

Molecular Analysis of Early Aquaporin Biogenesis and Folding Events

by

Teresa M. Buck

A DISSERTATION

Presented to the Department of Physiology and Pharmacology

and Oregon Health & Science University

School of Medicine

in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

May 2006

**School of Medicine
Oregon Health and Sciences University**

Certificate of Approval

This is to certify the Ph.D. thesis of

Teresa M. Buck

Has been approved



Professor in charge of thesis



Member



Member



Member



Member

Table of Contents

List of figures	iv
Abbreviations	vi
Acknowledgments	viii
I. Abstract	ix
II. Introduction	1
A. Aquaporin Physiology	1
1. Aquaporin Basics	
2. AQPs in the Kidney	
3. Other AQP Physiology	
B. Aquaporin Structure and Function	3
1. Structural selectivity (water vs. glycerol)	
2. AQP Gating and Regulation	
3. Other AQP substrates	
C. Aquaporin Biogenesis/Folding	6
1. Early events of AQP Translation	
2. Aquaporin Topogenesis	
3. Late Folding events of AQPs	
4. Oligomerization of AQPs	
D. Aquaporin Maturation and Degradation	12
1. AQP N-linked Glycosylation	
2. AQP Degradation/ERAD	
E. Specific Aims Adressed in this Thesis	15
1. How do NDI causing Mutations affect early biogenesis events?	
2. Is AQP1's unique biogenesis mechanism conserved in diverse expression systems?	
3. How is AQP1 topological maturation achieved in the ER?	
III. Results	
A. Manuscript 1: Evidence for stabilization of aquaporin-2 folding mutants by N-linked glycosylation in endoplasmic reticulum.	21

1. Abstract	
2. Introduction	
3. Materials and Methods	
4. Results	
5. Discussion	
B. Manuscript 2: Differential Stability of Biogenesis Intermediates Reveals a Common Pathway for Aquaporin-1 Topological Maturation	45
1. Abstract	
2. Introduction	
3. Materials and Methods	
4. Results	
5. Discussion	
C. Manuscript 3: A triad of hydrogen bonds is required for Aquaporin-1 early topogenesis, monomer folding, and tetramer stability	74
1. Abstract	
2. Introduction	
3. Materials and Methods	
4. Results	
5. Discussion	
IV. Summary and Future Directions	101
A. Summary	101
B. Future Directions	103
1. Can the glycosylated mutants traffic?	
2. Do chaperones stabilize glycosylated AQP2 protein?	
3. Why do topological isoforms have differential stability?	
4. What does N49M/K51L associate with?	
5. Is the Lys51/Asp185 interaction important for monomer folding too?	
6. Does D185 block heterotetramerization?	
7. In what cellular location does AQP1 mature?	
V. Appendix	107
A. Data Not Shown	107

List of Figures

III. Results

A. Manuscript 1

- Fig. 1-1:** Diagram of predicted aquaporin-2 (AQP2) topology and relative locations of 4 nephrogenic insipidus (NDI)-linked mutations, T126M, A147T, C181W and R187C. 38
- Fig. 1-2:** Glycosylation profiles for wild-type and mutant AQP2. 39
- Fig. 1-3:** Pulse-Chase analysis of AQP2 synthesis in *Xenopus* oocytes. 40
- Fig. 1-4:** Percentage of glycosylated wild-type and mutant AQP2 protein over time. 42
- Fig. 1-5:** Stability of glycosylated or non-glycosylated wild-type mutant AQP2 protein over time. 43
- Fig. 1-6:** Effect of trafficking mutants T126M, A147T, C181W or R187C on the cotranslational glycosylation of AQP2 protein. 44

B. Manuscript 2

- Fig. 2-1:** Two models of AQP1 biogenesis. 67
- Fig. 2-2:** Effect of reporter fusion site on early AQP1 biogenesis in oocytes. 68
- Fig. 2-3:** Stability of two biogenesis isoforms that are generated when AQP1 is truncated after TM3. 69
- Fig. 2-4:** 2-hour pulse-labeling of transfected HEK-293 cells to analyze the early topology of truncated AQP1 fusion proteins using glycosylation reporter. 70
- Fig. 2-5:** Pulse-Chase analysis of AQP1 (TM1-3) fusion proteins in HEK-293 cells. 71
- Fig. 2-6:** Effect of MG132 on stability of different early AQP1 biogenesis intermediates. 72
- Fig. 2-7:** Pulse-Chase analysis of the stability of full-length

wild-type AQP1 with or without MG132.	73
C. Manuscript 3	
Fig. 3-1: AQP1 structural model and biogenesis diagram.	94
Fig. 3-2: Assay of how mutations in AQP1 TM2 affect its stop transfer activity .	95
Fig. 3-3: Functional assay of AQP1 and AQP4 mutants	97
Fig. 3-4: Models of the hydrogen bonds between AQP1 Asn49, Lys51 and Asp185 in the monomer and tetramer.	98
Fig. 3-5: Analysis of WT and mutant AQP1 oligomerization state via sucrose density centrifugation.	99
Fig. 3-6: Coimmunoprecipitation of WT and mutant AQP1 to assess oligomerization state.	100
V. Appendices	
A. Data Not Shown	
Fig. A-1. Protease Protection assay of truncated AQP2 fusion proteins to test the effects of trafficking mutants (T126M, A147T, C181W and R187C) on early topogenesis events.	107
Fig. A-2. Time-course for depletion of dolichol bound oligosaccharide to block N-linked glycosylation of a substrate.	109
Fig. A-3. Protease protection assay to measure the effects of the addition of 8 hydrophobic amino acids to the C-terminus of AQP1 TM2.	110
Fig. A-4. Pulse-Chase analysis of early AQP1 biogenesis intermediates in the presence of proteasome inhibitor ALLN.	111

List of Abbreviations

AQP	Aquaporin
CNX	calnexin
CRT	calreticulin
DTT	dithiothreitol
EGFP	Enhanced fluorescent green protein
ER	Endoplasmic reticulum
ERAD	Endoplasmic reticulum associated degradation
IP	immunoprecipitation
MIP	Major Intrinsic Protein
NDI	Nephrogenic Diabetes Insipidus
OST	oligosaccharyltransferase
PK	proteinase K
PDI	Protein disulphide isomerase
PCR	polymerase chain reaction
PNGase F	N-glycosidase F
RI	Ribophorin I
SEM	Standard Error of the Mean
TM	transmembrane; spanning, membrane-spanning
TRAM	translocation chain-associated membrane protein
TRAP	translocon-associated protein
UGGT	UDP-glucose:glycoprotein glucosyl transferase
TX-100	Triton X-100

WT wild-type
OAP orthogonal array particles

Acknowledgements

I would like to thank everyone who has helped me complete this research. First, my family and friends for continuing to be supportive through the ups and many downs of graduate school. I would especially like to thank my fellow graduate students Yu-Wen Lin, Catherine Selph and Brenda McKee for creating a great support network. I would like to thank the many many members of the Skach lab that have provided both technical and emotional support (Jon Oberdorf, Colin Daniel, Vicki Anthony, Radhika Vaishnav, Eric Carlson, Dave Pitonzo, Joel Eledge, Jamie Knowles, Fred Larrabee, Steve Grund, Elisa Cooper, Barb Leighton, Prassana Kumar, Toru Shibatani, Simon Thompson, Heather Sadlish, Karl Rusterholz, and Justin Wagner). I would like to acknowledge my thesis committee and advisor for their input and their time spent (Bill Skach, Dennis Koop, Linda Musil, Showling Shyng, Zhengfeng Zhou, and Steve Smith) advising me in this process.

I. Abstract

Aquaporins (AQPs) are small (~29 kDa) hydrophobic proteins that belong to the Major Intrinsic Protein (MIP) family of glycerol and water channels. AQPs are 6-transmembrane spanning proteins with two additional helical regions flanked by conserved NPA motifs that fold inward within the plane of the membrane to create a selectivity filter. They are ubiquitously expressed and best characterized within the kidney water reabsorption system. Four AQPs are responsible for water reabsorption in the nephron; AQP1 in the proximal tubule and AQP2, AQP3, and AQP4 in the collecting duct. Over 30 mutations in AQP2 have been shown to cause hereditary nephrogenic diabetes insipidus (NDI), a disease characterized by the inability to produce concentrated urine. AQPs traffic as homotetramers and transport water passively across the plasma membrane according to concentration gradient. While the recently determined crystal structure of AQP family members has led to a more complete understanding of the architecture of the AQP water channel and pore, very little is known about the early folding events preceding the formation of a functional channel. For this reason my studies have focused on investigating the early steps of AQP topogenesis, processing and folding at the ER and how these events may ultimately affect trafficking and function.

AQPs and other polytopic membrane proteins are synthesized and oriented in the ER by the ribosome-translocon complex. N-linked glycosylation occurs cotranslationally as a protein is translocated into the ER. Only a subset of AQP2 is glycosylated (~20-30%). Interestingly, we found that while glycosylation had no apparent effect on WT protein it markedly stabilized the AQP2 trafficking mutants T126M, A147T, C181W and R187C in comparison

to the nonglycosylated population of protein (~25 h in comparison to ~5 h). Therefore N-linked glycans appear have a general mechanism to compensate in some general way for folding defects throughout the protein. Knowledge of whether diverse expression systems recognize topogenic information similarly is crucial for a complete understanding of the biogenesis process. Using a systematic comparison of oocyte and mammalian cell systems, we demonstrated that AQP1's unique topogenesis whereby it is initially synthesized as a 4-TM protein, which is then converted to the mature 6-spanning channel is conserved across systems, and that the information encoded with the nascent chain itself is primarily responsible for a protein's orientation within the lipid bilayer. Further, these studies also showed that truncated proteins lacking C-terminal information could generate multiple biogenesis intermediates that reflect different steps in early protein biogenesis. These intermediates exhibit very different stabilities depending on the stage of protein biosynthesis, and this should be taken into consideration when using truncated proteins to study topogenesis. Previous work has shown that two polar residues in AQP1 TM2, Asn49 and Lys51, were necessary for TM2 to slip into the ER lumen. We demonstrate that these two residues specifically interact with a charged residue in TM5, Asp185. Through intramolecular interactions, Asn49 and Asp185 are necessary for monomer folding, whereas intermolecular interactions between Lys51 and Asp185 help to stabilize the tetramer. These two residues are necessary to compensate for the charged Asp185, and thus AQP1's unique biogenesis is a consequence of this requirement. Importantly, these studies outline the complex relationship between topogenesis events, tertiary folding and oligomerization of a polytopic membrane protein. Together the results presented in this thesis advance our basic knowledge of early events of AQP synthesis and folding and may well be applied to other

multi-spanning proteins.

II. Introduction

A. Aquaporin Physiology

1) Aquaporin Basics

Aquaporins (AQPs) are members of the Major Intrinsic Protein (MIP) family of proteins, which are ubiquitously expressed throughout the bacterial, plant, and animal kingdoms (1,2). Most MIP family members can be classified as aquaporins (water conducting) or glyceroporins (glycerol conducting), and thirteen members of the aquaporin family (AQP0-AQP12) have been identified, to date, in humans. AQPs are small (~29 kD) 6 transmembrane spanning proteins that exist as homotetramers in cell membranes (3). AQP monomers have a two-fold inverted hour glass pseudosymmetry and two inverted repeats (NPA motifs) that fold inward within the plane of the membrane to help form the selectivity filter within the pore (4,5). AQPs passively transport water and/or glycerol down their concentration gradient and most have a rigid selectivity filter that excludes ions including hydronium and hydroxide (6). In humans, aquaporins are expressed in a wide variety of tissue and cell types including lens fiber cells, kidney tubules, red and white blood cells, brain, epithelium, lung, adipose tissue, testis, and heart (2). While the physiological importance in many cases remains to be elucidated, expression in most tissues confers a high capacity for water transport.

2) AQPs in the kidney

AQP expression and function have been most extensively characterized in the kidney where several AQP members are primarily involved in water reabsorption. AQP1 is expressed in the proximal tubule, which is responsible for reabsorbing approximately 75% of the total glomerular filtrate (150-180 L/day) (7). Consistent with this, microperfusion experiments on isolated proximal tubules from WT and AQP1-null mice revealed significantly higher water permeability for the WT tubules (0.033 cm/s versus 0.15 cm/s) (8). Interestingly, AQP1-null mice are normal in terms of survival, appearance, and organ morphology but become severely dehydrated when water deprived in contrast to WT (9). Humans that lack AQP1 (Colton Blood

Antigen) show a similar phenotype with a mildly impaired ability to concentrate urine (10). The symptoms are subclinical except under stress conditions. Although AQP1 normally reabsorbs the majority of glomerular filtrate, AQPs in the distal nephron are apparently able to compensate for the proximal AQP1 defect under nonstressed conditions.

AQP2, the best characterized renal AQP, is expressed in the kidney collecting duct and is responsible for the vasopressin induced water reabsorption. Vasopressin is released from the posterior pituitary in response to an increased osmolarity of the blood or reduced blood volume (11). It binds a G-protein coupled receptor (V2) in the basolateral membrane of the collecting duct principal cells and triggers a signaling cascade involving adenylate cyclase, cAMP, and protein kinase A (PKA) that ultimately phosphorylates Ser256 on AQP2 C-terminal tail. This phosphorylation stimulates AQP2 containing vesicles to fuse with the apical plasma membrane (12,13) thereby increasing apical water permeability and thus increasing fluid reabsorption from the distal nephron. Water passes through the cell and back into the bloodstream via AQP3 and AQP4, which are expressed on the basolateral membrane of the same cells. Plasma vasopressin levels regulate both AQP2 and AQP3, but not AQP4 at the transcriptional level (14). Thus AQPs are subject to intracellular regulation at multiple levels that include transcription and cellular trafficking.

Congenital nephrogenic diabetes insipidus (NDI) is characterized by a patient's inability to concentrate urine in response to vasopressin. NDI is caused both by mutations in the V2 receptor and AQP2 (15,16). At least thirty NDI mutations in AQP2 have been identified, most of which are autosomal recessive and interfere with AQP2 folding resulting in misfolded proteins that are recognized by the ER quality control machinery. Autosomal dominant mutations have also been identified in which the mutant protein oligomerizes with WT thus blocking its trafficking to the plasma membrane (7).

3) Other AQP physiology

Considering the widespread expression of AQPs, it is somewhat surprising that their physiological importance in many locations is not yet clear. In the brain, AQP4 is expressed in astroglial cells at the blood brain barrier as well as the choroid plexus, the organ that produces cerebral spinal fluid (CSF). Surprisingly, AQP4-null mice show no major neurological defect (17). However, in the case of ischemic injury, or cytotoxic edema in a meningitis model, the AQP4-null mice showed a significant reduction in the expected brain swelling (17-19). Very interestingly, the AQP4-null mice also showed increased brain swelling in models of vasogenic edema including brain tumors and abscesses (18,19) suggesting AQP4 plays a key role in water transport in the central nervous system.

AQP1 is also expressed in the corneal endothelium where it regulates corneal thickness and aids in recovery from injury induced swelling (20). The corneal thickness of null mutant mice is reduced by ~20% in comparison to WT mice (20) suggesting that AQP1 also regulates the steady-state volume of these cells. AQP0 comprises about 50% of total membrane protein in lens fiber cells and is the only AQP that is believed to play a role in cell adhesion as well as water transport (7). AQP5 is involved in secretion of saliva into the acinar lumen of the salivary gland (21,22). The importance of the glycerol transporting AQPs is only beginning to be appreciated. AQP7 is expressed in the plasma membrane of adipocytes and likely plays a role in glycerol export from fat cells. AQP7 null-mice reach a much higher fat mass than the WT mice as they age (23). AQP3, important for water reabsorption in the kidney, is also expressed in the basal membrane of keratinocytes where it maintains skin hydration by transporting glycerol from the bloodstream (24).

B. Aquaporin Structure and function

1) Structural selectivity, Water vs. Glycerol

In recent years the atomic structure of several AQPs including human and bovine AQP1 (5,25),

E.coli AqpZ (26), sheep and bovine AQP0 (27,28), and the glycerol transporter GlpF (4) have become available. They demonstrate a high degree of similarity even between the E.Coli glycerol transporter GlpF and the other AQPs providing insight into the mechanism and selectivity of water transport. These structures have explained a long standing question regarding AQP function, namely their ability to conduct water molecules while efficiently excluding protons and other ions (6). Molecular dynamics simulations (29-34) have revealed a novel mechanism in which water attains a configuration aligned in single file along the conducting pathway, and permeation occurs as one water molecule displaces the next (29,30,32,34). At the center of the pore a single water molecule hydrogen bonds simultaneously with the two amino groups of the asparagine side chains of the NPA motifs. This bipolar single file order self-propagates through water-water hydrogen bonds from the two central NPA motifs, and carbonyl groups of residues 59-62 and 182-185 serve as additional hydrogen bond acceptors for hydrogen atoms of water molecules along the pathway. As a water molecule passes these NPA motifs it must “flip” adopting the opposite orientation. This introduces a barrier to proton conduction that has been confirmed by multiple computational methods (1,29,33,35-38). Glycerol transport is facilitated in much the same way through a series of hydrogen bonds, although GlpF has a wider and less hydrophobic pore than the water conducting channels ($\sim 3.5\text{\AA}$ vs. 2.4\AA at the selectivity filter) and has a smaller asymmetric periplasmic projection that is proposed to be important for optimal substrate permeation (1,39). Collectively, these studies have elucidated the mechanism of water transport at a molecular level and defined small differences in the GlpF structure responsible for substrate selectivity (water vs. glycerol).

2) AQP Gating and Regulation

Water and glycerol transport through AQPs is a passive process primarily driven by concentration gradients, with transport regulation occurring mostly at the transcriptional and trafficking levels (2,7). Recent molecular dynamics simulations have suggested that aquaporin

channels may have two distinct conformational states for Arg-189 that may reflect a closed or open water channel (1). The crystal structure of the AqpZ tetramer also demonstrates Arg-189 in two states and predicts that alternating between these two states would disrupt water transport (40). The authors propose that the driving force for the closed-to-open conformation is electrostatic in nature (40). It is unknown whether this phenomenon occurs in the lipid bilayer, and the physiological importance of these conformations remains to be determined. Thus, while it is unknown whether mammalian AQPs exhibit physiological gating properties, a plant AQP, SoPIP2;1, is known to close by dephosphorylation of two conserved serine residues or protonating a conserved histidine residue. Structural data showed that Loop D caps the water channel in the absence of these events (41).

3) Other AQP substrates

Non-traditional roles for AQPs have been described including conduction of ions and gases. AQP6 is an unusual AQP residing in the intracellular vesicles of the kidney collecting duct intercalated cells (42). When expressed in *Xenopus* oocytes it produces a very low water permeability that is increased by treatments with mercurials (43). Oocytes injected with AQP6 also showed a marked anion conductance that was stimulated by mercurials (43). Because AQP6 is expressed in intercalated cells that are involved in renal acid/base regulation, it has been proposed that pH may regulate AQP6 function. In fact, at pH <5.5 water permeability and anion conductance through AQP6 are both increased and its selectivity for anions versus cations can be reversed by the mutation K72E. Further work has shown that Asn60, which corresponds to a Gly residue in all other AQPs, is responsible for AQP6's relatively low water permeability (44).

Many AQPs including AQP1 conduct water and glycerol to varying extents (45,46). In addition, Yool et al. (47) reported that forskolin treatment of AQP1 expressing oocytes induced a cation current through a pore formed at the center of the AQP1 tetramer that is distinct from

the water conducting pores. In addition, AQP1 cation conductance was also observed using purified AQP1 reconstituted in a planar bilayer (48), but with a stoichiometry of only one ion channel per 10^7 AQP1 molecules, and the relevance of this finding remains controversial (49,50). Studies in oocytes suggest that CO_2 may also permeate AQP1 (51,52), but in vivo data have yet to corroborate this role (53). Although AQP1 can conduct ions at very low rates, AQP6 is the only AQP that has been convincingly shown to conduct ions in a physiologically relevant fashion.

C. Aquaporin Biogenesis/Folding

1) Early events of AQP translation

Synthesis of AQPs and other polytopic transmembrane proteins requires the coordination of a complex set of events. The recent high-resolution crystal structures have been instrumental in understanding the mechanism by which AQPs transport water or glycerol, but there are few details as to how AQPs reach this final structure. Here I will discuss what is known about the mechanisms by which AQPs and other polytopic membrane proteins acquire secondary and tertiary structure, are integrated into the lipid bilayer, and acquire quaternary structure (oligomerization).

AQPs are synthesized and folded in the ER along with other proteins destined for the secretory pathway. As the signal sequence encoded within the nascent chain emerges from the ribosomal tunnel, it binds Signal Recognition Particle (SRP) and stalls translation. SRP then targets the ribosome nascent chain complex to the SRP receptor at the ER membrane (54,55) and transfers the translating ribosome to the Sec61 translocation channel (56). After docking, protein synthesis resumes and the signal sequence engages the translocon and gates open the protein conducting pore to allow newly synthesized polypeptide to be translocated into the ER lumen (57-60).

The translocon is a large protein complex that includes the Sec61 $\alpha\beta\gamma$ heterotrimer, which forms the aqueous channel through which proteins are translocated (60). Translocon gating is thought to be mediated by binding and release of the ER luminal Hsp70 homolog, BiP (61). However, recent crystal structures of an archaebacterial Sec61 α suggest a model whereby the pore is gated as a short helix is displaced from Sec61 α (62). Many other proteins are physically associated with the translocon complex. Oligosaccharyltransferase (OST) attaches N-linked sugars to proteins as they emerge from the translocon (63). TRAM (translocating chain-associated membrane protein) may play a role in formation of the ribosome-translocon junction (64-66), protein translocation (65,67,68), and integration of TM segments into the lipid bilayer (69-71). Another associated protein is TRAP (translocon-associated protein), which may influence nascent chain orientation and play a role in translocation reinitiation after targeting (72). Taken together, work defining the translocon has provided new insight into the complexity of biogenesis events. However, questions remain. How does the translocation pathway maintain the ER permeability barrier? How and when do TM segments integrate into the membrane? Where do TM helices begin to assemble and how is this process influenced by the translocon and ribosome?

2) *Aquaporin Topogenesis*

Polytopic transmembrane proteins utilize the same ER machinery for their synthesis and topogenesis as secretory and bitopic proteins, but they face the additional tasks of orienting and integrating multiple TM segments into the lipid bilayer and localizing cytosolic and luminal loops. Transmembrane segments are comprised of approximately 20 nonpolar amino acids, which is long enough to span the 30Å hydrocarbon core of the lipid bilayer as an α -helix (73). The von Heijne (74,75) group has shown that hydrophobicity alone does not necessarily define a TM segment, but that there is also a strong positional preference for certain amino acids within the TM segment. For example, arginine residues are poorly tolerated at the center of a TM segment, whereas moving them further away from the lipid core is better tolerated. There is

also a strong preference for aromatic residues at the edges of TM segments, which are thought to lie along the planar interface between the membrane lipid polar head groups (74,75). A major challenge in understanding basic mechanisms of polytopic protein biogenesis is understanding how the translocon is gated to properly localize luminal and cytosolic loops.

Early topogenesis events of AQP4 follow a relatively simple model for polytopic membrane topogenesis in that primary information for topology is encoded locally within TMs that alternate in their ability to open and close the translocon (76,77). TM1 functions as a signal anchor sequence to target the ribosome nascent chain complex to the ER, gate the translocon open, and initiate translocation of the first extracellular protein domain. As TM2 enters the translocon it acts as a stop transfer sequence, gating the translocon closed and directing the next peptide loop into the cytosol. TM3 reinitiates translocation, etc. and so the topology of AQP4 is established cotranslationally with each TM segment acting alternatively as either a signal anchor (TM1, TM3 and TM5) or stop transfer (TM2, TM4 and TM6).

In contrast to AQP4, AQP1 is initially synthesized as a four-spanning protein that undergoes an internal reorientation of 3 TMs to form a six-spanning functional channel. Studies of AQP1 topogenesis have shown that while TM1 acts as an efficient signal anchor sequence, TM2 does not act as a stop transfer but rather, is transiently translocated into the ER lumen. As a result TM3 terminates translocation and initially spans the membrane in a reverse orientation to its location in the mature protein. TM4 is unable to reinitiate translocation and remains in the cytosol. TM5 and TM6 act as signal anchor and stop transfer sequences respectively. Thus AQP1 is initially synthesized as a 4-spanning protein (76-80). Interestingly, synthesis of TMs5 and 6 stimulates a reorientation of TMs2-4 such that TM3 rotates 180° and TMs2 and 4 acquire their proper topology (81). Therefore, residues in the C-terminal half of AQP1 are needed to properly orient N-terminal TM segments. Other polytopic membrane proteins also have similar biogenesis mechanisms. For example, Sec61 α contains 10 TMs and also requires

C-terminal TMs to properly position the N-terminus in the membrane (82). Similarly, TM1 in both CFTR and Kv1.3 is unable to act as a signal sequence, but requires synthesis of TM2 for targeting and integration into the lipid bilayer (83,84). Thus, polytopic proteins can utilize a variety of folding pathways to acquire their topology depending on specific translocation events that are directed by the nascent polypeptide.

In addition to facilitating the orientation of transmembrane segments the translocon machinery must allow TM segments to exit laterally into the lipid bilayer. One model predicts that the translocon channel is open laterally into the lipid bilayer and that TM segments passively exit the translocon in accordance with their hydrophobicity (85). This model is supported by the recent crystal structure of SecYE β , which shows a lateral gate within a single translocon pore that could potentially accommodate a TM segment (86-88). Photocross-linking experiments also showed that a TM can cross-link lipids almost immediately after Sec61, and that lipid cross-linking was stimulated by increasing the hydrophobicity of the TM segment (71,85). A second model predicts that TM segment integration occurs in a regulated step-wise manner in which TM segments interact with specific binding sites within the Sec61 translocon and are released only upon the entry of the next TM or termination of translation. This model also predicts that the association of multiple helices may affect the exit of the TM from the translocation machinery (70,89-91). This model is supported by evidence that TM segments cross-link specifically to translocon components in a step-wise fashion at very precise moments during synthesis (70,91). Sadlish et al. systematically investigated the movement of all six AQP4 TM segments through the translocon and showed that TM segments contact Sec61 in strict succession. Notably, after losing contact with Sec61, some helices (TM1, 3 and 5) reestablish contact at an apparently different location (89). This suggests a more complicated mechanism than hydrophobic partitioning and raises key questions regarding complex folding such as AQP1, where TM reorientation is required for protein maturation.

3) Late Folding Events of AQPs and other Polytopic Membrane Proteins

As polytopic membrane proteins, AQPs must attain their final folded structure within the plane of the lipid bilayer. Folding of secondary structure likely begins even before the nascent chain exits the ribosomal tunnel (92-94). Specifically, TM6 of Kv1.3 channel was shown to form a compact helical structure before it exited the ribosomal tunnel (94). Although the mechanism of secondary structure folding in AQPs is unknown, these observations are especially interesting in the context of AQP1 topogenesis. Would AQP1 TM2 (which slips into the ER lumen) form a helix in the ribosome exit tunnel? If so, would this folded secondary structure be maintained as the protein slipped through the translocon and into the hydrophilic ER lumen? If the helix was not formed within the ribosome, when would it fold? The tertiary folding events of AQPs are also unknown. A particularly interesting question is how and where the TM segments begin to associate in AQP1. Helical packing of transmembrane segments in the lipid bilayer is driven primarily by the hydrophobic effect and stabilized by hydrogen bonds, van der Waals forces, and sometimes salt-bridges between helices (73). Depending on where early folding takes place, the environment of the translocon could significantly influence this process.

4) Oligomerization of AQPs and membrane proteins

AQPs traffic from the ER to their final cellular destination as tetramers, and tetramerization is prerequisite for functional expression of channels at the plasma membrane (2,95). Studies have demonstrated that AQP0 (96,97), AQP1 (98,99), AQP4 (100,101), as well as AQPcic (insect AQP) (102) all exist as homotetramers. However, little is known about the structural motifs that contribute to AQP tetramerization. There is some evidence suggesting that nonfunctional channels do not tetramerize and that proper folding of the water conducting channel is necessary for oligomerization. However, it is also possible that misfolding may cause both loss of function and lack of tetramerization. Mathai and Agre showed that two residues important for water conduction, one in Loop B, A73, and one in Loop E, C189, are also critical to AQP1

oligomerization (103). Similarly, a mutation in Loop E of AQPcic (S205D) also blocked tetramerization. Work by other groups showed that there is a strong correlation between loss of water channel function and inability to oligomerize (95). Importantly, these mutations are in Loop B and E, which fold inward into the plane of the lipid bilayer to form the water conducting pore. Therefore, amino acids on these loops could not be involved in direct subunit-subunit interactions, but rather may affect tetramerization through allosteric interactions.

Polytopic membrane proteins use a variety of strategies to promote oligomerization including N-terminal recognition domains (Kv (voltage gated potassium channels)), C-terminal CLZ domains (CNG channels) and GXXXG motifs within a TM segment (GPCRs) (104,105). Kv channels are six transmembrane spanning proteins that tetramerize in the ER to form functional channels through N-terminal T1 domains. The Deutsch group demonstrated that tetramerization through T1 domains can occur even before the nascent chain is released from the ribosome (106), and that oligomerization and tertiary folding are tightly coupled events (107). Unlike the Kv channels, AQPs do not have any obvious tetramerization domains but likely associate primarily through their TM segments. Interestingly, even when the T1 domains of Kv channels are deleted, oligomerization can occur through TM segment interaction but is less efficient and not subtype specific (105,108).

There is some evidence that AQPs have different oligomerization properties. Early studies predicted that GlpF trafficked and functioned as a monomer (109,110). However, cross-linking in intact membranes and cryoelectron microscopy demonstrated that GlpF does indeed tetramerize, but unlike AQP1 the intermolecular interactions are sensitive to nonionic detergents (111,112). It has also been suggested that differences in strength of intermolecular interactions may be related to their transport selectivity (113). Interestingly, the mutations Y222P/W223L not only convert AQP1 to a glycerol conducting channel, but also alters its migration on a sucrose gradient to more closely resemble GlpF (95). One possibility is that differences in

monomer association have evolved to prevent hetero-oligomerization. This may be important in cell types where multiple AQPs are co-expressed. For example, AQP2, AQP3 and AQP4 are all expressed in kidney collecting duct principal cells, yet do not form heterotetramers (114). AQP1 and AQP3 are both expressed on the erythrocyte membrane and do not hetero-oligomerize. Because members of the AQPs are overall highly homologous, a mechanism must therefore exist to prevent hetero-oligomerization and intermolecular interactions must therefore be highly specific.

In addition to tetramerization, AQP4 forms higher order orthogonal arrays in the plasma membrane (100). Freeze-fracture electron microscopy of kidney collecting duct, muscle, and astrocytes in the brain had revealed regular square arrays of intramembrane particles called orthogonal array particles (OAPs) (100). There is some evidence that OAPs are formed in response to vasopressin stimulation and that phosphorylation of a putative PKA site at Ser111 may be involved (115). The function of these structures is unknown, although it is possible that they are used to pack AQP4 into the highest density possible to facilitate rapid water transport (100).

D. AQP Maturation and Degradation

Transmembrane proteins are uniquely exposed to cytosolic, ER luminal and lipid environments. Therefore chaperone proteins in the cellular compartments assist in the folding and maturation process. This discussion will focus on the early folding events in the ER. The cytosolic chaperones Hsp70, Hsp40, Hsp90 all have homologues in the ER although their relevance and contribution to AQP1 biogenesis has not been explored. Disulfide bonds are formed in the ER with the assistance of the chaperone like protein disulfide isomerase (PDI) (116). However, AQPs do not have known disulfide bonds and have not been shown to interact with PDI. Several family members (e.g. AQP1, AQP4 and AQP2) have N-linked glycosylation consensus sites in their extracellular loops that are modified to different extents (76,77,117-120).

1) AQP N-linked glycosylation

Proteins undergo N-linked glycosylation as the nascent chain is translocated into the ER by oligosaccharyl transferase (OST). OST transfers a core sugar unit to an Asn residue in the consensus sequence NxS/T (x being anything but proline) (121). Blocking N-linked glycosylation with tunicamycin results in protein aggregation and induces the unfolded protein response (UPR), demonstrating the importance of N-glycans to productive protein folding (122-124). The effect of N-linked glycosylation on AQPs is not well known. Both AQP1 and AQP2 are inefficiently glycosylated, presumably because of the close proximity of the consensus sites to the membrane (~7 residues for AQP2 and ~5 in AQP1). Efficient glycosylation of a consensus site generally requires a distance of at least 12-14 residues from the membrane for accessibility to OST (125). Interestingly, the mutations N49M/K51L in AQP1 TM2, which converts TM2 to a stop transfer sequence, also completely block glycosylation (77). This is most likely because when extracellular loop 1 (which contains the consensus site) is now tightly tethered to the ER membrane, it is sterically inaccessible to OST. Blocking glycosylation of either AQP1 or AQP2 with tunicamycin has no effect on the trafficking or function of the channels (126,127), but AQP2 glycosylation may be important for trafficking of the tetramer out of the Golgi (128). Mutating the glycosylation site N123Q allowed AQP2 to tetramerize and still function in oocytes, but in mammalian cells this mutant colocalized with Golgi markers (128). The effect of N-linked sugars on AQP2 therefore, seems to be subtle and may involve one glycosylated subunit “escorting” unglycosylated monomers through the Golgi. The side effects of tunicamycin treatment could be responsible for the previous observations.

Glycosylated proteins are substrates for another family of ER chaperones (129,130). Calnexin and calreticulin (CNX/CRT), which are the major ER lectins, bind N-linked glycans likely to prevent aggregation and degradation (131). Different glycosylation profiles have been reported

for WT versus NDI causing AQP2 mutants (119,132,133) in membranes derived from microinjected oocytes. AQP2 mutants (T126M, A147T and R187C) were all found to be predominantly glycosylated in contrast to WT, which was primarily unglycosylated (132). A more recent study showed that the glycosylated form of AQP2 T126M had a longer half-life than the unglycosylated isoform (119) suggesting that the addition of N-linked sugars may promote proper folding. Properly folded proteins must be released from CNX/CRT prior to trafficking to their final cellular destination. However, misfolded conformers undergo additional rounds of CNX/CRT binding by attachment and removal of a terminal glucose residue (131,134).

2) AQP Degradation/ERAD

The majority of NDI-causing AQP2 mutants fail to fold in the ER and are subsequently degraded by the 26S proteasome (15). This process, called ER associated degradation (ERAD), can be broken down into three parts: recognition and targeting, dislocation from the ER to the cytosol, and proteasomal cleavage. Proteins can be targeted due to folding mutations (CFTR and AQP2) (15,135), failure to oligomerize (TCR α , possibly AQPs) (136,137), viral targeting (human cytomegalovirus US2 and US11 proteins target class I MHC molecules) (138), and cellular regulation (Hmg2p) (139). Interestingly, many chaperones involved in protein folding and maturation (BIP, CNX/CRT, PDI, cytosolic hsp70-hsp90 family) are also implicated in targeting proteins for degradation (140,141). For Hsp70 this is accomplished by co-chaperones (i.e. Bag1, CHIP) that can bring substrate to the proteasome and/or function as ubiquitin ligases (142).

Proteins are tagged for degradation by the covalent addition of ubiquitin. This is carried out via an enzyme cascade involving E1 (ubiquitin activating), E2 (ubiquitin conjugating) and E3 (ubiquitin ligase) enzymes transferring ubiquitin to the target (142). It has been proposed that ERAD proteins with luminal versus cytosolic lesions are recognized by different ubiquitinating

complexes (143-145). AQP2 degradation mutants have been localized to both the cytoplasmic and luminal surfaces of the protein, although it is unknown which set of factors (15) recognizes them for degradation.

Few details of AQP ERAD have been studied, therefore we will discuss ERAD in more general terms. Membrane proteins targeted for ERAD require specific mechanisms for ER extraction that involve active unfolding by ATPases. Evidence suggests that the most likely candidate for dislocation of secretory proteins is the Sec61 translocon (138,146). Additional proteins including Derlin-1 (Der1p in yeast), and the ligases Hrd1p and Doa10p are also involved (142,147-151). The AAA ATPase, p97 (Cdc48 in yeast) together with the 19S cap of the 26S proteasome likely provide the force needed to extract protein's TM segments from the lipid bilayer (148,152-154).

The 26S proteasome is a large, multicatalytic cytosolic protease complex responsible for the degradation of polyubiquitinated ERAD proteins. A barrel shaped protein with proteolytic machinery, the 26S proteasome consists of a 20S catalytic core flanked by two 19S CAPs. The 20S core contains the active sites for three different types of protease activity chymotrypsin-like, trypsin-like and peptidyl-glutamyl-peptide like hydrolyzing activity (155-158). The 19S cap has a base and a lid domain which collectively bind and unfold degradation substrates, cleave off the ubiquitin chains, gate the 20S core open, and finally drive the unfolded protein into the catalytic core where it is degraded (159). Details of the AQP degradation pathway are unknown; however, NDI can be a direct result of AQP2 degradation. Because of this understanding the details of AQP2 degradation may reveal a pharmacological target to potentially stabilize the AQP2 mutants.

E. Specific Aims Addressed in this Thesis

The body of work presented here focuses on understanding the early molecular events of AQP

biogenesis and folding. While crystal structures have provided important insight into the architecture and water channel function of AQPs, very little is known about the events that contribute to the folding process which gives rise to a mature water channel in the lipid bilayer.

1) How do NDI-causing mutations in AQP2 affect early biogenesis events?

Many of the NDI causing AQP2 mutants identified have been characterized as trafficking mutants that are retained in the ER and degraded by the ERAD machinery (15). These mutations are found throughout the protein in both luminal and cytosolic loops as well as in transmembrane segments. The specific folding defects that result and subsequently lead to recognition by the ER quality control machinery are unknown.

a. Do NDI causing AQP2 trafficking mutants affect early AQP2 topogenesis?

In Manuscript 1 I investigated whether trafficking mutants might disrupt early folding events that generate proper topology. Mutations at or near TM segments may affect the early topogenesis events and therefore the final folded structure of a protein. Four NDI-causing AQP2 trafficking mutants were examined (T126M, A147T, C181W and R187C) and early AQP2 biogenesis was investigated using a series of truncated proteins fused to a C-terminal translocation reporter. Manuscript 1 discusses the results of these experiments. Briefly, I found that none of the four AQP2 trafficking mutants we identified had a noticeable effect on early topogenesis, although we did note an unexpected role for N-linked glycosylation in stabilizing mutant proteins.

b. How does N-linked glycosylation affect AQP2 trafficking mutants?

In membranes from microinjected oocytes, the AQP2 mutants (T126M, A147T and R187C) were predominantly glycosylated whereas WT AQP2 was mostly unmodified (132). Manuscript 1 describes how N-linked glycosylation increases the stability of AQP2 trafficking mutants (T126M, A147T, C181W, and R187C) in comparison to WT protein.

2) *Is AQP1's unique biogenesis mechanism conserved in diverse expression systems?*

Studies of AQP1 in cell-free systems and microinjected *Xenopus* oocytes revealed a novel mechanism whereby only four of its six transmembrane segments span the lipid bilayer as the protein is synthesized (78). This four-spanning biogenesis intermediate later matures to the final six-spanning structure in the ER via an internal rotation that repositions TM2 and TM4 in the plane of the membrane (77,81). However, work by Dohke and Turner (160) suggested that biogenesis in HEK-293 cells was cotranslational in nature and did not involve a four-spanning intermediate. If mammalian cells handle topological information in a fundamentally different fashion, this would be of great importance for a number of reasons. First, what differences exist in the translocon machineries that would handle topological signals differently? Second, the in vitro RRL system and oocyte expression have been used to analyze the behavior of many polytopic membrane proteins. Therefore if these systems handle topological information differently it would be very important for examination of data generated in these systems. Manuscript 2 describes a systematic study undertaken to explain these divergent results. Specific questions asked are outlined below.

a. Does the reporter fusion site affect AQP1 topogenesis?

Residues within 10-15 amino acids of a transmembrane segment can effect the topology of TMs in the membrane (161). Dohke et al. (160) fused a C-terminal reporter (β -subunit of the H/K ATPase) directly after transmembrane segments in truncated versions of AQP1 whereas earlier studies fused a different reporter further downstream (76,78,81). I therefore tested whether the site of fusion was responsible for topologic differences observed. An examination of these sites in oocyte and mammalian systems is described in manuscript 2 and showed that fusion site did not explain topological differences.

b. Are glycosylation reporters and protease protection assays equivalent for assaying

the orientation of a polypeptide?

C-terminal translocation reporters have been widely used to study membrane protein topology and biogenesis (78,82,162-169). Two common methods to determine topology of the reporter are protease protection and N-linked glycosylation. Studies of AQP1 topology in cell-free and oocyte systems used protease protection assays to determine the reporter location (77,78,81), whereas the studies in HEK-293 cells assayed the glycosylation status of the reporter derived from the β -subunit of the H/K ATPase (160). Because these assays are widely used, it is important to determine if they yield equivalent results. Manuscript 2 describes studies comparing these systems. Briefly, I found topological results did not vary at early time points, but at later time points HEK-293 cells gave results similar to Dohke (160).

c. When multiple topological isoforms are generated do they have similar stabilities?

My studies revealed that in both oocyte and HEK-293 cells certain truncated AQP1 constructs generated proteins that exhibited two alternate topologies. These isoforms are “stuck” at a premature stage of folding and thus are unable to mature because they lack C-terminal residues. I therefore tested whether the ER degradation machinery may handle these isoforms differently. Manuscript 2 describes pulse-chase analysis examining the stability of these isoforms in the oocyte and mammalian cell systems. I found a significant difference in the stability of the major and minor isoform and conclude that the differences in topology were primarily caused by very rapid degradation of the 4-spanning AQP1 isoform such that at late time points only the (minor isoform) 6-spanning remained.

d. Are both isoforms stable in the context of full-length AQP1?

I observed very different stabilities for topological isoforms of truncated AQP1 fusion proteins. For fusion proteins truncated after TM3, the predominate 2- spanning

isoform was much less stable than the 3-spanning. I therefore wanted to determine if this differential stability was a consequence of the missing C-terminal information that is critical for folding or could be more generally applied to full-length AQP1. In the latter case only the small percentage of protein that is cotranslationally synthesized as a 6-spanning protein would be stable, but the majority of newly synthesized protein (4-spanning) would be degraded. Manuscript 3 describes pulse chase studies that show full-length AQP1 is stable and that differential stability is a consequence of using truncated proteins.

In summary, Manuscript 2 describes a systematic analysis demonstrating that oocyte and mammalian cell expression systems handle topological information in the same manner. I conclude that AQP1 topogenesis is conserved across three expression systems (*in vitro* RRL, oocyte, and mammalian) and proceeds through a 4-spanning intermediate.

3) How is AQP1 topological maturation achieved in the ER?

As previously described, AQP1 is initially synthesized as a 4-spanning immature protein that is converted to the mature 6-spanning channel (77,78,81). The mechanism for protein reorientation is unclear, although reorientation of N-terminal TMs requires synthesis of C-terminal TMs. Manuscript 3 defines the role of specific residues within the AQP1 N- and C-termini that are responsible for this unique biogenesis process. The questions we specifically addressed are outlined below.

a. What is the role of N-terminal residues Asn49 and Lys51 in AQP1 function and folding?

Previous work showed that two residues adjacent to TM2 (Asn49 and Lys51) are responsible for its poor stop transfer function. Replacing these residues with the corresponding AQP4 residues (N49M/K51L) converted TM2 into an efficient stop transfer, but destroyed AQP1 water channel function (77). Manuscript 3 describes the

relative contribution of these residues to AQP1 function and early biogenesis. I found that Asn49 or Lys51 exhibited synergistic effects on TM2 stop transfer activity.

Functional studies further showed that K51L did not alter water channel function, but the mutation N49M reduced function by ~50%.

b. What C-terminal structural features are required for N-terminal (TMs2-4) reorientation?

Specific interactions between the N-terminal residues Asn49 and Lys51 are important for protein maturation and folding. I therefore used the available crystal structure to identify a residue, Asp185, in the C-terminal portion of the protein that may productively interact with Asn49 and Lys51. Manuscript 3 describes functional assays that show how these three residues interact. I found that in contrast to the mutations in the C-terminus of AQP1 (i.e. N49M/K51L), D185N had WT water permeability. Strikingly, mutating Asp185 to the corresponding residue from AQP4 (D185N) along with the other two mutations in TM2 (N49M/K51L) restored function of the protein.

c. What is the contribution of N49, K51 and D185 to monomer folding and tetramer assembly?

Functional analysis and examination of the crystal structure for the AQP1 monomer and tetramer predicted that association between these residues was both intra- and intermolecular. Specifically, we tested whether Asn49 and Asp185 hydrogen bond within the monomer whereas Lys51 and Asp185 associate intermolecularly between adjacent monomers. In Manuscript 3 we examined how these mutations affected intermolecular interactions using sucrose gradients and native coimmunoprecipitations. These studies validated our model and I conclude that the intramolecular association between Asn49 and Asp185 was more crucial to channel function than the intermolecular interaction between Lys51 and Asp185. However, these interactions have a cumulative effect on the folding of a functional water channel.

III. Results

A. Manuscript 1

Evidence for Stabilization of Aquaporin-2 Folding Mutants by N-linked Glycosylation in the Endoplasmic Reticulum.

Teresa Buck, Joel Eledge, and William R. Skach

Molecular Medicine Division, Department of Medicine, Oregon Health Sciences University,
Portland Oregon. 97201

American Journal of Physiology-Cell Physiology, Vol 287, pp C1292-C1299, 2004

Address Correspondence to:

William R. Skach, M.D.
Division of Molecular Medicine
Oregon Health Sciences University
3181 SW Sam Jackson Park Rd.
Portland, Oregon 97201

Ph: 503-494-7322

Fax: 503-494-1933

Email: skachw@ohsu.edu

Abbreviations: AQP, Aquaporin; ER, Endoplasmic reticulum; NDI, Nephrogenic Diabetes Insipidus; OST oligosaccharyltransferase; TM, transmembrane; WT, wild type

1. Abstract

Aquaporin-2 (AQP2) is the vasopressin sensitive water channel that regulates water reabsorption in the distal nephron collecting duct. Inherited AQP2 mutations that disrupt folding lead to nephrogenic diabetes insipidus by targeting newly synthesized protein for degradation in the endoplasmic reticulum (ER). During synthesis, a subset of WT AQP2 is covalently modified by N-linked glycosylation at residue Asn123. To investigate the effect of glycosylation, we expressed WT and four NDI-related mutants in *Xenopus laevis* oocytes and compared stability of glycosylated and non-glycosylated isoforms. In all constructs, approximately 15-20% of newly synthesized AQP2 was covalently modified by N-linked glycosylation. At steady state, however, core glycosylated WT protein was nearly undetectable, whereas all mutants were found predominantly in the glycosylated form (60-70%). Pulse chase metabolic labeling studies revealed that glycosylated isoforms of mutant AQP2 were significantly more stable than their non-glycosylated counterparts. For non-glycosylated isoforms, the half-life of WT AQP2 was significantly greater (>48 h) than mutant AQP2 (T126M 4.1 ± 1.0 h, A147T 4.2 ± 0.60 h, C181W 4.5 ± 0.50 h, R187C 6.8 ± 1.2 h). This is consistent with rapid turnover in the ER as previously reported. In contrast, the half-lives of mutant proteins containing N-linked glycans were similar to WT ($T_{1/2} \approx 25$ h), indicating that differences in steady state glycosylation profiles are caused by increased stability of glycosylated mutant proteins. These results suggest that addition of a single N-linked oligosaccharide moiety can partially compensate for ER folding defects induced by disease related mutations.

Keywords: ERAD, Nephrogenic Diabetes Insipidus, oocytes, ER associated degradation

2. Introduction

Nephrogenic diabetes insipidus (NDI) is characterized by the inability of the collecting duct to concentrate urine in response to stimulation by 8-arginine vasopressin (AVP). Most cases of congenital NDI are caused by mutations in one of two key proteins, the vasopressin (V2) receptor, or aquaporin-2 (AQP2) (15,170). AQP2, a member of the Major Intrinsic Protein (MIP) family, (171) is a hydrophobic protein of approximately 29 kDa that exists as a homotetramer in cell membranes. It contains six transmembrane segments and two inverted repeats (NPA motifs) that are thought to fold inward within the plane of the membrane to form the water selective pore (4,5). AQP2 synthesis, membrane insertion, folding and maturation is facilitated by biosynthetic machinery in the endoplasmic reticulum (ER). Prior to its export from the ER, AQP2 must be properly folded and assembled into tetramers (128).

More than 26 AQP2 mutations have been causally linked to NDI, the majority of which are autosomal recessive and disrupt AQP2 function by altering stability of the immature protein (15). Four of these mutations T126M, A147T, C181W, and R187C have been extensively characterized in *Xenopus* oocyte, yeast, and mammalian systems and shown to exhibit defective intracellular trafficking and degradation by the ER-associated degradation (ERAD) pathway (117,119,172-175). While the precise mechanism of degradation remains unknown, it is believed that these mutations disrupt AQP2 folding and that the misfolded proteins are subsequently recognized by ER quality control machinery, ubiquitinated and degraded by the 26S proteasome (176,177). Certain mutations such as R187C and C181W that are located near the 2nd NPA box, disrupt the water conducting pathway (16,172,173,178), whereas other mutants, T126M and A147T, retain at least moderate water channel function (132,172,174,178,179). Thus, channel function per se is not a criterion for recognition by ER quality control, raising the possibility that restored trafficking of mutant proteins may partially

or completely correct the disease state.

Several AQPs (e.g. AQP1, AQP4 and AQP2) contain N-linked glycosylation consensus sites in their extracellular loops. Some of these sites are inefficiently recognized during protein synthesis by oligosaccharyltransferase (OST) to generate a mixture of glycosylated and nonglycosylated species (76,77,117-120,132,180). While N-linked glycosylation may be important for AQP2 trafficking through the Golgi (128), it has little effect on stability of WT AQP1 or AQP2, as both proteins exhibit normal function when glycosylation sites are removed or blocked by tunicamycin (126,127). An unusual feature of AQP2 is that different glycosylation patterns have been reported for wild type and mutant proteins (119,132,133). In membrane preparations derived from microinjected *Xenopus* oocytes, AQP2 mutants (T126M, A147T, and R187C) were found to be predominantly glycosylated whereas, WT AQP2 was present primarily in a non-glycosylated form (132). It was initially proposed that these differences were due to deglycosylation of WT protein as it passed through the Golgi, thus reducing the apparent glycosylation efficiency. However, it has recently been shown that WT and R187C AQP2 are both inefficiently glycosylated in oocytes (128). In addition, non-glycosylated forms of T126M AQP2 exhibited a shortened half-life compared to glycosylated forms when expressed in cultured hepatoma cells (119). Based on these findings we sought to determine whether N-linked glycosylation confers a general effect on AQP2 stability in the ER compartment. Using *Xenopus* oocytes and pulse-chase metabolic labeling, we examined four mutant AQP2 proteins that are known to cause NDI, T126M, A147T, C181W, and R187C. At early time points, all proteins were cotranslationally glycosylated to a similar extent as WT AQP2 (~20%). Interestingly, glycosylated forms of all mutant proteins were significantly more stable than their nonglycosylated counterparts, thus causing the percentage of glycosylated mutant protein to increase progressively over time. In contrast, N-linked oligosaccharides had little effect on the stability of the wild type protein. These findings indicate that the presence of N-linked sugars exert a generalized effect on AQP2 ERAD by

delaying the recognition of misfolded substrate by ER quality control machinery.

3. Materials and Methods

cDNA Construction. Plasmid pSP64.myc-AQP2 was constructed using a sense oligonucleotide (CCGGGTACATGTCTGAACAAAACTTATTTCTGAAGAAGATCTGTGGGACCTCCGCTCCATA) to fuse the 10 residue myc epitope (EKQEEDL) (181) to the N-terminus of AQP2 (provided by A.Verkmán) using standard cloning methods described previously (182). The missense mutants myc-AQP2 T126M, myc-AQP2 A147T, myc-AQP2 C181W, and myc-AQP2 R187C were constructed from pSP64.myc-AQP2 using site directed mutagenesis (PCR overlap extension (183)). The chimeric clones thus encode N-terminally myc-tagged AQP2, AQP2T126M, AQP2A147T, AQP2C181W, and AQP2R187C. All cloned fragments were verified by DNA sequencing. Truncated fusion proteins were generated by ligating the COOH terminal 142 residues from bovine prolactin (182) to the indicated codons in AQP2. This was accomplished by using a sense oligonucleotide (oligo 242 base pairs upstream of the SP6 promoter) and antisense oligonucleotides encoding a BstE II restriction site at AQP2 codons Val131 (GAGCTCGGTCACCACCGCCTGGCCAGCCGT), Pro157(AGCAGGGGTCACCGGGTCTCTCCGCGGCG), Trp202 (GATCCCAGGTCACCCAGTGGTCATCAAAT), and Arg267 (GGCCTTGGTCACCCGTGGCAGGCTCTGC). Fragments were digested with Nhe I and BstE II and ligated 5' to a Nhe I/BstE II -digested vector, S.LST.gG.P, as described previously (182). The chimeric constructs thus encode the AQP2 coding sequence specified and the reporter domain (diagrammed in Figure 6A). Importantly, this reporter contains no intrinsic topogenic information and faithfully follows the direction of upstream topogenic determinants (78,184).

Xenopus laevis Expression: mRNA was transcribed in vitro with SP6 RNA polymerase (New

England Biolabs) using 2 μ g of plasmid DNA in a 10 μ l volume at 40° C for 1 hour as previously described (184). Aliquots were used immediately or frozen in liquid nitrogen and stored at -80° C. 2 μ l of transcript was mixed with 50 μ Ci of ³⁵S-methionine (0.5 μ l of a 10X concentrated Tran ³⁵S-label, ICN Pharmaceuticals, Irvine, CA) and injected into stage VI *Xenopus* oocytes (50 nl/oocyte). Oocytes were incubated at 18° C in MBSH (88 mM NaCl, 1 mM KCl, 24 mM NaHCO₃, 0.82 mM MgSO₄, 0.33 mM Ca(NO₃)₂, 0.41 mM CaCl₂, 10 mM HEPES pH 7.4, 50 μ g/ml Gentamicin, 100 units/ml Penicillin, 100 mg/ml Streptomycin Sulfate). For pulse labeling experiments, oocytes were homogenized at indicated time points and radiolabeled protein was immunoprecipitated prior to SDS-PAGE. For pulse-chase experiments, media was removed from oocytes 1.5 hours after injection and replaced with MBSH supplemented with 2 mM unlabeled methionine. Incubation was then continued for an additional 30 min to equilibrate unlabeled methionine and complete translation of labeled AQP2 protein (185). This time was used as our initial time-point for pulse chase studies. At times indicated, oocytes (5/lane) were flash frozen in liquid nitrogen and stored at -80° C.

Immunoprecipitation: Groups of oocytes were thawed, solubilized in 100 μ l 0.1 M Tris-Cl pH 8.0, 1% SDS in a 1.5 ml microcentrifuge tube and incubated at 37° C for 30 minutes. Homogenates were then diluted in 10 volumes of Buffer A (1% Triton-X 100, 100 mM Tris-Cl pH 8.0, 100 mM NaCl, 5 mM EDTA), incubated on ice for 30 min, and insoluble material was removed by centrifugation (16,000 x g for 15 min at 4° C). 5.0 μ l of protein-A Affigel (BioRad, Hercules, CA) and 0.75 μ l of antibody Myc-9E10 (mouse ascites fluid, (181)) was added. The sample was mixed at 4° C for 10 hours prior to washing 3X with Buffer A and twice with 0.1 M NaCl and 0.1 M Tris-Cl pH 8.0. Samples were analyzed by SDS-PAGE, EN³HANCE (Perkin-Elmer Life Sciences, Boston, MA) fluorography and autoradiography. Band signals were quantitated using a BioRad personal Molecular PhosphorImager Fx (Kodak screens, Quantity-1 software).

PNGaseF Digest: After SDS solubilization at 37° C, samples were split and immunoprecipitated as described. Protein-A beads were washed, resuspended in 15 µl 0.1 M Tris-Cl pH 7.5 containing 0.3 µl PNGaseF (New England Biolabs, Beverly, MA) and incubated for 3 hrs at 37° C prior to SDS-PAGE.

Immunoblotting oocyte membranes: Oocytes were injected as above but without ³⁵S-methionine, incubated at 18° C for the times indicated, and homogenized in 0.25 M sucrose, 50 mM KOAc, 5 mM MgCl₂, 1 mM DTT, and 50 mM Tris-Cl pH 7.5 (5µl/oocyte). Homogenates were centrifuged 800 x g for 5 min. The supernatant was layered onto 400 µl of 0.5 M sucrose, 50 mM KOAc, 5 mM MgCl₂ and 50 mM Tris-Cl, pH 7.5 on top of 100 µl of 1.8M sucrose in the same buffer. Samples were centrifuged at 186,000 x g for 10 min. Total cellular membranes were collected at the 0.5M-1.8M sucrose interface, diluted with 400 µl of 50 mM KOAc, 5 mM MgCl₂ and 50 mM Tris-Cl, pH 7.5, and repelleted through 0.5 M sucrose, 50 mM KOAc, 5 mM MgCl₂ and 50 mM Tris-Cl, pH 7.5 by centrifugation at 186,000 x g for 10 min. Membrane pellets were solubilized and separated by SDS-PAGE and transferred to nitrocellulose. Blots were blocked overnight using 5% non-fat dry milk, probed for one hour with Myc 1-9E10 antibody (1:5,000, mouse ascites fluid), washed 6 times with 10mM Tris-Cl pH 8.0, 150 mM NaCl, .1% Tween-20 and probed with goat anti-mouse IgG HRP antibody (BioRad, Hercules, CA) 1:10,000 for one hour and washed as before. Bands were detected by ECL using pico west Supersignal (Pierce, Rockford, IL) per manufacturer's instructions.

Half-Life determination: To correct for variation in oocyte equilibration with the unlabeled methionine in the media, the 1.5 h time-point was normalized to 100% protein. Half-lives for the total and non-glycosylated protein were calculated assuming steady state first order elimination kinetics by fitting data to the equation $A=A_0e^{-kt}$ using Prism graphing software. Goodness of fit for the regression was analyzed by determining an F-statistic to determine the

ability of the regression equation to fit the data (SigmaPlot). Data with a $P < 0.05$ was eliminated from the calculations (two experiments out of 26). All data is shown as mean \pm SEM with a minimum of 5 experiments. Due to the stability of WT protein, an accurate half-life could not be calculated but was estimated to be >48 h. Data for the glycosylated protein did not follow first-order degradation kinetics and so half-lives were estimated from Figure 5B.

4. Results

Xenopus oocytes efficiently express a wide variety of AQP water channels and have been extensively used to study AQP2 function and trafficking (16,117,133,172,186). Importantly, oocytes also faithfully reconstitute quality control events at the level of the ER (185). Although this quality control is sometimes less stringent than in mammalian cells, oocytes maintain the ability to discriminate and degrade mutant proteins via the ERAD pathway and therefore provide a useful system for examining early events of protein biogenesis (185). We expressed myc-tagged versions of WT and four mutant AQP2 proteins (T126M, A147T, C181W, R187C) in microinjected oocytes to examine the role of N-linked glycosylation on AQP2 stability. These mutations were chosen because they exhibit similar trafficking defects in multiple expression systems including oocytes and because they reside in different regions of the protein, extracellular loop 2, intracellular loop 2, and extracellular loop 3, (see Figure 1) and therefore likely disrupt different aspects of AQP2 folding.

WT and Mutant AQP2 proteins exhibit different glycosylation profiles. AQP2 N-glycosylation status was initially examined using total oocyte membranes that were collected either 2 hr after injection when the proteins are predominantly localized to the ER (185) or 48 hr after injection when WT protein should be maximally expressed at the plasma membrane ((174,187). Immunoblot analysis revealed that both WT and mutant proteins were present as two distinct isoforms, a major band of 29 kDa and a minor band of 32 kDa (Figure 2A).

PNGaseF digestion confirmed that the difference in size was due to the covalent attachment of a single N-linked high mannose core oligosaccharide (Figure 2B). Initial glycosylation efficiency for both WT and mutant proteins was 10-20%. Thus, the only N-linked AQP2 consensus site, residue Asn123, which is located in extracellular loop 2, is relatively inaccessible to oligosaccharyltransferase (OST). Forty-eight hours after injection a distinctly different glycosylation pattern was noted. Wild type protein remained predominantly in the non-glycosylated 29 kDa form, while mutant proteins were found mainly in the glycosylated form. These results were observed for all four mutants regardless of the site of mutation. Thus they appear to represent a general pattern independent of the specific folding defect.

Pulse chase analysis of AQP2 synthesis in *Xenopus* oocytes. To determine why mutant and WT AQP2 exhibited different glycosylation profiles at later time points, pulse chase metabolic labeling studies were performed. Oocytes were co-injected with mRNA and ³⁵S-methionine to label newly synthesized protein. Previous studies have demonstrated that the ³⁵S-methionine is efficiently taken up into the aminoacyl-tRNA^{met} pool and incorporated into protein (185). Ninety minutes after injection, oocytes were placed in fresh media containing 2 mM unlabeled methionine and incubated for an additional 30 minutes (pre-chase period) to allow uptake and equilibration of cold methionine and to complete synthesis of labeled AQP2 polypeptides. Groups of oocytes were then homogenized at indicated time points, immunoprecipitated with anti-myc antibody, and analyzed by SDS-PAGE and autoradiography. As expected, initial glycosylation efficiency was low; ~20% of AQP2 was recovered as the 32 kDa glycosylated species for all constructs tested (Figure 3A). Non-glycosylated WT AQP2 was remarkably stable throughout the chase period, whereas its glycosylated form was stable for 10 h and then gradually decreased in intensity. This late disappearance likely represents Golgi processing of high mannose core carbohydrates, although it has been difficult to consistently detect Golgi processed forms in oocytes because of their low abundance and heterogeneous appearance (187). It is also possible that some

WT AQP2 may pass through the Golgi without further modification. For mutant proteins, non-glycosylated forms were rapidly degraded (after a brief lag period) with little protein detected after 24 h. Interestingly, glycosylated isoforms of mutant proteins were significantly more stable than their non-glycosylated counterparts, suggesting that the increase in AQP2 glycosylation observed at late time points primarily resulted from increased stability of mutant AQP2 induced by N-linked glycosylation.

AQP2 trafficking mutants display a reduced half-life in *Xenopus* oocytes. Data from multiple pulse chase experiments was quantitated by phosphorimaging to determine the half-lives of WT and mutant AQP2 proteins. The half-life of each protein was then determined by fitting a standard curve for first order degradation kinetics as described in methods. Data shown in Figure 3B clearly demonstrates a decrease in stability of mutant proteins compared to WT AQP2. Wild type protein had a half-life of >48 h, whereas the half-lives for mutant proteins were approximately 7.5±1.4 h (T126M), 7.1±1.6 h (A147T), 8.0±1.0 h (C181W), and 12 ±2.2 h (R187C).

We next examined the fraction of glycosylated proteins as a function of time (Figure 4). Two hours after injection, WT and mutant proteins were all glycosylated to the same extent consistent with Figure 2. All mutants showed a progressive increase in the percent glycosylation from approximately 20% to 65-70%, whereas the fraction of glycosylated WT protein remained nearly constant. Taken together, these results further demonstrate that N-linked glycans have a markedly different affect on mutant and WT AQP2.

Glycosylation stabilizes mutant protein. We next determined the half-lives of the 29 kDa and 32 kDa isoforms independently. Nonglycosylated WT AQP2 has a significantly longer half-life (>48 h) than any of the unglycosylated mutant proteins ($T_{1/2}$ of T126M=4.1±1.0 h, A147T= 4.2 ±0.60 h, C181W=4.5±0.50 h, R187C=6.8±1.2 h, Figure 5A). In contrast, the

half-life of glycosylated mutant proteins (15-26 h) was very similar to WT (20 h), and in each case, the glycosylated isoform was more stable than its non-glycosylated counterpart (Figure 5B). We also noted that for R187C there was an absolute increase in intensity of the 32 kDa band during early time points. The reason for this remains unclear but suggests that oligosaccharides may continue to be added to R187C polypeptide after synthesis is complete (discussed below).

AQP2 mutations do not affect cotranslational glycosylation and early folding events.

Although our data support a model in which N-linked glycosylation confers increased stability on mutant (and misfolded) AQP2 proteins, an alternative possibility is that these mutations might influence glycosylation on a subset of proteins that had already achieved a more stable conformation at the time of glycosylation. AQP topology is directed by topogenic determinants (e.g. signal anchor and stop transfer sequences) that target the ribosome nascent chain complex to the ER membrane, initiate and terminate translocation of peptide loops, and integrate TM segments into the lipid bilayer (77,78,81,160). Because N-linked glycosylation normally occurs cotranslationally as peptide loops translocate into the ER lumen, we tested whether accessibility of Asn123 by OST might be affected by NDI-related mutations that altered AQP2 folding.

Plasmids encoding WT and mutant AQP2 proteins were truncated at codons Val131, Pro157, Trp202, and Arg267 and ligated to a C-terminus reporter to generate fusion proteins containing 3, 4, 5 or 6 of the AQP2 TM segments (Figure 6A). In vitro transcribed mRNA was expressed in *Xenopus* oocytes, and reporter-containing polypeptides were recovered by immunoprecipitation and analyzed by autoradiography. These experiments revealed that WT and mutant constructs at any given truncation site were cotranslationally glycosylated to the same extent even though constructs encoded only a portion of the AQP2 protein. Thus synthesis of TMs 4, 5 and 6 are not required for glycosylation of Asn123 to occur. Moreover,

glycosylation of Asn123 occurs even before mutant residues are synthesized, confirming that AQP2 mis-folding must be a late event relative to the timing of initial oligosaccharide attachment. Of note, fusion proteins truncated at residue Val131, just prior to TM4 were more efficiently glycosylated (~50%, Figure 6B). This occurs because removal of TM4 increases the accessibility of Asn123 to OST by increasing its tethered distance from the ER membrane (see Figure 1 and(77)). Given that the fusion proteins lack one or more TM segments (TMs 5 and or 6), glycosylation observed here is not dependent on acquisition of tertiary folded structure.

Finally, to confirm that the extent of glycosylation observed (20%) reflects cotranslational attachment of oligosaccharide rather than posttranslational addition, we compared the timing of AQP2 glycosylation relative to protein synthesis. For these experiments, oocytes were harvested at very short time intervals following injection of mRNA and ³⁵S-methionine. Full length radiolabeled AQP2 first appeared 10-20 minutes after microinjection consistent with the time required for ³⁵S-met incorporation into the aa-tRNA pool and synthesis of the 29 kD protein (Figure 6C and D). Importantly, appearance of N-linked glycosylated species coincided precisely with the completion of synthesis, and levels of glycosylation (12-14%) are in good agreement with those observed at the initiation of the pulse chase experiments (See Figure 2). Due to technical constraints and the low abundance of protein present at these short time points, it was not possible to determine whether partial length translation intermediates were also glycosylated. However, taken together, our data support a model in which cotranslational attachment of N-linked sugars is primarily dictated by steric accessibility of the consensus site as ECL2 topology is established and not by the presence of specific mutations.

5. Discussion

The current study examines the effect of N-linked glycosylation on the early ER quality control of WT and mutant forms of the vasopressin regulated water channel AQP2. Because

AQP2 is inefficiently glycosylated in most cell types including mammalian cells, it affords a unique opportunity to directly compare, in the same cell, the fate of identical polypeptides that differ only in the presence of a single N-linked oligosaccharide. In the case of WT protein, non-glycosylated isoforms have relatively long half-lives that likely reflect productive processing and trafficking out of the ER compartment. This finding is consistent with previous studies demonstrating that glycosylation is not required for AQP function; some AQPs lack N-linked consensus sites (AQP6), while others have extracellular sites that are not utilized (AQP4 (76)). Still others, such as AQP1, AQP2, and AQP3, have sites that are utilized in only a subset of synthesized proteins (120,126,188), and removal or blockage of N-linked glycosylation from has no discernable effect on function (126-128). In contrast, our studies of NDI mutants provide evidence that the effects of AQP glycosylation can be both subtle and complex. In *Xenopus* oocytes, each of the four NDI mutations studied had a dramatic effect on the stability of AQP2 in the ER, decreasing the half-life of mutant proteins by >4 fold compared to WT. This was expected given that all mutants were chosen based on previous reports showing defective trafficking and rapid degradation in the ER (133,172,179). When stability of glycosylated and non-glycosylated species was examined separately however, we found that the presence of N-linked glycans markedly and selectively stabilized mutant proteins, with the half-life (in the ER) approaching that of WT. These results thus extend the previous findings of Hirano et al., (119) and demonstrate a general role for N-linked glycans in early AQP2 folding and ER quality control. It should be noted that glycosylated and non-glycosylated monomers of WT AQP2 readily formed hetero-tetramers. In contrast, recessive mutations that disrupt AQP2 folding and ER trafficking fail to tetramerize (16,128,189). Therefore it seems likely that the differential stability of mutant isoforms reflects that of monomeric proteins.

N-linked glycans have long been known to facilitate protein folding. Glycosylation inhibitors cause generalized protein misfolding in the ER lumen as demonstrated by the induction of a

strong unfolded protein response (190-198). N-linked glycans are also involved in targeting misfolded substrates to the ERAD pathway. Cleavage of a single mannose residue by ER mannosidase I, generates a Man₈ intermediate and provides a degradation signal recognized by the ER protein EDEM (199). This process may provide a timing mechanism for monitoring folding, because if folding is completed before Man₉ trimming, then the substrate is exported from the ER. In contrast, if mannose trimming occurs prior to export, then the substrate may become a degradation target. Our findings that glycosylation specifically improves the stability of trafficking mutants, but not WT AQP2, suggests a novel physiological role for N-linked glycans, namely that they provide a selective advantage for proteins slightly below the threshold for efficient folding. Such a finding is particularly intriguing given the large number of disease-related mutations that direct proteins into the ERAD pathway.

AQP2 is a highly hydrophobic protein with a significant fraction of its mass located within the lipid bilayer and only a small portion of lumenally exposed polypeptide. It is therefore unlikely that glycosylation would effect solubility or aggregation of AQP2 as has been proposed for soluble proteins. Consistent with this notion, none of the mutations examined affected protein translocation or early folding events that establish cotranslational topology of individual TM segments (Figure 6 and appendix Figure A-1), and addition of N-linked sugars occurs cotranslationally even before the mutant residues have been synthesized. This eliminates the possibility that a cohort of more stable mutant proteins is preferentially glycosylated. It therefore seems likely that the presence of the oligosaccharide moiety either directly facilitates late folding events, possibly within the lipid bilayer, or in some manner masks the misfolded protein from recognition by quality control machinery (200).

One possibility is that interactions with calnexin or calreticulin (CNX/CRT) may stabilize glycosylated AQP2 by binding to partially trimmed glucose residues. For example, when a truncated form of Ribophorin I, RI₃₃₂ was pharmacologically prevented from binding to

CNX/CRT, its half-life was reduced (201). Similarly, RI mutants lacking glycosylation consensus sites exhibited decreased half-lives. MHC Class I heavy chain was also degraded more rapidly when interactions with CNX were blocked, but in contrast to RI₃₃₂, removal of the N-linked consensus site had a stabilizing effect (202), suggesting a complex relationship between carbohydrate effects on folding and the CNX/CRT pathway. In our case, we were unable to detect a significant difference in the stability of either WT or mutant glycosylated AQP2 after treatment of oocytes with the glucosidase inhibitors castanospermine or deoxynojirimycin (data not shown). However, interpretation of these experiments was difficult because of the incubation times necessary to observe stability differences (10-20 hrs) and the toxicity of long term exposure to glucosidase inhibitors.

Because water channel function could not be detected for any of the mutant proteins (data not shown), it is unlikely that their proper trafficking ever occurs in oocytes. We therefore favor the possibility that recognition of mutant AQP2 by quality control machinery is somehow delayed by the presence of the glycan moiety. This implies that folding defects in different mutants are impacted by events in the ER lumen. AQP2 stabilization also appears to be generalized and independent of the folding defect because it is shared by missense mutations located in different physical and functional regions of the protein (Figure 1). Importantly, this effect is not limited to the *Xenopus* expression system because similar observations were made for the T126M mutant expressed in mammalian hepatocytes (119).

Most consensus sites in polytopic proteins must be located at least 12-14 residues from the ends of TM segments in order for OST to efficiently attach the carbohydrate moiety (125,203). Based on AQP1 and Glpf crystal structures (4,25,204), we estimate that Asn123 is only 6-9 residues from the N-terminus of TM4 (see Figure 1). This is likely responsible for the inefficient glycosylation observed for AQP2 because increasing this distance to ~16 residues (i.e. truncation prior to TM4) substantially improved cotranslational glycosylation

efficiency (Figure 6). AQP2 glycosylation also appears to be primarily cotranslational, as it takes place during translation and before AQP synthesis is completed. However, isolated reports have indicated prolonged ER residence time may increase potential exposure of N-linked consensus sites to ER-localized oligosaccharyltransferase and result in posttranslational glycosylation (205,206). To some extent this may occur for R187C in which the absolute amount of glycosylated material continued to increase after synthesis of radiolabeled protein was completed. ER retention per se, however, did not correlate well with the extent of post-translational glycosylation, because T126M, A147T and C181W mutants were all ER retained but did not exhibit the increase of total glycosylated protein observed for R187C. In addition, glycosylated proteins persisted after the non-glycosylated protein had been largely degraded making it unlikely that increased stability was due solely to posttranslational glycosylation. Thus our data support a model in which AQP glycosylation is primarily cotranslational and limited by steric accessibility of the consensus site. Unfortunately attempts to define the precise extent of post-translational glycosylation using tunicamycin were inconclusive. Oocytes stores of dolichol-bound high mannose oligosaccharides were only depleted after 8 hours of tunicamycin exposure (Appendix FigureA-2), and stability of both glycosylated and nonglycosylated AQP2 isoforms was decreased during subsequent incubation, consistent with a general toxic effect (data not shown).

In conclusion, we have shown that the N-linked glycosylated forms of four naturally occurring disease-related AQP2 mutants, T126M, A147T, C181W and R187C persist longer in the ER than their non-glycosylated counterparts. It was previously proposed that high steady state levels of glycosylated mutant AQP2 were caused by poor folding of the mutant proteins that interfered with the removal of the high-mannose carbohydrates in the early cis-Golgi complex (132). In contrast, our results are most consistent with the interpretation that addition of N-linked glycans result in a generalized increase in mutant AQP2 half-life, stabilizing the small pool of glycosylated (relative to non-glycosylated) species which accumulates and

predominates over time. This does not occur for WT protein in which the non-glycosylated form is already very stable. We can not completely rule out, however, the possibility that delayed glycosylation may contribute to the apparent half-life for some mutants, particularly R187C. While glycosylation of some proteins is necessary for them to obtain their native structure, wild type AQP2 does not require addition of N-linked glycosylation for proper folding or water channel function. Thus it is surprising that addition of N-linked sugars to AQP2 mutant proteins has such an effect. These studies raise the possibility that maneuvers to manipulate attachment of N-linked carbohydrates may provide one strategy for improving stability and possibly trafficking of mutant proteins in human disease.

Acknowledgements:

The authors thank Dr. A. Verkman for providing AQP2 cDNA and Dr. D. Koop for helpful discussion regarding degradation kinetics. Supported by the National Institutes of Health (GM53457, DK51818). W.R.S. is an established investigator of the American Heart Association.

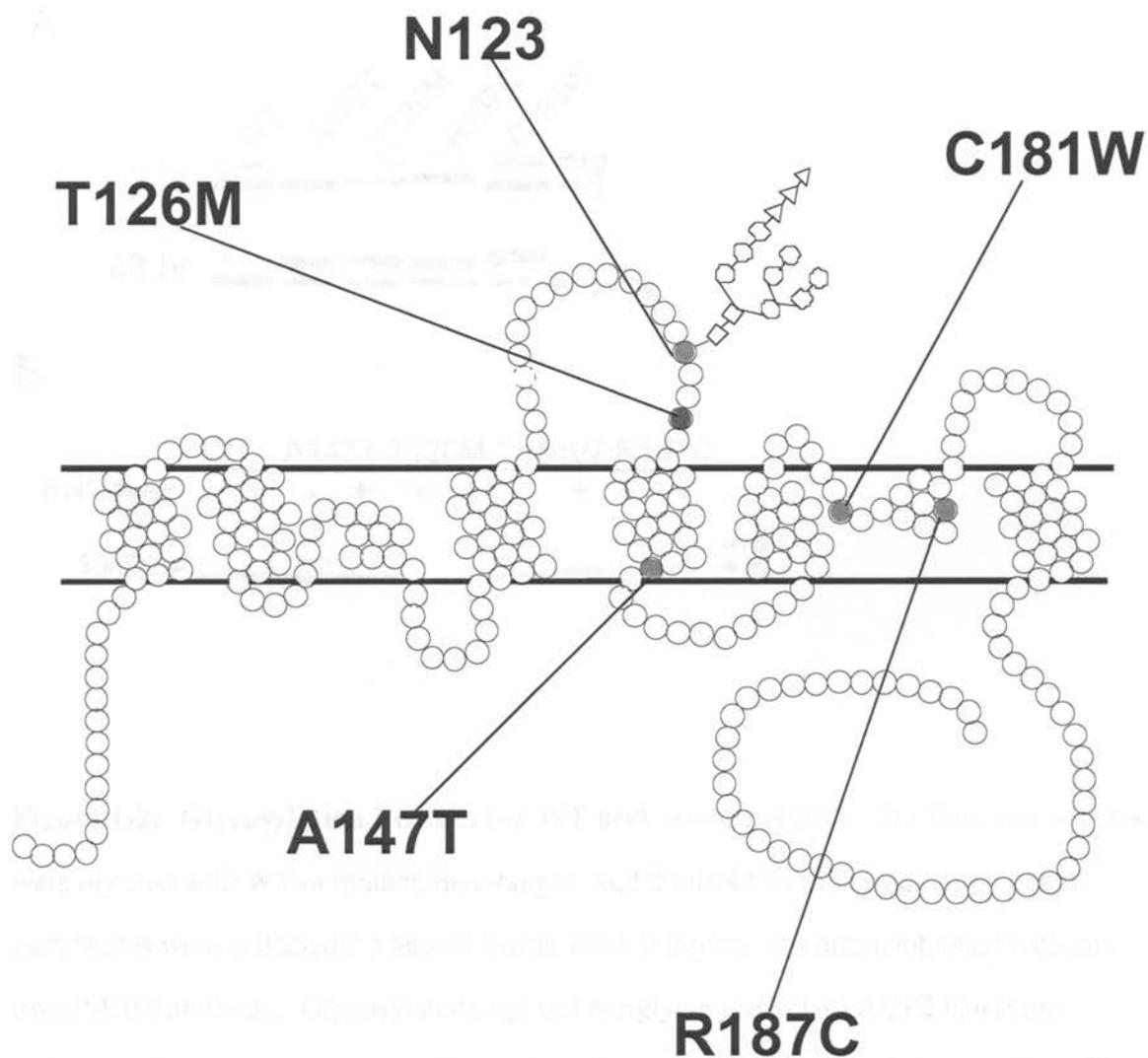


Figure 1-1: Diagram of predicted AQP2 topology and relative locations of four NDI-linked mutations, T126M, A147T, C181W, and R1817C. Topology and location of TM segments was based on crystal structure of AQP1 and GlpF (4,25,204), branching structure represents high mannose core oligosaccharide attached to Asn123.

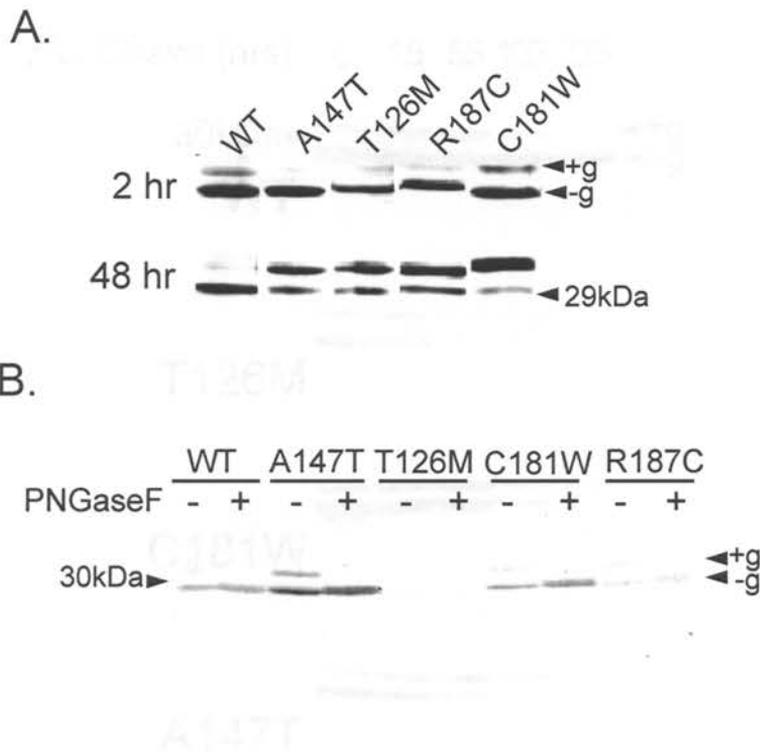


Figure 1-2: Glycosylation Profiles for WT and mutant AQP2. (A) *Xenopus* oocytes were injected with WT or mutant, myc-tagged AQP2 mRNA as indicated. Total oocyte membranes were collected 2 h and 48 h after RNA injection, and immunoblotted with anti-myc (9E10) antibody. Glycosylated (+g) and nonglycosylated (-g) AQP2 bands are indicated. Exposure times were adjusted to best demonstrate the glycosylation profile. (B) Oocytes were coinjected with mRNA and ^{35}S -methionine, incubated 8 h and homogenized. Immunoprecipitated samples (myc, 9E10 antibody) were analyzed directly (-) or after digestion with PNGaseF (+) by SDS-PAGE and autoradiography.

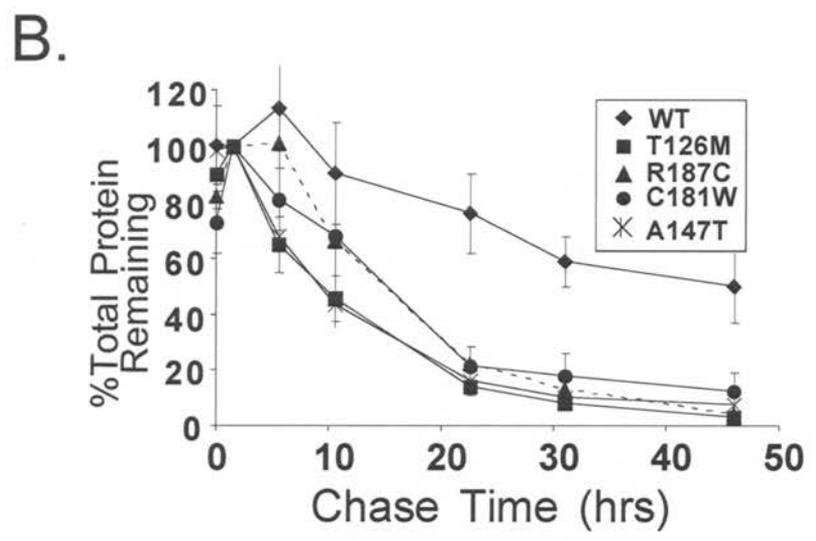
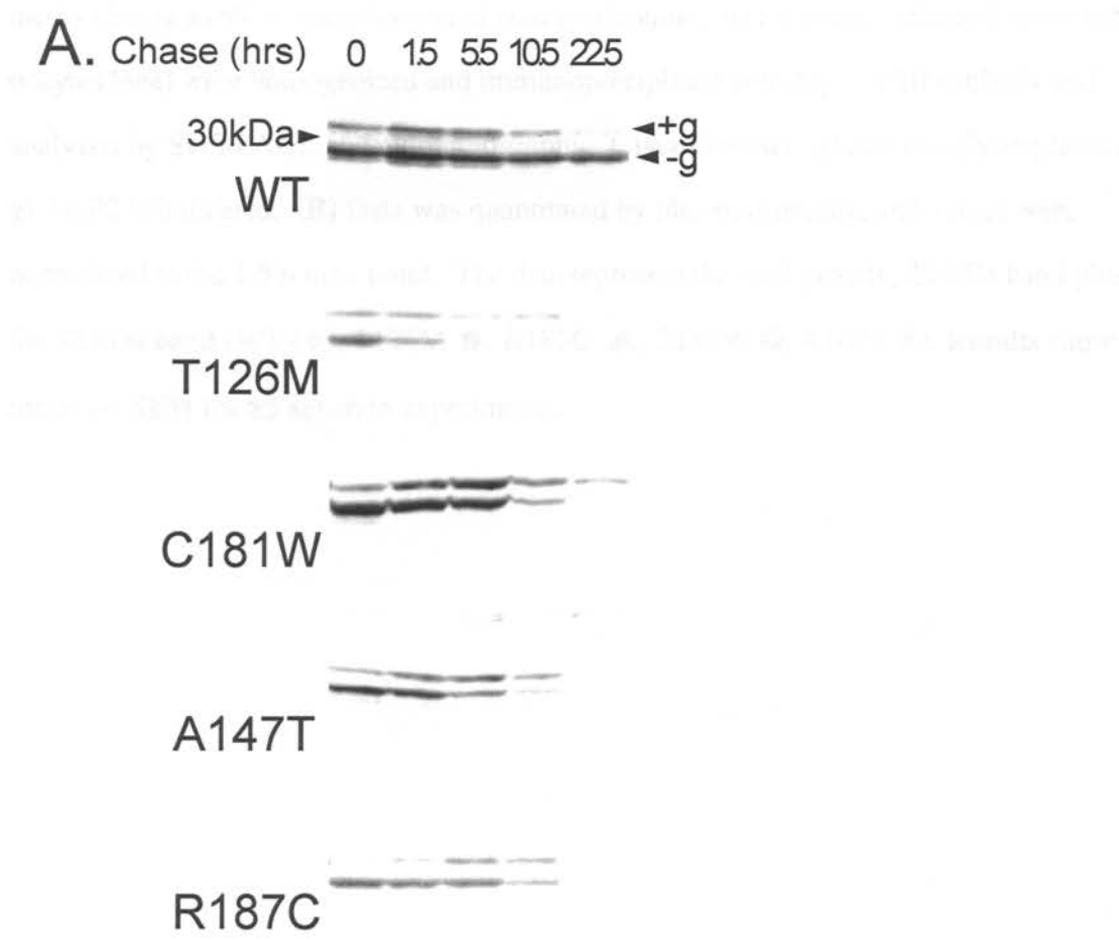


Figure 1-3: Pulse Chase analysis of AQP2 synthesis. (A) Xenopus oocytes were coinjected with myc-tagged AQP2 mRNA and ³⁵S-methionine, labeled for 1.5 hours and chased in fresh media containing unlabeled methionine. T=0 was taken 30 min after

media change to allow equilibration of cold methionine. At the times indicated oocytes (5 oocytes/lane) were homogenized and immunoprecipitated with Myc-9E10 antibody and analyzed by SDS PAGE and autoradiography. Glycosylated (+g) and nonglycosylated (-g) AQP2 is indicated. **(B)** Data was quantitated by phosphorimaging, and values were normalized to the 1.5 h time point. The data represent the total protein, 29 kDa band plus the 32 kDa band (WT (◆), T126M ■, R187C ▲, C181W ●, A147T ✱). Results show mean +/- SEM for ≥ 5 separate experiments.

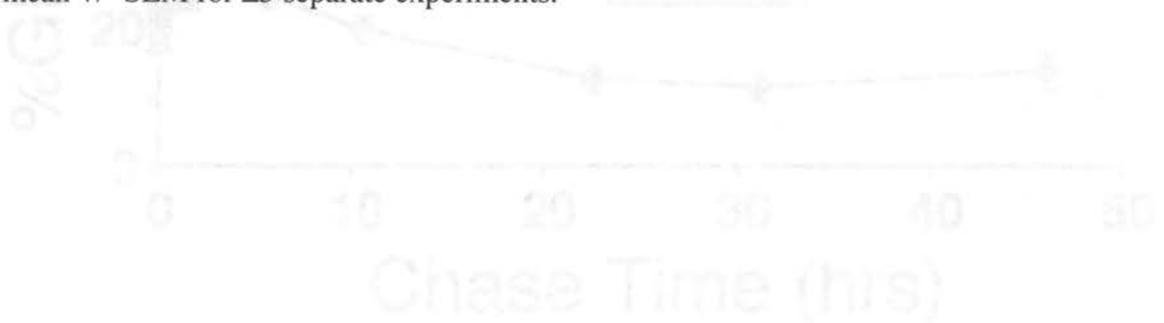


Figure 3B: Percentage of glycosylated AQP2. Oocytes were transfected with the indicated AQP2 constructs and chased for the indicated times. The percentage of glycosylated AQP2 in the media was determined as in Figure 3A. WT (◆), T126M (■), R187C (▲), C181W (●), A147T (✱). Results show mean +/- SEM of ≥ 5 separate experiments.

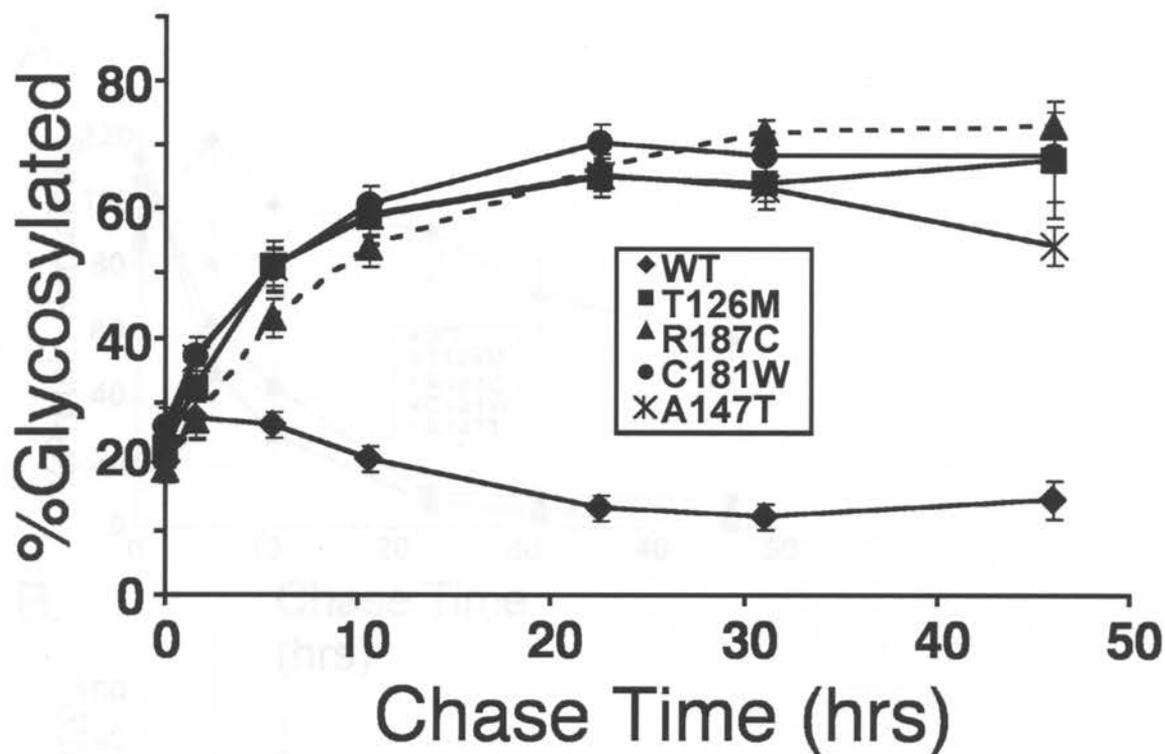


Figure 1-4: Percentage of glycosylated mutant AQP2 protein increases over time.

The percentage of glycosylated as a function of total protein at each time point was determined as in Figure 3 (WT \blacklozenge , T126M \blacksquare , R187C \blacktriangle , C181W \bullet , A147T $*$). Results show mean \pm SEM of ≥ 5 separate experiments.

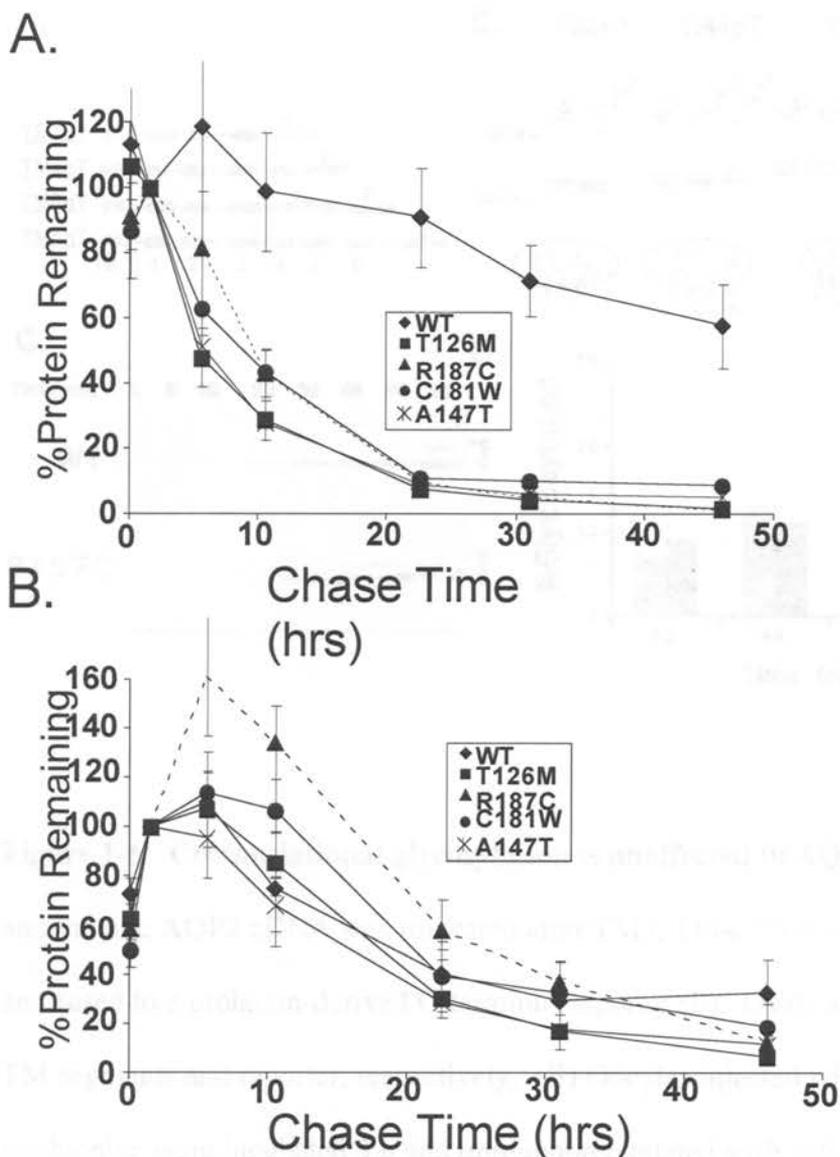


Figure 1-5: Glycosylation stabilizes mutant protein. Stability of non-glycosylated (A) and glycosylated (B) protein was determined as in Figures 4 and 5 by quantifying 32 kDa and 29 kDa bands, independently from separate pulse-chase experiments. Half-lives for mutants(T126M ■, R187C ▲, C181W ●, A147T ✱) were determined as described in methods. WT (◆) half-life was estimated to be >48 h.

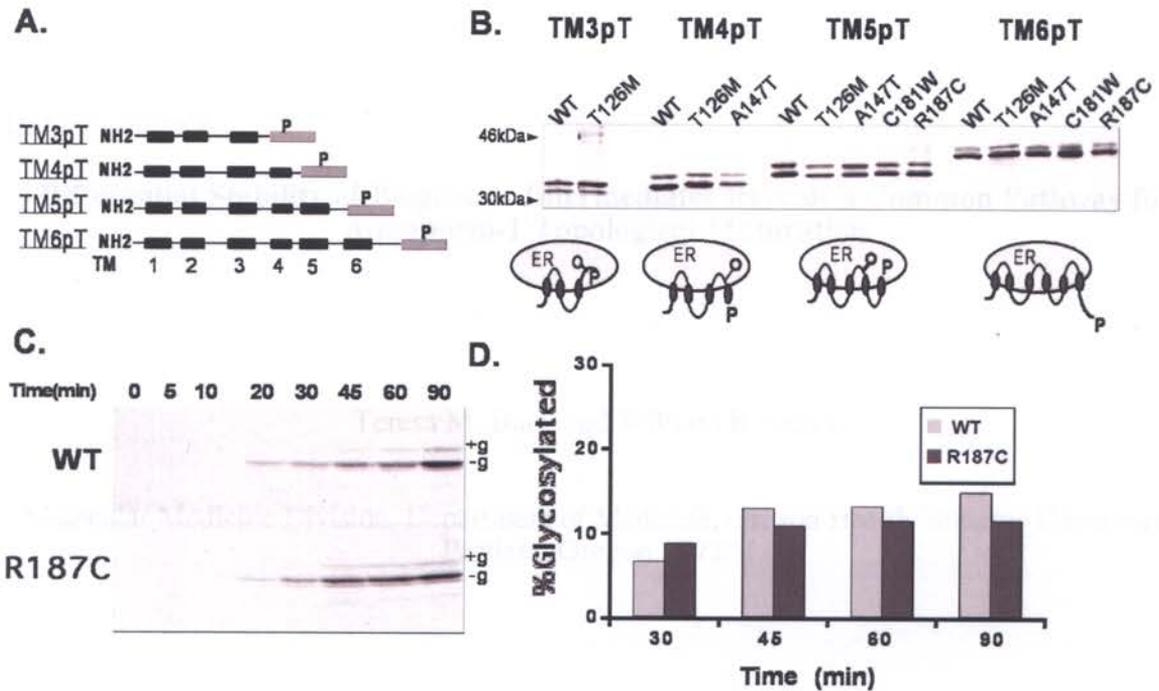


Figure 1-6: Cotranslational glycosylation is unaffected in AQP2 mutants. (A) WT and mutant, AQP2 cDNA was truncated after TM3, TM4, TM5 or TM6 as diagrammed and fused to a prolactin-derived C-terminus reporter (P). Ovals and rectangles represent TM segments and reporter, respectively. (B) Oocytes injected with mRNA and ^{35}S -methionine were incubated 3 h and immunoprecipitated with anti-prolactin antibody and analyzed by SDS-PAGE. Predicted topology for the AQP2 protein at each truncation is diagrammed beneath sets of autoradiograms. C: oocytes injected with mRNA [^{35}S]methionine were incubated for time specified and immunoprecipitated with 9E10 antibody. D: data in C were quantitated to determine the % glycosylation at each timepoint. Signals of glycosylated protein at time points earlier than 30 min were not sufficient for quantitation.

B. Manuscript 2

Differential Stability of Biogenesis Intermediates Reveals a Common Pathway for Aquaporin-1 Topological Maturation

Teresa M. Buck and William R. Skach

Molecular Medicine Division, Department of Medicine, Oregon Health Sciences University,
Portland Oregon. 97201

Journal of Biological Chemistry, Vol 280, pp261-269, 2005

Key words: Aquaporin, topology, protein biogenesis, polytopic protein, *Xenopus* oocytes, endoplasmic reticulum,

Address Correspondence to:
William R. Skach, M.D.
Division of Molecular Medicine
Oregon Health Sciences University
3181 SW Sam Jackson Park Rd.
Portland, Oregon 97201

Ph: 503-494-7322
Fax: 503-494-1933
Email: skachw@ohsu.edu

Abbreviations: AQP, Aquaporin; ER, Endoplasmic reticulum; Proteinase K, PK; Ribosome translocon complex, RTC; TM, transmembrane

1. Abstract

Topological studies of multi-spanning membrane proteins commonly use sequentially truncated proteins fused to a C-terminal translocation reporter to deduce TM segment orientation and key biogenesis events. Because these truncated proteins represent an incomplete stage of synthesis, they transiently populate intermediate folding states that may or may not reflect topology of the mature protein. For example, in *Xenopus* oocytes, the Aquaporin-1 (AQP1) water channel is cotranslationally directed into a 4-spanning intermediate, which matures into the 6-spanning topology at a late stage of synthesis (J. Cell Biol. 125:803, Mol. Biol. Cell 11:2973). The hallmark of this process is that TM3 initially acquires an N_{exo}/C_{cyto} (Type I) topology and must rotate 180° to acquire its mature orientation. In contrast, recent studies in HEK-293 cells have suggested that TM3 acquires its mature topology cotranslationally without the need for reorientation (J. Biol. Chem. 277:15215). Here we re-examine AQP1 biogenesis and show that irrespective of the reporter or fusion site used, oocytes and mammalian cells yielded similar topologic results. AQP1 intermediates containing the first three TM segments generated two distinct cohorts of polypeptides in which TM3 spanned the ER membrane in either an N_{cyto}/C_{exo} (mature) or N_{exo}/C_{cyto} (immature) topology. Pulse-chase analyses revealed that the immature form was predominant immediately after synthesis, but that it was rapidly degraded via the proteasome-mediated ERAD pathway with a half-life of less than 25 minutes in HEK cells. As a result, the mature topology predominated at later time points. We conclude that i) differential stability of biogenesis intermediates is an important factor for in vivo topological analysis of truncated chimeric proteins and ii) cotranslational events of AQP1 biogenesis reflect a common AQP1 folding pathway in diverse expression systems.

2. Introduction

Polytopic membrane proteins are synthesized and oriented in the endoplasmic reticulum (ER) by the ribosome and Sec61 translocon complex (RTC) (60,207,208). In the simplest model, topology of each transmembrane (TM) segment is established in a vectorial and sequential manner (N- to C-terminus) (209) as independent signal anchor and stop transfer sequences alternately gate the translocon and the ribosome-translocon junction and direct TM segment integration into the lipid bilayer (cotranslational model) (162,184,210). However, a growing body of evidence has demonstrated that the final topology of many native proteins is not necessarily established cotranslationally, but rather through cooperative interactions between topogenic determinants (TM segments) located within different regions of the polypeptide (post-translational model)(82,163,211-217).

One example of the post-translational model occurs during the biogenesis of Aquaporin-1 (AQP1), a hydrophobic membrane protein of approximately 29 kDa that exists as a homotetramer in cell membranes. AQP1 is a member of the Major Intrinsic Protein (MIP) family, (3,171). In its mature form it exhibits a characteristic topology with six TM segments and two additional short helical regions flanked by conserved NPA motifs that fold inward within the plane of the membrane to form a monomeric, water selective pore (4,5). AQP1 is expressed in diverse cell types and is localized in the kidney to the proximal tubule and descending limb of the loop of Henle where it plays a major role in renal water reabsorption (3,8).

Early biogenesis studies of AQP1 in cell free systems and microinjected *Xenopus* oocytes revealed a novel mechanism in which only four of its six transmembrane segments cotranslationally acquired a membrane spanning topology (78). This 4-spanning intermediate later matures in the ER membrane to form the final 6-spanning structure (77,81). AQP1 biogenesis differs from the cotranslational pathway utilized by a close homolog, AQP4 (76), in part because hydrophilic residues within the N-terminus of TM2 (Asn49 and K51) disrupt stop-

transfer sequence and allow TM2 to transiently pass through the translocon and into the ER lumen (77). As a result, when TM3 emerges from the ribosome, it acts as a stop-transfer sequence to terminate translocation and cotranslationally adopts an N_{exo}/C_{cyto} (Type I) topology. A 4-spanning intermediate is synthesized because TM4 does not reinitiate translocation, and TM5 and TM6 act as signal and stop transfer sequences, respectively (Figure 1A). The defining feature of this folding intermediate is that TM3 is initially oriented with its C-terminus positioned in the cytosol. In *Xenopus* oocytes, TM3 topology is “corrected”, as it rotates 180° about the plane of the membrane to acquire its mature topology during a later stage or following the completion of AQP1 synthesis (81).

Recent work has suggested that the AQP1 biogenesis pathway may be dependent on cell type (160). A study by the Turner group examined the topology of truncated constructs in which variable numbers of AQP1 TM segments were placed into a chimeric protein containing N- and C-terminal reporter domains derived from EGFP and the β -subunit of the H/K ATPase, respectively. The topology of TM segments was inferred by N-linked glycosylation of the C-terminal reporter in intact HEK-293 cells. In contrast to results in oocytes, the Turner group found that TM3 had already acquired its mature N_{cyt}/C_{exo} topology (e.g. the reporter was glycosylated) in constructs containing only the first three AQP1 TM segments and concluded that AQP1 biogenesis therefore occurs in a cotranslational fashion without reorientation from a 4-spanning intermediate (160).

These studies have lead to several unanswered questions regarding AQP1 biogenesis. Different reporter domains (prolactin verses H/K ATPase β -subunit derivatives), translocation assays (protease protection versus glycosylation) and fusion sites could potentially account for the different apparent topology observed for TM3. Because C-terminal reporters are routinely used to study polytopic protein biogenesis, these factors are of more than just academic interest. In addition, while truncated proteins provide important structural information at a relatively

defined point of synthesis, they lack C-terminal sequence information and therefore must be viewed as populating intermediate folding states. As such they are potential candidates for recognition by ER quality control machinery. Consistent with this we had previously observed that certain AQP1 fusion proteins are relatively unstable in oocytes (79). A final possibility proposed by the Turner group is that oocytes and mammalian cells handle topological information in fundamentally different ways. If true, this would have major implications for protein biogenesis.

Xenopus oocytes efficiently express diverse aquaporins and have been extensively used to study AQP biosynthesis, trafficking and function (16,117,133,172,186). However, a direct comparison of biosynthetic mechanisms in oocyte and mammalian cells has not previously been carried out. We therefore systematically examined AQP1 fusion proteins containing two, three or four TM segments in both Xenopus oocytes and HEK-293 cells to determine the origin of previous discrepancies. We now show that irrespective of the reporter or fusion site examined, both systems yielded similar topologic results. AQP1 intermediates containing the first three TM segments gave rise to two distinct cohorts of polypeptides in which TM3 spanned the ER membrane in either an N_{cyto}/C_{exo} (mature) or N_{exo}/C_{cyto} (immature) topology. Careful pulse-chase analyses revealed that the immature form predominated immediately after synthesis in both cell types, but that it was rapidly degraded via the proteasome-mediated ERAD pathway. As a result, the mature topology predominated 1-2 hours after synthesis in HEK cells. We conclude that differential stability of biogenesis intermediates is an important factor for in vivo analysis of truncated chimeric proteins and that cotranslational events of AQP1 biogenesis are conserved in diverse expression systems.

3. Materials and Methods

cDNA Construction: Plasmids AQP1P77.P, AQP1L139.P, and AQP1P169.P are described previously as clones 3, 6 and 7 (78); AQP1S66.P, AQP1T120.P, and AQP1L164.P were generated using the same strategy by PCR amplification of AQP coding region using antisense oligonucleotides: S66-GCTGATGGTCACCCCACTCTGCGCCAGCGTGGC, T120-AAGCGAGGTCACCGTCAGGGAGGAGGTGAT, L164-GGGGGCGGTCACCCCAAGGTCACGGCGCCTCCG. The resulting constructs contain AQP1 residues 1-66, 1-77, 1-120, 1-139, 1-164, or 1-169 followed by the C-terminal 142 amino acids of bovine prolactin, a passive translocation reporter (184). Plasmid pEGFP.AQP1T120.b was kindly provided by R. James Turner and is derived from pEGFP-C3 (Clontech, Palo Alto, CA) (160). It contains AQP1 codons 1-120 fused between EGFP and the C-terminal 177 amino acids of the H,K-ATPase β -subunit. pEGFP.AQP1S66. β , pEGFP.AQP1P77. β , pEGFP.AQP1L139. β , pEGFP.AQP1L164. β , pEGFP.AQP1P169. β were generated by PCR amplification of AQP1 cDNA using a sense oligonucleotide TGAGTAGATCTCATGGCCAGCGAGTTCAAG and corresponding antisense oligonucleotides S66 TGAGTAAGCTTCCA CTCTGCGCCAGCGTC, P77-CAGTGTGAAGCTTCCCGGGTTGAGGTGGGCCCCG, L139-ATCTCGAAGCTTCCAGGCCCTGGCCCCGAGTTC, L164-TGAGTAAGCTTCCAAGGTCACGGCGCCT, and P169-CCGATGAAGCTTCCGGGGGCTGAGCCACCAAGG encoding Bgl II (sense) and Hind III (antisense) restrictions sites that were used to ligate DNA fragments into the pEGFP.b cassette. The constructs encode AQP1 residues 1-66, 1-77, 1-139, 1-164, or 1-169 flanked by EGFP and β -subunit. The full-length AQP1 construct was generated by ligating AQP1 cDNA into the mammalian expression vector pEGFP-N3 (Clontech, Palo Alto, CA) between the Hind III and Not I restriction sites resulting in a full-length AQP1 construct without an EGFP tag. The sequence of all PCR amplified DNA was verified by automated DNA sequencing.

Xenopus laevis expression : mRNA was transcribed in vitro with SP6 RNA polymerase (New England Biolabs, Beverly, MA) using 2 μ g of plasmid DNA in a 10 μ l volume at 40 $^{\circ}$ C for 1 hour as previously described (78). Aliquots were used immediately or frozen in liquid nitrogen and stored at -80 $^{\circ}$ C. 2 μ l of transcript was mixed with 50 μ Ci of 35 S-methionine (0.5 μ l of a 10X concentrated Tran 35 S-label, ICN Pharmaceuticals, Irvine, CA) and 50 nl was injected into stage VI *Xenopus laevis* oocytes (50 nl/oocyte) on an ice cold stage. Oocytes were incubated at 18 $^{\circ}$ C in MBSH (88 mM NaCl, 1 mM KCl, 24 mM NaHCO₃, 0.82 mM MgSO₄, 0.33 mM Ca(NO₃)₂, 0.41 mM CaCl₂, 10 mM HEPES-pH 7.4, 50 mg/ml Gentamicin, 100 units/ml Penicillin, and 100 mg/ml Streptomycin Sulfate).

Protease Protection and Pulse/Chase assays in oocytes: Oocytes were injected as described above, incubated at 18 $^{\circ}$ C for 3 hrs, and homogenized by hand in 3 volumes 0.25 M sucrose, 50 mM KAc, 5 mM MgAc₂, 1.0 mM DTT, 50 mM Tris-Cl pH 7.5. Homogenates were divided into 3 x 10 ml aliquots on ice. Proteinase K (PK) (Roche Diagnostics, Indianapolis, IN) was added (0.2 mg/ml final concentration) in the presence or absence of 1% Triton X-100. Reactions were incubated on ice for 1 hr and rapidly mixed with PMSF (10 mM) and boiled in 10 volumes of 1% SDS, 0.1M Tris-HCl, pH 8.0 for 5 min. Samples were then diluted in 10 volumes of buffer A (100 mM Tris pH 8.0, 100 mM NaCl, 1% Triton X-100, 2 mM EDTA), incubated at 4 $^{\circ}$ C for 15 min and centrifuged at 16,000 x g for 10 min at 4 $^{\circ}$ C. to remove insoluble debris. Efficiency of the assay was regularly assessed using a known secretory control protein and was >80%. For pulse chase assays oocytes were injected as described, incubated at 18 $^{\circ}$ C for 2 hrs, and media was then replaced with fresh MBSH containing 2 mM methionine. Oocytes were harvested at the specified time-points and processed as above.

HEK-293 Pulse-Chase Studies: HEK-293 cells were cultured in Dulbecco's Modified Essential Medium (Fisher Scientific, Pittsburgh, PA) supplemented with 100 units/ml Penicillin, 100 mg/ml Streptomycin sulfate and 10% heat inactivated fetal bovine serum (GibcoBRL,

Grand Island, NY). Cells were grown on plastic in a humidified incubator at 37°C and 5% CO₂ and were passaged every 3-4 days. Cells were transfected at 50-60% confluence with *TransIT-LT* transfection reagent (Mirus, Madison, WI) according to manufacturers instructions (3 mg cDNA and 12 ml *TransIT-LT* reagent per 60 mm plate). Transfection efficiencies were approximately 40% as determined by EGFP fluorescence. 48 hrs after transfection media was replaced with 1 ml methionine-cysteine free media for 30 min. Cells were then pulse-labeled with 80 mCi ³⁵S-methionine for 15 min, and media was removed and replaced with fresh complete media for the indicated duration of the chase. Cells were lysed on ice with 1 ml RIPA buffer (0.1% SDS, 1% Triton X-100, 1% deoxycholate, 150 mM NaCl, 10 mM Tris-Cl, pH 8.0, 2.5 mM MgCl₂, 1X Protease Inhibitors III (CalBiochem, San Diego, CA) and 1mM PMSF) for 30 min on ice and passed through a 26-gauge needle x 3. Samples were then clarified at 16,000 x g for 12 min.

Immunoprecipitation: Clarified oocyte or HEK cell homogenates were incubated with anti-prolactin antisera (ICN Biomedicals, Costa Mesa, CA) 1:2000 dilution, polyclonal anti-EGFP antibody (Molecular Probes, Eugene, OR) 1:2000 dilution, or polyclonal rabbit anti-AQP1 antisera raised against purified AQP1 protein (generously provided by A. Verkman, UCSF, San Francisco, CA) 1:1000 dilution and 5.0 ml protein-A affigel (BioRad, Hercules, CA). Samples were rotated for 10 hrs at 4°C (oocyte homogenates) and 4 hr (HEK cells) prior to washing 3x's with solubilization buffer and twice with 100 mM NaCl, 100 mM Tris-Cl, pH 8.0. Samples were boiled in protein SDS loading buffer (218) and analyzed by SDS-PAGE, EN³HANCE (Perkin-Elmer Life Sciences, Boston, MA) fluorography. Intensity of recovered bands was quantitated using a BioRad personal Molecular PhosphorImager Fx (Kodak screens, Quantity-1 software).

PNGaseF Digests: Beads from immunoprecipitations were resuspended after the final wash in 15 ml 0.1 M Tris-Cl pH 7.5 and 0.3 ml PNGaseF (New England Biolabs, Beverly, MA), and

incubated at 37°C for 3 hrs. Samples were then analyzed by SDS-PAGE as described above.

4. Results

C-terminal translocation reporters placed after connecting peptide loops have been widely used to study membrane protein topology and biogenesis (78,82,162-169). The rationale for this approach is that topology of TM segments and their connecting loops is controlled by the action of topogenic determinants encoded within the nascent polypeptide. TM segments that function as signal (anchor) sequences open the Sec61 translocon channel and initiate translocation of polypeptide into the ER lumen. All else being equal, translocation will continue until a second TM segment is synthesized that functions as a stop transfer sequence to close the translocon, relax the ribosome-translocon junction and allow the next peptide loop to enter the cytosol. In this manner, topology is determined cotranslationally as the nascent polypeptide emerges from the ribosome. A passive reporter domain (i.e. one containing no intrinsic topogenic information) fused to a peptide loop downstream of a signal sequence will therefore follow the loop into the ER lumen, whereas a reporter located downstream of a stop transfer will, in turn, remain in the cytosol.

Two variations on this strategy have been employed to map the cotranslational topology of AQP1. In the first approach, a C-terminal reporter domain derived from the secretory protein prolactin was fused to each AQP1 peptide loop, and topology was determined *in vitro* and in microinjected *Xenopus* oocytes using a protease protection assay (78). The second approach used a different chimeric cassette containing an N-terminal EGFP domain and a C-terminal reporter derived from the β -subunit of the H/K ATPase. Topology of the β -subunit was then determined in HEK-293 cells based on N-linked glycosylation, which occurs exclusively in the ER lumen (160,219,220). Key differences between these studies are summarized in Figure 1. In oocytes, reporters fused to the TM2-3 loop (residue P77) or the TM3-4 loop (L139) were

located in the ER lumen and cytosol, respectively (Figure 1A) (78). Surprisingly this was different from their expected topology in the mature protein. Two subsequent studies confirmed these results and demonstrated that AQP1 TM segments 2-4 cotranslationally acquire an immature topology that is subsequently converted to the mature topology by an internal 180° rotation of TM3 (77,81). In contrast, glycosylation patterns of the β -subunit domain lead to the conclusion that the TM2-3 and TM3-4 loops acquire their proper topology cotranslationally in the ER of mammalian cells and therefore do not undergo a topological reorientation (Figure 1B).

While similar in many ways, these studies differed in several important aspects including the location of fusion sites, the reporter domain used, and the cell expression system. We therefore undertook a systematic comparison of AQP1 biogenesis in oocytes and mammalian cells using similar truncation sites and reporters to resolve these discrepancies. Because the major differences involved the initial orientation of TM3 (Type I in oocytes and Type II in HEK cells) we focused our attention primarily on the topology of TM3 and its immediate flanking residues. Topology of other regions such as the N- and C-termini and TM5-6 are well established and were therefore not re-tested here (Figure 1).

Reporter fusion site does not affect outcome of early biogenesis experiments in oocytes.

We first addressed whether topological differences might arise from the use of different fusion sites. In particular, up to 10-15 flanking residues can influence topogenic activities of TM segments (161). Thus, fusion sites very close to the TM C-terminus such as those used in the EGFP- β -subunit chimeras might delete potentially important topogenic information.

Alternatively, the fusion site might introduce new flanking residues (from the β -subunit) that could inadvertently alter TM2 behavior. AQP1 fusion sites were therefore tested head-to-head by placing the prolactin reporter at both truncations sites downstream of TM2 (S66 and P77), TM3 (T120 and L139), and TM4 (L164 and L169 (Figure 2A). Topology was determined by

protease protection as described in Methods. Consistent with our previous results, fusion to the TM2-3 peptide loop at either truncation S66 or P77 resulted in similar translocation of the reporter (80-83%, Figure 2B, lanes 1-6). Similarly, the β -subunit-TM2 fusion site used by Dohke et al. had only a modest effect on translocation of the P-reporter, decreasing translocation by 5-10% (Appendix Figure A-3). In contrast, when the reporter was placed in the TM3-4 loop (truncations T120 and L139) or the TM4-5 loop (truncations L164 and L169) it was ~85-90% and >95% cytosolic, respectively (Figure 2B, lanes 7-18). As was observed previously, ~50% of fusion proteins truncated after TM3 (TM3.P and TM4.P) underwent N-linked glycosylation at residue Asn42 which resulted in a 3 kDa shift in migration (78).

These results thus support previous topological findings in *Xenopus* oocytes and demonstrate that different topological orientations obtained for TM2 and TM3 (shown in Figure 1) were not likely due to subtle modification of downstream flanking residues caused by different fusion sites. In all cases, TM2 did not function as an efficient stop transfer sequence to terminate translocation, and TM3 initially spanned the membrane with its C-terminal flanking residues oriented towards the cytosol.

Truncated AQP1 constructs generate multiple biogenesis intermediates with differential stability in oocytes. One limitation to all topology mapping studies is that results are often not absolute. For example, variable degrees of glycosylation were observed for several EGFP-AQP- β -subunit chimeras in HEK cells, particularly for fusion sites in the TM2-3 and TM3-4 loops (160). Similarly, in our original analysis we used a cut-off of ~80% protease protection to define whether a peptide loop segment was translocated into the ER lumen (78). As shown in Figure 2B, the majority of the reporter fused to residues T120 or L139 was accessible to cytosolic protease. However, a small fraction (10-15%) appeared to be resistant to digestion. While this could potentially be caused by subtle variations in reporter folding and efficiency of proteolysis, even this minor fraction was completely degraded in the

presence of non-denaturing detergent (Figure 2B, lanes 9 and 12). This raised the possibility that nascent polypeptides with protected versus accessible reporters might represent separate cohorts of proteins with different topological structures.

We therefore used pulse chase metabolic labeling to test the topology of AQP1 fusion proteins truncated after TM3 at residues T120 and L139. Oocytes were simultaneously injected with mRNA and ^{35}S -methionine to initiate rapid radiolabeling of newly synthesized protein. Labeling gradually tapers off after several hours due to dilution by media and oocyte methionine stores (185). In some cases the chase was performed in the presence of excess unlabeled methionine but this had little effect given the long time course used (data not shown). Incubation times were chosen based on previous studies demonstrating that processing and transport proteins out of the oocyte ER is significantly slower than mammalian cells and occurs over the time frame of many hours (185,221). At the end of the incubation, reporter topology of remaining polypeptides was determined by protease protection. Two important results were obtained from these experiments. First, both constructs, regardless of the fusion site, were quite unstable and underwent significant degradation during the chase period (Fig 3A, compare lanes 1, 4, 7 and 10, 13, 16). After 24 hours the amount of protein remaining had decreased by 60-80% (Figure 3B). In contrast, there was little or no decrease in the amount of protein that yielded a proteaseprotected reporter. This resulted in a 3-4 fold increase in the fraction of remaining chains in which TM3 C-terminal residues appeared to reside in the ER lumen (Figure 3C and D).

While not conclusive in themselves, these results suggested a scenario in which synthesis of the first 3 AQP1 TM segments provides insufficient topogenic information to efficiently direct TM segments into their proper topology. Most polypeptides, (~85% in oocytes) can only achieve a 2-spanning topology in which TM2 does not yet stably span the membrane and in which TM3 is oriented in an $N_{\text{exo}}/C_{\text{cyto}}$ topology. However, in a relatively small fraction of polypeptides

TM3 does acquire its proper $N_{\text{cyt}}/C_{\text{exo}}$ topology which is reflected by translocation of the TM3-4 loop and its attached reporter (Figure 3C). Thus two different isoforms appear to be generated in oocytes when synthesis of TM3 and TM4 is completed and synthesis of TM5-6 has not yet begun. Whereas the 2-spanning topology initially predominates, selective degradation of this isoform leads to a relative increase in the more stable 3-spanning isoform over time.

AQP1 fusion proteins generate multiple topological isoforms in mammalian cells. If truncated AQP1 constructs gave rise to multiple topological isoforms in mammalian cells, then differential stability might explain some of the topological discrepancies demonstrated in Figure 1. We therefore re-evaluated chimeric fusion proteins previously used by the Turner group to better define AQP1 biogenesis in mammalian cells. AQP1 fragments were inserted into the EGFP... β -subunit chimera described previously at two different truncation sites after TM2, TM3 and TM4 (160). As shown in Figure 4A, the β -subunit contains five potential N-linked glycosylation sites that provide a read-out for translocation into the ER lumen. Constructs were transiently expressed in HEK-293 cells, pulse labeled with ^{35}S -methionine for two hours, and EGFP reactive proteins were immunoprecipitated and analyzed by SDS-PAGE before and after removal of N-linked sugars by PNGase F. When AQP1 was truncated at residues S66 and P77, 64% and 57% of EGFP-reactive protein was present in a glycosylated form (Figure 4B lanes 1-4). This is consistent with previous reports that found >50% glycosylation in the identical S66 construct at steady state. However in contrast to the previous conclusions that TM2 spanned the membrane cotranslationally, we would argue that TM2 does not efficiently terminate translocation in HEK cells. Rather its C-terminal residues pass into the ER lumen in a major fraction of nascent polypeptides.

The β -subunit reporter was also predominantly glycosylated when fused to residues T120 or L139. Most recovered protein (75-82%) migrated at a size of ~62-65 kDa whereas only faint

bands were observed at ~50-53 kDa (Figure 4B, lanes 5-8). PNGase F digestion confirmed that the larger bands represented glycosylated forms of the 50 kDa chimeric protein. These results are consistent with steady state immunoblots (160) of the T120 truncation in which the β -subunit reporter was glycosylated in ~80% of polypeptides. The simplest explanation for these results is that TM3 is indeed predominantly oriented in its mature $N_{\text{cyt}}/C_{\text{exo}}$ topology. However, a small proportion of nonglycosylated protein was also isolated (Figure 4 upward arrows). While it is possible that this might reflect inefficient glycosylation of the reporter (222), this seems unlikely because of the large number of consensus sites available in the β -subunit. Indeed, lack of glycosylation of this reporter has been used as a direct readout for its cytosolic location (160,168,223). Polypeptides with nonglycosylated reporters therefore more likely represent a minor cohort of AQP1 that exhibits an alternate topology in which the reporter remains in the cytosol. Thus like oocytes, mammalian cells appear to generate two different isoforms from constructs containing the first three AQP1 TM segments but the relative proportion of 2- and 3-spanning isoforms would appear to be reversed (Figure 4C).

Differential stability of AQP1 topological intermediates in HEK-293 cells. Based on our observations in oocytes, we further investigated the relationship between glycosylated and unglycosylated (3-spanning versus 2-spanning) AQP1 fusion proteins by pulse-chase labeling in HEK cells. Because of rapid processing, very short pulse times (15 min) were used to capture AQP1 chimeras as close to the completion of synthesis as possible. These experiments revealed that immediately after synthesis, AQP1 chimeras containing the first three TM segments (truncations T120 or L139) were glycosylated in less than 50% of total protein (Figure 5A and B). Glycosylated proteins were initially present as a series of bands between 53 and 65 kDa in size that became fully glycosylated within 1 hour. Remarkably, 50-60% of both chimeras were initially present as nonglycosylated protein that rapidly disappeared with a half-life of <25 min. Thus within 1 hour of synthesis, >80% of remaining protein contained a glycosylated reporter.

While it is possible that the reporter might have undergone a delayed glycosylation after reaching the ER lumen, we do not believe this is the case because the total amount of glycosylated protein did not change significantly during the chase period when all glycosylated bands were quantitated (Figure 5B and Figure 6). Rather our data indicate that a significant fraction of these chimeras is synthesized in a 2-spanning topology where TM3 is cotranslationally oriented in an $N_{\text{exo}}/C_{\text{cyt}}$ topology. Just as in oocytes, this topological isoform is very unstable and selectively degraded, leaving the minor isoform to predominate at later time points. Because of its very short half-life, it is likely that some of the 2-spanning isoform is degraded during the pulse period and that the protein present at $T=0$ represents an underestimate of the actual fraction of the 2-spanning isoform (see Figure 6). These data are therefore very similar to the results obtained in oocytes and demonstrate that regardless of the reporter, oocytes and mammalian cells both seem to handle AQP1 cotranslational biogenesis in a similar fashion.

Selective degradation of AQP1 biogenesis intermediates occurs via the proteasome-mediated ERAD pathway. Misfolded proteins in the ER membrane are usually recognized by quality control machinery and degraded by the ubiquitin-proteasome (ERAD) pathway. We therefore tested whether selective degradation of the 2-spanning isoform of AQP1 occurred via the proteasome. Pulse-chase studies performed in the presence of the proteasome inhibitor MG132 revealed that nonglycosylated (2-spanning) isoforms were dramatically stabilized upon proteasome inhibition (Figure 6A and C). Similar results were observed using another proteasome inhibitor ALLN (Appendix Figure A4). The relative fraction of the 2-spanning isoforms was increased to $\sim 70\%$ at initial time points, likely as a result of blocking degradation that took place even during the pulse period as mentioned above. In addition, the ratio of 2- and 3-spanning isoforms remained nearly constant in the presence of MG132 because selective degradation of the 2-spanning isoform was reduced (Figure 6B, 6D). Importantly, no

conversion of the nonglycosylated protein to the glycosylated form was observed. Therefore the predominance of the 3-spanning topology observed at late time points does not result from delayed glycosylation of a translocated β -subunit reporter. We conclude that the changes observed in the topological profiles of AQP1 chimeras containing the first three TM segments result from selective degradation of the 2-spanning, relative to the 3-spanning, isoform.

Full-length AQP1 is stable in HEK-293 cells. We previously showed that in *Xenopus* oocytes, completion of AQP1 synthesis provides additional folding information that converts TM3 from its immature to its mature conformation (81). Thus in the full-length protein, the six-spanning topology is not a result of selective degradation, but is derived from a rearrangement of TMs 2-4. However, it remained theoretically possible that in mammalian cells, full length AQP1 could be synthesized as a 4-spanning intermediate that was rapidly degraded and that this could result in the predominant mature six-spanning topology. Because the majority of AQP1 is initially synthesized with TM3 in reverse orientation (Figure 6), we would expect >50% of the protein to be rapidly degraded by the proteasome-ERAD pathway. To test this hypothesis, the stability of full length AQP1 was examined by pulse-chase analysis in HEK cells. Results in Figure 7 clearly demonstrate that in contrast to truncated constructs, full length AQP1 was very stable throughout the chase period. Moreover, no significant differences in expression level or stability were observed in the presence of MG132 under our conditions during the time course expected for topological maturation in the ER. Given the very short pulse labeling period used in our study, it is extremely unlikely that degradation of the unstable 4-spanning isoform could account for these findings. Thus, we conclude that protein initially synthesized with TM3 in the reverse orientation is converted into a more stable structure as additional folding information is provided by synthesis of C-terminus residues.

5. Discussion

Previous studies proposed two different pathways by which AQP1 acquires its mature six-

spanning topology in cells. In *Xenopus* oocytes, biogenesis proceeds from a four-spanning intermediate into a mature six-spanning topology during late stages of synthesis (77,78,81). The hallmark of this unusual folding pathway is that TM3 is initially oriented in an $N_{\text{exo}}/C_{\text{cyt}}$ topology and must rotate 180° to acquire its proper topology and position TM2 and TM4 within the membrane. In contrast, a recent study failed to observe this intermediate TM3 orientation and concluded that mammalian cells utilize a fundamentally different biogenesis mechanism in which the mature AQP1 topology is established directly by cotranslational events (160). The current study now reconciles these differences by directly comparing early AQP1 biogenesis events in oocytes and mammalian cells using systematically truncated fusion proteins and two different C-terminal reporter domains.

Our results demonstrate a high degree of conservation in the unique features of AQP1 topogenesis that was independent of the fusion site, reporter domain, topological assay or expression system utilized. Specifically, as the second TM segment emerges from the ribosome, it failed to efficiently terminate translocation of its C-terminus flanking residues in at least 50% of nascent polypeptides in both *Xenopus* oocytes and mammalian cells. As a result, TM3 subsequently functions as a stop transfer sequence to adopt an $N_{\text{exo}}/C_{\text{cyt}}$ topology in which its C-terminus is transiently oriented in the cytosol. This sequence of events gives rise to an intermediate topological state in which only two of the first three AQP1 TM segments synthesized, cotranslationally acquired a membrane-spanning conformation (Shown in Figure 1).

A second finding was that several AQP1 topogenic events are not carried out with absolute fidelity in either expression system. Fusion proteins containing only the first three AQP1 TM segments generated two different topological isoforms, a major 2-spanning isoform in which TM3 C-terminal flanking residues resided in the cytosol, and a minor 3-spanning isoform with the TM3 C-terminus in the ER lumen. Pulse chase metabolic labeling studies carried out in the

presence and absence of MG132 further demonstrated that these isoforms are differentially recognized and degraded by the proteasome-mediated ERAD pathway. Selective degradation of the 2-spanning isoform therefore resulted in a progressive increase in the percentage of remaining polypeptides in the 3-spanning (mature) topology. This phenomenon was particularly apparent in HEK cells where very short pulse labeling times were required to confirm the presence of the predominant 2-spanning isoform. Its very short half-life of <25 min explains why the Turner group was unable to observe this species at steady state and after a 1 hr cyclohexamide chase (160). At these longer time points we too found that the 3-spanning isoform predominated. While selective degradation also occurred in oocytes, the time scale of degradation ($T_{1/2} > 15$ h) was much slower relative to the time course of the experiments. Thus the 2-spanning isoform was easily observed. Importantly, our results are completely compatible with data reported by the Turner group, however we have reached different conclusions. The previous discrepancies in AQP1 biogenesis are primarily due to the relative rates of protein synthesis and degradation in oocytes versus mammalian cells rather than fundamentally different behavior of ER translocation machinery.

The findings that truncated fusion proteins can give rise to multiple topological isoforms has several important implications for our understanding of how polytopic protein topology is cotranslationally established by the ribosome translocon complex (RTC). Given the rate of protein synthesis (~ 5 aa/sec) in mammalian cells (224) and the length of the ribosome exit tunnel (100 Å) (225), TM segments and short connecting loops would normally reside within the ribosome only for a few seconds before entering the translocon. Cotranslational topogenesis of small polytopic proteins such as AQP1 therefore requires that the translocon rapidly recognize topogenic determinants and dynamically direct regions of peptide into their proper cellular compartment (90). In the case of AQP1 fusion proteins, the C-terminal reporter should theoretically reflect the last triage decision of the translocon and follow the peptide loop into its appropriate compartment. This assumption serves as the basis for a large number of

studies in which C-terminal reporters have been used to determine protein topology (163,168,184,214,226). Our findings that different reporter domains and AQP1 truncation sites yield similar topological outcomes in very different cell types support the argument that choice of the translocation pathway (either cytosolic or luminal) is strongly determined by TM segments present within the RTC. At the same time, however, certain truncations appear to provide the reporter access to both the ER lumen and cytosol (albeit at different efficiencies) as evidenced by the different topologic isoforms generated. This indicates that the translocation pathway is not gated absolutely in one direction or another at all time points during AQP1 synthesis. Rather, the translocon appears to provide a certain degree of ambiguity in which TM segments and their flanking residues may transiently sample multiple topological configurations as the nascent chain extends from the ribosome. This finding is consistent with recent studies demonstrating that N-terminal signal sequences can regulate cytosolic accessibility of translocating proteins (and hence translocation efficiency) in a graded manner (72,227,228).

The substitution of large reporters such as the prolactin domain or the β -subunit, in lieu of the native TM segments, likely forces the translocon to make a decision as to where the polypeptide should be directed based on the topogenic information available up to the particular truncation site. We do not yet know how flexible the environment is within the ribosome-translocon complex nor how long a TM segment may be allowed to sample different topological spaces. However, in the case of AQP1, the relative proportion of different isoforms may be viewed as reflecting relative accessibility of the reporter to ER and cytosolic compartments. Such flexibility in translocon function might also facilitate proper re-orientation of TM3 at later stages of synthesis as more folding information is provided. This is precisely what we observed using small epitope tags inserted into the TM2-3 and TM3-4 peptide loops. TM3 topology gradually changed into its mature orientation as additional TM segments and C-terminal folding information were synthesized (81,82). Similar results have been reported for

other polytopic proteins in which TM segment reorientation may take place within the translocon (217).

One question that arises from these studies is why do different AQP1 topological isoforms exhibit such markedly different stabilities. Because exposed hydrophobic patches can act as signals for ERAD (200), the 2-spanning isoform could be degraded because TM2 is exposed to the ER lumen (See Figure 1A). In contrast, topology of the 3-spanning isoform would more closely resemble the mature protein. The ERAD machinery could therefore provide an efficient mechanism for removing proteins that exhibit abnormal topology in the ER membrane.

Alternatively, stability may also be influenced by the reporter. Both reporters used here are normally expressed in ER lumen and localization to the cytosol might therefore disrupt specific folding requirements such as disulfide bond formation (prolactin) or glycosylation (β -subunit). While we do not know the extent to which this occurs, the choice of a seemingly "inert" reporter may have significant implications when topologic studies are carried out in intact cells with active ERAD pathways, and particularly when the reporter resides in both ER and cytosolic compartments. Our results demonstrate the importance of considering protein stability when performing topology studies with truncated intermediates because failure to examine topology at early time-points can lead to erroneous conclusions when multiple isoforms are generated.

This study provides the first direct comparison between the early biogenesis mechanisms of oocytes and mammalian cells using a unique membrane protein substrate. Because oocytes provide a useful and convenient system for heterologous expression of many biologically important proteins (229), it is reassuring that early biogenesis events concur with those in mammalian cells. Indeed, when differences in protein stability are taken into account, the translocation efficiency and relative proportion of different isoforms for all three truncations examined here were in remarkably good agreement. This observation held even when different reporters, reporter readouts and fusion sites were compared. In addition, the topology of all

other regions of AQP1 was essentially identical (78,160). Our findings are therefore consistent with the strong conservation of ER translocation machinery (62,230,231) and indicate that the unusual biogenesis mechanism utilized by AQP1 reflects a well conserved pathway of protein folding.

In this regard, we previously established that in *Xenopus* oocytes, once AQP1 synthesis is completed, and all 6 TM segments are present, the protein acquires a stable conformation due to reorientation of TMs2-4 and repositioning of the TM2-3 and TM3-4 connecting loops (23). Selective degradation is therefore observed only for truncated fusion proteins that lack C-terminal folding information and are therefore trapped in an unstable topology. Importantly, we also found this to be the case in mammalian cells. In contrast to constructs containing only 3 TMs where the major isoform was degraded with a $T_{1/2}$ of ~25 min, full length AQP1 protein was remarkably stable. Yet our data show that TM3 was cotranslationally directed into the unstable $N_{\text{exo}}/C_{\text{cyt}}$ topology in 70% of these polypeptides. Thus, generation of stable AQP1 protein in mammalian cells, as in oocytes, is not achieved by selective degradation of an unstable four-spanning intermediate. Instead, synthesis of C-terminal residues is able to confer stability on a protein region that is initially inserted into the membrane in an incorrect conformation. Given that mammalian cells and oocytes synthesize and process AQP1 into functional water channels with six TM segments (5,232,233), and that AQP1 TM segments are cotranslationally directed into the same immature topology in both systems, it seems reasonable to speculate that mammalian cells also have the capacity to complete AQP1 folding by facilitating topological reorientation of TMs 2-4 in the ER.

We should point out, however, that the purpose of our study was primarily to compare cotranslational translocation events and does not address the mechanism(s) by which AQP1 ultimately acquires its mature six-spanning topology. Specifically, our studies do not address the efficiency of late trafficking events, but rather they examine stability of protein during early

maturation in the ER which is where the 6-spanning topology is initially achieved (23). In this regard, our results contrast with previous findings by Leitch et al. that showed AQP1 degradation can be decreased by MG132 (and hypertonic stress) over longer time intervals (234). Additional work is therefore required to determine conclusively if and how AQP1 reorientation takes place in mammalian cells. The extent to which this unusual AQP biogenesis pathway is utilized by other proteins and the mechanism by which this process is carried out within the RTC also represent important future challenges for unraveling the complex process of membrane protein biogenesis.

Acknowledgments

The authors thank Mr. J. Eledge for his excellent technical assistance, Dr. R. J. Turner for generously providing the plasmid pEGFP.AQP1T120.b, Dr. A. S. Verkman for supplying AQP1 antisera, and members of the Skach lab for their helpful comments. This work was supported by NIH, GM53457, DK51818, and AHA Established Investigator Grant (W.R.S.) and HL007781 (T.M.B.)

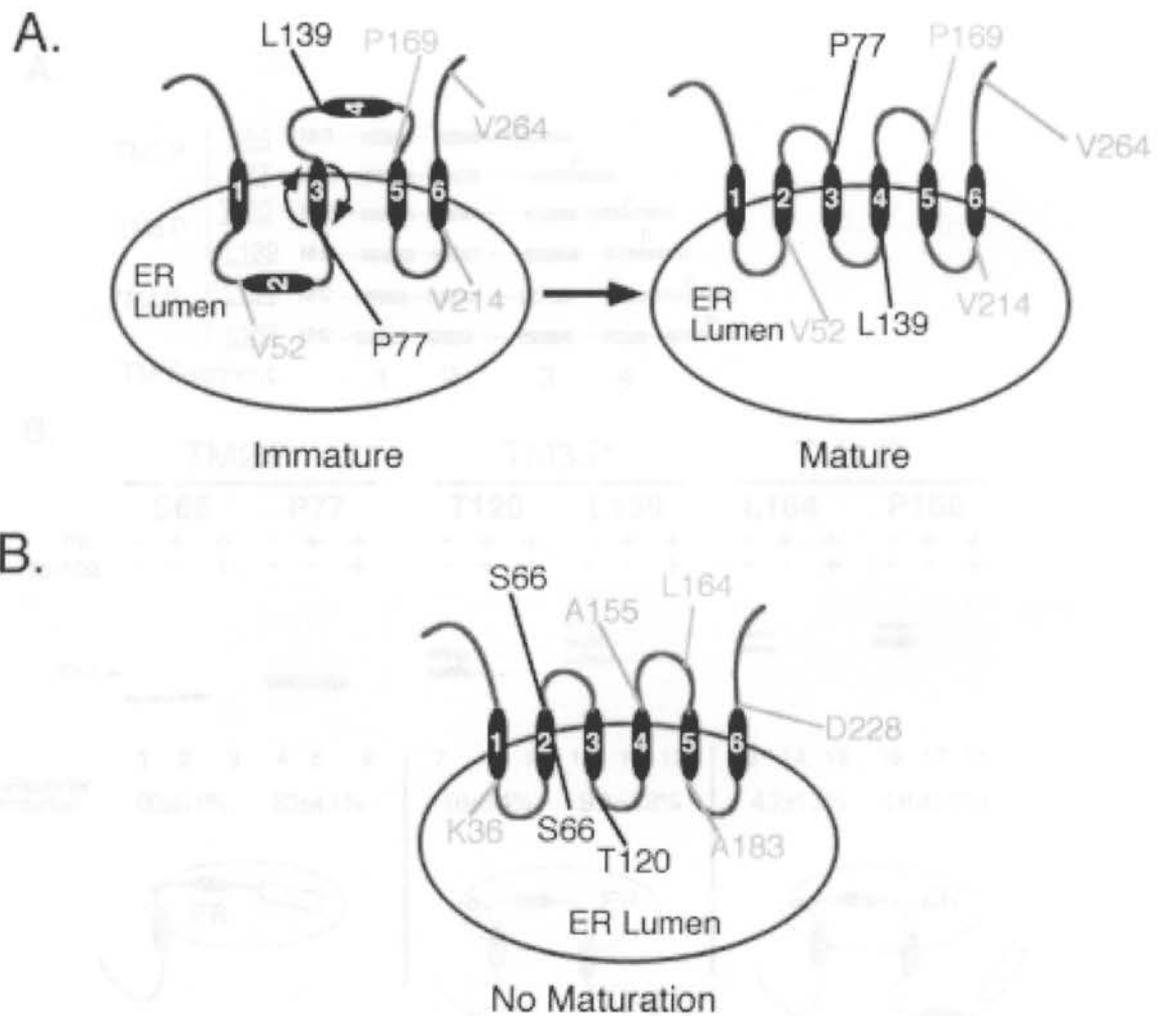


Figure 2-1: Two models for AQP1 biogenesis. Topology of reporter fusion sites used for experiments in *Xenopus* oocytes (**A**) and HEK-293 cells (**B**) are indicated. Fusion sites in which topology differs between these systems are shown in black while fusion sites with the same topology are shown in gray. In oocytes AQP1 is initially synthesized as a 4-transmembrane spanning intermediate that is converted into the mature 6-spanning form. In HEK-293 mammalian cell reporter topology did not identify the immature topology (160).

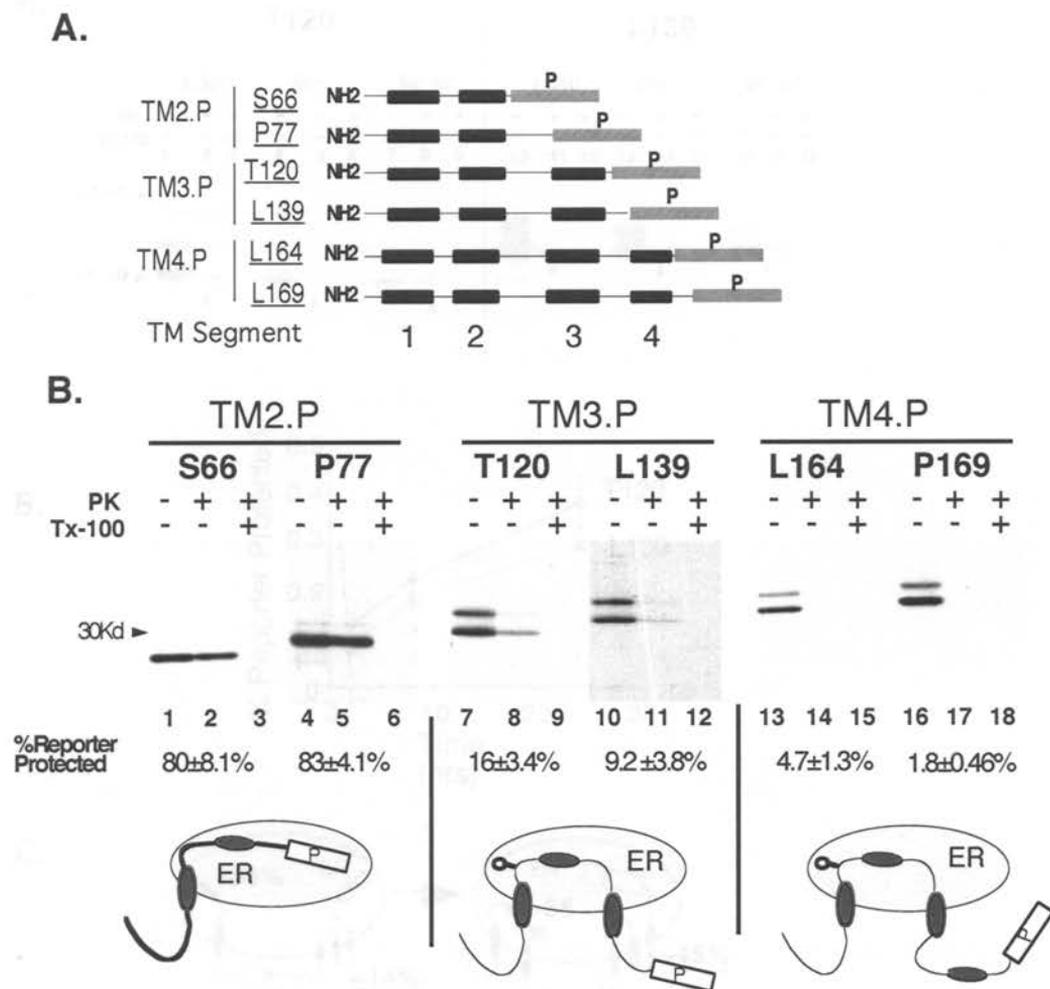


Figure 2-2: Location of fusion sites does not affect early AQP1 biogenesis events in oocytes. (A) Diagram of AQP1 fusion proteins showing TM segments (black rectangles) and prolactin-derived (P) reporter (Gray rectangles). (B) Constructs in A were expressed in oocytes and reporter topology was determined by Proteinase K (PK) digestion in the presence or absence of TX-100. Shown is a representative autoradiogram of prolactin reactive immunoprecipitated products. Percent of reporter protected from PK digestion was normalized to a control secretory protein. Topology of fusion sites are illustrated under autoradiogram. Data show average of 3 or more experiments +/-SEM.

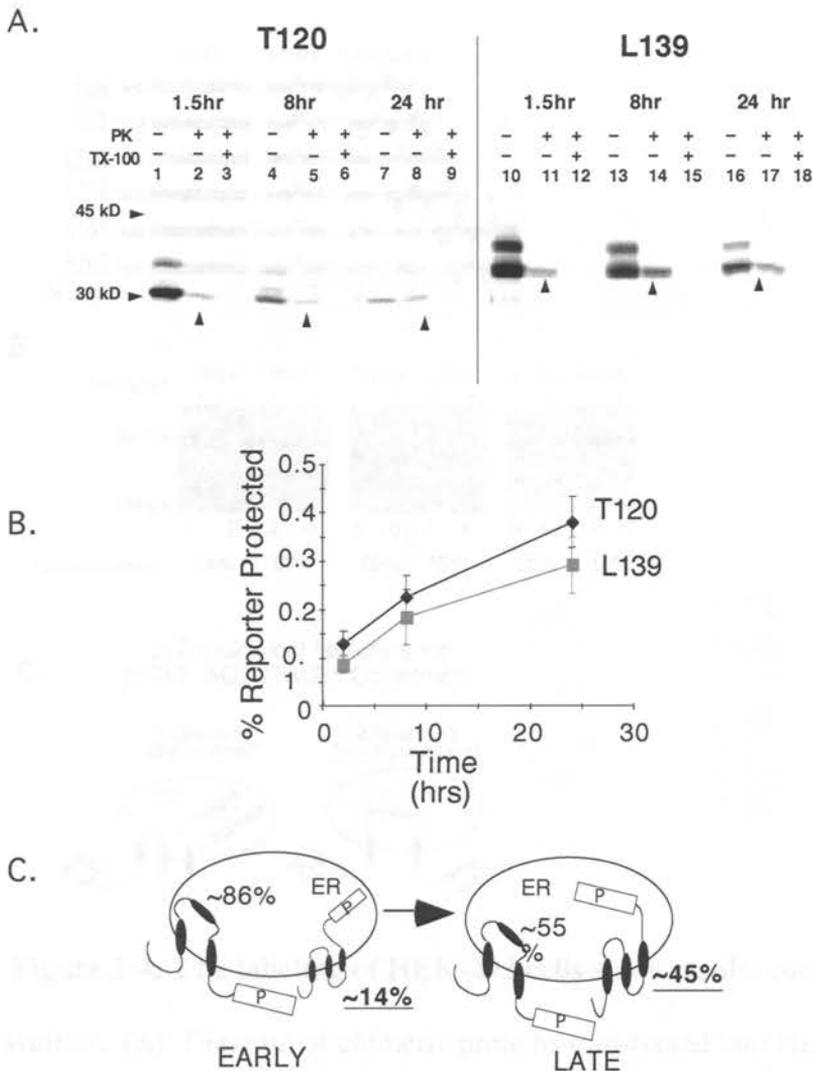
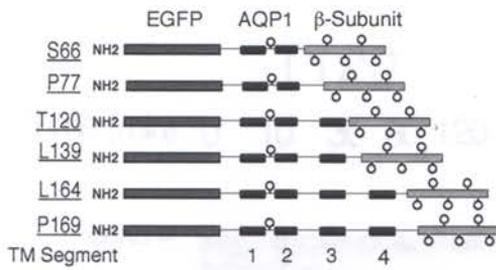


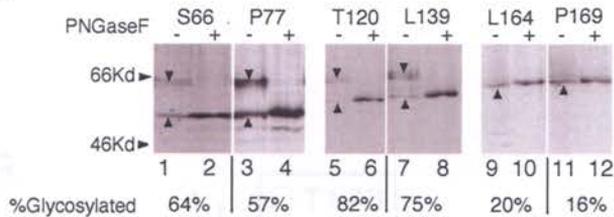
Figure 2-3: Differential stability of biogenesis isoforms truncated following TM3:

(A) Autoradiogram of AQP1 truncated in the TM3-4 loop and fused to the P reporter at residues T120 and L139 as indicated. Constructs were expressed in oocytes for the times indicated and homogenates were digested with PK in presence or absence of TX-100. **B.** Results from A were quantitated and the % of reporter protected at each time point was plotted. Results show averaged values from three independent experiments \pm SEM. **C.** Model of isoform profile. At early points the 2-spanning isoform predominates while the proportion of 3-spanning isoform increases \sim 3 fold over the time course of the experiment.

A.



B.



C. 2-Topological Isoforms for EGFP.AQP1TM2.β Constructs

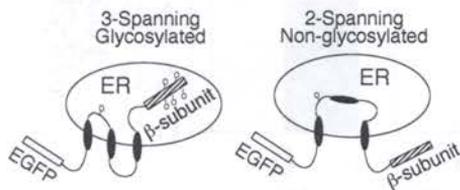
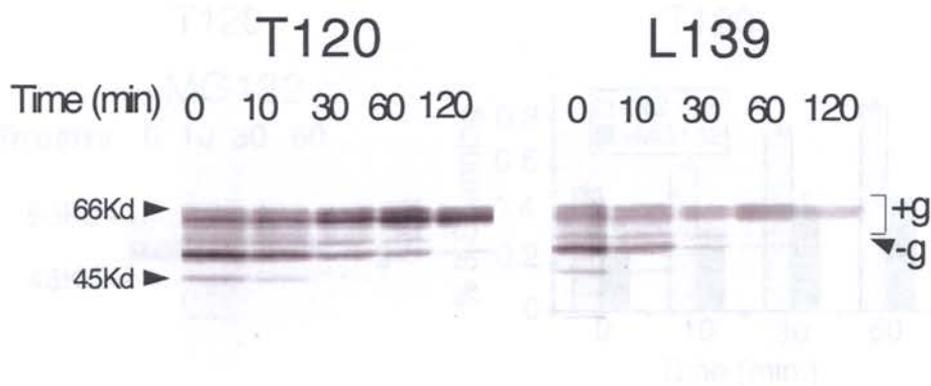


Figure 2-4: 2 hr labeling of HEK-293 cells show results consistent with previous studies. (A) Diagram of chimeric proteins transfected into HEK-293 cells. Open circles represent N-linked glycosylation sites on AQP1 and β -subunit. EGFP and β -subunit domains and AQP1 TM segments are indicated. (B) Autoradiogram of 2-hr pulse labeled HEK-293 cells immunoprecipitated with GFP antisera prior to and after PNGase F digestion. Downward arrows indicate glycosylated proteins as documented by PNGase F digestion. Upward arrows indicate nonglycosylated protein. Percent glycosylation of constructs is indicated under autoradiogram. (C) Diagram of the topological isoforms deduced from the T120 and L139 chimeric proteins.

A.



B.

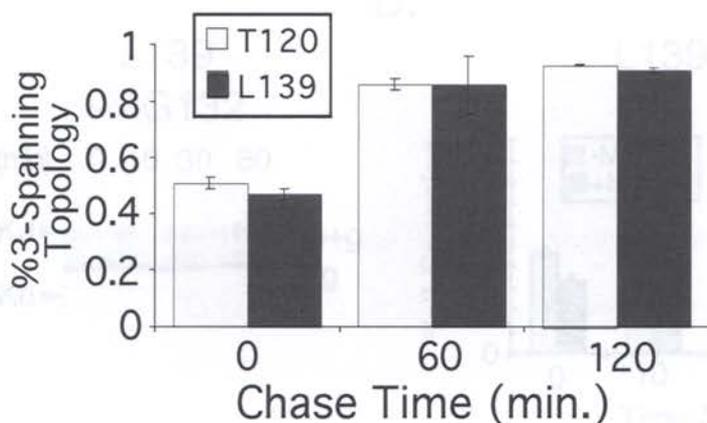


Figure 2-5: Differential stability of AQP1 isoforms in HEK-293 cells. (A) Cells were transfected with constructs (Figure4), labeled with ^{35}S -methionine for 15 min, chased with complete media for the indicated times, lysed and subjected to SDS-PAGE and autoradiography. Nonglycosylated (-g) and glycosylated (+g) proteins are indicated. (B) At T=0, ~50% of protein was recovered as a 50 kD nonglycosylated band which rapidly disappeared during the chase period. At later time-points the 3-spanning (glycosylated) protein predominated. Half-lives of the 2-spanning (nonglycosylated) and 3-spanning (glycosylated) isoforms were approximately 25 minutes and > 120 min, respectively. Data shows average of 3 or more experiments +/-SEM.

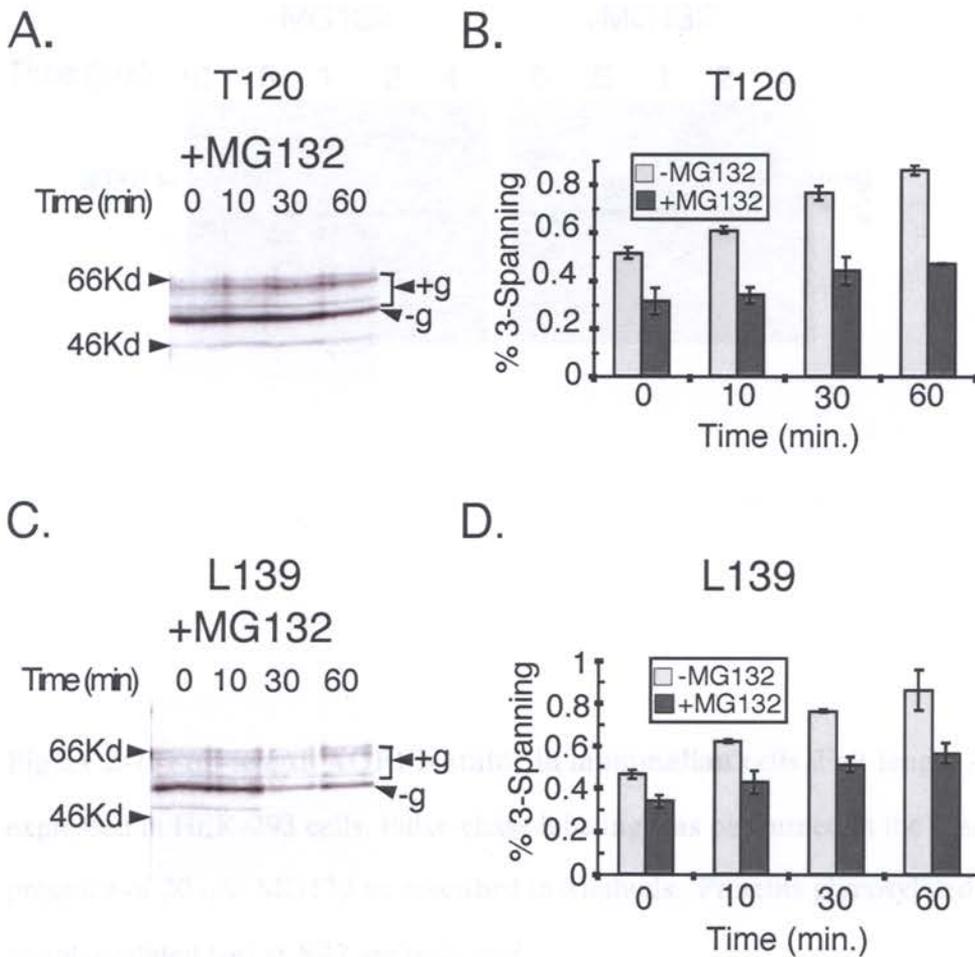


Figure 2-6: MG132 blocks selective degradation of non-glycosylated isoforms. (A, C). EGFP... β -subunit chimeras truncations 120 and 139 (see Figure 4) were expressed in HEK-293 cells. Pulse chase labeling was performed in the presence of 20 mM MG132 as described in Methods. Glycosylated (+g) and nonglycosylated (-g) proteins are indicated. (B, D) Percent of protein in the 3-spanning (glycosylated) topology was quantitated for experiments carried out with (dark gray) or without (light gray) MG132. Data shows average of 3 experiments +/-SEM.

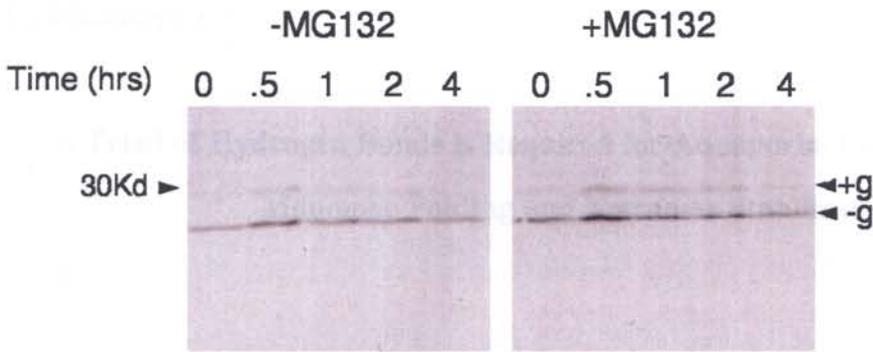


Figure 2-7: Full-length AQP1 is stable in mammalian cells. Full-length AQP1 was expressed in HEK-293 cells. Pulse-chase labeling was performed in the absence or presence of 20 μ M MG132 as described in Methods. Proteins glycosylated (+g) or nonglycosylated (-g) at N42 are indicated.

C. Manuscript 3

A Triad of Hydrogen Bonds is Required for Aquaporin-1 early Topogenesis, Monomer Folding and Tetramer Stabilization

Teresa M. Buck, Justin Wagner and William R. Skach

Department of Biochemistry and Molecular Biology, Oregon Health & Sciences University

Manuscript in Preparation, 2006

Address Correspondence to:

William R. Skach, M.D.
Professor of Medicine and Cell and Developmental Biology
Biochemistry and Molecular Biology
Mail Code: L224
3181 SW Sam Jackson Park Road
Portland, OR 97239

Phone: 503-494-7322
Fax: 503-494-8393
Email: skachw@ohsu.edu

1. Abstract

The topology of most polytopic membrane proteins is established cotranslationally at the ER membrane. The AQP1 water channel however, is initially synthesized as a 4-spanning intermediate that undergoes a reorientation of three internal TMs (TMs2-4) to acquire its mature 6-spanning structure. In contrast, a related protein, AQP4, acquires its topology cotranslationally as each TM emerges for the ribosome. Previous work has shown that this is due, in part, to two unique polar residues, Asn49 and K51, at the N-terminus of TM2 that prevents it from initially spanning the membrane, but allows it to pass into the ER lumen. Here we investigate the role of these residues in the early events of AQP1 maturation and function. We show that both Asn49 and Lys51 form hydrogen bonds with an Asp residue in TM5 (Asp185) that is critical for AQP1 folding and function. This creates four-fold symmetry in the AQP1 tetramer in which the interaction between Asn49 and Asp185 is intramolecular and required for correct AQP1 monomer folding. In contrast, Lys51 and Asp185 form hydrogen bonds between monomers and thereby stabilize the AQP1 tetramer. Functional studies of AQP1 mutants indicate that Asn49 and to a lesser extent Lys51 are needed to compensate for the presence of Asp185, in TM5 which in turn is responsible for AQP1's unique biogenesis mechanism. These data thus provide the structural and functional basis for AQP1's unique biogenesis mechanism and identify a novel structural motif involved in AQP tetramerization.

2. Introduction

Aquaporin (AQP) tetramers passively transport water and/or glycerol across cellular membranes. Each monomer forms a functional channel that exhibits a conserved 6-spanning structure arranged in an inverted hourglass with pseudo-twofold symmetry. Thirteen human aquaporins (AQP0-AQP12) have been identified and are expressed in diverse tissue types including lens fiber cells, kidney tubules, intestine, brain, skeletal muscles, liver and red and white blood cells (2). AQP function has been most extensively studied in the kidney where it plays a key role in water reabsorption. AQP1 is highly expressed in the proximal tubule and descending loop of Henle (3,171) and is responsible for reabsorbing ~70% of glomerular filtrate. AQP1-null humans as well as mice have a mild defect in urinary concentration as manifest by decreased tolerance to fluid withdrawal. Mice also exhibit additional phenotypes including reduced tumor growth and injury induced swelling of the cornea (2,9).

Mammalian AQPs exhibit a high degree of homology and a conserved structure. Recent crystal structures have defined the specific mechanism responsible for water and/or glycerol selectivity and have confirmed that each monomer forms a water conducting channel (2,95). AQPs fold and assemble in the ER and are trafficked through the cell as homotetramers. Despite many advances in AQP biology, little is known about how AQPs achieve their complex architecture in cells. In general folding of polytopic proteins such as AQPs can be broken into three stages, orientation of TMs into the membrane, helical packing/monomer folding, and oligomerization.

Early topogenesis events of two highly homologous AQPs (AQP1 and AQP4) exhibit striking differences. AQP4 topogenesis follows a relatively simple model of polytopic protein biogenesis (76,77) in which each TM segment acts alternatively as either a signal anchor (TM1,

TM3 and TM5) or stop transfer (TM2, TM4 and TM6). For example, TM1 targets the ribosome nascent chain complex to the ER, gates open the Sec61 translocon, and initiates translocation of the first extracellular protein loop. TM2 then functions as a stop transfer sequence to close the translocon and direct the next intracellular loop into the cytosol. TM3 acts as a signal anchor once again reinitiating translocation and so on. In this manner topology of AQP4 is established cotranslationally as the nascent polypeptide exits the ribosome. In contrast, studies of AQP1 topogenesis have shown that while TM1 acts as an efficient signal anchor sequence; TM2 does not terminate translocation but rather is transiently translocated into the ER lumen. As a result TM3 acts as a stop transfer and initially spans the membrane in a reverse orientation from that of the mature protein. TM4 is unable to reinitiate translocation and remains in the cytosol. Thus AQP1 is initially synthesized as a 4-spanning protein that must undergo an internal reorientation of TMs2-4 to establish a mature topology. Additionally, these studies identified two residues Asn49 and Lys51 in TM2 that when substituted for the corresponding AQP4 residues (N49M/K51L) converted TM2 into an efficient stop transfer but destroyed water channel function (77).

While few studies have examined the details of AQP tetramerization it is generally assumed that tetramers form in the ER membrane after the polypeptide is released from the translocon but before transport to the Golgi. AQPs traffic as homotetramers, yet the structural determinants for self-association remain undefined (109,110). Some evidence suggests that the strength of the intermolecular associations between AQP1 monomers may be linked to substrate specificity (glycerol vs. water) (113). These studies showed that unlike AQP, GlpF intermolecular interactions are sensitive to nonionic detergent (111,112), and interestingly AQP1 can be converted from a water to glycerol channel by making the mutations Y222P/W223L (95). These mutations also lead to weakened intermolecular interactions, which resemble those of the glycerol channel, GlpF. Residues in AQP loops B and E have been implicated in the oligomerization process. In fact two residues in AQP1,

A73 and C189, are necessary for both water conduction and oligomerization (103). Intriguingly, these residues are located in loops B and E, which fold inward to form the pore, and therefore can not be involved in direct interactions between subunits, but more likely have an indirect effect on the overall channel structure. These data suggest that intramolecular interactions that promote pore folding are also required for tetramerization.

To further define the mechanisms of AQP1 topogenesis, monomer folding, and oligomerization we investigated the role two unique residues, Asn49 and Lys51, play in these processes. We showed that Asn49 and Lys51 exhibit synergistic effects in disrupting TM2 stop transfer activity. In addition, both residues are required to form hydrogen bonds with a unique Aspartate at position 185 in TM5. Interactions between Asn49 and Asp185 are intramolecular and are needed to stabilize the monomer while Lys51 and Asp185 form hydrogen bonds between monomers that stabilize the tetrameric structure. Thus we define a structural relationship responsible for functional complementation that explains the unique biogenesis of AQP1. Using AQP1 as a model protein, we show that a complex arrangement of hydrogen bonds between Asn49, Lys51 and Asp185 affect protein folding at three different levels: early topogenesis, monomer folding, and tetramer association.

3. Materials and Methods

cDNA Construction. Plasmid pSP64.CHIP28 (AQP1) (78) or pSP64.MIWC (AQP4) were used as the templates for PCR amplifications to generate AQP1 and AQP4 derived constructs (76,78). The myc-tagged AQP1 constructs were constructed as previously described (81) with the 10 residue myc epitope (EQKLISEEDL) inserted at position T120 in the AQP1 sequence. Fusion proteins containing the C-terminal prolactin translocation reporter were generated by amplifying wild type or mutant AQP coding sequences using a sense oligonucleotide (SP6 promoter) and antisense oligonucleotides encoding a BstEII

restriction site at AQP1 codon Pro77 or AQP4 codon Gly72. Antisense oligonucleotides were identical to those described previously (76,78). PCR fragments were digested with NcoI and BstEII (AQP1 constructs) or HindIII and BstEII (AQP4 constructs) and ligated 5' to a translocation reporter (P), which encodes 142 C-terminal residues from the secretory protein bovine prolactin. This reporter contains no intrinsic topogenic information and faithfully follows topogenic determinants (78,184). The AQP1 (N49M/K51L, N49M/K51L/D185N, N49M, K51L, D185N, K51D, and K51D/D185N) and AQP4 (M48N/L50K, N184D, and M48N/L50K/N184D) mutants were generated with site-directed mutagenesis (PCR overlap extension; (183)). All cloned fragments were verified by DNA sequencing.

In Vitro Transcription/Translation. cDNA was transcribed in vitro using SP6 RNA polymerase (New England Biolabs, Beverly, MA) using 2 μ g of plasmid DNA per a 10 μ l reaction volume at 40 °C for 1 hr as previously described (78). Aliquots were used immediately or flash frozen with liquid nitrogen and stored at -85 °C. For *in vitro* translation RRL was prepared from New Zealand White Rabbits as described (235,236). Translation was performed at 25 °C for 1 hr in reactions containing 20% transcript mixture, 40% nucleated rabbit lysate, 1 mM ATP, 1 mM GTP, 12 mM creatine phosphate, 40 μ M each of 19 essential amino acids except methionine, 1 μ Ci/ μ l of Trans³⁵S-label (ICN Pharmaceuticals, Irvine, CA), 40 μ g/ml creatine kinase, 0.2 U/ μ l RNase inhibitor (Promega, Madison, WI), 10 mM Tris-Cl pH=7.5, 100 mM KOAc, and 2mM MgCl₂. Canine pancreas microsomal membranes, prepared as described (237) were added at the start of translation (final concentration, 7.0 OD₂₈₀).

Xenopus laevis Expression. Stage VI oocytes were harvested from *Xenopus laevis* frogs as described (238). 2 μ l of mRNA transcribed as described above was mixed with 50 μ Ci of [³⁵S] methionine (0.5 μ l of a 10X concentrated Trans³⁵S-label; ICN Pharmaceuticals, Irvine, CA), and 50 nl was injected into each stage VI *X. laevis* oocytes on an ice-cold stage.

The oocytes were then incubated at 18 °C in MBSH (88 mM NaCl, 1 mM KCl, 24 mM NaHCO₃, 0.82 mM MgSO₄, 0.33 mM Ca(NO₃)₂, 0.41 mM CaCl₂, 10 mM HEPES pH=7.4, 50 µg/ml gentamycin, 100 units/ml penicillin, and 100 µg/ml streptomycin sulfate). Oocyte microsome membranes were isolated by pelleting through 0.5 M sucrose in Buffer A (20 mM Tris-Cl pH=7.5, 5 mM MgCl₂) at 180,000 x g for 10 min as described (236,239) onto a 1.8 M sucrose cushion with the same salt concentrations for 15 min at 180,000 x g. 100 µl of the interface between the sucrose layers was taken and mixed with three volumes of 20 mM Tris-Cl pH=7.5, 5 mM MgCl₂ and centrifuged for 10 min at 180,000 x g. The resulting pellet was solubilized as described below.

Protease Protection Assay. Oocytes were injected as described above, incubated at 18 °C for 3 hr, and homogenized by hand in 3 volumes of 0.25 M sucrose, 50 mM KAc, 5 mM MgAc₂, 1.0 mM dithiothreitol, 50 mM Tris-Cl, pH=7.5. The homogenates were divided into three 10 µl aliquots on ice. Proteinase K (Roche Applied Science, Indianapolis, IN) was added (final concentration 0.2 mg/ml) in the presence or absence of 1% Triton X-100. The reactions were incubated on ice for 1 hr and then mixed with phenylmethylsulfonyl fluoride (10 mM) and boiled in 10 volumes of 1% SDS, 0.1M Tris-HCl pH=8.0 for 5 min to stop the reaction. Samples were then diluted in 10 volumes Buffer A (100 mM Tris-Cl pH=8.0, 100 mM NaCl, 1% Triton-X100, 2 mM EDTA), incubate at 4 °C for 15 min and centrifuged at 16,000 x g for 10 min at 4 °C to remove insoluble debris. Samples were then immunoprecipitated as described. Efficiency of the assay was regularly assessed using a known secretory control protein and was >80%.

Aquaporin Functional Assay-Oocyte Permeability. cRNA was transcribed from linearized plasmids using SP6 polymerase as described above. 2 µl of AQP or mock cRNA was mixed with 0.5 µl secretory alkaline phosphatase (SEAP) cRNA and injected into oocytes as described above. Oocytes were incubated at 18 °C for 24 hr and then transferred to

individual 96-well dishes with 150 μ l MBSH and incubated for another 24 h. Alkaline phosphatase assays were performed for each oocyte to determine cells functionally expressing protein. As described by Tate et al.(240) briefly, 100 μ l of MBSH from each oocyte was mixed with 100 μ l SEAP assay buffer(1M diethanolamine pH=9.8, 0.5 mM $MgCl_2$, 0.02 mM $ZnSO_4$, and 20 mM L-homoarginine) and 50 μ l p-nitrophenyl phosphate (Bio-Rad, Hercules, CA) in a 96-well plate and incubated at 37 °C for 1-2 hr. Plates were read at 405 nm, and oocytes giving at least two times background absorbance were chosen for functional assays.

Oocyte swelling was assayed in response to a 10-fold dilution of extracellular MBSH buffer with distilled water. Oocyte volume was measure in 1-s intervals by quantitative imaging and P_f was calculated from the initial swelling rate (first 30-s) $(dV/V_0)/dt$ by the relation $P_f = [d(V/V_0)/dt]/[(S/V_0)V_w(Osm_{out}-Osm_{in})]$, where $S/V_0=50cm^{-1}$, $V_w=18cm^3/mol$, and $Osm_{out}-Osm_{in}= 180$ mOsm. 4-7 oocytes were analyzed per experiment for each construct. Typical experiment is shown.

Sucrose Gradient Centrifugation. Microsomal pellets from oocyte preparations were solubilized in 20 mM Tris-Cl pH=7.4 containing either 1% Triton-X100 or 1% SDS by repeated pipetting, samples were then left at 4 °C for 15 min this process was repeated once and then insoluble debris was removed by centrifugation for 10 min at 16,000 x g. Velocity sedimentation on a sucrose gradient was performed essentially as described by Lagree (109). 2-20% linear sucrose gradients were prepared from 2 and 20% sucrose stocks in 20 mM Tris-Cl pH=7.4 containing 0.1% Triton-X100 or 0.1% SDS. Solubilized proteins were loaded on top of gradient and centrifuged at 100,000 x g for 14 hrs at 4 °C. 14 equal fractions were collected from each gradient and analyzed by SDS-PAGE following immunoprecipitation.

Immunoprecipitation. For immunoprecipitation, anti-prolactin antiserum (ICN Biomedicals, Costa Mesa, CA), mAb myc-9E19 (181) (mouse ascites), or polyclonal rabbit anti-AQP1 antisera raised against purified AQP1 protein (generously provided by A. Verkman, UCSF, San Francisco, CA) at 1:1000 dilution was added to proteins solubilized in Buffer A. For native immunoprecipitations membranes solubilized in 1% TX-100 as described above were directly added to Buffer A. For denaturing IPs proteins were initially denatured in 1% SDS prior to dilution with Buffer A. 5.0 μ l of protein A-Affigel (Bio-Rad, Richmond, CA) was added and samples were rotated at 4 °C for ~10 h before washing 3 times with buffer A and twice with 0.1 M NaCl, 0.1 M Tris-Cl pH=8.0. SDS loading buffer was added to remove proteins from beads followed by SDS-PAGE analysis.

PNGaseF Digestion. Beads from immunoprecipitations were resuspended after the final wash in 15 μ l 0.1 M Tris-Cl pH=7.5 and 0.3 μ l of PNGase F (New England Biolabs, Beverly, MA), and incubated at 37 °C for 3 hr. The samples were then analyzed by SDS-PAGE as described above.

Autoradiography and Quantitation. SDS-PAGE gels were analyzed using EN³HANCE (Perkin Elmer Life Sciences, Wellesley, MA) fluorography. The intensity of recovered bands was quantitated using a Bio-Rad personal Molecular PhosphorImager FX (Kodak screens, Quantity-1 software).

4. Results

Relative role of polar TM2 residues in AQP topogenesis and function. AQP1 and AQP4 acquire their mature 6-spanning topology by fundamentally different mechanisms (76,78). Topology of AQP4 is established cotranslationally by locally encoded topogenic information as each TM exits from the ribosome. In contrast, AQP1 is initially synthesized as a 4-spanning intermediate and undergoes an internal reorientation of TMs 2-4 to achieve its mature six-spanning structure (Figure1C). Using sequentially truncated AQP

polypeptides and a well characterized C-terminal translocation reporter assay, we previously showed that the topogenic activity of TM2 plays a key role in AQP folding pathways. TM2 efficiently terminates nascent chain translocation and cotranslationally spans the ER membrane in the case of AQP4, however AQP1-TM2 transiently passes through the translocon and initially resides in the ER lumen (Figure 2B, C). This latter behavior requires AQP1-TM2 to be inserted into the membrane at a late stage of synthesis (i.e. during or after synthesis of TMS 4-6) via cooperative interactions between AQP1 N- and C-terminal halves.

A comparison of sequences from all 13 mammalian aquaporins revealed high homology within the TM2 segment with the notable exception of two N-terminal polar residues (Asn49 and Lys51) that were unique to AQP1 (Figure 2A). Protease protection experiments further revealed that replacing these two residues with corresponding hydrophobic residues from AQP4 (Met and Leu, respectively) converted AQP1-TM2 into an efficient stop transfer sequence such that TM2 spanned the membrane in a type I topology with its C-terminus (and reporter) in the cytosol (Figure 2B, lanes 1-6 and (77)). However, the resulting mutant AQP1 protein, N49M/K51L, failed to generate functional water channels (Figure 3C and (77)) suggesting that these polar residues are involved both in the early stages of topogenesis and subsequent aspects of AQP1 folding and/or function. To understand the interrelationship between early AQP folding events in the ER membrane and acquisition of water channel function, we therefore examined the contribution of these unique residues, Asn49 and K51, towards cotranslational topogenesis, monomer formation, tetramerization and water channel function.

We initially suspected that the charged Lys residue at position 51 was primarily responsible for disrupting TM2 stop transfer activity. However, when N49M and K51L mutations were introduced independently, neither polar residue significantly changed TM2 topology

(Figure 2B, lanes 7-12). Reversing the charge at residue 51 from Lys to Asp also had no effect on TM2 topology (Figure 2B, lanes 13-15) even though this significantly changed the net cytosolic charge flanking TM2 from +2 to +4 which would be predicted to favor a type I topology according to the positive in rule (161). Thus both Asn49 and Lys51 residues contribute synergistically to the unexpected topogenic behavior of TM2 and hence play a primary role in the unusual folding pathway of AQP1. Consistent with this, introducing the complementary insertion of Asn and Lys into AQP4 (M48N/L50K) disrupted TM2 stop transfer activity and reoriented TM2 C-terminus flanking residues from a cytosolic to a predominantly luminal orientation (Figure 2C, lanes 1-6). Although both Asn and Lys are unique to AQP1 (Figure 2A), most AQPs have at least one polar residue at the position that corresponds to K51. Only AQP7 (Tyr) and AQP12 (Asp) have polar residues that correspond to N49. It is likely therefore, that the AQP1 folding pathway is relatively unique among the human AQP family. Interestingly, GlpF has a polar residue (Gln) similar to AQP1 suggesting that it too may have an unusual topogenesis mechanism.

Interactions between TM2 and TM5 are important for AQP1 function. A remarkable feature of AQP1 folding is that information within the C-terminal half of the protein is required to reorient the N-terminal half and properly position TMs 2, 3, and 4 within the plane of the ER membrane. Because Asn49 and Lys51 are highly polar, unique among the aquaporin family, and not directly involved in the water conduction pathway, we hypothesized that this reorientation might involve ionic or polar interactions between TM2 and specific residues in the C-terminal half of AQP1. The AQP1 crystal structure indicates that TM2 makes extensive physical contacts with TM5 such that its N-terminus is juxtaposed against the C-terminus of TM5 (Figure 1). Further, comparison of mammalian AQP sequences revealed a unique acidic residue in AQP1, Asp185, that was a prime candidate for such an interaction (Figure 3A). We therefore examined the reciprocal role of AQP1-TM2 and TM5 polar residues in generating functional water channels in

microinjected *Xenopus* oocytes by replacing residues in AQP1 with corresponding residues from AQP4.

As previously shown, oocytes expressing AQP1 N49M/K51L completely lacked water channel activity (Figure 3B). However, the K51L mutation alone had no effect on AQP1 function, and N49M reduced overall activity by only ~50%. Thus as was observed for topology, these residues also exhibit synergistic effects on AQP1 function. Remarkably, converting Asp185 to Asn (D185N) had no independent effect on AQP1 function, but completely restored function of the N49M/K51L mutant (Figure 3B). These results thus establish a functional interaction between N- and C-terminal halves of AQP1 that is localized to specific regions of TMs 2 and 5. It is unlikely that a simple salt bridge between these residues is responsible for functional complementation, because exchanging residues Lys51 and Asp185 (K51D/D185K) did not restore AQP1 activity (even though K51D completely disrupted AQP1 function (Figure 3B)). Rather, the functional effects observed for each individual mutation N49M, K51L and D185N strongly suggest that acquisition of AQP1 function involves a complex tertiary interaction between all three residues.

To test whether residues Asn49, Lys51 and Asp185 play a general role in AQP biology, we examined the functional effects of introducing these residues into the corresponding locations in AQP4. In striking contrast to AQP1, introducing polar residues into TM2 (M48N/L50K) had no effect on AQP4 function (Figure 3C). However, the N184D mutation reduced function by nearly 60% and activity was fully complemented by M48N/L50K mutations. These results indicate that the structure/function relationship between TM2 and TM5 reflects conserved architecture among different AQPs. In addition, they suggest that the primary requirement for polar residues in AQP1 TM2 (N49/K51) is to compensate for Asp185 at the C-terminus of TM5. Both proteins tolerate Asn at this location regardless of TM2 residues. In addition, both proteins require polar TM2 residues

for full function when Asp is present. Thus, in order to retain AQP1 function, Asp185 requires the presence of Asn49 and Lys51 in TM2, which in turn disrupts TM2 stop transfer activity, force TM2 to transiently enter the ER lumen during synthesis, and require that AQP1 utilize its characteristic complex folding pathway.

Our results identify a novel and complex interaction that physically couples residues Asn49, Lys51 and Asp185 within the primary AQP1 sequence. The structural mechanism for this functional complementation is elegantly illustrated by the crystallographic structure shown in Figure 4. Asn49 forms a hydrogen bond with one carbonyl oxygen of Asp185 on adjacent TMs within the AQP1 monomer (Figure 4A). In contrast, the ϵ -amino group of Lys51 is precisely aligned with the second carbonyl oxygen of Asp185 on an adjacent monomer. Thus Asp185 forms two hydrogen bonds, one in “cis” with Asn49 and the other in “trans” with Lys51 across the tetrameric interface (Figure 4B-C). Our functional studies demonstrate that the hydrogen bond between Asn49 and Asp185 in AQP1 (or corresponding Asn48 and Asp184 in AQP4) is fundamentally required for water channel activity, whereas the Lys51-Asp185 interaction is not (Figure 3). This suggests that the Asn49-Asp185 interaction may contribute to topological maturation of the AQP1 monomer by stabilizing both polar residues at the membrane interface via hydrogen bonding. In contrast, the Lys51-Asp185 hydrogen bonds in between AQP1 monomers should therefore be involved in stabilizing the tetrameric structure. It is interesting in this regard that TM2 and TM5 are the shortest TMs in AQP1 and exhibit extensive helical packing and relatively minor protein-lipid interactions. Thus the rotation of TM3 and orientation of TM2 and TM4 into their proper transmembrane topology may involve early helical packing needed to properly align the TM2 N-terminus and TM5 C-terminus. This model further predicts that failure to form TM2-TM5 hydrogen bonds within monomers should disrupt nascent tertiary structure, perhaps by trapping AQP1 in a 4-spanning topology, and thus disrupt early stages of monomer folding. At the same time, failure to form hydrogen bonds

between Lys51 and Asp185 across the tetrameric interface should weaken quaternary interactions needed to stabilize AQP1 tetramers.

N49M/K51L fails to fold and is retained in a high molecular weight complex. To determine how N49M/K51L affect AQP1 function we tested whether misfolded monomers might be misprocessed. AQP1 proteins were expressed in oocytes and cellular membranes were solubilized and analyzed by sucrose gradient centrifugation (Figure 5). WT protein was recovered primarily in two peaks in fractions 7-11 in nondenaturing detergent (Figure 5A) and as a single peak in fractions 5-6 in SDS (Figure 5B). This pattern is consistent with the majority of AQP1 forming stable tetramers. In contrast, the N49M/K51L mutant was almost entirely recovered in a very large complex that pelleted to the bottom of the gradient (Figure 5C). Thus the N49M/K51L mutation has a major effect on AQP processing that is not solely limited to tetramerization. Rather it would appear that disruption of helical packing between TM2 and TM5 results in an entirely different processing pathway in which the immature form of AQP1 is unable to fold properly and is not freely released into the lipid bilayer. The corrected (N49M/K51L/D185N) protein migrates well out of the gradient in fractions 6-8 (Figure 5D) which is consistent with correction of the primary defect in monomer folding; yet most corrected protein was found in slightly lighter fractions than WT suggesting weaker tetrameric interactions.

Interactions between TM2 and TM5 contribute to AQP1 homo-oligomerization. Finally, we tested whether interactions between Lys51 and Asp185 affect homooligomerization by coimmunoprecipitating WT and myc-tagged AQP1 constructs. For these studies we generated AQP1 constructs that contained a myc epitope inserted at residue Thr120 in the second extracellular loop, which has previously been shown to be well tolerated and yield functional water channels with near wild type efficiency (81). As shown in Figure 6A and B ~50% of WT AQP1 undergoes N-linked glycosylation and deglycosylated (PNGaseF

digested) myc-tagged protein migrates slightly larger than non-tagged protein. When expressed in oocytes and immunoprecipitated under nondenaturing conditions, WT AQP1 was recovered with myc tagged protein (Figure 6B, lane 4 and 7), indicating that these proteins form detergent stable heterooligomers. In contrast, little untagged N49M/K51L mutant protein was recovered with the myc antisera consistent with sequestration of this protein by the ER biosynthetic machinery (Figure 6B lane 5). AQP1 (N49M/K51L/D185N) also exhibited inefficient coimmunoprecipitation under these conditions (Figure 6B, lane 6). These results predict that while the corrected protein retains function, its intermolecular associations are more labile than those of WT protein as manifest by both its smaller size on glycerol gradients as well as its failure to coimmunoprecipitate. The functional K51L mutant also should exhibit normal hydrogen bonding within the monomer (N49-D185), but lack the intermolecular K51-D185 bond involved in tetramer stability. Consistent with this, coimmunoprecipitation studies revealed minimal detergent stable association (Figure 6B, lane 8). Because corrected and K51L mutant proteins exhibit WT levels of water channel activity, we conclude that at least in oocytes intermolecular TM2-TM5 interactions are not required for later aspects of intracellular processing, trafficking and membrane insertion.

5. Discussion

Polytopic protein biogenesis is a multistep process that includes orientation and integration of TM segments (topogenesis), helical packing and monomer folding, and tetramerization. In this study we have identified a unique primary structural feature in AQPs that can impact all three stages of folding via an elaborate tripartite arrangement of polar residues at the membrane interface. We show that two residues from TM2 (Asn49 and Lys51) specifically interact with an aspartate in TM5 (Asp185) to form a symmetric tetrahedron of hydrogen bonds both within the AQP1 monomer and across the tetrameric interface. Based on functional and folding data and crystal structure predictions, we propose that the “cis”

interaction between Asn49 and Asp185 within the monomer serves a primary role in the stability and folding of the monomer. In contrast the “trans” interaction between Lys51 and Asp185 between adjacent monomers stabilizes the aquaporin tetramer. Importantly, both Asn49 and Lys51 contribute to TM2 topogenic activity and thus play a key role in cotranslational translocation events. As a result TM2 initially translocates into the ER lumen and must be positioned within the membrane via an internal reorientation of TMs 2-4 during and/or after completion of synthesis. Until now the forces that drive AQP1 topological maturation have remained largely speculative, but our data suggest that they involve intramolecular hydrogen bonds and possibly neutralization of polar residues at the membrane interface in addition to helical packing within the bilayer.

Most transmembrane segments in polytopic proteins are comprised of stretches of approximately 20 contiguous nonpolar amino acids (73). Polar residues, Asn49 and K51, thus decrease the length and hydrophobicity of AQP1 TM2 as compared to AQP4. In addition, there is also a strong positional preference for amino acids within the TM segment. For example, both Lysine and Asparagine are better tolerated near the membrane interface than at the center of a TM segment and would be predicted to be ~ 4 kcal/mol less favorable than AQP4's Met and Leu in the N49M/K51L mutant (74,75). However, alone either Asn49 or Lys51 can disrupt TM2 stop transfer activity and individually N49M would be predicted to only change the ΔG of the TM segment by $\sim +1$ kcal/mol (74). Thus it is unlikely that this small energy difference would significantly alter TM2 topology. Similarly the distribution of flanking charges does not appear to determine TM2 initial topology because the K51D mutation had no effect. Thus it seems likely that initial TM2 topology is determined by how the TM interacts with the translocon and may involve specific structural aspects of the helix or timing of helix formation.

Association of transmembrane segments within a lipid bilayer is driven primarily by helical packing and the hydrophobic effect and stabilized by hydrogen bonds, van der Waals forces, and occasionally salt-bridges (73). In AQP1 hydrogen bonding between Asn49 and Asp185 plays an important role in early monomer folding. This interaction requires that TM2 span the membrane in order to align with TM5. When this bond fails to form (N49M mutant) the unpaired residue, Asp185 in TM5, results in reduced function and the protein remains trapped in a high molecular weight complex, indicating that folded AQP1 is readily released into the bilayer, whereas the misfolded (immature) protein is not. In some ways this behavior resembles the polytopic membrane protein CFTR (Cystic Fibrosis Transmembrane Conductance Regulator) that also remains associated with ER biosynthetic machinery and is released in an ATP and cytosol sensitive fashion (154). In both cases the final stage of membrane integration appears to be influenced by the folded state of the nascent chain. In contrast to CFTR however, WT AQP1 is rapidly released at the completion of synthesis consistent with rapid and efficient folding. Based on the crystal structure we propose that the N49M/K51L mutation disrupts folding and maturation of the aquaporin channel and that this explains its lack of membrane release. Given that TM2 initially resides in the ER lumen the N49-Asp185 interaction will occur only after topological reorientation is complete and both TM2 and TM5 span the membrane. This hydrogen bond would contribute 4-5 kcal/mole (73,241) stabilizing energy to the monomer. In addition the unpaired Asp185 residue could further destabilize TM5 in the absence of Asn49. Combined with the effect of the Asn49-Asp185 interactions likely contributes significantly to the reorientation needed for AQP1 maturation. An interesting future question is whether this reorientation occurs freely in the bilayer or within the environment of the translocon as has been proposed previously (89).

Despite numerous structural and functional studies demonstrating that AQPs traffic and functions as a homotetramers (98-101) little is known regarding the factors that determine

tetrameric specificity and stability. In general, polytopic membrane proteins use a variety of strategies to promote oligomerization including N-terminal recognition domains (Kv (voltage gated potassium channels)), C-terminal CLZ domains (CNG channels) and GXXXG motifs within TM segments (GPCRs) (104,105). Kv channels contain 6-TMs and utilize both cytosolic T1 (tetramerization) domains as well as hydrophobic interactions between TMs to tetramerize (105,108). The Deutsch group demonstrated that T1 domains could interact even before the nascent chain is released from the ribosome (106) and that oligomerization and tertiary folding are tightly coupled events (107). Unlike the Kv channels, AQPs do not have obvious tetramerization domains and thus likely associate primarily through helical packing in the membrane. Evidence suggests that nonfunctional AQP channels do not tetramerize and that proper folding of the water conducting pore is necessary for oligomerization. Our data further indicate that for the N49M/K51L mutant, defective in folding, results in sequestration (of monomers) by the ER biosynthetic machinery (or other large complex) and that this may explain why tetramerization does not occur. Mathai and Agre showed that two residues important for water conduction, one in Loop B, A73, and one in Loop E, C189, are also critical for AQP1 oligomerization. Mutating either residue (A73M or C189M) prevented oligomerization and blocked function possibly by disrupting intracellular trafficking (103). Similarly, a mutation in Loop E of AQPcic (S205D) blocked tetramerization as well. Other groups have also shown a strong correlation between loss of water channel function and inability to oligomerize (95). Importantly, these mutations are in Loop B and E, which fold into the interior of the protein and thus are not involved in direct contacts, but rather affect the overall protein architecture. In contrast the trio of residues investigated in this study reside at the monomer-monomer interface and are thus the first residues shown to directly contribute to tetramer stability.

While early studies predicted that GlpF functioned as a monomer (109,110) cross-linking and cryoelectron microscopy confirmed the tetrameric structure. However GlpF tetramers

are sensitive to nonionic detergents (111,112). It has further been suggested that tetrameric stability may be influenced by transport selectivity (i.e. water vs. glycerol) (113), because mutations Y222P/W223L not only convert AQP1 to a glycerol conducting channel, but also alter its migration on a sucrose gradient to resemble GlpF (95). Interestingly, these weak detergent sensitive intermolecular interactions are similar to those we observe for the “corrected” AQP1 (N49M/K51L/D185N) lacking the Lys51-Asp185 hydrogen bond. Thus, the hydrogen bond between Lys51 and Asp185 across the tetrameric interface stabilizes the AQP1 tetramer even though it is not obligatorily required for water channel function. Given the extensive hydrophobic interactions between AQP1 monomers it seems unlikely that K51L mutation completely disrupts tetramer formation. Rather we believe that the loss of hydrogen bonds ($\sim 4\text{-}5 \text{ kcal/mol}^{-1}$ per bond $\times 4 = 16\text{-}20 \text{ kcal/mol}^{-1}$ for the tetramer) reduces energetics to a level similar to that of GlpF, thus allowing the protein to still traffic as a tetramer, but to disassociate upon TritonX-100 solubilization.

We have identified a novel tetramerization motif for AQP1. Notably, a sequence comparison of other human AQPs revealed potential intermolecular hydrogen bonds at positions that correspond to Lys51 and Asp185 in AQP1 (See Figure 2A and 3A) for other AQPs. For example AQP0 and AQP5 contain a Gln-Tyr pair that could potentially hydrogen bond if presented in the proper architecture. AQP2 has Gln-His and AQP11 has His-Ser both potential hydrogen bonding partners. In contrast, GlpF, which has already been shown to have weak monomer-monomer associations, has a Glu-Pro pair that would not be expected to strongly interact. It remains to be seen how these hydrogen bond pairs will affect the tetrameric stability of other AQP family members.

Intriguingly, differential tetramer stability within the AQP family may also provide a mechanism to discourage hetero-oligomerization. Given the high degree of sequence and structural homology between family members, oligomerization may well include specialized

residues that alter the strength of intermolecular interactions. Data from the current study supports this hypothesis as substituting residues from AQP4 into AQP1 dramatically reduces the strength of AQP1 monomer association; moreover, Asp185 is unique to AQP1 and thus may play a role in preventing heterotetramerization of other family members by preferentially stabilizing AQP1 homotetramers. For example GlpF has a negatively charged residue in TM2 and this would not form a stable intermolecular association with the negatively charged Asp185 in AQP1 TM5, thus AQP1 and GlpF would not be expected to heterooligomerize. Given AQP1's biogenesis mechanism, and the unique role of Asp185 we cannot rule out the possibility that intermolecular associations between Lys51-Asp185 may also help drive AQP1 reorientation. Thus if reorientation takes place after release from the translocon this would provide a means to mechanistically couple reorientation of TMs2-4 and tetramerization. Clearly, given their diversity of expression, functional importance and medical relevance of AQPs additional studies are needed to determine any general relationship between AQP topogenesis, folding and oligomerization.

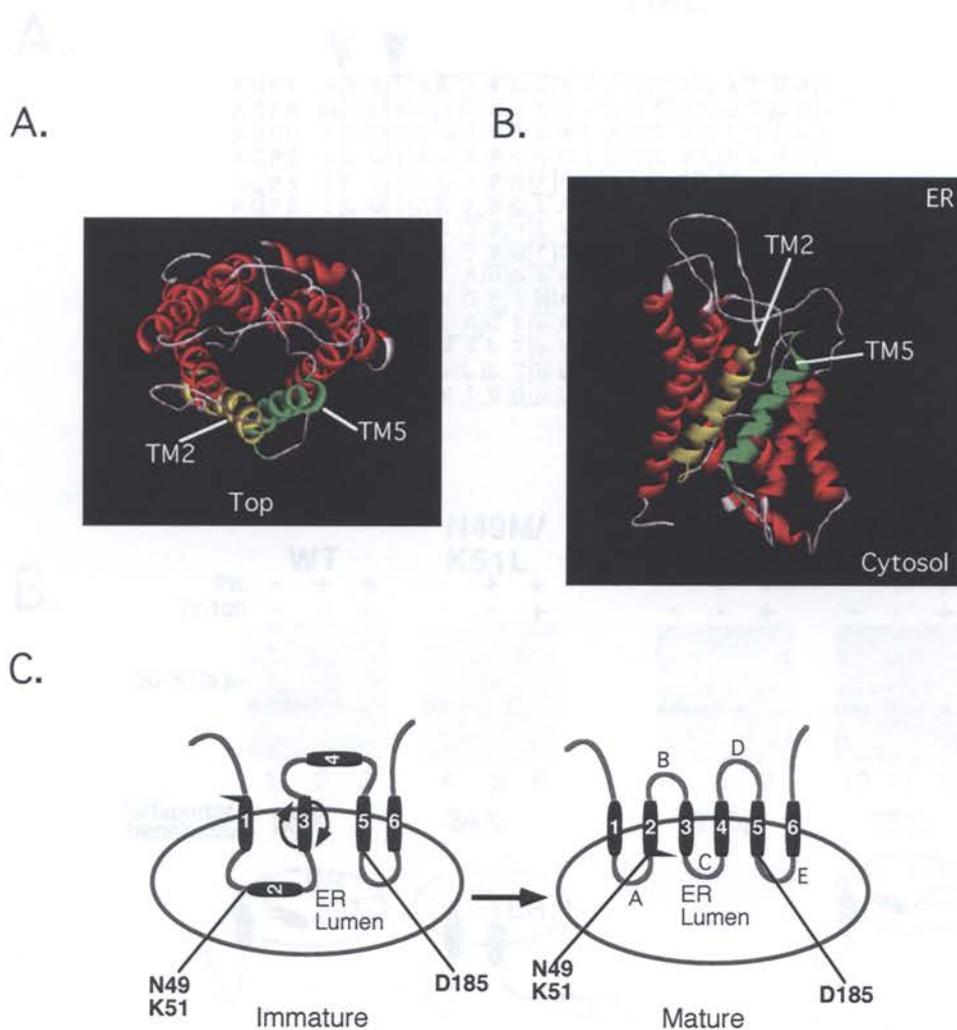


Figure 3-1: AQP1 biogenesis model and crystal structure.

A. and **B.** Top and side view ribbon diagram of AQP1 channel displaying adjacent location of TM2 (yellow) and TM5 (green). Model derived from Sui et al. crystal structure (5). **C.** Model of AQP1 topological maturation from 4 to 6 spanning mature channel and location of polar TM2 residues (Asn49 and Lys51) and charged TM5 residue (Asp185). Extra- and Intracellular loops are as labeled A-E.

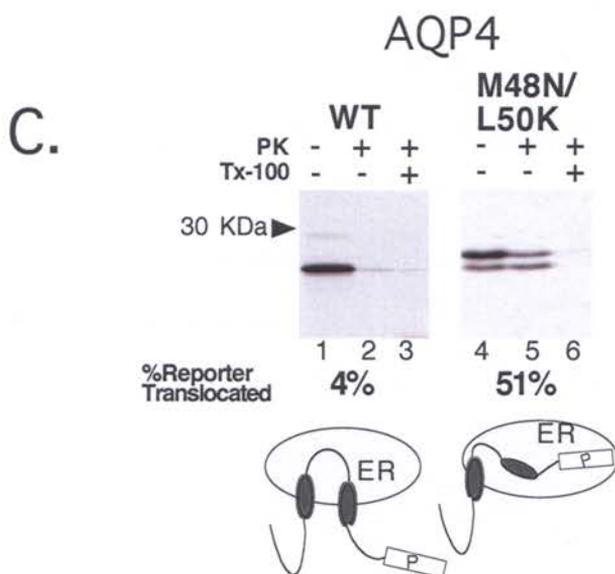
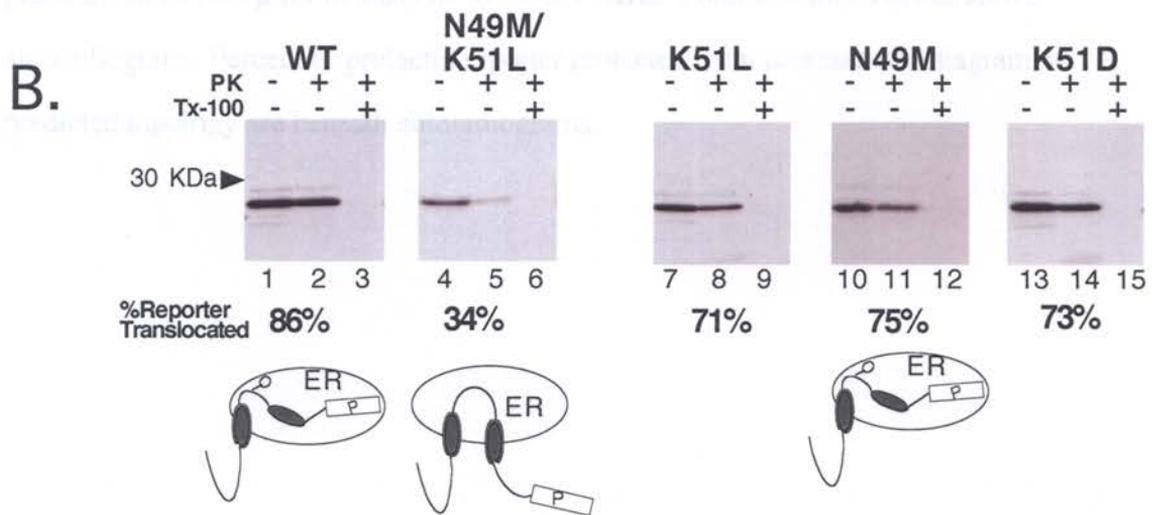
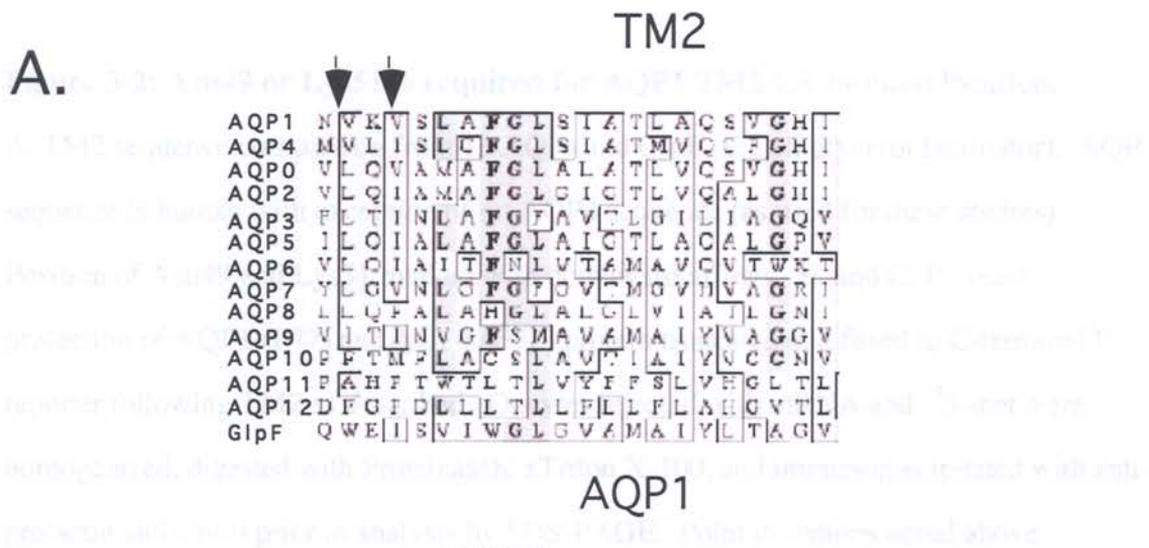


Figure 3-2: Asn49 or Lys51 is required for AQP1 TM2 ER luminal location.

A. TM2 sequence comparison from 13 AQPs and GlpF (E.Coli glycerol facilitator). AQP sequence is human with exception of rat AQP4 sequence (as used for these studies).

Position of Asn49 and Lys51 marked with downward arrows. B. and C. Protease protection of AQP1 (P77) and AQP4 (G72) protein truncated and fused to C-terminal P-reporter following TM2 as described. Oocytes injected with mRNA and ³⁵S-met were homogenized, digested with ProteinaseK ±Triton X-100, and immunoprecipitated with anti-prolactin antiserum prior to analysis by SDS-PAGE. Point mutations noted above autoradiogram. Percent of prolactin reporter protected from protease and diagram of predicted topology are beneath autoradiograms.

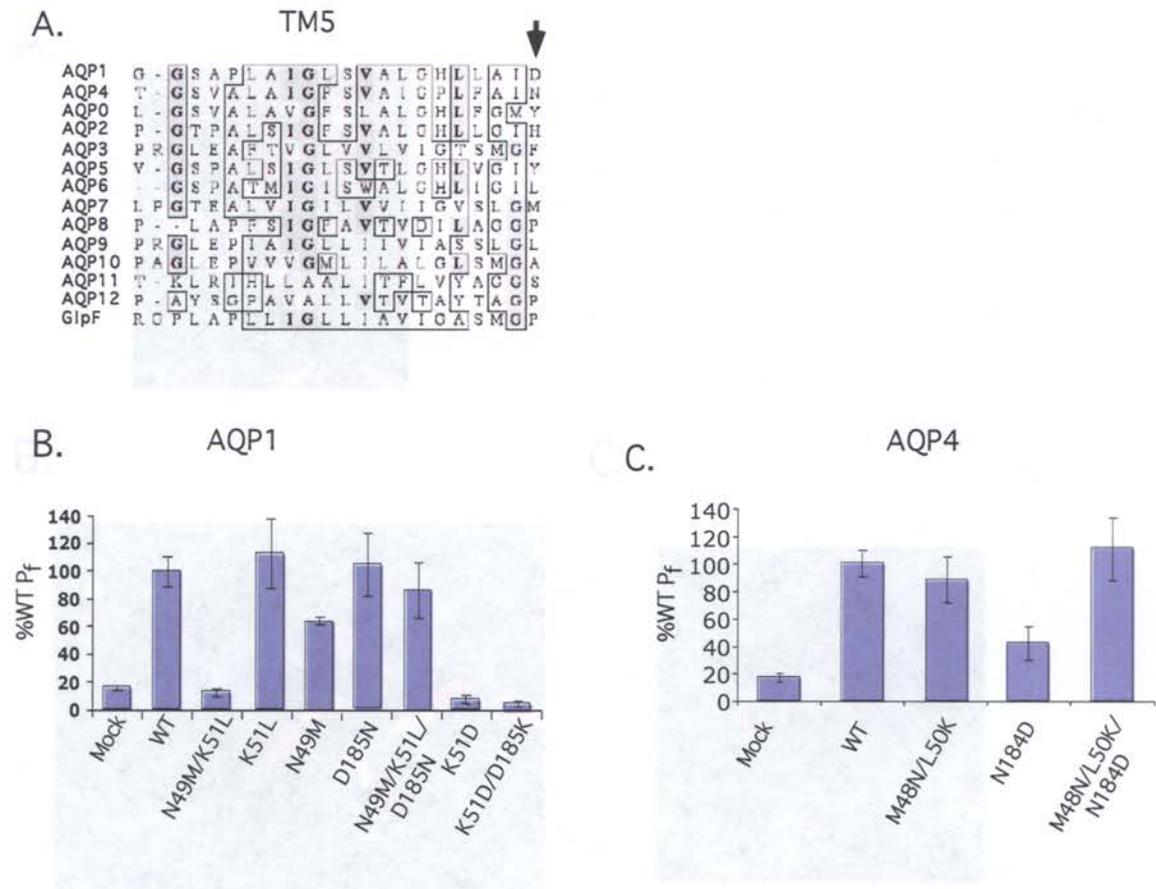


Figure 3-3: Interactions between TM2 and TM5 are important for AQP1 Function.

A. TM5 sequence comparison from 13 AQPs and GlpF (E.Coli glycerol facilitator). AQP sequence is human with exception of rat AQP4 sequence (as used for these studies). Position Asp185 marked with downward arrow. **B.** and **C.** Water permeability of oocytes 48 hr post injection with mRNA. Assayed as described in Methods. Point mutations are as indicated under bar graph. Shown are results from a representative experiment n=5 oocytes in each group \pm SEM.

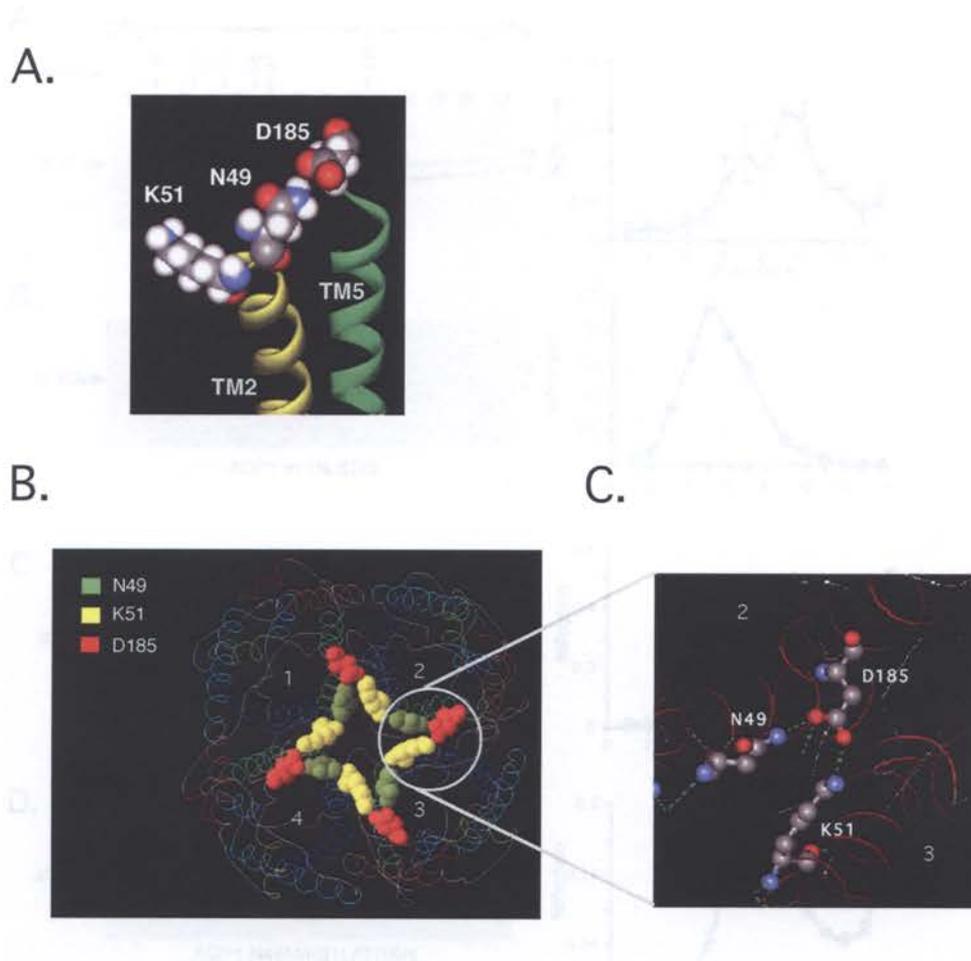


Figure 3-4: Asn49, Lys51 and Asp185 form a complex triad of intra- and intermolecular Hydrogen Bonds. **A.** Model of TM2-TM5 interactions with Asn49, Lys51 and Asp185 represented as space-filling molecules produced using published crystal structure (5). **B.** Ribbon diagram of AQP1 tetramer model, four interacting monomers labeled 1-4. Location of 3 residues noted. **C.** Enlarged detail of hydrogen bonds between two monomers. Green dotted lines represent hydrogen bonding between Asn49 and Asp185 (within monomer) as well as Lys51 and Asp185 (between monomers).

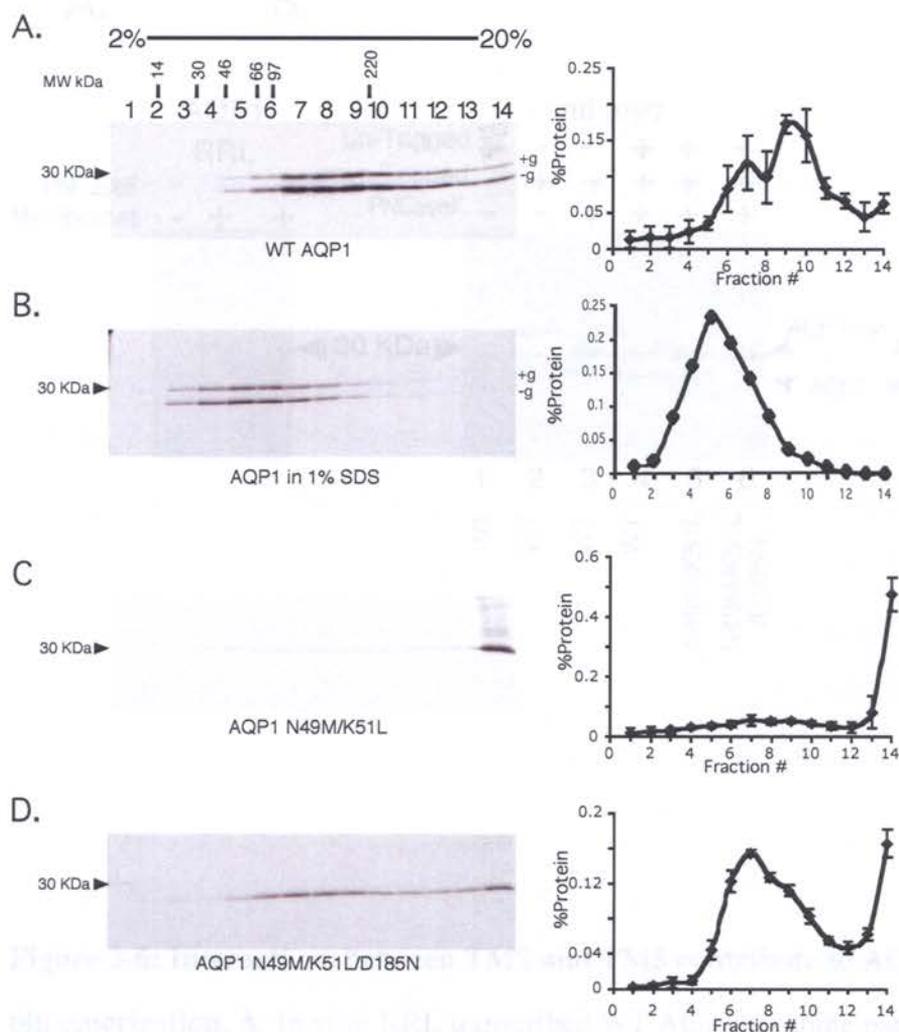


Figure 3-5: N49M/K51L fails to fold and is retained in a high molecular weight complex. Oocytes were injected with mRNA and ^{35}S -methionine incubated for 5 hr at 18°C , homogenized and microsomes were isolated and solubilized in 1% Tx-100 (**A,C,D**) or 1% SDS (**B**). Solubilized microsomes were loaded on a 2-20% sucrose gradient and subjected to centrifugation at $100,000 \times g$ for 14 h. 14 equal fraction were taken and subjected to immunoprecipitation with anti-AQP1 antisera preceding analysis by SDS-PAGE. Glycosylated (+g) and unglycosylated (-g) protein is indicated. Data were quantitated by phosphorimaging A, C and D average of 3 experiments \pm SEM. B. Data from one representative experiment.

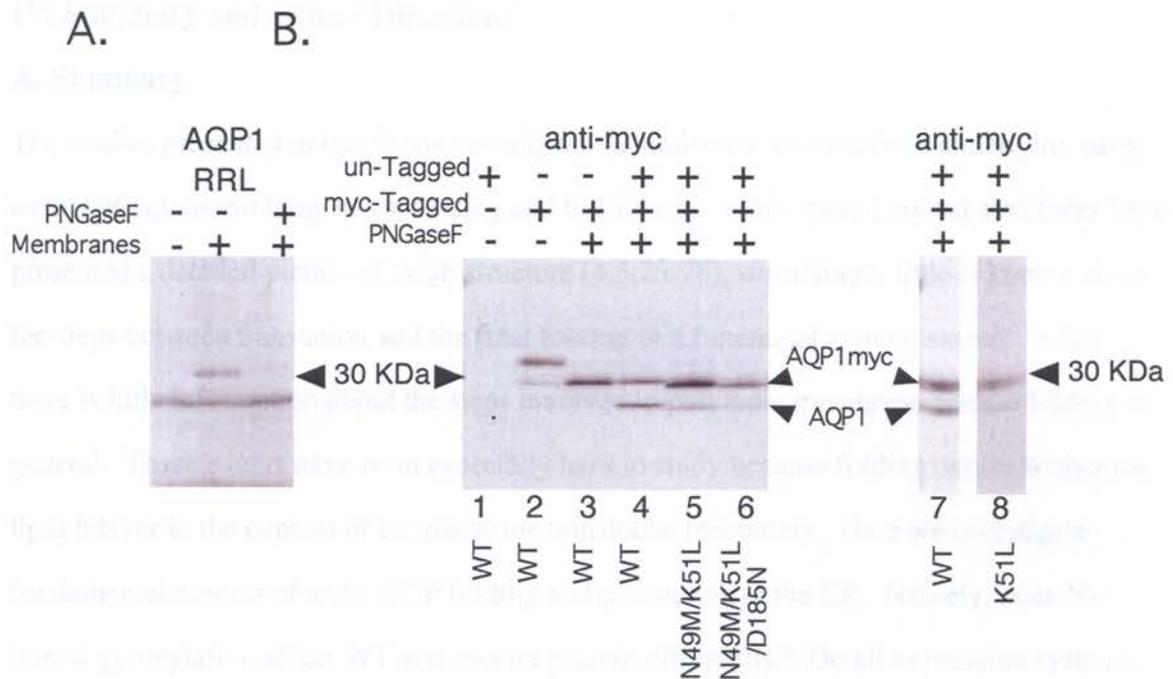


Figure 3-6: Interactions between TM2 and TM5 contribute to AQP1 homo-oligomerization. **A.** In vitro RRL transcribed WT AQP1 \pm canine pancreas membranes and \pm PNGase F to remove N-linked glycans. **B.** Oocytes injected with ^{35}S -met and mRNA, homogenized, and microsomes were isolated and solubilized in 1% Tx-100 immunoprecipitated with Myc-9E10 antibody. Oocytes were injected with untagged AQP1 (lane1), AQP1mycT120 (lane 2), or both myc-tagged and untagged protein together (lanes4-8). mRNA injected was either WT (lanes 4 and 7), AQP1 N49M/K51L (lane5), AQP1 N49M/K51L/D185N (lane 6) or AQP1 K51L (lane 8). After immunoprecipitation protein was treated with PNGaseF (lanes 3-8) prior to analysis by SDS-PAGE. Immunoprecipitated myc-tagged and untagged protein mobility is as indicated.

IV. Summary and Future Directions

A. Summary

The studies presented in this thesis investigate the molecular mechanisms underlying early events of aquaporin biogenesis, folding and trafficking. While recent crystal structures have presented a detailed picture of AQP structure (4,5,25-28), surprisingly little is known about the steps between translation and the final folding of a functional water channel. In fact there is little information about the steps involved in polytopic membrane protein folding in general. These events have been especially hard to study because folding occurs within the lipid bilayer in the context of the ribosome translocon machinery. Here we investigate fundamental aspects of early AQP folding and processing at the ER. Namely, does N-linked glycosylation affect WT and mutant protein differently? Do all expression systems recognize topogenic information similarly? Lastly, what is the relationship between early inter- and intramolecular folding and topogenesis?

I first examined the role of N-linked glycosylation on AQP2 trafficking mutants. Because AQP2 is inefficiently glycosylated (~30%) it provides us with an opportunity to directly compare glycosylated and unglycosylated protein within the same cell. These studies confirm that NDI causing trafficking mutants T126M, A147T, C181W and R187C all have reduced half-lives in oocytes. Interestingly, when the glycosylated and nonglycosylated isoforms were examined separately we found that the presence of N-linked sugars markedly stabilized mutant proteins, whereas WT protein stability was unaffected. N-linked glycosylation has long been known to facilitate protein folding, although it can also target proteins for degradation (201,202). In our hands, mutations, with the possible exception of R187C, did not affect the extent to which the protein was glycosylated. However, N-linked glycosylation had a general effect on the stability of trafficking mutants. As the four trafficking mutants examined are found in very different regions of AQP2, TM segments (T126M, A147T), cytosolic (T126M), extracellular (A147T) and within the pore itself

(C181W and R187C) they would be expected to disrupt protein folding in different ways. Yet N-linked sugars stabilize all of them similarly. These studies raise the possibility that manipulating the attachment of N-linked glycans may provide a strategy for improving stability of mutant protein in human disease.

Topological studies of multi-spanning membrane proteins commonly use sequentially truncated proteins fused to a C-terminal translocation reporter. Using an assay for reporter location (e.g. protease protection, glycosylation) allows one to define TM segment orientation and key biogenesis events. Because truncated proteins represent an intermediate in the folding pathway, they may transiently occupy multiple folding states (i.e. more than one topological isoform). We have demonstrated that two such AQP1 biogenesis intermediates truncated after TM3 have markedly different stabilities. The predominant isoform, a 2-spanning intermediate with TM2 in the ER lumen, was degraded rapidly whereas the 3-spanning intermediate with TM2 spanning the membrane was not. This difference in stability was due to the truncated nature of the construct and not a general property of AQP1, as full-length protein was quite stable. Importantly, these studies further confirmed that the translocation machinery in oocyte and mammalian cell systems handle topological information in much the same way and that the location of a reporter is dictated by the information encoded within the polypeptide itself and not an artifact of the expression system. These results also highlight the importance of considering stability when carrying out topological studies with truncated proteins.

AQP1's unique biogenesis has been well established, but its molecular mechanism and role in the mature water channel are poorly characterized. The current study investigates how AQP1's biogenesis mechanism affects both tertiary and quaternary protein folding. Polar residues near TM2 (Asn49 and Lys51), that allow TM2 to slip into the ER lumen hydrogen bond to Asp185 in TM5. Specifically, hydrogen bonding between Asp185 and Asn49 is

important for stable monomer folding whereas intermolecular hydrogen bonding between Lys51 and Asp185 is more important for stabilizing the AQP1 tetramer. This trio of residues contributes to monomer folding as well as oligomer stabilization. The conclusion of these experiments is that AQP1's unique biogenesis mechanism is required to compensate for the charged Asp185 in TM5. Asp185 is unique to AQP1 within the AQP family and it is unclear what advantage to water channel function or regulation it imparts. In conclusion these studies give us new insight into the molecular mechanism of AQP1 biogenesis and highlight how early biogenesis, tertiary and quaternary folding events can be intimately linked by specific structural features.

B. Future Directions

1) Is the degradation of glycosylated AQP2 mutants simply delayed or does the addition of N-linked sugars promote proper folding of the trafficking mutants?

Our work demonstrates that AQP2 trafficking mutants are stabilized by N-linked glycosylation at the level of the ER. It is unclear whether the stabilization we observe allows the glycosylated protein to become a functional water channel and traffic to the plasma membrane or just delays its recognition by the ER quality control machinery. Because only a small percentage of the overall AQP2 population is glycosylated (~20%) oocyte functional assays may not provide a sensitive enough measure of trafficking.

Immunocytochemistry in mammalian cells may provide a better means to investigate this question.

2) Are chaperone proteins such as calnexin and calreticulin involved in stabilizing the glycosylated mutant AQP2 proteins?

We showed that N-linked glycosylation stabilizes AQP2 trafficking mutants but the basic mechanism behind this stabilization remains unexplored. The ER lectins CNX/CRT bind and have been shown to stabilize glycosylated proteins (201). In contrast MHC class I

molecule was actually stabilized in the absence of N-linked glycosylation (202). It would therefore be of interest to determine if these chaperones affected the mutant AQP2 stability either by using biochemical assays in oocytes, or pharmacological studies in mammalian cells.

3) Why do different AQP1 topological isoforms exhibit markedly different stabilities?

We have shown that different AQP1 topological isoforms have differential stabilities. There are a few potential explanations. First, hydrophobic patches can act as signals for degradation (200). Therefore the two spanning isoform, in which the hydrophobic TM2 is transported into the ER lumen, may be more easily recognized. Another possibility is reporter location. Both reporters used in our studies (prolactin and β -subunit of H/K ATPase) are normally expressed in the ER lumen, therefore cytosolic localization may target the reporter for degradation. Understanding how two populations of the same protein are first directed into different topologies by the translocation machinery and second differentially recognized by the ER quality control system will be important for understanding the early folding events of multi-spanning proteins in general.

4) What cellular components is the AQP1 N49M/K51L mutant associated with?

Converting the two polar residues Asn49 and Lys51 at the boundary of AQP1 TM2 to the corresponding AQP4 residues (N49M and K51L) allows TM2 to stop in the ER membrane, but destroys water channel function. Further analysis showed that in contrast to WT AQP1 protein N49M/K51L was associated with a large molecular weight complex that was soluble in 1% Tx-100 presumably because it fails to properly fold. Candidate protein complexes include the ribosome translocon or a proteasomal degradation complex, but this is yet to be determined. Cross-linking and coimmunoprecipitations studies could be used to identify members of this protein complex.

5) Is the intermolecular interaction between Lys51 and Asp185 important for the final folding of the monomer as well as tetramer stability?

We have shown that the intermolecular association between Lys51 and Asp185 is important for stabilizing the tetramer. Work by other labs has suggested that properly folded water channel was a prerequisite for tetramerization, however it is possible that instead tetramerization is a prerequisite for a functional channel. While there is substantial evidence that each AQP1 monomer contains a functional water channel, no studies have shown AQP monomer functioning outside of the tetramer. It is therefore possible that tetramerization and possibly the intermolecular interaction between Lys51 and Asp185 are necessary to stabilize a functional AQP monomer. Further mutagenesis studies, expression of mutant concatamers, and functional assays in a system that doesn't require cellular trafficking may prove useful in answering these questions.

6) Does the charged residue in AQP1 TM5 (Asp185) prevent heterotetramerization with other MIP family members?

Intramolecular interactions between Asp185 and Lys51 strengthen AQP1 monomer association. Asp185 is unique to AQP1 within the AQP family. Because it is involved in intermolecular interactions, this unique residue may help block heterotetramerization in cells where multiple AQPs are expressed. Heterotetramerization has not been observed for any AQPs, yet AQP1 and AQP3 are coexpressed in erythrocyte membranes and AQP2, AQP3 and AQP4 are coexpressed in the principal cells of the kidney collecting duct. Therefore it is of interest to determine what mechanisms these very homologous proteins use for specificity in oligomerization and what role Asp185 plays. Sequence comparison of AQPs and GlpF revealed that other AQPs have polar or charged residues at this position. For example, AQP2 has a His at this position and a Gln at the corresponding position in TM2. If the position of these residues is similar to that in AQP1, then they could block the formation of an AQP1-AQP2 heterotetramer as the His-Lys interaction would not be as

favorable as the Asp-Lys interaction in AQP1. Coexpression and mutagenesis studies will allow us to determine to what extent AQPs use residues at these positions to block heterotetramerization.

7) Where does the reorientation of AQP1 TM3 occur and what facilitates it?

AQP1's unique biogenesis mechanism has been validated in multiple expression systems. Our recent work has given some insight into why this process occurs and identified some requirements at the protein sequence level. However, we have little if any information as to where within the ER membrane this process occurs. Does this process occur while the protein is still associated with the translocon? If so, where in the translocon environment would this occur? Recent models suggest that the translocation pore is only wide enough to accommodate one helix at a time ($\sim 10 \text{ \AA}$) (62,86). Would AQP1 then associate with multiple translocon complexes at once, or is some other mechanism/machinery involved? Alternatively, this maturation may occur after the protein exits the translocon into the lipid bilayer. Insertion of cross-linking probes into the AQP1 sequence will help us track the movements of TM2 in relation to other proteins and lipids.

V. Appendix

A. Data not Shown

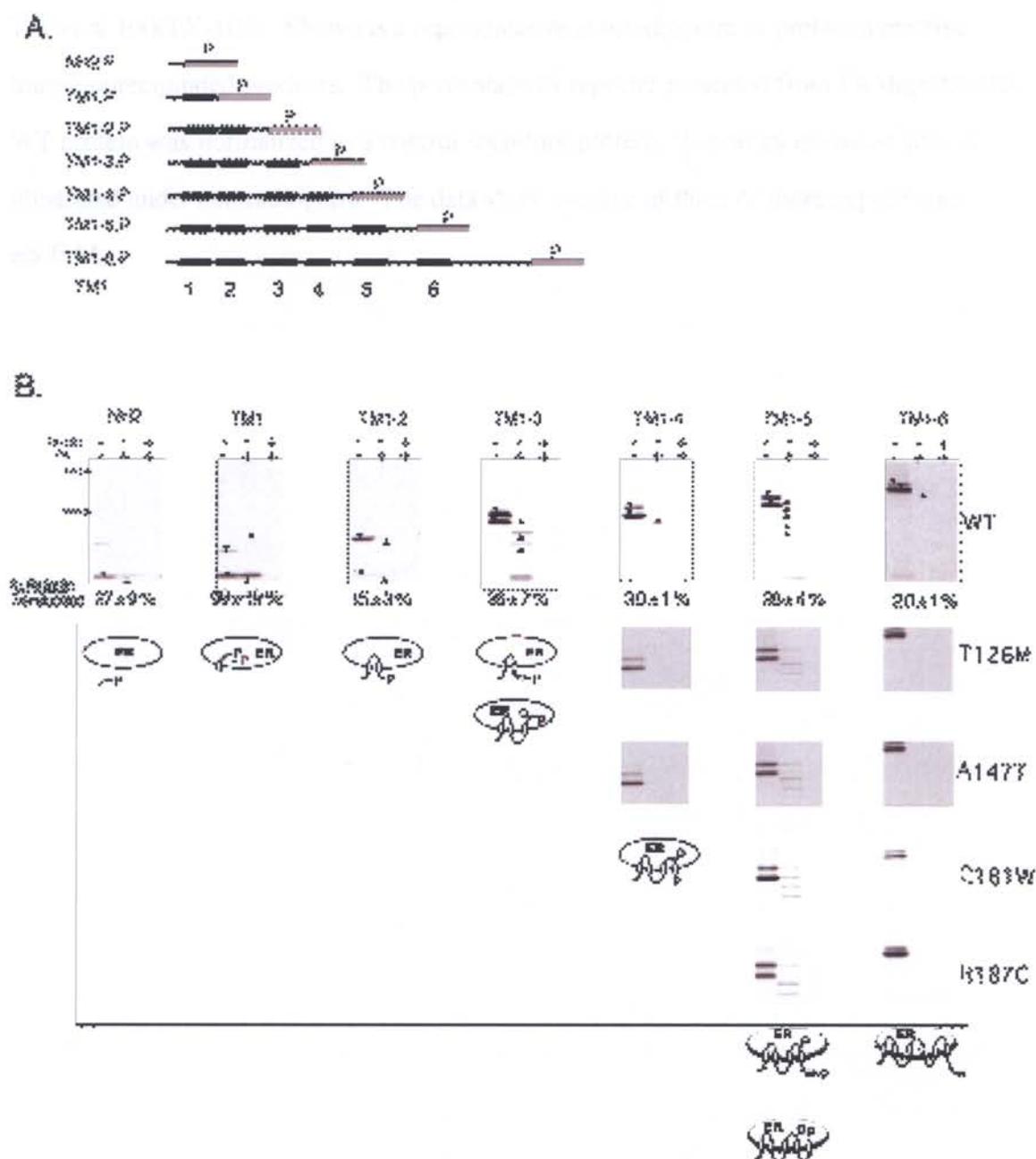
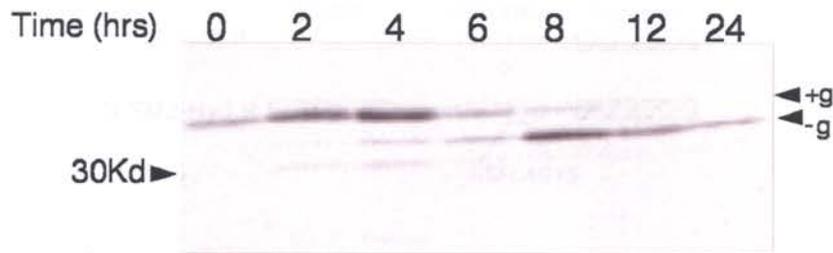


Figure A-1: Trafficking Mutants do not affect early topogenesis of AQP2. A. WT

and mutant AQP2 cDNA was truncated before TM1 and after TM1, TM2, TM3, TM4, TM5 and TM6 as diagrammed and fused to a prolactin-derived COOH terminus reporter (P). B. Oocytes injected with mRNA and [³⁵S]methionine were incubated 3 h, reporter topology was determined by proteinase K (PK) digestion in the presence or absence of Triton X-100(TX-100). Shown is a representative autoradiogram of prolactin reactive immunoprecipitated products. The percentage of reporter protected from PK digestion for WT protein was normalized to a control secretory protein. Topology of fusion sites is illustrated under autoradiogram. The data show average of three or more experiments \pm S.E.M.

A.



B.

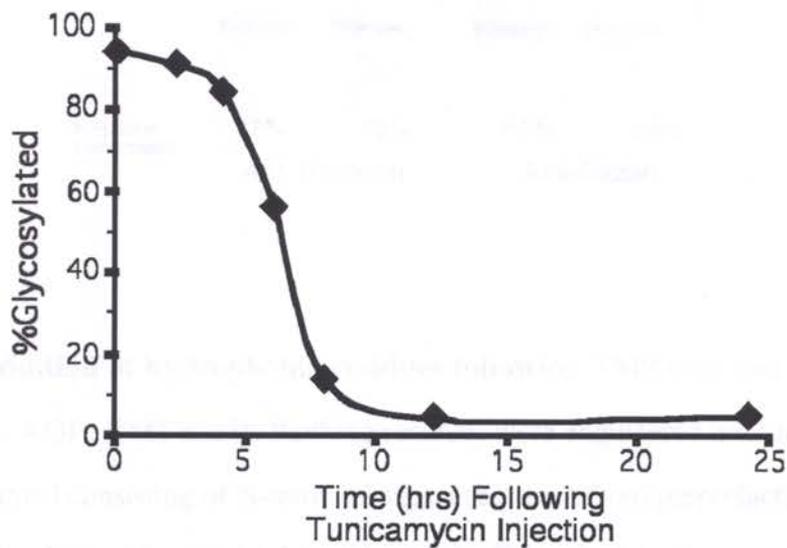


Figure A-2: Tunicamycin blocks N-linked glycosylation of proteins in oocytes 8 hrs

after injection. A. *Xenopus* oocytes were injected with 50 nl of 400 μ g/ml Tunicamycin in 8% DMSO and incubated for the time indicated (0,2,4,6,8,12,or 24 hr). Oocytes were then injected a second time with Sp+1gGSTpT (a chimeric protein with a well-used glycosylation site that has been previously characterized (166,184)) mRNA and [³⁵S]methionine and incubated for 2 h at 18° C. B. Results from A were quantitated using Bio-Rad personal Molecular PhosphorImager Fx (Kodak screens, Quantity-1 software) and the percentage of glycosylated protein at each time point was plotted.

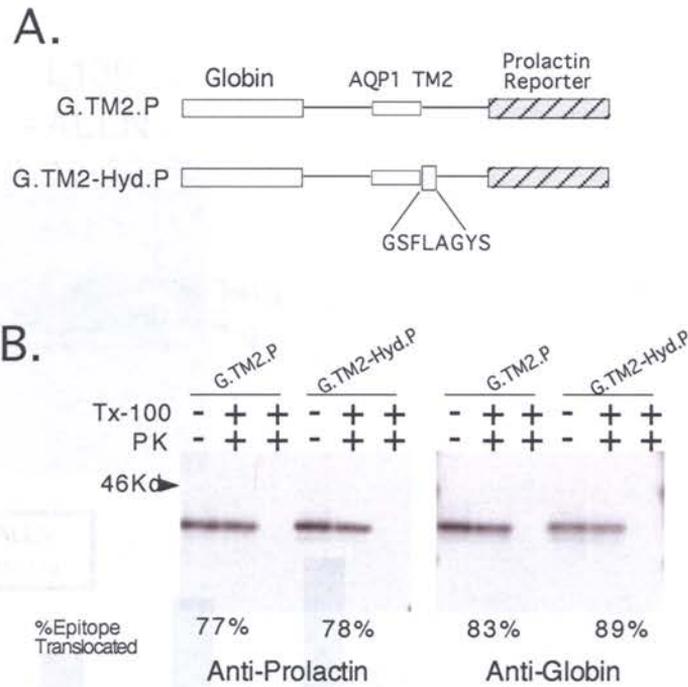


Figure A-3: Addition of hydrophobic residues following TM2 does not affect

topogenesis. A. AQP1 TM1 and its flanking residues were engineered into the chimeric construct diagramed consisting of N-terminal signal sequence from preprolactin (S), a passenger domain derived from α -globin (G), and the C-terminal prolactin reporter (P). B.

Autoradiograms show pretease protection of *Xenopus* oocyte products

immunoprecipitated with anti-prolactin or anti-globin. %Epitope Translocated is average of two experiments.

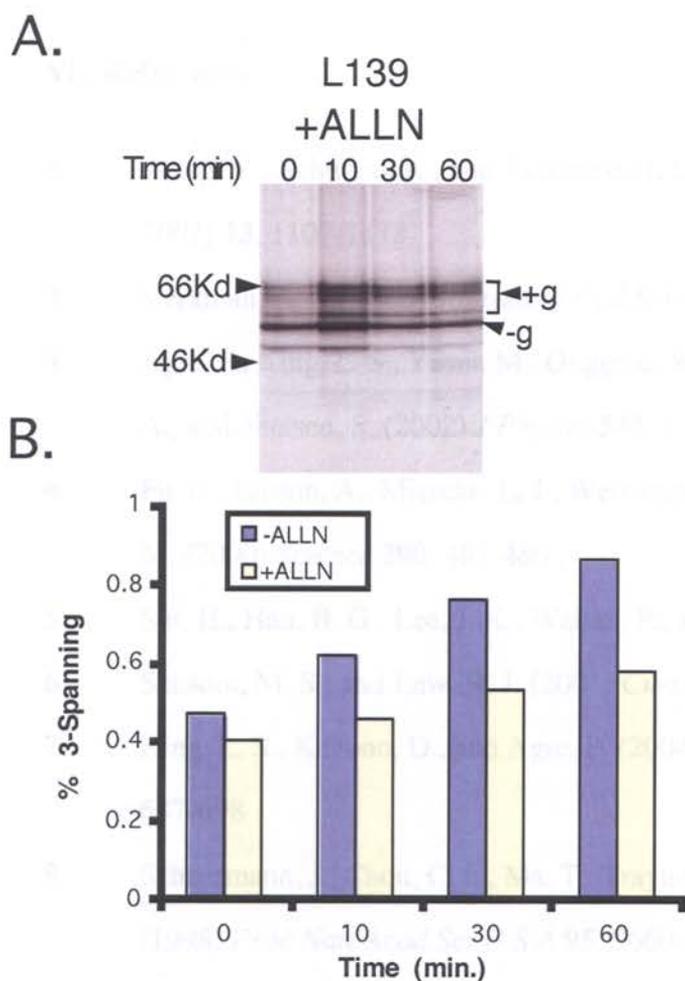


Figure A-4: ALLN blocks selective degradation of nonglycosylated isoforms. A. EGFP- β subunit chimera truncated at L139 (see Figure 4 Manuscript 2) was expressed in HEK-293 cells. Pulse-chase labeling was performed in the presence of 20 μ M ALLN and immunoprecipitated with anti-EGFP antisera. Glycosylated (+g) and nonglycosylated (-g) are indicated. B. Percentage of protein in the three-spanning (glycosylated) topology was quantitated for experiments carried out with or without ALLN.

VI. References

1. Wang, Y., Schulten, K., and Tajkhorshid, E. (2005) *Structure (Cambridge, Mass. : 2001)* **13**, 1107-1118
2. Verkman, A. S. (2005) *Journal of Cell Science* **118**, 3225-3232
3. Agre, P., King, L. S., Yasui, M., Guggino, W. B., Ottersen, O. P., Fujiyoshi, Y., Engel, A., and Nielsen, S. (2002) *J Physiol* **542**, 3-16
4. Fu, D., Libson, A., Miercke, L. J., Weitzman, C., Nollert, P., Krucinski, J., and Stroud, R. M. (2000) *Science* **290**, 481-486
5. Sui, H., Han, B. G., Lee, J. K., Walian, P., and Jap, B. K. (2001) *Nature* **414**, 872-878
6. Sansom, M. S., and Law, R. J. (2001) *Current Biology* **11**, R71-73
7. King, L. S., Kozono, D., and Agre, P. (2004) *Nature reviews. Molecular cell biology*. **5**, 687-698
8. Schnermann, J., Chou, C. L., Ma, T., Traynor, T., Knepper, M. A., and Verkman, A. S. (1998) *Proc Natl Acad Sci U S A* **95**, 9660-9664
9. Ma, T., Yang, B., Gillespie, A., Carlson, E. J., Epstein, C. J., and Verkman, A. S. (1998) *Journal of Biological Chemistry* **273**, 4296-4299
10. King, L. S., Choi, M., Fernandez, P. C., Cartron, J. P., and Agre, P. (2001) *New England Journal of Medicine* **345**, 175-179
11. Noda, Y., and Sasaki, S. (2005) *Biology of the Cell* **97**, 885-892
12. Fushimi, K., Sasaki, S., and Marumo, F. (1997) *Journal of Biological Chemistry* **272**, 14800-14804
13. Katsura, T., Gustafson, C. E., Ausiello, D. A., and Brown, D. (1997) *American Journal of Physiology* **272**, F817-822
14. Terris, J., Ecelbarger, C. A., Nielsen, S., and Knepper, M. A. (1996) *American Journal of Physiology* **271**, F414-422
15. Morello, J. P., and Bichet, D. G. (2001) *Annu Rev Physiol* **63**, 607-630

16. Deen, P. M., Verdijk, M. A., Knoers, N. V., Wieringa, B., Monnens, L. A., van Os, C. H., and van Oost, B. A. (1994) *Science* **264**, 92-95
17. Manley, G. T., Fujimura, M., Ma, T., Noshita, N., Filiz, F., Bollen, A. W., Chan, P., and Verkman, A. S. (2000) *Nature Medicine* **6**, 159-163
18. Papadopoulos, M. C., Manley, G. T., Krishna, S., and Verkman, A. S. (2004) *FASEB Journal* **18**, 1291-1293
19. Bloch, O., Papadopoulos, M. C., Manley, G. T., and Verkman, A. S. (2005) *Journal of Neurochemistry* **95**, 254-262
20. Thiagarajah, J. R., and Verkman, A. S. (2002) *Journal of Biological Chemistry* **277**, 19139-19144
21. Ma, T., Song, Y., Gillespie, A., Carlson, E. J., Epstein, C. J., and Verkman, A. S. (1999) *The Journal of biological chemistry*. **274**, 20071-20074
22. Krane, C. M., Melvin, J. E., Nguyen, H. V., Richardson, L., Towne, J. E., Doetschman, T., and Menon, A. G. (2001) *Journal of Biological Chemistry* **276**, 23413-23420
23. Hara-Chikuma, M., Sohara, E., Rai, T., Ikawa, M., Okabe, M., Sasaki, S., Uchida, S., and Verkman, A. S. (2005) *The Journal of biological chemistry*. **280**, 15493-15496
24. Ma, T., Hara, M., Sougrat, R., Verbavatz, J. M., and Verkman, A. S. (2002) *The Journal of biological chemistry*. **277**, 17147-17153
25. Murata, K., Mitsuoka, K., Hirai, T., Walz, T., Agre, P., Heymann, J. B., Engel, A., and Fujiyoshi, Y. (2000) *Nature* **407**, 599-605
26. Savage, D. F., Egea, P. F., Robles-Colmenares, Y., O'Connell, J. D., 3rd, and Stroud, R. M. (2003) *Plos Biology* **1**, E72
27. Harries, W. E., Akhavan, D., Miercke, L. J., Khademi, S., and Stroud, R. M. (2004) *Proceedings of the National Academy of Sciences of the United States of America* **101**, 14045-14050
28. Gonen, T., Sliz, P., Kistler, J., Cheng, Y., and Walz, T. (2004) *Nature* **429**, 193-197

29. Tajkhorshid, E., Nollert, P., Jensen, M. O., Miercke, L. J., O'Connell, J., Stroud, R. M., and Schulten, K. (2002) *Science* **296**, 525-530
30. de Groot, B. L., Engel, A., and Grubmuller, H. (2001) *FEBS Letters* **504**, 206-211
31. Jensen, M. O., Park, S., Tajkhorshid, E., and Schulten, K. (2002) *Proceedings of the National Academy of Sciences of the United States of America* **99**, 6731-6736
32. Jensen, M. O., Tajkhorshid, E., and Schulten, K. (2001) *Structure* **9**, 1083-1093
33. Jensen, M. O., Tajkhorshid, E., and Schulten, K. (2003) *Biophysical Journal* **85**, 2884-2899
34. Zhu, F., Tajkhorshid, E., and Schulten, K. (2001) *FEBS Letters* **504**, 212-218
35. Zhu, F., and Schulten, K. (2003) *Biophysical Journal* **85**, 236-244
36. Chakrabarti, N., Roux, B., and Pomáes, R. (2004) *Journal of molecular biology.* **343**, 493-510
37. Ilan, B., Tajkhorshid, E., Schulten, K., and Voth, G. A. (2004) *Proteins* **55**, 223-228
38. de Groot, B. L., Frigato, T., Helms, V., and Grubmuller, H. (2003) *Journal of Molecular Biology* **333**, 279-293
39. Lu, D., Grayson, P., and Schulten, K. (2003) *Biophysical journal.* **85**, 2977-2987
40. Jiang, J., Daniels, B. V., and Fu, D. (2006) *Journal of Biological Chemistry* **281**, 454-460
41. Teornroth-Horsefield, S., Wang, Y., Hedfalk, K., Johanson, U., Karlsson, M., Tajkhorshid, E., Neutze, R., and Kjellbom, P. (2006) *Nature.* **439**, 688-694
42. Yasui, M., Kwon, T. H., Knepper, M. A., Nielsen, S., and Agre, P. (1999) *Proceedings of the National Academy of Sciences of the United States of America* **96**, 5808-5813
43. Yasui, M., Hazama, A., Kwon, T. H., Nielsen, S., Guggino, W. B., and Agre, P. (1999) *Nature* **402**, 184-187
44. Liu, K., Kozono, D., Kato, Y., Agre, P., Hazama, A., and Yasui, M. (2005) *Proceedings of the National Academy of Sciences of the United States of America.* **102**, 2192-2197

45. Meinild, A. K., Klaerke, D. A., and Zeuthen, T. (1998) *Journal of Biological Chemistry* **273**, 32446-32451
46. Abrami, L., Tacnet, F., and Ripoche, P. (1995) *Pflugers Archiv - European Journal of Physiology* **430**, 447-458
47. Yool, A. J., and Weinstein, A. M. (2002) *News in physiological sciences : an international journal of physiology produced jointly by the International Union of Physiological Sciences and the American Physiological Society.* **17**, 68-72
48. Saparov, S. M., Kozono, D., Rothe, U., Agre, P., and Pohl, P. (2001) *Journal of Biological Chemistry* **276**, 31515-31520
49. Agre, P., Lee, M. D., Devidas, S., and Guggino, W. B. (1997) *Science* **275**, 1490; author reply 1492
50. Nielsen, S., Frøkjaer, J., Marples, D., Kwon, T. H., Agre, P., and Knepper, M. A. (2002) *Physiological reviews.* **82**, 205-244
51. Nakhoul, N. L., Davis, B. A., Romero, M. F., and Boron, W. F. (1998) *American Journal of Physiology* **274**, C543-548
52. Cooper, G. J., and Boron, W. F. (1998) *American Journal of Physiology* **275**, C1481-1486
53. Yang, B., Fukuda, N., van Hoek, A., Matthay, M. A., Ma, T., and Verkman, A. S. (2000) *Journal of Biological Chemistry* **275**, 2686-2692
54. Walter, P., and Blobel, G. (1981) *Journal of Cell Biology* **91**, 557-561
55. Keenan, R. J., Freymann, D. M., Stroud, R. M., and Walter, P. (2001) *Annual Review of Biochemistry* **70**, 755-775
56. Song, W., Raden, D., Mandon, E. C., and Gilmore, R. (2000) *Cell* **100**, 333-343
57. Plath, K., Mothes, W., Wilkinson, B. M., Stirling, C. J., and Rapoport, T. A. (1998) *Cell.* **94**, 795-807
58. Jungnickel, B., and Rapoport, T. A. (1995) *Cell.* **82**, 261-270

59. Belin, D., Bost, S., Vassalli, J. D., and Strub, K. (1996) *The EMBO journal*. **15**, 468-478
60. Sadlish, H., and Skach, W. R. (2004) *J Membr Biol* **202**, 115-126
61. Alder, N. N., Shen, Y., Brodsky, J. L., Hendershot, L. M., and Johnson, A. E. (2005) *Journal of Cell Biology* **168**, 389-399
62. Van den Berg, B., Clemons, W. M., Jr., Collinson, I., Modis, Y., Hartmann, E., Harrison, S. C., and Rapoport, T. A. (2004) *Nature* **427**, 36-44
63. Nilsson, I., Kelleher, D. J., Miao, Y., Shao, Y., Kreibich, G., Gilmore, R., von Heijne, G., and Johnson, A. E. (2003) *The Journal of cell biology*. **161**, 715-725
64. Hegde, R. S., Voigt, S., and Lingappa, V. R. (1998) *Molecular cell*. **2**, 85-91
65. Hegde, R. S., Voigt, S., Rapoport, T. A., and Lingappa, V. R. (1998) *Cell*. **92**, 621-631
66. Voigt, S., Jungnickel, B., Hartmann, E., and Rapoport, T. A. (1996) *The Journal of cell biology*. **134**, 25-35
67. Gèorlich, D., and Rapoport, T. A. (1993) *Cell*. **75**, 615-630
68. Gèorlich, D., Hartmann, E., Prehn, S., and Rapoport, T. A. (1992) *Nature*. **357**, 47-52
69. High, S., Martoglio, B., Gèorlich, D., Andersen, S. S., Ashford, A. J., Giner, A., Hartmann, E., Prehn, S., Rapoport, T. A., Dobberstein, B., and et al. (1993) *The Journal of biological chemistry*. **268**, 26745-26751
70. Do, H., Falcone, D., Lin, J., Andrews, D. W., and Johnson, A. E. (1996) *Cell* **85**, 369-378
71. Martoglio, B., Hofmann, M. W., Brunner, J., and Dobberstein, B. (1995) *Cell* **81**, 207-214
72. Fons, R. D., Bogert, B. A., and Hegde, R. S. (2003) *J Cell Biol* **160**, 529-539
73. White, S. H., and Wimley, W. C. (1999) *Annual Review of Biophysics & Biomolecular Structure* **28**, 319-365
74. Hessa, T., White, S. H., and von Heijne, G. (2005) *Science* **307**, 1427

75. Hessa, T., Kim, H., Bihlmaier, K., Lundin, C., Boekel, J., Andersson, H., Nilsson, I., White, S. H., and von Heijne, G. (2005) *Nature* **433**, 377-381
76. Shi, L. B., Skach, W. R., Ma, T., and Verkman, A. S. (1995) *Biochemistry* **34**, 8250-8256
77. Foster, W., Helm, A., Turnbull, I., Gulati, H., Yang, B., Verkman, A. S., and Skach, W. R. (2000) *J Biol Chem* **275**, 34157-34165
78. Skach, W. R., Shi, L. B., Calayag, M. C., Frigeri, A., Lingappa, V. R., and Verkman, A. S. (1994) *Journal of Cell Biology* **125**, 803-815
79. Skach, W. R., and Verkman, A. S. (1995) *Biophysical Journal* **68**, A334 [Abstract]
80. Buck, T. M., and Skach, W. R. (2005) *J Biol Chem* **280**, 261-269
81. Lu, Y., Turnbull, I. R., Bragin, A., Carveth, K., Verkman, A. S., and Skach, W. R. (2000) *Mol Biol Cell* **11**, 2973-2985
82. Wilkinson, B. M., Critchley, A. J., and Stirling, C. J. (1996) *J Biol Chem* **271**, 25590-25597
83. Lu, Y., Xiong, X., Helm, A., Kimani, K., Bragin, A., and Skach, W. R. (1998) *J Biol Chem* **273**, 568-576
84. Tu, L., Wang, J., Helm, A., Skach, W. R., and Deutsch, C. (2000) *Biochemistry*. **39**, 824-836
85. Heinrich, S. U., Mothes, W., Brunner, J., and Rapoport, T. A. (2000) *Cell* **102**, 233-244
86. Menetret, J. F., Hegde, R. S., Heinrich, S. U., Chandramouli, P., Ludtke, S. J., Rapoport, T. A., and Akey, C. W. (2005) *Journal of Molecular Biology* **348**, 445-457
87. Van den Berg, B., Clemons, W. M., Jr., Collinson, I., Modis, Y., Hartmann, E., Harrison, S. C., and Rapoport, T. A. (2004) *Nature* **427**, 36-44
88. Cannon, K. S., Or, E., Clemons, W. M., Jr., Shibata, Y., and Rapoport, T. A. (2005) *Journal of Cell Biology* **169**, 219-225
89. Sadlish, H., Pitonzo, D., Johnson, A. E., and Skach, W. R. (2005) *Nat Struct Mol Biol* **12**, 870-878

90. Johnson, A. E., and van Waes, M. A. (1999) *Annu Rev Cell Dev Biol* **15**, 799-842
91. McCormick, P. J., Miao, Y., Shao, Y., Lin, J., and Johnson, A. E. (2003) *Molecular Cell* **12**, 329-341
92. Kowarik, M., Kung, S., Martoglio, B., and Helenius, A. (2002) *Molecular Cell* **10**, 769-778
93. Woolhead, C. A., McCormick, P. J., and Johnson, A. E. (2004) *Cell* **116**, 725-736
94. Lu, J., and Deutsch, C. (2005) *Biochemistry* **44**, 8230-8243
95. Duchesne, L., Deschamps, S., Pellerin, I., Lagree, V., Froger, A., Thomas, D., Bron, P., Delamarche, C., and Hubert, J. F. (2001) *Kidney International* **60**, 422-426
96. Aerts, T., Xia, J. Z., Slegers, H., de Block, J., and Clauwaert, J. (1990) *Journal of Biological Chemistry* **265**, 8675-8680
97. Konig, N., Zampighi, G. A., and Butler, P. J. (1997) *Journal of Molecular Biology* **265**, 590-602
98. Verbavatz, J. M., Brown, D., Sabolic, I., Valenti, G., Ausiello, D. A., Van Hoek, A. N., Ma, T., and Verkman, A. S. (1993) *Journal of Cell Biology* **123**, 605-618
99. Smith, B. L., and Agre, P. (1991) *Journal of Biological Chemistry* **266**, 6407-6415
100. Verbavatz, J. M., Ma, T., Gobin, R., and Verkman, A. S. (1997) *Journal of Cell Science* **110**, 2855-2860
101. Neely, J. D., Christensen, B. M., Nielsen, S., and Agre, P. (1999) *Biochemistry* **38**, 11156-11163
102. Beuron, F., Le Caherec, F., Guillam, M. T., Cavalier, A., Garret, A., Tassan, J. P., Delamarche, C., Schultz, P., Mallouh, V., and Rolland, J. P. (1995) *Journal of Biological Chemistry* **270**, 17414-17422
103. Mathai, J. C., and Agre, P. (1999) *Biochemistry* **38**, 923-928
104. Overton, M. C., Chinault, S. L., and Blumer, K. J. (2005) *Eukaryotic Cell* **4**, 1963-1970
105. Deutsch, C. (2003) *Neuron* **40**, 265-276

106. Lu, J., Robinson, J. M., Edwards, D., and Deutsch, C. (2001) *Biochemistry* **40**, 10934-10946
107. Robinson, J. M., and Deutsch, C. (2005) *Neuron* **45**, 223-232
108. Sheng, Z., Skach, W., Santarelli, V., and Deutsch, C. (1997) *Biochemistry* **36**, 15501-15513
109. Lagree, V., Froger, A., Deschamps, S., Pellerin, I., Delamarche, C., Bonnac, G., Gouranton, J., Thomas, D., and Hubert, J. F. (1998) *J Biol Chem* **273**, 33949-33953
110. Bron, P., Lagree, V., Froger, A., Rolland, J. P., Hubert, J. F., Delamarche, C., Deschamps, S., Pellerin, I., Thomas, D., and Haase, W. (1999) *Journal of Structural Biology* **128**, 287-296
111. Manley, D. M., McComb, M. E., Perreault, H., Donald, L. J., Duckworth, H. W., and O'Neil, J. D. (2000) *Biochemistry* **39**, 12303-12311
112. Braun, T., Philippsen, A., Wirtz, S., Borgnia, M. J., Agre, P., Kuhlbrandt, W., Engel, A., and Stahlberg, H. (2000) *EMBO Reports* **1**, 183-189
113. Duchesne, L., Pellerin, I., Delamarche, C., Deschamps, S., Lagree, V., Froger, A., Bonnac, G., Thomas, D., and Hubert, J. F. (2002) *Journal of Biological Chemistry* **277**, 20598-20604
114. van Balkom, B. W., van Raak, M., Breton, S., Pastor-Soler, N., Bouley, R., van der Sluijs, P., Brown, D., and Deen, P. M. (2003) *The Journal of biological chemistry*. **278**, 1101-1107
115. Silberstein, C., Bouley, R., Huang, Y., Fang, P., Pastor-Soler, N., Brown, D., and Van Hoek, A. N. (2004) *American Journal of Physiology - Renal Physiology* **287**, F501-511
116. Bulleid, N. J., and Freedman, R. B. (1988) *Nature* **335**, 649-651
117. Deen, P. M., Croes, H., van Aubel, R. A., Ginsel, L. A., and van Os, C. H. (1995) *J Clin Invest* **95**, 2291-2296
118. Han, Z. B., Yang, J. B., Wax, M. B., and Patil, R. V. (2000) *Curr Eye Res* **20**, 242-247

119. Hirano, K., Zuber, C., Roth, J., and Ziak, M. (2003) *American Journal of Pathology* **163**, 111-120
120. Smith, B. L., Preston, G. M., Spring, F. A., Anstee, D. J., and Agre, P. (1994) *J Clin Invest* **94**, 1043-1049
121. Gavel, Y., and von Heijne, G. (1990) *Protein Engineering* **3**, 433-442
122. Leavitt, R., Schlesinger, S., and Kornfeld, S. (1977) *Journal of Virology* **21**, 375-385
123. Lehle, L., and Tanner, W. (1976) *FEBS Letters* **72**, 167-170
124. Helenius, A. (1994) *Molecular Biology of the Cell* **5**, 253-265
125. Nilsson, I. M., and von Heijne, G. (1993) *Journal of Biological Chemistry* **268**, 5798-5801
126. Baumgarten, R., Van De Pol, M. H., Wetzels, J. F., Van Os, C. H., and Deen, P. M. (1998) *Journal of the American Society of Nephrology* **9**, 1553-1559
127. van Hoek, A. N., Wiener, M. C., Verbavatz, J. M., Brown, D., Lipniunas, P. H., Townsend, R. R., and Verkman, A. S. (1995) *Biochemistry* **34**, 2212-2219
128. Hendriks, G., Koudijs, M., Van Balkom, B. W., Oorschot, V., Klumperman, J., Deen, P. M., and Van Der Sluijs, P. (2003) *J Biol Chem* **279**, 2975-2983
129. van Leeuwen, J. E., and Kearsse, K. P. (1996) *Proceedings of the National Academy of Sciences of the United States of America* **93**, 13997-14001
130. Peterson, J. R., Ora, A., Van, P. N., and Helenius, A. (1995) *Mol Biol Cell* **6**, 1173-1184
131. Anken, E., Braakman, I., and Craig, E. (2005) *Critical Reviews in Biochemistry & Molecular Biology* **40**, 191-228
132. Marr, N., Kamsteeg, E. J., van Raak, M., van Os, C. H., and Deen, P. M. (2001) *Pflugers Archiv - European Journal of Physiology* **442**, 73-77
133. Mulders, S. M., Knoers, N. V., Van Lieburg, A. F., Monnens, L. A., Leumann, E., Wühl, E., Schober, E., Rijss, J. P., Van Os, C. H., and Deen, P. M. (1997) *J Am Soc Nephrol* **8**, 242-248
134. Sousa, M. C., Ferrero-Garcia, M. A., and Parodi, A. J. (1992) *Biochemistry* **31**, 97-105

135. Gelman, M. S., and Kopito, R. R. (2002) *Journal of Clinical Investigation* **110**, 1591-1597
136. Yu, H., Kaung, G., Kobayashi, S., and Kopito, R. R. (1997) *Journal of Biological Chemistry* **272**, 20800-20804
137. Huppa, J. B., and Ploegh, H. L. (1997) *Immunity* **7**, 113-122
138. Wiertz, E. J., Tortorella, D., Bogyo, M., Yu, J., Mothes, W., Jones, T. R., Rapoport, T. A., and Ploegh, H. L. (1996) *Nature* **384**, 432-438
139. Shearer, A. G., and Hampton, R. Y. (2004) *Journal of Biological Chemistry* **279**, 188-196
140. Meacham, G. C., Patterson, C., Zhang, W., Younger, J. M., and Cyr, D. M. (2001) *Nature Cell Biology* **3**, 100-105
141. Connell, P., Ballinger, C. A., Jiang, J., Wu, Y., Thompson, L. J., Hohfeld, J., and Patterson, C. (2001) *Nature Cell Biology* **3**, 93-96
142. Meusser, B., Hirsch, C., Jarosch, E., and Sommer, T. (2005) *Nature Cell Biology* **7**, 766-772
143. Biederer, T., Volkwein, C., and Sommer, T. (1997) *Science* **278**, 1806-1809
144. Vashist, S., and Ng, D. T. (2004) *Journal of Cell Biology* **165**, 41-52
145. Romisch, K. (2005) *Annual Review of Cell & Developmental Biology* **21**, 435-456
146. Pariyarath, R., Wang, H., Aitchison, J. D., Ginsberg, H. N., Welch, W. J., Johnson, A. E., and Fisher, E. A. (2001) *Journal of Biological Chemistry* **276**, 541-550
147. Knop, M., Finger, A., Braun, T., Hellmuth, K., and Wolf, D. H. (1996) *EMBO Journal* **15**, 753-763
148. Ye, Y., Shibata, Y., Yun, C., Ron, D., and Rapoport, T. A. (2004) *Nature* **429**, 841-847
149. Lilley, B. N., and Ploegh, H. L. (2004) *Nature* **429**, 834-840
150. Hitt, R., and Wolf, D. H. (2004) *FEMS Yeast Research* **4**, 721-729
151. Taxis, C., Hitt, R., Park, S. H., Deak, P. M., Kostova, Z., and Wolf, D. H. (2003) *Journal of Biological Chemistry* **278**, 35903-35913

152. Lee, R. J., Liu, C. W., Harty, C., McCracken, A. A., Latterich, M., Romisch, K., DeMartino, G. N., Thomas, P. J., and Brodsky, J. L. (2004) *EMBO Journal* **23**, 2206-2215
153. Ye, Y., Meyer, H. H., and Rapoport, T. A. (2003) *Journal of Cell Biology* **162**, 71-84
154. Oberdorf, J., Pitzonzo, D., and Skach, W. R. (2005) *J Biol Chem* **280**, 38193-38202
155. Heinemeyer, W., Trondle, N., Albrecht, G., and Wolf, D. H. (1994) *Biochemistry* **33**, 12229-12237
156. Groll, M., Ditzel, L., Lowe, J., Stock, D., Bochtler, M., Bartunik, H. D., and Huber, R. (1997) *Nature* **386**, 463-471
157. Heinemeyer, W., Fischer, M., Krimmer, T., Stachon, U., and Wolf, D. H. (1997) *Journal of Biological Chemistry* **272**, 25200-25209
158. Arendt, C. S., and Hochstrasser, M. (1997) *Proceedings of the National Academy of Sciences of the United States of America* **94**, 7156-7161
159. Wolf, D. H., and Hilt, W. (2004) *Biochimica et Biophysica Acta* **1695**, 19-31
160. Dohke, Y., and Turner, R. J. (2002) *J Biol Chem* **277**, 15215-15219
161. Hartmann, E., Rapoport, T. A., and Lodish, H. F. (1989) *Proc Natl Acad Sci U S A* **86**, 5786-5790
162. Lipp, J., Flint, N., Haeuptle, M. T., and Dobberstein, B. (1989) *J Cell Biol* **109**, 2013-2022
163. Xie, Y., Langhans-Rajasekaran, S. A., Bellovino, D., and Morimoto, T. (1996) *J Biol Chem* **271**, 2563-2573
164. San Millan, J. L., Boyd, D., Dalbey, R., Wickner, W., and Beckwith, J. (1989) *J Bacteriol* **171**, 5536-5541
165. Boyd, D., Traxler, B., and Beckwith, J. (1993) *J Bacteriol* **175**, 553-556
166. Skach, W. R., Calayag, M. C., and Lingappa, V. R. (1993) *J Biol Chem* **268**, 6903-6908
167. Tector, M., and Hartl, F. U. (1999) *Embo J* **18**, 6290-6298
168. Bayle, D., Weeks, D., and Sachs, G. (1997) *J Biol Chem* **272**, 19697-19707

169. Ota, K., Sakaguchi, M., von Heijne, G., Hamasaki, N., and Mihara, K. (1998) *Mol Cell* **2**, 495-503
170. van Lieburg, A. F., Verdijk, M. A., Knoers, V. V., van Essen, A. J., Proesmans, W., Mallmann, R., Monnens, L. A., van Oost, B. A., van Os, C. H., and Deen, P. M. (1994) *Am J Hum Genet* **55**, 648-652
171. Fujiyoshi, Y., Mitsuoka, K., de Groot, B. L., Philippsen, A., Grubmèuller, H., Agre, P., and Engel, A. (2002) *Curr Opin Struct Biol* **12**, 509-515
172. Tamarappoo, B. K., and Verkman, A. S. (1998) *Journal of Clinical Investigation* **101**, 2257-2267
173. Yamauchi, K., Fushimi, K., Yamashita, Y., Shinbo, I., Sasaki, S., and Marumo, F. (1999) *Kidney Int* **56**, 164-171
174. Goji, K., Kuwahara, M., Gu, Y., Matsuo, M., Marumo, F., and Sasaki, S. (1998) *Journal of Clinical Endocrinology & Metabolism* **83**, 3205-3209
175. Werner, E. D., Brodsky, J. L., and McCracken, A. A. (1996) *Proc Natl Acad Sci U S A* **93**, 13797-13801
176. Ellgaard, L., and Helenius, A. (2003) *Nat Rev Mol Cell Biol* **4**, 181-191
177. Zwickl, P., Seemèuller, E., Kapelari, B., and Baumeister, W. (2001) *Adv Protein Chem* **59**, 187-222
178. Shinbo, I., Fushimi, K., Kasahara, M., Yamauchi, K., Sasaki, S., and Marumo, F. (1999) *Am J Physiol* **277**, F734-741
179. Tamarappoo, B. K., Yang, B., and Verkman, A. S. (1999) *J Biol Chem* **274**, 34825-34831
180. Han, Z., and Patil, R. V. (2000) *Biochem Biophys Res Commun* **273**, 328-332
181. Evan, G. I., Lewis, G. K., Ramsay, G., and Bishop, J. M. (1985) *Mol Cell Biol* **5**, 3610-3616
182. Skach, W. R., and Lingappa, V. R. (1993) *J Biol Chem* **268**, 23552-23561

183. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) *Gene* **77**, 51-59
184. Rothman, R. E., Andrews, D. W., Calayag, M. C., and Lingappa, V. R. (1988) *Journal of Biological Chemistry* **263**, 10470-10480
185. Xiong, X., Bragin, A., Widdicombe, J. H., Cohn, J., and Skach, W. R. (1997) *Journal of Clinical Investigation* **100**, 1079-1088
186. Sasaki, S., Fushimi, K., Saito, H., Saito, F., Uchida, S., Ishibashi, K., Kuwahara, M., Ikeuchi, T., Inui, K., Nakajima, K., and et al. (1994) *J Clin Invest* **93**, 1250-1256
187. Mulders, S. M., Bichet, D. G., Rijss, J. P., Kamsteeg, E. J., Arthus, M. F., Lonergan, M., Fujiwara, M., Morgan, K., Leijendekker, R., van der Sluijs, P., van Os, C. H., and Deen, P. M. (1998) *J Clin Invest* **102**, 57-66
188. Baum, M. A., Ruddy, M. K., Hosselet, C. A., and Harris, H. W. (1998) *Pediatric Research* **43**, 783-790
189. Kamsteeg, E. J., Bichet, D. G., Konings, I. B., Nivet, H., Lonergan, M., Arthus, M. F., van Os, C. H., and Deen, P. M. (2003) *J Cell Biol* **163**, 1099-1109
190. Parodi, A. J. (2000) *Biochem J* **348**, 1-13
191. Helenius, A., and Aebi, M. (2001) *Science* **291**, 2364-2369
192. Ronnett, G. V., Knutson, V. P., Kohanski, R. A., Simpson, T. L., and Lane, M. D. (1984) *J Biol Chem* **259**, 4566-4575
193. Slieker, L. J., Martensen, T. M., and Lane, M. D. (1986) *Journal of Biological Chemistry* **261**, 15233-15241
194. Kèonig, R., Ashwell, G., and Hanover, J. A. (1988) *J Biol Chem* **263**, 9502-9507
195. Marquardt, T., and Helenius, A. (1992) *J Cell Biol* **117**, 505-513
196. Williams, A. M., and Enns, C. A. (1991) *J Biol Chem* **266**, 17648-17654
197. Hammond, C., Braakman, I., and Helenius, A. (1994) *Proceedings of the National Academy of Sciences of the United States of America* **91**, 913-917

198. Pipe, S. W., Morris, J. A., Shah, J., and Kaufman, R. J. (1998) *Journal of Biological Chemistry* **273**, 8537-8544
199. Jakob, C. A., Burda, P., Roth, J., and Aebi, M. (1998) *J Cell Biol* **142**, 1223-1233
200. McCracken, A. A., and Brodsky, J. L. (2003) *Bioessays* **25**, 868-877
201. de Virgilio, M., Kitzmuller, C., Schwaiger, E., Klein, M., Kreibich, G., and Ivessa, N. E. (1999) *Molecular Biology of the Cell* **10**, 4059-4073
202. Wilson, C. M., Farmery, M. R., and Bulleid, N. J. (2000) *J Biol Chem* **275**, 21224-21232
203. Li, J., Quilty, J., Popov, M., and Reithmeier, R. A. (2000) *Biochemical Journal* **349**, 51-57
204. de Groot, B. L., and Grubmuller, H. (2001) *Science* **294**, 2353-2357
205. Compton, T., and Courtney, R. J. (1984) *J Virol* **52**, 630-637
206. Hasler, U., Greasley, P. J., von Heijne, G., and Geering, K. (2000) *J Biol Chem* **275**, 29011-29022
207. Alder, N. N., and Johnson, A. E. (2004) *J Biol Chem* **279**, 22787-22790
208. Turner, R. J. (2003) *J Membr Biol* **192**, 149-157
209. Blobel, G. (1980) *Proc Natl Acad Sci U S A* **77**, 1496-1500
210. Wessels, H. P., and Spiess, M. (1988) *Cell* **55**, 61-70
211. Moss, K., Helm, A., Lu, Y., Bragin, A., and Skach, W. R. (1998) *Mol Biol Cell* **9**, 2681-2697
212. Skach, W. R., and Lingappa, V. R. (1994) *Cancer Res* **54**, 3202-3209
213. Carveth, K., Buck, T., Anthony, V., and Skach, W. R. (2002) *J Biol Chem* **277**, 39507-39514
214. Lin, J., and Addison, R. (1995) *J Biol Chem* **270**, 6942-6948
215. Beguin, P., Hasler, U., Beggah, A., Horisberger, J. D., and Geering, K. (1998) *J Biol Chem* **273**, 24921-24931
216. Beguin, P., Hasler, U., Staub, O., and Geering, K. (2000) *Mol Biol Cell* **11**, 1657-1672

217. Goder, V., Bieri, C., and Spiess, M. (1999) *J Cell Biol* **147**, 257-266
218. Laemmli, U. K. (1970) *Nature* **227**, 680-685
219. Hubbard, S. C., and Ivatt, R. J. (1981) *Annu Rev Biochem* **50**, 555-583
220. Kaplan, H. A., Welply, J. K., and Lennarz, W. J. (1987) *Biochim Biophys Acta* **906**, 161-173
221. Buck, T. M., Eledge, J., and Skach, W. R. (2004) *Am J Physiol Cell Physiol* **In Press**
222. Nilsson, I., and von Heijne, G. (2000) *J Biol Chem* **275**, 17338-17343
223. Dohke, Y., Oh, Y. S., Ambudkar, I. S., and Turner, R. J. (2004) *J Biol Chem* **279**, 12242-12248
224. Braakman, I., Hoover-Litty, H., Wagner, K. R., and Helenius, A. (1991) *J Cell Biol* **114**, 401-411
225. Ban, N., Freeborn, B., Nissen, P., Penczek, P., Grassucci, R. A., Sweet, R., Frank, J., Moore, P. B., and Steitz, T. A. (1998) *Cell* **93**, 1105-1115
226. Manoil, C., Mekalanos, J. J., and Beckwith, J. (1990) *J Bacteriol* **172**, 515-518
227. Kim, S. J., and Hegde, R. S. (2002) *Mol Biol Cell* **13**, 3775-3786
228. Kim, S. J., Mitra, D., Salerno, J. R., and Hegde, R. S. (2002) *Dev Cell* **2**, 207-217
229. Wagner, C. A., Friedrich, B., Setiawan, I., Lang, F., and Broer, S. (2000) *Cell Physiol Biochem* **10**, 1-12
230. Matlack, K. E., Mothes, W., and Rapoport, T. A. (1998) *Cell* **92**, 381-390
231. Eichler, J., and Duong, F. (2004) *Trends Biochem Sci* **29**, 221-223
232. Katsura, T., Verbavatz, J. M., Farinas, J., Ma, T., Ausiello, D. A., Verkman, A. S., and Brown, D. (1995) *Proc Natl Acad Sci U S A* **92**, 7212-7216
233. Toriano, R., Ford, P., Rivarola, V., Tamarappoo, B. K., Verkman, A. S., and Parisi, M. (1998) *J Membr Biol* **161**, 141-149
234. Leitch, V., Agre, P., and King, L. S. (2001) *Proceedings of the National Academy of Sciences of the United States of America* **98**, 2894-2898

235. Jackson, R. J., Campbell, E. A., Herbert, P., and Hunt, T. (1983) *Eur J Biochem* **131**, 289-301
236. Oberdorf, J., and Skach, W. R. (2002) *Methods Mol Med* **70**, 295-310
237. Walter, P., and Blobel, G. (1983) *Methods Enzymol* **96**, 84-93
238. Zhang, R. B., and Verkman, A. S. (1991) *American Journal of Physiology* **260**, C26-34
239. Oberdorf, J., Carlson, E. J., and Skach, W. R. (2001) *Biochemistry* **40**, 13397-13405
240. Tate, S. S., Urade, R., Micanovic, R., Gerber, L., and Udenfriend, S. (1990) *The FASEB journal : official publication of the Federation of American Societies for Experimental Biology*. **4**, 227-231
241. White, S. H., and von Heijne, G. (2005) *Current opinion in structural biology*. **15**, 378-386