions that govern its subsequent trafficking and regulation surrounding these cancer processes. This reveals a gap in the research knowledge, with only a few protein interacting partners identified through molecular biology and biochemistry methods, but very few direct results proving any kind of physical interactions. To fully understand these mechanisms, these protein-protein interactions must be fully explored to gain a deeper understanding into not only AQP1's involvement in disease processes, but also into the other 12 AQPs that are involved in physiological processes that extend far beyond just water transport.

Of recent interest in the literature has been AQP1's crucial involvement and upregulated expression in lung, breast, and colon cancer cells [123–125]. Increases of AQP1 expression have especially been found in breast cancer cells, where this high expression has been linked to a poor prognosis factor in this specific cancer [37]. AQP1 gene knockdown studies in mice and breast cancer cell lines have also revealed that removal of AQP1 expression suppressed the proliferation and migration abilities of the cancer cells *in vitro* [126]. Given this increasing evidence, some work has also been done on trying to establish the mechanisms behind AQP1's involvement in breast cancer. A study using anthracyclines – a common chemotherapy drug – found that breast cancer patients with a markedly higher AQP1 expression ended up with better clinical outcomes to anthracycline treatment as opposed to patients with lower AQP1 expression [127]. This was followed up with an exploration into the specific protein-protein interactions that governed this behaviour. It was found that AQP1 may be interacting with β-catenin, a central protein component to the Wnt signalling pathway, and towards the progression of tumours [128,129]. β-catenin itself contains 12 armadillo repeats, and its these repeats that AQP1 competes for interaction alongside glycogen synthase kinase 3 (GSK3) – another protein integral to the Wnt/β-catenin signalling pathway. The overexpression of AQP1 and its binding to β-

catenin ultimately prevents its degradation; how this interaction modulates the Wnt signalling pathway in the end is still to be explored, however.

Another recent discovery has expanded upon AQP1's involvement in the membrane and cytoskeleton of erythrocytes. Two landmark papers published on the intracellular ankyrin complex have revealed the structure of a multi-protein complex anchored by the spectrin-ankyrin mediated cytoskeleton. The cryo-EM structures had revealed the unexpected participation of aquaporin-1 within the complex, however not entirely surprising given studies involving Forster resonance energy transfer (FRET) that show that AQP1 is located within 8 nm of the Band-3 within cells [130]. Further evidence that AQP1 also contains two CAII binding motifs that are homologous to CAII binding sites on Band-3, shows that a proposed potential "CO_2 metabolon" exists within this cluster of proteins that ankyrin organizes together in the red blood cell membrane [105,108,131]. This localized cluster of proteins links the CAII catalyzed reaction of CO_2 and H_2O to bicarbonate (HCO_3^-), which is then channelled by band-3 either to or from the cytosol [132–134]. AQP1 is known to transport both substrates of CAII, water and CO_2, and its location in the ankyrin complex and its potential CAII binding motifs signal a potential role in this proposed CO_2 metabolon [135]. However, some studies have suggested that the facilitated diffusion of CO_2 either by AQP1 or Rh-associated glycoprotein (RhAG – another protein found in the ankyrin complex) may be unlikely, as the experimental CO_2 permeability was unchanged regardless if AQP1 was expressed in the membranes of epithelial cells [136]. Molecular dynamics simulations have also demonstrated that the bilayer lipid membrane itself can facilitate CO_2 diffusion at a much lower free energy cost than through AQP1's pore [96].

Finally, it is worth mentioning that some aquaporins may be involved in Alzheimer's disease, the crippling condition characterized primarily by the association of senile plaques in the brains of the individuals afflicted by the disease. It has been found that both AQP1 and AQP4 are abnormally expressed in the early stages of Alzheimer's disease, suggesting a dysregulation of water flux at some point in the disease progression[137–139].

1.5 Membrane mimetic systems

Unlike for soluble proteins, a discussion surrounding membrane proteins must always equally emphasise the role of the encasing lipid environment, especially considering integral MPs. This is because without the hydrophobic lipid environment, integral MPs, like aquaporins and rhodopsins, are either entirely non-functional or completely mis-folded into aggregation without lipids or a detergent that mimics the same properties. To isolate and study this class of MPs either biophysically or biochemically, one must be familiar with the available types of detergents or membrane mimetic systems.

Classically, MPs are isolated with detergents [11,14,140–142]. Detergents, a synonym for surfactants, are wetting agents that lower the surface tension of a liquid, as well as the interfacial tension between two liquids. Detergents are organic compounds that consist of a long non-polar alkyl chain, and a charged or polar headgroup [140]. These two parts of the molecule can vary to a great extent depending on whether the detergent is ionic, non-ionic, or a zwitterionic compound. Ultimately, detergents are used to destabilize and solubilize the membrane in which the target MP resides, which frees the MP and suspends it into solution [12]. Detergents solubilize MPs by forming micelles, with the hydrophobic tail segments of the detergent associating with the hydrophobic regions of the MP, meanwhile the hydrophilic headgroups of

the detergent interfacing with the water molecules, thereby constructing a ring-like structure around the MP [143–145]. Detergents only form micelles above the critical micellar concentration (CMC), which is dependent on both temperature and concentration [**Figure 1-5].** Below or above this concentration the detergent exists in either a monomeric or a liquid crystalline state, respectively [140]. These structural properties of detergents enable them to solubilize cell membranes efficiently and effectively. Choosing which type of detergent to use and at what concentration is initially a trial-and-error process, as each membrane protein behaves differently, and whichever cellular membrane system one is using will also impact the choice in detergent.

Although detergents have been successfully used in the past towards the isolation and study of membrane proteins, they are not without their own drawbacks. For one, the detergent often completely removes the surrounding lipids from the MP during its isolation – apart from some cases in which lipids are bound very tightly to the MP [146]. This removal of the lipids disrupts possibly crucial protein-lipid interactions, which may impact the overall fold of the membrane protein, which in turn could impact its function. The crystallization of membrane proteins, either through the lipidic cubic phase, bicelles, or the high lipid-detergent concentration methods, requires the use of detergents, and these structures could be misleading as the impact of delipidation is not fully realized yet [147,148]. As well, since detergents surround the MP in a sphere-like fashion, they could obscure or hinder flexible loop regions which are often exposed to solvent, so certain functional or protein-protein interaction studies may be infeasible using this system. As well, some detergents may not be compatible with several downstream processes in purification, and so must often be removed from the solution, either by dilution or dialysis methods [149].

These issues surrounding detergent-based isolation methods have paved the way for the realization of new membrane mimetics, systems that retain the protein-lipid interactions to varying extents. A few of these systems will be discussed and introduced, with the most relevant and commonly used in the literature being the amphipols, MSP-based nanodiscs, and the polymer nanodiscs, the last category of which includes the styrene-maleic acid (SMA) polymers which will be a significant focus of this thesis.

Firstly, although not a new system, and used in conjunction with detergents, proteoliposomes have been a standard system in which to study MPs. It requires first isolating the MP from the membranes via detergents, but then reconstituting the MP back into liposomes designed with one's choice of lipids and corresponding protein: lipid ratios [150,151]. In this method, a MP can be reconstituted into a proteoliposome with a choice of lipids that are more suitable or more native for the MP, rather than incorporating the lipids found from the expression systems membranes. These proteoliposomes have been used in solid-state NMR and cryo-electron microscopy studies, among others [150].

Amphipols are short amphipathic polymers that have been designed to solubilize MP's [152]. Since they contain both hydrophilic and hydrophobic elements, they also belong to the wider class of surfactants, and function in a similar way to detergents. These amphipols self-assemble in solution into globular complexes consisting of usually 3-4 molecules [153]. These high molecular weight macromolecules have been shown to have a high affinity for hydrophobic particles, making them very suitable for the solubilization of transmembrane proteins [154]. A8-35 is likely the most well-studied amphipol. It is a poly-acrylate-based polymer and contains alternating carboxylate groups between the octyl chains and acrylate moieties. These carboxylate

groups have a pK_a of around 6.0, ensuring that A8-35 is soluble in water at a pH range that is amenable for many membrane proteins [152,155].

Thus far, amphipols have been used to isolate a variety of membrane proteins, including the sarcoplasmic reticulum Ca^{2+}-ATPase, the *E.coli* outer membrane protein A (OmpA) and an *E. coli* diacylglycerol kinase to name a few [156–160]. The native fold and activity were retained to various degrees by the amphipols, but most often performing better than their detergent counterparts. Amphipols have also been explored for use in solution NMR studies of membrane proteins, particularly in combination with cell-free protein synthesis methods [161–163].

Another widely used membrane mimetic system is the membrane scaffold protein (MSP) nanodisc. This method is more involved and laborious, however yields a protein-lipid environment that is more akin to the lipid bilayer than amphipols and detergents. MSP nanodiscs are a non-covalent, discoidal assembly of phospholipids, the target membrane protein, and MSP [164,165]. MSP itself is a genetically engineered protein based on the human serum apolipoprotein, with two MSP molecules wrapping around the nanodisc. Depending on the type of MSP used, nanodiscs can be also tuned in size as they range from roughly 9 to 17 nm in diameter [165].

Nanodiscs confer several advantages over the strictly detergent method, albeit still requiring the use of detergents in the process. Firstly, they are more stable than detergent micelles and bicelles, and have already been used to successfully reconstitute rhodopsins, ion-transporters, and several cytochrome P450s [166–168]. Nanodiscs can reconstitute membrane proteins within a similar lipid bilayer environment, and these lipids can be chosen and customized to a great degree. These discs are often quite small, averaging around 10 nm in diameter, and are compatible with many biochemical and biophysical methods. To this end, a

great deal of effort has been put forth to incorporate this method into isolating membrane proteins and characterizing them with this new system. So far, MSP nanodiscs have proven beneficial towards studying MPs in both solution and solid-state NMR, yielding spectra well-resolved enough to perform residue backbone assignments [169–171]. The discs themselves are both dynamic enough for solution NMR, and massive enough to be ultracentrifuged at high speeds to be pelleted for solid-state NMR. Nanodiscs have also been used in cryo-electron microscopy studies, further highlighting how versatile this system is for membrane protein characterization [172–174].

Although MSP nanodiscs have great potential towards biophysical studies of membrane proteins, there are some disadvantages inherent to this method as well. Even though the final nanodisc product is void of any detergent, detergent is still needed to initially solubilize the MP from its over-expressed membrane environment before assembly into the disc [175]. After solubilization, the nanodisc must be self-assembled using the appropriate ratio of membrane protein to lipids to MSP, and the amount of lipids being displaced by the membrane protein in the disk must also be accounted for [150]. This reconstitution does require some trial and error after an initial ratio estimate is determined for the MP of interest. As well, MSP nanodiscs interfere with protein concentration estimation using UV-Vis spectroscopy, as they absorb at the same wavelength as the encapsulated protein within the discs; 215 nm and 280 nm. Similar, secondary structure determination via circular dichroism (CD) spectroscopy can also be impacted given the helical nature of the MSPs.

Finally, another recently developed group of membrane mimetics is the polymer nanodiscs. This group of nanodiscs can vary widely in their properties as polymers can be easily derivatized with functional groups, however the premise of the resulting nanodiscs is the same as

the MSP nanodiscs – the polymer wraps around a discoidal lipid bilayer that entraps the MP target [176]. One key caveat with the polymer nanodiscs however, is the complete absence of detergent in the isolation of MPs – these polymers can solubilize the cell membranes directly [177,178]. The first significant co-polymer used for the direct solubilization of cell membranes was the styrene-maleic acid co-polymer, and since then a multitude of other polymers have been synthesized to expand the repertoire of solubilizing polymers.

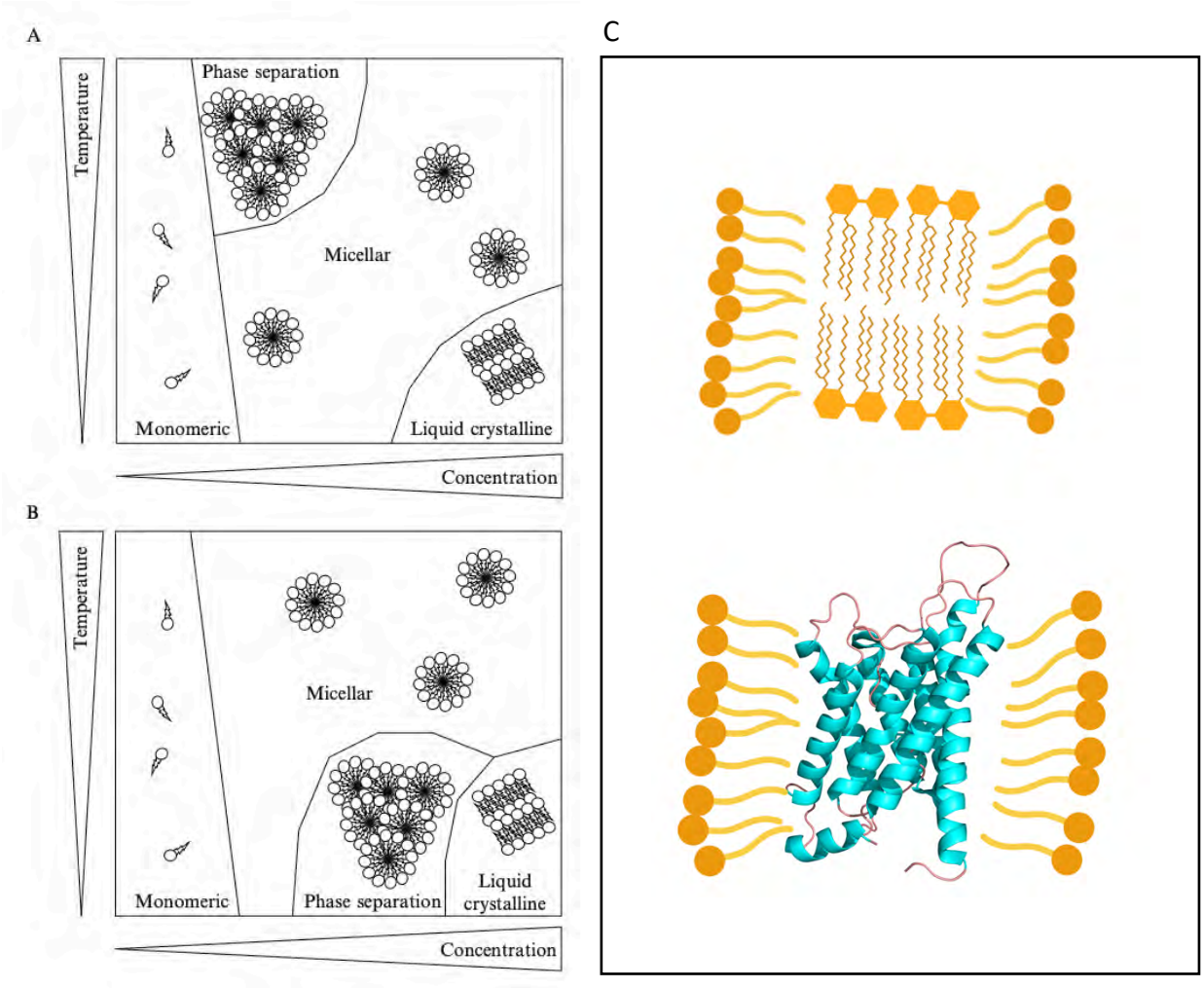

Figure 1-5:(A-B) *A phase diagram outlining the minimal critical micellar concentration of detergents. (**A**) a detergent phase separation with a lower consolute boundary, a category which*

*non-ionic detergents typically fall into. (**B**) a detergent with a higher consolute boundary, a group containing mostly zwitterionic and glycosidic detergents. Above this CMC higher order nonmicellar structures can be observed and are usually liquid crystalline in nature. Figure modified from Linke, 2009, with permission from Elsevier. (**C**): Schematic showing lipid-detergent (top) and protein-detergent (bottom) complexes formed after solubilizing cell membranes with a surfactant.*

1.6 Styrene maleic acid co-polymer and lipid nanoparticles

The SMA polymers are a new class of polymers that were first reported in 2009, when they were used to solubilize the 7-TM protein bacteriorhodopsin into 11 nm sized discs, while retaining native secondary structure and activity [179]. This provided a new means towards safely isolating MPs while also retaining the lipid environment, and these discs have been since been termed the styrene-maleic acid lipid nanoparticles (SMALPs) [179,180]. The water-soluble SMA polymer itself is derived from styrene-maleic acid anhydride (SMAnh) co-polymers, which occurs through a hydrolysis reaction which converts the anhydride to two carboxyl/carboxylate groups [181,182]. The SMAnh co-polymers are synthesized via a free-radical polymerization reaction from styrene and maleic acid monomers.

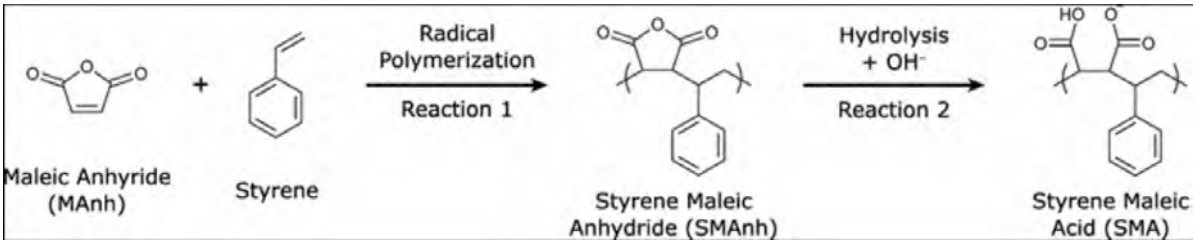

Figure 1-6: *Reaction pathway for the synthesis for the styrene-maleic acid co-polymer. The schematic follows a two-step reaction scheme, where the first involves a radical polymerization of styrene and maleic anhydride. The resulting styrene maleic anhydride (SMAnh) is then hydrolyzed with a strong base, often NaOH is used, which affords the final product. Figure was adapted from Dorr et al., 2015, [175], with permission from Springer.*

The properties and solubilizing capabilities of the SMA polymer are also dependent on the ratio of the styrene to maleic acid units. This ratio of styrene to maleic acid is determined through the initial amount of each in the polymerization reaction **[Figure 1-6]**. There are a few ratios available through various chemical suppliers, these being the 1:1, 2:1, 3:1, and 4:1, although the 2:1 SMA is the most widely used and the most versatile towards membrane protein solubilization. These ratios have implications towards the hydrophobicity and hydrophilicity of the polymer overall. A higher mol% of styrene units within the polymer impart a higher hydrophobic nature on the polymer. As well, the two carboxyl groups of the maleic acid groups have differing pK_as, with the first pK_a being ~5.5, and the second closer to 10. Of course, these pK_as are also impacted by the fraction of maleic acid monomers present in the polymer, with the lower 1.4:1 SMA ratio having carboxyl pK_as of ~4.4 and 9.0, respectively [183]. The charges on these carboxylate groups are thus also pH-dependant, and the pH and ionic strength of a solution will govern whether the polymers are soluble in water or not. At neutral and high pH, electrostatic repulsion between the negatively charged carboxyl groups allows the polymer to adopt a random coil conformation, and thus becoming easily solved in water. When the pH is low, and this effect starts to occur as high as a pH of 6, the carboxyl groups become protonated

and lose their charge, and the hydrophobic effect causes the polymer to fold into a globular conformation, leading to aggregation [180,181].

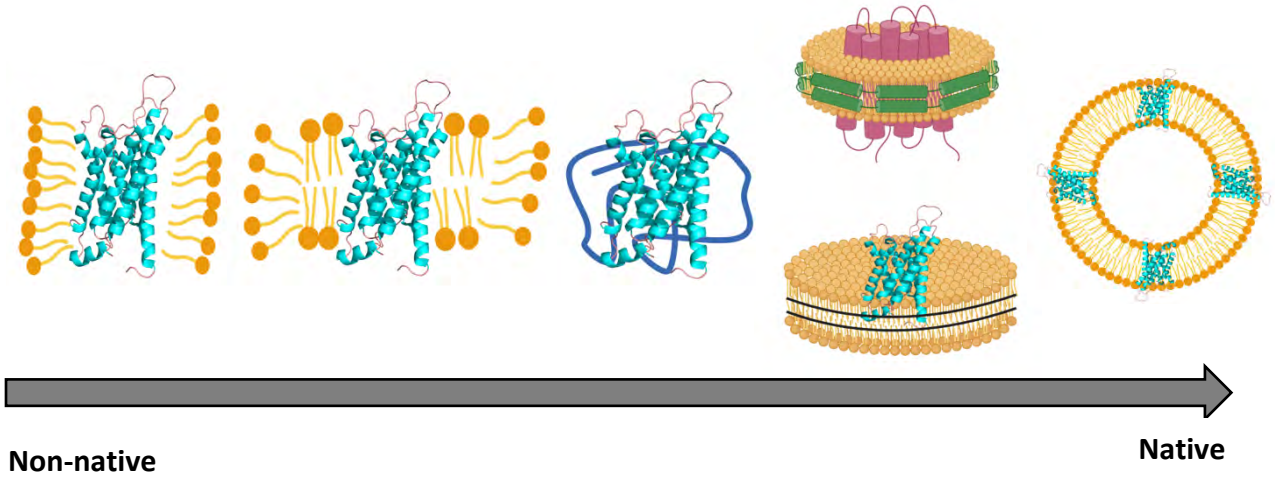

Figure 1-7*: Outlook of the different membrane mimetic systems currently, ranging from least native lipid environment to most native lipid environment. From left to right: protein-detergent micelles, protein-lipid-detergent bicelles, amphipols, nanodiscs (Top: MSP, bottom: SMALPs) and liposomes.*

Various studies have observed the properties of the different SMA polymer ratios and their solubilization efficacies. In terms of overall structure of the SMALPs, small angle neutron scattering and various transmission electron microscopy images have shown that the resulting SMALPs are between 10-25 nm in diameter **[Figure 1-7]**, depending on the specific ratio of SMA used and the conditions during solubilization [184]. Quite consistently, the 2:1 SMA ratio has the highest solubilization efficiency when solubilizing cell membranes [185]. Many studies have also demonstrated SMA's effectiveness in extracting MPs from a wide variety of bacteria,

yeast (including *E. coli* and *Pichia pastoris),* insect, plant and even mammalian cells [186–188]. Although SMAs can solubilize a wide variety of differing membranes, their efficacy varies depending on the membrane and how densely packed the target protein is in the membrane [186].

A few experimental and molecular modelling approaches have investigated the mechanisms behind how the SMA co-polymers solubilize membranes. Molecular modelling studies involving both the 2:1 and 3:1 SMA co-polymers found that the SMA initially exists in solution as a globular cluster, and upon interacting with the membrane, begins to disaggregate within the membranes [189–191]. On the timescales used, different pathways were observed for the 2:1 and 3:1 ratios, with the 3:1 co-polymer forming nanometer-sized protrusions from the membrane, postulated to be the early-stages of the SMALP discs. The 2:1 co-polymer had induced large-scale deformations with the bilayer, often forming nanometer-sized water pores [191]. Building on these theoretical simulations, Bjornestad *et al.* set out to determine these mechanisms experimentally using small angle X-ray scattering [192]. They used the 3:1 co-polymer and found from SAXS data that indeed the polymer formed globular aggregates in solution initially. The hydrophobic styrene groups on the polymer favour a more collapsed conformation, shielding these non-polar groups from unfavourable interactions with water. Upon the polymer being added to either 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC) or 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (POPC) vesicles, the SMA immediately associates with the vesicles and adapts a longer extended chain conformation, where the styrene groups associate with the acyl groups of the lipids. This leads to breaks in the bilayer, where then more co-polymer can insert itself, forming mixed polymer-lipid vesicles. At the final stage, excess polymer then wraps around these mixed lipid-SMA discs, finally forming the SMALPs. This

intermediate mixed state is especially important, as it was found that higher concentrations of SMA than estimated were actually needed to appropriately solubilize the vesicles, and this is likely due to the fact that the polymers first saturate the bilayer before wrapping into discs [192].

Thus far, SMALPs have been used to characterize membrane proteins using cryo-electron microscopy, mass-spectrometry, and solution and solid-state NMR spectroscopy. So far, several classes of therapeutic targets have been isolated and characterized using SMALPs, including G-protein coupled receptors (GPCRs), ion-transporters, and channels [193]. The first GPCR to be isolated was the human adenosine $A2_AR$ receptor, which displayed similar ligand binding activity in SMALPs to that within *P. pastoris* membranes [194]. The SMALP platform has also yielded several cryo-EM structures. The structure of *E. coli* multidrug efflux transporter AcrB was solved to sub-nm resolution and was in agreement with other higher-resolution structures solved by X-ray crystallography [195,196].

Although the SMALPs are too large for full structural characterization using solution NMR, as they do not tumble rapidly enough in solution, some recent work has achieved well-resolved spectra of proteins containing large soluble domains, as these are solvent-exposed and may be flexible enough to be resolved using solution NMR. From the work of Ramamoorthy *et al.,* both a cytochrome-b5 and CYP450-reductase were extracted from *E.coli* membranes and yielded well-resolved spectra [197,198]. In both cases, the SMA polymer was derivatized with certain functional groups so that favourable electrostatic repulsions remained, as the outer membrane domains were likely to interact with the polymer. In the case of CYP450-reductase, a non-ionic pentyl-inulin polymer, a derivative of the SMA co-polymer was chosen to solubilize this protein as well. This further highlights the functionality of SMA, as the polymer backbone

itself can be derivatized with both charged and non-charged functional groups, yielding a polymer that can be customized exactly to the MP of interest [199].

In addition, there are quite a few more advantages in choosing the SMA co-polymer in solubilizing a MP of choice. Firstly, SMA polymers are able to directly solubilize MPs from their membrane environments without the need for detergents [189]. These resulting SMALPs are then easily purified using affinity chromatography with an N- or C-terminal tag, such as a polyhistidine tag (His-tag). As such, they are also amenable to a variety of biophysical studies, as opposed to the peptide bearing MSP nanodiscs, SMALPs are compatible with CD-spectroscopy. SMALPs also allow access to both sides of the MP, whereas more traditional liposome or 2D-crystalline systems have one or more faces buried. This is advantageous towards protein-ligand or protein-protein interaction studies, where the N- and C-terminal tails must be solvent-accessible, for any binding to occur. Some detergent systems are also not perfect for interaction studies, as the detergent may bury more of the protein than when the MP is inside a lipid-embedded system.

Although the SMALP system is indeed quite versatile, there are some disadvantages to working with this system as well. One major disadvantage for the SMA system is its incompatibility towards divalent cations, such as Mg^{2+} or Ca^{2+} ions, which could interfere with proteins that require these ions for their function. SMA has an affinity for these ions and once bound, the nanodiscs will aggregate out of solution. Aggregation was also observed for a 4 mM concentration of Ni^{2+}, Co^{2+}, and Zn^{2+} [187]. This is particularly impactful towards nickel ions, as this could cause some aggregation during the nickel-binding step of protein purification, prior to affinity chromatography. Likewise, SMALPs tend to aggregate out of solution in an acidic pH, and typically the pH of the buffer is retained above 7 to prevent any aggregation from occurring

[200]. Finally, in terms of protein quantification of the SMALPs, the styrene-moieties do absorb at the same wavelength of 280 nm as is typically used for whole protein concentration determination [201,202].

1.7 Scope of thesis

This thesis will present the cumulation of work from two distinct projects, with aquaporin-1 being the consistent focus of the biophysical characterization work. Before summarizing and discussing the results, *Chapter 2* will provide an overview of the theory behind the various experimental methods used to acquire the data on the biological samples. This will include a thorough discussion on Fourier transform infrared spectroscopy, and more brief discussions surrounding differential scanning calorimetry and solution nuclear magnetic resonance spectroscopy.

After this, *Chapter 3* will present the novel assembly and reconstitution of AQP1 into the styrene-maleic acid lipid nanoparticles. This will include a summary on the various solubilization trials and subsequent stability assays before resulting in the final, 37˚C solubilization protocol. These discs were also characterized both biochemically and biophysically, ensuring that AQP1 was indeed present inside the discs after purification, and that the appropriate size for solution NMR studies was achieved. AQP1 was then isotopically labelled with NMR-active nuclei nitrogen-15, and the resulting NMR spectra were well-resolved, with multiple dispersed peaks appearing in the region corresponding to a random coil structure, heavily suggesting these peaks belong to AQP1's C-terminal tail. This novel method resulted in a successful membrane mimetic platform that can be used towards future protein-protein

interaction experiments that involve this exposed and flexible C-terminal tail, a region in which there are many important physiologically relevant partners to choose from.

Finally, *Chapter 4* will detail the second body of work outlined in this thesis, work that involves the mutagenesis of certain amino acids within the AQP1 sequence, that is not only conserved among other AQP1 homologs in different species, but also among other classical and some unorthodox AQPs. These residues were chosen and rationalized from prior solid-state NMR experiments performed by Xiao Peng and from close analysis of the high-resolution structures available. These studies highlighted several crucial regions where hydrogen-bonding may contribute significantly towards the stability of the overall fold of AQP1. These mutagenesis findings have consequences on the entire AQP family and may have implications towards future drug discovery efforts targeting this family of proteins.

Chapter 2

A review of the biophysical methods

Chapter 2: A review of the biophysical methods

Overview

In this chapter, an overview of the experimental methods used to acquire the results in later chapters will be presented. A formal, but succinct description will be provided on Fourier transform infrared spectroscopy and its application towards determining and analyzing the secondary structure of proteins, as well as their dynamics. As well, a discussion on differential scanning calorimetry to probe the thermal stability of membrane proteins will be introduced, as well as a section devoted to dynamic light scattering. Lastly, a brief overview of nuclear magnetic resonance spectroscopy will be given, along with its applications towards membrane proteins and nanodiscs, and the specific experiments used in this thesis.

2.1 Basic principles of vibrational spectroscopy

Infrared spectroscopy has over the years become an increasingly powerful analytical technique used across scientific fields encompassing pharmaceuticals, materials science, and astronomy. Infrared (IR) spectroscopy is used to analyze and provide a molecular fingerprint of a material or substance, based on the vibrations of the atoms in the molecules present in the analyte [203]. An IR spectrum is acquired by passing IR light through the material and measuring the fraction of the intensity that is absorbed at a particular energy level. This absorption of IR radiation is the result of different chemical bonds absorbing at different wavelengths, which allows a high degree of chemical sensitivity towards structural elucidation.

IR spectroscopy is thus a powerful tool, and the instrument, theory, and its application towards characterizing protein secondary structure will be discussed thoroughly in this section.

2.1.1 Fourier transform infrared spectroscopy

The most predominantly used FTIR spectrometer in the modern era is based on the Michelson interferometer. Simply put, the Michelson interferometer is composed of a heat source – the source of infrared light – a beam splitter, and the two mirrors, one stationary and the other mobile [204,205]. A key difference between the Michelson interferometer and older spectrometers is the beamsplitter, which is made of a material that especially suited for a region of infrared to be examined. For the mid-infrared region, the beamsplitter is normally composed of a substrate of infrared-transparent material of either potassium bromide or caesium iodide, which is then coated with a layer of germanium or iron oxide. This beamsplitter splits the oncoming IR radiation into two equal beams, where 50% of the incident light will be directed towards the fixed mirror, and the other 50% of light will be directed towards the movable mirror **[Figure 2-1].** These beams of light are then reflected to the beamsplitter, where they recombine and interfere, and then pass through the sample. For each wavelength within the initial source beam, the beams returning from both mirrors undergo either constructive or destructive interference, where the result is that the light that reaches the detector results in the interferogram [206].

The Fourier transformation of the interferogram results in a transmission spectrum, which is the intensity of transmitted light at each wavenumber. The advent of the Fourier transformation and increasingly fast computing resulted in the quick advancement of the FTIR method, and its adoption within many disciplines. The Fourier transformation allows for the

digitization of the interferogram into a conventional spectrum that can be easily understood. Once the spectra are collected, the scale can be presented as a percentage of total transmittance (% transmittance) or by absorbance. The transmittance view is often used for spectral interpretation, meanwhile the absorbance is used for more quantitative work. When two transmission spectra are taken, one for the sample and one for of an empty window for reference, the absorbance spectra can be calculated as follows

$$A = -\log\left(\frac{T_s}{T_r}\right) \quad \textbf{(1)}$$

where T_s and T_r are the transmission values for the sample and reference, respectively.

In terms of IR sources, typically a Globar or Nernst source is used for the mid-infrared region, while for the far- and near-infrared regions, either a high-pressure mercury lamp or a tungsten-halogen lamp would be used. Regarding detectors in the mid-infrared region, the standard detector of choice would be a pyroelectric device consisting of deuterium triglycine sulfate (DTGS) in an alkali halide window that is temperature resistant. A detector composed of mercury cadmium telluride (MCT) can be used for work requiring a higher sensitivity, although this detector needs to be cooled to liquid nitrogen temperatures.

The infrared spectrum is divided into three regions, the far-infrared (<400 cm^{-1}), the mid-infrared (4000 – 400 cm^{-1}) and the near-infrared (13,000 – 4000 cm^{-1}). Infrared spectroscopy measures the vibrations of molecules within this region in the electromagnetic spectrum. Specifically, it measures asymmetric molecules or pairs of atoms that have an intrinsic dipole moment. These asymmetric vibrations differ depending on the functional group, and thus each group has a characteristic infrared absorption band [207,208]. These band frequencies are determined by the energy gap between the ground state and the first excited vibrational state, and

only occur via a change in the dipole moment of the asymmetric molecule upon absorbance of the incident infrared light. If a molecule or functional group is symmetrical, there will be no absorbance of IR light, and so it is termed IR-inactive or translucent.

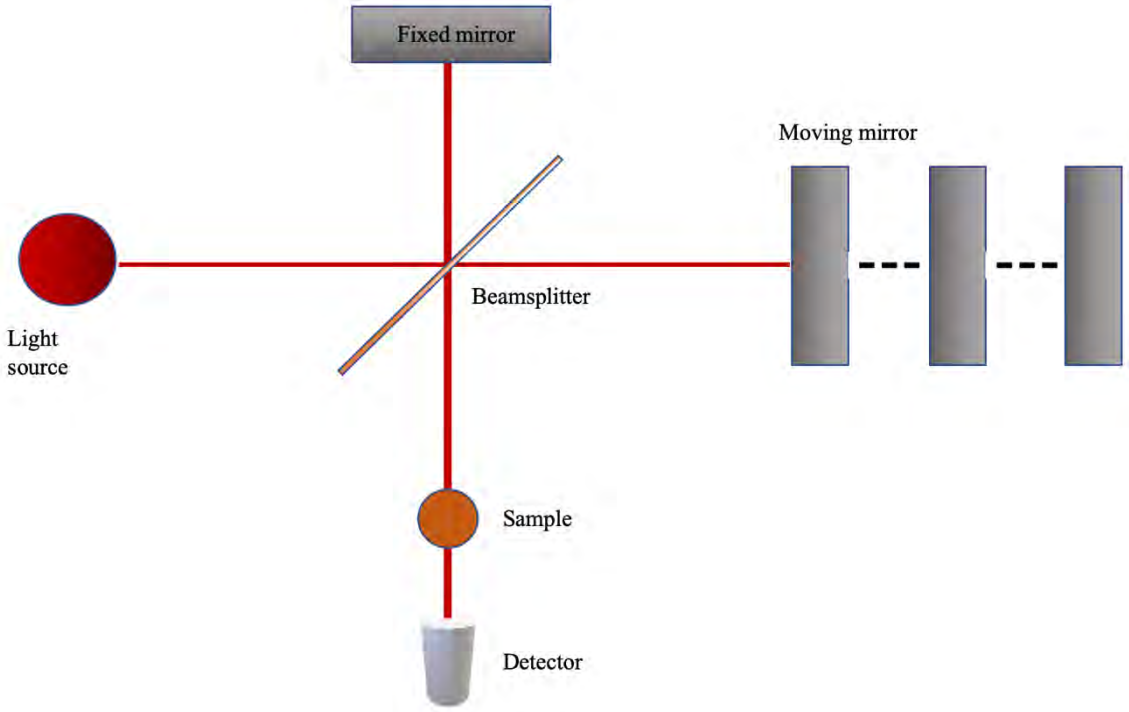

Figure 2-1: *A simple schematic of a Michelson interferometer.*

There are two main types of vibrations that occur in the mid-infrared region (4000-1000 cm^{-1}) – the main region for proteins and amino acid characterization – stretching vibrations and bending vibrations. Stretching vibrations occur along the chemical bond between two substituent atoms, and involve bond-length changes, while bending vibrations evoke changes involving the bond angles, with δ denoting in-plane bends and π denoting out-of-plane bends [209]. These stretching vibrations of a diatomic molecule can thus be modelled by a harmonic oscillator

$$v = (1/2\pi c)\sqrt{\left(\frac{k(m_1 + m_2)}{m_1 m_2}\right)} \quad \textbf{(2)}$$

where the bond strength is represented by the spring constant k, the two point masses m_1 and m_2, the speed of light c, resulting in the wavenumber of vibration v (cm^{-1}). Consequently then, the vibration frequency depends on the relative bond strength between the two atoms. Triple bonds and double bonds thus have higher frequencies than those of single bonds. These frequencies are also sensitive to the atom's environment, as nearby electronegative groups or the formation of hydrogen-bonds can affect these frequencies. A carbonyl group involved in a hydrogen-bond can downshift the relative frequency up to 20 cm^{-1}, and even more so in proteins [205]. As well, the larger the change in dipole moment, the larger and more intense the absorption band will be. Highly polar groups like carbonyl or amide groups will tend to have a larger absorption band, due to the presence of the relatively highly electronegative oxygen or nitrogen atoms. Since these vibrational frequencies are also a function of the atoms' masses, the frequencies can shift via isotopic labelling. A proton exchanged for a deuterium will experience a shift in its associated frequency – a technique that has been used extensively to probe exchangeable protons within proteins in solution, and to explore protein dynamics.

2.1.2 Attenuated total reflectance FTIR spectroscopy

Attenuated total reflectance FTIR spectroscopy (ATR-FTIR) is a contact sampling IR method that involves using an internal reflection element (IRE), often in the form of a crystal with a high-refractive index and good IR-transmitting properties. This crystal is usually a

material made up either diamond, ZnSe, or Ge. In this method, the IR beam is passed through this crystal at an angle greater than the critical angle so that total internal reflection occurs. This internal reflection causes an "evanescent wave" that probes beyond the crystals surface, and penetrates through the sample by a depth of 0.5 – 5 μm [205,210]. This IR radiation is then attenuated when absorbed by the sample, resulting in the IR spectrum that is collected by the detector. The critical angle is a function of the refractive indices of both the crystal and the sample, determined by the equation

$$\theta_C = sin^{-1}\left(\frac{\eta_2}{\eta_1}\right) \quad \textbf{(3)}$$

with η_2 and η_1 being the refractive indices of the sample and crystal, respectively.

Two different methods of ATR have been used to study both soluble and membrane proteins. The first, is thin-film ATR (TF-ATR) in which the protein of interest is dried onto the ATR crystal forming a hydrated thin film layer. This thin film consists of a very high protein concentration and increases the signal-to-noise ratio, even with as little as 10 μg deposited onto the crystal [211]. Evidence has shown that the secondary structure of the protein is not perturbed with this method, and even though the sample is dried, the film itself is still quite hydrated. The other ATR method is allowing the protein to adsorb onto the crystal through drying as well (Ads-ATR) [212]. It has been shown that the interaction between the surface of the crystal and the protein is strong enough to keep the protein bound. Ads-ATR has been used to study conformational changes of proteins in changing environments but also in measuring protein-ligand binding.

ATR-FTIR has been proven useful towards characterizing secondary structure in not only soluble proteins but also that of membrane proteins reconstituted into lipid environments [213]. Reconstituted membrane proteins can be dried onto the crystal and the resulting spectra contain a wealth of information from both protein and lipids, without the need of any labelling, isotopic or otherwise. Membrane proteins can thus be studied using either transmission FTIR or ATR-FTIR, with ATR-FTIR being particularly useful in that as little as 100 μg of lipid bilayer films can be dried onto the IDE, resulting in a high signal-to-noise with even less protein. As well, stacked bilayers can be prepared to increase the signal even further.

2.1.3 FTIR of proteins

FTIR spectroscopy is a powerful analytical technique in measuring the secondary structure of proteins, and extensive review of this application exists in the literature [208,214–216]. The IR spectrum of proteins is dominated by the Amide I and Amide II absorption bands, however nine such amide bands occur throughout the mid-infrared region in the spectrum **[Table 2.1]**. The Amide I band consists mainly of the C=O stretching vibration, with weaker contributions from the amide C-N stretching and the N-H bending. The Amide I band is also sensitive to hydrogen-bonding of the C=O group, downshifting the natural frequency of the amide [214]. This Amide I band is generally in the range of 1660-1645 cm^{-1} when the protein is mainly α-helical **[Figure 2-2]**, while β-sheet structures can have amide I absorption bands between 1640-1615 cm^{-1} and between 1670-1695 cm^{-1}. Random coil structures have been shown to occur around 1644 cm^{-1} **[Table 2.2]** [217]. Thus, the amide I band can span from 1700 – 1620 cm^{-1}, and applying various curve-fitting methods can resolve the separate amide I components that correspond to the different secondary structures. This structural sensitivity of the amide I

band allows performing experiments that can monitor secondary structure changes in solution and even in a time-resolved manner. The amide II vibration consists mainly of the out-of-phase N–H bending, along with the C–N stretching vibration. The amide II band is also sensitive to secondary structure, but less so than the amide I band. For mainly helical structures, amide II occurs around 1555 – 1545 cm^{-1}, and for mainly β-sheet structures, these occur around 1535 – 1525 cm^{-1}.

Table 2.1: *Characteristic absorption bands for the amide group on the polypeptide backbone. Wavenumbers are an approximate, as they are influenced by secondary structure, solvent, and environment. Information compiled from Stuart, 2004 and Lorenz-Fonfria, 2020 [205,214].*

Amide band	***Wavenumber (cm^{-1})***	***% assignment contribution***
A	3300	N–H stretching
B	3110	N–H stretching
I	1653	80% C=O stretching, 10% C–N stretching, 10% N–H bending
II	1567	60% N-H bending, 40% C–N stretching
III	1299	30% C–N stretching, 30% N–H bending, 10% C=O stretching, 10% O=C–N bending, 20% other
IV	627	40% O=C–N bending, 60% other
V	725	N–H bending
VI	600	C=O bending
VII	200	C–N torsion

Of note as well, the amide II band also appears as a weaker, much broader band than the amide I band [214]. Most importantly however, is the amide II band's sensitivity to hydrogen-deuterium (H/D) exchange. Deuterium (^2H), the heavier isotope of hydrogen, has twice the mass of the naturally abundant ^1H isotope and has almost identical chemical properties, but will affect the vibrational frequencies of the amide bands when exchanged in solution. Upon deuteration of the

amide N–H group, a new amide II' band appears, downshifted by ~100 cm^{-1}. The deuteration of the amide group uncouples the vibration contributions from the C–N stretching and the N–H bending, which results in largely a C–N stretching mode at 1445 cm^{-1} [218,219].

The amide III and A bands are broader bands that convey less information about secondary structure than the amide I and II bands. The amide A band is made up of entirely the N–H stretching vibration and is quite sensitive towards hydrogen-bonding. However, to observe this band one would need to measure the protein in a dry film state, as otherwise the water O–H stretching vibration obscures this band in solution **[Table 2.3]**. Lastly, the amide III band occurs around 1300 cm^{-1} and is a complex vibrational frequency with multiple contributions from different chemical groups. Due to its relatively low amplitude, less is known about the amide III band as compared to amide I, however secondary structure differences can also be predicted from this absorption frequency [220].

Table 2.2: *Amide vibration frequencies as they relate to proteins' secondary structure, and the frequencies observed when deuterated. (s) and (w) notations refer to strong and weak signals. Information compiled from Lorenz-Fonfria, 2020 and Tatulian 2012 [206,214].*

Secondary structure	***Wavenumber (cm^{-1}) in H_2O***	***Wavenumber (cm^{-1}) in D_2O***
α-helical	1655-1645	1655-1638
β-sheet	1635-1625	~1630
Parallel β-sheet	1638-1632	1636-1630
Antiparallel β-sheet	1638-1623 (s) 1695-1675 (w)	1636-1630 (s) 1680-1670 (w)
β-turn	1685-1655	1675-1640
Random coil	1660-1648	1648-1640

Certain amino acid sidechains also have characteristic absorption frequencies in the mid-infrared region. The most notable frequencies of which occur from the aspartate and glutamate, both of which contains carboxylic side chain groups that can participate in hydrogen-bonding. The carboxylic C=O stretching vibration spans from 1780 – 1690 cm^{-1}, and is directly proportional to the C=O bond length. There are four possible H-bonding interactions, with the **H**O–C=O acting as the H-bond donor being the most energetically favourable conformation [214]. This H-bond where the hydroxyl acts as the donor, downshifts the C=O vibration by 30-40 cm^{-1}. These carboxylic frequencies can be used to measure proton transfer reactions using IR difference spectroscopy, as Asp or Glu are primarily involved in these types of reactions.

Other notable amino acids absorption frequencies include cysteine, which contains a thiol group that has a main absorption band between 2590 - 2480 cm^{-1} in proteins but disappears upon deprotonation or oxidation of the Cys residue. This makes it a unique probe to monitor the redox changes or protonation state of cysteine in IR. A table outlining a few other notable amino acids frequencies is also included **[Table 2.4].**

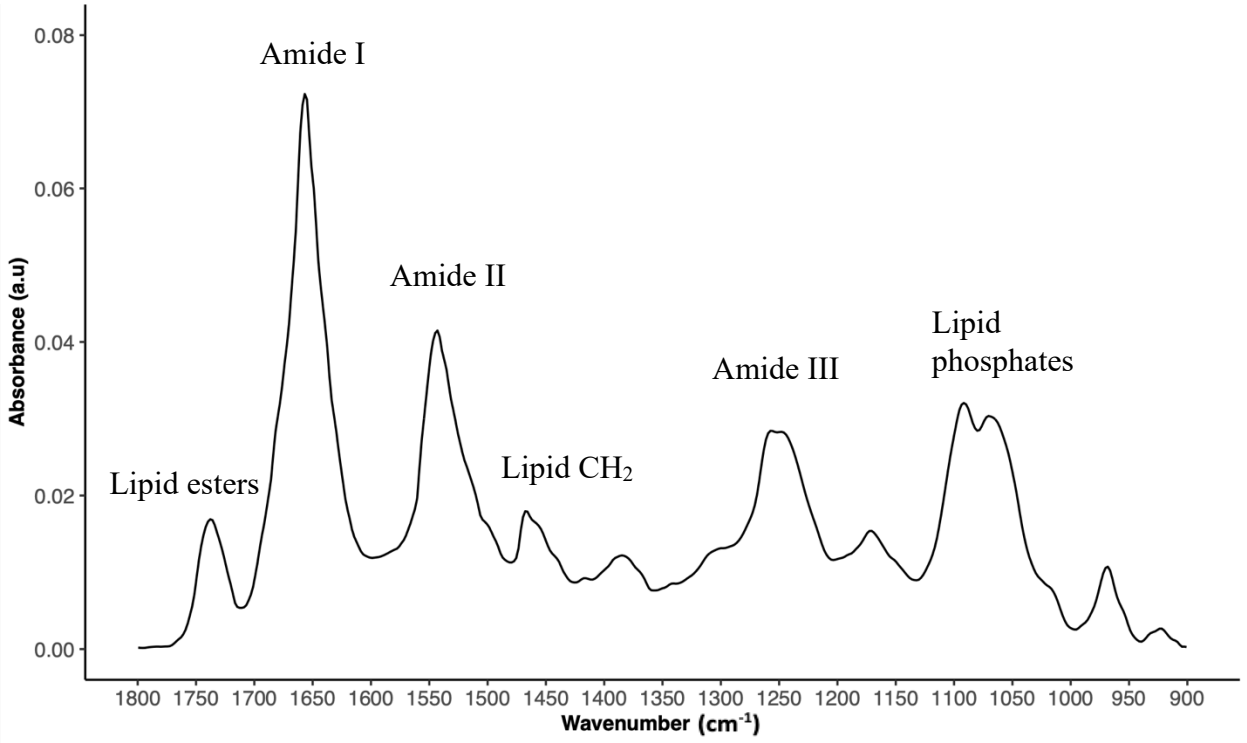

Figure 2-2: *FTIR spectra of the α-helical integral membrane hAQP1 protein reconstituted into liposomes, at a 2:1 protein:lipid ratio, and subsequently dried onto a germanium crystal for ATR-FTIR. Dominant amide I, II, and III bands are observed, as well as the smaller intensity lipid ester, CH_2. Lipid phosphate head groups may have absorbance bands near 1060, 1086, and 1222 cm^{-1}, with the last wavenumber likely contributing to the amide III peak as well.*

Table 2.3: *The most significant absorption bands of H_2O and D_2O in the mid-infrared region. Information compiled from Stuart, 2004 and Lorenz-Fonfria, 2020 [205,214]*

Assignment	***Wavenumber (cm^{-1})***
O-H stretching	3920
O-H stretching	3490
O-H stretching	3280
H-O-H bending	1645
O-D stretching	2900
O-D stretching	2540
O-D stretching	2450
D-O-D bending	1215

Table 2.4: *An example of a few absorption frequencies for notable amino acids and their functional groups. Table modified from Tatulian 2012 [206].*

Amino acid	***Vibration mode***	***Wavenumber (cm^{-1})***
Aspartic acid	COO^- stretching	1574
Glutamic acid	COO^- stretching	1560
Histidine	Imidazole ring	1596
Tyrosine	Aromatic ring	1518, 1602, 1498
Phenylalanine	Ring	1494
Lysine	NH_3^+ asymmetric deformation	1629

2.1.4 Hydrogen-Deuterium Exchange FTIR

Hydrogen-Deuterium exchange (H/DX) experiments involving proteins measure the exchangeability of solvent-exposed protons on a proteins surface. Both the amide I and II bands are sensitive to deuterium exchange, where the amide II exchange can be directly related to a protein's stability. Upon increasing deuteration of solvent-exposed protons, the amide II band decreases while the new amide II' band at ~1445 cm^{-1} increases with intensity. A protein with fixed structure and buried residues will exchange much slower than a protein that has many flexible and open domains, allowing the passage of water through. In this way, protein stability can be monitored through the kinetics of H/D exchange. Full exchange occurs when the amide II band completely disappears from the spectra, as then all the backbone N-H groups will have exchanged to N-D, and just the C-N stretching frequency will remain at 1445 cm^{-1} [221,222].

2.2 Differential Scanning Calorimetry

Differential scanning calorimetry is an analytical technique that measures the specific heat capacity difference between a sample chamber and a reference cell. DSC can provide thermodynamic information on both reversible and irreversible processes, one such example being the conformational equilibrium between the folded and unfolded states of a protein [223]. The heat capacity of the sample is measured as a function of temperature, whereby the instrument contains two identical cells, one with sample and the other filled with just buffer to act as a reference. These twin cells are then heated, or cooled, at a constant rate. The power difference needed to be supplied to maintain the temperature between the cells in Joules s^{-1}, after normalization by the scanning rate in K s^{-1}, is thus a direct measure of the heat capacity difference between sample and reference (ΔC_P) [224]. This difference in temperature is caused by the difference in specific heat capacity between the sample and buffer chambers. Given the C_P of the protein with respect to the C_P of just the buffer solution, the absolute heat capacity (C_{Ppr}) of the protein can then be calculated when the mass of the protein (m_{pr}) is known, as well as the partial specific volumes ($\bar{v}_{pr}$ and v_{sol}), the heat capacity of the solvent used ($C_{p_{sol}}$), and the measured excess heat capacity ($C_{p_{exp}}$).

$$C_{p_{pr}} = C_{p_{sol}} \frac{\bar{v}_{pr}}{v_{sol}} - \frac{C_{p_{exp}}}{m_{pr}} \quad \textbf{(4)}$$

For simple protein systems, often just one peak is observed in the DSC trace. When the protein solution is heated, if the unfolding pathway is mainly cooperative, the resulting thermogram will result in a sharp peak-shaped curve **[Figure 2-3]**. The protein denatures between T_1 and T_2, while the melting temperature T_m is the mid-point of the peak, where 50% of

the protein exists in the unfolded state while the rest still exists in its native state. The resulting peak can then be analyzed further to extract certain thermodynamic parameters of the unfolding pathway. Thus, the enthalpy function can be calculated by integration of the heat capacity curve, the area under the transition temperature.

$$\Delta H = \int_{T_1}^{T_2} C_p dT \quad \textbf{(5)}$$

The ability to measure enthalpy changes and monitor the stability of macromolecules makes DSC a versatile tool towards the study of proteins. Processes such as protein unfolding, aggregation, and re-folding can all be studied using DSC. Protein unfolding information can also be determined by varying the conditions of the DSC experiment. Proteins have differing stability in different buffer conditions, where ionic strength and pH are major factors, and so would have different unfolding profiles and melting temperatures in a DSC experiment. As well, kinetics of unfolding can be measured by varying the scan rates, as proteins may display different T_m through different rates of heating and cooling [225]. Lastly, the effects of mutations in a protein on its overall stability can also be analyzed through DSC. Mutational analyses coupled with DSC can report on the enthalpy values of unfolding, and more unstable mutants may have lower corresponding melting temperatures or display a broader curve due to the loss of cooperativity [226]. As it is for helical proteins, one would expect a fairly sharp curve depending on the nature of cooperativity in the helical packing, as is the case for bacteriorhodopsin [227,228].

There are four models of protein unfolding generally accepted; the reversible two-state, irreversible two-state, partially reversible three-state denaturation with equilibrium, and the general partially reversible three-state denaturation, with the last being the classical Lumry-

Eyring model, where the first step contains two rate constants, k_1 for the forward reaction and k_{-1} for the reverse [225]. This model is written as follows:

$$N \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} I \xrightarrow{k_2} D \qquad \textbf{(6)}$$

The three-state model thus assumes the reversibility of the first stage of unfolding, from native state to an intermediate state (which by itself can be populated by many different conformations or intermediates), where then the reaction from intermediate to the denatured state follows an irreversible scheme. Interestingly, a DSC experiment that involves a re-heating scan can probe some of this reversibility, where the first heating scan stops at a point between T_1 and T_2, where some protein is not yet irreversibly unfolded [221].

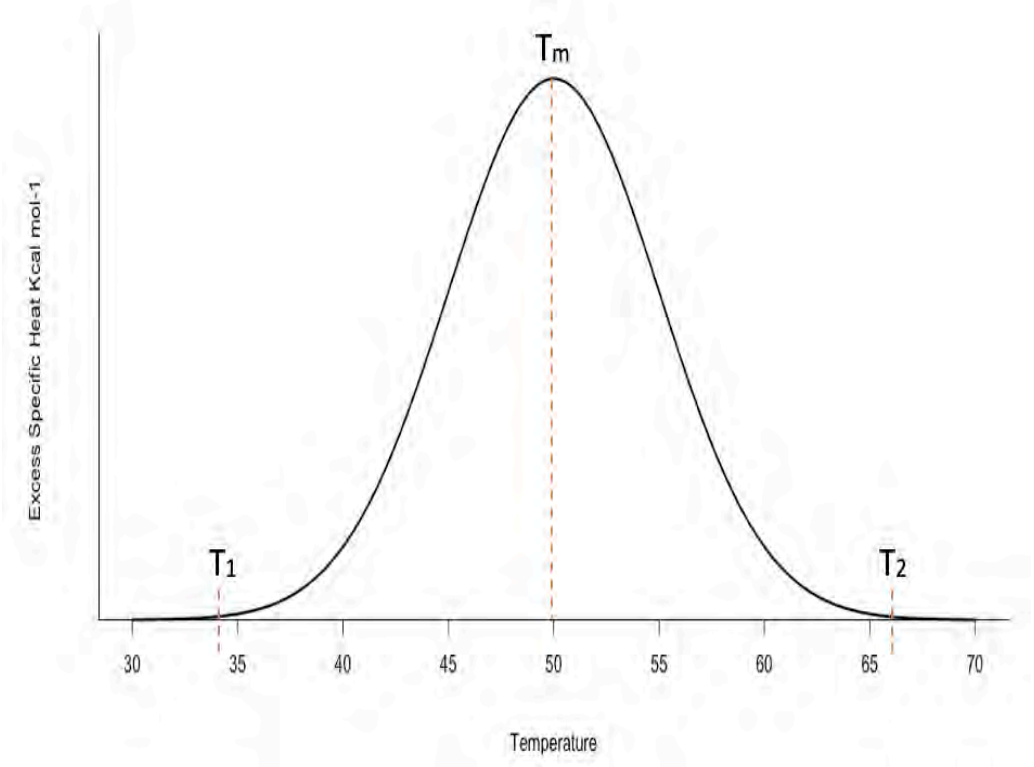

Figure 2-3: *Example hypothetical DSC thermogram of the cooperative unfolding of a protein, after subtraction of the baseline. T_1, T_2, and T_m shown at respective positions on curve.*

2.3 Dynamic Light Scattering

Dynamic light scattering (DLS), also called Photon correlation spectroscopy (PCS), is a technique that studies the diffusion behaviour of certain analytes, and can report on the nanometer sizes of macromolecules [229,230]. DLS measures the amount of light scattered by a molecule in solution, where light encounters the macromolecule, and the scattering intensity is measured by a detector. The intensity fluctuation of the scattered light is then correlated to time (ns - μs) to determine the kinetics of the fluctuation, which is directly related to the diffusion of the macromolecule in solution. Since the diffusion coefficient of the protein can be determined, the hydrodynamic radius (R_h), defined as the radius of a sphere that has the same diffusion coefficient as the protein, can be found through rearrangement of the Stokes-Einstein equation, where k_B is the Boltzmann coefficient (1.38×10^{-23} kg·m^2 ·s^{-2}·K^{-1}), T the temperature, η the viscosity of the medium, and D_t the calculated translational diffusion coefficient,

$$R_h = \frac{k_B T}{6\pi\eta D_t} \quad \textbf{(7)}$$

After data acquisition, there are a few approaches in terms of data analysis that can then determine the size and homogeneity of the protein or macromolecule in question. Namely, for size distribution analysis of the hydrodynamic radius, there is the general intensity-weighted distribution, that can then be converted numerically to either the volume-weighted or number-weighted distribution. Since the intensity distribution is dependent on the intensity fluctuations

given by the measured macromolecules, DLS measurements give the z-average R_h, which is the ensemble collection of particles measured by the scattering.

$$\%I_a = \frac{a^6 N_a 100}{N_a a^6 + N_b b^6} \quad \textbf{(8)}$$

Equation **(8)** gives the intensity distribution ($\%I_a$) for a solution containing a certain amount of N_a and N_b molecules of sizes *a* and *b,* respectively. The intensity-weighted distribution can thus be converted into the volume-weighted distribution, which will give the relative proportion of multiple sizes in a sample, based on their volume or size rather than their scattering intensity. Since higher MW particles will have a higher refractive index and will thus scatter more than lower molecular weight particles, the volume-weighted distribution may be more practically useful, as it will show the abundance of each size of molecule. Equation **(9)** details the volume-distribution.

$$\%V_a = \frac{a^3 N_a 100}{N_a a^3 + N_b b^3} \quad \textbf{(9)}$$

Since mass is proportional to ($size^3$), and mass can be approximated to volume if the sample solution density is uniform, this implies that the volume-weighted distribution is proportional to ($size^3$). Although the volume-distribution is indeed more practically useful, it has some disadvantages as it assumes that all molecules are homogenous and spherical in shape, and that the intensity distribution data it is being converted from is absent of any errors.

Finally, the intensity distribution can also be converted to the number-weighted distribution, which outlines the number of molecules in each bin sorted in a given histogram. This is represented by equation **(10)**, for a given system of N_a and N_b sizes.

$$\%N_a = \frac{N_a 100}{N_a + N_b} \quad \textbf{(10)}$$

DLS is thus a powerful technique to measure the hydrodynamic radius of a protein or macromolecular complex and can also accurately monitor aggregation in solution. Higher molecular weight peaks will occur if aggregation is observed, and the heterogeneity (polydispersity) of a sample can also be monitored.

2.4 Nuclear magnetic resonance spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy is a powerful technique, one that allows the analysis of proteins at an atomic resolution, and may very well be one of the most impactful techniques in modern science. NMR has been used to solve the structures of proteins and small molecules, as well as to study their dynamics. Both solid-state and solution NMR have been used to characterize proteins at length, however in this review just the background needed to understand solution NMR will be discussed, and how it applies to protein-protein or protein-ligand interaction studies will also be covered in the introduction to *Chapter 3.*

2.4.1 Basic overview of NMR spectroscopy

NMR is a widely used technique that takes advantage of the inherent nuclear spin states that nuclei with odd mass numbers have, such as ^{1}H, ^{13}C, ^{15}N, ^{19}F, ^{31}P, ^{29}Si – apart from some

exceptions. Simply put, these spin states arise as each nucleus has an intrinsic angular momentum. A given nuclei with a spin quantum number *I,* will have $2I + 1$ possible spin states in the presence of a magnetic field, which is defined as directional quantization. The ^{1}H and ^{13}C nuclei for example, are ½ spin nuclei, and so in the presence of an external magnetic field B_0, will only have two allowed spin states, -½ and + ½. In the absence of a magnetic field, these spin states will be of equal energy, but once the external magnetic field is applied the degeneracy is removed and energy levels of the spin states will split – this is defined as the Zeeman effect. At thermal equilibrium relative populations of these states, e.g., the proportion of nuclei in each state, follows the Boltzmann distribution, thereby resulting in the net magnetic moment. If perturbed out of equilibrium, it will precess around the direction of the magnetic field with the Larmor frequency, v_L, which is proportional to the magnetic field B_0:

$$v_L = -\frac{\gamma}{2\pi} B_o \tag{11}$$

where γ is the gyromagnetic ratio of a nuclear spin. In NMR experiments, transitions between the energy levels are induced through application of radiofrequency (r.f) pulses. A transition between energy levels will only occur when the frequency of the electromagnetic radiation matches the Larmor frequency. Once perturbed from equilibrium, the magnetization will precess around the magnetic field and this precession is detectable in an NMR experiment. The non-equilibrium state of the magnetization induced by the r.f. pulses, would eventually relax back to the thermal equilibrium state [231]. The detected signal from an NMR experiment is referred to as the free induction decay (FID). A typical protein NMR experiment has two or more dimensions and this along with poor sensitivity, results in long acquisition times. Therefore, the

FIDs of many pulses are accumulated and then transformed. Through this accumulation, the electronic noise averages out, while the true signal increases, proportional to the square root of the number of scans, given by equation **(12).** This often results in NMR experiments being on the length of hours to multiple days.

$$S{:}\,N \sim \sqrt{N_{scans}} \tag{12}$$

2.4.2 The Chemical shift

Nuclear resonance frequencies are affected by their environment, where the electrons surrounding these will influence the resonance. This is observed via chemical shielding, where the effective magnetic field at the nucleus is always less than the applied field B_o. The resonance frequency ν_1 of a given nucleus is thus determined by the following equation:

$$\nu_1 = \frac{\gamma}{2\pi}(1-\sigma)B_0 \tag{13}$$

where $\sigma\ is$ the shielding constant. The shielding constant is a dimensionless quality that is around 10^{-5} for protons but increases as the mass of an atom increases. As well, this constant also increases with the number of electrons surrounding the nucleus. The important conclusion from this is that the chemically non-equivalent nuclei would have distinct resonance signals in the NMR spectrum since they are shielded to different extents by their respective electrons. This nuclear shielding offers the incredible resolution that an NMR spectrum can provide for a sample.

The resulting resonance frequency can then in turn be used to calculate the relative chemical shift δ (ppm) of each nucleus as compared to the reference:

$$\delta_{sample}[ppm] = \frac{v_{sample} - v_{reference}}{v_{reference}} \frac{[Hz]}{[MHz]} \quad \textbf{(14)}$$

where the numerator is in units of Hz, and the denominator in MHz, to give a more convenient number for ppm. Generally, the x-axis of an one-dimensional NMR spectrum is labelled as parts per million (ppm), with respect to an established standard.

2.4.3 Solution NMR and Proteins

Proteins have been studied using NMR for decades now, and the technique has allowed the high-resolution determination of hundreds of protein structures. Most importantly, however, is its accessibility towards studying proteins in solution, which until recently has been applied to only soluble proteins. With the onset of various membrane mimetics, such as nanodiscs, the idea of using solution NMR to study membrane proteins has increasingly gained interest in the literature [197,198,232]. Generally, proteins are expressed in media that contain either ^{13}C or ^{15}N-labelled isotopes, or both. These labelled isotopes, along with naturally abundant ^{1}H, provide the means to detect the amino acids within the protein chain, and can provide bond correlations using 2D and 3D experiments. These multidimensional NMR experiments use a variety of different pulse sequences which enable a transfer of magnetization that produces correlations between signals of different nuclei. These correlations result in chemical shift cross-peaks in the spectrum, which form the basis for resonance assignments, and have allowed for the 3D structural determination of many proteins using NMR [165, 227].

Applications of solution NMR requires rapid tumbling to average out dipolar couplings and chemical shift anisotropy and to establish high spectral resolution. Since larger molecular weight proteins have longer rotational correlation times, and thus have a short transverse relaxation time (T_2) which leads to increased linewidths, the NMR signal from large proteins (40 kDa or larger) results in a poor spectral resolution [233]. Often, the use of deuteration is needed to remove the effect of proton couplings on the spectrum, and establish higher signal-to-noise and resolution. Smaller molecular weight proteins are more amenable to solution NMR due to these tumbling dynamics. Proteins reconstituted into liposomes are often too large and tumble very slowly in solution and are generally studied using solid-state NMR instead. Proteins in detergents or nanodisc systems are sometimes small enough to tumble fast, and therefore more amenable to solution NMR. Whole structure determination using solution NMR can be difficult, and often times requires alternate isotope labelling strategies. One example involves deuteration along with specific labeling of the methyl groups. The latter is achieved through the incorporation of biosynthetic methyl precursors during protein expression [234]. After overcoming the expression and membrane mimetic system challenges, acquiring the NMR spectrum involves some 2D and 3D experiments, of which the ones relevant to this thesis will be discussed briefly.

The primary experiment usually done with ^{15}N-labelled proteins is the ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) experiment. HSQC experiments provide correlations between the chemical shifts of ^{15}N and their attached amide protons $^{1}H_N$ [235]. HSQC can also be applicable towards other nuclei, like ^{1}H-^{13}C. Since every amino acid along the chain, apart from proline, contains a backbone amide with a N-H bond, this experiment will yield correlations for each amide, and thus each residue. In addition, the NH_2 groups from asparagine

and glutamine will also appear in this spectrum as doublets. Another experiment used in conjunction with HSQC, is transverse relaxation-optimized spectroscopy (TROSY), and it aims to increase the resolution and sensitivity of the spectrum from a protein in solution. Typical HSQC experiments which have not been decoupled lead to peaks appearing as multiplets due to J-coupling to protons (also known as through bond coupling). The TROSY experiment selects for the different relaxation mechanisms that cancel each other out, leading to a single sharp peak in the spectrum for each correlation [236]. This experiment thus leads to an increase in sensitivity and resolution of the overall spectra.

Chapter 3

The characterization of human Aquaporin-1 reconstituted into polymer nanodiscs

---

3.1 Introduction

The recent developments of various membrane mimetic systems have provided a new means towards studying membrane proteins in as close to a native state as possible [150,151,175,177]. Membrane proteins often do not fold correctly without the peripheral lipids the native membrane bilayer provides, which in turn impacts their function. Detergents have historically been used to solubilize MPs from membranes, as they did this with very high yields and resulted in a protein-detergent micelle complexes that were often quite stable [11,140]. However, the removal of lipids closely associated with the purified MP often results in a decrease or an outright loss in function. Membrane mimetics such as the SMA co-polymer-based nanodiscs are thus a very intriguing platform, as this system can solubilize MP's directly from the cellular membrane, purified easily with conventional chromatography methods, and result in a nanodisc that is theoretically small enough to be amenable for solution NMR experiments. These characteristics make the SMA co-polymer the prime choice for developing a platform that is needed to observe the mobile elements of a protein, and in this work described here, a platform that was used to create nanodiscs that encapsulate the human aquaporin-1 integral membrane protein.

To this effect, the goal in this body of work was to assemble these AQP1-SMA nanodiscs, assess their size and stability, and then record 1D and 2D solution NMR spectra on a ^{15}N-isotope labelled sample expressed from *P. pastoris* cells. A novel platform to visualize the flexible C-terminal tail of aquaporin-1 was established, observing what is likely the C-terminal tail region for the first time using solution NMR spectroscopy. These experiments then set the foundation for future protein-protein binding events to be measured directly using this system, as

the direct biophysical measurement of a protein-protein interaction event has not yet been reported for hAQP1.

Human aquaporin-1, as illustrated in detail in *Chapter 1*, has become increasingly medically significant, as recent studies provide further evidence that it may be involved in several physiologically important processes within the cell [94,130,133]. As well, it's becoming increasingly evident that hAQP1 is also involved in several cancer processes, most notably in several types of breast, lung, and colon cancers [116,123,124,126,127]. These processes are often governed by protein-protein interaction events that modulate AQP1's trafficking to the cellular membrane, where it can fulfill its primary function as a water transporter. These interaction events within the cell seem to consistently involve the ~ 30 residue long C-terminal tail, a region of AQP1 that is intrinsically disordered, and consequently never structurally characterized. Previous solid-state NMR studies have unable to observe this tail, as the disordered nature results in many different conformations in solution [111]. Of course, the actual structure may differ within the crowded cellular environment as well, without even mentioning the structural changes that may occur once this region is interacting with another protein partner.

The structural changes induced upon a protein binding event can be observed using solution NMR, by probing the chemical shift changes in the spectra [233]. Solution NMR is thus a powerful technique that can monitor these dynamic changes in chemical shifts in solution, and has been used extensively in the drug discovery industry to monitor protein-drug binding [238]. These binding events can be probed using chemical shift perturbations (CSPs) during a titration experiment, where a series of NMR spectra

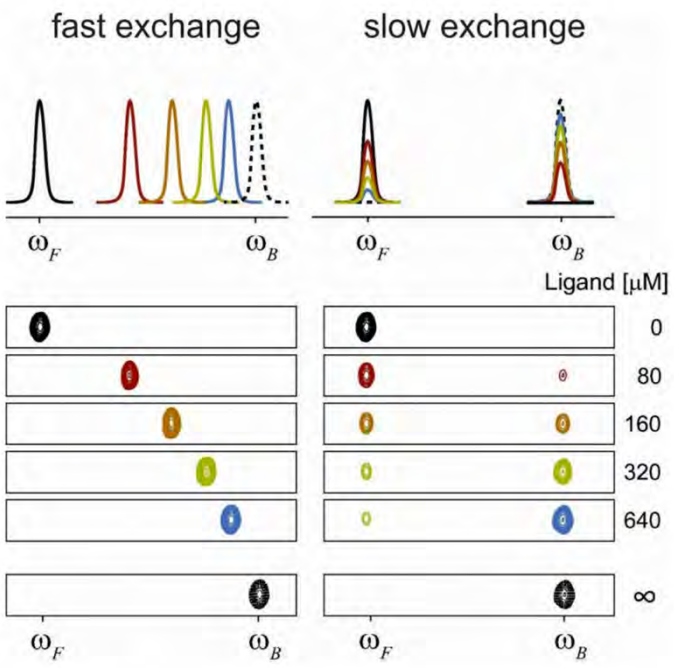

Figure 3-1: *The chemical shift perturbation behavior in both the fast and slow exchange regimes in a protein-ligand NMR titration experiment. Figure reproduced from Becker et al., 2018, with permission from Wiley.*

can be recorded both in the absence and presence of varying, most times increasing, amounts of the binding partner [239]. These perturbations are mainly readily observed in two kinetic regimes, in the fast or slow exchange timescales [**Figure 3-1**]. In the fast exchange regime, where the binding event is fast and relatively weak, the time of transition between the free and the bound state of the protein-ligand complex is too quick for the observation of the individual signals. In this case the signal collapses into a single peak that progressively shifts toward the fully-ligand saturated peak during the titration. For the slow exchange regime, this occurs when the protein-ligand binding event is strong, and so the transition time between the free and bound states is long enough to resolve both signals [**Figure 3-1**]. In this case both resonances

corresponding to the bound and free states appear in the spectrum, but at varying intensities. Upon increasing the concentration of the ligand, the free state signal decays while the bound state signal increases. Thus, monitoring the CSPs during a titration of a specific binding partner for hAQP1 may show changes in the structure of this C-terminal region if the protein interaction partner does indeed bind. The HSQC spectrum cross-peaks observed from the free AQP1's C-terminal tail region are expected to correspond to a disordered state. However, once the protein partner is titrated in, the regions and residues which are directly involved in the protein-protein interaction may become semi- or fully structured, leading to either a disappearance or a shift in the impacted chemical shifts.

These chemical shift titration experiments are thus the end goal when assembling this nanodisc platform for solution NMR. However, first the underlying ^{1}H-^{15}N HSQC spectra need to be recorded as most protein-ligand titration experiments are commonly recorded in an HSQC spectrum given the good spectral resolution and signal dispersion.

The overall goal of the following work in this chapter was to assemble nanodiscs containing AQP1 using the SMA co-polymer system for solution NMR. These resulting AQP1-SMALPs were then characterized biochemically and biophysically, to ensure the nanodiscs were properly assembled and stable enough at room temperature for solution NMR studies. AQP1 was expressed in the *P. pastoris* using ^{15}N-enriched growth media, to then observe the AQP1-SMALPs in both an 1D proton and 2D ^{1}H-^{15}N HSQC solution NMR spectra. Finally, a series of DLS measurements were performed on the AQP1-SMALPs to assess stability and their aggregation tendencies. The results of these measurements then guided further probing into various buffer conditions for the AQP1-SMALPs to inhibit or stop the aggregation of the nanodiscs.

3.2 Materials and Methods

3.2.1 Materials

Common chemicals were acquired from either Fisher Scientific (Mississauga, Ontario, Canada) or Sigma-Aldrich (Oakville, Ontario, Canada). The isotopically labelled compound $(^{15}NH_4)_2SO_4$ was obtained from Cambridge Isotope Laboratories (Andover, MA) or Sigma Aldrich.

3.2.2 Transformation and plasmid preparation

The plasmid construct pPICZB encoding hAQP1 with an 6x-His, a c-myc tag and a TEV protease cleavage tag was transformed into *E. coli* DH5α cells. 100 ng of plasmid was added to the cells and then heat-shocked at 42˚C for 45 s. Transformed *E. coli* cells were then grown on 2XYT (1.6% tryptone, 1% yeast extract, 0.5% NaCl) plates with 2 mg/mL Zeocin to select for transformants resistant to antibiotic. Isolated colonies were scraped and resuspended in 2XYT liquid culture media, with 25 μg/mL zeocin, grown overnight at 37˚C and shaken at 240 rpm. The plasmids were then isolated using the Qiaprep Spin Miniprep Kit (Qiagen, Mississauga, Ontario, Canada). The concentration of DNA was determined using a UV-VIS spectrophotometer (Cary 50, Varian), using a wavelength of 260 nm. DNA purity was assessed using the ratio of absorbances of 260/280 nm, with an optimal ratio of 1.8 for pure DNA. The plasmid was then linearized using the restriction enzyme PmeI. 8.5 ug of DNA was linearized using a PCR thermocycler (Mastercycler, Eppendorf). The reaction ran at 36˚C for 4 hours, and then stopped at 65˚C for 20 mins. Following linearization, the plasmid DNA was cleaned using the QIAquick nucleotide removal kit (Qiagen, Mississauga, Ontario, Canada). Then, to

concentrate the DNA the ethanol precipitation method was used. To the eluted DNA, a 1/10th volume of 3 M sodium acetate, pH 6.3 was added, and then 2.5 volumes of 100% cold ethanol followed. The DNA was incubated overnight at -20˚C and then spun down at 17,700xg in a Eppendorf tabletop centrifuge [Eppendorf, Hamburg, Germany]. The resulting pellet was washed in 80% ethanol and spun down again at 17,700xg. The DNA pellet was air-dried completely before resuspending in sterile Milli-Q water. A 0.8% agarose gel was cast for agarose gel electrophoresis at 100 V. The DNA was assayed to check for fully linearized DNA before transformation into competent protease-deficient *P. pastoris* strain SMD1168H. Finally, the DNA and competent *P. pastoris* cells were loaded onto cold sterile 0.2 cm electroporation cuvettes (Bio-Rad, Mississauga, Ontario, Canada). Cells were pulsed twice at 2.00 kV, 5.6 ms. Immediately after pulsing cells were resuspended in 1 M sorbitol and incubated at 30˚C without shaking for 2 hrs. Following this, cells were incubated at 30˚C for 1 hr at 100 rpm, and then finally plated on YPDS plates with varying zeocin concentrations which were as follows: 0, 100, 200, and 500 μg/mL. These cells were incubated for 5 days at 30˚C.

3.2.3 LT-SEVIN *Pichia pastoris* expression

Isolated transformed colonies were then re-plated on new YPD plates for 5 days at 30˚C, before inoculating 12 ml of buffered minimal dextrose media (BMD) (0.8% (w/v) ammonium sulfate, 2.5% glucose, 0.34% (w/v) YNB without amino acids and ammonium sulfate, 4 x 10^{-5} % (w/v) biotin, 100 mM potassium phosphate buffer, pH 6.0), which was incubated overnight at 30˚C and 300 rpm. The following day, the culture was diluted to 50 ml with additional BMD the next morning, and left to incubate again at the same conditions for 24 hrs. Then, a final dilution to 250 mL with additional BMD was achieved the next day, and again the cultures were

incubated at 30˚C, 300 rpm, for another 24 hrs. Following this, the cells were spun down at 1500xg, for 15 minutes to collect the cells to resuspend them in 250 mL buffered minimal methanol media (BMM) (0.8% (w/v) ammonium sulfate, 2.5% glucose, 0.34% (w/v) YNB without amino acids and ammonium sulfate). The cells were then shaken at 300 rpm, 30˚C for 1 hour for any leftover glucose to be metabolized before adding the methanol and sorbitol. Methanol and sorbitol were added to the cultures to a final concentration of 0.5% MeOH and 1.1 mM, respectively. After induction, the cells incubated for another 24 hours at 20.5˚C and 240 rpm. The cells were then collected, spinning down at 1500xg for 15 mins, with several washes of the pellet with 150 mM NaCl. The pellet was stored at -80˚C before cell breakage.

3.2.4 Styrene maleic acid solubilization

Several solubilizing conditions were performed before establishing the final solubilization and buffer conditions for solution NMR experiments. The main variables tuned over the course of the purification trials were the time and temperature of solubilization, the solubilizing and purification buffer conditions, the duration and centrifugation speed during concentrating of the sample, and the final buffer conditions for NMR spectra recording. At all stages protein yield was assessed through SDS-PAGE and densitometry analysis using the ImageJ software or by UV-Vis spectroscopy by measuring absorbance at 280 nm but accounting for the extra absorbance by the copolymer.

3.2.5 General solubilization protocol

The general expression and then solubilization of AQP1 with the SMA co-polymer is outlined hereafter. Firstly, small-scale cultures were grown at 100 mL volumes to produce

enough AQP1 to be solubilized and visualized on an SDS-PAGE gel, but also to remain a small enough culture to save on polymer use and to experiment with multiple solubilizing conditions at once.

To break the cells, 500 μm acid-washed glass beads (Thermofisher, Mississauga, Canada) were used. To the pellet, glass beads were added along with breaking buffer (50 mM potassium phosphate, pH 7.5, 5% glycerol, 2 mM EDTA, 1 mM PMSF) and the cells were vortexed for 16 mins consisting of 1 min "on" cycles and 1 min "off" cycles, keeping the cells on ice while in the "off" cycle. The cells were then spun down at 700xg for 5 mins to pellet the unbroken cells, while the supernatant carrying broken membrane fragments were collected. Cells were then vortexed once again but this time in 8 min cycles, before spinning down the unbroken cells. This process was generally repeated 10-12 times, until the pellets lost all hue and became a white colour.

After cell breakage, the supernatants were collected, and ultracentrifugation was used to spin down the broken membranes. A Sorvall F37L rotor (Thermofisher Scientific, Mississauga, ON, Canada) was used to spin the supernatant down at 100,000xg, for 30 mins. The membrane pellets were once again collected and re-suspended in solubilization buffer (50 mM HEPES, pH 7.5, 10-100 mM NaCl, 10% glycerol, 1 Roche cOmplete EDTA-free protease inhibitor cocktail tablet (Millipore Sigma, Oakville, Canada)) containing 2.5% SMALP 200 (2:1 styrene to maleic acid) copolymer (CubeBiotech, Wayne, PA, USA). Solubilization occurred at varying temperatures (4, 24, 37, 41˚C), varying time lengths (2, 4, and 16 hours), and different pH (7.0, 7.5, 8.0, 9.0). The varying conditions were trialed to increase the yield and to assemble the most stable nanodiscs. To this end, the final conditions for solubilization were 37˚C, 16 hours, and a pH of 7.5.

After solubilization, the solubilized cell membranes were loaded into a F50L rotor and once again ultracentrifuged at 120,000xg for 30 mins. This time the supernatant was collected, containing the self-assembled SMALPs, and was diluted 10-fold using Ni^{2+}-binding buffer (50 mM HEPES, pH 7.5, 10 mM NaCl, 10% glycerol, 1 protease inhibitor cocktail tablet) reducing the final SMA concentration to 0.25%, thus preventing any non-specific interactions of the SMA copolymer with the Ni^{2+}-NTA beads. These diluted SMALPs were then added to 1 mL of Ni^{2+}-NTA agarose beads (Qiagen, Toronto, Canada), and incubated at 4˚C with stirring for 6-8 hrs.

After nickel binding, the SMALP-lysate was added to an Econo-Pac chromatography gravity flow column (Bio-Rad, Mississauga, Ontario, Canada). The resulting beads were washed with 2.5x the bed volume of wash buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 10% glycerol, 30 mM imidazole), until the absorbance at 280 nm for the eluent was below 0.05, indicating that the remaining protein is pure enough for solution NMR studies. Purified SMALPs were eluted with 300 mM imidazole, and subsequently concentrated using the 10 kDa MWCO Amicon concentrators (Millipore Sigma, Oakville, ON, Canada) spinning at 3000-3700xg for 6-15 min intervals. The concentrated sample was then buffer exchanged into NMR-suitable buffer, which consisted of 20 mM phosphate buffer, pH 7.0, 10 mM NaCl. Finally, the tobacco mosaic etch viral (TEV) protease was used to cleave the His and c-Myc-tag off the C-terminal tail. The AQP1-SMALPs were incubated overnight at 4˚C at a ratio of 1:25 AQP1 to enzyme. Following this, a brief buffer wash, using centrifugation with a concentrator, with the same buffer was used to wash off the enzyme and the cleaved peptide.

3.2.6 SDS-PAGE

AQP1-SMALPs prepared from the varying solubilization trials were analyzed on SDS-PAGE gels, and the band density was compared to a known concentration control to determine approximate yields. SDS-PAGE gels were prepared using the preparation of three buffer solutions: the acrylamide solution (29.2% (w/v) acrylamide, 0.8% (w/v) bisacrylamide), the 1.5 M Tris-buffer (1.5 M Tris-base, pH 8.8, 3.75% (w/v) SDS) and the 0.5 M Tris buffer (0.5 M Tris-base, pH 6.8, 3.75% (w/v) SDS). These buffer components made up the final acrylamide gel solution **[Table 3.1].** First the resolving gel component was prepared, with the tetramethylethylenediamine (TEMED) and ammonium persulfate (APS) added last, and this was poured between two glass plates with a 5 mm thickness and incubated for 1 hour for the gel to polymerize. Next, the stacking solution was poured on top, with a 1 mm comb inserted into the top to form the wells and allowed to polymerize for 30 mins.

Table 3.1: *Buffer and reagent components for preparation of SDS-PAGE gels.*

Component	***12.5% resolving gel***	***5% stacking gel***
Acrylamide solution	5 mL	0.67 mL
1.5 M Tris-buffer	3 mL	-
0.5 M Tris-buffer	-	1 mL
Water	3.95 mL	2.32 mL
10% (w/v) APS	40 uL	12 uL
TEMED	12 uL	6 uL

For running the gel electrophoresis, the entire gel cassette was placed in a Bio-Rad electrophoresis chamber, and SDS running buffer (25 mM Tris, pH 8.3, 192 mM glycine, 0.1%

SDS) was added. For purified samples, protein was loaded with ¼ volume of 5X SDS-PAGE loading dye (10% (w/v) SDS, 500 mM DTT, 50% (w/v) glycerol, 250 mM Tris-HCl, pH 6.8, 0.5% (w/v) bromophenol blue) and loaded into the well. For membrane samples, 10% (w/v) SDS was added to the sample and incubated at 65˚C for 10 mins. The membranes were then vortexed briefly and then spun down at 13,300xg for 5 mins, and the resulting supernatant was loaded onto the gel with ¼ volume of 5X loading dye. After loading samples, the gels were run for 1-2 hours at 120V. For visualization of the bands, the gel was briefly washed in mqH_2O to wash off SDS, and then stained for 1 hour or overnight in Coomassie stain (0.1% (w/v) Coomassie Brilliant Blue, 20% (v/v) methanol, 10% (v/v) glacial acetic acid). To de-stain the gel, the gel was added to de-stain solution (50% (v/v) methanol, 10% (v/v) glacial acetic acid) and microwaved for 30-60s to heat the solution. The gel was incubated for 1-12 hours to de-stain. Bands were analyzed using ImageJ software using densitometry analysis of band intensity peaks as compared to a control with a known concentration run on the same gel.

3.2.7 Western blotting

For Western blot analysis, the gels were transferred onto PVDF membranes using the iBlot (Invitrogen) following the manufacturer's instructions. The PVDF membrane with transferred protein was incubated with the primary antibody Pierce 6x-His epitope tag IgG2b mouse antibody (Thermofisher), and the Pierce SuperSignal West Pico Fast Western blot kit (Thermofisher) was used according to the manufacturer's instructions. The blot was imaged with a UVP ChemiDoc-It TS2, along with the GelAnalyzer 19.1 software. Relative band intensities were also analyzed using the ImageJ software.

3.2.8 Stability gel assays

The thermostability of AQP1-SMALPs was monitored using a gel-based assay. The AQP1-SMALPs were heated for 10 mins at varying temperatures (4 – 70˚C), followed by centrifugation at 10,000xg for 10 mins to remove any aggregated nanodiscs. The supernatant was then run on an SDS-PAGE gel and stained with Coomassie blue afterwards to visualize the protein bands. Band intensities were analyzed using ImageJ normalized to that of a control band.

The time stability gel assay was run similarly, however the treatment just consisted of incubating the AQP1-SMALPs at 4˚C over a course of 96 hours, and at room temperature for a total of 72 hours. These samples were then also subsequently run on the gel and analyzed as described previously.

3.2.9 Raman spectroscopy

AQP1-SMALPs were prepared as described previously and packed into a metal sample compartment containing 5 uL. FT-Raman spectra were measured in a Bruker FRA106/s accessory to the IFS66vs spectrometer, with excitation at 1024 nm, at 2 cm^{-1} resolution, in back-scattering geometry.

3.2.10 Transmission electron microscopy

The AQP1-SMALP discs were prepared as discussed prior, before using negative staining transmission electron microscopy to directly visualize the discs and their respective sizes. AQP1-SMALPs were dialyzed into TEM buffer (20 mM Tris-HCl, pH 7.0, 10 mM NaCl) and concentrated to 0.05 -0.15 mM. Two different preparations were used before imaging, one set of discs solubilized at 4˚C while the other solubilized at 37˚C. An FEI Tecnai F20 G2

transmission electron microscope (FEI Company, Hillsboro, Oregon, USA) was used to acquire the images (at the Advanced Analysis Centre, University of Guelph). To achieve negative staining, 0.5-3% uranyl acetate was used, and the AQP1-SMALPs were administered onto copper grids using the "on-drop method", whereby the sample was first adhered to the grid for 2 minutes, followed by wicking off the excess and then a wash with filtered water. Next the grid was placed on a drop of stain, and immediately transferred to the second drop and incubated for 30 s. Excess stain was wicked off and dried for 2 mins before proceeding to imaging. Grids were then imaged immediately and at room temperature.

3.2.11 Solution Nuclear Magnetic Resonance spectroscopy

^{15}N-AQP1-SMALPs were concentrated to 0.11 – 0.21 mM of hAQP1, in a volume of 300 uL and aliquoted into a 5 mm Shigemi tube (Bruker, Milton, Ontario, Canada) for solution NMR experiments. Spectra were recorded using a 600 MHz Bruker Avance III NMR spectrometer with cryoprobe (at the Advanced Analysis Centre, University of Guelph). A 1D proton, 2D ^{1}H-^{15}N (HSQC), and a 2D TROSY-HSQC spectra were obtained at 25˚C, with recording time ranging from 1 to 8 hours. The final HSQC spectra shown was the result of an 8-hour recording.

3.2.12 Dynamic light scattering

100 uL of AQP1-SMALPs of the various temperature or buffer treatments were added to a 1.0 cm path length disposable cuvette. Measurements were taken on a Malvern Zetasizer Nano ZS (Malvern, UK) at 25˚C, with a backscatter angle of 173˚, an equilibration time of 120 s, and an average of 15 readings per sample. Each measurement was repeated in triplicate, and the triplicates were averaged before plotting the intensity percentage and volume percentage size

profiles using R, and the plotting package ggplot2. For the buffer condition experiments, the AQP1-SMALPs were purified as described above, and split into 5 equal samples (0.1 mM) for further exchange into subsequent buffers. The buffers were as follows: control (NMR buffer as described previously), EDTA (NMR buffer + 2 mM EDTA), glycerol (NMR buffer + 1% glycerol), high-salt (20 mM potassium phosphate, 200 mM NaCl), low salt (25 mM HEPES, 5 mM NaCl). All buffers were set to a pH of 7.0.

3.3 Results and Discussion

Aquaporin-1, through the validation of differing solubilizing and buffer content conditions, was successfully reconstituted into the SMA nanodiscs for biophysical characterization.

3.3.1 AQP1 solubilization efficacy

Several solubilizing conditions were experimented with to determine those that provided the highest protein yield, while also retaining the structural stability of the nanodisc complex. AQP1 was first solubilized as recommended by the manufacturer CubeBiotech, for a duration of 2 hours and at 4˚C, a concentration of 2.5% SMA, and a high concentration of salt was used throughout both the solubilization and purification process (400 and 300 mM NaCl respectively), as established previously with detergent solubilizations of AQP1. AQP1 was successfully solubilized and purified to a high degree of purity after just the affinity chromatography step, as determined by both SDS-PAGE and UV-Vis spectroscopy **[Figure 3-2].** The same AQP1-containing membranes were solubilized by n-Octyl-β-D-glucopyranoside (OG) to function as a control, where this OG-AQP1 band appeared as one uniform band at around ~28 kDa, the

expected size. AQP1-SMALPs also appeared as one uniform band in the location where we would expect, not dissimilar from the OG-AQP1. The yield of this first solubilization was 0.22 mg/0.1 L of culture, which when extrapolated to 1 L (which would be the basis for solution NMR studies), was estimated to be 2.2 mg/L. It was later found that when scaling up from 100 mL to 1 L, the yield estimates remained true and scaled linearly. Although SMA was able to solubilize AQP1 from *P. pastoris* membranes, it was clear that OG was much more efficient at solubilizing the membranes, comparing the insoluble and soluble lanes of the SMA solubilization **[Figure 3-2].** From previous trials, OG was able to solubilize nearly 80% of all AQP1 while SMA was only able to solubilize around 20% of the protein present in the insoluble fraction.

To boost yields and overall solubilization efficacy, a few variables were modified during the solubilization process. Firstly, the concentration of salt in the buffer was drastically decreased, from 400 mM to 10 mM NaCl. Although high salt was used for detergent extractions of AQP1 previously, for SMA a low salt buffer was believed to increase the solubilizing efficacy. SMA has a high density of negative charges surrounding the polymer to the maleic acid groups and is known to bind Na^+ ions. It is also well known that a high salt concentration can induce protein aggregation, as is the case with "salting-out" proteins. An increase in ionic strength will decrease the net-charge repulsion of solutes within the solvent, and hydrophobic protein-protein interactions will be favoured, but this is also protein-dependant [240]. This low salt solubilization subsequently increased the yields to 3 mg/L while keeping the time and temperature the same.

Next, the total time of incubation time during solubilization was varied as well, with 2, 4, 5, and 16 hrs time points. Varying the time course of solubilization did not appreciably increase

protein yields at all, with yields ranging from 2.0 – 2.2 mg/L of culture. Even so, a 16-hr incubation time was chosen to be the definitive time point as it was later observed it proved beneficial in conjunction with higher temperatures during solubilization. Subsequently, a few temperature points were trialed as well, with 4˚C, 24˚C, and 37˚C chosen while keeping the incubation time at 16 hrs. There was no meaningful difference observed between the 4˚C and 24˚C solubilization with yields ranging again from 2.0 – 2.4 mg/L, however the largest increase was observed in the 37˚C solubilization, with the total yield of 8.5 mg/L. Although the yield was very high, a significant amount of degradation was observed on the gel **[Figure 3-3].** The 4˚C and 24˚C preparations showed a uniform band, but the 37˚C preparation showed two distinct bands, with the major AQP1 band migrating slightly lower than the others and the control as well.

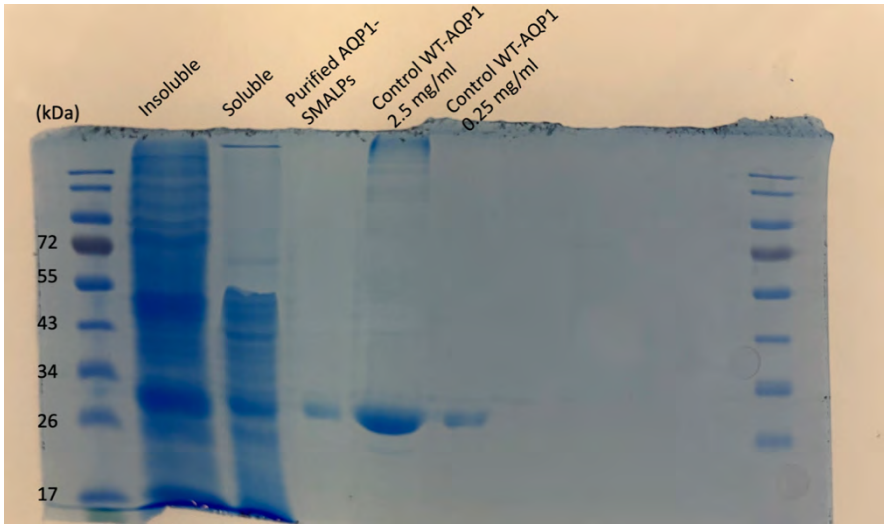

Figure 3-2: *Solubilization of AQP1 with the SMA copolymer, with the insoluble fragments and soluble (before-purification) also ran on the gel. AQP1-SMALPs ran as one band where the size of AQP1 was to be expected (~28 kDa).*

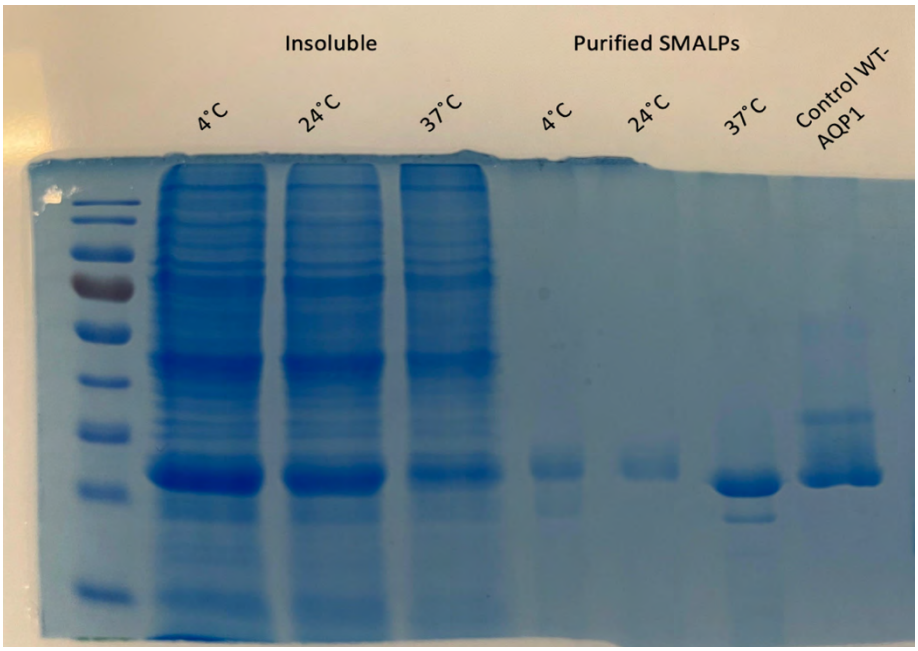

Figure 3-3: *SMA solubilization of P. pastoris cells at differing temperatures, while incubation time was kept consistent at 4 hours just to assay the temperature differences.*

This signalled a high degree of protein degradation, likely through the existence of proteases at some point in the solubilization and purification process. This encouraged the use of more protease inhibitors at various stages and the addition of EDTA during solubilization to inhibit any metalloproteases also present in the membranes. This addition would later seem help alleviate the protease problem.

After establishing that a longer time of incubation, as well as a higher temperature yields the greatest protein yield, the last variable to be experimented with was the pH. For these trials, a solubilization temperature of 37˚C and an incubation time of 4 hrs was used, to observe the

different solubilization efficacies of the following pHs: 7.0, 8.0, and 9.0. The highest solubilization efficiency was observed at pH 7.0, and these results were to be expected [176,179], with the yields decreasing by 50% at pH 8.0, and by 65% at pH 9.0 **[Figure 3-4]**.

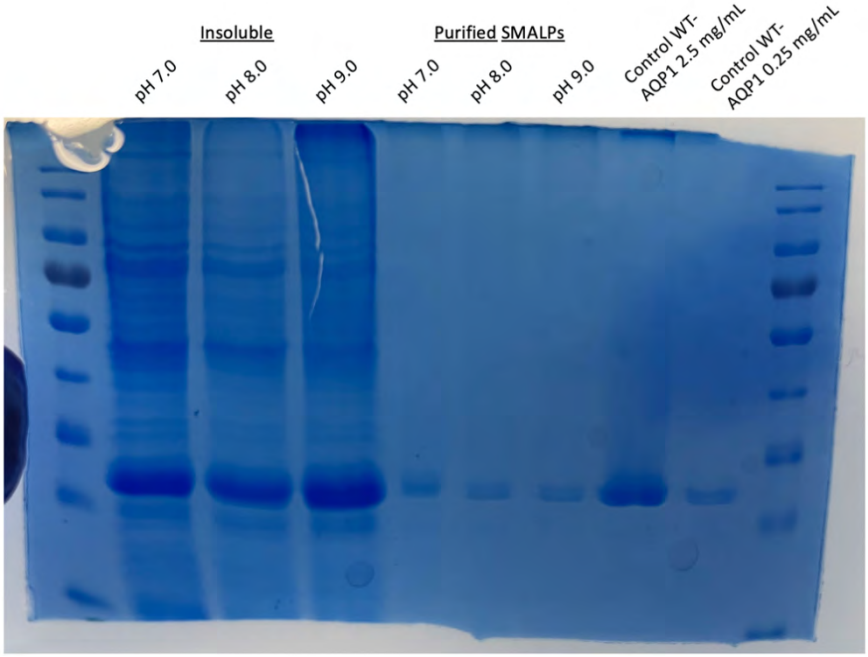

Figure 3-4: *pH solubilization trial of AQP1, comparing the solubilization efficiency of SMA at the different pHs with the insoluble fragments shown also to compare total solubilization from the cell membranes. The total yield was lower than in the previous 16 hr solubilization trial due to the fact that these were solubilized in a 4 hr incubation time.*

3.3.2 Gel-based stability assays

After the solubilizing conditions were optimized to increase overall yield, the next step was to assay the stability of these nanodiscs in solution. Since the primary goal for these SMALPs was to study them using solution NMR, these discs would have to be stable for a

certain amount of time to complete the NMR measurements. A benchmark for about 5-7 days of observable stability was established at first, enough to complete a few 2D NMR experiments, and enough time to also perform any future titration experiments with a protein partner of interest. A second benchmark of 2 weeks of stability was set so that the discs were stable enough to perform all 2D, 3D, and possibly some 4D NMR experiments to assign all residues observed within the initial 2D spectrum. These assignments would provide crucial information towards exactly which residues are interacting with the protein binding partner during the titration, giving an atomic level resolution towards the protein-interaction event. Initially SDS-PAGE gel-based assays were performed to determine the discs stability over time, as well as its thermostability. The thermostability of the discs is also valuable information considering solution NMR experiments are normally conducted at room (or higher) temperature. The temperature during experiment time can be increased to provide faster tumbling and higher signal-to-noise, given the nanodisc system is stable enough at those higher temperatures.

The thermostability gel assayed the stability of the AQP1-SMALPs at a range of temperatures from 4˚C to 70˚C, with 4˚C being the standard storage temperature for the discs **[Figure 3-5]**. Gel densitometry analysis revealed that the band intensities did not differ or decrease significantly even up to 70˚C, past the transition temperature of AQP1 reconstituted in liposomes, which was 61˚C as previously confirmed. Denatured protein is more prone to aggregation, and so unstable protein likely would have aggregated and ran just at the top of the gel, leaving a less intense band at ~28 kDa where the AQP1 monomer rests. This was not observed to any significant degree even up to 70˚C, confirming that the temperature range of 24˚C to 40˚C needed for solution NMR would not initially be a problem unless for longer

timescales. This was later further analyzed using a more sensitive technique, dynamic light scattering.

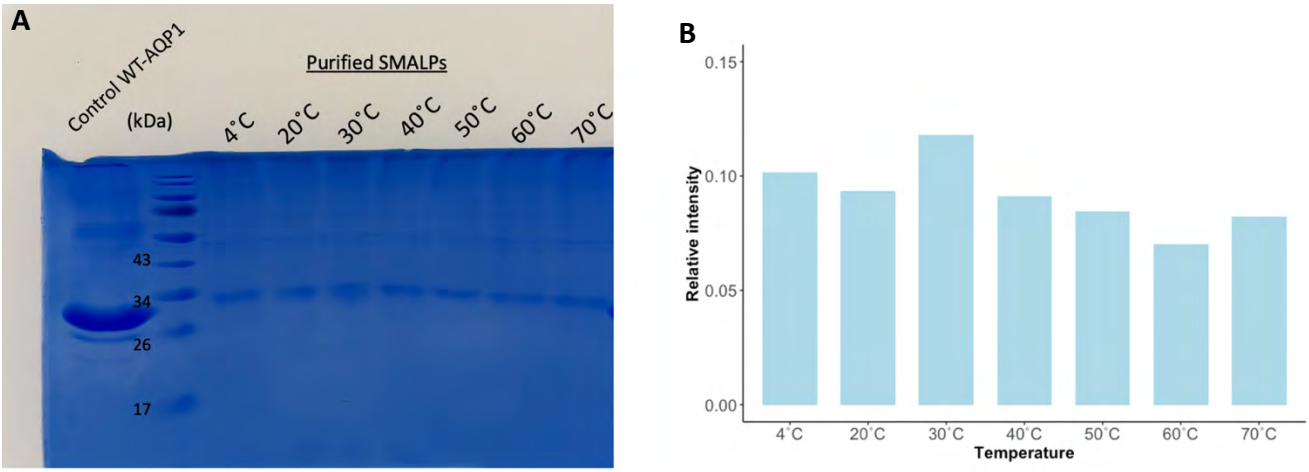

Figure 3-5: (A) *Thermostability of purified AQP1-SMALPs by SDS-PAGE gel assay.* ***(B)*** *Plot of relative intensities of AQP1-SMALP bands as compared to a standard. No significant decrease in intensity was observed as temperature of incubation increased.*

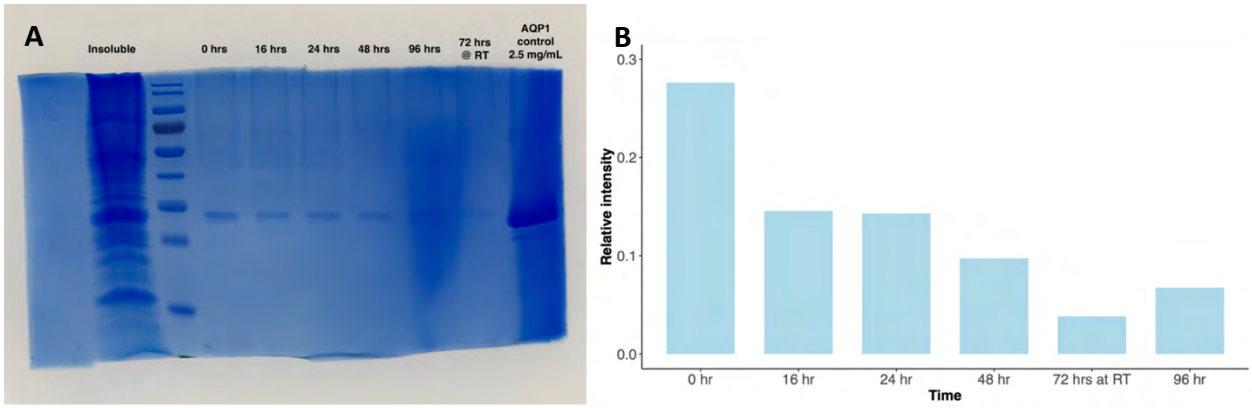

Figure 3-6: (A) *The time stability of AQP1-SMALPs incubated at 4˚C over the course of 96 hours, and run on an SDS-PAGE gel at the 0, 16, 24, 48, and 96 hour time points. As well, one*

sample was incubated at room temperature for 72 hours to observe stability at ~24˚C. ***(B)*** *Bar plot of relative band intensities.*

Next, the time stability of AQP1-SMALPs was monitored over the course of 96 hours. After initially observing degradation in the SMALPs after the first 37˚C preparation **[Figure 3-3]** and in previous gels with older samples [data not shown], the exact timescale of their stability needed to be determined. Over 96 hours, incubated at 4˚C, the relative intensity of the AQP1 bands dropped by nearly 70%, with a 50% decrease in band intensity observed after just 16 hours **[Figure 3-6]**. After the first 24 hours, band intensity decrease was less dramatic, but continued linearly. Since there was no sign of smaller molecular weight fragments as time went on, the intensity loss was attributed to aggregation. The possibility of protease degradation could not be completely ruled out however, as smaller weight peptide fragments could have been too low of a concentration to be fully observed given the low sensitivity of the SDS-PAGE gel. As well, the AQP1-SMALPs were also incubated at room temperature (RT) (~24˚C) for 72 hours, and this band displayed the highest decrease in intensity, losing 85% of the original signal. These results indicated that whether it was protease degradation or aggregation, the mechanism was accelerated at higher temperatures of incubation. Given that starting solution NMR experiments begin at 25˚C, this was a property that would need to be optimized fully.

The original hAQP1 plasmid construct contained a 6xHis-tag and additional Myc-tag with a spacer between these, contributing to an additional 21 residues on top of the already long C-terminal tail. In addition, this poly-histidine tag is charged at a pH higher than 7.0 and so could be resulting in unwanted, non-specific electrostatic interactions with neighbouring nanodiscs

given the maleic-acid moieties are also negatively charged. Thus, the idea was to then assess the stability of the nanodiscs after using the TEV enzyme to cleave the tags off the C-terminal tail.

Finally, the AQP1-SMALPs stability over time was assessed post TEV cleavage, up to a time frame of 120 hours [**Figure 3-7**]. It was found that the nanodiscs, incubated at 4˚C during this time span, were indeed stable. No significant degradation or loss of signal attributed to aggregation was observed on the gel.

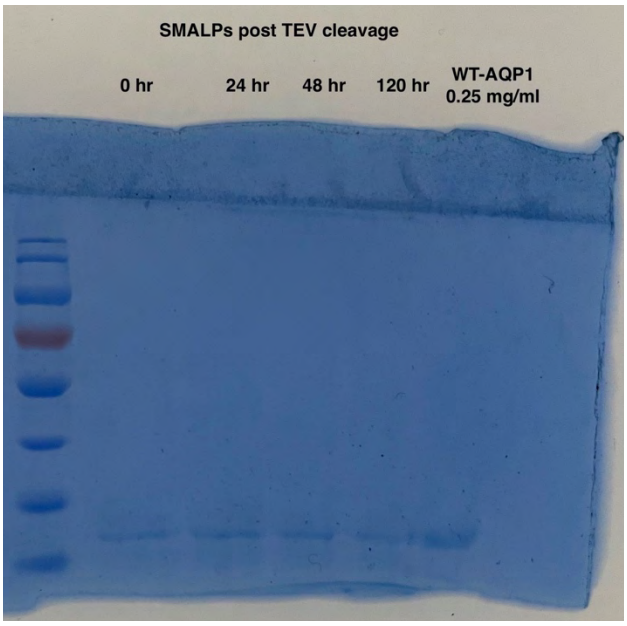

Figure 3-7: *Time stability of AQP1-SMALPs after TEV reaction overnight. Relative band intensities did not decrease significantly, indicating the stability in solution.*

3.3.3 Size and diameter of SMALPs

Many reported values in the literature on the size and diameter of the SMALPs have determined them to be around 10-15 nm in size, with some groups reporting larger sizes as well [243,244]. Sizes of the discs seem to be determined by the copolymer itself; the functional

groups it contains but also the length. Sizes also vary greatly between the SMA and DIBMA copolymers, with DIBMA-based nanodiscs being much larger with an average diameter of ~20 nm [243]. To determine the size of the AQP1-SMALPs, we used negative staining transmission electron microscopy (TEM) to directly image these discs and ascertain the size.

Two different preparations were used before imaging, with the main difference being the solubilization temperatures at 4˚C and 37˚C. Interestingly, large differences were observed in the overall diameters and in morphology from both preparations. 4˚C SMALP preparation resulted in much smaller disc diameters, with an average mean size of 16 nm **[Figure 3-8]** and a bimodal-like distribution with the largest disc counts in the 14 nm and 18 nm sizes. Interestingly, the SMALPs solubilized at 37˚C resulted in much higher disc diameters **[Figure 3-9],** with the mean size just over 20 nm. The largest counts were still in the 14 nm size, however the size distribution also stretched into much higher sizes, with appreciable counts occurring in the 30-40 nm range as well. These much higher sizes may be explained by the fact that nanodisc dimers were also seen in this preparation, as observed in some of the TEM images **[Figure 3-9]**. These dimers and other aggregates that might have formed on the grids could explain the higher sizes, as it is unlikely the solubilized discs were much larger than 20 nm. Although these higher sizes possibly resulting from solubilization at the higher temperature cannot be completely ruled out either.

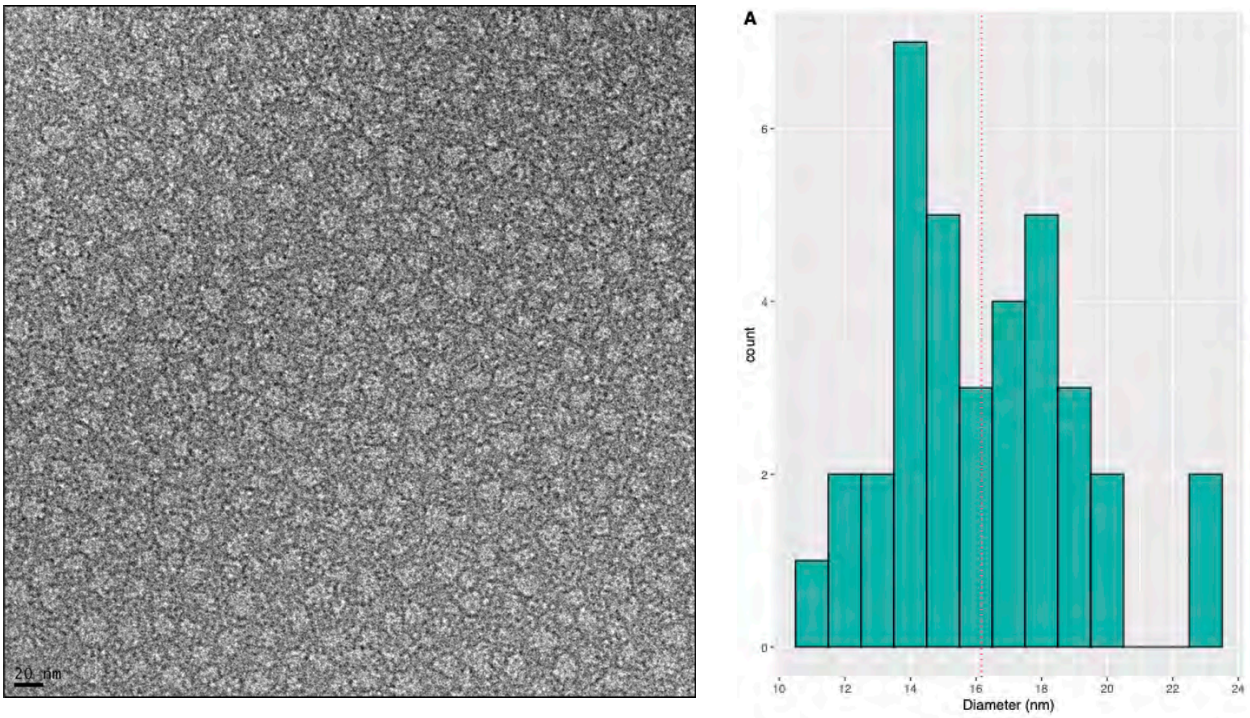

Figure 3-8: (A) *4˚C preparation of AQP1-SMALPs TEM image, acquired at 25˚C.* ***(B)*** *Discs observed were smaller in diameter, with a mean average size of 16 nm, indicated by the dashed line.*

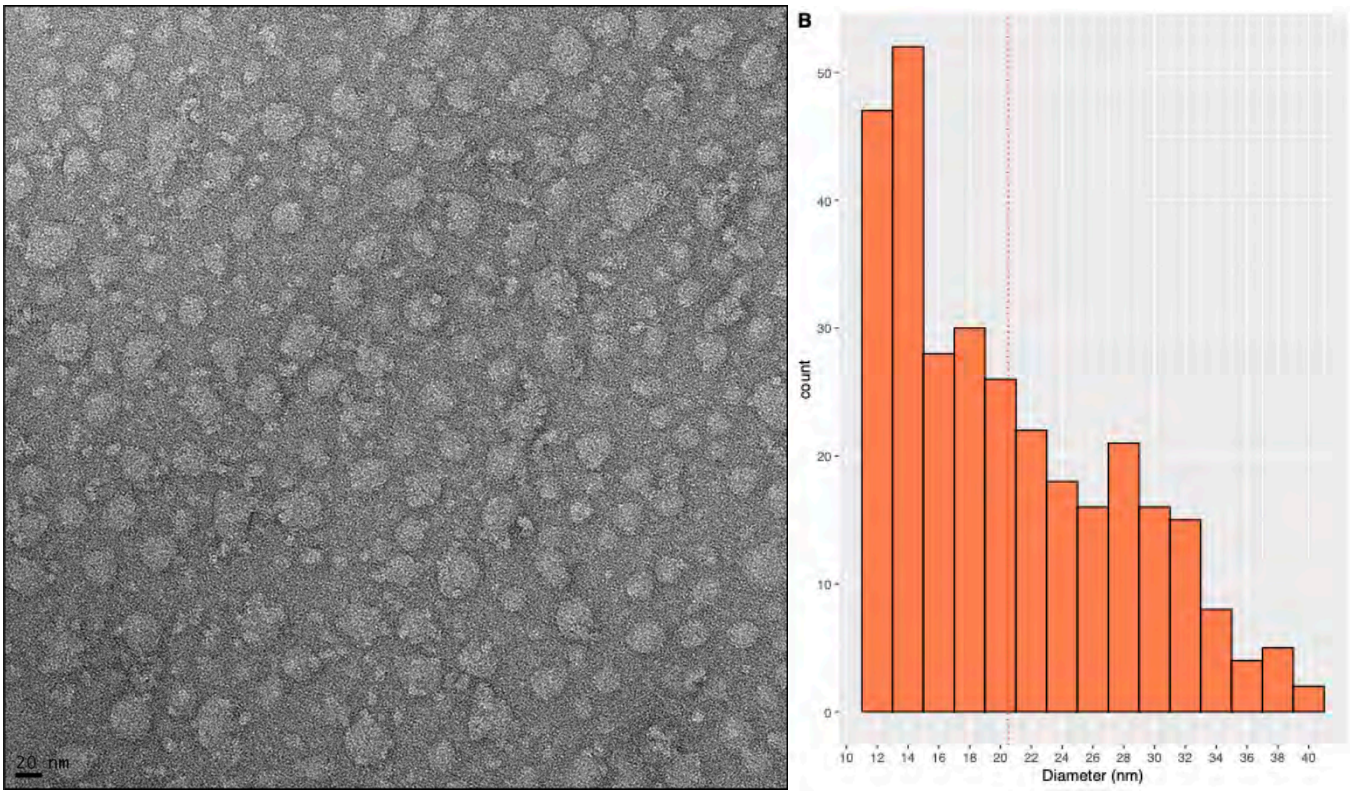

Figure 3-9: (A) *37˚C solubilization temperature preparation of AQP1-SMALPs TEM image, acquired at 25˚C.* ***(B)*** *Discs observed had a larger size distribution, ranging from 10 – 40 nm size discs, with a mean average size of 20.5 nm, as indicated by the dashed line on the plot.*

3.3.4 SMALP components

Raman spectroscopy is a spectroscopic technique that relies on the phenomenon of Raman scattering, which describes the inelastic scattering of photons. It is a very useful technique towards the study of proteins suspended in water, as water does not interfere near the amide I, II, and III regions in a Raman spectrum, unlike FTIR, where water will crowd the signal in the amide I region. As well, since the AQP1-SMALPs are not easily pelleted unless ultracentrifuged at high speeds, Raman was the chosen method to detect the different

components of the discs, as they could be studied in solution. The AQP1-SMALPs contain three chemically different components, all with differing absorption frequencies in both FTIR and Raman spectroscopy, with some absorption overlap.

The resulting Raman spectrum showed only a few bands, in the 900-2000 cm^{-1} region, where one would expect the amide and lipid frequencies to occur **[Figure 3-10]**. A sharp peak was observed at ~1002 cm^{-1}, and another prominent peak near ~1605 cm^{-1} . These two sharp peaks, in both amplitude and shape, corroborated very well with previously published spectra on styrene-maleic acid thin films [245]. As well, these two sharp peaks have been observed in a

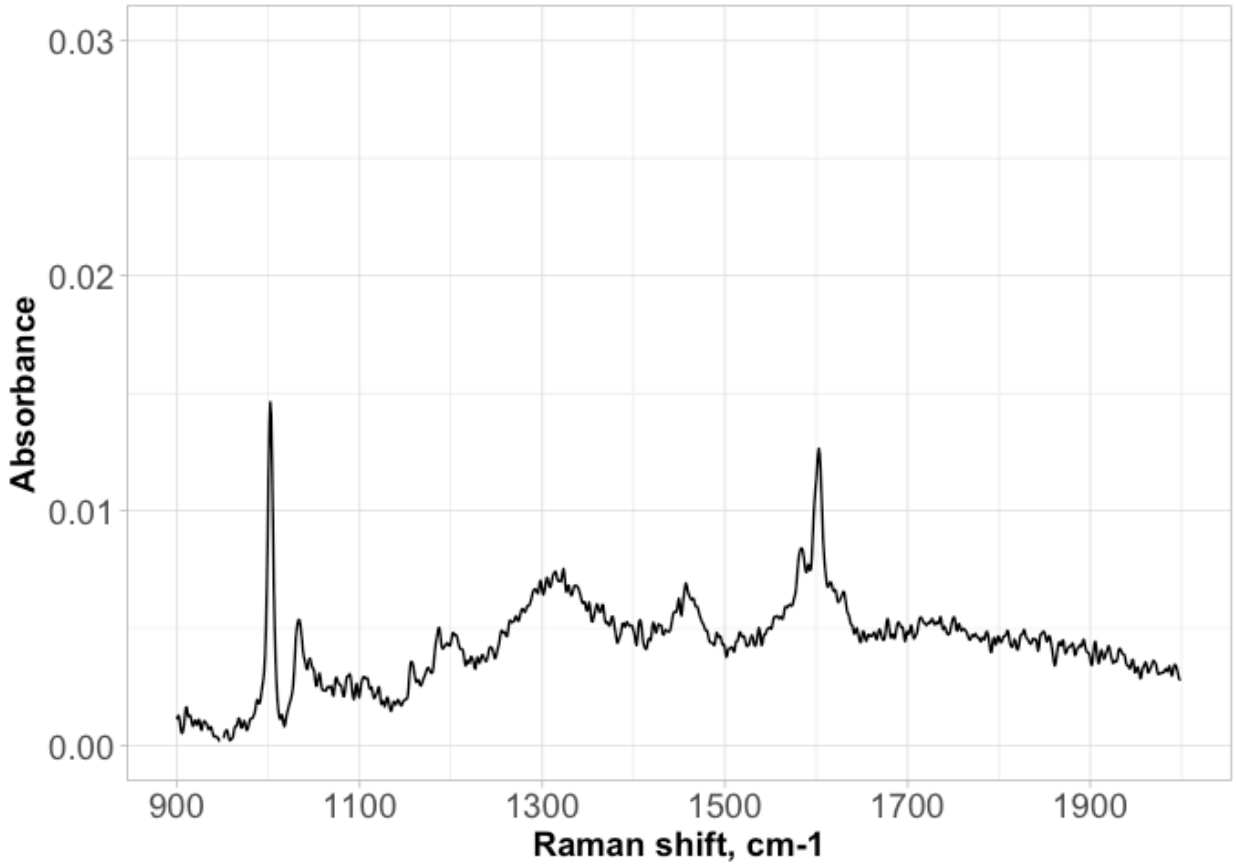

Figure 3-10: *Raman spectra of the AQP1-SMALPs. Predominant bands occur from the presence of the styrene-maleic acid copolymers, meanwhile the concentration of protein was too low to detect any discernible amide bands.*

Raman spectra of a voltage-gated cation Kv7.1 channel encapsulated into SMALPs [246]. Although assigning peaks unambiguously can be difficult, bands can be approximated based on prior published spectra. The Raman shift at ~1000 cm^{-1} can very likely be assigned to the aromatic C-H in-plane bending modes of the styrene group of the polymer. This region could also generally be assigned to either Phe residues or C-H stretching of any of the protein, lipid, or polymer component. As well, the prominent band at ~1600 cm^{-1} might be attributed to the ring vibration mode (C-C stretching), also belonging to the styrene group on the polymer. The broad peak at ~1450 cm^{-1} could be attributed to lipids as the CH_2 scissoring mode, but the band is weak and surrounded by noise in the spectrum. The absence of any discernible protein bands, most prominent being the Amide I band, was odd, although not unexplainable. The amide I band in the Kv7.1 channel spectrum was also quite weak in relation to the styrene and lipid absorption bands [246]. As well, depending on the overall size of the discs, ranging from 10 – 40 nm from the literature, this would impact just how many lipids are enclosed inside the nanodisc. The AQP1-SMALPs have an average diameter of 20 nm, while the AQP1 tetramer itself is close to 10 nm. This leaves 10 nm of space surrounding the protein for lipids to fill the volume, and so could result in a higher lipid to protein ratio than initially expected for the 10 nm size discs assembled from prior preparations.

In summary, to prove that the nanodiscs had indeed self-assembled, contained both protein and lipids, and were stable at room temperature, the nanodiscs were analyzed using negative-staining TEM, Raman spectroscopy, and various gel-based assays. These experiments

confirmed the presence of the SMA co-polymer and lipids in the spectra, the existence of the protein at the expected size when ran on the gel, and the direct imaging of the nanodiscs using TEM confirmed the expected size of the discs and that they had been successfully self-assembled, to a moderate degree of stability as confirmed by gel-based stability assays.

3.3.5 Solution NMR spectroscopy

Solution NMR experiments of ^{15}N-labelled-AQP1-SMALPs were recorded to determine if any ^{1}H-^{15}N peak correlations could be resolved, and in turn if any residues from AQP1 could be observed.

3.3.5.1 1D proton spectrum

Firstly, a 1D ^{1}H spectrum was taken to see if any protons from the protein, polymer or lipids could be resolved [**Figure 3-11**]. In the highlighted 6.5 – 9 ppm region in the spectrum, there is a cluster of peaks normally attributed to the amide backbone protons [247]. Aliphatic protons usually appear in the ~3 – 5 ppm region, and in this spectrum a few peaks are also observed here. The large peak at ~ 4.9 ppm is attributed to water. Aromatic protons are typically found in the 6.5 – 7.5 ppm region, but there are no peaks observed here in the AQP1-SMALPs spectrum. Given the presence of multiple styrene groups within the nanodisc, one would expect a noticeable signal here, however due to the size of the AQP1-SMALPs, the tumbling dynamics may not be entirely favourable for these largely immobilized groups. The nanodiscs tumble slowly and so the aromatic proton peak linewidths may be broadened significantly, and unable to

be differentiated from the background noise. Prior groups were able to resolve these aromatic signals in lipid-only DMPC-based SMALPs, however these peaks were weak as well [244]. The same DMPC-SMALPs were reported to be smaller in size compared to the AQP1-SMALPs, with DLS measurements reporting a 10 nm diameter. A smaller size for the DMPC-SMALPs would mean slightly faster tumbling dynamics of the discs in solution, resulting in a better resolution of these aromatic peaks.

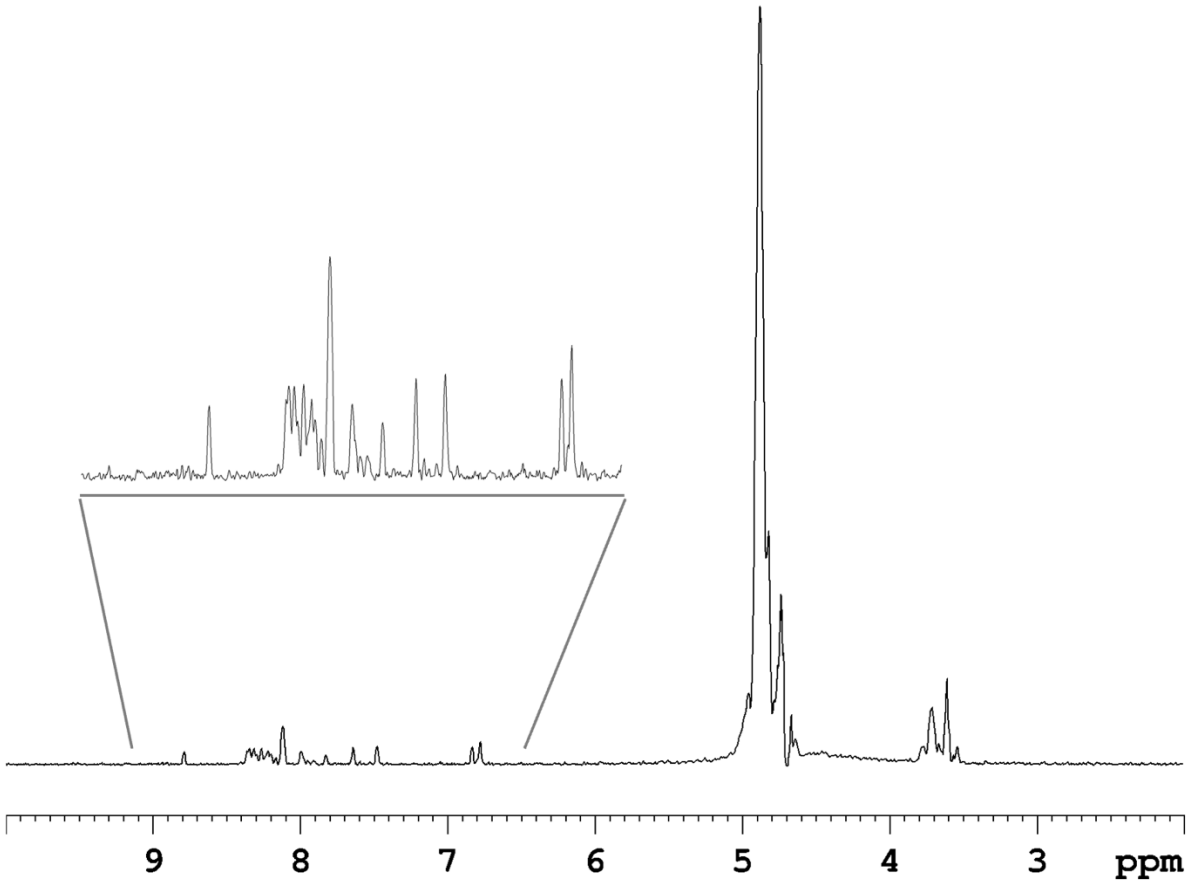

Figure 3-11: *1D proton spectra of AQP1-SMALPs at 600 MHz. The spectrum was recorded at room temperature (24˚C), in NMR buffer (as described above) with 5% D_2O.*

Although no aromatic proton signal was observed, there were plenty of $^{1}H_{N}$ proton signals, corresponding entirely to protein. As well, the aliphatic proton signals could be attributed to either the protein or lipid, however given that no signal belonging to the polymer were observed, it is reasonable to assume the lipid may not be observable either within the spectrum. As it will be seen in the 2D ^{1}H-^{15}N HSQC NMR spectrum, all the peaks between the 6.5 – 9.0 ppm region are attributable to amino acids within AQP1.

3.3.5.2 2D ^{1}H-^{15}N HSQC

Next, a 2D HSQC spectrum was recorded for the AQP1-SMALPs. Here a cluster of well-resolved and well-spaced cross-peaks were observed in the 7.5 – 9.0 ppm region of the proton dimension **[Figure 3-12]**. This corresponds well with the typical ^{1}H-^{15}N HSQC cross-peaks found in a mostly unstructured and disordered protein, comparable to similar 2D spectra from intrinsically disordered proteins (IDPs) [248]. Outside of this region there were a few cross-peaks of interest, mainly the one peak near 8.8 ppm, and a pair of mirrored peaks appearing around 6.8 ppm and 7.4 – 7.7 ppm. The mirrored peaks likely correspond to the NH_2 side chain groups of either Asn or Gln, and are connected on the spectrum with dashed line [249]. There are potentially 4 of these mirrored peaks, however only two with a very strong intensity, while the other two peaks are much weaker.

Close to 34 individual peaks could be resolved in the 2D HSQC spectrum at 25˚C. If we take the prior lack of evidence of any peaks correlating towards the aromatic protons of styrene, or the absence of any strong aliphatic protons from the lipids, it would be safe to assume that the entire nanodisc is too large and tumbling too slow to resolve any substantial cross-peaks from the disc itself. This would mean that any structured regions within AQP1, so most of the helical

structures, would be invisible in this spectrum. This would leave just the loop regions, LA – LE, and the N- and C-termini. In terms of the loop regions, apart from LA and LC, the loops are short linkers between the helices, and may not actually be flexible structures. As well, ssNMR data suggests that loop C is structured as well, forming a significant hydrogen bond network with surrounding helices and residues in the pore [250]. Additionally, earlier ssNMR evidence showed that all the residues of loop C could be assigned in solid-state, and 8 out of 12 residues could be assigned in loop A [**Figure 3-13**] [110]. Most of loop E could also be assigned, leaving just loops B and D that could not be assigned in solid-state NMR, hinting at their high mobility. Loops B and D contain 12 residues overall, which could make up some of the amide cross-peaks in the HSQC spectra. In terms of the terminal tail regions, the N-terminus is very short, consisting of only 6 residues, while the C-terminal tail, after cleavage of the His- and Myc-tags with the TEV enzyme, consists of ~44 residues. Adding up the potentially flexible and thus observable residues from the loop and termini regions, there are a total of ~60 residues, of which approximately 35 gave resolved cross-peaks in the HSQC. However, this is only an estimate of the total that are mobile enough to be observed in solution NMR, and future work will aim to both increase the amount of cross-peaks observed in the 2D spectra and to increase the time until significant signal degradation. As well, further NMR experiments on uniformly ^{15}N,^{13}C- labelled hAQP1 will focus on eventually assigning these cross-peaks to amino acids in the sequence.

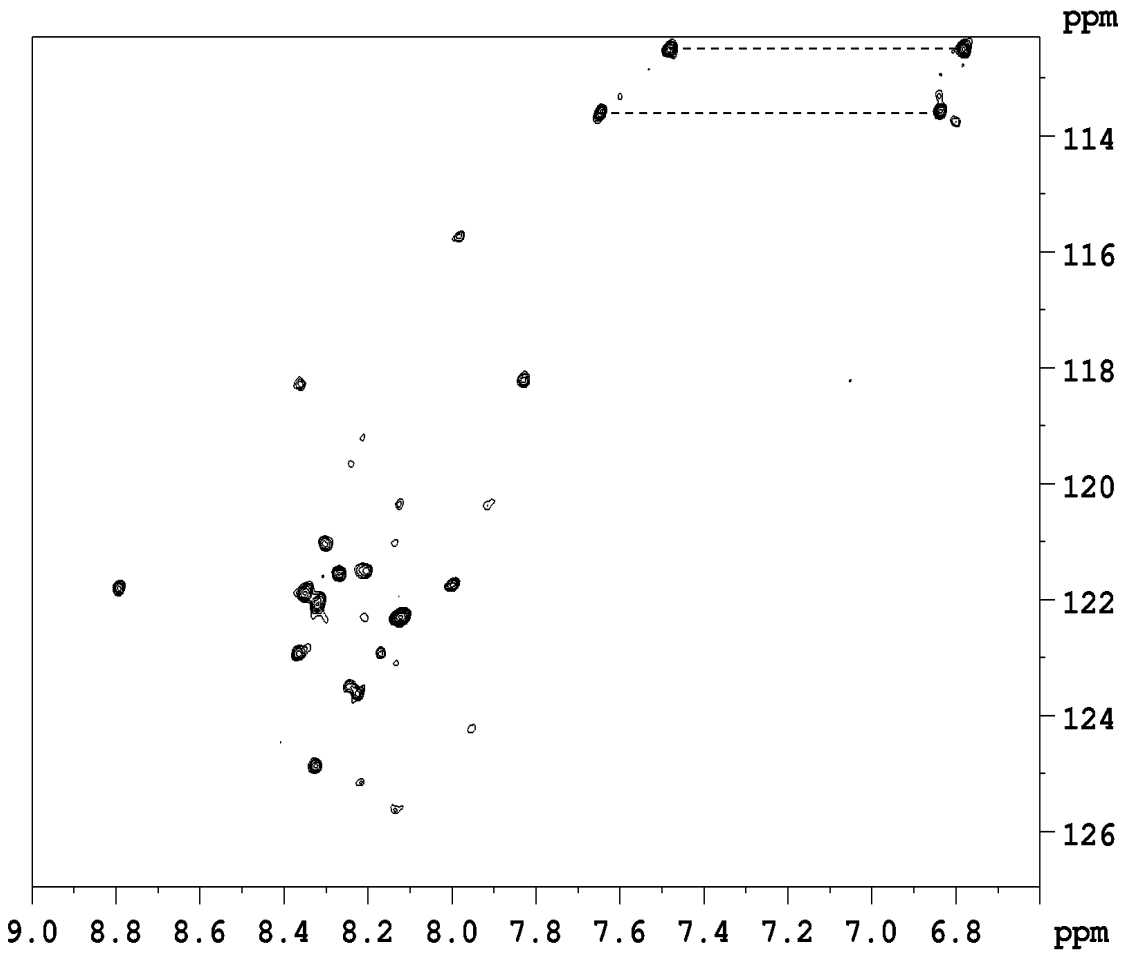

Figure 3-12: *^{1}H-^{15}N Heteronuclear single quantum coherence spectra of ^{15}N- labelled human aquaporin-1 in styrene-maleic acid-based nanodiscs, prepared in NMR buffer with 5% D_2O as described above. The spectrum was recorded at 25˚C at 600 MHz. The ^{1}H signals are shown on the x-axis, while the ^{15}N signals are on the y-axis. Linked pairs of either Asn or Gln side chain NH_2 are connected with a dashed-line. The bulk of cross peaks reside in the region known for intrinsically disordered sequences, hinting that the origin of these cross peaks may be from the loop or terminus regions of hAQP1.*

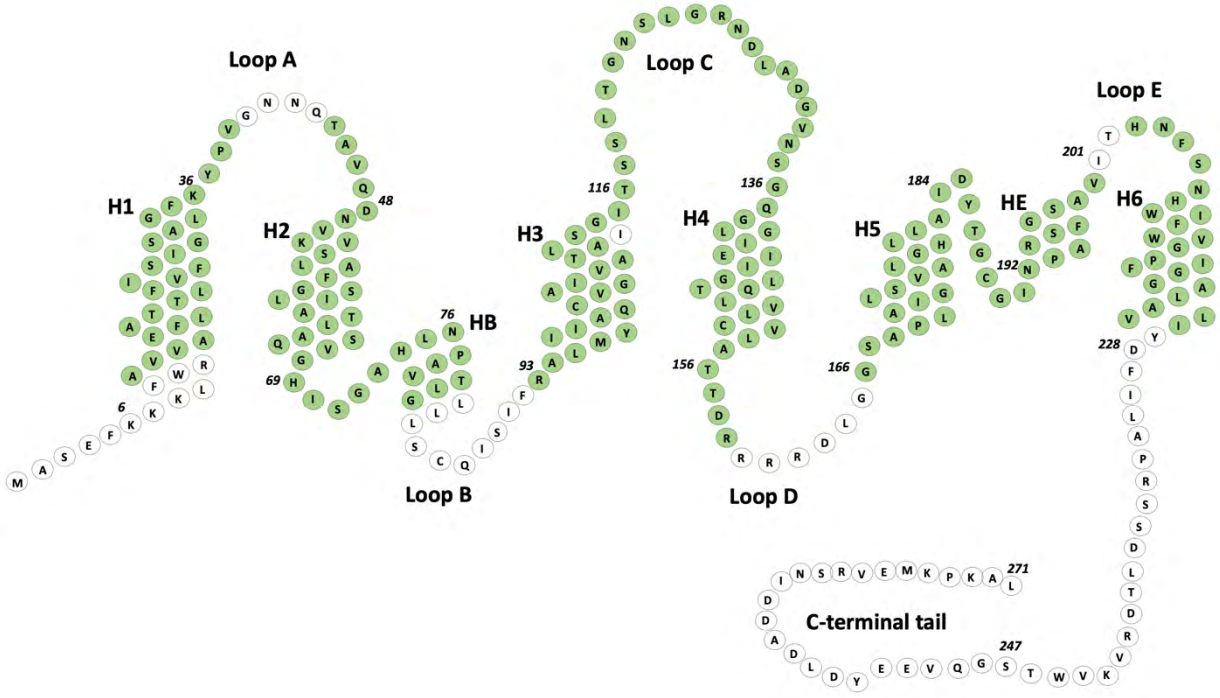

Figure 3-13: *Topological model of hAQP1, with helices and amino acid positions of note detailed. The native long C-terminal tail is also included without the additional tags. All residues that could be resolved in solid-state, and thus structured enough, are coloured in green. Residues unable to be resolved in ssNMR left white. Figure adapted from Wang et al., 2016 [109].*

3.3.6 Nanodisc Aggregation

After performing the solution NMR experiments on the AQP1-SMALPs, it was determined that although the nanodiscs yielded well-resolved spectra, and enough peaks to potentially correspond with the long C-terminal tail region, the signal had started to noticeably decay after ~5 days. The signal strength lasting for 5 days was enough for simple 2D HSQC

experiments and would theoretically be enough time to observe chemical shift changes in a protein-protein titration experiment but would not be enough to perform more complex 3D and 4D experiments that would assist in assigning amino acids to these cross-peaks. As assessed later, 20 days of relative stability would enable the longer 3D NMR experiments for residue assignment of peaks, performed on a uniformly ^{13}C- and ^{15}N- labelled hAQP1 sample. DLS was the technique of choice to determine if indeed the loss of signal was attributed to the nanodiscs aggregating, and later, if the aggregation could be hindered or inhibited completely by changing the final buffer conditions for NMR. The immediate next goal then was to increase the stability of the nanodiscs in solution.

3.3.6.1 Temperature effects on aggregation

Firstly, DLS was used to monitor the AQP1-SMALPs in the same buffer conditions as was performed in the NMR experiments. This meant a buffer consisting of 50 mM potassium phosphate, 10 mM NaCl, and a pH of 7.0. The AQP1-SMALPs were then monitored over the course of 7 days, with measurements on the 1st, 3rd, and 8th day, at incubation temperatures of 4˚C and 25˚C.

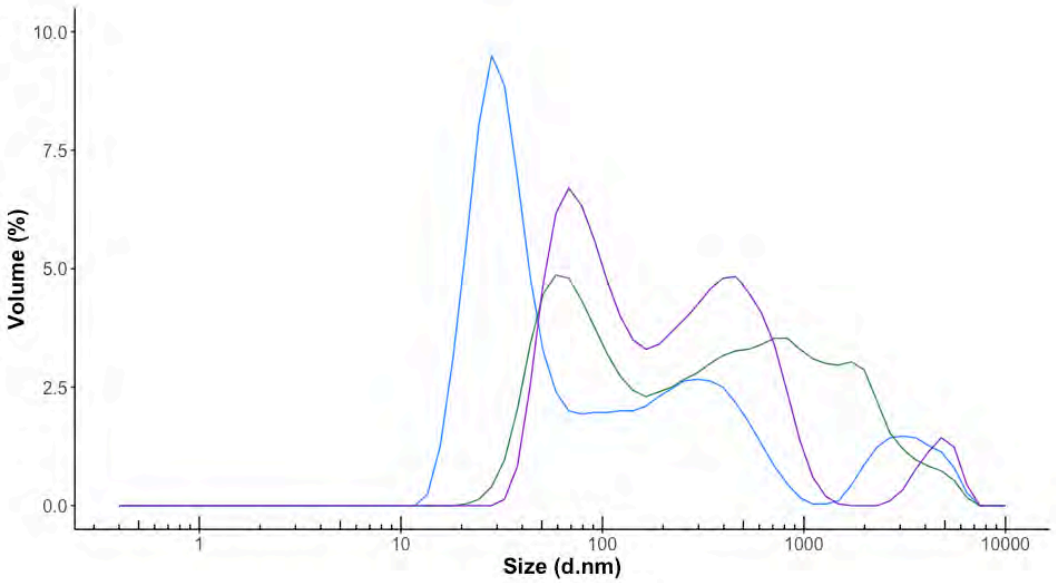

Figure 3-14: *Volume percentage distribution of the AQP1-SMALPs in NMR buffer (20 mM potassium phosphate, pH 7.0, 10 mM NaCl) incubated at 4˚C, and analyzed at days 1 (blue), 4 (green), and 8 (purple).*

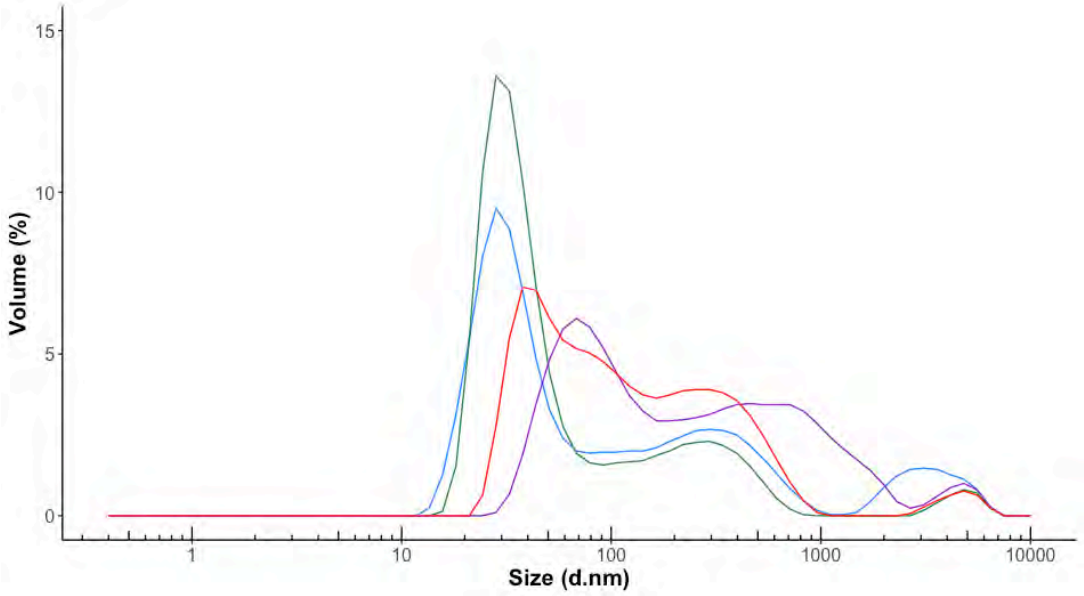

Figure 3-15: *Volume percentage distribution of the AQP1-SMALPs in NMR buffer (20 mM potassium phosphate, pH 7.0, 10 mM NaCl) incubated at 25˚C, and analyzed at days 1 (blue), 4 (green), 8 (purple), and 20 (red).*

The rationale for these temperatures being that 25˚C was the temperature at which the NMR experiments were conducted, and the 4˚C being chosen as a possible temperature to keep the AQP1-SMALPs stored between NMR experiments, as a deterrent to aggregation. Additionally, the 25˚C sample was measured on the 20th day as well, after displaying less overall aggregation after the first 8 days. First, the nanodiscs, incubated overnight at 4˚C following the TEV reaction to cleave the C-terminal His-tag, were measured and found to have an almost normal volume percentage distribution of discs at 28 nm, with a broad range from 15 – 50 nm. Of course, DLS measures the hydrodynamic radius of a particle, which is the equivalent size of a spherical particle that diffuses through solution at the same rate as the measured protein. Knowing this, the SMALPs are not spherical but rather discoidal. This could explain the discrepancy in which the DLS measurements do not completely agree with the measured diameters from the electron microscopy images, which were slightly smaller at ~20 nm.

Some higher molecular weight aggregates were also observed in this Day 1 measurement at around 300 nm. This first day distribution is the same for both 4˚C and 25˚C samples, as they were prepared from the same purification. Then on, the aggregation dynamics were monitored as they were incubated separately. Already at Day 4, the 4˚C nanodiscs had started to aggregate **[Figure 3-14]**. The highest amplitude peak shifted to nearly 60 nm, and higher molecular weight aggregates appeared, nearing 500 and 1000 nm sized species. On the 8th day, the distribution remained comparable to the size profile on Day 4, with the highest amplitude peak at 60 nm, albeit an increase in higher size aggregates forming a sharper peak at ~500 nm. The size profile of the AQP1-SMALPs incubated at 4˚C thus showed a trend towards increasing aggregation, both in the shift in the smallest size distribution of the discs, but also in the increasing appearance of aggregates.

The size profile across a further range of 20 days was examined for the AQP1-SMALPs at 25˚C, as this was the temperature used for previous NMR measurements, and likely the temperature for future experiments. Notably, the Day 1 and Day 4 size profiles for this incubation at higher temperature resulted in a higher stability, whereby the highest amplitude peaks were at 28 and 32 nm, respectively **[Figure 3-15]**. On Day 8, these discs had started to aggregate as well, with the discs displaying higher sizes at almost 70 nm, and a broad range of aggregates from 500 to 1000 nm. On Day 20, this size profile did not change remarkably. The disc sizes were now around 40 nm, however, this distribution was large and included diameters up to 100 nm, and then another large broad peak once again at 400 nm, indicating aggregates. These results indicated that the AQP1-SMALPs at 25˚C, while marginally more stable than 4˚C, were still aggregating past 4 days. This timeline roughly coincides with the appreciable loss of NMR signal at 5 days. Although already speculated to be the cause, the loss in NMR signal could be attributed to the increase in aggregation. Higher molecular weight aggregates would tumble even slower in solution and would yield a broader signal likely even before precipitating completely out of solution. Even though protease degradation could have been another cause as well, the DLS size profiles likely would have shown an increase in much smaller sizes, in the 1-5 nm range if the nanodiscs were degrading to the point of falling apart. Since the reverse was observed, and larger size volume distributions were observed along a function of time, aggregation can be assumed to be the likeliest cause.

This phenomenon is not entirely surprising, given the complex assembly of the SMALPs, and its inherently “sticky” nature. The discs themselves are composed of surrounding negative charges from the maleic acid subunits from the SMA polymer belt and the lipid headgroups, but also of a collection of positive charges on one face of AQP1 [180,184,189,251]. These

electrostatic interactions from neighbouring nanodiscs in solution may be accelerating the aggregation process into higher molecular weight species. As well, it has been reported that collisions between SMALPs in solution can also facilitate the diffusion of lipids between the discs [252]. This highly dynamic lipid-exchange behavior within SMALPS is unique among the other membrane mimetic systems, but so far it is unclear how this affects the overall stability and tendency of aggregation of SMALPs within solution. To further assess the aggregation dynamics of these SMALPs, and to circumvent it, a series of conditions for the AQP1-SMALPs were experimented with that were still amenable to solution NMR, but could yield more stable discs in solution.

3.3.6.2 Buffer conditions

The solution conditions tested for the AQP1-SMALPs include a high and low salt buffer, a glycerol buffer, a buffer with EDTA, and the prior NMR buffer as a control, in which the size profiles were already determined for with DLS. The rationale for these were to assess whether a high or low salt environment would reduce aggregation in the nanodiscs, given that the ionic strength of a solution will impact the electrostatic interactions between colliding nanodiscs. The low salt conditions, in addition to having a lower salt content (5 mM NaCl), also had the previous potassium phosphate buffer switched to HEPES. The potassium ions from the previous buffer are known to contribute to the ionic strength of a solution, whereas HEPES is neutral in this regard, and buffers well still within the pH 7.0 range. As well, the glycerol buffer condition was to test whether the addition of an osmolyte would prevent the onset of aggregation. Since glycerol binds to the backbone amides of proteins, its presence provides a stabilizing effect. Lastly, a buffer with added EDTA was also tested, given that SMA binds metal ions quite well, it

was believed that EDTA could sequester any metal ions possibly present within the nanodisc complex.

The size profiles of both volume and intensity percentage from the various buffer conditions showed some interesting findings **[Figure 3-16]**, **[Table 3.2].**

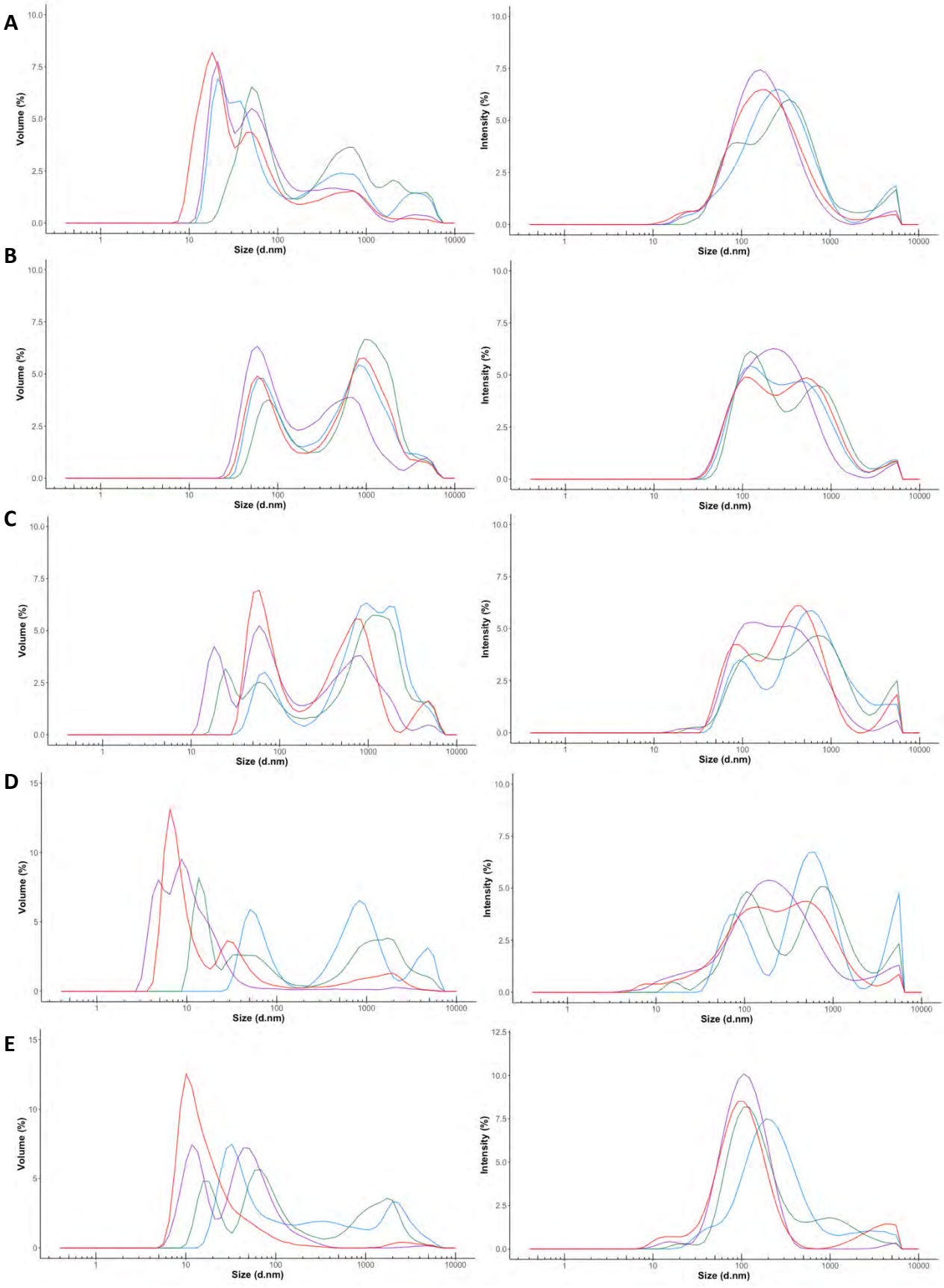
A
B
C
D
E
Volume (%)
Intensity (%)
Size (d.nm)

Figure 3-16: *Volume and intensity percentage distributions of the AQP1-SMALPs in differing buffer conditions, incubated at 25˚C and analyzed at days 1 (blue), 4 (green), 8 (purple), and 20 (red). Each DLS measurement on each day is an average of a triplicate, with each triplicate containing 13-16 individual scans. The control NMR buffer (**A**), same conditions as the previous NMR experiments (20 mM potassium-phosphate, pH 7.0).* ***(B)*** *EDTA buffer.* ***(C)*** *1% glycerol buffer* ***(D)*** *high salt (200 mM NaCl) buffer* ***(E)*** *and low salt (5 mM NaCl) + HEPES.*

Firstly, the control NMR buffer displayed similar findings as before, where the overall size of the discs remained around 20-40 nm, with some aggregation near higher size species appearing over time. *Figure 3-16:B* shows the size profiles of the AQP1-SMALPs buffer exchanged to the same buffer but with the addition of EDTA. Interestingly this resulted in a large distribution of sizes at and above 1000 nm in diameter, a higher amount present than its smaller sized species at ~60 nm. The intensity percentage trace showed a similar distribution profile. In *Figure 3-16:C,* the buffer containing 1% glycerol also showed its highest peaks towards 1000 nm in size, with a trend towards even larger aggregates over time, as the size profile on Day 20 shows a substantial peak even at diameters of 5000 nm, evident in both the volume and intensity percentage profiles.

Although it was not expected for the high salt buffer to alleviate the aggregation, it was found that the presence of a higher salt concentration, and therefore a higher ionic strength, proved even more detrimental towards the stability of the AQP1-SMALPs as compared to the control **[Figure 3-16:D]**. The volume percentage plot, on day 1, shows a somewhat even distribution between 50 nm size discs and 1000 nm size aggregates. However, as incubation time increases, even smaller sized discs and larger aggregates appear. By day 20, it appears sizes as

low as 5 nm are observed, indicating a large amount of degradation occurring, with the nanodiscs losing their integrity and falling apart. It is also possible given the much smaller sizes observed, that AQP1 itself may be crashing out of the nanodiscs, no longer staying solubilized within the nanodiscs but aggregating out of solution entirely. Overall, degradation seems to be the most plausible explanation for these size changes over time given the equal disappearance of the higher size species as incubation time increases.

Finally, *Figure 3-16E* shows the size distributions of the AQP1-SMALPs suspended in the low ionic strength buffer. The volume size distribution looks more comparable to the control, with what initially seems to be a binomial distribution of nanodiscs appearing after Day 1. On Day 4, these two peaks occur at 15 nm and 60 nm, whereas on day 7 they are 10 nm and 50 nm. After 20 days, one large peak with an amplitude at 10 nm, but ranging from 7 – 50 nm diameters is observed. This seems to indicate possibly some degradation is occurring, but not nearly at the same rate as observed in the high salt conditions. Of particular note, however, is the absence of higher molecular weight aggregates past the 7-day mark. There appears to be some initial aggregation during the first 4 days, however the amplitude of these peaks are the smallest of all the buffer conditions tried. As well, the intensity percentage plots for the low salt buffer appear to be the most consistent across the entire span of 20 days when compared to the other conditions and the control.

Table 3.2: *Summary of DLS volume intensity measurements from the temperature trials and the differing buffer conditions. First row displays measurements from the 4˚C and 25˚C temperature incubation experiments. The second row details the buffer condition measurements. The control sample contained the same buffer treatment as the temperature trials.*

AQP1-SMALP sample type	*Highest volume % amplitude peak size (diameter-nm)*	*Second highest volume % amplitude peak size (diameter-nm)*
4˚C		
Day 1	28	~40
Day 4	59	~700
Day 8	71	~500
25˚C		
Day 1	28	~300
Day 4	29	~300
Day 8	70	~500
Day 20	38	~350
Control		
Day 1	21	40
Day 4	50	54
Day 8	21	52
Day 20	18	50
EDTA		
Day 1	62	~900
Day 4	75	~1100
Day 8	60	~700
Day 20	55	~950
Glycerol		
Day 1	71	~1000
Day 4	25	~1200
Day 8	18	~900
Day 20	60	~750
High salt		
Day 1	48	~800
Day 4	14	~2000
Day 8	9	4.5
Day 20	6	30

Low salt	32	~2000
Day 1	60	16
Day 4	10	45
Day 8	10	-
Day 20		

3.4 Conclusions and future remarks

Throughout this work outlined in *Chapter 3*, AQP1 was successfully solubilized with the SMA co-polymer, resulting in high purity AQP1 containing nanodiscs. A few different solubilization conditions were experimented with to increase the yield and to prevent protease degradation or aggregation from occurring. These changes resulted in a high temperature, low salt solubilization of AQP1 from *P. pastoris* membranes, conditions that increased the solubilizing efficiency of SMA by nearly 100%. The final yield per litre of culture amounted to approximately ~3.3 mg, which when concentrated for solution NMR resulted in 0.21 mM of AQP1-SMALPs. The subsequent ^{15}N-labelled AQP1-SMALPs gave well-resolved and dispersed ^{1}H-^{15}N HSQC spectra, with a total of approximately 35 H-N correlations observed. The absence of any notable aromatic protons from the styrene groups from the polymer, and of any aliphatic protons from the lipids, suggests that the nanodiscs themselves tumble very slowly. However, this evidence further suggests that the cross-peaks indeed observed may come from only the most flexible regions of AQP1, a combination of the loop and termini regions in AQP1. Just how many and which residues appear from the C-terminal tail needs to be elucidated still.

Although these discs are likely too large to give any appreciable signal from the structured and non-flexible elements in the complex, and the fact that the NMR signal degrades after 5 days seems to suggest there is aggregation occurring, which was further confirmed by the DLS measurements. The DLS results from the AQP1-SMALPs suspended in the original buffer

for NMR suggest that the discs may form dimers or even higher molecule weight structures that are still suspended in solution, nearing size distributions from 20 – 50 nm. Given that TEM images confirmed that the individual disc sizes are on average 20 nm in diameter, these higher size species may indeed be some nanodiscs sticking to one another, forming oligomers. These oligomeric species likely catalyze the formation of even higher mass aggregates, as also confirmed by DLS measurements as the aggregation continues as time goes on. This aggregation is likely the cause of the decay in NMR signal after 5 days.

Finally, to inhibit the aggregation, a series of buffer conditions were experimented with to see if they would have any effect on the AQP1-SMALPs. DLS measurements showed that indeed changing the conditions even slightly had drastic changes on the aggregation mechanisms, and as the case for the high salt condition, could also introduce massive instability in the nanodiscs themselves, causing them to degrade into smaller sized species. Interestingly, there is some potential given the low salt, low ionic strength condition, where the AQP1-SMALPs showed the least amount of aggregation, eventually reaching an equilibrium of ~10 nm sized discs.

Once the aggregation of the AQP1-SMALPs in solution is inhibited, then further NMR experiments can be performed on a uniformly ^{13}C, ^{15}N-labelled AQP1. 3D and 4D experiments involving different types of bond correlations will need to be recorded to assign cross-peaks in the ^{1}H-^{15}N HSQC spectrum to specific residues within the AQP1 sequence. Once all the peaks have been assigned, then the AQP1-SMALP platform is ready for protein-protein titration experiments, recorded again using just the ^{1}H-^{15}N HSQC spectrum, but this time with all the peak positions known.

Chapter 4

Investigation into a conserved hydrogen-bond network in AQP1 using FTIR spectroscopy and H/D exchange

---

4.1 Introduction

The aquaporins are a family of proteins that is found ubiquitously throughout the kingdoms of life, given the passage of water molecules through biological membranes is a requirement for cellular life.

Aquaporin-1 homologs across different species share a high structural homology. Across humans, rats, bovine, spinach, yeast, and *E. coli,* aquaporins retain high conservation among the two NPA regions (Asn76-Pro77-Ala78 and Asn192-Pro193-Ala194 in hAQP1) [**Figure 4-1**]. Partial conservation exists in the aromatic/arginine selectivity filter, which includes the following residues: Phe56, His180, Arg195, Gly188, and Cys189. F56, H180, R195 are entirely conserved among these species, while Gly188 is almost conserved apart from AQPZ from *E. coli*, which instead contains an Asn at this position. The residue responsible for the well-known mercury inhibition of aquaporins, Cys189, is only present in the mammalian AQP1s, but replaced with a Thr in the yeast, bacterial, and spinach AQPs. These residues define the pore in which water transport occurs, and differences can be observed when comparing the classical AQPs to the aquaglyceroporins. A combination of these differences dictates the diameter of the pore which can allow the passage of glycerol through the AQGPs. For the following discussion, human AQP1 sequence numbering for all residues and motifs will be used.

As a quick overview from the introduction to aquaporins in *Chapter 1*, the AQPs numbering 0, 1, 2, 4, 5, 6, and 8 belong to the classical group, whereas AQPs 3 ,7 ,9, 10 belong to the aquaglyceroporins family. AQP11 and 12 belong to their own superaquaporins family given their large deviations in the pore region. A structural alignment [**Figure 4-2**] shows these differences between the human AQPs and the AQGPs.

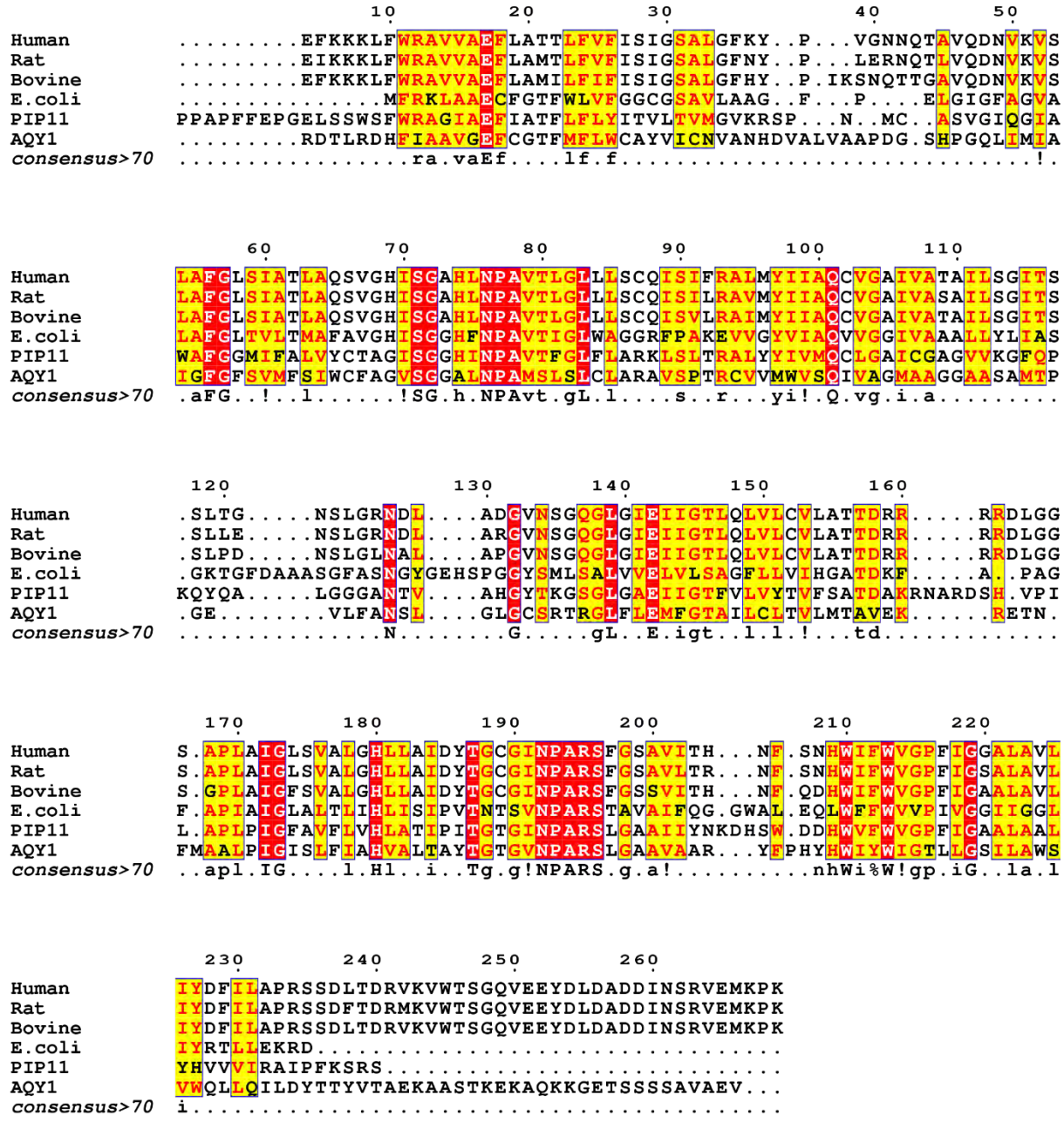

Figure 4-1: *Structural sequence alignment of aquaporins across a few different species (human, rat, and bovine AQP1, E. coli AQPZ, PIP11 aquaporin from Arabidopsis thaliana, and AQY1 from Pichia pastoris). Positions of interest related to the mutagenesis of hAQP1 in this work are located as follows on the alignment chart: N127, V133, E142, T187, R195, and S196.*

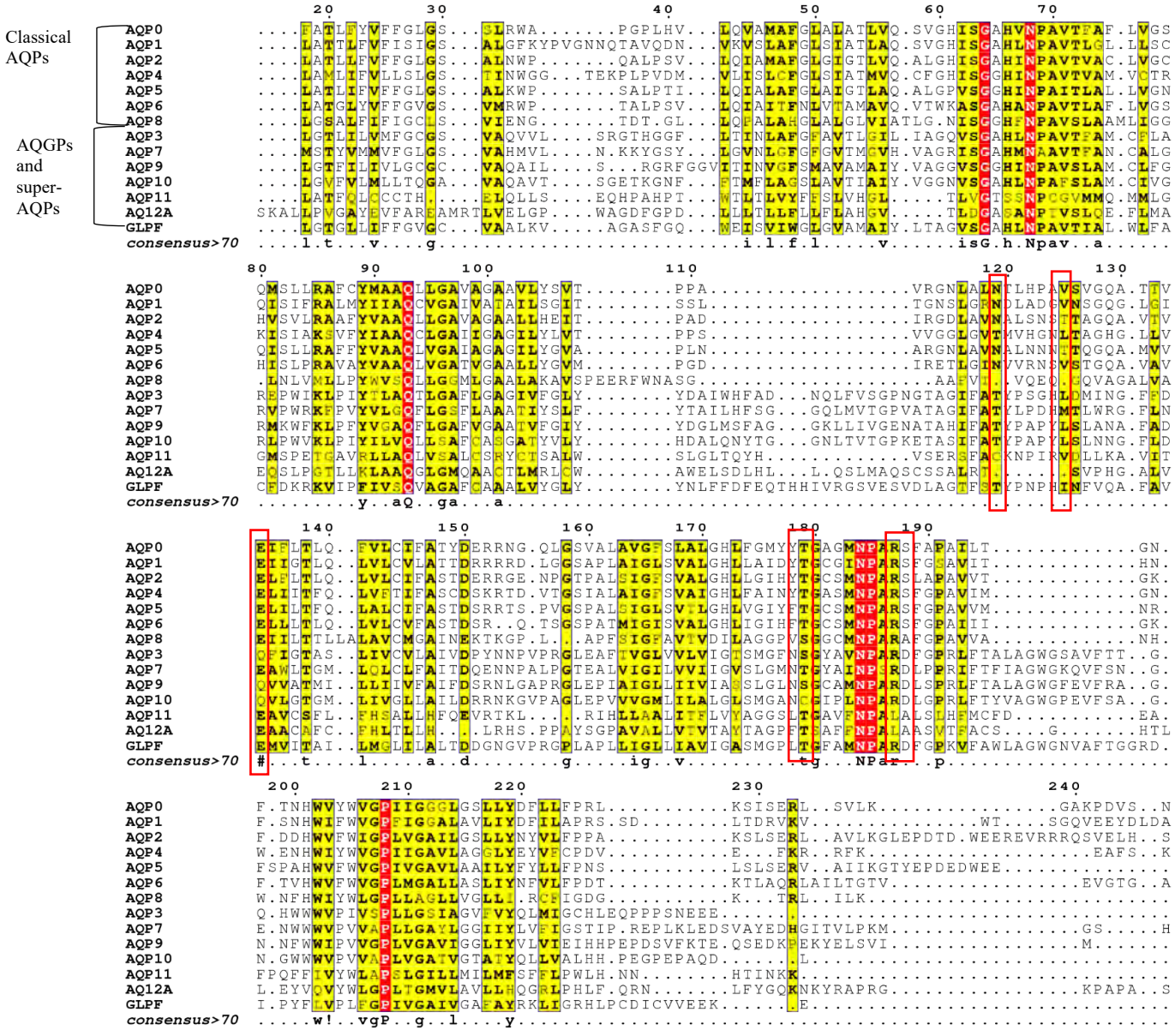

Figure 4-2: *Multiple sequence alignment of all the mammalian aquaporins (1-12) and bacterial glycerol transporter GLPF. Sequences shown are from position ~20-248 to highlight most relevant regions. Positions of interest related to the mutagenesis of hAQP1 in this work are located and bolded as follows on the chart: N127(**119**), V133(**125**), E142(**134**), T187(**179**), R195(**187**), and S196(**189**). Relevant positions are also highlighted with a red box.*

For the most part, the dual NPA motifs are entirely conserved among all aquaporins, apart from AQP11 and 12, albeit the numbering is different considering it is a structural alignment.

The largest differences between the classical AQPs and the AQGPs occur in the aromatic/Arginine selectivity filter, which is comprised of residues Phe56, His180, Arg195, Gly188 and Cys189 (on the structural alignment, they correspond to positions *48, 172, 187, 180, 181 due to the alignment*). For the AQPs, the ^{192}NPARS^{196} region is very conserved. However for the AQGPs, this motif is instead ^{192}NPARD^{196}, with the exception of AQP8, 11, and 12 containing an Ala196 instead of serine. Interestingly, the *E. coli* GLPF also contains an Asp196 instead of a serine. His180 *(172)* is completely conserved in the classical AQPs apart from AQP8 once again. The glyceroporins do not have this histidine conserved, but instead contain non-polar, smaller side-chain residues such as Ile, Gly, Val, or Ala, suggesting that for glycerol transport the specific residue is less important than the lack of the bulky imidazole ring. Gly188*(180)* is completely conserved in both the aquaporins and the glyceroporins, apart from AQP12. Lastly, Cys189*(181)* is only conserved among AQP1, 2, 5, 6, and 9.

Interestingly, when looking at AQP1 homologs across species, and across the different types of aquaporins in humans, there are residues that are strongly conserved in regions that are not directly involved with water transport. These regions are thus an unexplored area in the aquaporin structure, even though this strong conservation has evolutionary implications towards the native fold, and in turn its function. Some of these residues have not yet been fully investigated experimentally either, through mutagenesis coupled with activity assays, a decidedly intriguing gap in the research. However, there does exist some evidence that hints that these amino acids indeed have a role in the overall stability of the native fold, as previously discussed in *Chapter 1* as well [110,250]. Now with more recent work, performed by fellow lab member

Xiao Peng, a study involving solid-state NMR and H/D exchange to probe the unfolding pathway of AQP1, leads the way toward an investigation into a potential hydrogen-bond network that stabilizes the fold of AQP1 – a network that likely exists among the broad AQP family as well. This final chapter will then discuss the mutagenesis of AQP1, involving N127, V133, E142, T187, R195, and S196. All the amino acids, apart from Val133, were replaced with an alanine residue to disrupt the potential hydrogen-bond patterns involving their sidechains. Val133 was replaced with a Pro, with the same idea to disrupt local hydrogen-bonding, but within the backbone (with Ala130 carbonyl), which enacts a sharp turn constituting the type II β-turn motif found near the end of loop C, between helices H3 and H4. This structural motif is completely conserved for AQP1 homologs among the different species, and even though the exact identity of Ala130 and Val133 are not conserved, structurally similar amino acids are often conserved instead. In addition, this β-turn also shows strong conservation among other human aquaporins as well. This high conservation in these specific regions suggests an involvement in the overall structural stability in AQPs, and therefore an indirect involvement in the water channel function. The mutants in this chapter were analyzed using attenuated total reflectance FTIR spectroscopy, coupled with hydrogen/deuterium exchange to probe the extent of residues exposure to solvent upon increasing temperature. As well, secondary structure changes were also observed by monitoring the changes in the amide I and II bands.

β-turn and Loop C

The first region that shows not only residue specific conservation, but also a conserved structural feature, is the β-turn in loop C, between helices H3 and H4. Loop C is a long, extracellular loop that is a unique feature among aquaporins, given the presence of two flanking

conserved β-turns, and may interact with residues that define the selectivity filter, or with neighbouring residues [109,110]. The β-turn at the H3 helix end of loop C may contribute to stabilizing the loop C structure down towards the pore and is possibly stabilized by a few hydrogen-bonds.

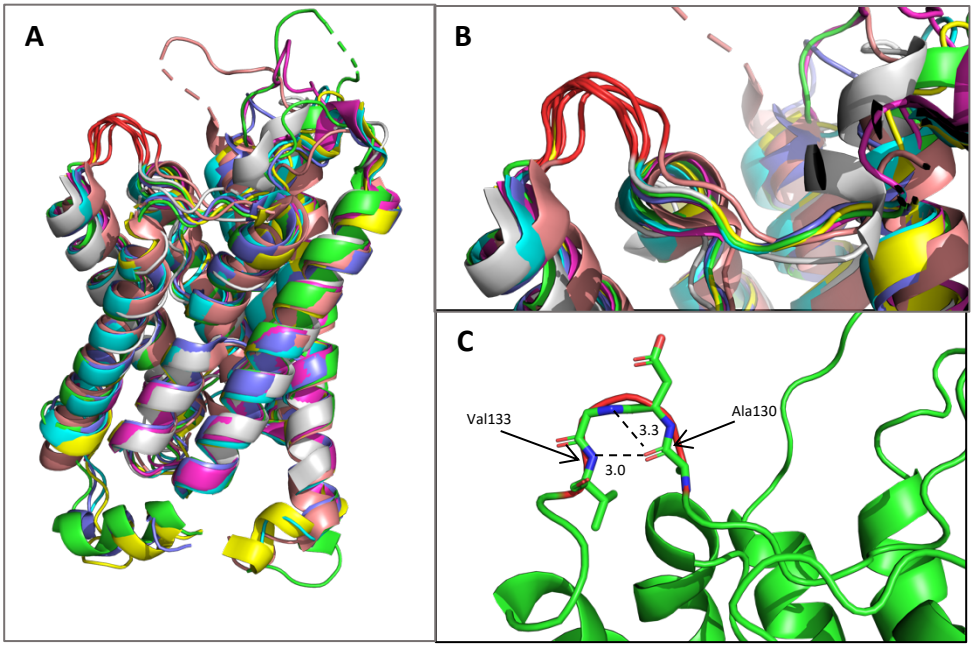

Figure 4-3: *Structural alignment of several aquaporins, showing the β-turn motif coloured in red. (**A**) Structural alignment of the following AQPs: bovine AQP1 (color: green, PDB ID: 1J4N), hAQP2 (blue, 4NEF), hAQP4 (purple, 3GD8), hAQP5 (yellow, 3D9S), AQY1 from P. pastoris (salmon, 3Z0J), hAQP7 (gray, 6QZI), and bovine AQP0 (navy, 2B6O). (**B**) A zoomed-in picture into the alignment of the β-turn motif. All the structures possess this sharp turn in the loop joining helices H3 and H4. (**C**) Potential hydrogen-bonding interactions in the β-turn illustrated on a 1J4N homology model of AQP1 with accurate human AQP1 numbering, showing the residue contacts that stabilize these β-turn. Backbone carbonyl of Ala130 has a potential hydrogen-bond with the backbone amide of Val133. Ala130 carbonyl also potentially hydrogen-bonds with Gly132.*

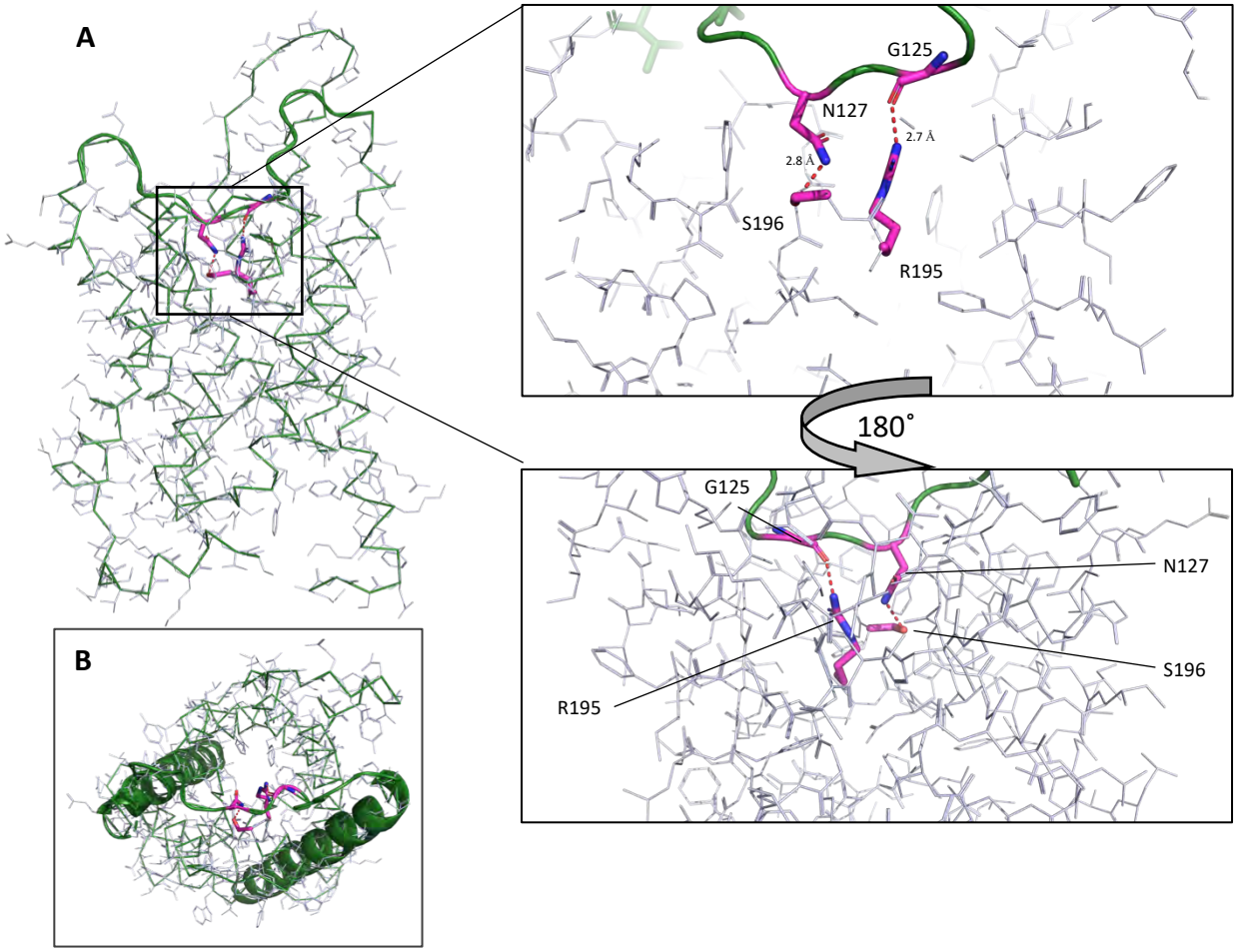

Figure 4-4: (A) *The potential hydrogen-bond network involving loop C and the sidechains of the residues of N127, S196, R195, and the backbone of G125. Secondary structure of AQP1 shown as ribbons, with the involved residues shown as sticks for clarity. Shown amino acids are buried deep in the pore of AQP1, with N127's side chain amine group forming a potential hydrogen-bond with S196's hydroxyl group, anchoring in the loop C within the pore. The N127 sidechain can also possibly hydrogen-bond with R195, and the backbones of G190 and I191.The R195 – G125 interaction is also shown.* ***(B)*** *Top down view of loop C and its connecting H3 and H4 helices. Structures created with a homology model based on 1J4N coordinates with hAQP1 sequence and numbering, modelled via the SWISS-MODEL via the Expasy web server.*

This β-turn motif occurs between Ala130 and Val133, and is structurally conserved among many aquaporins, both across species and other human homologs [**Figure 4-3**]. To this end, Val133 was chosen as a site of mutation to proline, which not only would disrupt the hydrogen-bonding between the backbone N–H donor with the Ala backbone C=O acceptor, but also introduce a sharp kink in the region, likely disrupting the local secondary structure. This hydrogen-bond between residues in the β-turn may help with stabilizing the entire region, as this turn ensures loop C is aimed down towards the pore in the structure. Without this β-turn motif, the entire region may in turn become more flexible, as loop C is known to be quite structured and rigid according to prior solid-state NMR measurements.

Loop C anchoring interactions

Along the middle region of loop C there exists another network that potentially stabilizes the loop, as well as the adjacent H3 and H4 helices [**Figure 4-4**]. The residues involved in this network are N127, R195, S196, and G125. The amine from the sidechain of N127 likely forms a hydrogen-bond with S196's side-chain hydroxyl group, as the distance is 2.8 Å in the homology model of AQP1. This interaction is also present in many structures of other AQPs, including several human homologs [47,66,253–255] Of significant interest as well, is when looking at the highest resolution structure available for an aquaporin, AQY1 from *Pichia pastoris* at a resolution of 0.88 Å, the conformation of the N127 sidechain differs slightly [253,256]. In the AQY1 structure, instead of the amine, N127's carbonyl group potentially hydrogen-bonds with S196's hydroxyl group. The N127 sidechain is free to rotate, and so in this conformation the amino group has two hydrogen bond donors, which are in close enough proximity to the

backbone carbonyls of Gly220 and Ala221 (in human numbering this corresponds to Gly188 and Cys189). With this high-resolution structural information from AQY1, it is quite possible the N127 equivalent in hAQP1 has 3 hydrogen-bond partners, further illustrating how important this single residue may be towards stability in the loop C region. Due to the proximity and long sidechain of R195, this residue is likely interacting with N127 as well [253,254].

As well, the amine from the guanidino group of R195's sidechain could form a hydrogen-bond with the backbone carbonyl of G125, with a distance of 2.7 Å between H-bond donor and acceptor. Among the mutants explored in this work, R195 is the only position in which prior mutagenesis studies have been performed [257]. A mutation of R195 to Val was reported to have no effect on overall water permeability but had been shown to allow the passage of protons through the pore. As well, a double mutation of H180A – R195V in AQP1 allowed urea to pass, consistent with previous findings that the increase in the diameter of the pore would allow the permeation of other solutes. These mutations were not characterized structurally, however another paper had reported that a R195C mutation disrupted oligomerization and water permeability [39]. However, it is still unclear whether any local secondary structure changes also impact the passage of protons or other solutes. This chapter will include the first structural investigation into these mutants. N127A, R195A, and S196A were the mutations explored here, all of which are replaced with an alanine that cannot form any hydrogen-bonds with its sidechain. Since these residues reside in loop C and the surrounding region, any disruption of stabilizing factors could potentially disrupt the entire loop C structure, and in turn also destabilize the connecting helices of H3 and H4. This region could very well be critical in stabilizing the overall helical packing in the structure.

H-bond network involving E142, T187, G190, and I191

Lastly, there exists another possible hydrogen-bond network involving residues from the H4 helix and the loop region linking H5 and HE. E142, from the H4 helix, has a few potential hydrogen-bonding partners in close proximity [**Figure 4-5**]. E142's carboxyl sidechain can form potential H-bonds with the backbone amides of G190 and I191. As well, the sidechain of T187 from H5 linker loop can form another possible hydrogen-bond with E142. This hydrogen-bond network thus simultaneously connects helices H4, H5, and the NPARS motif-containing HE. These interactions potentially stabilize the selectivity filter in the pore, namely R195 which also potentially interacts with G125 from the neighbouring loop C. E142 may especially be integral towards the stability of this entire region, forming up to 3 hydrogen-bonds which may keep the connected helices bundled together. In this work, both E142 and T187 were replaced with an alanine residue to disrupt this hydrogen-bonding network, and to observe the effects of these mutations on the overall stability of the protein.

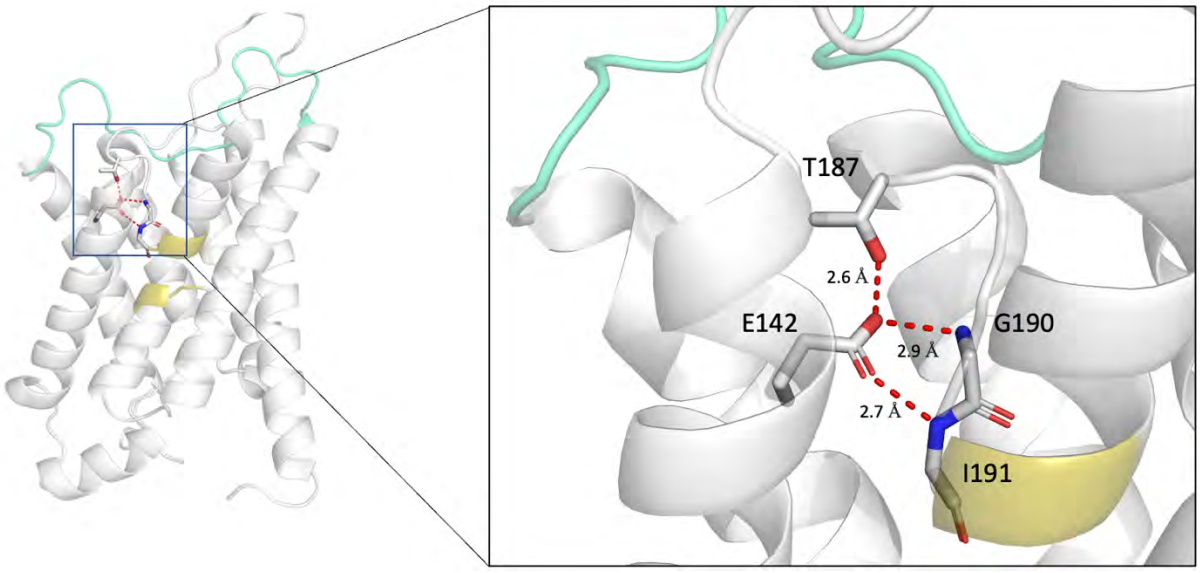

Figure 4-5: *The proposed hydrogen-bond network involving the sidechains of E142 and T187, and the backbones of G190 and I191. Residues of importance coloured by element, while the nearby NPA motif is represented by green line structure. Structure shown is a homology model of 1J4N with hAQP1 sequence residues and numbering, produced using SWISS-MODEL*

Objective

In this final chapter, the work relating to the mutagenesis of hAQP1, and the subsequent FTIR coupled with H/D exchange experiments will be discussed, as well as some DSC experiments involving the WT and the N127A mutant. The objective of these experiments was to monitor two crucial elements in probing the overall stability of these new mutants spectroscopically. Firstly, using hydrogen/deuterium exchange the overall solvent accessibility of the amino acids in the structure could be monitored upon increasing temperature. Using prior solid-state NMR and differential scanning calorimetry data of the WT, a transition temperature between 60 – 65˚C was used as the final temperature endpoint for all of the mutants – the point at which the WT is fully exchanged and most residues are entirely exposed to solvent, and thus most N–H groups of the backbone amides now exchanged to N–D. This exchangeability of the backbones can be measured using FTIR, where upon exposure to D_2O, the amide II band will uncouple from the N-H vibrations and the subsequent amide II band's amplitude will increase. Consequently, the extent of exchangeability can be monitored using the total amplitude or the integration of these two bands.

Secondly, FTIR can also measure secondary structure changes in the mutants upon increasing temperature. After accounting for the small deuteration shift on the amide I band,

further changes in the amplitude and shapes of this band can suggest the existence of secondary structure changes, both in the initial state and upon increasing temperature.

Lastly, differential scanning calorimetry was performed on the WT and the N127A mutant to observe changes in the melting temperature, and to calculate the calorimetric enthalpy of unfolding. Overall, the mutagenesis of hAQP1 at specific hydrogen-bonding networks in loop C and neighbouring helices have illustrated the importance of these residues towards the overall stability of the protein fold. ATR-FTIR spectroscopy coupled with H/D exchange has provided valuable insights into the structure and stability of hAQP1, and future experiments will likely continue to explore the mutants' stability through DSC, and possibly investigate whether these mutations also impact the intrinsic water transport function.

4.2 Materials and Methods

4.2.1 Pichia pastoris expression of hAQP1 and its mutants

The protease-deficient *Pichia pastoris* strain SMD1168H (Invitrogen) was transformed via electroporation with the expression vector pPICZB-hAQP1-Myc-His6, which encodes full-length human aquaporin-1 with a Myc and 6xHis-tag. The mutant constructs (N127A, V133P, S196A, E142A, T187A, and R195A) were transformed the same way and the following protocol was the same for the wild type and the mutants. Cells expressing the mutants were handled as described in *Chapter 3*, and so the procedure will be summarized here briefly. *P. pastoris* cells transformed with this vector were re-plated on yeast extract-peptone-dextrose (YPD) plates with 2 mg/mL Zeocin. Isolated cell colonies were scraped and suspended in buffered minimal dextrose (BMD) for large scale growth. Before induction with methanol, the cells were spun

down and resuspended in buffered minimal media (BMM), in which they were incubated for another 24 hours before pelleting for collection and storing at -80˚C.

4.2.2 Solubilization and purification

Cells were broken as described previously in *Chapter 3*, but will be mentioned briefly here once more with some minor modifications. Glass beads were used along with vortexing on 1 min "on" and 1 min "off" cycles to break the yeast cell wall and membranes. The resulting supernatant was ultracentrifuged at 100,000xg for 30 mins to collect the broken membranes. The membranes were then suspended in membrane storage buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 20% (v/v) glycerol, 2 mM EDTA, 1 Roche cOmplete EDTA-free protease inhibitor cocktail tablet per ~40 mL of buffer) and stored at -80˚C until proceeding to the solubilization step.

Prior to solubilization, the membranes were thawed on ice for ~30 mins, and then an equal volume of solubilization buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 20% (v/v) glycerol, 5% n-dodecyl-β-D-maltopyranoside (DDM)) was added, resulting in a final DDM concentration of 2.5%. The membrane/detergent mixture was incubated at 4˚C for 2 hours. Following this, the solubilized membranes were again ultracentrifuged at 100,000xg for 30 mins. The supernatant was collected and combined with 4 mL of Ni-NTA agarose beads (Qiagen) and left to stir overnight at 4˚C to promote as much binding to the nickel beads as possible. The following morning the protein solution was added to a gravity column and the entire solution flowed through. The Ni-NTA beads were washed with wash buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 20% (v/v) glycerol, 0.05% DDM, 30 mM imidazole) for a total of 6x column bed volume. Purification was completed using the elution buffer (20 mM Tris-HCl, pH 8.0, 100 mM

NaCl, 20% (v/v) glycerol, 0.05% DDM, 300 mM imidazole, 1 Roche protease tablet/30 mL of buffer). The resulting purified protein was then concentrated and buffer exchanged into reconstitution buffer (50 mM potassium phosphate, pH 7.0, 100 mM NaCl, 0.02% DDM). The final protein concentrations ranged from 2.8 to 4.5 mg/mL for the WT and the mutants, but consistently had to be above ~3 mg/mL to allow for a successful reconstitution into liposomes.

4.2.3 Reconstitution into lipids

The resulting concentrated protein samples suspended in detergent were reconstituted with a mixture of PC and PS lipids (9:1 w/w). To reconstitute, protein and lipids were mixed at a ratio of 2.33:1.16 mg of protein:lipid. This suspension was stirred at 4˚C for 16 hours for the proteoliposomes to form. After this, bio-beads at a concentration of 0.8 g/mL of suspension were added in three equal but separate additions, to prevent the detergent withdrawal from being too quick. Before each addition, the bio-beads were "pre-wetted" with reconstitution buffer lacking any DDM, to further prevent any sticking of the protein to the bio-beads. These additions were then performed over the course of ~9 hours, with one addition every 3 hours, and the suspension was nutated for a total of 24 hrs at 4˚C. The following day, an additional 0.2 g/mL of bio-beads was added to the proteoliposome suspension to facilitate the total removal of detergent. After another 24 hr incubation at 4˚C while nutating, the proteoliposome suspension was collected with both 18- and 22-gauge needles, to prevent any collection of bio-beads. The collected proteoliposomes were then ultracentrifuged at 300,000xg for 20 mins. After a wash with mqH_2O, the proteoliposomes were ultracentrifuged again at 150,000xg for 30 mins. The resulting pellets were dried and then stored at -20˚C prior to structural characterization.

4.2.4 Western blot

Western blots were used to confirm the expression of the different mutants following transformation of the cells. 100 mL small-scale cultures were broken as described above, and the Western blots were performed as described in *Chapter 3*.

4.2.5 ATR-FTIR spectroscopy and H/D exchange

FTIR measurements of AQP1 in proteoliposomes were conducted on a temperature-controlled Germanium Attenuated Total Reflectance (ATR) accessory (Pike Technologies, Madison WI) installed in a Vertex 70 FTIR spectrometer (Bruker, Milton ON). One hundred spectra were averaged at 4 cm^{-1} resolution, using a DTGS detector and transmission spectra of the empty ATR cell as reference. The control spectrum of WT-AQP1 was recorded at 25˚C by drying 15 µL (100 – 150 µg of protein) of the proteoliposomes, resuspended in buffer (20 mM Bis-tris propane, pH 7.0, 10 mM NaCl) on the surface of the germanium crystal over a flow of dry nitrogen gas. After drying, the hydrogen/deuterium exchange at room temperature was accomplished by aliquoting 2 µL of D_2O directly on the sample, which was left to incubate for 10 mins, before drying again with N_2 gas. The extent of H/D exchange at 25˚C was then recorded once the sample was fully dry. For each temperature point thereafter, another 2 µL of D_2O was aliquoted on the crystal before increasing the temperature to the desired point. Once the temperature equilibrated, the sample was incubated for 2 mins, then cooled back down to 25˚C and dried again with N_2 gas. Then another spectrum was taken to observe any secondary structure changes and the extent of the H/D exchange on the Amide II band, and the appearance of the Amide II' band.

The resulting spectra of the WT and all mutants were baseline corrected via OPUS (Bruker, Milton, Ontario, Canada), and normalized by the amplitude of amide I from zero using the Quasar spectroscopy package [258]. Spectra were then plotted using R with the following packages: ggplot2, ggtext, ggpubr and ggsci.

4.2.6 Differential Scanning Calorimetry

Differential scanning calorimetry was used to measure the T_M and the calorimetric enthalpy of unfolding for the WT and the N127A mutant. DSC was performed on a TA Nano DSC 602000.901 with 1 and 2 mg/mL protein concentrations, and in both H_2O and D_2O based buffers. The sample cell was filled with the protein pellet resuspended in buffer, at concentrations of either 1 or 2 mg/mL, while the reference cell contained only the related buffer. Both cells were pressurized to 3 atm, and heating of the sample and reference cells was achieved at a rate of 1˚C/min, while cooling was at a rate of 2˚C/min. The total temperature range sampled was 20˚C – 90˚C.

4.3 Results and Discussion

4.3.1 Expression and mutant yields

The WT and mutants were all expressed in 1 L *P. pastoris* cultures and solubilized using DDM as the detergent. Initially, OG was the chosen detergent, however upon noticing that the solubilized N127A mutant was aggregating in OG-micelles, the detergent was switched to DDM and kept consistent for the rest of the mutants and the wild-type. This instability within OG-micelles was the first hint towards the N127 mutant being marginally stable, as OG is reported to be a slightly more harsh detergent than DDM [259,260].

To confirm the protein expression of *P. pastoris* cells, Western blots were performed on each of the mutants using small scale cultures. Indeed, all of the mutants expressed well but to varying degrees [**Figure 4-6**]. The varying level of expression was further confirmed through the yields of purified proteins, as determined through UV-Vis absorbance at the common wavelength of 280 nm for protein quantitation [**Table 4.1**].

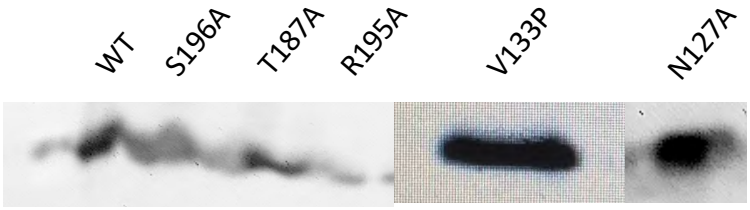

Figure 4-6: *Western blot analysis confirming expression of all mutants. Blots confirmed the presence of AQP1-containing His-tag. Mutants were run on different blots so were concatenated here. All bands ran at the expected size of ~28 kDa, to varying degrees of intensity.*

Table 4.1: *The WT and mutant yields per L of culture after purifying and before reconstitution into lipids.*

AQP1	***Yield per L of culture***
WT-AQP1	11.9 mg
N127A	9.9 mg
V133P	3.77 mg
E142A	5.12 mg
T187A	2.33 mg
R195A	1.4 mg
S196A	17.5 mg

4.3.2 ATR-FTIR H/DX temperature series

4.3.2.1 WT-AQP1

To observe the secondary structure changes and the exposure of backbone amides to solvent upon increasing temperature, ATR-FTIR was used to monitor the changes occurring. Using hydrogen/deuterium exchange, the amount of solvent exposed backbone amides can be measured as a function of temperature [214,215,218]. This is done by observing the changes in the amide II and II' bands and can be used to measure the stability of both the WT and the mutants [215, 218].

Firstly, the WT was measured at increasing temperature points to its eventual melting temperature (25, 40, 50, 55, 60, and 65˚C) which through prior experiments was determined to be around 62˚C (T_m). This was done in strictly H_2O first to observe whether any changes to the amide I and II bands would occur along the way to reaching the T_m. Interestingly, no such changes occurred in either of the bands. The amide I band, characteristically sensitive to secondary structure elements such as α-helices, β-sheets, turns, and loops, was unchanged for the WT [**Figure 4-7**]. This would suggest that although the WT does indeed unfold, according to the previous H/D exchange data collected by ssNMR, it is still mostly helical in this state. As well, a DSC thermogram of the WT in an H_2O-based buffer further confirmed that the T_m occurs at 62˚C, and is essentially unfolded, but not entirely unstructured, at 65˚C in the membranes [**Figure 4-8**]. This is consistent with some models of membrane protein unfolding, which describe the unfolded state in the membrane as still mostly helical even though most native tertiary structure is lost [261,262]. In the H/DX spectrum [**Figure 4-7**], the small shift in the peak of the amide I band is most likely due to the deuteration of the backbone amide N–H, which

results in a small shift that increases with temperature as more N–H groups are exchanged to N–D. At the final recorded temperature of 65˚C, the deuterated amide I occurs at 1651 cm^{-1}, resulting in a ~5 cm^{-1} wavenumber shift, corresponding to likely most of the backbone amides in the helices fully exchanged at this temperature [263].

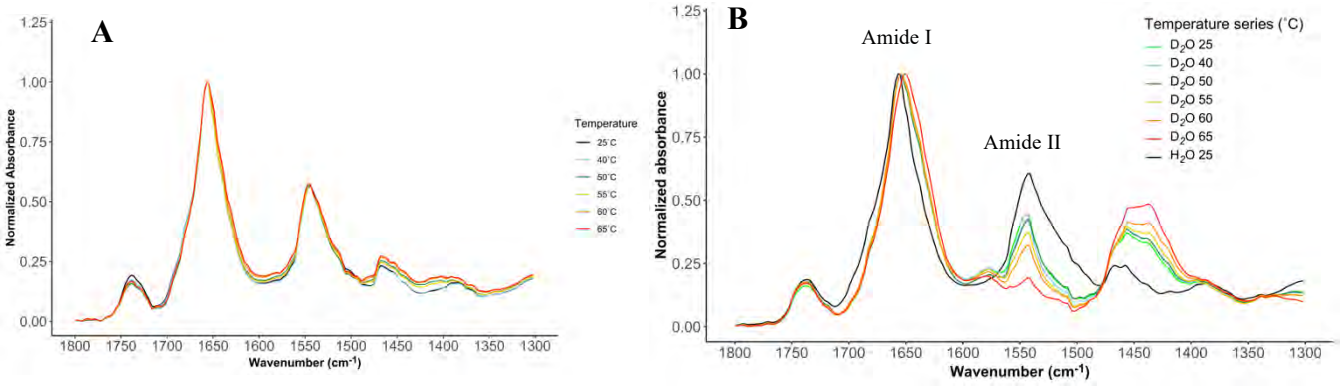

Figure 4-7: (A) *Overlaid normalized temperature series spectra of WT-AQP1 in H_2O and* ***(B)*** *normalized H/D exchange spectra of WT-AQP1. The WT remains completely helical after unfolding even at 65˚C, as observed in the H_2O series. The minor amide I shift is observed in the H/D exchange spectrum, as a result of the backbone deuteration evident from the Amide II to Amide II' transition, as the shift progressively increases as temperature increases.*

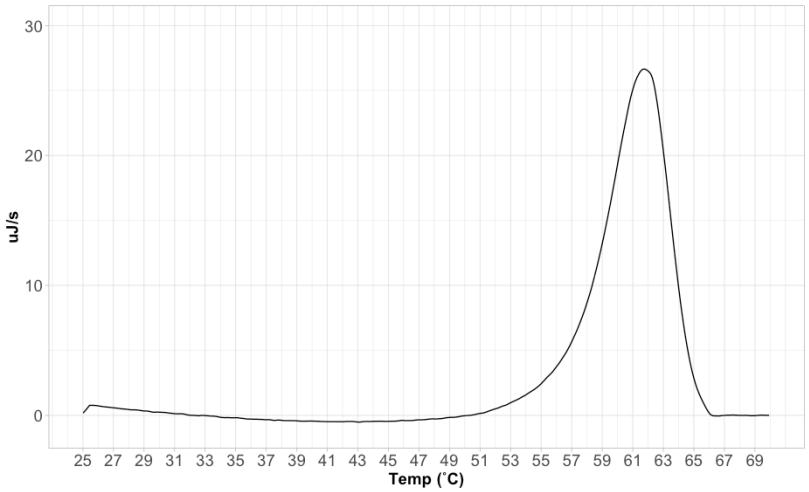

Figure 4-8: *DSC thermogram of 1 mg of WT-AQP1 in 20 mM BTP, 10 mM NaCl, pH 7.0. Heating was done at a rate 1˚C/min and cooling at 2˚C/min. A primary peak corresponding to the cooperative unfolding of WT-AQP1 has a* T_m *of 62˚C.*

4.3.2.2 N127A and V133P

Next, the mutants N127A and V133P were assayed using the same H/D exchange experimental setup. As described for previous attempts to solubilize N127A, it was evident that this mutant was indeed less stable. The FTIR spectroscopy measurements further confirmed this [**Figure 4-9**]. Immediately it was observed that the spectrum dried from H_2O revealed a much broader amide I band, as compared to the WT. As well, a broad shoulder on the amide I centered around ~1630 cm^{-1} would indicate other secondary structure contributions towards the total intensity of the amide I band. This overall shape of the amide I continued to broaden and change shape as the temperature increased to 65˚C. As well, upon the first addition of D_2O to the thin protein film, the amide II band almost completely disappears. It further decreases at 40˚C but exhibits very few changes past this temperature.

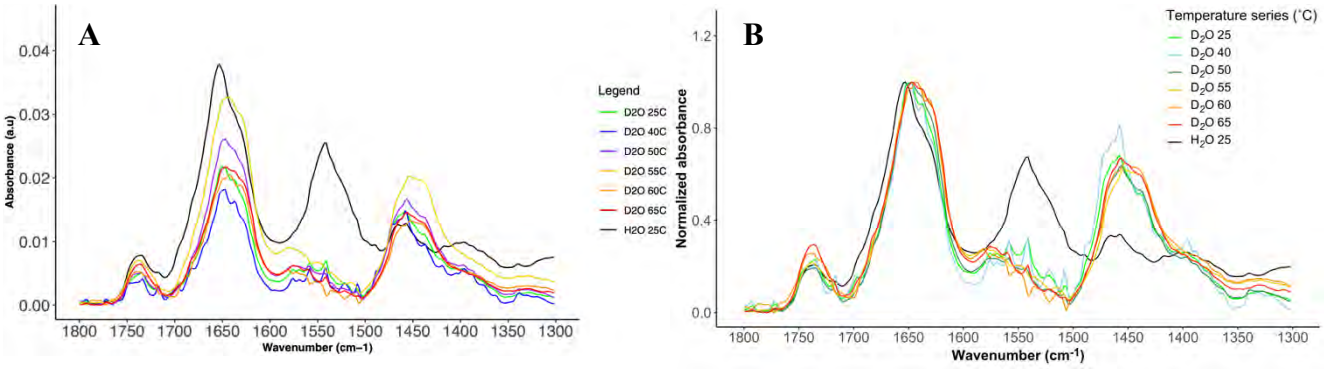

Figure 4-9: *FTIR spectrum of the H/D exchange of N127A,* ***(A)*** *raw data, and* ***(B)*** *data normalized to the amide I band.*

The proportional appearance of the amide II' band at room temperature confirmed that most of the backbone amides had been exchanged at this first temperature point. Past this temperature, no drastic changes occurred in amplitude as it is likely all backbone amide N–H groups were completely exchanged. Interestingly, although the amplitude of the amide II' does not further increase, the band shape itself becomes broader. This may provide further proof that secondary structure changes occur upon increasing temperature, as this amide II' band is also sensitive to secondary structure, although less sensitive than the amide I band.

The data from the H/DX experiments were to be expected. Since N127 is positioned downwards towards the pore in several high-resolution structures, and is closest in proximity with S196 to form a hydrogen-bond [250,253,256]. As well, prior ssNMR and H/DX experiments confirming the non-exchangeability of the N127 sidechain further highlight how buried this residue likely is, and how likely it is participating in hydrogen-bonding. Aside from S196, N127 could also hydrogen-bond with neighbouring residues such as R195, or the backbones of G188 and C189.

For the V133P mutant, a similar profile was observed, however with more drastic changes [**Figure 4-10**]. The H/D exchange was also quite extensive even just at room temperature, and continued to decrease upon increasing temperature, however only marginally after the first exchange at 25˚C. As well, the amide I band also appeared to be much broader than the WT, further indicating this mutant contained some other secondary structures that the WT did not. Interestingly, the lipid ester bands (~1740 cm^{-1}) seemed to increase as a function of temperature as well. This increase in lipid intensity seems to mirror the decrease in protein intensity as well. These changes in intensity could be explained by macro-scale changes in the thin film, in a way that the lipid starts to preferentially bind to the crystal, meanwhile the protein begins to crash out of the 2D lipid membrane. This phenomenon was even observed in some preparations of the N127A mutant, where the lipid bands started to increase upon temperature increases. It is difficult to identify this anomaly precisely, but if the V133P and N127A mutations impacts oligomerization in some manner, then re-organization of this thin film lattice may occur, affecting the ATR measurement.

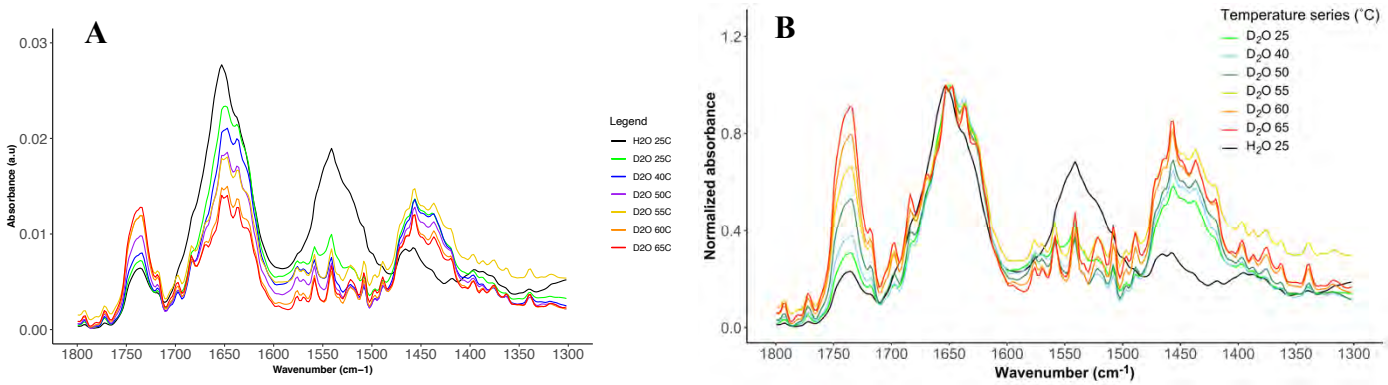

Figure 4-10: *FTIR spectrum of the H/D exchange of V133P,* ***(A)*** *raw data, and* ***(B)*** *data normalized to the amide I band.*

4.3.2.3 R195A and S196A

Rather interestingly, with the S196 sidechain hypothesized to hydrogen-bond with N127, the S196A mutant is quite stable in comparison to both N127A and the other mutants [**Figure 4-11**]. S196A follows a similar exchange pattern to that of the WT, where incubation of D_2O at 25˚C does not impart a large amount of exchange on its residues. Only the original number of exposed residues, such as loops and the C-terminal tail seem to be exchanged, comparable to the WT. Further exchange does occur as temperature increases, and full exchange does seem to happen earlier in the S196A mutant when compared to the WT – at 60˚C instead of 65˚C. These differences may not be large, but it could be predicted that the S196 sidechain is indeed important for stability, as the melting temperature does seem to decrease with the S196A mutant. However, the protein is still relatively stable, as no immediately obvious secondary structure changes are observed from the spectrum either, at both lower and higher temperatures. When compared to the WT H_2O spectrum, it would be reasonable to assume that S196A also remains helical at higher temperatures as well, which was a stark contrast from the N127A and V133P mutants.

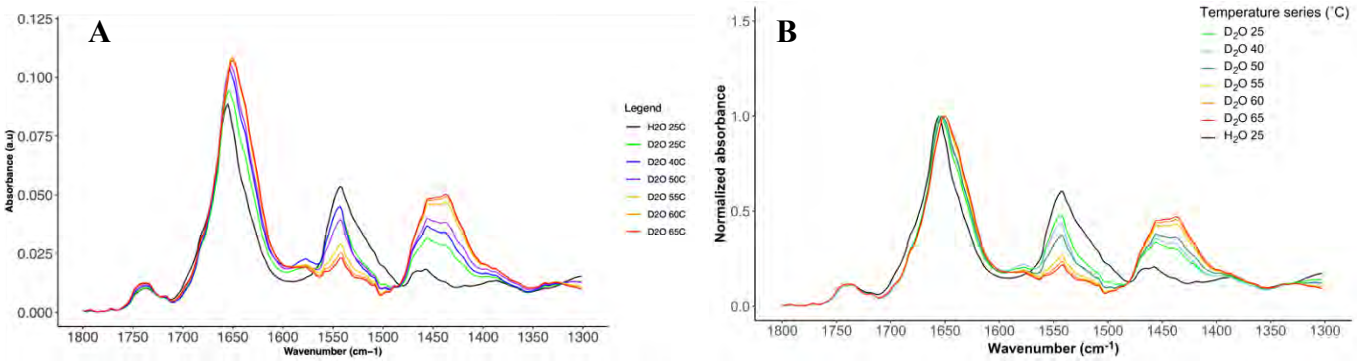

Figure 4-11: *FTIR spectrum of the H/D exchange of S196A,* ***(A)*** *raw data, and* ***(B)*** *data normalized to the amide I band.*

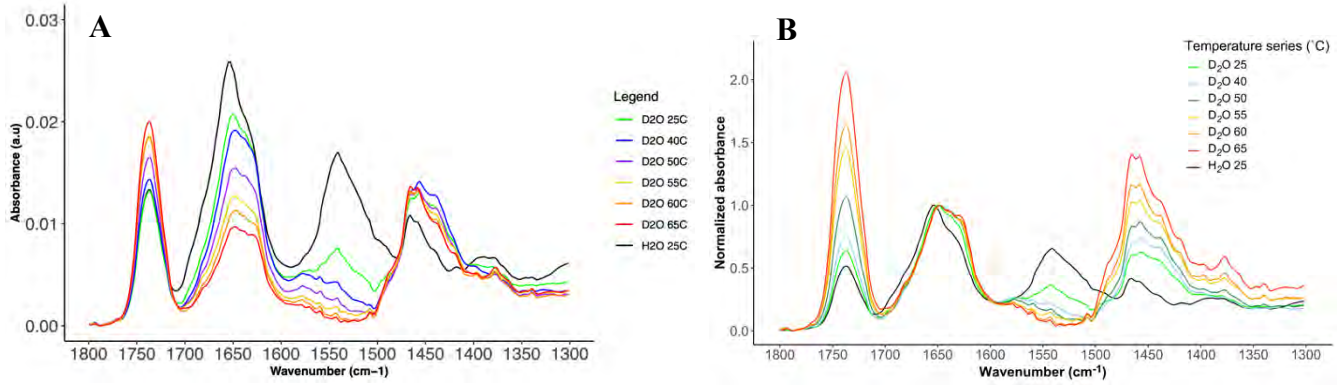

Figure 4-12: *FTIR spectrum of the H/D exchange of R195A,* ***(A)*** *raw data, and* ***(B)*** *data normalized to the amide I band.*

The H/DX experiments on the R195A mutant showed that it was also indeed quite unstable [**Figure 4-12**]. With prior structural and conservation analysis, this mutant was hypothesized to be unstable as the R195 sidechain potentially forms a hydrogen-bond with the backbone of Gly125, providing a further stabilizing interaction for loop C, keeping it in place down towards the pore. The R195 sidechain may also interact with N127 as well, offering

another crucial stabilizing point for loop C [**Figure 4-4**]. The R195A mutant exchanged to a large extent after just the 25˚C D_2O incubation. The exchange continued further upon increasing temperatures, but the rate of change was marginal past 50˚C or so. The most noticeable changes in the spectra occurred in the amide I region, where initially it is a sharp peak comprising mainly of α-helical contributions, but becomes much broader and the shoulder at ~1535 cm^{-1} intensifies. As well, the entire amide I signal itself decreases, whether due to the simultaneous disappearance of α-helical and appearance of β-sheet structure, or due to an entirely different mechanism is unclear. As seen before with the V133P mutant as well, the lipid peaks increase while the amide I decreases along the temperature series. If this phenomenon is real, and not an artifact from the instrument itself, it could also indicate that R195 has a role in oligomerization, or at least its stabilizing interactions towards loop C may have a role. Although this specific mutant has not been explored before, an R195V replacement has previously shown that this mutation allows proton conductance through the pore [251]. As well, a prior R195C mutation also demonstrated a decreased water permeability as compared to the WT [264]. Since R195 is a part of the ar/R selectivity filter, this mutation could have implications towards water transport, as the bulky arginine sidechain is replaced with the small non-polar side chain of alanine.

4.3.2.4 E142A and T187A

Lastly, the mutants E142A and T187A were explored using the same H/D exchange experimental setup. The sidechains of E142 and T187 likely form hydrogen-bonds, and the sidechain of E142 has further potential interactions with the backbone amides of G190 and I191. G190 and I191 are the residues that make up the preceding loop region before the NPA filter, so

these interactions may stabilize this entire region, and the E142 residue may especially have indirect interactions with the 2nd NPA motif in AQP1. The H/DX experiments confirmed that E142A is indeed less stable than the WT, as it exchanges almost completely after 40˚C [**Figure 4-13**]. Surprisingly, the exchange at 25˚C is not incredibly dramatic, but the full exchange at 40˚C indicates that it is likely mostly unfolded at this point and thus all the residues are solvent exposed. The E142A mutant's secondary structure, as observed with the amide I band, does not remarkably change across the temperature series. The amide I band shape does indeed shift and become broader, but this might be a combination of deuteration effects with whatever small secondary structure changes are occurring [265].

Lastly, the T187A mutant temperature series spectra also indicated large changes in its structure [**Figure 4-14**]. It exchanged almost completely at 25˚C, with smaller decreases of amide II following this temperature. The amide II exchangeability pattern closely resembled to that of the V133P, R195A, and N127A spectra [**Figure 4-15**]. The amide I band also noticeably becomes broader upon increasing temperature, revealing a sharper shoulder at around 1635 cm^{-1}, which is usually the frequency in which the β-sheet conformation absorbs. This is comparable to some of the other mutants, namely the most unstable, with N127A, V133P, R195A, and E142A all showing the appearance of an extra peak or shoulder on the amide I at around 1635 cm^{-1} [**Figure 4-16**]. This would indicate that for these mutants, some part of the structure unfolds and potentially refolds into a β-sheet conformation. Or perhaps that when these helices disassemble, they reassemble into a lower energy β-sheet intermediate structure. Since the WT remains helical even past its transition temperature, this would indicate a loss in helical conformation and stability in these mutants. It is unknown whether these new β-structures exist still inside the proteoliposome membrane or if it only interacts with one face, like that of peripheral membrane

proteins. It is also possible that at these higher temperatures where the β-sheets appear, the protein has aggregated and perhaps crashed out of the membrane. This could explain the phenomenon present in some of the FTIR spectra where the protein amide I signal shrinks in intensity while the lipid ester bands increase.

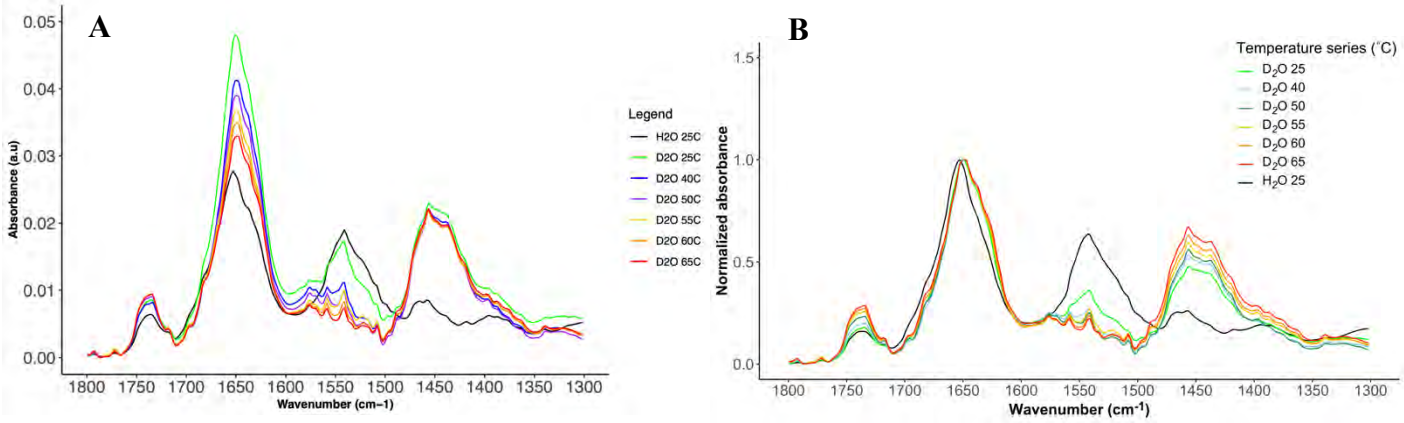

Figure 4-13: *Hydrogen-deuterium exchange experiments with the E142A mutant, with (**A**) raw data and **(B)** normalized spectra. Immediate exchange is observed at room temperature, and full deuterium exchange of all backbone protons occurs at 40˚C and above. As well, secondary structure changes are observed as the temperature increases.*

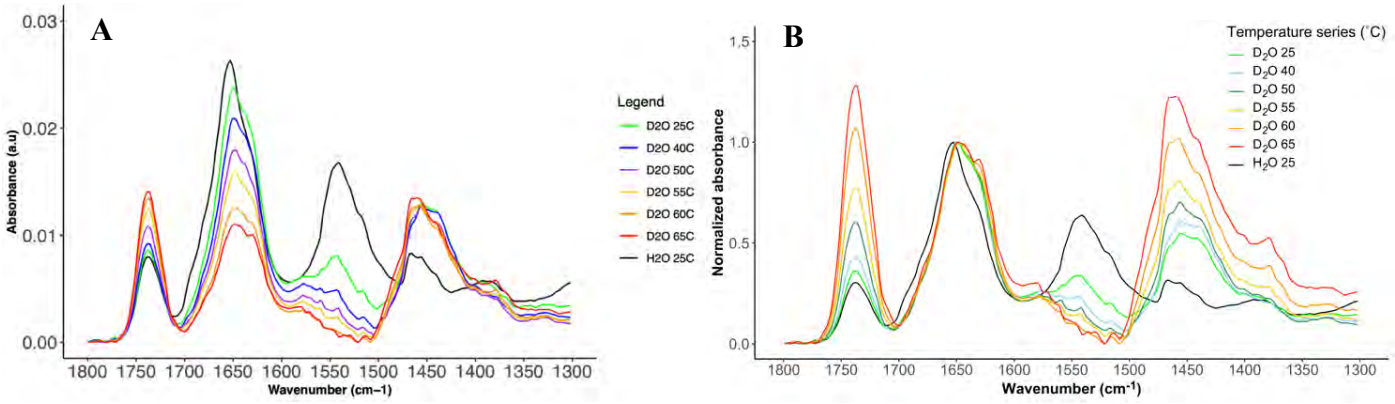

Figure 4-14: *FTIR spectrum of the H/D exchange of T187A,* ***(A)*** *raw data, and* ***(B)*** *data normalized to the amide I band.*

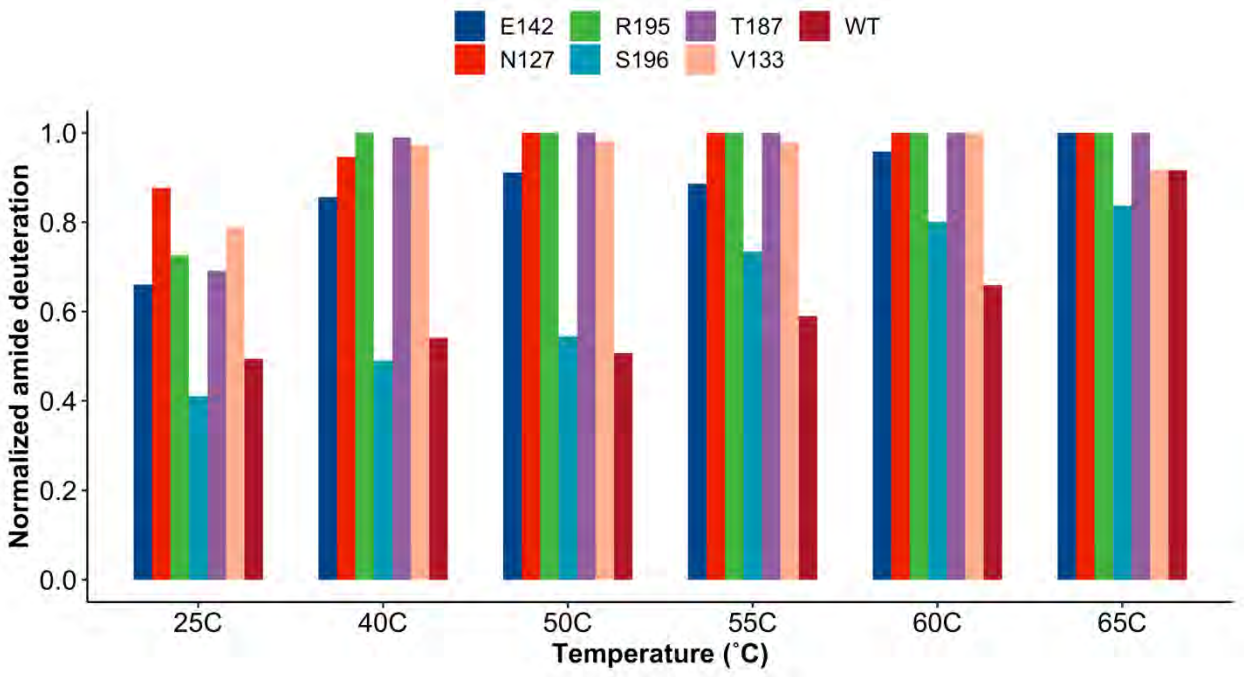

Figure 4-15: *The relative extent of amide II deuteration across all of the mutants and the WT. Extent of deuteration was determined by plotting the ratio of the amide II' maximum peak amplitude over the total amplitude contribution of both amide II and II' peaks.*

The evidence of secondary structural changes of the mutants when compared to the WT is also apparent when looking at just the relative amide I band maximum position [**Table 4.2**]. For the WT, the frequency shift due to deuteration is only ~2 cm^{-1} at 25˚C. At 65˚C, this shift amounts to just 5 cm^{-1}. The only other mutant that is comparable is the S196A mutant, and this mutant was also the most protected against amide II exchange [**Figure 4-15]**. The most unstable mutants show the largest shifts in the amide I position, further indicating that the shifts observed are more than just the effect of deuteration [**Figure 4-16**]. It is also interesting that the N127A mutant

experiences its entire amide I shift at 25˚C, as there is no further shift past this temperature all the way up to 65˚C.

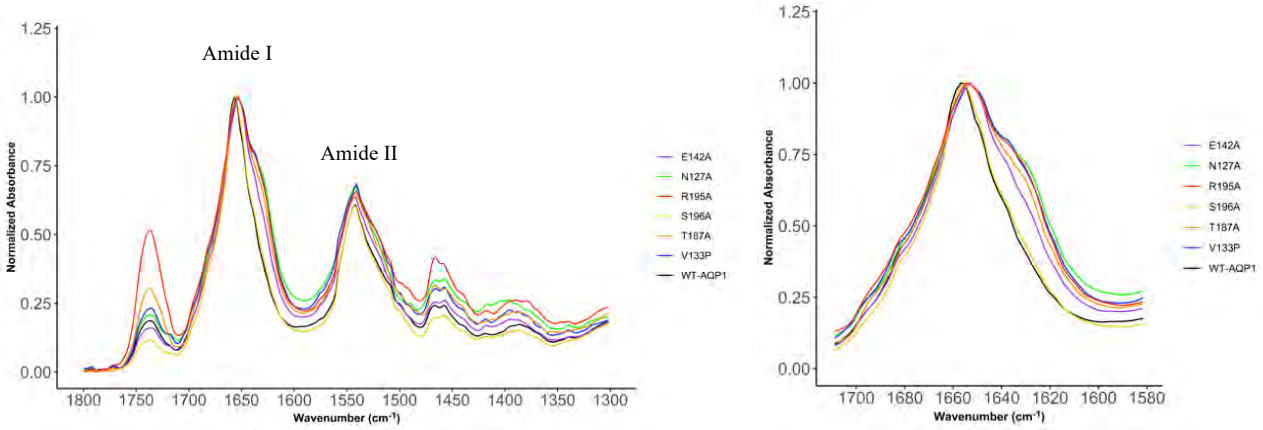

Figure 4-16: *An overlay of the mutants and the WT, dried from H_2O and measured at 25˚C, showing the entire spectrum of lipid and protein bands (left), and an overlay of just the amide I region (right). The overlay of the amide I band shows the stark differences between the WT and the rest of the mutants, with most exhibiting a shift in the amide I and subsequent broadening of the shoulder. Interestingly, S196A overlays well with the WT.*

Table 4.2: *Summary of the results from the H/D exchange FTIR experiments, indicating the amide I band peak positions at 25˚C for both H_2O and D_2O measurements, and 65˚C. Bolded value is the typical amide I band wavenumber for WT-hAQP1, comprised mainly of the α-helical component. Italicized values are the lowest amide I peak positions recorded for all the mutants.*

hAQP1 sample	***Amide I positions at 25˚C, H_2O (wavenumber cm^{-1})***	***Amide I positions at 25˚C, D_2O (wavenumber cm^{-1})***	***Amide I positions at 65˚C, D_2O (wavenumber cm^{-1})***
WT	**1656**	1654	1651
N127A	1653	1647	*1647*
V133P	1653	1649	*1647*
S196A	1655	1655	1651
R195A	1653	1651	*1647*
E142A	1653	1651	1649
T187A	1653	1649	1649

4.3.3 Differential scanning calorimetry

Lastly, differential scanning calorimetry was performed on the WT and the N127A mutant. DSC was performed on the WT in an H_2O buffer, to confirm prior transition temperature measurements [**Figure 4-8]**. The transition temperature found (62˚C) was consistent with prior DSC measurements made by another student in our group. Although not shown, the WT was also run in D_2O, however resulted in both a lower transition temperature and lower enthalpy of

unfolding, however this could be due to some instrumental error and needs to be investigated further.

The N127A mutant was also measured using DSC, with both 1 mg and 2 mg amounts in D_2O buffer [**Figure 4-17**]. This mutant displayed a very different thermogram, one in which the transition peak was much broader and stretched over a longer range of temperatures. In the trace collected from the 1 mg amount, the T_m was around 64˚C, while the trace from the 2 mg sample was even broader, with the transition temperature around 70˚C. This was strange, given that the higher concentration of protein should not yield a differing melting temperature, and would only impact the total calorimetric enthalpy. This discrepancy, and additional prior discrepancies given by a measurement of the WT-AQP1 in D_2O-based buffer which gave a similarly broad peak (results not shown), may seem to indicate that it is possible the DSC instrument used was not reliable, and would ultimately need to be repeated to confirm.

Nevertheless, if these results are indeed real and not a result of unreliable equipment, the prior FTIR exchange experiments that had shown N127A is indeed already unfolded at 25˚C given its large deuterium exchange, coupled with these DSC results may hint that N127A is indeed much less stable than the WT. The transition temperature may be similar, but it is hard to determine it precisely as the peak is so broad, and this broadness may be a result of uncooperative unfolding, where instead of the structure unfolding entirely in a narrow range of temperatures like the WT, N127A displays a positive specific heat capacity difference starting from 45˚C to as high as 85˚C.

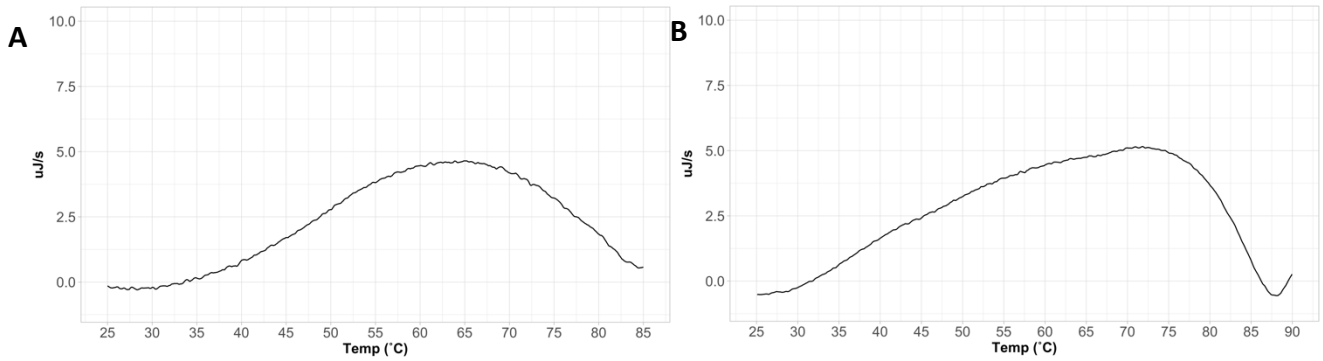

Figure 4-17: *DSC thermogram detailing the specific heat capacity differences of N127A, with 1 mg (A) and 2 mg (B) experiments. Both thermograms yielded a broad curve, with a transition temperature between 60 – 70˚C, however due to the broad shape, and the seemingly uncooperative nature of the unfolding pathway of the mutant, the exact* T_M *is uncertain.*

Prior studies have demonstrated this uncooperative nature and the resulting DSC trace, as the molten globule intermediate state has been explored, and also revealed a very broad peak due to its uncooperative unfolding and the absence of any tertiary structure [266,267,268,269]. This may be a similar situation with N127A, especially so if this mutant disrupts oligomerization. The monomers themselves would have even less cooperative unfolding than the WT tetramer. As well, the broadness of the peak could be attributed to the possibility that N127A exists as a mixture of differently folded and misfolded states. This combination of states may elicit a broad range in transitions, thus resulting in the very broad DSC trace for the mutant.

4.4 Conclusions and Future remarks

The ATR-FTIR spectroscopy coupled with hydrogen/deuterium exchange experiments confirmed the presence of a very intriguing network of amino acids that contribute to the overall

stability of the AQP1 fold. The importance of these hydrogen-bond networks was previously hypothesized based on prior solid-state NMR coupled with H/D exchange, as well as conservation and structural analysis. These prior experiments indicated that some of these residues explored may participate in hydrogen-bonding interactions that prevent them from being as easily exposed to solvent. As well, bioinformatics analysis confirmed these residue positions were conserved through other aquaporins, and namely other human aquaporins as well. In turn, the work outlined in this chapter has now established that these residue positions; N127, V133, E142, T187, R195 and S196 all contribute to the overall stability of the hAQP1 fold, albeit to varying degrees. The N127A, V133P, T187A, and R195A mutants are especially disrupting towards the overall fold. Not only did these mutants almost completely exchange at 25˚C, but the shapes of the amide I bands were heavily altered from the initial helical contribution as the incubation temperature increased.

This work outlines the first structural look into this hydrogen-bond network with the use of mutagenesis and ATR-FTIR experiments coupled with H/D exchange. The information provided from this work has resulted in even more questions relating to these amino acid positions. Although the exchangeability and overall stability was probed using FTIR, further analysis using DSC on the rest of the mutants could be performed to confirm that these mutants result in a lower transition temperature than the WT. After this, the question of how these amino acid substitutions affect the primary water function of the pore must be explored as well, whether these mutants may be more permeable to water than the WT, and whether they now allow other solutes to pass through the pore. These permeability assays could likely be performed via stopped-flow spectroscopy, or through an *in vitro* cellular water transport assay. Another interesting question to posit would be whether any of these mutants also allow proton

conductance. Finally, some of these mutants may be affecting oligomerization, and assessing whether these mutants are either tetramers like the WT or are now monomeric would provide crucial information into how these residues stabilize the overall structure.

Of course, these future experiments on the residue positions not only provide information into hAQP1, but into other aquaporins that also share these residues and motifs. The structural β-turn motif and the loop C hydrogen-bonding networks are incredibly conserved among other AQP1 homologs and in some human water-transporting AQP's, namely AQP2, 4, 5 and 7. AQP2 and 4 are especially relevant towards certain physiological processes and involved in diseases such as diabetes and several cancers [55,56,60,63]. Structural insights such as these in *Chapter 4* provide the necessary information towards developing eventual drugs that target aquaporins. Many aquaporins are highly expressed in certain cancers, and establishing which areas are most crucial towards AQP stability may provide a foundation for the eventual drug targeting of aquaporins.

References

[1] A. Engel, H.E. Gaub, Structure and Mechanics of Membrane Proteins, Annu. Rev. Biochem., 77 (2008) 127–148.

[2] A.S. Verkman, Water Channels in Cell Membranes, (n.d.).

[3] P.J.F. Henderson, Ion Transport by Energy-Conserving Biological Membranes, Annu. Rev. Microbiol., 25 (1971) 393–428.

[4] T. Harayama, H. Riezman, Understanding the diversity of membrane lipid composition, Nat. Rev. Mol. Cell Biol., 19 (2018) 281–296.

[5] S. Tan, H.T. Tan, M.C.M. Chung, Membrane proteins and membrane proteomics, PROTEOMICS, 8 (2008) 3924–3932.

[6] G. van Meer, D.R. Voelker, G.W. Feigenson, Membrane lipids: where they are and how they behave, Nat. Rev. Mol. Cell Biol., 9 (2008) 112–124.

[7] J.F. Nagle, S. Tristram-Nagle, Structure of lipid bilayers, Biochim. Biophys. Acta BBA - Rev. Biomembr., 1469 (2000) 159–195.

[8] F. Bezanilla, How membrane proteins sense voltage, Nat. Rev. Mol. Cell Biol., 9 (2008) 323–332.

[9] J.N. Israelachvili, Refinement of the fluid-mosaic model of membrane structure, Biochim. Biophys. Acta BBA - Biomembr., 469 (1977) 221–225.

[10] P. Klein, M. Kanehisa, C. DeLisi, The detection and classification of membrane-spanning proteins, Biochim. Biophys. Acta BBA - Biomembr., 815 (1985) 468–476.

[11] A.M. Seddon, P. Curnow, P.J. Booth, Membrane proteins, lipids and detergents: not just a soap opera, Biochim. Biophys. Acta BBA - Biomembr., 1666 (2004) 105–117.

[12] M. le Maire, P. Champeil, J.V. Møller, Interaction of membrane proteins and lipids with solubilizing detergents, Biochim. Biophys. Acta BBA - Biomembr., 1508 (2000) 86–111.

[13] A.L. Lomize, I.D. Pogozheva, M.A. Lomize, H.I. Mosberg, The role of hydrophobic interactions in positioning of peripheral proteins in membranes, BMC Struct. Biol., 7 (2007) 44.

[14] G.G. Privé, Detergents for the stabilization and crystallization of membrane proteins, Methods, 41 (2007) 388–397.

[15] K.N. Allen, S. Entova, L.C. Ray, B. Imperiali, Monotopic Membrane Proteins Join the Fold, Trends Biochem. Sci., 44 (2019) 7–20.

[16] K. Balali-Mood, P.J. Bond, M.S.P. Sansom, Interaction of Monotopic Membrane Enzymes with a Lipid Bilayer: A Coarse-Grained MD Simulation Study, Biochemistry, 48 (2009) 2135–2145.

[17] F. Forneris, A. Mattevi, Enzymes Without Borders: Mobilizing Substrates, Delivering Products, Science, 321 (2008) 213–216.

[18] H. Chiba, M. Osanai, M. Murata, T. Kojima, N. Sawada, Transmembrane proteins of tight junctions, Biochim. Biophys. Acta BBA - Biomembr., 1778 (2008) 588–600.

[19] T. Stora, J.H. Lakey, H. Vogel, Ion-Channel Gating in Transmembrane Receptor Proteins: Functional Activity in Tethered Lipid Membranes, Angew. Chem. Int. Ed., 38 (1999) 389–392.

[20] G.E. Tusnady, Z. Dosztanyi, I. Simon, Transmembrane proteins in the Protein Data Bank: identification and classification, Bioinformatics, 20 (2004) 2964–2972.
[21] D. Nietlispach, A. Gautier, Solution NMR studies of polytopic α-helical membrane proteins, Curr. Opin. Struct. Biol., 21 (2011) 497–508.
[22] M.S. Almén, K.J. Nordström, R. Fredriksson, H.B. Schiöth, Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin, BMC Biol., 7 (2009) 50.
[23] O.P. Ernst, D.T. Lodowski, M. Elstner, P. Hegemann, L.S. Brown, H. Kandori, Microbial and Animal Rhodopsins: Structures, Functions, and Molecular Mechanisms, Chem. Rev., 114 (2014) 126–163.
[24] H. Kandori, Ion-pumping microbial rhodopsins, Front. Mol. Biosci., 2 (2015).
[25] J.L. Spudich, The multitalented microbial sensory rhodopsins, Trends Microbiol., 14 (2006) 480–487.
[26] X.E. Zhou, K. Melcher, H.E. Xu, Structure and activation of rhodopsin, Acta Pharmacol. Sin., 33 (2012) 291–299.
[27] R. Fredriksson, M.C. Lagerström, L.-G. Lundin, H.B. Schiöth, The G-Protein-Coupled Receptors in the Human Genome Form Five Main Families Phylogenetic Analysis, Paralogon Groups, and Fingerprints, Mol. Pharmacol., 63 (2003) 1256–1272.
[28] T. Gonen, T. Walz, The structure of aquaporins, Q. Rev. Biophys., 39 (2006) 361–396.
[29] A.S. Verkman, Aquaporins, Curr. Biol., 23 (2013) R52–R55.
[30] E. Beitz, P. Agre, eds., Aquaporins, Springer, Berlin, 2009.
[31] M. Hara-Chikuma, A.S. Verkman, Physiological roles of glycerol-transporting aquaporins: the aquaglyceroporins, Cell. Mol. Life Sci., 63 (2006) 1386–1392.
[32] M.C. Papadopoulos, S. Saadoun, A.S. Verkman, Aquaporins and cell migration, Pflüg. Arch. - Eur. J. Physiol., 456 (2008) 693–700.
[33] M.A. Knepper, J.B. Wade, J. Terris, C.A. Ecelbarger, D. Marples, B. Mandon, C.-L. Chou, B.K. Kishore, S. Nielsen, Renal aquaporins, Kidney Int., 49 (1996) 1712–1717.
[34] D. Ribatti, G. Ranieri, T. Annese, B. Nico, Aquaporins in cancer, Biochim. Biophys. Acta BBA - Gen. Subj., 1840 (2014) 1550–1553.
[35] M.M. Salman, P. Kitchen, A.J. Yool, R.M. Bill, Recent breakthroughs and future directions in drugging aquaporins, Trends Pharmacol. Sci., 43 (2022) 30–42.
[36] H. Alkhalifa, F. Mohammed, S. Taurin, K. Greish, S. Taha, S. Fredericks, Inhibition of aquaporins as a potential adjunct to breast cancer cryotherapy, Oncol. Lett., 21 (2021) 458.
[37] Z. Shi, T. Zhang, L. Luo, H. Zhao, J. Cheng, J. Xiang, C. Zhao, Aquaporins in human breast cancer: Identification and involvement in carcinogenesis of breast cancer: Aquaporins Proteins and Breast Cancer, J. Surg. Oncol., 106 (2012) 267–272.
[38] P. Agre, G.M. Preston, B.L. Smith, J.S. Jung, S. Raina, C. Moon, W.B. Guggino, S. Nielsen, Aquaporin CHIP: the archetypal molecular water channel, Am. J. Physiol.-Ren. Physiol., 265 (1993) F463–F476.
[39] J.S. Jung, G.M. Preston, B.L. Smith, W.B. Guggino, P. Agre, Molecular structure of the water channel through aquaporin CHIP The hourglass model, J. Biol. Chem., 269 (1994) 14648–14654.
[40] S. Ahmed, S. Kouser, M. Asgher, S.G. Gandhi, Plant aquaporins: A frontward to make crop plants drought resistant, Physiol. Plant., 172 (2021) 1089–1105.

[41] J.P. Bezerra-Neto, F.C. de Araújo, J.R.C. Ferreira-Neto, M.D. da Silva, V. Pandolfi, F.F. Aburjaile, T. Sakamoto, R.L. de Oliveira Silva, E.A. Kido, L.L. Barbosa Amorim, J.M. Ortega, A.M. Benko-Iseppon, Plant Aquaporins: Diversity, Evolution and Biotechnological Applications, Curr. Protein Pept. Sci., 20 (2019) 368–395.
[42] S. Törnroth-Horsefield, Y. Wang, K. Hedfalk, U. Johanson, M. Karlsson, E. Tajkhorshid, R. Neutze, P. Kjellbom, Structural mechanism of plant aquaporin gating, Nature, 439 (2006) 688–694.
[43] S.D. Tyerman, C.M. Niemietz, H. Bramley, Plant aquaporins: multifunctional water and solute channels with expanding roles: Function and expression of plant aquaporins, Plant Cell Environ., 25 (2002) 173–194.
[44] J.B. Heymann, A. Engel, Aquaporins: Phylogeny, Structure, and Physiology of Water Channels, Physiology, 14 (1999) 187–193.
[45] Structure and Function of Aquaporins: the Membrane Water Channel Proteins, Biointerface Res. Appl. Chem., 12 (2021) 690–705.
[46] A.S. Verkman, Mammalian aquaporins: diverse physiological roles and potential clinical significance, Expert Rev. Mol. Med., 10 (2008) e13.
[47] D.F. Savage, P.F. Egea, Y. Robles-Colmenares, J.D.O. Iii, R.M. Stroud, Architecture and Selectivity in Aquaporins: 25 Å X-Ray Structure of Aquaporin Z, PLoS Biol., 1 (2003) e72.
[48] P. Agre, Aquaporin Water Channels (Nobel Lecture), Angew. Chem. Int. Ed., 43 (2004) 4278–4290.
[49] W.E.C. Harries, D. Akhavan, L.J.W. Miercke, S. Khademi, R.M. Stroud, The channel architecture of aquaporin 0 at a 22-Å resolution, Proc. Natl. Acad. Sci., 101 (2004) 14045–14050.
[50] T. Gonen, P. Sliz, J. Kistler, Y. Cheng, T. Walz, Aquaporin-0 membrane junctions reveal the structure of a closed water pore, 429 (2004).
[51] K.L. Németh-Cahalan, J.E. Hall, pH and Calcium Regulate the Water Permeability of Aquaporin 0, J. Biol. Chem., 275 (2000) 6777–6782.
[52] M. Ozu, J.J. Alvear-Arias, M. Fernandez, A. Caviglia, A. Peña-Pichicoi, C. Carrillo, E. Carmona, A. Otero-Gonzalez, J.A. Garate, G. Amodeo, C. Gonzalez, Aquaporin Gating: A New Twist to Unravel Permeation through Water Channels, Int. J. Mol. Sci., 23 (2022) 12317.
[53] E.T.B. Olesen, R.A. Fenton, Aquaporin-2 membrane targeting: still a conundrum, Am. J. Physiol.-Ren. Physiol., 312 (2017) F744–F747.
[54] M. Judith Radin, M.-J. Yu, L. Stoedkilde, R. Lance Miller, J.D. Hoffert, J. Frokiaer, T. Pisitkun, M.A. Knepper, Aquaporin-2 regulation in health and disease, Vet. Clin. Pathol., 41 (2012) 455–470.
[55] L. Bai, K. Fushimi, S. Sasaki, F. Marumo, Structure of Aquaporin-2 Vasopressin Water Channel, J. Biol. Chem., 271 (1996) 5171–5176.
[56] A.J.M. Loonen, N.V.A.M. Knoers, C.H. van Os, P.M.T. Deen, Aquaporin 2 Mutations in Nephrogenic Diabetes Insipidus, Semin. Nephrol., 28 (2008) 252–265.
[57] J.-P. Morello, D.G. Bichet, Nephrogenic Diabetes Insipidus, (2001).
[58] B.W.M. van Balkom, P.J.M. Savelkoul, D. Markovich, E. Hofman, S. Nielsen, P. van der Sluijs, P.M.T. Deen, The Role of Putative Phosphorylation Sites in the Targeting and Shuttling of the Aquaporin-2 Water Channel, J. Biol. Chem., 277 (2002) 41473–41479.

[59] M.C. Papadopoulos, A.S. Verkman, Aquaporin-4 and brain edema, Pediatr. Nephrol., 22 (2007) 778–784.
[60] S. Saadoun, Aquaporin-4 expression is increased in oedematous human brain tumours, J. Neurol. Neurosurg. Psychiatry, 72 (2002) 262–265.
[61] J.D. Neely, B.M. Christensen, S. Nielsen, P. Agre, Heterotetrameric Composition of Aquaporin-4 Water Channels, Biochemistry, 38 (1999) 11156–11163.
[62] Y. Hiroaki, K. Tani, A. Kamegawa, N. Gyobu, K. Nishikawa, H. Suzuki, T. Walz, S. Sasaki, K. Mitsuoka, K. Kimura, A. Mizoguchi, Y. Fujiyoshi, Implications of the Aquaporin-4 Structure on Array Formation and Cell Adhesion, J. Mol. Biol., 355 (2006) 628–639.
[63] A.S. Verkman, A.J. Smith, P. Phuan, L. Tradtrantip, M.O. Anderson, The aquaporin-4 water channel as a potential drug target in neurological disorders, Expert Opin. Ther. Targets, 21 (2017) 1161–1170.
[64] S.K. Kang, Y.K. Chae, J. Woo, M.S. Kim, J.C. Park, J. Lee, J.C. Soria, S.J. Jang, D. Sidransky, C. Moon, Role of Human Aquaporin 5 In Colorectal Carcinogenesis, Am. J. Pathol., 173 (2008) 518–525.
[65] I. Direito, A. Madeira, M.A. Brito, G. Soveral, Aquaporin-5: from structure to function and dysfunction in cancer, Cell. Mol. Life Sci., 73 (2016) 1623–1640.
[66] R. Horsefield, K. Nordén, M. Fellert, A. Backmark, S. Törnroth-Horsefield, A.C. Terwisscha van Scheltinga, J. Kvassman, P. Kjellbom, U. Johanson, R. Neutze, High-resolution x-ray structure of human aquaporin 5, Proc. Natl. Acad. Sci., 105 (2008) 13327–13332.
[67] L. Janosi, M. Ceccarelli, The Gating Mechanism of the Human Aquaporin 5 Revealed by Molecular Dynamics Simulations, PLOS ONE, 8 (2013) e59897.
[68] M. Ikeda, E. Beitz, D. Kozono, W.B. Guggino, P. Agre, M. Yasui, Characterization of Aquaporin-6 as a Nitrate Channel in Mammalian Cells, J. Biol. Chem., 277 (2002) 39873–39879.
[69] K. Liu, D. Kozono, Y. Kato, P. Agre, A. Hazama, M. Yasui, Conversion of aquaporin 6 from an anion channel to a water-selective channel by a single amino acid substitution, Proc. Natl. Acad. Sci., 102 (2005) 2192–2197.
[70] A. Warshel, Inverting the selectivity of aquaporin 6: Gating versus direct electrostatic interaction, Proc. Natl. Acad. Sci., 102 (2005) 1813–1814.
[71] M. Yasui, A. Hazama, T.-H. Kwon, S. Nielsen, Wm.B. Guggino, P. Agre, Rapid gating and anion permeability of an intracellular aquaporin, Nature, 402 (1999) 184–187.
[72] S. Ma, H. Xie, K. Yu, J. Yang, Mechanism of unusual AQP6 activation by mercury binding to a pore-external residue C155, Biochem. Biophys. Res. Commun., 618 (2022) 1–7.
[73] H. Xie, S. Ma, Y. Zhao, H. Zhou, Q. Tong, Y. Chen, Z. Zhang, K. Yu, Q. Lin, L. Kai, M. Liu, J. Yang, Molecular Mechanisms of Mercury-Sensitive Aquaporins, J. Am. Chem. Soc., 144 (2022) 22229–22241.
[74] S.M. Saparov, Fast and Selective Ammonia Transport by Aquaporin-8*, 282 (2007).
[75] L.R. Soria, E. Fanelli, N. Altamura, M. Svelto, R.A. Marinelli, G. Calamita, Aquaporin-8-facilitated mitochondrial ammonia transport, Biochem. Biophys. Res. Commun., (2010).
[76] K. Liu, H. Nagase, C.G. Huang, G. Calamita, P. Agre, Purification and functional characterization of aquaporin-8, Biol. Cell, 98 (2006) 153–161.

[77] S. Bestetti, I. Medraño-Fernandez, M. Galli, M. Ghitti, G.P. Bienert, G. Musco, A. Orsi, A. Rubartelli, R. Sitia, A persulfidation-based mechanism controls aquaporin-8 conductance, Sci. Adv., 4 (2018) eaar5770.
[78] D. Fu, A. Libson, L.J.W. Miercke, C. Weitzman, P. Nollert, J. Krucinski, R.M. Stroud, Structure of a Glycerol-Conducting Channel and the Basis for Its Selectivity, Science, 290 (2000) 481–486.
[79] M.Ø. Jensen, S. Park, E. Tajkhorshid, K. Schulten, Energetics of glycerol conduction through aquaglyceroporin GlpF, Proc. Natl. Acad. Sci., 99 (2002) 6731–6736.
[80] J. Hénin, E. Tajkhorshid, K. Schulten, C. Chipot, Diffusion of Glycerol through Escherichia coli Aquaglyceroporin GlpF, Biophys. J., 94 (2008) 832–839.
[81] R.M. Stroud, L.J. Miercke, J. O'Connell, S. Khademi, J.K. Lee, J. Remis, W. Harries, Y. Robles, D. Akhavan, Glycerol facilitator GlpF and the associated aquaporin family of channels, Curr. Opin. Struct. Biol., 13 (2003) 424–431.
[82] J. Badaut, Aquaglyceroporin 9 in brain pathologies, Neuroscience, 168 (2010) 1047–1057.
[83] K. Ishibashi, T. Morinaga, M. Kuwahara, S. Sasaki, M. Imai, Cloning and identification of a new member of water channel (AQP10) as an aquaglyceroporin, Biochim. Biophys. Acta BBA - Gene Struct. Expr., 1576 (2002) 335–340.
[84] A. Rodríguez, V. Catalán, J. Gómez-Ambrosi, G. Frühbeck, Aquaglyceroporins serve as metabolic gateways in adiposity and insulin resistance control, Cell Cycle, 10 (2011) 1548–1556.
[85] I.V. da Silva, J.S. Rodrigues, I. Rebelo, J.P.G. Miranda, G. Soveral, Revisiting the metabolic syndrome: the emerging role of aquaglyceroporins, Cell. Mol. Life Sci., 75 (2018) 1973–1988.
[86] U. Laforenza, C. Bottino, G. Gastaldi, Mammalian aquaglyceroporin function in metabolism, Biochim. Biophys. Acta BBA - Biomembr., 1858 (2016) 1–11.
[87] Z. Liu, J. Shen, J.M. Carbrey, R. Mukhopadhyay, P. Agre, B.P. Rosen, Arsenite transport by mammalian aquaglyceroporins AQP7 and AQP9, Proc. Natl. Acad. Sci., 99 (2002) 6053–6058.
[88] Z. Liu, J.M. Carbrey, P. Agre, B.P. Rosen, Arsenic trioxide uptake by human and rat aquaglyceroporins, Biochem. Biophys. Res. Commun., 316 (2004) 1178–1185.
[89] P. Kitchen, R.E. Day, M.M. Salman, M.T. Conner, R.M. Bill, A.C. Conner, Beyond water homeostasis: Diverse functional roles of mammalian aquaporins, Biochim. Biophys. Acta BBA - Gen. Subj., 1850 (2015) 2410–2421.
[90] T. Walz, T. Hirai, K. Murata, J.B. Heymann, K. Mitsuoka, Y. Fujiyoshi, B.L. Smith, P. Agre, A. Engel, The three-dimensional structure of aquaporin-1, Nature, 387 (1997) 624–627.
[91] J.D. Ho, R. Yeh, A. Sandstrom, I. Chorny, W.E.C. Harries, R.A. Robbins, L.J.W. Miercke, R.M. Stroud, Crystal structure of human aquaporin 4 at 18 Å and its mechanism of conductance, Proc. Natl. Acad. Sci., 106 (2009) 7437–7442.
[92] D. Kozono, M. Yasui, L.S. King, P. Agre, Aquaporin water channels: atomic structure molecular dynamics meet clinical medicine, J. Clin. Invest., 109 (2002) 1395–1399.
[93] K. Tani, T. Mitsuma, Y. Hiroaki, A. Kamegawa, K. Nishikawa, Y. Tanimura, Y. Fujiyoshi, Mechanism of Aquaporin-4's Fast and Highly Selective Water Conduction and Proton Exclusion, J. Mol. Biol., 389 (2009) 694–706.

[94] G. Benga, The first discovered water channel protein, later called aquaporin 1: Molecular characteristics, functions and medical implications, Mol. Aspects Med., 33 (2012) 518–534.
[95] G.V.R. Prasad, L.A. Coury, F. Finn, M.L. Zeidel, Reconstituted Aquaporin 1 Water Channels Transport CO2 across Membranes, J. Biol. Chem., 273 (1998) 33123–33126.
[96] J.S. Hub, B.L. de Groot, Does CO2 Permeate through Aquaporin-1?, Biophys. J., 91 (2006) 842–848.
[97] S. Prak, S. Hem, J. Boudet, G. Viennois, N. Sommerer, M. Rossignol, C. Maurel, V. Santoni, Multiple Phosphorylations in the C-terminal Tail of Plant Plasma Membrane Aquaporins, Mol. Cell. Proteomics, 7 (2008) 1019–1030.
[98] M. Kuwahara, T. Asai, Y. Terada, S. Sasaki, The C-terminal tail of aquaporin-2 determines apical trafficking, Kidney Int., 68 (2005) 1999–2009.
[99] K. Kalman, K.L. Németh-Cahalan, A. Froger, J.E. Hall, Phosphorylation Determines the Calmodulin-mediated Ca2+ Response and Water Permeability of AQP0, J. Biol. Chem., 283 (2008) 21278–21283.
[100] K.M. Lindsey Rose, Z. Wang, G.N. Magrath, E.S. Hazard, J.D. Hildebrandt, K.L. Schey, Aquaporin 0−Calmodulin Interaction and the Effect of Aquaporin 0 Phosphorylation, Biochemistry, 47 (2008) 339–347.
[101] K. Nemoto, T. Niinae, F. Goto, N. Sugiyama, A. Watanabe, M. Shimizu, K. Shiratake, M. Nishihara, Calcium-dependent protein kinase 16 phosphorylates and activates the aquaporin PIP2;2 to regulate reversible flower opening in *Gentiana scabra*, Plant Cell, 34 (2022) 2652–2670.
[102] I. Kadohira, Y. Abe, M. Nuriya, K. Sano, S. Tsuji, T. Arimitsu, Y. Yoshimura, M. Yasui, Phosphorylation in the C-terminal domain of Aquaporin-4 is required for Golgi transition in primary cultured astrocytes, Biochem. Biophys. Res. Commun., 377 (2008) 463–468.
[103] H.B. Moeller, J. Praetorius, M.R. Rützler, R.A. Fenton, Phosphorylation of aquaporin-2 regulates its endocytosis and protein–protein interactions, Proc. Natl. Acad. Sci., 107 (2010) 424–429.
[104] S. Lindskog, Structure and mechanism of carbonic anhydrase, Pharmacol. Ther., 74 (1997) 1–20.
[105] G. Vilas, D. Krishnan, S.K. Loganathan, D. Malhotra, L. Liu, M.R. Beggs, P. Gena, G. Calamita, M. Jung, R. Zimmermann, G. Tamma, J.R. Casey, R.T. Alexander, Increased water flux induced by an aquaporin-1/carbonic anhydrase II interaction, Mol. Biol. Cell, 26 (2015) 1106–1118.
[106] G. Ren, A. Cheng, P. Melnyk, A.K. Mitra, Polymorphism in the Packing of Aquaporin-1 Tetramers in 2-D Crystals, J. Struct. Biol., 130 (2000) 45–53.
[107] J.W. Vince, U. Carlsson, R.A.F. Reithmeier, Localization of the Cl^-/HCO_3^- Anion Exchanger Binding Site to the Amino-Terminal Region of Carbonic Anhydrase II, Biochemistry, 39 (2000) 13344–13349.
[108] J.W. Vince, R.A.F. Reithmeier, Identification of the Carbonic Anhydrase II Binding Site in the Cl^-/HCO_3^- Anion Exchanger AE1, Biochemistry, 39 (2000) 5527–5533.
[109] M. Ota, R. Koike, T. Amemiya, T. Tenno, P.R. Romero, H. Hiroaki, A.K. Dunker, S. Fukuchi, An assignment of intrinsically disordered regions of proteins based on NMR structures, J. Struct. Biol., 181 (2013) 29–36.

[110] S. Wang, C. Ing, S. Emami, Y. Jiang, H. Liang, R. Pomès, L.S. Brown, V. Ladizhansky, Structure and Dynamics of Extracellular Loops in Human Aquaporin-1 from Solid-State NMR and Molecular Dynamics, J. Phys. Chem. B, 120 (2016) 9887–9902.
[111] D.A. Dingwell, L.S. Brown, V. Ladizhansky, Structure of the Functionally Important Extracellular Loop C of Human Aquaporin 1 Obtained by Solid-State NMR under Nearly Physiological Conditions, J. Phys. Chem. B, 123 (2019) 7700–7710.
[112] D.F. Savage, R.M. Stroud, Structural Basis of Aquaporin Inhibition by Mercury, J. Mol. Biol., 368 (2007) 607–617.
[113] Y. Hirano, N. Okimoto, I. Kadohira, M. Suematsu, K. Yasuoka, M. Yasui, Molecular Mechanisms of How Mercury Inhibits Water Permeation through Aquaporin-1: Understanding by Molecular Dynamics Simulation, Biophys. J., 98 (2010) 1512–1519.
[114] G.M. Preston, J.S. Jung, W.B. Guggino, P. Agre, The mercury-sensitive residue at cysteine 189 in the CHIP28 water channel, J. Biol. Chem., 268 (1993) 17–20.
[115] R. Zhang, A.N. van Hoek, J. Biwersi, A.S. Verkman, A point mutation at cysteine 189 blocks the water permeability of rat kidney water channel CHIP28k, Biochemistry, 32 (1993) 2938–2941.
[116] T. Yoshida, S. Hojo, S. Sekine, S. Sawada, T. Okumura, T. Nagata, Y. Shimada, K. Tsukada, Expression of aquaporin-1 is a poor prognostic factor for stage II and III colon cancer, Mol. Clin. Oncol., 1 (2013) 953–958.
[117] Y. Tomita, H. Dorward, A. Yool, E. Smith, A. Townsend, T. Price, J. Hardingham, Role of Aquaporin 1 Signalling in Cancer Development and Progression, Int. J. Mol. Sci., 18 (2017) 299.
[118] X. Wei, J. Dong, Aquaporin 1 promotes the proliferation and migration of lung cancer cell in vitro, Oncol. Rep., 34 (2015) 1440–1448.
[119] L. Luo, R. Yang, S. Zhao, Y. Chen, S. Hong, K. Wang, T. Wang, J. Cheng, T. Zhang, D. Chen, Decreased miR-320 expression is associated with breast cancer progression, cell migration, and invasiveness via targeting Aquaporin 1, Acta Biochim. Biophys. Sin., 50 (2018) 473–480.
[120] C. Clapp, G.M. de la Escalera, Aquaporin-1: a novel promoter of tumor angiogenesis, Trends Endocrinol. Metab., 17 (2006) 1–2.
[121] S. Saadoun, M.C. Papadopoulos, M. Hara-Chikuma, A.S. Verkman, Impairment of angiogenesis and cell migration by targeted aquaporin-1 gene disruption, Nature, 434 (2005) 786–792.
[122] G.P. Nicchia, C. Stigliano, A. Sparaneo, A. Rossi, A. Frigeri, M. Svelto, Inhibition of aquaporin-1 dependent angiogenesis impairs tumour growth in a mouse model of melanoma, J. Mol. Med., 91 (2013) 613–623.
[123] M.O. Hoque, J.-C. Soria, J. Woo, T. Lee, J. Lee, S.J. Jang, S. Upadhyay, B. Trink, C. Monitto, C. Desmaze, L. Mao, D. Sidransky, C. Moon, Aquaporin 1 Is Overexpressed in Lung Cancer and Stimulates NIH-3T3 Cell Proliferation and Anchorage-Independent Growth, Am. J. Pathol., 168 (2006) 1345–1353.
[124] Y. Jiang, Aquaporin-1 activity of plasma membrane affects HT20 colon cancer cell migration, IUBMB Life, 61 (2009) 1001–1009.
[125] P.H. Chow, M. Kourghi, J.V. Pei, S. Nourmohammadi, A.J. Yool, 5-Hydroxymethyl-Furfural and Structurally Related Compounds Block the Ion Conductance in Human Aquaporin-1 Channels and Slow Cancer Cell Migration and Invasion, Mol. Pharmacol., 98 (2020) 38–48.

[126] Y. Ji, X. Liao, Y. Jiang, W. Wei, H. Yang, Aquaporin 1 knockdown inhibits triple-negative breast cancer cell proliferation and invasion *in vitro* and *in vivo*, Oncol. Lett., 21 (2021) 437.
[127] W. Chong, H. Zhang, Z. Guo, L. Yang, Y. Shao, X. Liu, Y. Zhao, Z. Wang, M. Zhang, C. Guo, L. Fu, Y. Ma, F. Gu, Aquaporin 1 promotes sensitivity of anthracycline chemotherapy in breast cancer by inhibiting β-catenin degradation to enhance TopoIIα activity, Cell Death Differ., 28 (2021) 382–400.
[128] J. Liu, Q. Xiao, J. Xiao, C. Niu, Y. Li, X. Zhang, Z. Zhou, G. Shu, G. Yin, Wnt/β-catenin signalling: function, biological mechanisms, and therapeutic opportunities, Signal Transduct. Target. Ther., 7 (2022) 3.
[129] S. Shang, F. Hua, Z.-W. Hu, The regulation of β-catenin activity and function in cancer: therapeutic opportunities, Oncotarget, 8 (2017) 33972–33989.
[130] K. Hsu, T. Lee, A. Periasamy, F. Kao, L. Li, C. Lin, H. Lin, M. Lin, Adaptable interaction between aquaporin-1 and band 3 reveals a potential role of water channel in blood CO_2 transport, FASEB J., 31 (2017) 4256–4264.
[131] J.W. Vince, R.A.F. Reithmeier, Carbonic Anhydrase II Binds to the Carboxyl Terminus of Human Band 3, the Erythrocyte Cl−/HCO3−Exchanger *, J. Biol. Chem., 273 (1998) 28430–28437.
[132] D.N. Wang, Band 3 protein: structure, flexibility and function, FEBS Lett., 346 (1994) 26–31.
[133] R.A.F. Reithmeier, A Membrane Metabolon Linking Carbonic Anhydrase with Chloride/Bicarbonate Anion Exchangers, Blood Cells. Mol. Dis., 27 (2001) 85–89.
[134] R. Occhipinti, W.F. Boron, Role of Carbonic Anhydrases and Inhibitors in Acid-Base Physiology: Insights from Mathematical Modeling, Int. J. Mol. Sci., 20 (2019) 3841.
[135] Evidence that aquaporin 1 is a major pathway for CO2 transport across the human erythrocyte membrane, (n.d.).
[136] A. Missner, P. Kügler, S.M. Saparov, K. Sommer, J.C. Mathai, M.L. Zeidel, P. Pohl, Carbon Dioxide Transport through Membranes, J. Biol. Chem., 283 (2008) 25340–25347.
[137] I. Silva, J. Silva, R. Ferreira, D. Trigo, Glymphatic system, AQP4, and their implications in Alzheimer's disease, Neurol. Res. Pract., 3 (2021) 5.
[138] J. Park, M. Madan, S. Chigurupati, S.H. Baek, Y. Cho, M.R. Mughal, A. Yu, S.L. Chan, J.V. Pattisapu, M.P. Mattson, D.-G. Jo, Neuronal Aquaporin 1 Inhibits Amyloidogenesis by Suppressing the Interaction Between Beta-Secretase and Amyloid Precursor Protein, J. Gerontol. Ser. A, 76 (2021) 23–31.
[139] T. Misawa, K. Arima, H. Mizusawa, J. Satoh, Close association of water channel AQP1 with amyloid-β deposition in Alzheimer disease brains, Acta Neuropathol. (Berl.), 116 (2008) 247–260.
[140] D. Linke, Chapter 34 Detergents, in: Methods Enzymol., Elsevier, 2009: pp. 603–617.
[141] S. Clarke, The size and detergent binding of membrane proteins, J. Biol. Chem., 250 (1975) 5459–5469.
[142] D.J. Slotboom, R.H. Duurkens, K. Olieman, G.B. Erkens, Static light scattering to characterize membrane proteins in detergent solution, Methods, 46 (2008) 73–82.
[143] J. Lipfert, L. Columbus, V.B. Chu, S.A. Lesley, S. Doniach, Size and Shape of Detergent Micelles Determined by Small-Angle X-ray Scattering, J. Phys. Chem. B, 111 (2007) 12427–12438.

[144] P.J. Bond, J.M. Cuthbertson, S.S. Deol, M.S.P. Sansom, MD Simulations of Spontaneous Membrane Protein/Detergent Micelle Formation, J. Am. Chem. Soc., 126 (2004) 15948–15949.
[145] J.M. Corkill, J.F. Goodman, S.P. Harrold, Thermodynamics of micellization of non-ionic detergents, Trans. Faraday Soc., 60 (1964) 202.
[146] J.E. de Vlugt, P. Xiao, R. Munro, A. Charchoglyan, D. Brewer, M.S. Al-Abdul-Wahid, L.S. Brown, V. Ladizhansky, Identifying lipids tightly bound to an integral membrane protein, Biochim. Biophys. Acta BBA - Biomembr., 1862 (2020) 183345.
[147] Y. Guo, Be Cautious with Crystal Structures of Membrane Proteins or Complexes Prepared in Detergents, Crystals, 10 (2020) 86.
[148] Y. Guo, Detergent-free systems for structural studies of membrane proteins, Biochem. Soc. Trans., 49 (2021) 1361–1374.
[149] R. Skrzypek, S. Iqbal, R. Callaghan, Methods of reconstitution to investigate membrane protein function, Methods, 147 (2018) 126–141.
[150] T.K. Ritchie, Y.V. Grinkova, T.H. Bayburt, I.G. Denisov, J.K. Zolnerciks, W.M. Atkins, S.G. Sligar, Reconstitution of Membrane Proteins in Phospholipid Bilayer Nanodiscs, in: Methods Enzymol., Elsevier, 2009: pp. 211–231.
[151] P.J. Loll, Membrane proteins, detergents and crystals: what is the state of the art?, Acta Crystallogr. Sect. F Struct. Biol. Commun., 70 (2014) 1576–1583.
[152] J.-L. Popot, T. Althoff, D. Bagnard, J.-L. Banères, P. Bazzacco, E. Billon-Denis, L.J. Catoire, P. Champeil, D. Charvolin, M.J. Cocco, G. Crémel, T. Dahmane, L.M. de la Maza, C. Ebel, F. Gabel, F. Giusti, Y. Gohon, E. Goormaghtigh, E. Guittet, J.H. Kleinschmidt, et al., Amphipols From A to Z, Annu. Rev. Biophys., 40 (2011) 379–408.
[153] M. Zoonens, J.-L. Popot, Amphipols for Each Season, J. Membr. Biol., 247 (2014) 759–796.
[154] C. Tribet, R. Audebert, J.-L. Popot, Amphipols: Polymers that keep membrane proteins soluble in aqueous solutions, Proc. Natl. Acad. Sci., 93 (1996) 15047–15050.
[155] J.-L. Popot, E.A. Berry, D. Charvolin, C. Creuzenet, C. Ebel, D.M. Engelman, M. Fl□tenmeyer, F. Giusti, Y. Gohon, P. Herv□, Q. Hong, J.H. Lakey, K. Leonard, H.A. Shuman, P. Timmins, D.E. Warschawski, F. Zito, M. Zoonens, B. Pucci, C. Tribet, Amphipols: polymeric surfactants for membrane biology research, Cell. Mol. Life Sci. CMLS, 60 (2003) 1559–1574.
[156] J.H. Kleinschmidt, J.-L. Popot, Folding and stability of integral membrane proteins in amphipols, Arch. Biochem. Biophys., 564 (2014) 327–343.
[157] C.L. Pocanschi, J.-L. Popot, J.H. Kleinschmidt, Folding and stability of outer membrane protein A (OmpA) from Escherichia coli in an amphipathic polymer, amphipol A8-35, Eur. Biophys. J., 42 (2013) 103–118.
[158] B.M. Gorzelle, A.K. Hoffman, M.H. Keyes, D.N. Gray, D.G. Ray, C.R. Sanders, Amphipols Can Support the Activity of a Membrane Enzyme, J. Am. Chem. Soc., 124 (2002) 11594–11595.
[159] M. Picard, T. Dahmane, M. Garrigos, C. Gauron, F. Giusti, M. le Maire, J.-L. Popot, P. Champeil, Protective and Inhibitory Effects of Various Types of Amphipols on the Ca^{2+}-ATPase from Sarcoplasmic Reticulum: A Comparative Study, Biochemistry, 45 (2006) 1861–1869.

[160] P. Champeil, T. Menguy, C. Tribet, J.-L. Popot, M. le Maire, Interaction of Amphipols with Sarcoplasmic Reticulum Ca2+-ATPase, J. Biol. Chem., 275 (2000) 18623–18637.
[161] N. Planchard, É. Point, T. Dahmane, F. Giusti, M. Renault, C. Le Bon, G. Durand, A. Milon, É. Guittet, M. Zoonens, J.-L. Popot, L.J. Catoire, The Use of Amphipols for Solution NMR Studies of Membrane Proteins: Advantages and Constraints as Compared to Other Solubilizing Media, J. Membr. Biol., 247 (2014) 827–842.
[162] P. Bazzacco, E. Billon-Denis, K.S. Sharma, L.J. Catoire, S. Mary, C. Le Bon, E. Point, J.-L. Banères, G. Durand, F. Zito, B. Pucci, J.-L. Popot, Nonionic Homopolymeric Amphipols: Application to Membrane Protein Folding, Cell-Free Synthesis, and Solution Nuclear Magnetic Resonance, Biochemistry, 51 (2012) 1416–1430.
[163] S. Elter, T. Raschle, S. Arens, A. Viegas, V. Gelev, M. Etzkorn, G. Wagner, The Use of Amphipols for NMR Structural Characterization of 7-TM Proteins, J. Membr. Biol., 247 (2014) 957–964.
[164] T.H. Bayburt, S.G. Sligar, Membrane protein assembly into Nanodiscs, FEBS Lett., 584 (2010) 1721–1727.
[165] S. Inagaki, R. Ghirlando, R. Grisshammer, Biophysical characterization of membrane proteins in nanodiscs, Methods, 59 (2013) 287–300.
[166] F.J.D. Alvarez, C. Orelle, A.L. Davidson, Functional Reconstitution of an ABC Transporter in Nanodiscs for Use in Electron Paramagnetic Resonance Spectroscopy, J. Am. Chem. Soc., 132 (2010) 9513–9515.
[167] N. Shirzad-Wasei, J. van Oostrum, P.H.M. Bovee-Geurts, L.J.A. Kusters, G.J.C.G.M. Bosman, W.J. DeGrip, Rapid transfer of overexpressed integral membrane protein from the host membrane into soluble lipid nanodiscs without previous purification, Biol. Chem., 396 (2015) 903–915.
[168] A. Nath, W.M. Atkins, S.G. Sligar, Applications of Phospholipid Bilayer Nanodiscs in the Study of Membranes and Membrane Proteins, Biochemistry, 46 (2007) 2059–2069.
[169] L. Sušac, R. Horst, K. Wüthrich, Solution-NMR Characterization of Outer-Membrane Protein A from *E coli* in Lipid Bilayer Nanodiscs and Detergent Micelles, ChemBioChem, 15 (2014) 995–1000.
[170] T. Raschle, S. Hiller, T.-Y. Yu, A.J. Rice, T. Walz, G. Wagner, Structural and Functional Characterization of the Integral Membrane Protein VDAC-1 in Lipid Bilayer Nanodiscs, J. Am. Chem. Soc., 131 (2009) 17777–17779.
[171] T. Viennet, M.M. Wördehoff, B. Uluca, C. Poojari, H. Shaykhalishahi, D. Willbold, B. Strodel, H. Heise, A.K. Buell, W. Hoyer, M. Etzkorn, Structural insights from lipid-bilayer nanodiscs link α-Synuclein membrane-binding modes to amyloid fibril formation, Commun. Biol., 1 (2018) 44.
[172] N. Akkaladevi, S. Mukherjee, H. Katayama, B. Janowiak, D. Patel, E.P. Gogol, B.L. Pentelute, R. John Collier, M.T. Fisher, Following Natures Lead: On the Construction of Membrane-Inserted Toxins in Lipid Bilayer Nanodiscs, J. Membr. Biol., 248 (2015) 595–607.
[173] K. Janson, F.L. Kyrilis, C. Tüting, M. Alfes, M. Das, T.K. Träger, C. Schmidt, F. Hamdi, C. Vargas, S. Keller, A. Meister, P.L. Kastritis, Cryo-Electron Microscopy Snapshots of Eukaryotic Membrane Proteins in Native Lipid-Bilayer Nanodiscs, Biomacromolecules, 23 (2022) 5084–5094.

[174] A.K. Khan, M. Jagielnicki, Brad.C. Bennett, M.D. Purdy, M. Yeager, Cryo-EM structure of an open conformation of a gap junction hemichannel in lipid bilayer nanodiscs, Structure, 29 (2021) 1040-1047.e3.
[175] T. Ravula, N.Z. Hardin, A. Ramamoorthy, Polymer nanodiscs: Advantages and limitations, Chem. Phys. Lipids, 219 (2019) 45–49.
[176] M.D. Farrelly, L.L. Martin, S.H. Thang, Polymer Nanodiscs and Their Bioanalytical Potential, Chem. – Eur. J., 27 (2021) 12922–12939.
[177] M.C. Fiori, Y. Jiang, W. Zheng, M. Anzaldua, M.J. Borgnia, G.A. Altenberg, H. Liang, Polymer Nanodiscs: Discoidal Amphiphilic Block Copolymer Membranes as a New Platform for Membrane Proteins, Sci. Rep., 7 (2017) 15227.
[178] K. Yasuhara, J. Arakida, T. Ravula, S.K. Ramadugu, B. Sahoo, J. Kikuchi, A. Ramamoorthy, Spontaneous Lipid Nanodisc Fomation by Amphiphilic Polymethacrylate Copolymers, J. Am. Chem. Soc., 139 (2017) 18657–18663.
[179] T.J. Knowles, R. Finka, C. Smith, Y.-P. Lin, T. Dafforn, M. Overduin, Membrane Proteins Solubilized Intact in Lipid Containing Nanoparticles Bounded by Styrene Maleic Acid Copolymer, J. Am. Chem. Soc., 131 (2009) 7484–7485.
[180] J.M. Dörr, S. Scheidelaar, M.C. Koorengevel, J.J. Dominguez, M. Schäfer, C.A. van Walree, J.A. Killian, The styrene–maleic acid copolymer: a versatile tool in membrane research, Eur. Biophys. J., 45 (2016) 3–21.
[181] A.H. Kopf, M.C. Koorengevel, C.A. van Walree, T.R. Dafforn, J.A. Killian, A simple and convenient method for the hydrolysis of styrene-maleic anhydride copolymers to styrene-maleic acid copolymers, Chem. Phys. Lipids, 218 (2019) 85–90.
[182] T. Alfrey, E. Lavin, The Copolymerization of Styrene and Maleic Anhydride, J. Am. Chem. Soc., 67 (1945) 2044–2045.
[183] S. Scheidelaar, M.C. Koorengevel, C.A. van Walree, J.J. Dominguez, J.M. Dörr, J.A. Killian, Effect of Polymer Composition and pH on Membrane Solubilization by Styrene-Maleic Acid Copolymers, Biophys. J., 111 (2016) 1974–1986.
[184] M. Jamshad, V. Grimard, I. Idini, T.J. Knowles, M.R. Dowle, N. Schofield, P. Sridhar, Y. Lin, R. Finka, M. Wheatley, O.R.T. Thomas, R.E. Palmer, M. Overduin, C. Govaerts, J.-M. Ruysschaert, K.J. Edler, T.R. Dafforn, Structural analysis of a nanoparticle containing a lipid bilayer used for detergent-free extraction of membrane proteins, Nano Res., 8 (2015) 774–789.
[185] D.J.K. Swainsbury, S. Scheidelaar, N. Foster, R. van Grondelle, J.A. Killian, M.R. Jones, The effectiveness of styrene-maleic acid (SMA) copolymers for solubilisation of integral membrane proteins from SMA-accessible and SMA-resistant membranes, Biochim. Biophys. Acta BBA - Biomembr., 1859 (2017) 2133–2143.
[186] L. Unger, A. Ronco-Campaña, P. Kitchen, R.M. Bill, A.J. Rothnie, Biological insights from SMA-extracted proteins, Biochem. Soc. Trans., 49 (2021) 1349–1359.
[187] A.A. Gulamhussein, D. Meah, D.D. Soja, S. Fenner, Z. Saidani, A. Akram, S. Lallie, A. Mathews, C. Painter, M.K. Liddar, Z. Mohammed, L.K. Chiu, S.S. Sumar, H. Healy, N. Hussain, J.H. Patel, S.C.L. Hall, T.R. Dafforn, A.J. Rothnie, Examining the stability of membrane proteins within SMALPs, Eur. Polym. J., 112 (2019) 120–125.

[188] A.J. Horsey, D.A. Briggs, N.D. Holliday, S.J. Briddon, I.D. Kerr, Application of fluorescence correlation spectroscopy to study substrate binding in styrene maleic acid lipid copolymer encapsulated ABCG2, Biochim. Biophys. Acta BBA - Biomembr., 1862 (2020) 183218.
[189] P.S. Orekhov, M.E. Bozdaganyan, N. Voskoboynikova, A.Y. Mulkidjanian, H.-J. Steinhoff, K.V. Shaitan, Styrene/Maleic Acid Copolymers Form SMALPs by Pulling Lipid Patches out of the Lipid Bilayer, Langmuir, 35 (2019) 3748–3758.
[190] S. Scheidelaar, M.C. Koorengevel, J.D. Pardo, J.D. Meeldijk, E. Breukink, J.A. Killian, Molecular Model for the Solubilization of Membranes into Nanodisks by Styrene Maleic Acid Copolymers, Biophys. J., 108 (2015) 279–290.
[191] M. Xue, L. Cheng, I. Faustino, W. Guo, S.J. Marrink, Molecular Mechanism of Lipid Nanodisk Formation by Styrene-Maleic Acid Copolymers, Biophys. J., 115 (2018) 494–502.
[192] V.A. Bjørnestad, M. Orwick-Rydmark, R. Lund, Understanding the Structural Pathways for Lipid Nanodisc Formation: How Styrene Maleic Acid Copolymers Induce Membrane Fracture and Disc Formation, Langmuir, 37 (2021) 6178–6188.
[193] M. Overduin, C. Trieber, R.S. Prosser, L.-P. Picard, J.G. Sheff, Structures and Dynamics of Native-State Transmembrane Protein Targets and Bound Lipids, Membranes, 11 (2021) 451.
[194] S.J. Routledge, M. Jamshad, H.A. Little, Y.-P. Lin, J. Simms, A. Thakker, C.M. Spickett, R.M. Bill, T.R. Dafforn, D.R. Poyner, M. Wheatley, Ligand-induced conformational changes in a SMALP-encapsulated GPCR, Biochim. Biophys. Acta Biomembr., 1862 (2020) 183235.
[195] M. Parmar, S. Rawson, C.A. Scarff, A. Goldman, T.R. Dafforn, S.P. Muench, V.L.G. Postis, Using a SMALP platform to determine a sub-nm single particle cryo-EM membrane protein structure, Biochim. Biophys. Acta BBA - Biomembr., 1860 (2018) 378–383.
[196] S. Gulati, M. Jamshad, T.J. Knowles, K.A. Morrison, R. Downing, N. Cant, R. Collins, J.B. Koenderink, R.C. Ford, M. Overduin, I.D. Kerr, T.R. Dafforn, A.J. Rothnie, Detergent-free purification of ABC (ATP-binding-cassette) transporters, Biochem. J., 461 (2014) 269–278.
[197] B. Krishnarjuna, T. Ravula, A. Ramamoorthy, Detergent-free extraction, reconstitution and characterization of membrane-anchored cytochrome-b5 in native lipids, Chem. Commun., 56 (2020) 6511–6514.
[198] B. Krishnarjuna, T. Ravula, A. Ramamoorthy, Detergent-free isolation of CYP450-reductase's FMN-binding domain in *E coli* lipid-nanodiscs using a charge-free polymer, Chem. Commun., 58 (2022) 4913–4916.
[199] V. Schmidt, J.N. Sturgis, Modifying styrene-maleic acid co-polymer for studying lipid nanodiscs, Biochim. Biophys. Acta BBA - Biomembr., 1860 (2018) 777–783.
[200] N.L. Pollock, S.C. Lee, J.H. Patel, A.A. Gulamhussein, A.J. Rothnie, Structure and function of membrane proteins encapsulated in a polymer-bound lipid bilayer, Biochim. Biophys. Acta BBA - Biomembr., 1860 (2018) 809–817.
[201] G. Grimsley, C. Pace, Current Protocols in Protein Science, Curr. Protoc. Protein Sci. Editor. Board John E Coligan Al, Chapter 3 (2004) Unit 3.1.
[202] T. Cao, S.E. Webber, Free-Radical Copolymerization of Fullerenes with Styrene, Macromolecules, 28 (1995) 3741–3743.
[203] A.M.C. Davies, An Introduction to near Infrared Spectroscopy, NIR News, 16 (2005) 9–11.
[204] W.D. Perkins, Fourier transform-infrared spectroscopy: Part l Instrumentation, J. Chem. Educ., 63 (1986) A5.

[205] Infrared spectroscopy: Fundamentals and applications, in: Anal. Tech. Sci., John Wiley & Sons, Ltd, Chichester, UK, 2005: pp. i–xvi.
[206] S.A. Tatulian, Structural Characterization of Membrane Proteins and Peptides by FTIR and ATR-FTIR Spectroscopy, in: J.H. Kleinschmidt (Ed.), Lipid-Protein Interact. Methods Protoc., Humana Press, Totowa, NJ, 2013: pp. 177–218.
[207] N. Colthup, Introduction to Infrared and Raman Spectroscopy, Elsevier, 2012.
[208] C. Berthomieu, R. Hienerwadel, Fourier transform infrared (FTIR) spectroscopy, Photosynth. Res., 101 (2009) 157–170.
[209] P.R. Griffiths, Fourier Transform Infrared Spectrometry, Science, 222 (1983) 297–302.
[210] M.-M. Blum, H. John, Historical perspective and modern applications of Attenuated Total Reflectance – Fourier Transform Infrared Spectroscopy (ATR-FTIR), Drug Test. Anal., 4 (2012) 298–302.
[211] K.A. Oberg, A.L. Fink, A New Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy Method for the Study of Proteins in Solution, Anal. Biochem., 256 (1998) 92–106.
[212] J.D. Kubicki, L.M. Schroeter, M.J. Itoh, B.N. Nguyen, S.E. Apitz, Attenuated total reflectance Fourier-transform infrared spectroscopy of carboxylic acids adsorbed onto mineral surfaces, Geochim. Cosmochim. Acta, 63 (1999) 2709–2725.
[213] S.A. Tatulian, Attenuated Total Reflection Fourier Transform Infrared Spectroscopy: A Method of Choice for Studying Membrane Proteins and Lipids, Biochemistry, 42 (2003) 11898–11907.
[214] V.A. Lorenz-Fonfria, Infrared Difference Spectroscopy of Proteins: From Bands to Bonds, Chem. Rev., 120 (2020) 3466–3576.
[215] D.M. Byler, H. Susi, Examination of the secondary structure of proteins by deconvolved FTIR spectra, Biopolymers, 25 (1986) 469–487.
[216] W. Gallagher, FTIR Analysis of Protein Structure, (n.d.).
[217] S.E. Glassford, B. Byrne, S.G. Kazarian, Recent applications of ATR FTIR spectroscopy and imaging to proteins, Biochim. Biophys. Acta BBA - Proteins Proteomics, 1834 (2013) 2849–2858.
[218] L.C. Mayne, B. Hudson, Resonance Raman spectroscopy of N-methylacetamide: overtones and combinations of the carbon-nitrogen stretch (amide II') and effect of solvation on the carbon-oxygen double-bond stretch (amide I) intensity, J. Phys. Chem., 95 (1991) 2962–2967.
[219] C. Vigano, M. Smeyers, V. Raussens, F. Scheirlinckx, J. m. Ruysschaert, E. Goormaghtigh, Hydrogen-deuterium exchange in membrane proteins monitored by IR spectroscopy: A new tool to resolve protein structure and dynamics, Biopolymers, 74 (2004) 19–26.
[220] S. Cai, B.R. Singh, A Distinct Utility of the Amide III Infrared Band for Secondary Structure Estimation of Aqueous Protein Solutions Using Partial Least Squares Methods, Biochemistry, 43 (2004) 2541–2549.
[221] H. Susi, S.N. Timasheff, L. Stevens, Infrared Spectra and Protein Conformations in Aqueous Solutions, J. Biol. Chem., 242 (1967) 5460–5466.
[222] E. Goormaghtigh, Infrared Spectroscopy of Protein Dynamics: H/D Exchange, in: G.C.K. Roberts (Ed.), Encycl. Biophys., Springer, Berlin, Heidelberg, 2013: pp. 1081–1083.

[223] P. Gabbott, A Practical Introduction to Differential Scanning Calorimetry, in: P. Gabbott (Ed.), Princ. Appl. Therm. Anal., Blackwell Publishing Ltd, Oxford, UK, 2008: pp. 1–50.
[224] P. Gill, T.T. Moghadam, B. Ranjbar, Differential Scanning Calorimetry Techniques: Applications in Biology and Nanoscience, 21 (2010).
[225] S. Mazurenko, A. Kunka, K. Beerens, C.M. Johnson, J. Damborsky, Z. Prokop, Exploration of Protein Unfolding by Modelling Calorimetry Data from Reheating, Sci. Rep., 7 (2017) 16321.
[226] P. Farber, H. Darmawan, T. Sprules, A. Mittermaier, Analyzing Protein Folding Cooperativity by Differential Scanning Calorimetry and NMR Spectroscopy, J. Am. Chem. Soc., 132 (2010) 6214–6222.
[227] V. Muñoz, L.A. Campos, M. Sadqi, Limited cooperativity in protein folding, Curr. Opin. Struct. Biol., 36 (2016) 58–66.
[228] M.B. Jackson, J.M. Sturtevant, Phase transitions of the purple membranes of Halobacterium halobium, Biochemistry, 17 (1978) 911–915.
[229] J. Stetefeld, S.A. McKenna, T.R. Patel, Dynamic light scattering: a practical guide and applications in biomedical sciences, Biophys. Rev., 8 (2016) 409–427.
[230] M. Kaszuba, D. McKnight, M.T. Connah, F.K. McNeil-Watson, U. Nobbmann, Measuring sub nanometre sizes using dynamic light scattering, J. Nanoparticle Res., 10 (2008) 823–829.
[231] H. Friebolin, Basic one- and two-dimensional NMR spectroscopy, 4th completely rev. and updated ed, WILEY-VCH, Weinheim, 2005.
[232] K.S. Mineev, K.D. Nadezhdin, Membrane mimetics for solution NMR studies of membrane proteins, Nanotechnol. Rev., 6 (2017) 15–32.
[233] M.P. Foster, C.A. McElroy, C.D. Amero, Solution NMR of large molecules and assemblies, Biochemistry, 46 (2007) 331–340.
[234] V. Tugarinov, V. Kanelis, L.E. Kay, Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy, Nat. Protoc., 1 (2006) 749–754.
[235] A.A. Yee, A. Savchenko, A. Ignachenko, J. Lukin, X. Xu, T. Skarina, E. Evdokimova, C.S. Liu, A. Semesi, V. Guido, A.M. Edwards, C.H. Arrowsmith, NMR and X-ray Crystallography, Complementary Tools in Structural Proteomics of Small Proteins, J. Am. Chem. Soc., 127 (2005) 16512–16517.
[236] K. Pervushin, R. Riek, G. Wider, K. Wüthrich, Transverse Relaxation-Optimized Spectroscopy (TROSY) for NMR Studies of Aromatic Spin Systems in 13C-Labeled Proteins, J. Am. Chem. Soc., 120 (1998) 6394–6400.
[237] P. Rajagopal, E.B. Waygood, J. Reizer, M.H. Saier, R.E. Klevit, Demonstration of protein-protein interaction specificity by NMR chemical shift mapping, Protein Sci., 6 (2008) 2624–2627.
[238] E. Barile, M. Pellecchia, NMR-Based Approaches for the Identification and Optimization of Inhibitors of Protein–Protein Interactions, Chem. Rev., 114 (2014) 4749–4763.
[239] W. Becker, K.C. Bhattiprolu, N. Gubensäk, K. Zangger, Investigating Protein-Ligand Interactions by Solution Nuclear Magnetic Resonance Spectroscopy, ChemPhysChem, 19 (2018) 895–906.

[240] T. Arakawa, S.N. Timasheff, Mechanism of protein salting in and salting out by divalent cation salts: balance between hydration and salt binding, Biochemistry, 23 (1984) 5912–5923.
[241] A.H. Kopf, J.M. Dörr, M.C. Koorengevel, F. Antoniciello, H. Jahn, J.A. Killian, Factors influencing the solubilization of membrane proteins from Escherichia coli membranes by styrene–maleic acid copolymers, Biochim. Biophys. Acta BBA - Biomembr., 1862 (2020) 183125.
[242] S. Scheidelaar, M.C. Koorengevel, C.A. van Walree, J.J. Dominguez, J.M. Dörr, J.A. Killian, Effect of Polymer Composition and pH on Membrane Solubilization by Styrene-Maleic Acid Copolymers, Biophys. J., 111 (2016) 1974–1986.
[243] A.A. Gulamhussein, R. Uddin, B.J. Tighe, D.R. Poyner, A.J. Rothnie, A comparison of SMA (styrene maleic acid) and DIBMA (di-isobutylene maleic acid) for membrane protein purification, Biochim. Biophys. Acta BBA - Biomembr., 1862 (2020) 183281.
[244] A.F. Craig, E.E. Clark, I.D. Sahu, R. Zhang, N.D. Frantz, M.S. Al-Abdul-Wahid, C. Dabney-Smith, D. Konkolewicz, G.A. Lorigan, Tuning the size of styrene-maleic acid copolymer-lipid nanoparticles (SMALPs) using RAFT polymerization for biophysical studies, Biochim. Biophys. Acta BBA - Biomembr., 1858 (2016) 2931–2939.
[245] P. Samyn, G. Schoukens, Morphologies and Thermal Variability of Patterned Polymer Films with Poly(styrene-co-maleic anhydride), Polymers, 6 (2014) 820–845.
[246] M. Karlova, D. Bagrov, M. Vorobyova, K. Mamatkulov, G. Arzumanyan, O. Sokolova, K. Shaitan, Raman spectroscopy reveals lipids in protein-containing SMA-stabilized lipodiscs, Microsc. Microanal., 27 (2021) 1714–1715.
[247] F.M. Poulsen, Introduction to NMR spectroscopy of proteins, (n.d.).
[248] S. Kosol, S. Contreras-Martos, C. Cedeño, P. Tompa, Structural Characterization of Intrinsically Disordered Proteins by NMR Spectroscopy, Molecules, 18 (2013) 10802–10828.
[249] L.P. McIntosh, E. Brun, L.E. Kay, Stereospecific assignment of the NH2 resonances from the primary amides of asparagine and glutamine side chains in isotopically labeled proteins, (n.d.).
[250] D.A. Dingwell, L.S. Brown, V. Ladizhansky, Structure of the Functionally Important Extracellular Loop C of Human Aquaporin 1 Obtained by Solid-State NMR under Nearly Physiological Conditions, J. Phys. Chem. B, 123 (2019) 7700–7710.
[251] A. Grethen, A.O. Oluwole, B. Danielczak, C. Vargas, S. Keller, Thermodynamics of nanodisc formation mediated by styrene/maleic acid (2:1) copolymer, Sci. Rep., 7 (2017) 11517.
[252] R. Cuevas Arenas, B. Danielczak, A. Martel, L. Porcar, C. Breyton, C. Ebel, S. Keller, Fast Collisional Lipid Transfer Among Polymer-Bounded Nanodiscs, Sci. Rep., 7 (2017) 45875.
[253] G. Fischer, U. Kosinska-Eriksson, C. Aponte-Santamaría, M. Palmgren, C. Geijer, K. Hedfalk, S. Hohmann, B.L. de Groot, R. Neutze, K. Lindkvist-Petersson, Crystal Structure of a Yeast Aquaporin at 115 Å Reveals a Novel Gating Mechanism, PLOS Biol., 7 (2009) e1000130.
[254] H. Sui, B.-G. Han, J.K. Lee, P. Walian, B.K. Jap, Structural basis of water-specific transport through the AQP1 water channel, Nature, 414 (2001) 872–878.

[255] A. Frick, U.K. Eriksson, F. de Mattia, F. Öberg, K. Hedfalk, R. Neutze, W.J. de Grip, P.M.T. Deen, S. Törnroth-Horsefield, X-ray structure of human aquaporin 2 and its implications for nephrogenic diabetes insipidus and trafficking, Proc. Natl. Acad. Sci., 111 (2014) 6305–6310.
[256] U. Kosinska Eriksson, G. Fischer, R. Friemann, G. Enkavi, E. Tajkhorshid, R. Neutze, Subangstrom Resolution X-Ray Structure Details Aquaporin-Water Interactions, Science, 340 (2013) 1346–1349.
[257] E. Beitz, B. Wu, L.M. Holm, J.E. Schultz, T. Zeuthen, Point mutations in the aromatic/arginine region in aquaporin 1 allow passage of urea, glycerol, ammonia, and protons, Proc. Natl. Acad. Sci., 103 (2006) 269–274.
[258] J. Demsˇar, T. Curk, A. Erjavec, J. Demsar, T. Curk, A. Erjave, C. Gorup, T. Hocevar, M. Milutinovic, M. Mozina, M. Polajnar, M. Toplak, A. Staric, M. Stajdohar, L. Umek, L. Zagar, J. Zbontar, M. Zitnik, B. Zupan, Orange: Data Mining Toolbox in Python, (n.d.).
[259] Z. Yang, C. Wang, Q. Zhou, J. An, E. Hildebrandt, L.A. Aleksandrov, J.C. Kappes, L.J. DeLucas, J.R. Riordan, I.L. Urbatsch, J.F. Hunt, C.G. Brouillette, Membrane protein stability can be compromised by detergent interactions with the extramembranous soluble domains, Protein Sci. Publ. Protein Soc., 23 (2014) 769–789.
[260] A. Stetsenko, A. Guskov, An Overview of the Top Ten Detergents Used for Membrane Protein Crystallization, Crystals, 7 (2017) 197.
[261] J.T. Marinko, H. Huang, W.D. Penn, J.A. Capra, J.P. Schlebach, C.R. Sanders, Folding and Misfolding of Human Membrane Proteins in Health and Disease: From Single Molecules to Cellular Proteostasis, Chem. Rev., 119 (2019) 5537–5606.
[262] G.H.M. Huysmans, S.A. Baldwin, D.J. Brockwell, S.E. Radford, The transition state for folding of an outer membrane protein, Proc. Natl. Acad. Sci., 107 (2010) 4099–4104.
[263] T. Heimburg, D. Marsh, Investigation of secondary and tertiary structural changes of cytochrome c in complexes with anionic lipids using amide hydrogen exchange measurements: an FTIR study, Biophys. J., 65 (1993) 2408–2417.
[264] P. Kitchen, M.M. Salman, S.U. Pickel, J. Jennings, S. Törnroth-Horsefield, M.T. Conner, R.M. Bill, A.C. Conner, Water channel pore size determines exclusion properties but not solute selectivity, Sci. Rep., 9 (2019) 20369.
[265] J. De Meutter, E. Goormaghtigh, Evaluation of protein secondary structure from FTIR spectra improved after partial deuteration, Eur. Biophys. J., 50 (2021) 613–628.
[266] Y.M. Efimova, S. Haemers, B. Wierczinski, W. Norde, A.A. van Well, Stability of globular proteins in H2O and D2O, Biopolymers, 85 (2007) 264–273.
[267] G.I. Makhatadze, G.M. Clore, A.M. Gronenborn, Solvent isotope effect and protein stability, Nat. Struct. Biol., 2 (1995) 852–855.
[268] M. Dreydoppel, J. Balbach, U. Weininger, Monitoring protein unfolding transitions by NMR-spectroscopy, J. Biomol. NMR, 76 (2022) 3–15.
[269] H.-J. Hinz, T. Vogl, R. Meyer, An alternative interpretation of the heat capacity changes associated with protein unfolding, Biophys. Chem., 52 (1994) 275–285.