“

Now, a living organism is nothing but a wonderful machine endowed with the most marvellous properties and set going by means of the most complex and delicate mechanism

”

Claude Bernard

An Introduction to the Study of Experimental Medicine (1865)

(translated by Henry Copley Green, 1957)

University of Alberta

Towards a Deeper Structural Understanding of
Eukaryotic Na^+/H^+ Exchangers

by

Grant Alexander Kemp

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Biochemistry

Fall 2013
Edmonton, Alberta

To my loving wife.

Now the adventure begins...

Examining Committee

Dr. Howard Young, Biochemistry, *Supervisor*

Dr. Larry Fliegel, Biochemistry, *Supervisor*

Dr. M. Joanne Lemieux, Biochemistry

Dr. James Young, Physiology

Dr. Michael Walsh, Biochemistry and Molecular Biology, University of Calgary

Abstract

Sodium proton exchangers (NHEs) are polytopic membrane proteins that, in archaea, bacteria, yeast and plants, provide increased salt tolerance by removing excess *toxic* sodium, and in mammals regulate cell volume, growth, differentiation, proliferation, migration and apoptosis in relation to changes in either pH or sodium concentration. As an essential player in cellular physiology it is not surprising that NHE1 dysregulation in the body has been implicated in several diseases, which result from pathological regulation of NHE1 activity. Indeed, the scarcity of structural data has prevented the elucidation of a precise molecular mechanism of ion transport and despite recent technical advances, only a small handful of eukaryotic membrane protein structures have been uncovered. Herein are presented our advances in developing and optimizing an expression system for producing both a full-length human NHE1 protein and larger portions of the transmembrane domain thought to be responsible for ion transport. The three dimensional molecular envelope of human NHE1 was determined by single particle reconstruction electron microscopy, and progress towards determining the structure of transmembrane segments V-VII has been completed by NMR. The data presented in this thesis contribute to an improved understanding of NHE1 function at the molecular level and will help inform future therapeutic development.

Additionally, I present my contribution towards characterizing transmembrane segment IV (TM IV) of sod2, a primary sodium proton exchanger in *Schizosaccharomyces pombe*. Functional analysis of TM IV have uncovered that Thr144–Val147 are critical to competent ion transport and the structural basis for this functional effect was analyzed by NMR, revealing a partially unwound helical conformation of TM IV in the centre of the membrane. To better hypothesize what role TM IV may play in the full length protein, we created a homology model of sod2, which indicated that sod2 TM IV is likely analogous to *E. coli* NhaA TM IV and human NHE1 TM VI. This study further confirmed the importance of partially unwound helices in the transport mechanism of sodium proton exchangers and provides a basis for further experimentation.

Acknowledgements

Completing a PhD requires the input of countless people. Advice, technical help, outside expert opinions and, of course, social time to 'collaborate'. Beginning in my undergraduate career, I want to acknowledge: Dr. Rachel Milner, and Dr. Adrienne Wright who taught me my basic Biochemistry lab skills and have continued to be great supporters into my graduate program; and Dr. Charles Holmes, who was brave enough to accept me as an honours student in his lab. Next I want to thank Dr. Howard Young and Dr. Larry Fliegel who recognized my excitement about science and helped me cultivate it into a PhD. I have learned many lessons about how to be a scientist from them and now, as a colleague, I hope I can take this experience and grow it into my own independent research. Several other faculty members have also been integral as mentors, both personally as well as through their roles in my graduate studies: Dr. Joanne Lemieux, who sat on my committee and was my source for membrane protein purification and crystallization advice; Dr. Joe Casey, who was an external examiner at my candidacy exam and was always willing to offer academic advice; and Dr. Dave Stuart, who was my graduate student adviser and my go–to yeast expert. I also enjoyed many interesting conversations about scientific life with Dr. Leo Spyracopoulos, Dr. Mark Glover and Dr. Brian Sykes who were always willing to spare a moment to talk despite their busy schedules.

I have also been lucky to have been a part of two amazing labs. After two years as a summer student and 6 years as a graduate student many people have passed through the doors of the Young and Fliegel labs. For technical support I thank Angela, Heather, Yong, Leah, and Jodi. They were always willing to help me find what I needed and help me complete my experiments. Also Cathy, Xiuju, Ray, Pratap, Asad, Ludo and Paramita were incredible sources of knowledge that were always available for me to troubleshoot my experimental design. Outside of the lab experimental support was always available from Ross Edwards, Dave Mannock, Richard Rothery, and Jack Moore. There have also been a lot of amazing students that have passed through the lab. Ersillia, Jen, and Aruna were always smiling and making the lab a fun place to be. In my direct sphere I am incredibly indebted to John Paul, Delaine, and Shem for being such amazing lab mates. They made the core of my graduate experience amazing. They were always able to offer a helping hand, a shoulder to cry on, an ear to complain to and most importantly a beer mug to celebrate with. We have had too many great times to include here and I hope many more in the years to come. More recently added to their roster is Joe Primeau. We've become such good friends in

such a short time and I hope that we too can 'collaborate' in the future.

There was also an incredible community of peers and friends that made grad school fun. Stephen Campbell, Charles Leung, Joyce Wong, Jeff Odenbach, Roshani Payoe, Angela Fung, Mridula Swayampakula, Tamara Arnold, Ply Pasarj, Allison Kraus, Greg Workun, Matt Solomonson, Vic Cheng, Justin Fedor, Michelina Kierzek, Jon Lacasse, Kristen Conn, Rory Shott, Rebecca Gibeault, Bernard Kok, Andrew Locke, Brian Lee, Ian Robertson, Olivier Julien, Sue-Ann Mok, Ryan Hoffman, Jordi, Christelle Lazareno-Saez, Nicolas Coquelle, Pam Bonar, Melissa Morrison, Gina Thede, Jason Lamaroux, and any others that I have forgotten here.

Of course grad school would not have been grad school without the help of my beer Friday peers. Dan *Pringles* Prins, Jen *Wastey-Ann* Wang, Curtis Hodge, and Mike Burkat. In the last couple years we've become such good friends (cake at Avenue!) in and outside of the lab. Thanks for celebrating victories and helping numb the failures.

Also a shout out to the muffins! If anyone thought that a bunch of grad students running around a field for 2 hours a week trying to catch a frisbee would result in fame and fortune, they were wrong but, we had a heck of a lot of fun doing it anyway.

Thank you to my parents. You always believed in me even when I wasn't sure and gave me all the opportunities I needed to succeed and finish my PhD. I hope I can continue to make you proud.

Finally, a huge thank you goes to my wife, Sarah. I am the luckiest man alive. You have been generous with your time, energy and love when I was busy with the lab and with other distractions in life. I love you so much. As Howard once said, I've been a scientist for as long as we've been together. Thank you for supporting me in my exploration of the world around me and for putting up with me in all the times when it seemed to take over. I know that we have a lot of fun adventures on the way in the coming decades, and I can't imagine having these with any one else.

Contents

List of Tables

List of Figures

Abbreviations

Akt	Protein kinase B
AOXI	Alcohol oxidase 1 of *Pichia pastoris*
$C_{12}E_8$	Detergent: octaethylene glycol mono-n-dodecyl ether
CAIV	Cabonic annhydrase IV
CaM	Calmodulin
CD	Circular dichroism spectroscopy
CHP1/2/3	Calcineurin homology protein isoforms 1, 2 and 3 (which is also known as Tescalcin)
CPA	Cation Proton Antiporter superfamily of transporters
CPA1/2	Cation Proton Antiporter 1/2 families of transporters
Daxx	Death-associated protein 6
DDM	Detergent: n-dodecyl-β-D-maltoside
DNA	Deoxyribonucleic acid
DPC	Detergent: dodecylphosphocholine = Fos-choline 12
ERK1/2	Extracellular signalling related kinases 1 and 2
ERM	Ezrin, radixin and moesin proteins
EYPC	Lipid: egg yolk l-α-phosphatidylcholine

FC-10	Detergent: Fos-choline 10
FC-12	Detergent: Fos-choline 12
FC-14	Detergent: Fos-choline 14
HEPES-KOH	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid, pH adjusted with KOH
IMAC	Immobilized metal affinity chromatography
KefB	Potassium efflux protein B, CPA2 ancestral gene
LB	Lysogeny Broth
LPC	Detergent: α–lysophosphatidylcholine
MadA	Methylmalonyl-CoA decarboxylase A, NaT-DC ancestral gene
mRNA	Messenger ribonucleic acid
NaT-DC	Na^+-Transporting Decarboxylase family of transporters
NhaA	Sodium hydrogen antiporter A, CPA2 ancestral gene
NHE1	The human Na^+/H^+ exchanger isoform-1
Ni-NTA	Nickel-nitrilotriacetic acid agarose, IMAC resin
p38	p38 mitogen-activated protein kinase
p90	p90 ribosomal s6 kinase
PCR	Polymerase chain reaction
pH_i	Intracellular pH
PIP_2	Phosphatidylinositol-4,5-bisphosphate
PWO DNA polymerase	DNA polymerase from *Pyrococcus woesei*
SDS-PAGE	Sodium dodecylsulfate-polyacrylamide gel electrophoresis
Taq DNA polymerase	DNA polymerase from *Thermus aquaticus*

TEV	Tobacco etch virus
TM	Transmembrane
TMS	Transmembrane segment
Tris	tris(hydroxymethyl)aminomethane

Chapter 1

Introduction

Many sections of this chapter were significantly inspired by the NHE1 structure and function review I wrote with Drs. Howard Young and Larry Fliegel (1). I wrote about 65% of the text and designed all the figures. The majority of this review. This chapter covers background information relevant to understanding and interpreting my results and hypotheses.

The membrane, ion transport and biodiversity

Life on earth ranges from microscopic single celled organisms to complex multicellular mammals. This range of organisms inhabit a similarly diverse range of habitats and metabolic niches, from scorching deep sea vents to the frozen tundra and from acidic soil to the mammalian intestine. How differences in molecular details result in this incredible biodiversity is a fundamental biological question. Through the advent of whole genome sequencing and metagenomics, researchers are discovering that this diversity is even broader than originally thought (2). One of the most important contributors to this biodiversity is the ability of an organism to create and maintain an appropriate intracellular environment for cellular metabolism. Therefore, evolution has generated unique biostructures that are able to deal with temperature extremes, nutrient scarcity, and environmental toxins.

One of these biostructures is the cell membrane, which is composed of a phospholipid bilayer with embedded and associated proteins. These membrane proteins allow the selective accumulation of nutrients and ions in the cell. Since many of these solutes are being trans-

	Cytosolic	Extracellular
K^+	139 mM	4 mM
Na^+	12 mM	145 mM
Cl^-	4 mM	116 mM
HCO_3^-	12 mM	29 mM
Ca^{2+}	<1 μM	1.8 mM

Table 1.1: The concentrations of various abundant ions in the cytosol and the extracellular space of the average human cell.

ported against their concentration gradients, many membrane transport proteins require energy to function—that is, active transporters. This energy can come from the hydrolysis of the γ–phosphoanhydride bond of adenosine 5'–triphosphate (ATP) by so called ATPases. In metazoans, the Na^+/K^+-ATPase maintains a low cytosolic Na^+ concentration and an opposing high K^+ concentration. The relatively low Na^+ concentration provides an inward-facing gradient that is then used as an alternative energy source for secondary active transport.

In bacteria, yeast, and plants, which lack the Na^+/K^+-ATPase, a plasma membrane H^+-ATPase creates a H^+ gradient to power secondary active transporters. However, a consequence of not having a Na^+/K^+-ATPase, is a build up of excess Na^+ and K^+ in the cell, driven by passive diffusion across the membrane. This excess changes the osmolality of the cell, so it must be either removed or sequestered in order for the cell to survive. This is true of all abundant ions in the environment. Therefore, active ion pumps exist that maintain appropriate concentrations of K^+, Ca^{2+}, Cl^-, HCO_3^-, Na^+, H^+ (Table 1.1), and other trace organic and inorganic ions.

Intracellular pH homoeostasis in animals

One of the ions mentioned above is the H^+. Intracellular pH (pH_i) in metazoans is typically maintained at 7.1–7.3, a vanishingly small H^+ concentration of 50–80 nM (3, 4). This limiting concentration is crucial because biomolecules such as proteins and phospholipids are highly protonatable. Globally even small pH changes can affect the surface charge of proteins and membranes in the cell, which in turn affects functional and structural interactions between proteins, proteins and the membrane, ligands and receptors, and ions or solutes with channels and transporters.

A beautiful proof of this principle comes from the expanding field of protein engineering. It has long been understood that protein structure and function are correlated and that the primary

structure—amino acid sequence—of a protein encodes all the information required for protein function (5). The recognition of the nascent polypeptide chain by folding chaperones, trafficking of proteins, the recognition of transmembrane segments by the translocon, the assembly of multisubunit complexes, and synthesizing a fully assembled functional macromolecular complex is not a trivial task. Recent protein engineering efforts using modified simple protein scaffolds to study protein structure-function relationships *ab initio* have produced simple enzymes that are able to catalyse novel chemical reactions (6–8). While these enzymes create products much more efficiently than simple chemical synthesis (i.e in less time and with fewer steps) they are several orders of magnitude less efficient than naturally evolved enzymes. However, when these initial proteins were modified using directed evolution, significant improvements in the rate of catalysis were observed in mutations of the *outer shells* of amino acids—those that interact with the substrate through one or more degrees of separation (7, 9). This demonstrates the importance of the complete folded structure to the function of a protein. In fact there is mounting evidence that the *dynamics* or *plasticity* of a protein structure can be a significant contributor to protein function (10–12). Therefore anything that affects protein structure and dynamics, including pH, will have an overall negative effect on the functioning of a cell, which has evolved to function at a specific pH. In fact a folded protein can significantly affect the pK_a of important residues by creating a unique chemical microenvironment (13, 14). Therefore changing the global pH could affect the ability of a protein to maintain a buried functional microenvironment.

pH_i regulation mechanisms

As cellular metabolism continually alters pH_i (i.e. creating CO_2 (HCO_3^-) and lactate), and the membrane potential drives the passive diffusion of H^+ into the cell, multiple mechanisms for quickly restoring pH_i are central to cell viability. Changes in pH_i are rapidly attenuated using various intracellular physicochemical and biochemical buffers (previously discussed in detail (15, 16)). However, while this buffering capacity decreases excess free H^+ in the cytoplasm, it would quickly become saturated without a mechanism of removing excess H^+. Several acid-extruding membrane transporters exist in the plasma membrane (reviewed (3)); the Na^+/H^+ exchanger (NHE), the Na^+/HCO_3^- cotransporter (NBC), the Na^+-dependent Cl^-/HCO_3^- exchanger (NDCBE), the monocarboxylic acid transporter (MCT), and the vacuolar ATPase (V-ATPase) (Figure 1.1). On the opposing side, the acid-loader (base-extruder) anion exchanger (AE) plays

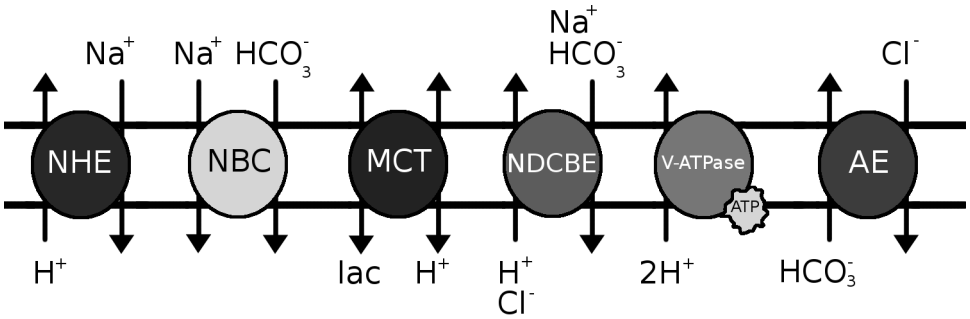

Figure 1.1: The pH rectifying transport proteins found in the plasma membrane of animal cells. NHE, Na^+/H^+ exchanger; NBC, Na^+/HCO_3^- cotransporter; NDCBE, Na^+-dependent Cl^-/HCO_3^- exchanger; MCT, monocarboxylic acid transporter; V-ATPase, vacuolar ATPase; AE, anion exchanger. Exchange stiochiometry is not exact as this is not the same for all isoforms.

the primary role of decreasing cell pH if too much base or too little acid is accumulated[1].

My doctoral work has focused on determining the structure of the human Na^+/H^+ exchanger isoform 1 (NHE1) and correlating this information to function. The first hypothesis of active Na^+/H^+ antiport came in the 1950s (18) and subsequently pH-sensitive sodium transport was described in hepatic mitochondria (19). Ten years later the first measurements of Na^+/H^+ exchange activity were recorded (20, 21) and today thirteen human Na^+/H^+ exchangers have been reported with a variety of distributions, localizations and physiological roles (22).

Evolutionary origins of Na^+/H^+ exchangers

CPA and SLC families

To understand how aspects of cellular physiology are regulated, such as pH_i, and how structure-function relationships that allow for this regulation are shaped, it is appropriate to trace how these arose during evolution (23). The Cation Proton Antiporter (CPA) superfamily, as defined by the Transporter Classification Database (TCDB) (24, 25), includes four families of Electrochemical Potential-driven Transporters (subclass 2.A) and one family of Decarboxylation-driven Transporters (subclass 3.B). Specifically these are: 2.A.36, The Monovalent Cation:Proton Antiporter-1 (CPA1); 2.A.37, The Monovalent Cation:Proton Antiporter-2 (CPA2); 2.A.70, The Malonate:Na^+ Symporter (MSS); 2.A.98, The Putative Sulfate Exporter (PSE); and 3.B.1, The Na^+-transporting Carboxylic Acid Decarboxylase (NaT-DC) families (Figure 1.2). Na^+/H^+ ex-

[1] A Cl^-/OH^- exchanger (CHE) has also been described, but it remains poorly characterized (17)

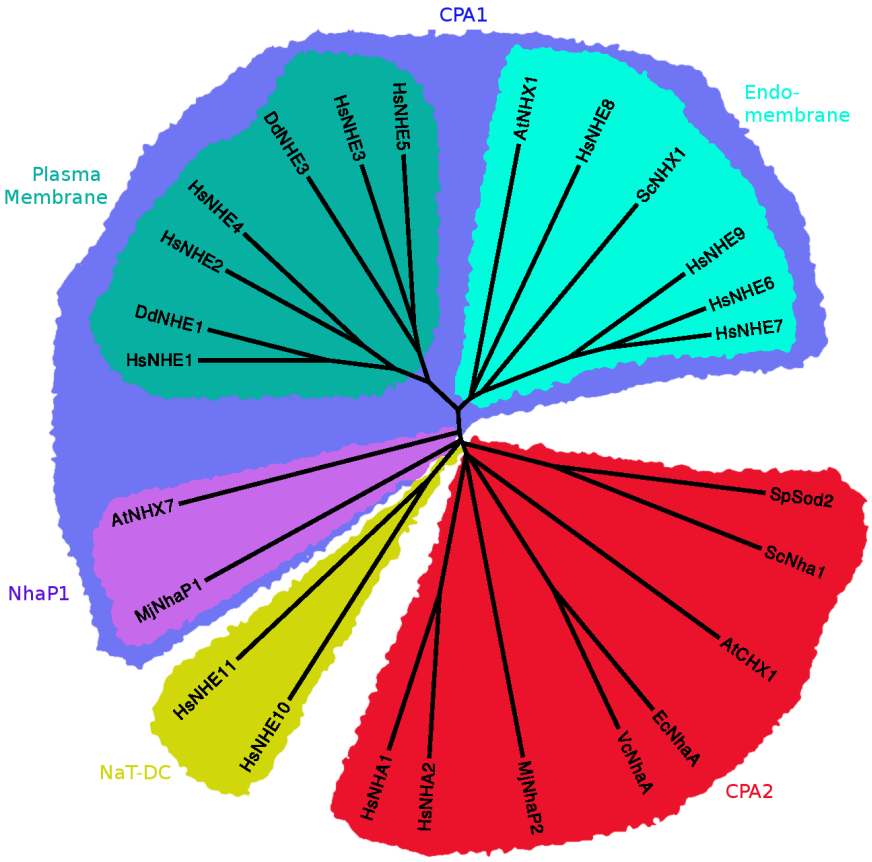

Figure 1.2: An incomplete unrooted CPA superfamily tree. Shown here are the evolutionary relationships of all the human CPA members and several other key and representative CPA transporters. The first two letters of each branch label indicates the species (Hs, *Homo sapiens*; Dr, *Danio rerio*; At, *Arabidopsis thaliana*; Sc, *Saccharomyces cerevisiae*; Mj, *Methanococcus jannaschii*; Sp, *Schizosaccharomyces pombe*; Ec, *Escherichia coli*; Vc, *Vibrio cholerae*. CPA2 members are highlighted in red, NaT-DC members in yellow and CPA1 in blue. CPA1 members are further divided into NhaP1-like (purple), endomembrane (cyan) and plasma membrane (teal).

changers are found in the CPA1, CPA2 and NaT-DC families. Each of these families is classified by the ancestral prokaryotic gene from which it has arisen: CPA1 from NhaP, CPA2 from NhaA and KefB, and NaT-DC from MadA (23).[2] Here I will focus on the evolution of CPA1 and CPA2 families as these both have well characterized members that are pertinent to this thesis.

The CPA1 and CPA2 families are composed of secondary active transporters that tightly couple the stoichiometic transport of either Na^+, Li^+, K^+, NH_4^+ or Ca^{2+} to H^+ transport. While members of these two families are related in function, and likely in structure (see more below), two important differences exist: stoichiometry and primary substrate. The archetypical CPA1

[2]Kef, K^+ efflux; Mad, methylmalonyl-CoA decarboxylase

transporter is defined by the TCDB as electroneutral, able to exchange one H^+ for one Na^+ (or Li^+), although notably many CPA1 transporters, including the intracellular mammalian isoforms, are able to transport either Na^+ or K^+ efficiently. Conversely the archetypical CPA2 member is an electrogenic exchanger using 2 H^+ to move 1 Na^+. While both CPA1 and CPA2 transporters are $Na^+(K^+)/H^+$ exchangers, they do not actually perform the same function in all organisms. Since it is thought that these proteins are capable of bi-directional exchange (26), depending on the magnitude of the prevailing electrochemical gradient, they are tuned for different cellular physiology. In prokaryotes, yeast, and plants a plasma membrane H^+-ATPase creates an inward pH gradient that is used as the driving force for CPA proteins to export Na^+ or K^+. In animals a Na^+/K^+-ATPase creates an extracellular Na^+ gradient that is used by CPA members to alkalinlize the cell. Analyzing the available functional and phylogenetic data, it appears that CPA2 members typically work as Na^+ or K^+ exporters coming from the more ancient NhaA and KefB genes. Evolving from here the ancestral NhaP motif was still able to transport both Na^+ and K^+ but began to primarily use a pH gradient. In eukaryotes, the early evolved CPA1 genes, which still have K^+ transport ability, localize to intracellular membranes using the high cytosolic K^+ concentration to regulate the pH of subcellular compartments. Therefore, the last CPA members to appear were Na^+ selective plasma membrane transporters seemingly after the appearance of the Na^+/K^+ ATPase on the plasma membrane that creates the Na^+ gradient used by many plasma membrane secondary transporters to import solutes. Additionally this phenomenon allows animal cells, which lack a rigid cell wall, to use the Na^+/K^+ ratio to drive water movement thereby maintaining the appropriate cell volume and shape. CPA members contribute to this process by controlling Na^+ or K^+ movement in response to cell stimuli. A very thorough analysis, including expanded phylogenetic trees, for the CPA superfamily has been published by Brett *et al.* (27).

While the TCDB provides a universal nomenclature for transport proteins, similar to the Enzyme Commission, an early convention for human genes also exists. In the 1960s, as the field of human genetics was quickly expanding, concerns arose about possible confusion caused by inconsistent gene nomenclature. By 1979 a standard convention was created and is currently maintained by the Human Genome Organization naming conventions (see HGNC). Under this system all human genes that encode a secondary active transporter are placed into the Solute Carrier family (SLC). The diversity of transport substrates includes sugars, nucleotides, amino acids, drugs and ions making the SLC family the second largest family of human membrane proteins

with at least 384 members (28). The SLC9 subfamily contains all the known human Na^+/H^+ exchangers thereby forming the human subfamily of the TCDB CPA family. As HGNC divisions are based solely on protein function the SLC9 subfamily has been divided into three subgroups to better reflect phylogenetic relationships defined in the TCDB (22). SLC9A contains NHE isoforms 1-9 (CPA1); SLC9B contains the less well studied NHA1 and NHA2 (CPA2); and SLC9C contains recently discovered Sperm-NHEs (NaT-DC). As noted above, while each of these three subgroups share the same function ($Na^+(K^+)/H^+$ antiport) they have travelled different evolutionary paths leaving them without significant sequence homology. These many differences make the exploration of structure-function relationships increasingly interesting as it is thought that the CPA family likely shares a similar three dimensional structure and cation coordination sites (29). Differences in cation selectivity, stoichiometry, and physiological functions likely result from subtle structural changes (30), whereas functional regulation is more closely associated with differences in the regulatory domains (31–33). To avoid confusion I will use NHE when referring to SLC9A isoforms, SLC9 for all human Na^+/H^+ exchangers and CPA when referring to Na^+/H^+ antiporter relatives across different organisms (including humans).

SLC9—the human NHEs

An excellent description of NHE physiology has been published (3); a brief overview demonstrating the way that these isoforms work together is presented here (Figure 1.3).

SLC9A: Plasma membrane NHEs

As Na^+/H^+ exchangers have a central role in cellular regulation at least one NHE isoform is present in every cell in the body. Some isoforms are ubiquitously expressed while others have a limited tissue distribution. They also vary significantly in their sequence homology (Table 1.2). In the kidney and gastrointestinal (GI) tract, where polarized epithelial cells control absorption, secretion, and reabsorption of Na^+, H^+ and fluid, different apical and basolateral NHE isoforms that can be independently regulated are required. In these tissues, NHE1 and 4 are found on the basolateral membrane and NHE2 and 3 are found on the apical membrane (34). NHE2, probed using *Nhe2-/-* mice, plays a significant role in the viability of gastric parietal cells and parotid gland fluid secretion (35) while not significantly affecting intestinal and kidney ab-

1	1								
2	44.9	1							
3	38.6	39.9	1						
4	39.9	51.8	36	1					
5	35.5	37.6	47.7	34.2	1				
6	27.4	28.7	28.9	26.9	26	1			
7	26.6	28.6	27.3	26.6	26.8	63.3	1		
8	28.3	27.2	27.1	27.8	25	31.7	31.1	1	
9	27.3	27.8	27	27.8	26.2	53.8	53.8	33.3	1
NHE	1	2	3	4	5	6	7	8	9

Table 1.2: SLC9A sequence identity comparison table. Values are percent identity and were calculated using individual Clustal W pairwise alignments.

sorption (35–37). NHE3, on the other hand, significantly contributes to Na^+ absorption in the intestine following feeding stimulus (38, 39) and is responsible for the bulk of Na^+ and fluid reabsorption in the proximal tubule of the kidney (40, 41). The resultant acidification of the urine by NHE3 also drives up to two-thirds of HCO_3^- resorption (42–44). Conversely, NHE4 in colonic crypts, is upregulated in response to aldosterone stimulation while no activation of NHE1 or NHE3 was detected, providing a unique mechanism to increase fluid reabsorption specifically in the intestine (45). Similarly in the kidney NHE4 appears to account for basolateral NHE activity in cells lacking NHE1 expression (34, 46), perhaps also allowing aldosterone-specific stimulation in these cells. NHE5 is expressed nearly exclusively in the brain (47, 48) in a unique pool of recycling vesicles that are moved to the plasma membrane upon stimulation (49–51). Although they do not appear to be involved in regulating neuronal signalling (52), they do have a role in the growth of dendritic spines (51).

SLC9A: Intracellular NHEs

While NHE1-5 are active on the plasma membrane, with some having a limited tissue distribution, NHE6-9 have distinct intracellular localizations and are expressed ubiquitously. The discovery of the first intracellular Na^+/H^+ exchanger, *Saccharomyces cerevisiae* Nhx1 (53), the (mis)identification of a mitochondrial NHE (54)[3], and the fact that the acidic pH of various subcellular compartments is essential for proper membrane trafficking (56), posited the idea that NHEs may have a role in vesicular acidification (55). While the vacuolar ATPase (V-ATPase) has been shown to be the the primary acidification mechanism (57), recent work has shown

[3] A later analysis identified that this was more likely endosomal NHE6 (55)

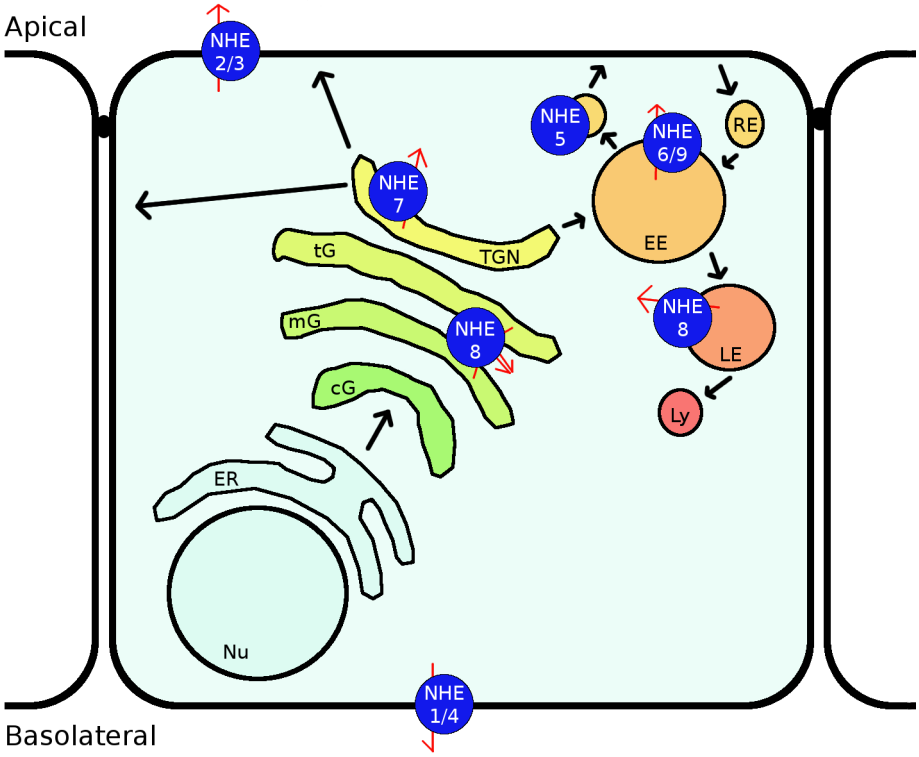

Figure 1.3: Subcellular localization and pH regulation by the nine SLC9A isoforms in a polarized human cell. The apical and basolateral plasma membrane labelled, separated by a tight junction between the adjacent cells (small black circle). All SLC9A isoforms are shown in the compartment in which they are thought to most often reside and each compartment is coloured from blue to red with decreasing pH. A single representative cell is shown for simplicity, though not all isoforms exist in all cell types nor co-exist together (see text for more detail). The compartments (and pH) are: Cy, cytoplasm (7.2); Nu, nucleus (7.2); ER, endoplasmic reticulum (7.2); cG, *cis*-Golgi (6.7); mG, *medial*-Golgi (6–6.7); tG, *trans*-Golgi (6–6.7) ;TGN, Trans Golgi Network (6–6.7); EE, early endosome (6.3); RE, recycling endosomes (6.5); LE, late endosome (5.5); and Ly, lysosome (4.7). Compartment pH data was taken from Casey *et al.* (4).

that several compartments have Na^+/H^+ and Cl^-/H^+ exchangers that may precisely tune the pH (4, 58, 59). NHE6 and NHE9 are found in early and recycling endosomes (55), NHE7 in the *trans*-Golgi (60), and NHE8 in the *medial-trans*-Golgi (58) and late endosomes (61). However, there is also significant evidence that suggests that these *intracellular* NHEs may appear in small tightly regulated pools on the plasma membrane in response to specific stimuli (55, 62–64).

SLC9B: the other cousin

While the main focus of research on mammalian Na^+/H^+ exchangers has been on the SLC9A subfamily, a few interesting and exciting studies have been published about transporters in the

other two SLC9 subfamilies. It has long been known that erythrocytes have a mysterious Na^+/Li^+ counter transport capability (65) that is elevated in hypertension (66, 67). While NHEs are capable of this function (68), Na^+/Li^+ antiport in human erythrocytes is not sensitive to NHE inhibitors (69). For many years this question remained unanswered until an inhibitor-resistant splice variant of NHE1 was identified that was capable of Na^+/Li^+ but not Na^+/H^+ exchange (70). However, while this may be a candidate for the mysterious function first described by Haas *et al.*—a discussion of the potential structural and mechanistic consequences has been published (1)—no further data have been collected to corroborate this conclusion. More recently, SLC9B2 has been identified as a candidate for essential hypertension based on tissue distribution, cellular localization, and inhibitor sensitivity (71). The first study to functionally characterize NHA2, exogenously expressed in mammalian cells lacking any NHEs, was unable to find any Na^+-driven H^+ efflux (72). Later work, however, measured H^+-driven Na^+ antiport in concert with a plasma membrane V-ATPase (73). This is initially a surprising result for a mammalian plasma membrane Na^+/H^+ exchanger that typically use the strong Na^+ gradient for secondary active transport. It is consistent, though with its closer genetic cousin *E. coli* NhaA and other members of the CPA2 family. Although these transporters are bi-directional, as mentioned above, and work with the prevailing chemiosmotic gradient, there is abundant evidence that indicates that they are strongly regulated to work in one particular direction (3, 74, 75), consistent with activity of SLC9B2. This transporter along with SLC9B1 provide a new opportunity to study structure-function relationships in mammalian membrane proteins. Do these transporters from the same superfamily differentially regulate cellular activity? Do they work together or are they opposed? And how do potentially subtle structural differences manifest these functionally large changes?

SLC9A1: a cellular butler

NHE1, often referred to as the *housekeeping enzyme* due to its role in cell volume and pH_i regulation, is expressed on the plasma membrane of nearly all cells in the body. In the decades following its discovery, it has been shown to be a central regulator of cellular status with direct contributions to cell volume, growth, proliferation, migration and apoptosis (76–79). NHE1 participates in these diverse physiological roles through modulation in response to a multitude of cellular signals (80–82) (Figure 1.4). Perhaps then NHE1 could be more aptly described as a *cellular butler* rather than a simple *housekeeper*. The regulatory interactome of NHE1 is vast and cannot be covered here in detail, but several excellent functional reviews (both general and tissue-specific) are available (3, 80, 83–85). Briefly, there are three primary ways that NHE1 interacts with the cell: protein-protein interactions, protein-lipid interactions, and protein phosphorylation.

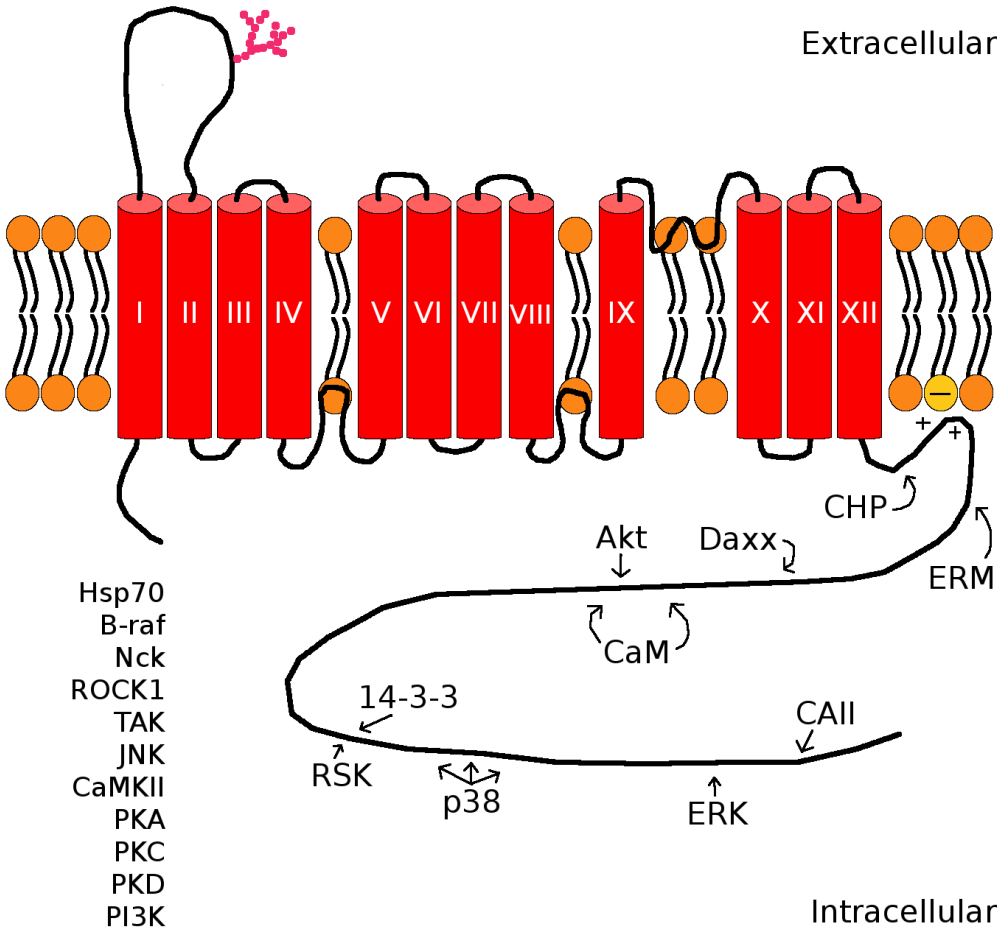

Figure 1.4: A two dimensional model of NHE1 showing the approximate location of the regulatory factors mentioned in the text. The topology is taken from Wakabayashi *et al.* with the transmembrane segments numbered with Roman numerals. The phospholipid bilayer is coloured with orange head groups except one labelled with a (-), which represents NHE1 tail interaction with phosphatidylinositol–4,5–bisphosphate (PIP_2), diacylglycerol, and phosphatidylethanolamine. The approximate interaction location is shown by an arrow (except for those for which the location is unknown). CHP, calcineurin homology protein (1/2/3); ERM, ezrin/radixin/moesin/4.1R; Daxx, death-associated protein 6; CaM, calmodulin (two sites); Akt, protein kinase B; 14–3–3, scaffold protein; RSK, p90-ribosomal S6 kinase; p38, p38-mitogen activated protein kinase; ERK, extracellular signal-regulated kinase; CAII, carbonic anhydrase II; Hsp70, heat shock protein 70; B-raf, B-raf kinase; Nck, Nck-interacting kinase; ROCK1, Rho-associated, coiled-coil containing protein kinase 1; TAK, transforming growth factor β-activated kinase; JNK, c-Jun N-terminal kinase; CaMKII, Ca^{2+}/calmodulin-dependent kinase II; PKA, protein kinase A; PKC, protein kinase C; PKD, protein kinase D; PI3K, phosphoinositide–3–kinase.

Protein-protein interactions

Several NHE1 protein-protein interaction partners have been well characterized (80). Four of these are calcium binding proteins: calmodulin (CaM), Calcineurin homologous Protein (CHP) 1, CHP2, and CHP3 (also known as tescalcin). This indicates that NHE1 is connected to cellular calcium signalling. Calmodulin, in response to calcium binding, binds two contiguous sites, one of high affinity and one of low affinity, on NHE1 (86). It is thought that calmodulin binding blocks a H^+-sensor *auto-inhibitory* binding region thereby activating NHE1 (87, 88), in response to angiotensin hormone signalling (89). CHP1 is a ubiquitously expressed (90) constitutive activator of NHE1 (91). Mutations that prevent their interaction decrease NHE1 proton affinity thereby decreasing basal NHE1 activity by 80% (92). Similarly, CHP2 is over-expressed in cancer cells and prevents cell death by activating NHE1 under serum-deprived conditions (93). Conversely CHP3 is highly expressed in cardiomyocytes and has been shown to interact with NHE1 (94) inhibiting serum-dependent activation (95) and promoting maturation and retention on the plasma membrane (96). NHE1 half-life on the plasma membrane is further enhanced by CHP3 myristoylation (97). Interestingly, all CHP isoforms appear to interact with the same binding site suggesting a tissue-specific regulatory mechanism (95, 96), although CHP3 also appears to have another more distal binding region (95).

The function of NHE1 and the acid-loader anion exchanger AE are functionally coupled by carbonic anhydrases, which interconvert $HCO_3^- + H^+ \rightleftharpoons CO_2 + H_2O$. They are also physically tethered together in a *transport metabolon* with NHE1 interacting with carbonic anhydrase CAII (98), which interacts with both AE1 (99) and AE3 (100). These interactions have also been implicated in cardiac hypertrophy (101, 102).

NHE1 also binds to the ezrin/radixin/moesin (ERM) proteins of the cytoskeleton that regulate cell shape, migration, and focal adhesion (76, 103). Phosphorylation of NHE1 by protein kinase B (see more below) recruits ERM proteins to promote cell survival by maintaining cell volume and shape (104). NHE1 has also been noted to interact with 4.1R another major structural component of the cytoskeleton (105). A contrasting situation arises during ischaemic stress, where death-associated protein 6 (Daxx) is activated and competes with ERM binding, prevents the cell survival signal, and leads to cell death (106). Evidence of binding of anti-apoptotic heat shock protein 70 (Hsp70) also exists (107, 108) and while acidification by NHE1 has been

implicated indirectly (109, 110), recent experiments suggest that Hsp70-NHE1 interactions mediate the inflammatory response to lipopolysaccharide challenge (111). Interestingly, knock-out of a chaperone stress 70 protein, which is Hsp70-like, results in impaired pH_i recovery but it is unknown whether this is a result of a direct interaction (112). The protein signalling scaffold protein 14–3–3 binds at Ser703 in response to p90-ribosomal S6 kinase (see below) phosphorylation, activating NHE1 activity (113). While no further downstream partners have been identified, 14–3–3 could act to prevent dephosphorylation thereby sustaining NHE1 stimulation. Several other interaction partners have also been identified by an antibody array but have not been functionally characterized (108).

Protein-lipid interactions

While initially regarded as a simple barrier and protein trestle, membrane lipids have been recognized to play important structural and functional roles (114). While this includes superstructures like lipid microdomains or *rafts*, individual lipid molecules can also be essential for protein function (115, 116). Phosphatidylinositol–4,5–bisphosphate (PIP_2) is a plasma membrane signalling lipid that becomes depleted when cellular ATP levels decrease (117). The observation that NHE1 activity also decreases upon ATP depletion, though it does not require ATP to function, led to the discovery that NHE1 binds PIP_2 (118). PIP_2 is known to bind to several other plasma membrane ion transporters (117). It has also been found that direct interaction of phosphatidylethanolamine and diacylglycerol with NHE1 is the mechanism by which phorbol-esters stimulate NHE1 activity (119).

Kinases

Protein kinases comprise the most diverse NHE1 regulatory mechanism. NHE1 is stimulated by extracellular signal-regulated kinase (ERK) (120) and p38-mitogen activated protein kinase phosphorylation (121). Phosphorylation of amino acids Thr718, Ser723, Ser726, and Ser729 by these kinases in pro-B cells stimulates pro-apoptotic pathways (121) and mutating residues Ser726 and 729 to alanine protects against serum-deprived cell death (122). ERK 1/2 in cardiac myocytes phosphorylates Ser770 and Ser771 of NHE1 stimulating NHE1 during acidosis (120, 123), while downstream p90-ribosomal S6 kinase (RSK) phosphorylates at S703 (124)

forming a 14–3–3 binding site (113). Interestingly, while RSK stimulation has been linked to ischaemic injury (125), RSK is not activated during sustained intracellular acidosis in cardiac cells (126), indicating that acidosis during ischaemia is not what leads to RSK activation. Protein kinase B (Akt) phosphorylates Ser648 of NHE1 but this has been shown to have different effects in different cells. In cardiac myocytes it inhibits NHE1, possibly by interfering with CaM binding (127), while in fibroblasts phosphorylation stimulates NHE1 activity increasing cell survival and possibly metastasis (104). Other kinases are known to directly phosphorylate NHE1, but they are less well characterized. These include Nck-interacting kinase (128), Ca^{2+}/calmodulin-dependent kinase II (129), transforming growth factor β-activated kinase 1 (81), phosphoinositide–3–kinase, and c-Jun N-terminal kinase (130). Activation of the RhoA kinase cascade facilitates focal adhesion by stimulation of NHE1 through the downstream Rho-associated, coiled-coil containing protein kinase 1, although it is unknown if this phosphorylation happens directly (131). NHE1 is also activated indirectly by protein kinase A (132), C (133), D (134), and phosphoinositide–3–kinase (135). Recently B-raf, a kinase widely implicated in cancer, has been shown to interact with and stimulate NHE1 activity, though this may not be by direct phosphorylation (136).

Countering the regulation of NHE1 by protein kinases are a handful of protein phosphatases. Although specific sites have not been characterized, protein phosphatase 1 (PP1) interacts with and seemingly dephosphorylates all sites *in vitro* and a PP1 inhibitor prevented NHE1 kinase-dependent activation *in vivo* (137). Protein phosphatase 2A appears to be more specific. It poorly dephosphorylates NHE1 *in vitro*, but inhibits the α-adenosine-receptor stimulation of NHE1, which likely corresponds to a specific ERK or RSK site (138). SHP-2, a Src-homology domain containing tyrosine phosphatase, has been shown to have a functional linkage to NHE1 regulation but the precise mechanism is still unknown (108). Very recently calcineurin has been shown to bind to NHE1 leading to the activation of nuclear factor of activated T-cells signalling and cardiac hypertrophy (139).

While traditionally it is thought that membrane transporters are activated by phosphorylation events, there is evidence that NHE1 actually acts as a *kinase scaffold* that directs the propagation of a signal (81, 140). This highlights the importance of NHE1 function in the overall balance between cell survival and cell death; and also why it has been so strongly implicated in tumour development and metastasis.

Nhe1-/- mouse model

Despite the significant role of NHE1 in the body, a *Nhe-/-* mouse is viable and indistinguishable from its littermates until two weeks of age, suggesting that NHE1 is not required for proper embryogenesis and neonatal development (141). The knockout animals are also able to mate and bear offspring, although the mothers die a few days postpartum and the knockout animals have retarded growth after two weeks of age. These results recapitulate an earlier study on a spontaneous mutant in the *Nhe* allele (termed *swe*) (142). The phenotype includes: an ataxic gait, excitability followed by total inactivity for a brief period, increased mortality before weaning, and probable death by convulsive seizure. Interestingly, it was later found that the NHE1 promoter is activated most highly at 12–15 days of age (143), consistent with the previously observed phenotype (141, 142). There is also evidence that the seizures are directly correlated to increased neuronal excitability, caused by an increased Na^+ current density, induced by higher expression of Na^+ channels (144). Many other tissues of the *swe* or knockout mouse, including gastric parietal cells (36), the renal thick ascending limb (145), and astrocytes (146), have specific alternations in organ function, and altered gene expression has been observed in the brain (147).

NHE1 and disease

As described above, NHE1 is involved in preserving ion homoeostasis and coordinating a variety of cell signalling events that require elaborate regulatory mechanisms. Perhaps then it is not surprising that conditions that upset this homoeostasis lead to NHE1 dysregulation and pathophysiology, either in response to erroneous stimuli or a drastic change in pH_i. Research has revealed that NHEs play a significant role in many diseases including hypertrophy and ischaemia/reperfusion injury in the heart (3) as well as tumour metastasis (148). In ischaemia, increased Na^+/H^+ exchange activity, induced by hypoxia, leads to activation of the Na^+/Ca^{2+} exchanger (NCX) and subsequent calcium-induced cell death (3). Cardiac hypertrophy results from pathological activation of NHE1 leading to increased cell growth, thereby thickening the ventricular walls (80). In some forms of cancer, NHE1 is activated leading to the creation of an acidic microenvironment surrounding tumour cells, which promotes metastasis (82).

Studies in mammalian cells and mice have shown that NHE1 blockers can be effective treat-

ments for ischaemic organ disease (146, 149–151) and pre-clinical trial experiments on a small cohort of patients receiving coronary angioplasty also indicated some beneficial effects (152). However, large scale clinical trials were unsuccessful (153–157). It has been suggested that this may have been due to a combination of dose and treatment timing, where too small a dose had no significant effect and too large a dose given after the onset of ischaemia actually aggravated reperfusion recovery leading to a negative outcome (83). Additionally, there was a small but significant increase in cerebrovascular events (strokes) in the treatment group of one study, which corresponded with increased mortality (157). These events are not consistent with previous work on NHE1 inhibitors, which generally demonstrated cerebroprotection, and the mechanism remains unknown (158). Despite this, the data supporting the beneficial effects of NHE inhibitors in human treatment continue to grow (83, 158, 159) and a better understanding of NHE structure, function, and regulation may be necessary for the development of safe and effective therapies (160). It may also be possible that a better structural understanding of the conformational changes that occur in response to NHE1 regulation may provide the basis for designing an inhibitor that only interacts with the *stimulated* conformation. This type of inhibitor would be very effective in tissues that contain chronically overstimulated NHE1 without interrupting basal activity in any other cells. Additionally, NHE1 has been identified as playing a major role in tumour metastasis by acidifying the tumour microenvironment and increasing the invasiveness of cancer cells (148). Again, a better understanding of the structural changes that result upon NHE1 activation in tumours may provide therapeutic options for preventing the spread of cancer in the body. A large amount of research has been published on the clinical relevance of NHE1 in several human diseases and some excellent reviews have been published that summarize these findings (80, 82, 161–163).

NHE structure and function

NHE1 is a 90 kDa transmembrane protein comprised of 815 amino acids divided into two functional domains: an amino-terminal membrane domain, responsible for ion transport; and a carboxy-terminal cytosolic regulatory domain (or *tail*), containing the known sites of activity modulation (Figure 1.4). The membrane domain also contains an allosteric H^+ binding site responsible for basal NHE1 activation at low pH (164) and has residues that are known to interact

with various NHE1 inhibitors (165–169). Likewise all activity modulation by cellular signalling is achieved through the cytosolic regulatory domain, also called the *tail*, allowing different isoforms (i.e. different tails) to have different responses to signalling events. Elegant experiments with NHE1 and NHE3, which are found on the basolateral and apical membranes, respectively, show that NHE1/3 chimeras respond to cellular signalling via the tail and to inhibitors via the membrane domain (31–33). NHE1 is a dimer both in the membrane (170) and in mild detergent solution (171) and although oligomerization may be synergistic and important for stability, it is not essential for activity (172). The first structural predictions were presented in a topology model from Wakabayshi *et al.* (173). The model used hydropathy scales (174) to determine an initial possible topology that was then refined using Substituted Cysteine Accessibility Modelling (SCAM) to indicate twelve transmembrane segments and three intramembrane loops (Figure 1.8-A). This model provided a template for additional directed mutagenesis studies that led to the identification of several functionally critical residues. It has also been somewhat helpful in the design of peptide constructs for NMR structural studies and to date our group has determined the structures of five transmembrane segments and two extracellular loops by NMR (168, 175–180). Additionally, we have determined a low resolution molecular envelope by single particle reconstruction of electron microscopy images which is covered in detail in *Chapter 2*. While this data was being collected and attempts to determine a high resolution structure of NHE1 were underway, a novel computational approach was taken by Landau *et al.* to produce an alternative three-dimensional model (181), using the recently published crystal structure of *E. coli* NhaA, a CPA2 member (182) (Figure 1.8-B). This was a very exciting new addition to the developing but limited NHE1 structural knowledge base and we set out to analyze the differences between the two models. As my PhD progressed, we assessed these two models using the data that I have collected and we have attempted to ascertain which model, or more likely which parts of each model, best fit the data collected to date.

The carboxy-terminal regulatory tail

In addition to structural data on the transmembrane domain, some biochemical and biophysical analysis of regulatory domain has also been completed. As reviewed above, the carboxy-terminal tail is responsible for the specific regulation of NHE isoforms and therefore, its structure may be useful for understanding NHE modulation. To date three high resolution structures of short

regions of the tail bound to the calcium binding regulatory partners CHP1, CHP2 and CaM have been elucidated. These structures reveal that these regulators bind to α–helical regions of NHE1 (183–185). As noted above, these proteins bind to the proximal region of the tail (i.e. residues 503–685), consistent with its role as a scaffold for protein–protein interactions. However, analysis of the entire tail by circular dichroism has indicated that it contains a significant proportion of random coil structure (186). Recently, the distal region of the tail (residues 686–815), which contains the sites of phosphorylation, has been found to be intrinsically disordered and this disorder is well conserved over several species (187). This same report describes conserved molecular recognition motifs that become transiently structured upon interaction with a regulator. When the authors disrupted one of these motifs with a point mutation, which prevented the formation of a transient α–helix, the resultant full length protein was poorly trafficked to the plasma membrane. As this is an intrinsically disorder region, there cannot be a folding defect *per se*. However, the chaperone proteins Hsp70 and chaperone stress 70 protein bind to the distal tail (107, 112), so a disruption in their binding could result in impaired trafficking.

NhaA and NhaP1 structure and function

In order to properly compare these two models—one a two-dimensional hydropathy model informed by mutagenesis experiments in living cells, the other a three-dimensional computational model constructed *in silico* informed by bioinformatics and scrutinized by the available mutagenesis data—we must examine the data on *E. coli* NhaA, the only Na^+/H^+ exchanger whose high resolution structure has been determined. NhaA was identified and cloned from bacteria in 1987 (188), only two years before NHE1, and an overproduction and purification system was quickly developed making a plethora of *in vivo* and *in vitro* experiments possible (189). On the other hand, the difficulty of expressing and purifying mammalian membrane proteins for equivalent purposes has prevented the same rapid advancement in structural knowledge. In less than ten years a two-dimensional projection map was achieved at 4 Å resolution (190) and a three-dimensional structure at 7 Å resolution (191). This was further improved by attaining three-dimensional crystals leading to the elucidation of an x-ray structure at 3.75 Å resolution (182) (Figure 1.5-A). At the time, a simple transport mechanism was also proposed (Figure 1.5-B) that indicated that the charged residues Asp163 and Asp164 located in the centre of the membrane are likely involved in ion binding and transport. Informed by the crystal structure, more functional

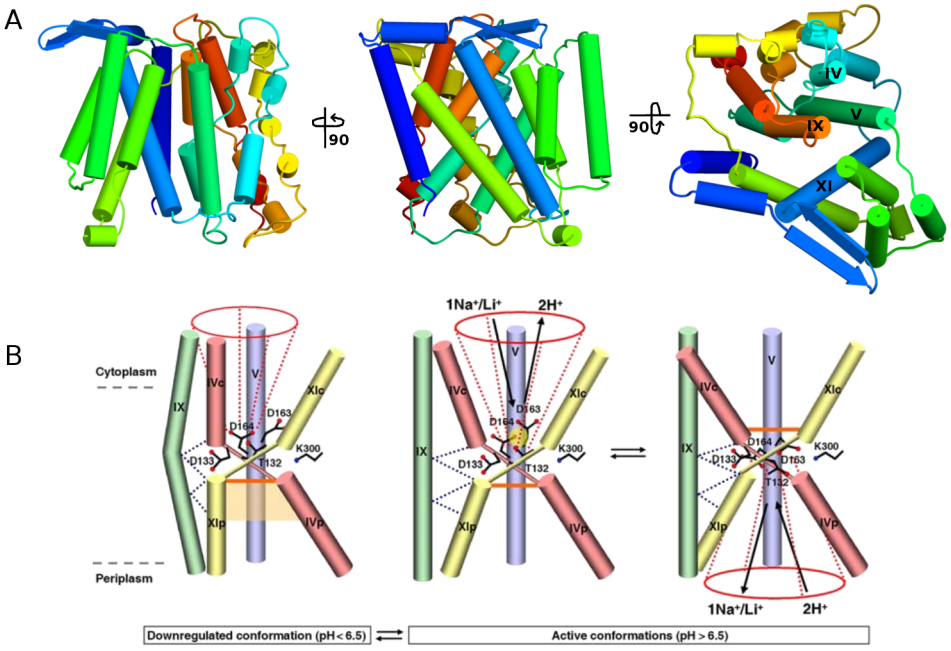

Figure 1.5: Structure and key functional residues of *E. coli* NhaA. ***A***, Three views of a cartoon representation of *E. coli* NhaA coloured like the rainbow starting with transmembrane helix I in blue going to transmembrane helix XII in red (PDB code: 1ZCD (182)). In the far right image the transmembrane helices shown in the next panel are labelled for reference. ***B***, Key residues and initial transport mechanism. The far left panel shows the arrangement of helices in the transport bundle (IV, V, IX, and XI) in the crystal structure (inactive). The middle and left panels represent putative conformational changes made during transport. Residues are numbered and their interactions with other helices or ions are labelled with dotted lines. Panel ***B*** was reprinted with permission from Macmillan Publishers Ltd: *Nature* (2005) **435**, 1197–1202, ©2005 (182).

data was collected (reviewed here (26, 192)) and, along with molecular molecular dynamics experiments, led to a more refined transport mechanism (193) demonstrating that accessibility of the ion binding pocket changes with the protonation state of Asp163 and Asp164 (Figure 1.6-A) leading to a putative transport cycle (Figure 1.6-B).

As the crystal structure likely represents an acid-locked state (the crystals were grown at pH 4), data describing the physiological conformations of NhaA have been obtained by closely examining transporter function at physiological pH (194–199). Applying this knowledge along with electrophysiology experiments resulted in a refined kinetic model of the Na^+/H^+ exchange cycle (200) (Figure 1.6-B). It indicates that acidic pH_i slows the binding of intracellular Na^+ (C_iNa) and likewise an alkaline periplasmic pH decreases H^+ binding, thereby decreasing the speed tat which the H^+-loaded conformation returns to the cytoplasmic side. Despite this progress, challenges in producing quality x-ray diffraction data for NhaA crystals at alkaline pH has, to date, prevented the elucidation of a high resolution structure of the periplasmic-facing antiporter (unpublished communication from E. Padan). However, a computational model of the periplasmic-facing NhaA conformation has been published based on all the previous observations (201). To summarize this large body of work, NhaA is a 388 residue electrogenic antiporter that transports $1Na^+$ out of the cell in exchange for $2H^+$ acting as a mechanism of salt tolerance and pH regulation (189). Its activity is modulated by pH but it lacks an additional regulatory domain such as that found in NHE1 (75, 202, 203). The transmembrane domain has twelve transmembrane segments with pseudo two-fold symmetry in the so-called transport bundle between segments III-V and X-XII, with IV and XI being partially unwound and crossed in the membrane exposing the ion binding residues on segment V (182). Conformational changes during transport have been measured in segments II, VI, VIII indicating that helices outside of the transport bundle are directly implicated in antiport (194, 196, 197).

Conversely, NHE1, which shares only 18% sequence identity with NhaA, is predicted to have the same number of transmembrane helices while the membrane domain is 30% larger. It has a large regulatory domain, is electroneutral, and works primarily as a H^+ efflux pump. NHE1 and NhaA have evolved from different ancestral proteins with NHE1 being one of the most recently evolved Na^+/H^+ exchangers and NhaA much older (27). However, it is thought that ancient gene duplication and fusion events followed by divergent evolution resulted in the commonly observed pseudo two-fold symmetry that is a general feature of secondary active transporters

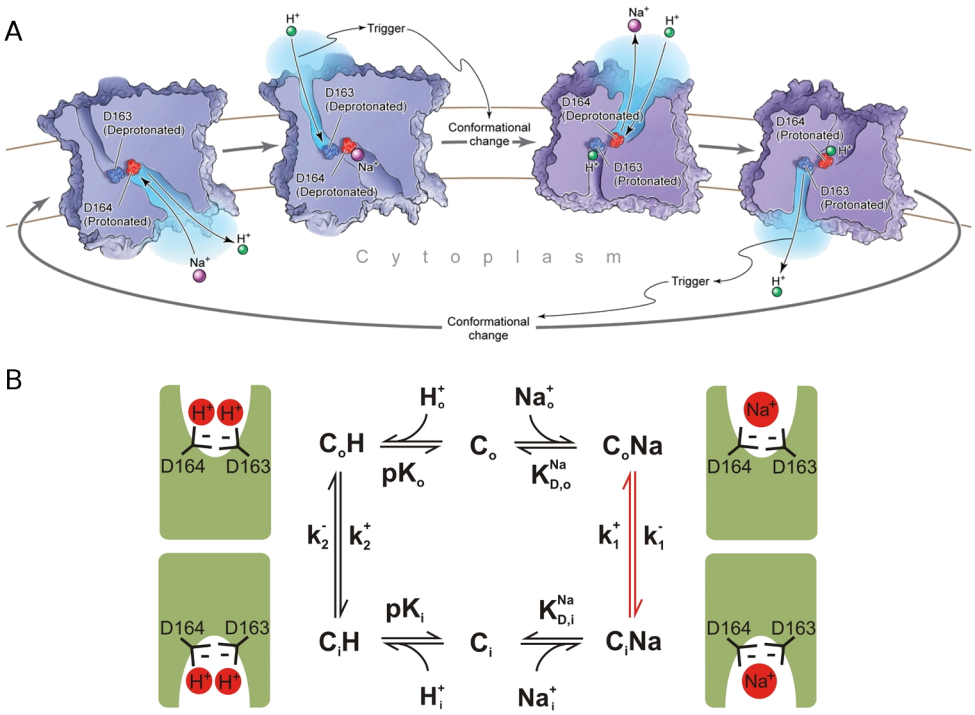

Figure 1.6: Mechanism of Na^+/H^+ antiport in *E. coli* NhaA. ***A***, Transport cycle of NhaA based on the crystal structure and molecular dynamics. The ion binding residues Asp163 and Asp164 and their respective protonation states are labelled. The timing of Na^+ binding and deprotonation of Asp163 or Asp164, triggering a conformational change (trigger) during transport are also labelled. From Arkin *et al. Science* (2007) **317**(5839), 799–803 (193). Reprinted with permission from AAAS. ***B***, A kinetic model of the NhaA transport cycle. The dissociation constants for Na^+ (intracellular, $K^{Na}_{D,i}$; periplasmic $K^{Na}_{D,o}$) and H^+ (intracellular, pK_i; periplasmic, pK_o) for the *outward*-facing (C_o) and *inward*-facing (C_i) NhaA conformations along with the translocation rate constants for Na^+ (forward, k^+_1; reverse, k^-_1) and H^+ (forward, k^+_2; reverse, k^-_2) are shown. The red rate arrows (k_1) represent the net charge difference implicated in returning one Na^+ for $2H^+$. This figure was originally published in The Journal of Biological Chemistry. Thomas Mager *et al. J. Biol. Chem.* 2011; 286:23570–23581. © the American Society for Biochemistry and Molecular Biology.

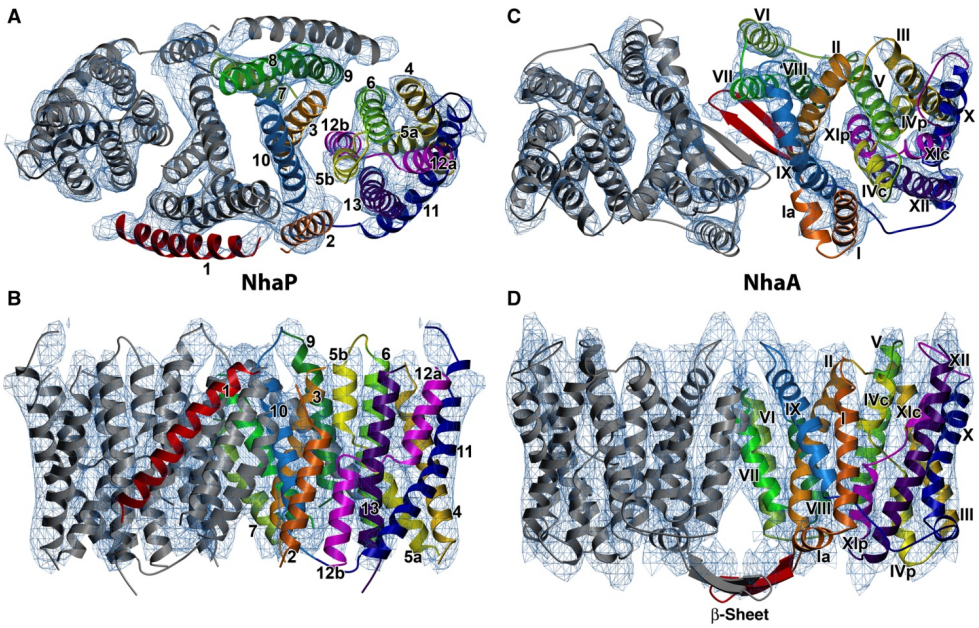

Figure 1.7: Structural model of NhaP1 compared to NhaA. This figure was originally published by Goswami *et al.* (209) and is reprinted with permission from Macmillan Publishers Ltd: *The EMBO Journal* (2011) **30**, 439–449, ©2010. ***A***, View of the electron density (blue mesh) and structural model of *Methanococcus jannaschi* NhaP1 from the cytoplasm. ***B***, Side view of NhaP1 (cytoplasmic side at the top). ***C***, View of electron density (191) and crystal structure (182) of *Escherichia coli* NhaA from the cytoplasm. ***D***, Side view of NhaA (cytoplasmic side at the top).

(23, 204–206). Therefore, although NHE1 and NhaA do not share a common ancestor, they may share a common fold via an ancient duplication event, though this may not become evident until a high resolution structure is solved (30). In fact it was initially expected that a number of unique secondary active transporter folds would account for the variety of substrates transported by these proteins; however, to date most secondary active transporters can be classed into only a few shared folds (30, 206, 207). Indeed the recent elucidation of the bile acid symporter ASBT indicated a structural architecture that was very similar to NhaA despite being from a different TCDB superfamily (208).

Additional evidence supporting a common CPA fold comes from NhaP1, a Na^+/H^+ exchanger from the archea *Methanococcus jannaschii*. This transporter belongs to the CPA1 family and, despite being from a much simpler organism, shares more sequence identity with NHE1 (20%) than *E. coli* NhaA (16%). A two-dimensional projection map of NhaP1 solved to 8 Å resolution indicates an NhaA-like arrangement of helical segments with an additional thirteenth transmembrane segment at the amino-terminus (210). A three dimensional structure at 7 Å reso-

lution further indicates an NhaA-like fold with relatively subtle changes to the overall molecular shape (209) (Figure 1.7). Like NHE1, NhaP1 is an electroneutral exchanger that is activated at acidic pH (211). With both phylogeny and function in common it is very likely that NHE1 will also have a NhaP1-like, and therefore a NhaA-like, fold. This justifies using the NhaA crystal structure as a basis for three-dimensional homology modelling (181). More comparison of the NhaA and NhaP1 structures with NHE1 is included in *Chapter 2*. The strengthening nexuses of biochemistry, structural biology, and bioinformatics may offer insight into the origins of structural motifs allowing a better correlation of structure and function. This has possibilities in both predicting the structure or function of uncharacterized genes and aiding protein engineering in the creation of *de novo* protein designs.

Introducing the models

As stated above, the model (Figure 1.8-A) by Wakabayashi *et al.* was created by testing a NHE1 hydropathy model using mutagenesis (173). The membrane domain of NHE1 is more than a simple hydrophobic anchor. It contains several polar and charged residues that facilitate ion coordination and the conformational changes that are carried out during the transport cycle. Therefore, a simple hydropathy analysis that calculates the likelihood of a sequence being a transmembrane segment based on the hydrophobicity of amino acids within a fixed-length window, may not be sufficient to accurately predict membrane topology. While more sophisticated methods for predicting membrane protein topology informed by empirically measured data are improving the accuracy of these predictions (5, 212), any model should be scrutinized by additional experiments. Generally putative CPA1 transporters found during analysis of sequenced genomes are predicted to have ten or twelve transmembrane segments with both termini in the cytoplasm (213). However NhaP1 has been shown to have thirteen transmembrane segments (209) and Nhx1 from *Saccharomyces cerevisiae* also has an odd number of segments resulting in the carboxy-terminus being found on the extracellular side of the membrane (214). Nonetheless, a hydropathy prediction of twelve transmembrane segments is reasonable. Functional evidence at the time demonstrated that the carboxy-terminal regulatory domain must be located in the cytosol (87, 88, 90, 215); with an even number of putative transmembrane segments this placed the amino-terminus in the cytosol as well. After the publication of this model, this localization was confirmed as immunoreactivity of carboxy-terminal tags required the permeabilization of cells (216). To refine this rough

topology Wakabayahi *et al.* tested the accessibility of various residues spread over the predicted twelve transmembrane segments and interspersed extramembrane loops using SCAM (217). Beginning with a functional *cysteine-less* NHE1, whereby all native cysteines had been mutated to alanine, individual cysteine residues were introduced at strategic positions and cells expressing the mutant transporter were treated with the sulfhydryl-reactive reagents biotin maleimide or MTSET ([2-(Trimethylammonium)Ethyl]Methanethiosulfonate). This reaction was repeated with permeabilized cells. Introduced cysteines that were labelled in intact cells were deemed accessible from the extracellular surface of the membrane, while additional residues labelled in permeabilized cells were accessible from the cytosol. If a residue remained unlabelled it was deemed inaccessible and therefore located within a transmembrane segment. The results are shown in Figure 1.8-A. This represents the first detailed analysis of NHE1 topology. A unique feature of the Wakabayashi model is the inclusion of a large intramembrane loop between transmembrane segments IX and X. While this loop is long enough to be a transmembrane segment it would have created topological inconsistencies and was therefore proposed to reside in the membrane possibly playing a functional role. This hypothesis was supported by an independent group using an *in vitro* cell-free expression and glycosylation technique (218, 219). Confusingly, two intracellular loops, amino acids 176–189 and amino acids 320–330, had contiguous residues that were inside-accessible and outside-accessible. To explain this phenomenon the authors reasoned that the loops could be possibly dipping into the membrane lining part of the transport pore and so during exchange would be accessible to both sides.

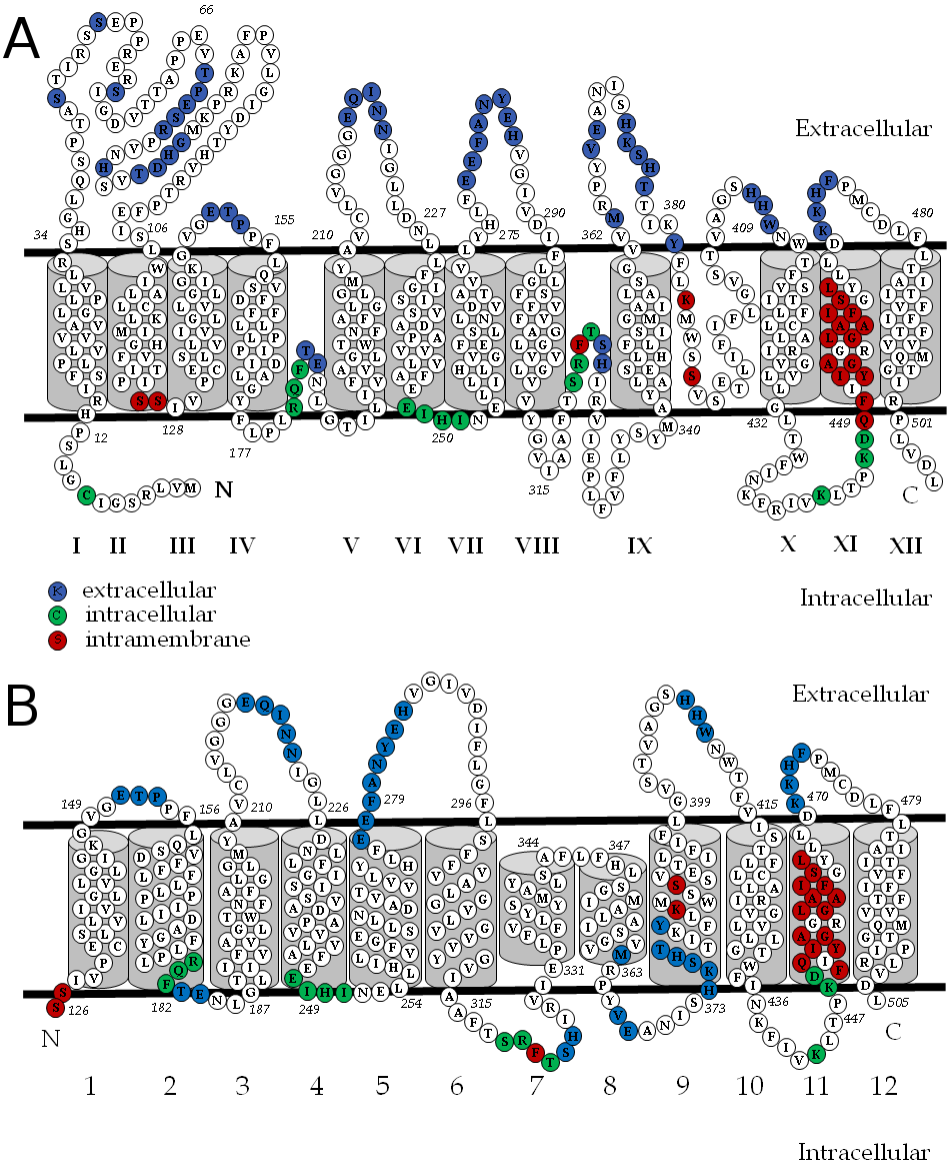

Figure 1.8: The Wakabayshi and Landau topology models of NHE1. ***A***, Topology model by Wakabayashi *et al.* (173). All amino acids in the putative membrane domain are labelled with single letter code, the membrane boundaries are marked with black lines, and transmembrane segments are highlighted by grey cylinders. Residue positions where native and introduced cysteines where subjected to the Substituted Cysteine Accessibility Method (SCAM) are coloured as follows: blue, accessible on the extracellular side; green, accessible on the intracellular side; red, inaccessible (see experimental details in the *Chapter* body). ***B***, Topology model by Landau *et al.* (181), labelled as in ***A***.

While the Wakabayashi model provided a useful tool to guide mutagenesis studies, seven years later another group, Landau *et al.*, proposed an alternative topology (181) (Figure 1.8-B). As introduced above, this model used a computational *in silico* approach beginning with the NhaA crystal structure and combining bioinformatics and fold alignment algorithms to devise a hypothetical NHE1 structure. They analyzed their model against other transmembrane segment predictions and the known phylogenetic, biochemical, and biophysical data. Using this approach their model also indicates twelve transmembrane segments with both termini in the cytosol. However, the first 125 amino acids were not included and the topological arrangement of amino acids 330–410 is radically different. The possible conclusions and insights that these two conflicting models present will be compared in more detail in *Chapter 6*.

Structural biology of mammalian membrane proteins

These models of NHE1 are an important tool that can be used to take a more directed approach to further biochemical and biophysical characterization. The development of efficient molecular biology techniques has made protein mutagenesis a relatively quick method to gather a large amount of functional and structural information. Furthermore, genome sequencing projects have allowed the identification of conserved motifs that are good candidates for mutagenesis studies. In combination with bioinformatics, a good topology model can identify polar and charged residues within transmembrane helices which likely play a structural or functional role (220). Of course, while obtaining high resolution structures of secondary active transporters is crucial to our understanding of translocation mechanisms, and structural techniques continue to improve, relatively few unique membrane protein structures have been elucidated. Of these structures, few have been sourced from mammals and only a handful from humans. Most of these structures are of prokaryotic proteins, many of which do not have an identifiable eukaryotic homologue or the sequences are too divergent to be able to properly identify conserved motifs. In addition, motifs between transporters that have a shared fold but have little sequence similarity may be too difficult to identify. Secondary structure and topology modelling provides another tool that can help researchers find interesting residues or possibly related motifs that are not found by automated sequence alignment programs. An example of this is provided in *Chapter 4* using NhaA and *Schiziosaccharomyces pombe* Na^+/H^+ exchanger sod2, where an improved sequence alignment

was created using secondary structure and topology predictions that facilitated the creation of a reasonable homology model of sod2 using the crystal structure of NhaA (221).

Thesis Overview

The goal of my PhD studies was to deepen our understanding of Na^+/H^+ exchanger function by increasing our structural knowledge. I have made several steps towards this goal. *Chapter 2* discusses our development of a NHE1 expression system in yeast, the difficulties encountered with optimizing and improving the protein yield and purity, and the structural and functional data we were able to collect. *Chapter 3* discusses our implementation and optimization of the bacterial maltose binding protein fusion expression system to study single and multiple transmembrane segments of membrane proteins. *Chapter 4* describes the successful application of the techniques in the previous chapter to a single transmembrane segment of the yeast Na^+/H^+ exchanger sod2 as well as structure determination and homology modelling of the transporter. *Chapter 5* describes recent results in the expression, purification, and initial structural characterization of a contiguous three transmembrane segment of human NHE1 using the techniques described in the previous two chapters. The ideas and conclusions from this work are then combined with the other published data in a final model of NHE1 structure and function in *Chapter 6*.

References

1. Kemp, G., Young, H. S. & Fliegel, L. (2008). Structure and function of the human Na^+/H^+ exchanger isoform 1. *Channels (Austin)* **2**, 329–336.

2. Rinke, C. *et al.* (2013). Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **499**, 431–437.

3. Karmazyn, M., Avkiran, M. & Fliegel, L. *The sodium-hydrogen exchanger: from molecule to its role in disease* (Kluwer Academic Publishers, 2003).

4. Casey, J. R., Grinstein, S. & Orlowski, J. (2010). Sensors and regulators of intracellular pH. *Nat Rev Mol Cell Biol* **11**, 50–61.

5. Bernsel, A. *et al.* (2008). Prediction of membrane-protein topology from first principles. *Proc Natl Acad Sci U S A* **105**, 7177–7181.

6. Jiang, L. *et al.* (2008). De novo computational design of retro-aldol enzymes. *Science* **319**, 1387–1391.

7. Röthlisberger, D. *et al.* (2008). Kemp elimination catalysts by computational enzyme design. *Nature* **453**, 190–195.

8. Siegel, J. B. *et al.* (2010). Computational design of an enzyme catalyst for a stereoselective bimolecular diels-alder reaction. *Science* **329**, 309–313.

9. Giger, L. *et al.* (2013). Evolution of a designed retro-aldolase leads to complete active site remodeling. *Nat Chem Biol* **9**, 494–498.

10. Vinson, V. J. (2009). Proteins in motion. *Science* **324**, 197–197.

11. Markin, C. J., Xiao, W. & Spyracopoulos, L. (2010). Mechanism for recognition of polyubiquitin chains: balancing affinity through interplay between multivalent binding and dynamics. *J Am Chem Soc* **132**, 11247–11258.

12. Bahar, I., Lezon, T. R., Yang, L.-W. & Eyal, E. (2010). Global dynamics of proteins: bridging between structure and function. *Annu Rev Biophys* **39**, 23–42.

13. García-Moreno, B. E. *et al.* (1997). Experimental measurement of the effective dielectric in the hydrophobic core of a protein. *Biophys Chem* **64**, 211–224.

14. Isom, D. G., Castaeda, C. A., Cannon, B. R. & Garca-Moreno, B. (2011). Large shifts in pKa values of lysine residues buried inside a protein. *Proc Natl Acad Sci U S A* **108**, 5260–5265.

15. Roos, A. & Boron, W. F. (1981). Intracellular pH. *Physiol Rev* **61**, 296–434.

16. Boron, W. F. *Cellular Buffering and Intracellular pH*, chap. 2, 33–56. The Regulation of Acid-Base Balance (Raven Press, New York, 1989).

17. Sun, B., Leem, C. H. & Vaughan-Jones, R. D. (1996). Novel chloride-dependent acid loader in the guinea-pig ventricular myocyte: part of a dual acid-loading mechanism. *J Physiol* **495 (Pt 1)**, 65–82.

18. Davies, R. E. & Krebs, H. A. (1952). Biochemical aspects of the transport of ions by nervous tissue. *Biochem J* **50**, xxiii–xxv.

19. Mitchell, P. & Moyle, J. (1967). Respiration-driven proton translocation in rat liver mitochondria. *Biochem J* **105**, 1147–1162.

20. Murer, H., Hopfer, U. & Kinne, R. (1976). Sodium/proton antiport in brush-border-membrane vesicles isolated from rat small intestine and kidney. *J Am Soc Nephrol* **9**, 143–150.

21. Aickin, C. C. & Thomas, R. C. (1977). An investigation of the ionic mechanism of intracellular pH regulation in mouse soleus muscle fibres. *J Physiol* **273**, 295–316.

22. Donowitz, M., Ming Tse, C. & Fuster, D. (2013). SLC9/NHE gene family, a plasma membrane and organellar family of Na/H exchangers. *Mol Aspects Med* **34**, 236–251.

23. Chang, A. B., Lin, R., Studley, W. K., Tran, C. V. & Saier, M. H., Jr (2004). Phylogeny as a guide to structure and function of membrane transport proteins (review). *Mol Membr Biol* **21**, 171–181.

24. Saier, J., Milton H, Tran, C. V. & Barabote, R. D. (2006). TCDB: the transporter classification database for membrane transport protein analyses and information. *Nucleic Acids Res* **34**, D181–186.

25. Saier, J., Milton H, Yen, M. R., Noto, K., Tamang, D. G. & Elkan, C. (2009). The transporter classification database: recent advances. *Nucleic Acids Res* **37**, D274–278.

26. Padan, E., Kozachkov, L., Herz, K. & Rimon, A. (2009). NhaA crystal structure: functional-structural insights. *J Exp Biol* **212**, 1593–1603.

27. Brett, C. L., Donowitz, M. & Rao, R. (2005). Evolutionary origins of eukaryotic sodium/proton exchangers. *Am J Physiol Cell Physiol* **288**, C223–239.

28. Hoglund, P. J., Nordstrom, K. J., Schioth, H. B. & Fredriksson, R. (2011). The solute carrier families have a remarkably long evolutionary history with the majority of the human families present before divergence of bilaterian species. *Mol Biol Evol* **28**, 1531–1541.

29. Dibrov, P. & Fliegel, L. (1998). Comparative molecular analysis of Na^+/H^+ exchangers: a unified model for Na^+/H^+ antiport? *FEBS letters* **424**, 1–5.

30. Shi, Y. (2013). Common folds and transport mechanisms of secondary active transporters. *Annu Rev Biophys* **42**, 51–72.

31. Borgese, F., Malapert, M., Fievet, B., Pouyssegur, J. & Motais, R. (1994). The cytoplasmic domain of the Na^+/H^+ exchangers (NHEs) dictates the nature of the hormonal response: behavior of a chimeric human NHE1/trout beta NHE antiporter. *Proc Natl Acad Sci U S A* **91**, 5431–5435.

32. Wakabayashi, S. *et al.* (1995). Cytoplasmic domain of the ubiquitous Na^+/H^+ exchanger NHE1 can confer Ca^{2+} responsiveness to the apical isoform NHE3. *J Biol Chem* **270**, 26460–26465.

33. Kurashima, K. *et al.* (1999). The apical Na^+/H^+ exchanger isoform NHE3 is regulated by the actin cytoskeleton. *J Biol Chem* **274**, 29843–29849.

34. Orlowski, J. & Grinstein, S. (1997). Na^+/H^+ exchangers of mammalian cells. *J Biol Chem* **272**, 22373–22376.

35. Schultheis, P. J. *et al.* (1998). Renal and intestinal absorptive defects in mice lacking the NHE3 Na^+/H^+ exchanger. *Nat Genet* **19**, 282–285.

36. Park, K. *et al.* (2001). Defective fluid secretion and NaCl absorption in the parotid glands of Na^+/H^+ exchanger-deficient mice. *J Biol Chem* **276**, 27042–27050.

37. Ledoussal, C. *et al.* (2001). Renal salt wasting in mice lacking NHE3 Na^+/H^+ exchanger but not in mice lacking NHE2. *Am J Physiol Renal Physiol* **281**, F718–727.

38. Maher, M. M., Gontarek, J. D., Bess, R. S., Donowitz, M. & Yeo, C. J. (1997). The Na^+/H^+ exchange isoform NHE3 regulates basal canine ileal na+ absorption in vivo. *Gastroenterology* **112**, 174–183.

39. Yeo, C. J., Barry, K., Gontarek, J. D. & Donowitz, M. (1994). Na^+/H^+ exchange mediates meal-stimulated ileal absorption. *Surgery* **116**, 388–394; discussion394–395.

40. Schultheis, P. J. *et al.* (1998). Targeted disruption of the murine Na^+/H^+ exchanger isoform 2 gene causes reduced viability of gastric parietal cells and loss of net acid secretion. *J Clin Invest* **101**, 1243–1253.

41. Preisig, P. A. & Rector, J., F C (1988). Role of na+-h+ antiport in rat proximal tubule NaCl absorption. *Am J Physiol* **255**, F461–465.

42. Preisig, P. A., Ives, H. E., Cragoe, J., E J, Alpern, R. J. & Rector, J., F C (1987). Role of the Na^+/H^+ antiporter in rat proximal tubule bicarbonate absorption. *J Clin Invest* **80**, 970–978.

43. Good, D. W. & Watts, 3rd, B. A. (1996). Functional roles of apical membrane Na^+/H^+ exchange in rat medullary thick ascending limb. *Am J Physiol* **270**, F691–699.

44. Wang, T., Hropot, M., Aronson, P. S. & Giebisch, G. (2001). Role of NHE isoforms in mediating bicarbonate reabsorption along the nephron. *Am J Physiol Renal Physiol* **281**, F1117–1122.

45. Arena, E. A. *et al.* (2012). Functional role of NHE4 as a pH regulator in rat and human colonic crypts. *Am J Physiol Cell Physiol* **302**, C412–C418.

46. Pizzonia, J. H. *et al.* (1998). Immunochemical characterization of Na^+/H^+ exchanger isoform NHE4. *Am J Physiol* **275**, F510–517.

47. Baird, N. R. *et al.* (1999). Molecular cloning, genomic organization, and functional expression of Na^+/H^+ exchanger isoform 5 (NHE5) from human brain. *J Biol Chem* **274**, 4377–4382.

48. Attaphitaya, S., Park, K. & Melvin, J. E. (1999). Molecular cloning and functional expression of a rat Na^+/H^+ exchanger (NHE5) highly expressed in brain. *J Biol Chem* **274**, 4383–4388.

49. Szabó, E. Z., Numata, M., Lukashova, V., Iannuzzi, P. & Orlowski, J. (2005). beta-arrestins bind and decrease cell-surface abundance of the Na^+/H^+ exchanger NHE5 isoform. *Proc Natl Acad Sci U S A* **102**, 2790–2795.

50. Onishi, I., Lin, P. J. C., Diering, G. H., Williams, W. P. & Numata, M. (2007). RACK1 associates with NHE5 in focal adhesions and positively regulates the transporter activity. *Cell Signal* **19**, 194–203.

51. Diering, G. H., Mills, F., Bamji, S. X. & Numata, M. (2011). Regulation of dendritic spine growth through activity-dependent recruitment of the brain-enriched Na/H exchanger NHE5. *Mol Biol Cell* **22**, 2246–2257.

52. Lukashova, V. *et al.* (2013). The Na^+/H^+ exchanger NHE5 is sorted to discrete intracellular vesicles in the central and peripheral nervous systems. *Adv Exp Med Biol* **961**, 397–410.

53. Nass, R., Cunningham, K. W. & Rao, R. (1997). Intracellular sequestration of sodium by a novel Na^+/H^+ exchanger in yeast is enhanced by mutations in the plasma membrane H^+-ATPase: Insights into mechanisms of sodium tolerance. *J Biol Chem* **272**, 26145–26152.

54. Numata, M., Petrecca, K., Lake, N. & Orlowski, J. (1998). Identification of a mitochondrial Na^+/H^+ exchanger. *J Biol Chem* **273**, 6951–6959.

55. Brett, C. L., Wei, Y., Donowitz, M. & Rao, R. (2002). Human Na^+/H^+ exchanger isoform 6 is found in recycling endosomes of cells, not in mitochondria. *Am J Physiol* **282**, C1031–1041.

56. Mellman, I. (1992). The importance of being acid: the role of acidification in intracellular membrane traffic. *J Exp Biol* **172**, 39–45.

57. Futai, M. *et al.* (2000). Luminal acidification of diverse organelles by v-ATPase in animal cells. *J Exp Biol* **203**, 107–116.

58. Nakamura, N., Tanaka, S., Teko, Y., Mitsui, K. & Kanazawa, H. (2005). Four Na^+/H^+ exchanger isoforms are distributed to golgi and post-golgi compartments and are involved in organelle pH regulation. *J Biol Chem* **280**, 1561–1572.

59. Smith, A. J. & Lippiat, J. D. (2010). Direct endosomal acidification by the outwardly rectifying CLC-5 Cl^-/H^+ exchanger. *J Physiol* **588**, 2033–2045.

60. Numata, M. & Orlowski, J. (2001). Molecular cloning and characterization of a novel $(Na^+,K^+)/H^+$ exchanger localized to the trans-golgi network. *J Biol Chem* **276**, 17387–17394.

61. Lawrence, S. P., Bright, N. A., Luzio, J. P. & Bowers, K. (2010). The Sodium/Proton exchanger NHE8 regulates late endosomal morphology and function. *Mol Biol Cell* **21**, 3540–3551.

62. Kagami, T. *et al.* (2008). Identification and biochemical characterization of the SLC9A7 interactome. *Mol Membr Biol* **25**, 436–447.

63. Lin, P. J. C., Williams, W. P., Kobiljski, J. & Numata, M. (2007). Caveolins bind to $Na^+(K^+)/H^+$ exchanger NHE7 by a novel binding module. *Cell Signal* **19**, 978–988.

64. Ohgaki, R., Fukura, N., Matsushita, M., Mitsui, K. & Kanazawa, H. (2008). Cell surface levels of organellar Na^+/H^+ exchanger isoform 6 are regulated by interaction with RACK1. *J Biol Chem* **283**, 4417–4429.

65. Haas, M., Schooler, J. & Tosteson, D. C. (1975). Coupling of lithium to sodium transport in human red cells. *Nature* **258**, 425–427.

66. Canessa, M., Adragna, N., Solomon, H. S., Connolly, T. M. & Tosteson, D. C. (1980). Increased sodium-lithium countertransport in red cells of patients with essential hypertension. *N Engl J Med* **302**, 772–776.

67. Zerbini, G., Podesta, F., Meregalli, G., Deferrari, G. & Pontremoli, R. (2001). Fibroblast Na^+/Li^+ countertransport rate is elevated in essential hypertension. *J Hypertens* **19**, 1263–1269.

68. Busch, S., Burckhardt, B. C. & Siffert, W. (1995). Expression of the human sodium/proton exchanger NHE-1 in *Xenopus laevis* oocytes enhances sodium/proton exchange activity and establishes sodium/lithium countertransport. *Pflügers Arch* **429**, 859–869.

69. Kahn, A. M. (1987). Difference between human red blood cell Na^+/Li^+ countertransport and renal Na^+/H^+ exchange. *Hypertension* **9**, 7–12.

70. Zerbini, G., Maestroni, A., Breviario, D., Mangili, R. & Casari, G. (2003). Alternative splicing of NHE-1 mediates Na-Li countertransport and associates with activity rate. *Diabetes* **52**, 1511–1518.

71. Xiang, M., Feng, M., Muend, S. & Rao, R. (2007). A human Na^+/H^+ antiporter sharing evolutionary origins with bacterial NhaA may be a candidate gene for essential hypertension. *Proc Natl Acad Sci U S A* **104**, 18677–18681.

72. Fuster, D. G. *et al.* (2008). Characterization of the Sodium/Hydrogen exchanger NHA2. *Journal of the American Society of Nephrology* **19**, 1547–1556.

73. Kondapalli, K. C., Kallay, L. M., Muszelik, M. & Rao, R. (2012). Unconventional chemiosmotic coupling of NHA2, a mammalian Na^+/H^+ antiporter, to a plasma membrane H^+ gradient. *J Biol Chem* **287**, 36239–36250.

74. Wakabayashi, S., Hisamitsu, T., Pang, T. & Shigekawa, M. (2003). Kinetic dissection of two distinct proton binding sites in Na^+/H^+ exchangers by measurement of reverse mode reaction. *J Biol Chem* **278**, 43580–43585.

75. Krulwich, T. A., Sachs, G. & Padan, E. (2011). Molecular aspects of bacterial pH sensing and homeostasis. *Nat Rev Microbiol* **9**, 330–343.

76. Denker, S. P., Huang, D. C., Orlowski, J., Furthmayr, H. & Barber, D. L. (2000). Direct binding of the Na-H exchanger NHE1 to ERM proteins regulates the cortical cytoskeleton and cell shape independently of H^+ translocation. *Mol Cell* **6**, 1425–1436.

77. Fliegel, L., Sardet, C., Pouyssegur, J. & Barr, A. (1991). Identification of the protein and cDNA of the cardiac Na^+/H^+ exchanger. *FEBS letters* **279**, 25–29.

78. Orlowski, J. & Grinstein, S. (2004). Diversity of the mammalian sodium/proton exchanger SLC9 gene family. *Pflügers Arch* **447**, 549–565.

79. Petrecca, K., Atanasiu, R., Grinstein, S., Orlowski, J. & Shrier, A. (1999). Subcellular localization of the Na^+/H^+ exchanger NHE1 in rat myocardium. *Am J Physiol* **276**, H709–717.

80. Fliegel, L. (2009). Regulation of the Na^+/H^+ exchanger in the healthy and diseased myocardium. *Expert Opin Ther Targets* **13**, 55–68.

81. Bandyopadhyay, S. *et al.* (2010). A human MAP kinase interactome. *Nat Methods* **7**, 801–805.

82. Amith, S. R. & Fliegel, L. (2013). Regulation of the Na^+/H^+ exchanger (NHE1) in breast cancer metastasis. *Cancer research* **73**, 1259–1264.

83. Karmazyn, M., Kilić, A. & Javadov, S. (2008). The role of NHE-1 in myocardial hypertrophy and remodelling. *J Mol Cell Cardiol* **44**, 647–653.

84. Malo, M. E. & Fliegel, L. (2006). Physiological role and regulation of the Na^+/H^+ exchanger. *Can J Physiol Pharmacol* **84**, 1081–1095.

85. Luo, J. & Sun, D. (2007). Physiology and pathophysiology of Na^+/H^+ exchange isoform 1 in the central nervous system. *Curr Neurovasc Res* **4**, 205–215.

86. Bertrand, B., Wakabayashi, S., Ikeda, T., Pouysségur, J. & Shigekawa, M. (1994). The Na^+/H^+ exchanger isoform 1 (NHE1) is a novel member of the calmodulin-binding proteins. identification and characterization of calmodulin-binding sites. *J Biol Chem* **269**, 13703–13709.

87. Wakabayashi, S., Bertrand, B., Shigekawa, M., Fafournoux, P. & Pouysségur, J. (1994). Growth factor activation and H^+-sensing of the Na^+/H^+ exchanger isoform 1 (NHE1). evidence for an additional mechanism not requiring direct phosphorylation. *J Biol Chem* **269**, 5583–5588.

88. Wakabayashi, S., Ikeda, T., Iwamoto, T., Pouysségur, J. & Shigekawa, M. (1997). Calmodulin-binding autoinhibitory domain controls "pH-sensing" in the Na^+/H^+ exchanger NHE1 through sequence-specific interaction. *Biochemistry* **36**, 12854–12861.

89. Eguti, D. M. N., Thieme, K., Leung, G. P., Mello-Aires, M. & Oliveira-Souza, M. (2010). Regulation of Na^+/H^+ exchanger isoform 1 (NHE1) by calmodulin-binding sites: role of angiotensin II. *Cell Physiol Biochem* **26**, 541–552.

90. Lin, X. & Barber, D. L. (1996). A calcineurin homologous protein inhibits GTPase-stimulated Na^+/H^+ exchange. *Proc Natl Acad Sci U S A* **93**, 12631–12636.

91. Pang, T., Su, X., Wakabayashi, S. & Shigekawa, M. (2001). Calcineurin homologous protein as an essential cofactor for Na^+/H^+ exchangers. *J Biol Chem* **276**, 17367–17372.

92. Pang, T., Hisamitsu, T., Mori, H., Shigekawa, M. & Wakabayashi, S. (2004). Role of calcineurin b homologous protein in pH regulation by the Na^+/H^+ exchanger 1: tightly bound Ca2+ ions as important structural elements. *Biochemistry* **43**, 3628–3636.

93. Pang, T., Wakabayashi, S. & Shigekawa, M. (2002). Expression of calcineurin b homologous protein 2 protects serum deprivation-induced cell death by serum-independent activation of Na^+/H^+ exchanger. *J Biol Chem* **277**, 43771–43777.

94. Mailänder, J., Müller-Esterl, W. & Dedio, J. (2001). Human homolog of mouse tescalcin associates with Na^+/H^+ exchanger type-1. *FEBS letters* **507**, 331–335.

95. Li, X., Liu, Y., Kay, C. M., Müller-Esterl, W. & Fliegel, L. (2003). The Na^+/H^+ exchanger cytoplasmic tail: structure, function, and interactions with tescalcin. *Biochemistry* **42**, 74487456.

96. Zaun, H. C., Shrier, A. & Orlowski, J. (2008). Calcineurin b homologous protein 3 promotes the biosynthetic maturation, cell surface stability, and optimal transport of the Na^+/H^+ exchanger NHE1 isoform. *J Biol Chem* **283**, 12456–12467.

97. Zaun, H. C., Shrier, A. & Orlowski, J. (2012). N-myristoylation and Ca_2^+ binding of calcineurin b homologous protein CHP3 are required to enhance Na^+/H^+ exchanger NHE1 half-life and activity at the plasma membrane. *J Biol Chem* **287**, 36883–36895.

98. Li, X., Alvarez, B., Casey, J. R., Reithmeier, R. A. F. & Fliegel, L. (2002). Carbonic anhydrase II binds to and enhances activity of the Na^+/H^+ exchanger. *J Biol Chem* **277**, 36085–36091.

99. Vince, J. W. & Reithmeier, R. A. (2000). Identification of the carbonic anhydrase II binding site in the Cl^-/HCO_3^- anion exchanger AE1. *Biochemistry* **39**, 5527–5533.

100. Li, X., Liu, Y., Alvarez, B. V., Casey, J. R. & Fliegel, L. (2006). A novel carbonic anhydrase II binding site regulates NHE1 activity. *Biochemistry* **45**, 2414–2424.

101. Alvarez, B. V. *et al.* (2007). Carbonic anhydrase inhibition prevents and reverts cardiomyocyte hypertrophy. *J Physiol* **579**, 127–145.

102. Brown, B. F., Quon, A., Dyck, J. R. B. & Casey, J. R. (2012). Carbonic anhydrase II promotes cardiomyocyte hypertrophy. *Can J Physiol Pharmacol* **90**, 1599–1610.

103. Denker, S. P. & Barber, D. L. (2002). Cell migration requires both ion translocation and cytoskeletal anchoring by the Na-H exchanger NHE1. *J Cell Biol* **159**, 1087–1096.

104. Meima, M. E., Webb, B. A., Witkowska, H. E. & Barber, D. L. (2009). The sodium-hydrogen exchanger NHE1 is an Akt substrate necessary for actin filament reorganization by growth factors. *J Biol Chem* **284**, 26666–26675.

105. Nunomura, W., Denker, S., Barber, D., Takakuwa, Y. & Gascard, P. (2012). Characterization of cytoskeletal protein 4.1R interaction with NHE1 Na^+H^+ exchanger isoform 1. *Biochem J* **446**, 427–435.

106. Jung, Y.-S. *et al.* (2008). Physical interactions and functional coupling between Daxx and sodium hydrogen exchanger 1 in ischemic cell death. *J Biol Chem* **283**, 1018–1025.

107. Silva, N. L., Haworth, R. S., Singh, D. & Fliegel, L. (1995). The carboxyl-terminal region of the Na^+/H^+ exchanger interacts with mammalian heat shock protein. *Biochemistry* **34**, 10412–10420.

108. Xue, J. *et al.* (2007). Novel functional interaction between Na^+/H^+ exchanger 1 and tyrosine phosphatase SHP-2. *Am J Physiol* **292**, R2406–2416.

109. Lachapelle, G., Radicioni, S. M., Stankiewicz, A. R. & Mosser, D. D. (2007). Acute acidification or amiloride treatment suppresses the ability of Hsp70 to inhibit heat-induced apoptosis. *Apoptosis* **12**, 1479–1488.

110. Nylandsted, J., Jttel, M., Hoffmann, E. K. & Pedersen, S. F. (2004). Heat shock protein 70 inhibits shrinkage-induced programmed cell death via mechanisms independent of effects on cell volume-regulatory membrane transport proteins. *Pflügers Arch* **449**, 175–185.

111. Huang, C., Wang, J., Chen, Z., Wang, Y. & Zhang, W. (2013). 2-phenylethynesulfonamide prevents induction of pro-inflammatory factors and attenuates LPS-induced liver injury by targeting NHE1-Hsp70 complex in mice. *PloS one* **8**, e67582.

112. Bae, J.-S. *et al.* (2013). Chaperone stress 70 protein (STCH) binds and regulates two acid/base transporters NBCe1-B and NHE1. *J Biol Chem* **288**, 6295–6305.

113. Lehoux, S., Ji, A., Florian, J. A. & Berk, B. C. (2001). 14–3–3 binding to Na^+/H^+ exchanger isoform-1 is associated with serum-dependent activation of Na^+/H^+ exchange. *J Biol Chem* **276**, 15794–15800.

114. Hunte, C. & Richers, S. (2008). Lipids and membrane protein structures. *Curr Opin Struct Biol* **18**, 406–411.

115. Chini, B. & Parenti, M. (2009). G-protein-coupled receptors, cholesterol and palmitoylation: facts about fats. *J Mol Endocrinol* **42**, 371–379.

116. Gustavsson, M., Traaseth, N. J. & Veglia, G. (2011). Activating and deactivating roles of lipid bilayers on the Ca2+-ATPase/Phospholamban complex. *Biochemistry* **50**, 10367–10374.

117. Hilgemann, D. W., Feng, S. & Nasuhoglu, C. (2001). The complex and intriguing lives of PIP2 with ion channels and transporters. *Science's STKE: Signal Transduction Knowledge Environment* **2001**, re19.

118. Aharonovitz, O. *et al.* (2000). Intracellular pH regulation by Na^+/H^+ exchange requires phosphatidylinositol-4,5-bisphosphate. *J Cell Biol* **150**, 213–224.

119. Wakabayashi, S., Nakamura, T. Y., Kobayashi, S. & Hisamitsu, T. (2010). Novel phorbol ester-binding motif mediates hormonal activation of Na^+/H^+ exchanger. *J Biol Chem* **285**, 26652–26661.

120. Malo, M. E., Li, L. & Fliegel, L. (2007). Mitogen-activated protein kinase-dependent activation of the Na^+/H^+ exchanger is mediated through phosphorylation of amino acids Ser770 and Ser771. *J Biol Chem* **282**, 6292–6299.

121. Khaled, A. R. *et al.* (2001). Trophic factor withdrawal: p38 mitogen-activated protein kinase activates NHE1, which induces intracellular alkalinization. *Mol Cell Biol* **21**, 7545–7557.

122. Grenier, A. L. *et al.* (2008). Apoptosis-induced alkalinization by the Na^+/H^+ exchanger isoform 1 is mediated through phosphorylation of amino acids Ser726 and Ser729. *Am J Physiol* **295**, C883–896.

123. Coccaro, E., Karki, P., Cojocaru, C. & Fliegel, L. (2009). Phenylephrine and sustained acidosis activate the neonatal rat cardiomyocyte Na^+/H^+ exchanger through phosphorylation of amino acids Ser770 and Ser771. *Am J Physiol* **297**, H846–858.

124. Takahashi, E. *et al.* (1999). p90(RSK) is a serum-stimulated Na^+/H^+ exchanger isoform-1 kinase. regulatory phosphorylation of serine 703 of Na^+/H^+ exchanger isoform-1. *J Biol Chem* **274**, 20206–20214.

125. Maekawa, N. *et al.* (2006). Inhibiting p90 ribosomal s6 kinase prevents Na^+/H^+ exchanger-mediated cardiac ischemia-reperfusion injury. *Circulation* **113**, 2516–2523.

126. Karki, P., Coccaro, E. & Fliegel, L. (2010). Sustained intracellular acidosis activates the myocardial Na^+/H^+ exchanger independent of amino acid Ser703 and p90(RSK). *Biochim Biophys Acta* **1798**, 1565–1576.

127. Snabaitis, A. K., Cuello, F. & Avkiran, M. (2008). Protein kinase B/Akt phosphorylates and inhibits the cardiac Na^+/H^+ exchanger NHE1. *Circ Res* **103**, 881–890.

128. Yan, W., Nehrke, K., Choi, J. & Barber, D. L. (2001). The Nck-interacting kinase (NIK) phosphorylates the Na^+/H^+ exchanger NHE1 and regulates NHE1 activation by platelet-derived growth factor. *J Biol Chem* **276**, 31349–31356.

129. Fliegel, L., Walsh, M. P., Singh, D., Wong, C. & Barr, A. (1992). Phosphorylation of the C-terminal domain of the Na^+/H^+ exchanger by Ca^{2+}/calmodulin-dependent protein kinase II. *Biochem J* **282**, 139–145.

130. Pedersen, S. F., Darborg, B. V., Rasmussen, M., Nylandsted, J. & Hoffmann, E. K. (2007). The Na^+/H^+ exchanger, NHE1, differentially regulates mitogen-activated protein kinase subfamilies after osmotic shrinkage in ehrlich lettre ascites cells. *Cell Physiol Biochem* **20**, 735–750.

131. Tominaga, T., Ishizaki, T., Narumiya, S. & Barber, D. L. (1998). p160ROCK mediates RhoA activation of Na-H exchange. *EMBO J* **17**, 4712–4722.

132. Cardone, R. A. *et al.* (2005). Protein kinase a gating of a pseudopodial-located RhoA/ROCK/p38/NHE1 signal module regulates invasion in breast cancer cell lines. *Mol Biol Cell* **16**, 3117–3127.

133. Maly, K. *et al.* (2002). Critical role of protein kinase C alpha and calcium in growth factor induced activation of the Na^+/H^+ exchanger NHE1. *FEBS letters* **521**, 205–210.

134. Haworth, R. S., Sinnett-Smith, J., Rozengurt, E. & Avkiran, M. (1999). Protein kinase D inhibits plasma membrane Na^+/H^+ exchanger activity. *Am J Physiol* **277**, C1202–1209.

135. Reshkin, S. J. *et al.* (2000). Phosphoinositide-3-kinase is involved in the tumor-specific activation of human breast cancer cell Na^+/H^+ exchange, motility, and invasion induced by serum deprivation. *J Biol Chem* **275**, 5361–5369.

136. Karki, P., Li, X., Schrama, D. & Fliegel, L. (2011). B-raf associates with and activates the NHE1 isoform of the Na^+/H^+ exchanger. *J Biol Chem* **286**, 13096–13105.

137. Misik, A. J., Perreault, K., Holmes, C. F. B. & Fliegel, L. (2005). Protein phosphatase regulation of Na^+/H^+ exchanger isoform I. *Biochemistry* **44**, 5842–5852.

138. Snabaitis, A. K., D'Mello, R., Dashnyam, S. & Avkiran, M. (2006). A novel role for protein phosphatase 2A in receptor-mediated regulation of the cardiac sarcolemmal Na^+/H^+ exchanger NHE1. *J Biol Chem* **281**, 20252–20262.

139. Hisamitsu, T., Nakamura, T. Y. & Wakabayashi, S. (2012). Na^+/H^+ exchanger 1 directly binds to calcineurin A and activates downstream NFAT signaling, leading to cardiomyocyte hypertrophy. *Mol Cell Biol* **32**, 3265–3280.

140. Baumgartner, M., Patel, H. & Barber, D. L. (2004). Na^+/H^+ exchanger NHE1 as plasma membrane scaffold in the assembly of signaling complexes. *Am J Physiol Cell Physiol* **287**, C844–850.

141. Bell, S. M. *et al.* (1999). Targeted disruption of the murine nhe1 locus induces ataxia, growth retardation, and seizures. *Am J Physiol* **276**, C788–795.

142. Cox, G. A. *et al.* (1997). Sodium/Hydrogen exchanger gene defect in slow-wave epilepsy mutant mice. *Cell* **91**, 139–148.

143. Rieder, C. V. & Fliegel, L. (2002). Developmental regulation of Na^+/H^+exchanger expression in fetal and neonatal mice. *Am J Physiol* **283**, H273–H283.

144. Gu, X. Q., Yao, H. & Haddad, G. G. (2001). Increased neuronal excitability and seizures in the Na^+/H^+ exchanger null mutant mouse. *American Journal of Physiology - Cell Physiolog* **281**, C496–C503.

145. Good, D. W., Watts, B. A., George, T., Meyer, J. W. & Shull, G. E. (2004). Transepithelial HCO_3^- absorption is defective in renal thick ascending limbs from Na^+/H^+ exchanger NHE1 null mutant mice. *American Journal of Physiology-Renal Physiology* **287**, F1244–F1249.

146. Kintner, D. B. *et al.* (2004). Increased tolerance to oxygen and glucose deprivation in astrocytes from Na^+/H^+ exchanger isoform 1 null mice. *Am J Physiol* **287**, C12–21.

147. Zhou, D., Xue, J., Gavrialov, O. & Haddad, G. G. (2004). Na^+/H^+ exchanger 1 deficiency alters gene expression in mouse brain. *Physiological Genomics* **18**, 331–339.

148. Cardone, R. A., Casavola, V. & Reshkin, S. J. (2005). The role of disturbed pH dynamics and the Na^+/H^+ exchanger in metastasis. *Nat Rev Cancer* **5**, 786–795.

149. Avkiran, M. & Marber, M. S. (2002). Na^+/H^+ exchange inhibitors for cardioprotective therapy: progress, problems and prospects. *J Am Coll Cardiol* **39**, 747–753.

150. Wang, Y., Meyer, J. W., Ashraf, M. & Shull, G. E. (2003). Mice with a null mutation in the NHE1 Na^+/H^+ exchanger are resistant to cardiac ischemia-reperfusion injury. *Circ Res* **93**, 776–782.

151. Yamashita, J. *et al.* (2007). Role of Na^+/H^+ exchanger in the pathogenesis of ischemic acute renal failure in mice. *J Cardiovasc Pharmacol* **49**, 154–160.

152. Rupprecht, H. J. *et al.* (2000). Cardioprotective effects of the Na^+/H^+ exchange inhibitor cariporide in patients with acute anterior myocardial infarction undergoing direct PTCA. *Circulation* **101**, 2902–2908.

153. Theroux, P. *et al.* (2000). Design of a trial evaluating myocardial cell protection with cariporide, an inhibitor of the transmembrane sodium-hydrogen exchanger: the Guard during ischemia against necrosis (GUARDIAN) trial. *Current controlled trials in cardiovascular medicine* **1**, 59–67.

154. Zeymer, U. *et al.* (2001). The Na^+/H^+ exchange inhibitor eniporide as an adjunct to early reperfusion therapy for acute myocardial infarction. results of the evaluation of the safety and cardioprotective effects of eniporide in acute myocardial infarction (ESCAMI) trial. *J Am Coll Cardiol* **38**, 1644–1650.

155. Boyce, S. W. *et al.* (2003). Impact of sodium-hydrogen exchange inhibition by cariporide on death or myocardial infarction in high-risk CABG surgery patients: results of the CABG surgery cohort of the GUARDIAN study. *J Thorac Cardiovasc Surg* **126**, 420–427.

156. Chaitman, B. R. (2003). A review of the GUARDIAN trial results: clinical implications and the significance of elevated perioperative CK-MB on 6-month survival. *J Card Surg* **18**, 13–20.

157. Mentzer Jr., R. M. (2003). Effects of Na^+/H^+ exchange inhibition by cariporide on death and nonfatal myocardial infarction in patients undergoing coronary artery bypass graft surgery: The expedition study. *Circulation* **108**, 2723–2723.

158. Karmazyn, M. (2013). NHE-1: still a viable therapeutic target. *J Mol Cell Cardiol* **61**, 77–82.

159. Cook, A. R. *et al.* (2009). Paradoxical resistance to myocardial ischemia and age-related cardiomyopathy in NHE1 transgenic mice: a role for ER stress? *J Mol Cell Cardiol* **46**, 225–233.

160. Avkiran, M., Cook, A. R. & Cuello, F. (2008). Targeting Na^+/H^+ exchanger regulation for cardiac protection: a RSKy approach? *Curr Opin Pharmacol* **8**, 133–140.

161. Kraut, J. A. & Madias, N. E. (2012). Treatment of acute metabolic acidosis: a pathophysiologic approach. *Nat Rev Nephrol* **8**, 589–601.

162. Boedtkjer, E., Bunch, L. & Pedersen, S. F. (2012). Physiology, pharmacology and pathophysiology of the pH regulatory transport proteins NHE1 and NBCn1: similarities, differences, and implications for cancer therapy. *Current pharmaceutical design* **18**, 1345–1371.

163. Daniel, C. *et al.* (2013). The role of proton dynamics in the development and maintenance of multidrug resistance in cancer. *Biochim Biophys Acta* **1832**, 606–617.

164. Aronson, P. S., Nee, J. & Suhm, M. A. (1982). Modifier role of internal H^+ in activating the Na^+/H^+ exchanger in renal microvillus membrane vesicles. *Nature* **299**, 161–163.

165. Counillon, L., Franchi, A. & Pouyssgur, J. (1993). A point mutation of the Na^+/H^+ exchanger gene (NHE1) and amplification of the mutated allele confer amiloride resistance upon chronic acidosis. *Proc Natl Acad Sci U S A* **90**, 4508–4512.

166. Khadilkar, A., Iannuzzi, P. & Orlowski, J. (2001). Identification of sites in the second exomembrane loop and ninth transmembrane helix of the mammalian Na^+/H^+ exchanger important for drug recognition and cation translocation. *J Biol Chem* **276**, 43792–43800.

167. Touret, N., Poujeol, P. & Counillon, L. (2001). Second-site revertants of a low-sodium-affinity mutant of the Na^+/H^+ exchanger reveal the participation of TM4 into a highly constrained sodium-binding site. *Biochemistry* **40**, 5095–5101.

168. Slepkov, E. R. *et al.* (2005). Structural and functional characterization of transmembrane segment IV of the NHE1 isoform of the Na^+/H^+ exchanger. *J Biol Chem* **280**, 17863–17872.

169. Pedersen, S. F., King, S. A., Nygaard, E. B., Rigor, R. R. & Cala, P. M. (2007). NHE1 inhibition by amiloride- and benzoylguanidine-type compounds. inhibitor binding loci deduced from chimeras of NHE1 homologues with endogenous differences in inhibitor sensitivity. *J Biol Chem* **282**, 19716–19727.

170. Hisamitsu, T., Ben Ammar, Y., Nakamura, T. Y. & Wakabayashi, S. (2006). Dimerization is crucial for the function of the Na^+/H^+ exchanger NHE1. *Biochemistry* **45**, 13346–13355.

171. Moncoq, K., Kemp, G., Li, X., Fliegel, L. & Young, H. S. (2008). Dimeric structure of human Na^+/H^+ exchanger isoform 1 overproduced in *Saccharomyces cerevisiae*. *J Biol Chem* **283**, 4145–4154.

172. Fafournoux, P., Noël, J. & Pouysségur, J. (1994). Evidence that Na^+/H^+ exchanger isoforms NHE1 and NHE3 exist as stable dimers in membranes with a high degree of specificity for homodimers. *J Biol Chem* **269**, 2589–2596.

173. Wakabayashi, S., Pang, T., Su, X. & Shigekawa, M. (2000). A novel topology model of the human Na^+/H^+ exchanger isoform 1. *J Biol Chem* **275**, 7942–7949.

174. Kyte, J. & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**, 105–132.

175. Ding, J., Rainey, J. K., Xu, C., Sykes, B. D. & Fliegel, L. (2006). Structural and functional characterization of transmembrane segment VII of the Na^+/H^+ exchanger isoform 1. *J Biol Chem* **281**, 29817–29829.

176. Reddy, T. *et al.* (2008). Structural and functional characterization of transmembrane segment IX of the NHE1 isoform of the Na^+/H^+ exchanger. *J Biol Chem* **283**, 22018–22030.

177. Lee, B. L., Li, X., Liu, Y., Sykes, B. D. & Fliegel, L. (2009). Structural and functional analysis of transmembrane XI of the NHE1 isoform of the Na^+/H^+ exchanger. *J Biol Chem* **284**, 11546–11556.

178. Lee, B. L., Li, X., Liu, Y., Sykes, B. D. & Fliegel, L. (2009). Structural and functional analysis of extracellular loop 2 of the Na^+/H^+ exchanger. *Biochim Biophys Acta* **1788**, 2481–2488.

179. Tzeng, J., Lee, B. L., Sykes, B. D. & Fliegel, L. (2010). Structural and functional analysis of transmembrane segment VI of the NHE1 isoform of the Na^+/H^+ exchanger. *J Biol Chem* **285**, 36656–36665.

180. Lee, B. L., Liu, Y., Li, X., Sykes, B. D. & Fliegel, L. (2012). Structural and functional analysis of extracellular loop 4 of the nhe1 isoform of the na(+)/H(+) exchanger. *Biochim Biophys Acta* **1818**, 2783–2790.

181. Landau, M., Herz, K., Padan, E. & Ben-Tal, N. (2007). Model structure of the Na^+/H^+ exchanger 1 (NHE1): functional and clinical implications. *J Biol Chem* **282**, 37854–37863.

182. Hunte, C. *et al.* (2005). Structure of a Na^+/H^+ antiporter and insights into mechanism of action and regulation by pH. *Nature* **435**, 1197–1202.

183. Ammar, Y. B., Takeda, S., Hisamitsu, T., Mori, H. & Wakabayashi, S. (2006). Crystal structure of CHP2 complexed with NHE1-cytosolic region and an implication for pH regulation. *The EMBO Journal* **25**, 2315–2325.

184. Mishima, M., Wakabayashi, S. & Kojima, C. (2007). Solution structure of the cytoplasmic region of Na^+/H^+ exchanger 1 complexed with essential cofactor calcineurin B homologous protein 1. *J Biol Chem* **282**, 2741–2751.

185. Köster, S., Pavkov-Keller, T., Kuhlbrandt, W. & Yildiz, o. (2011). Structure of human Na^+/H^+ exchanger NHE1 regulatory region in complex with calmodulin and Ca^{2+}. *J Biol Chem* **286**, 40954–40961.

186. Gebreselassie, D., Rajarathnam, K. & Fliegel, L. (1998). Expression, purification, and characterization of the carboxyl-terminal region of the Na^+/H^+ exchanger. *Biochem Cell Biol* **76**, 837–842.

187. Nørholm, A.-B. *et al.* (2011). The intracellular distal tail of the Na^+/H^+ exchanger NHE1 is intrinsically disordered: implications for NHE1 trafficking. *Biochemistry* **50**, 3469–3480.

188. Goldberg, E. B. *et al.* (1987). Characterization of a Na^+/H^+ antiporter gene of *Escherichia coli*. *Proc Natl Acad Sci U S A* **84**, 2615–2619.

189. Taglicht, D., Padan, E. & Schuldiner, S. (1991). Overproduction and purification of a functional Na^+/H^+ antiporter coded by nhaA (ant) from *Escherichia coli*. *J Biol Chem* **266**, 11289–11294.

190. Williams, K. A., Geldmacher-Kaufer, U., Padan, E., Schuldiner, S. & Kühlbrandt, W. (1999). Projection structure of NhaA, a secondary transporter from *Escherichia coli*, at 4.0 å resolution. *EMBO J* **18**, 3558–3563.

191. Williams, K. A. (2000). Three-dimensional structure of the ion-coupled transport protein NhaA. *Nature* **403**, 112–115.

192. Padan, E. (2008). The enlightening encounter between structure and function in the NhaA Na^+/H^+ antiporter. *Trends Biochem Sci* **33**, 435–443.

193. Arkin, I. T. *et al.* (2007). Mechanism of Na^+/H^+ antiporting. *Science* **317**, 799–803.

194. Herz, K., Rimon, A., Olkhova, E., Kozachkov, L. & Padan, E. (2010). Transmembrane segment II of NhaA Na^+/H^+ antiporter lines the cation passage, and asp65 is critical for pH activation of the antiporter. *J Biol Chem* **285**, 2211–2220.

195. Kozachkov, L. & Padan, E. (2011). Site-directed tryptophan fluorescence reveals two essential conformational changes in the Na^+/H^+ antiporter NhaA. *Proc Natl Acad Sci U S A* **108**, 15769–15774.

196. Diab, M., Rimon, A., Tzubery, T. & Padan, E. (2011). Helix VIII of NhaA Na^+/H^+ antiporter participates in the periplasmic cation passage and pH regulation of the antiporter. *J Mol Biol* **413**, 604–614.

197. Rimon, A., Kozachkov-Magrisso, L. & Padan, E. (2012). The unwound portion dividing helix IV of NhaA undergoes a conformational change at physiological pH and lines the cation passage. *Biochemistry* **51**, 9560–9569.

198. Kozachkov, L. & Padan, E. (2013). Conformational changes in NhaA Na^+/H^+ antiporter. *Mol Membr Biol* **30**, 90–100.

199. Mager, T. *et al.* (2013). Differential effects of mutations on the transport properties of the Na^+/H^+-antiporter NhaA from *Escherichia coli*. *J Biol Chem* .

200. Mager, T., Rimon, A., Padan, E. & Fendler, K. (2011). Transport mechanism and pH regulation of the Na^+/H^+ antiporter NhaA from *Escherichia coli*: an electrophysiological study. *J Biol Chem* **286**, 23570–23581.

201. Schushan, M. *et al.* (2012). A model-structure of a periplasm-facing state of the NhaA antiporter suggests the molecular underpinnings of pH-induced conformational changes. *J Biol Chem* **287**, 18249–18261.

202. Padan, E. & Schuldiner, S. (1994). Molecular physiology of the Na^+/H^+ antiporter in *Escherichia coli*. *J Exp Biol* **196**, 443–456.

203. Padan, E. (2011). Regulation of NhaA by protons. *Comprehensive Physiology* **1**, 1711–1719.

204. Lolkema, J. S., Dobrowolski, A. & Slotboom, D.-J. (2008). Evolution of antiparallel two-domain membrane proteins: tracing multiple gene duplication events in the DUF606 family. *J Mol Biol* **378**, 596–606.

205. Kamiya, T. & Maeshima, M. (2004). Residues in internal repeats of the rice cation/H^+ exchanger are involved in the transport and selection of cations. *J Biol Chem* **279**, 812–819.

206. Abramson, J. & Wright, E. M. (2009). Structure and function of Na^+-symporters with inverted repeats. *Curr Opin Struct Biol* **19**, 425–432.

207. Oberai, A., Ihm, Y., Kim, S. & Bowie, J. U. (2006). A limited universe of membrane protein families and folds. *Protein Sci* **15**, 1723–1734.

208. Hu, N.-J., Iwata, S., Cameron, A. D. & Drew, D. (2011). Crystal structure of a bacterial homologue of the bile acid sodium symporter ASBT. *Nature* **478**, 408–411.

209. Goswami, P. *et al.* (2011). Structure of the archaeal Na^+/H^+ antiporter NhaP1 and functional role of transmembrane helix 1. *EMBO J* **30**, 439–449.

210. Vinothkumar, K. R., Smits, S. H. J. & Kühlbrandt, W. (2005). pH-induced structural change in a sodium/proton antiporter from *Methanococcus jannaschii*. *EMBO J* **24**, 2720–2729.

211. Hellmer, J., Pätzold, R. & Zeilinger, C. (2002). Identification of a pH regulated Na^+/H^+ antiporter of *Methanococcus jannaschii*. *FEBS Letters* **527**, 245–249.

212. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *J Mol Biol* **305**, 567–580.

213. Saier, J., M H *et al.* (1999). Phylogenetic characterization of novel transport protein families revealed by genome analyses. *Biochim Biophys Acta* **1422**, 1–56.

214. Wells, K. M. & Rao, R. (2001). The yeast Na^+/H^+ exchanger Nhx1 is an N-linked glycoprotein. topological implications. *J Biol Chem* **276**, 3401–3407.

215. Sardet, C., Fafournoux, P. & Pouysségur, J. (1991). Alpha-thrombin, epidermal growth factor, and okadaic acid activate the Na^+/H^+ exchanger, NHE-1, by phosphorylating a set of common sites. *J Biol Chem* **266**, 19166–19171.

216. Murtazina, R., Booth, B. J., Bullis, B. L., Singh, D. N. & Fliegel, L. (2001). Functional analysis of polar amino-acid residues in membrane associated regions of the NHE1 isoform of the mammalian Na^+/H^+ exchanger. *Eur J Biochem* **268**, 4674–4685.

217. Akabas, M. H., Kaufmann, C., Cook, T. A. & Archdeacon, P. (1994). Amino acid residues lining the chloride channel of the cystic fibrosis transmembrane conductance regulator. *J Biol Chem* **269**, 14865–14868.

218. Sato, Y., Ariyoshi, N., Mihara, K. & Sakaguchi, M. (2004). Topogenesis of NHE1: direct insertion of the membrane loop and sequestration of cryptic glycosylation and processing sites just after TM9. *Biochem Biophys Res Commun* **324**, 281–287.

219. Sato, Y. & Sakaguchi, M. (2005). Topogenic properties of transmembrane segments of *Arabidopsis thaliana* NHX1 reveal a common topology model of the Na^+/H^+ exchanger family. *J Biochem* **138**, 425–431.

220. Dibrov, P., Young, P. G. & Fliegel, L. (1998). Functional analysis of amino acid residues essential for activity in the Na^+/H^+ exchanger of fission yeast. *Biochemistry* **37**, 8282–8288.

221. Ullah, A. *et al.* (2013). Structural and functional analysis of transmembrane segment IV of the salt tolerance protein sod2. *J Biol Chem* **288**, 24609–24624.

Chapter 2

Expression, purification and characterization of NHE1 expressed in yeast

This research was originally published in The Journal of Biological Chemistry. Karine Moncoq, Grant Kemp, Xiuju Li, Larry Fliegel and Howard S. Young. Dimeric Structure of Human Na^+/H^+ Exchanger Isoform 1 Overproduced in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 2008; 283:4145–4154. © the American Society for Biochemistry and Molecular Biology. It has been modified and expanded here to reflect further experimental optimization and different approaches explored during the tenure of my PhD studies. The roles of Moncoq and Kemp have highlighted in the text.

Introduction

The physiological and pathological significance of NHE1, outlined in *Chapter 1*, spurred us to attempt to determine the three-dimensional structure of this transporter. However, no abundant natural source of this protein exists, requiring the development of an overproduction and purification system to begin structural studies. Though the Fliegel laboratory has previously had some success in high level expression and structural analysis of isolated short segments of the

mammalian NHE1 protein (1–5), the initial production of larger transmembrane fragments of eukaryotic Na^+/H^+ exchangers in *Escherichia coli* remained unsuccessful (6). While later work during my PhD was able to surmount some of these challenges and successfully produce larger transmembrane fragments (see *Chapter 3* and *Chapter 5*), these approaches were not useful for the entire full length membrane domain. This chapter discusses our advances in producing functional full-length NHE1 protein. Prior work by many groups has established yeast as a useful overexpression host for eukaryotic membrane proteins (see *Chapter 1*). Perhaps most famously, the yeast *Pichia pastoris* was used to express Kv1.2 from *Rattus norvegicus*, a Shaker-family potassium channel (7). This enabled the determination of the first structure of a mammalian membrane protein obtained without a natural source (8). A few mammalian Na^+/H^+ exchanger isoforms have been heterologously expressed in *S. cerevisiae*, though only in relatively small amounts, with NHE1 expressed either as a functionally inactive or mistargeted protein (9–11). These results suggested that, with modification, *S. cerevisiae* might be a suitable host for large scale overexpression. Herein we describe the expression, purification and characterization of the human NHE1 isoform of the Na^+/H^+ exchanger in a *S. cerevisiae* system, following an approach similar to that used for overproduction of an anion exchanger (*Band 3*, AE1) in *S. cerevisiae* (12). Our results present a system that can be used to produce milligram quantities of human NHE1 (hNHE1) suitable for preliminary structural studies, where the protein is fully functional in reconstituted vesicles following affinity purification from detergent solubilized membranes. Single particle electron microscopy using detergent solubilized NHE1 reveals a compact dimer at 22 Å resolution. These data provide the first structural insight into NHE1 and its oligomeric state.

Experimental Procedures

Materials

Biochemicals were from Sigma unless specified otherwise. Taq DNA polymerase was purchased from Invitrogen. DNA restriction and modification enzymes were from New England BioLabs. Products for culture of *S. cerevisiae* and *P. pastoris* were purchased from DIFCO Media Prod-

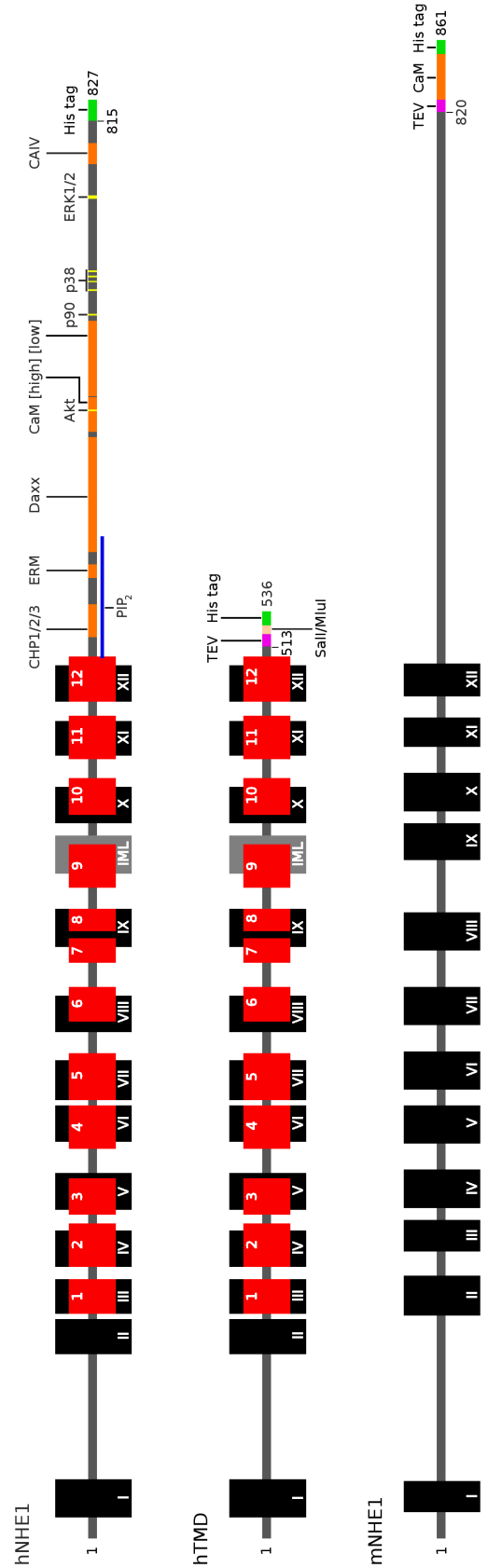

Figure 2.1: 2D diagrams of NHE1 constructs for expression in yeast. The relative position and length of each feature in the primary structure is shown to scale. **hNHE1**: the full length human NHE1 gene expressed with a flexible His-tag (Gly_2His_{10}, shown in green) for IMAC purification. The putative transmembrane segments defined by Wakabayashi *et al.* (13) are labelled as black boxes with Roman numerals. The long intramembrane loop proposed is shown as a grey box (IML). Transmembrane segments predicted by Landau *et al.* (14) are shown as red boxes labelled with Arabic numbers. Sites of protein-protein interaction are shown in orange, phosphorylation sites in yellow and the PIP_2 binding region underlined in blue. The interaction partner acronyms are: CHP, calcineurin homologous protein; PIP_2, phosphatidylinositol-4,5-bisphosphate; ERM, erzin-radixin-moesin; Daxx, death-associated protein-6; Akt, protein kinase B; CaM, calmodulin ([high] and [low] affinity sites); p90 (or $p90^{rsk}$ or Rsk), p90 ribosomal S6 kinase or (mitogen-activated-protein-kinase)-activated protein kinase-1; p38, p38 mitogen-activated protein kinase; ERK, extracellular signal-regulated kinase; CAIV, carbonic anhydrase IV. **hTMD**: transmembrane domain construct of **hNHE1**, containing: a TEV cleavage site (magenta) for the removal of purification tags; a *SalI/MluI* unique cloning site for addition of new purification tags (beige); and a flexible His-tag (Gly_2His_{10}, green). **mNHE1**: full length murine NHE1 with added TEV site (magenta), CaM affinity tag (orange) and 8xHis-tag (green). Predicted transmembrane segments (TMHMM2 (15)) are shown in black.

ucts. Zeocin® antibiotic was purchased from Invitrogen. Ni-NTA agarose was purchased from Qiagen. Calmodulin-agarose resin, Sephadex G-50, and Superdex 200 HR 10/30 were from Amersham Biosciences (Piscataway, NJ). The detergents octaethylene glycol mono-n-dodecyl ether ($C_{12}E_8$), n-dodecyl β-d-maltoside (DDM), and Fos-choline (FC) 10, 12, and 14 were purchased from Anatrace (Maumee, OH). l-α-lysophosphatidylcholine (LPC) and egg yolk l-α-phosphatidylcholine (EYPC) were from Avanti Polar Lipids (Alabaster, AL). Bio-Beads (SM-2) were from Bio-Rad. HisProbe-HRP was from Pierce, anti-NHE1 monoclonal antibody was from Chemicon Int. Inc. (Temecula, CA).

Yeast strains and plasmids

S. cerevisiae strains W303.1b (*leu2–3,112; his3–11,15; trp1–1; ura3–1; ade2–1; can1–100^r; cir$^+$*) and DSY864 (*pep4; prb1–112; leu2; trp1; ura3–52; his3*::GAL1-GAL4) were a gift from Dr. David Stuart (Department of Biochemistry, University of Alberta). W303.1b is a common laboratory strain used for various genetic experiments, analogous to *E. coli* DH5α. DSY864, originally NKY879 (16, 17), lacks two major vacuolar proteases (*pep4* and *prb1*) and elicits a strong response to galactose induction (*his3*::GAL1-GAL4), leading to increased protein overexpression from galactose inducible promoters and decreased degradation. The expression plasmid pYeDP60 (Figure 2.5) was generously supplied by Denis Pompon (CGM, Gif-sur-Yvette, France). *P. pastoris* strains GS115(*his4*) and X-33 (wild type), and expression vector pPICZ-A (Figure 2.6) were a gift from Dr. M. Joanne Lemieux (Department of Biochemistry, University of Alberta).

Site-directed mutagenesis and plasmid construction

Completed by Dr. Karine Moncoq.

The primers used for the following steps are listed in Figure 2.2. The complete cDNA of hNHE1, in the pYN4$^+$ plasmid (1), was subjected to site-directed mutagenesis using a QuikChange™ site-directed mutagenesis kit (Stratagene, La Jolla, CA) as recommended by the manufacturer. Briefly, PWO DNA polymerase (Roche Applied Sciences) and specific primers (Figure 2.2a) were used to mutate the non-essential glycosylation site at amino acid Asn75 to aspartic acid and simultaneously create a novel silent restriction site (*SalI*) for screening transformants. Plas-

mid DNA purified from positive transformants was then subsequently mutated, using the above method with the primers in Figure 2.2b, to remove the unique *BglII* restriction site in the hNHE1 coding sequence of pYN4$^+$-NHE1(N75D). This mutation was required to allow the insertion of hNHE1 into the multiple cloning site of pYeDP60 (Figure 2.5). hNHE1(N75D)$^{-BglII}$ was amplified by PCR using TaqDNA polymerase and specific primers (Figure 2.2c) incorporating a 5' *BglII* restriction site and a protein carboxy-terminal Gly_2His_{10} tag followed by a stop codon and a 3' *KpnI* restriction site. This amplified fragment was cut with *BglII* and *KpnI*, and inserted in-frame into the plasmid pYeDP60 after its digestion with the compatible *BamHI* and *KpnI* restriction enzymes. DNA sequencing verified the correct insertion, N75D mutation, and the absence of other mutations. The resultant construct contained human NHE1(N75D)-Gly_2His_{10} (hNHE1, Figure 2.1) under the control of the inducible GAL10-CYC1 hybrid promoter and a phosphoglycerate kinase terminator.

fwd: 5'–GAGCCGCCCTGTcgAcCATTCCGTCACTG–3'

rev: 5'–CAGTGACGGAATGgTcgACAGGGCGGCTC–3'

(a)

fwd: 5'–CGCACCCCCTTCGAAATCTCCCTCTGG–3'

rev: 5'–CCAGAGGGAGATTTCGAAGGGGGTGCG–3'

(b)

fwd: 5'–GGATAGATCTATGGTTCTGCGGTCTGGCATCTGTGGCCTC–3'

rev: 5'–CCATGGTACC**TCA***ATGATGATGATGATGATGATGATGATGATG-ACCACCC*CTGCCCCTTGGGGAAGAACGGTTCTC–3'

(c)

Figure 2.2: Primers used for construction of yeast expression vector pYeDP60 containing full-length human NHE1(N75D). Forward (fwd) and reverse (rev) primers are labelled. (a) Primers used for site-directed mutagenesis of Asn75 to Asp within the pNY4$^+$ plasmid. Mutated nucleotides shown in lower case text. (b) Primers used to remove the *BglII* restriction site from coding region of pYN4$^+$-NHE1(N75D). (c) Primers used to amplify hNHE1(N75D) from the pNY4$^+$ plasmid. Additional nucleotides incorporating a *BglII* restriction site (fwd, underlined), a *KpnI* restriction site (rev, underlined), a stop codon (rev, bold) and a C-terminal Gly_2His_{10} tag (rev, italics) are highlighted.

	2° codon U			2° codon C			2° codon A			2° codon G		
	UUU	26.1	170666	UCU	23.5	153557	UAU	18.8	122728	UGU	8.1	52903
1°	UUC	18.4	120510	UCC	14.2	92923	UAC	14.8	96596	UGC	4.8	31095
U	UUA	26.2	170884	UCA	18.7	122028	UAA	1.1	6913	UGA	0.7	4447
	UUG	27.2	177573	UCG	8.6	55951	UAG	0.5	3312	UGG	10.4	67789
	CUU	12.3	80076	CCU	13.5	88263	CAU	13.6	89007	CGU	6.4	41791
1°	CUC	5.4	35545	CCC	6.8	44309	CAC	7.8	50785	CGC	2.6	16993
C	CUA	13.4	87619	CCA	18.3	119641	CAA	27.3	178251	CGA	3	19562
	CUG	10.5	68494	CCG	5.3	34597	CAG	12.1	79121	CGG	1.7	11351
	AUU	30.1	196893	ACU	20.3	132522	AAU	35.7	233124	AGU	14.2	92466
1°	AUC	17.2	112176	ACC	12.7	83207	AAC	24.8	162199	AGC	9.8	63726
A	AUA	17.8	116254	ACA	17.8	116084	AAA	41.9	273618	AGA	21.3	139081
	AUG	20.9	136805	ACG	8	52045	AAG	30.8	201361	AGG	9.2	60289
	GUU	22.1	144243	GCU	21.2	138358	GAU	37.6	245641	GGU	23.9	156109
1°	GUC	11.8	76947	GCC	12.6	82357	GAC	20.2	132048	GGC	9.8	63903
G	GUA	11.8	76927	GCA	16.2	105910	GAA	45.6	297944	GGA	10.9	71216
	GUG	10.8	70337	GCG	6.2	40358	GAG	19.2	125717	GGG	6	39359

Table 2.1: *Saccharomyces cerevisiae* Codon Usage Table. The data in the table is calculated from sequencing data deposited in NCBI-Genbank (www.ncbi.nlm.nih.gov/genbank) Flat File Release 160.0 (June 15, 2007), containing 14411 Coding DNA Sequences (6534504 codons). The columns of each 3x4 section are (from left to right): mRNA codon, frequency of occurrence (per thousand), actual number of codons found.

Alternate construct development

Completed by Grant Kemp.

Optimized cDNA corresponding to the transmembrane domain of hNHE1(N75D) (hTMD, Figure 2.1) was synthesized by Biomatik Corporation (Cambridge, ON, Canada) and delivered in a pUC19 vector. Sequence optimization for expression in *S. cerevisiae* was performed using OPTIMIZER (http://genomes.urv.es/OPTIMIZER) and the codon usage table from Kazusa DNA Research Institute (Japan), Table 2.1. Engineered onto the C-terminus was a TEV protease site for tag removal following purification, a *SalI*/*MluI* unique cloning site for the addition of other purification tags if required and a Gly_2His_{10} tag (Figure 2.3). Following *E. coli* transformation and plasmid purification, the synthesized gene was cut out of pUC19 with *BamHI* and *KpnI*, and cloned into pYeDP60 (Figure 2.5) as above. Additionally the cDNA of murine NHE1 (mNHE1, Figure 2.1) was synthesized by GeneArt (Regensburg, Germany) using codon harmonization based on the codon usage table generously provided by Dr. Ina Urbatsch (Texas Tech University Health Sciences Center, USA), Table 2.2. A TEV site, CaM purification tag and octahistidine tag were engineered on the C-terminus. Following *E. coli* transformation and plasmid purification, the synthesized gene was cut out of GeneArt's proprietary vector (pMA-RQ) with *EcoRI* and *KpnI* (Figure 2.4) and cloned into pPICZ-A (Figure 2.6) at the same sites.

	Codon	Kazusa *pastoris*[A]	Genomic *pastoris*[B]	H.E. *cerevisiae*[C]	H.E. *pastoris*[D]	Codon	Kazusa *pastoris*[A]	Genomic *pastoris*[B]	H.E. *cerevisiae*[C]	H.E. *pastoris*[D]
Ala	GCA	23.4	27.5	4.0	8.6	GCU	44.8	40.1	67.7	59.0
	GCG	6.1	7.9	1.0	1.2	GCC	25.7	24.5	27.3	31.3
Pro	CCA	41.6	37.9	85.4	58.9	CCU	34.8	33.9	12.2	35.0
	CCG	8.6	10.2	0.0	1.6	CCC	15.0	18.1	2.4	4.5
Cys	UGU	63.6	62.2	88.9	82.0	UGC	36.4	37.8	11.1	18.0
Gln	CAA	60.9	60.4	97.1	68.0	CAG	39.1	39.6	2.9	32.0
Asp	GAU	58.0	61.7	44.0	41.0	GAC	42.1	38.3	56.0	59.0
Arg	AGA	48.0	45.6	82.7	68.9	AGG	15.8	18.1	1.9	3.6
	CGA	10.0	12.0	0.6	2.3	CGU	16.5	14.1	15.4	23.4
	CGG	4.5	5.3	0.3	0.0	CGC	5.3	5.1	1.1	1.9
Glu	GAA	56.3	59.5	92.2	39.1	GAG	43.7	40.5	7.8	60.9
Phe	UUU	53.9	57.8	23.1	29.5	UUC	46.1	42.2	76.9	70.5
Ser	AGU	15.0	15.9	3.0	5.6	AGC	9.1	10.4	3.0	3.6
	UCA	18.2	20.8	6.0	9.9	UCU	29.2	26.0	52.2	46.1
	UCG	8.9	9.5	1.5	3.9	UCC	19.7	17.4	34.3	31.0
Gly	GGA	32.7	36.5	1.5	20.4	GGU	43.6	35.8	94.2	70.8
	GGG	9.9	12.2	0.1	2.2	GGC	13.9	15.6	4.4	6.7
His	CAU	56.5	61.8	31.6	32.2	CAC	43.5	38.2	68.4	67.8
Thr	ACA	24.3	27.7	4.7	8.8	ACU	39.5	36.3	50.0	47.5
	ACG	10.6	12.3	1.6	3.1	ACC	25.6	23.8	43.8	40.6
Val	GUA	15.5	17.8	2.5	5.0	GUU	42.0	38.8	56.3	49.1
	GUG	19.2	21.7	3.8	9.5	GUC	23.3	21.7	37.5	36.5
Lys	AAA	46.9	52.0	20.7	22.1	AAG	53.1	48.0	79.3	77.9
Leu	UUA	16.2	17.9	18.3	9.8	UUG	32.7	29.1	67.1	50.8
	CUA	11.1	12.3	9.8	5.8	CUU	16.5	17.1	2.4	13.7
	CUG	15.5	15.3	1.2	15.7	CUC	7.9	8.4	1.2	4.2
Tyr	UAU	46.9	51.4	15.2	23.5	UAC	53.1	48.6	84.9	76.5
Asn	AAU	48.5	53.5	17.5	24.4	AAC	51.5	46.5	82.5	75.6
Met	AUG	100.0	100.0	100.0	100.0					
Trp	UGG	100.0	100.0	100.0	100.0					
Ile	AUA	18.0	23.6	3.1	2.6					
	AUU	50.5	46.7	46.9	52.6					
	AUC	31.5	29.8	50.0	44.8					
STOP	UAA	50.0	40.0	75.0	59.4					
	UAG	31.3	34.0	25.0	34.4					
	UGA	18.8	24.9	0.0	6.3					

Table 2.2: *Saccharomyces cerevisiae* and *Pichia pastoris* Codon Usage. This table of codon usage in DNA coding sequences is modified from Bai *et al.* (18). The values are shown as percent codon usage. The data was sources as follows: *A* uses codon usage information from the Kazusa database (www.kazusa.or.jp/codon/ (c. 2007)); *B* uses codon usage from the *Pichia pastoris* sequencing project (19); *C* uses codon usage data of highly expressed proteins in *S. cerevisiae* (20); *D* uses proteomic data of 30 of the mostly higly expressed proteins in *P. pastoris* compiled from (21–23). Preferred codon usage in highly expressed proteins are highlighted in red, infrequently codon usage in highly expressed proteins are highlighted in yellow, and codons favourable for *S. cerevisiae* over *P. pastoris* are highlighted in green.

Yeast Transformation, Expression, and Culture

Completed by Dr. Karine Moncoq and Grant Kemp.

For constructs in pYeDP60 (Figure 2.5), the yeast strain W303.1b or DSY864 was transformed using a lithium acetate procedure (24) and selected on supplemented Synthetic Dropout medium, SDaa (0.67% yeast nitrogen base, 2% glucose, 0.5% casamino acids, ± 40 mg/L Trp (for W303.1b)). For large scale expression, a single yeast colony was inoculated into SDaa and grown overnight at 30 °C. A 1:50 dilution was made into YPDE-2 medium (2% yeast extract, 1% glucose, 2% bactopeptone, 2.7% ethanol) and cells were grown at 28 °C with shaking until an A_{600} of 8.0 was obtained (~24–36 h). The expression of recombinant hNHE1 was induced by addition of sterile galactose to a final concentration of 2% (w/v) and 30 °C incubation continued for 12–14 h. After induction, cells were harvested yielding ~25 g wet cells per litre of culture.

For constructs in pPICZ-A (Figure 2.6), protocols from EasySelect™ *Pichia* Expression Kit (Invitrogen) manual were followed. Briefly, the plasmid was linearized within the 5' AOX1 gene by digestion with *SacI* and transformed into *P. pastoris* by electroporation. The electroporated cells were incubated in 1 M sorbitol at 30 °C for 1 hour without shaking to allow DNA integration and cell recovery. The recovered cells were plated on YPDS (1% yeast extract, 2% peptone, 2% glucose, 1 M sorbitol) agar containing 100 μg/mL Zeocin antibiotic and incubated until colonies appeared (3-10 days). Since the linearized plasmid DNA may integrate into the AOX1 gene of *P. pastoris* multiple times, and increase the target expression levels (25, 26), individual transformed colonies were screened for multi-copy recombination by selection on plates containing increasing amounts of Zeocin (500, 1000, or 2000 μg/mL). Colonies that grew well at higher concentrations were screened for hNHE1 protein expression. A single colony was selected and grown in either YPD (2% yeast extract, 1% glucose, 2% bactopeptone), MGY (1.34% yeast nitrogen base, 1% glycerol, 4 x 10^{-5} biotin, ± 0.004% histidine) or BMGY (100 mM potassium phosphate buffer pH 6.0, 1.34% yeast nitrogen base, 1% glycerol, 4 x 10^{-5} biotin, ± 0.004% histidine) at 30 °C with shaking until OD_{600} of >2. This culture was harvested in sterile centrifuge bottles and resuspended in MM (1.34% yeast nitrogen base, 4 x 10^{-5} biotin, 0.5% methanol, ± 0.004% histidine) or BMM (100 mM potassium phosphate buffer pH 6.0, 1.34% yeast nitrogen base, 4 x 10^{-5} biotin, 0.5% methanol, ± 0.004% histidine) induction media and incubated at 30 °C with shaking for at least 24 h.

hTMD

GGATCCAAAA AA**ATG**GTTTT GAGATCTGGT ATTTGTGGTT TGTCTCCACA TAGAATTTTC - 60
CCATCTTTGT TGGTTGTTGT TGCTTTGGTT GGTTTGTTGC CAGTTTTGAG ATCTCATGGT - 120
TTGCAATTGT CTCCAACTGC TTCTACTATT AGATCTTCTG AACCACCAAG AGAAAGATCT - 180
ATTGGTGATG TTACTACTGC TCCACCAGAA GTTACTCCAG AATCTAGACC AGTTGATCAT - 240
TCTGTTACTG ATCATGGTAT GAAACCAAGA AAAGCTTTCC CAGTTTTGGG TATTGATTAT - 300
ACTCATGTTA GAACTCCATT CGAAATTTCT TTGTGGATTT TGTTGGCTTG TTTGATGAAA - 360
ATTGGTTTCC ATGTTATTCC AACTATTTCT TCTATTGTTC CAGAATCTTG TTTGTTGATT - 420
GTTGTTGGTT TGTTGGTTGG TGGTTTGATT AAAGGTGTTG GTGAAACTCC ACCATTCTTG - 480
CAATCTGATG TTTTCTTCTT GTTCTTGTTG CCACCAATTA TTTTGGATGC TGGTTATTTC - 540
TTGCCATTGA GACAATTCAC TGAAAATTTG GGTACTATTT TGATTTTCGC TGTTGTTGGT - 600
ACTTTGTGGA ATGCTTTCTT CTTGGGTGGT TTGATGTATG CTGTTTGTTT GGTTGGTGGT - 660
GAACAAATTA ATAATATTGG TTTGTTGGAT AATTTGTTGT TCGGTTCTAT TATTTCTGCT - 720
GTTGATCCAG TTGCTGTTTT GGCTGTTTTC GAAGAAATTC ATATTAATGA ATTGTTGCAT - 780
ATTTTGGTTT TCGGTGAATC TTTGTTGAAT GATGCTGTTA CTGTTGTTTT GTATCATTTG - 840
TTCGAAGAAT TCGCTAATTA TGAACATGTT GGTATTGTTG ATATTTTCTT GGGTTTCTTG - 900
TCTTTCTTCG TTGTTGCTTT GGGTGGTGTT TTGGTTGGTG TTGTTTATGG TGTTATTGCT - 960
GCTTTCACTT CTAGATTCAC TTCTCATATT AGAGTTATTG AACCATTGTT CGTTTTCTTG - 1020
TATTCTTATA TGGCTTATTT GTCTGCTGAA TTGTTCCATT TGTCTGGTAT TATGGCTTTG - 1080
ATTGCTTCTG GTGTTGTTAT GAGACCATAT GTTGAAGCTA ATATTTCTCA TAAATCTCAT - 1140
ACTACTATTA AATATTTCTT GAAAATGTGG TCTTCTGTTT CTGAAACTTT GATTTTCATT - 1200
TTCTTGGGTG TTTCTACTGT TGCTGGTTCT CATCATTGGA ATTGGACTTT CGTTATTTCT - 1260
ACTTTGTTGT TCTGTTTGAT TGCTAGAGTT TTGGGTGTTT TGGGTTTGAC TTGGTTCATT - 1320
AATAAATTCA GAATTGTTAA ATTGACTCCA AAAGATCAAT TCATTATTGC TTATGGTGGT - 1380
TTGAGAGGTG CTATTGCTTT CTCTTTGGGT TATTTGTTGG ATAAAAAACA TTTCCCAATG - 1440
TGTGATTTGT TCTTGACTGC TATTATTACT GTTATTTTCT TCACTGTTTT CGTTCAAGGT - 1500
ATGACTATTA GACCATTGGT TGATTTGTTG GCTGTTAAAA AAAAACAA*GA AAATTTGTAT* - 1620
*TTCCAAGGT*G TCGACATTAC GCGTGGTGGT CATCACCATC ACCATCACCA TCACCATCAC - 1680
TAAGGTACC - 1690

Figure 2.3: Human NHE1 transmembrane domain construct (hTMD) DNA sequence. The gene was synthesized by Biomatik Corporation (Cambridge, ON, Canada) with codons optimized for expression in *S. cerevisiae*. Important features of the gene are indicated as follows. At the 5' end a *BamHI* restriction site (highlighted in black), a Kozak sequence (italics), and a start codon (bold). At the 3' end a TEV protease site (italics), a calmodulin affinity tag (highlighted in grey), a His_8 tag (underlined), a stop codon (bold) and a *KpnI* restriction site.

mNHE1

GAGCTCGAAT TCAAAA**ATG**G TTTTGAGATG GTCTGGTGTC TGGGGATTCC ACCCACCTCG - 60
TATTTTTCCA TCCCTGCTTG TTGTTGTCGC TTTGGTGGGT TTGTTACCTG TCCTAAGATC - 120
ACACGGTCTG CAACATTCTC CAACTGCCAG TACCATCAGA GGTTCCGAGC CACCTAGAGA - 180
ACGTTCTATT GGCGACGTTA CTACCGCTCC AAGCGAGCCC TTGCACAGAC CTGACGATCA - 240
CAACTTGACA AATCTCATTA TCGAACATGG TGGAAAGCCA TCCAGAAAGG CATTCCCAGT - 300
ACTTGACATC GATTACCCTC ACGTTCGTAC TCCATTCGAG ATTTCTTTGT GGATCTTGCT - 360
GGCTTGTTTA ATGAAGATTG GTTTTCACGT CATTCCCACC ATCTCTTCGA TAGTTCCAGA - 420
GTCCTGCTTG CTTATCGTTG TCGGTTTGTT AGTGGGTGGA CTGATCAAAG GTGTCGGTGA - 480
AACTCCACCT TTCTTGCAGT CTGACGTTTT CTTCTTGTTT CTGCTCCCAC CTATTATTTT - 540
GGATGCCGGT TATTTCTTGC CTCTTAGACA ATTCACCGAG AACTTGGGCA CTATTCTGAT - 600
CTTTGCTGTT GTCGGAACCC TATGGAACGC CTTCTTTTTG GGTGGTTTAC TTTACGCTGT - 660
TTGTTTGGTG GGTGGTGAAC AAATCAACAA TATTGGATTG CTGGACACAT TGCTGTTCGG - 720
TTCCATTATC TCAGCTGTTG ATCCAGTCGC CGTTCTGGCA GTATTCGAGG AAATTCACAT - 780
CAACGAGTTG TTGCATATTC TCGTCTTTGG CGAGTCTTTG CTGAACGACG CTGTTACTGT - 840
TGTCCTTTAC CACCTGTTCG AAGAGTTCGC TTCCTACGAC TCTGTGGGTA TTAGTGATAT - 900
CTTCTTGGGT TTCCTGAGCT TCTTTGTTGT CGCCTTGGGA GGTGTTTTTG TTGGTGTTGT - 960
CTATGGTGTG ATCGCTGCAT TCACTTCTAG ATTCACCTCC CACATTCGTG TCATCGAACC - 1020
CCTTTTCGTC TTTCTGTACT CGTACATGGC TTACTTGTCT GCCGAGTTGT TCCACCTATC - 1080
CGGAATTATG GCTCTGATCG CCTCTGGTGT TGTAATGCGT CCATATGTCG AGGCTAACAT - 1140
TTCCCACAAG TCACATACAA CCATCAAGTA CTTTCTGAAA ATGTGGTCTA GTGTCTCCGA - 1200
AACTCTGATT TTCATCTTCC TGGGCGTTTC TACCGTTGCT GGTTCTCATC AGTGGAATTG - 1260
GACTTTTGTG ATTTCCACCT TGTTGTTCTG TCTCATAGCC AGAGTTTTGG GTGTTTTGGT - 1320
CCTGACTTGG TTTATCAACA AGTTCAGAAT TGTCAAGCTC ACCCCTAAAG ACCAATTCAT - 1380
TATCGCATAC GGTGGACTGA GAGGTGCTAT TGCTTTCAGC CTTGGTTACC TTCTGGATAA - 1440
GAAACACTTC CCAATGTGTG ACTTGTTTCT GACTGCCATC ATCACCGTCA TCTTTTTCAC - 1500
TGTTTTCGTC CAGGGCATGA CCATTCGTCC TTTGGTTGAC TTGCTGGCTG TGAAGAAGAA - 1560
GCAAGAGACT AAGAGATCGA TTAACGAGGA AATCCACACC CAATTCTTGG ATCACTTGCT - 1620
GACAGGCATT GAGGACATTT GCGGTCACTA TGGACACCAC CATTGGAAAG ATAAGCTGAA - 1680
TAGATTCAAC AAGAAATACG TCAAGAAATG TCTGATCGCC GGCGAGCGTT CTAAAGAACC - 1740
ACAGCTGATC GCTTTCTACC ACAAGATGGA AATGAAGCAA GCCATCGAGC TGGTTGAATC - 1800
AGGTGGTATG GGTAAGATTC CCTCTGCTGT TTCTACTGTG TCCATGCAAA ACATCCACCC - 1860
AAAGGCAGTT ACCTCCGACA GAATTTTGCC TGCTCTCTCC AAGGACAAAG AGGAAGAGAT - 1920
TAGAAAGATC CTTCGTTCCA ACCTGCAAAA GACTAGACAG AGACTGAGAT CCTACAATCG - 1980
TCATACCCTG GTCGCCGATC CATATGAAGA GGCTTGGAAC CAGATGTTGC TGCGTAGACA - 2040
AAAGGCCAGA CAACTCGAGC AAAAGATCAC TAACTACCTG ACCGTCCCTG CTCACAAGTT - 2100
AGACTCTCCA ACTCTGTCTA GAGCCCGTAT TGGAAGTGAT CCTTTGGCTT ACGAGCCAAA - 2160
GGCAGACCTG CCAGTTATTA CAATCGACCC AGCTTCTCCT CAGTCTCCCG AATCAGTTGA - 2220
TCTGGTCAAT GAGGAACTGA AGGGTAAAGT CTTGGGTCTG AACAGAGGTC CTCGTGTTAC - 2280
TCCAGAGGAA GAAGAAGAGG ACGAGGACGG CATTATTATG ATCAGAAGTA AAGAGCCATC - 2340
TAGCCCTGGA ACCGATGACG TTTTTACTCC AGGTTCTTCC GATTCACCCT CCTCGCAAAG - 2400
AATCCAACGT TGTTTGTCTG ACCCAGGTCC ACATCCTGAG CCTGGTGAGG GTGAACCTTT - 2460
TATTCCAAAG GGTCAA*GAGA ACCTCTACTT CCAGGG*TAAG AGAAGATGGA AGAAGAACTT - 2520
CATTGCTGTC TCCGCTGCCA ATCGTTTCAA GAAAATCTCT TCTTCCGGTG CCCTGCACCA - 2580
CCACCATCAC CATCATCAC**T AA**GGTACC - 2608

Figure 2.4: Murine NHE1 construct (mNHE1) DNA sequence. The gene was synthesized by GeneArt (Regensburg, Germany) using codon harmonization. Important features of the gene are indicated as follows. At the 5' end an *EcoRI* restriction site (highlighted in black), a Kozak sequence (italics), and a start codon (bold). At the 3' end a TEV protease site (italics), *SalI* and *MluI* restriction enzyme sites (highlighted in grey), a Gly_2His_{10} tag (underlined), a stop codon (bold) and a *KpnI* restriction site.

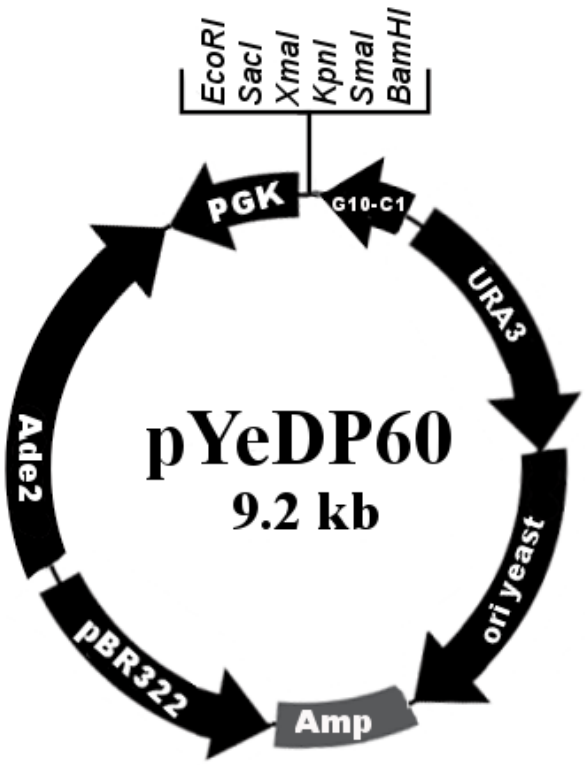

Figure 2.5: Plasmid map of yeast expression vector, pYeDP60. The plasmid can be replicated in both *E. coli* and *S. cerevisiae* using the *pBR322* and *ori yeast* elements, respectively. Bacterial selection is achieved with ampicillin (Amp, encodes β–lactamase). Nutritional selection of auxotrophic *S. cerevisiae* in minimal media can be obtained via adenine of uracil Ade2 or URA3 genotypes, respectively). The multiple cloning site, indicated, is located between a GAL10-CYC1 hybrid promoter (galactose inducible) and a phosphoglycerate kinase (PGK) terminator.

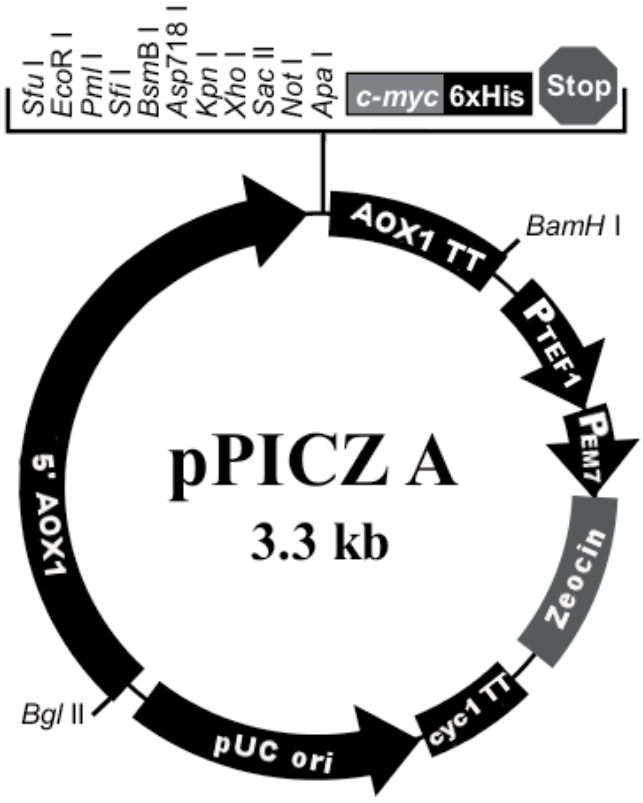

Figure 2.6: Plasmid map of yeast expression vector, pPICZ-A. The plasmid integrates the insert into the yeast genome within the potent methanol-inducible AOX1 gene. The plasmid can be replicated in *E. coli* (pUC ori) and selected in both yeast and bacteria using the antibiotic Zeocin® (Invitrogen). The multiple cloning site indicated is followed by a myc tag, a hexahistidine tag, and a stop codon; however, these elements were not used and our own construct containing two purification tags followed by a stop condon was cloned into the multiple cloning site.

Preparation of Membrane Fractions

Completed by Dr. Karine Moncoq and Grant Kemp.

Membrane fractions were prepared immediately from fresh culture (Figure 2.7). All manipulations of cells, membrane fractions and purified protein were carried out at 4 °C, or on ice, using pre-chilled solutions. After induction, the cells were collected by centrifugation at 6000g for 15 min and were resuspended in 10 mM Tris pH 7.5 and centrifuged again. The cells were then resuspended in lysis buffer (25 mM HEPES-KOH pH 7.5, 1.5 M sorbitol, 1 mM EGTA, 1 mM EDTA, 1 mM dithiothreitol) supplemented with: 1 mM 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride[1], 2 μg/mL leupeptin, 2 μg/mL aprotinin, and 1 μg pepstatin at a concentration of 0.6 g wet cells/mL. Cells were disrupted using a microfluidizer (Emulsiflex C3, Avestin Inc.) with 7 passes at 20,000 psi. The lysate, crude extract (CE), was centrifuged at 2500g for 15 min to pellet unbroken cells, yielding supernatant (S1) and pellet (P1). Further usable sample was extracted from the cell debris by resuspending P1 in the same volume of lysis buffer used above and centrifuging again at 2500g for 15 min (P1'and S1'). S1 and S1' were combined and centrifuged at 12,000g for 15 min to remove the heavy membrane (P2: plasma membrane and mitochondria). Further contaminating heavy membrane was removed by decanting the supernatant, S2, into a clean tube and repeating the centrifugation (12,000g /15 min). This S2' supernatant was then centrifuged at 100,000g for 90 min. The final high speed pellet (P3) contained the membrane fractions enriched in hNHE1 (endoplasmic reticulum, Golgi and other intracellular trafficking membranes). The membrane fractions were resuspended in Storage Buffer (25 mM HEPES-KOH pH 7.5, 20% glycerol) at a protein concentration of ~12 mg/mL. The yield of membranes was about 300 mg of total protein per litre of culture. Membrane fractions were frozen in liquid nitrogen and stored at -80 °C until use.

Increasing Membrane Recovery

Completed by Grant Kemp.

To improve membrane recovery, the lysis protocol was later modified. This protocol has been adapted from the one published by Crotti *et al.* (27). Concentrations listed below as g/mL correspond to grams of cells per mL of buffer. "Harvesting" refers to centrifugation at 6000g

[1] Although similar protease inhibitor phenylmethylsulfonyl fluoride can be substituted (also at 1 mM), it is less water soluble and may form small crystals that can damage high pressure lysis machinery (e.g. Avestin Emulsiflex).

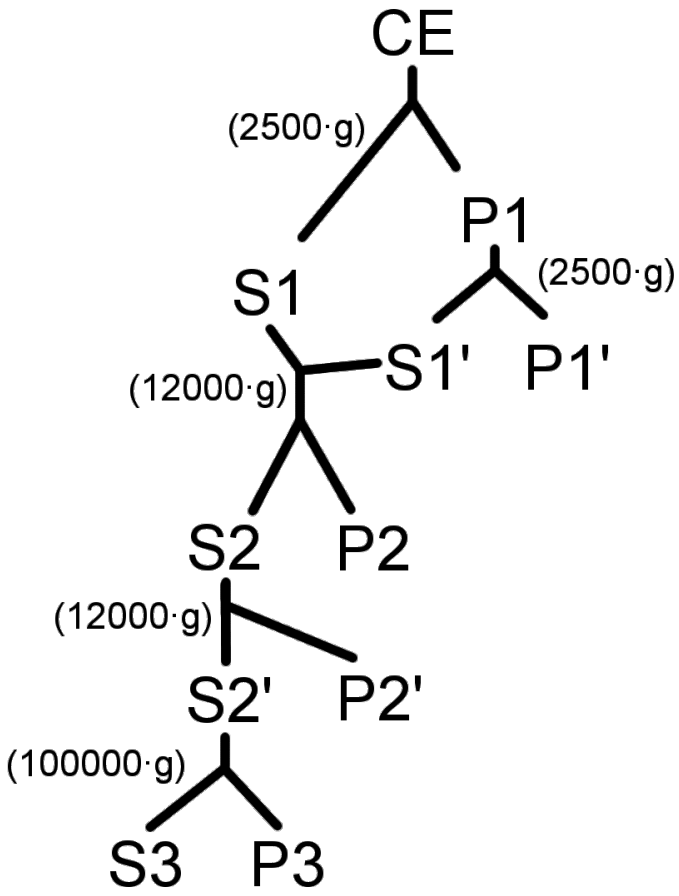

Figure 2.7: Schematic representation of the differential centrifugation scheme to isolate hNHE1 containing membranes. Following lysis the crude extract (CE) is centrifuged at 2500g . The pellet (P1) is resuspended in lysis buffer and centrifuged again. The supernatants (S1 and S1') are combined and centrifuged at 12,000g . S2 is centrifuged again to further remove contaminants. S2' is centrifuged at 100,000g to pellet the hNHE1 containing membranes.

for 15 min at 4 °C. After harvesting the cell pellet was resuspended in room temperature Pre-incubation Buffer (400 mM EDTA, 6.86% (v/v) β–mercaptoethanol) at 0.3 g/mL and placed in a shaking incubator (120 rpm) at 30 °C for 30 min. The cells were harvested and then resuspended in ice-cold Osmolysis Buffer (1 mM Citrate-Phosphate pH 6.3, 1 mM EDTA, 0.75 M sorbitol) at 0.15 g/mL and incubated on ice for 30 min. Following harvesting the cells were resuspended in ice-cold 0.05 M Tris pH 7.5 at 0.03 g/mL and allowed sit for 5 min at 4 °C. The cells were harvested and the lysis procedure continued as outlined above.

Solubilization and Purification of hNHE1

Completed by Dr. Karine Moncoq and Grant Kemp.

Prior to solubilization and hNHE1 purification, membrane fractions were stripped of peripheral membrane proteins by treating with high ionic strength solution. This step greatly enhanced the purity of the final protein. Frozen membrane fractions (at 12 mg/ml) were rapidly thawed to 4 °C and diluted 6 times in Stripping Buffer (60 mM Tris, pH 8.0, 12% glycerol, 0.7 M KCl)(B_{KCl}). The suspension was centrifuged at 100,000g for 90 min (S_{KCl}). The pellet was resuspended in Storage Buffer at ~4 mg/mL (P_{KCl}).

Alternatively, to decrease the time required for protein purification, a final concentration of 0.6 M KCl and 100 mM Tris pH 8.0 was added to supernatant S2' during membrane preparation (see above) before the final high speed centrifugation step. This stripped membrane pellet was then resuspended, diluted to ~12 mg/mL and flash frozen as above. To proceed with purification the membranes were rapidly thawed to 4 °C then diluted to 4 mg/mL with Storage Buffer and the procedure was continued at the solubilization step (see below).

NHE1 was solubilized from the stripped yeast membranes by mixing an equal volume of membrane (at 4 mg/ml) and Solubilization Buffer (25 mM HEPES-KOH pH 7.5, 20% glycerol, 100 mM NaCl, 5 mM imidazole, 1.2% FC-14 (w/v) or 1.2% LPC (w/v)). The suspension was incubated for 45 min with constant, gentle stirring. Insoluble material was removed by centrifugation at 100,000g for 45 min.

For purification of hNHE1, the supernatant from detergent solubilization was incubated in batch with 2.5 mL of Ni-NTA agarose resin equilibrated in Binding Buffer (25 mM HEPES-KOH pH 7.5, 20% glycerol, 150 mM NaCl, 20 mM imidazole, 0.1% DDM). The column was washed at 0.5 ml/min with four column volumes of Binding Buffer and then four volumes of Wash

Buffer (25 mM HEPES-KOH pH 7.5, 20% glycerol, 150 mM NaCl, 50 mM imidazole, 0.1% DDM). hNHE1 was eluted with Elution Buffer (25 mM HEPES-KOH pH 7.5, 20% glycerol, 150 mM NaCl, 250 mM imidazole, 0.075% DDM) at a flow rate of 0.2 ml/min. This purification procedure could be carried out using the detergents LPC, FC-14, or DDM (Figure 2.11). The hNHE1 resulting from this purification procedure was used for all structure and function studies described herein, with the exception of the circular dichroism (CD) spectroscopy studies.

For CD spectroscopy, the hNHE1 protein purified as described above, was further purified using Calmodulin (CaM) affinity resin. To reduce the imidazole concentration, the Ni-NTA eluate was dialyzed overnight against CaM Binding Buffer (25 mM HEPES-KOH 7.5, 20% glycerol, 100 mM NaCl, 4 mM $CaCl_2$, 0.05% FC-14).[2] The dialyzed suspension was loaded onto a 1 mL CaM-Sepharose column equilibrated with Binding Buffer at a flow rate of 0.5 ml/min. After washing with 3 column volumes of Binding Buffer, the protein was eluted with 2 volumes of CaM Elution Buffer (25 mM HEPES-KOH 7.5, 20% glycerol, 100 mM NaCl, 5 mM EGTA, 0.05% FC-14) at a flow rate of 0.2 ml/min.

Biophysical and Biochemical Characterization

Size Exclusion Chromatography

Completed by Dr. Karine Moncoq.

Size exclusion chromatography was used to examine the hydrodynamic volume of hNHE1. Samples of purified hNHE1 were run on a Superdex 200 HR 10/30 gel filtration column in 25 mM HEPES-KOH pH 7.5, 100 mM NaCl, 10% glycerol, 0.05% FC-14. The column was calibrated by measuring the elution volumes of proteins of known apparent molecular mass: thyroglobulin 679,000 (void volume); myosin, 200,000; bovine serum albumin, 66,000; carbonic anhydrase, 29,000; cytochrome c, 12,400.

Circular Dichroism Spectroscopy

Completed by Dr. Karine Moncoq and Grant Kemp.

For CD spectroscopy, purified hNHE1 protein was dialyzed in phosphate buffer (20 mM

[2]FC-14 was originally used as the detergent for solubilization and purification. Following optimization DDM became the detergent of choice for all purification steps while FC-14 continued to be used for solubilization. The CD and size exclusion experiments were completed before we switched to DDM.

phosphate pH 7.5, 10% glycerol, 0.05% FC-14) overnight. CD was performed at 20 °C on a Jasco J-500C spectropolarimeter (Jasco, Easton, MD) (28). CD spectra were recorded from 250 to 190 nm in quartz cells (path length of 0.05 cm) as the average of 8 scans at 0.1 nm intervals. Protein concentration was determined by amino acid analysis (0.85 μm), and the CD spectrum obtained in millidegrees was converted to molar ellipticity and analyzed using CDPro (29, 30).

Membrane Reconstitution

Completed by Dr. Karine Moncoq and Grant Kemp.

Reconstitution of hNHE1 into proteoliposomes follows established protocols (31). Briefly, 2.34 mg of egg yolk PC was dried to a thin film under nitrogen gas, and lyophilized for at least 2 h. Dried lipids were rehydrated with Reconstitution Buffer (20 mM BTP-MES pH 7.5, 25 mM $(NH_4)_2SO_4$, 10% glycerol) containing 2.5 mM pyranine (figure 2.8-A)[3] using constant vigourous Vortex mixing for 3x 2 min. The lipid was then solubilized by adding 20 μl of 20% n-octyl-d-glucoside followed by vigorous mixing. The best results were observed when 4x 5 μl additions of detergent were added interspersed with 2 min of vortex mixing. Following complete solubilization the mixture changed from cloudy to clear. Then 10 μg of purified hNHE1 was added giving a final volume of 220 μl. The solubilized protein/lipid/detergent mixture was applied to a 2 mL Sephadex G-50 fine column that was preloaded with Reconstitution Buffer containing pyranine. As the sample progressed through the resin small detergent monomers were removed resulting in hNHE1 reconstitution into EYPC vesicles (Figure 2.8-C). These large proteoliposomes, observed as cloudy elution fractions, were collected in the void volume and incubated for 30 min at room temperature with 100 mg of wet SM-2 Bio-Beads to remove residual detergent. The sample was applied to a 2 mL Sephadex G-50 fine column equilibrated with pyranine-free Reconstitution Buffer to remove external pyranine from the sample. The proteoliposomes were collected, observed as cloudy yellow fractions, and monitored for Na^+/H^+ exchanger activity via pyranine fluorescence.

Alternatively, reconstitution can be carried out by taking the solubilized lipid, hNHE1 mixture from above (immediately before the first G-50 column), adding 2.5 mM pyranine and incubating it with 100 mg of SM-2 Biobeads with gentle stirring at 4 °C for 1h. The external pyranine is removed by carefully pipetting the reconstitution (using a thin gel-loading tip to avoid taking

[3] fluorescent pH indicator trisodium 8-hydroxypyrene-1,3,6-trisulfonate

Biobeads) into a small centrifuge tube. The proteoliposomes are pelleted at 41,000g /4 °C for 30 min. The supernatant is carefully removed, the proteoliposomes are resuspended in pyranine-free Reconstitution Buffer and the centrifugation is repeated.

To measure NHE1 activity proteoliposomes (100 µl) were diluted into 2 mL of Reaction Buffer (20 mM BTP-MES pH 7.5, 10% glycerol) and incubated at 25 °C. The absence of ammonium in the Reaction Buffer generates a pH gradient—large intravesicular [H^+]—by uncharged ammonia quickly diffusing out of the proteoliposomes.

$$NH_4^+ \rightleftharpoons NH_3 + H^+ \xrightarrow{\text{dilution into Reaction Buffer}} NH_4^+ \rightleftharpoons NH_3\uparrow + H^+$$

Na^+/H^+ exchange was initiated by NaCl addition and activity monitored by a change in pyranine fluorescence in a PTI Deltascan spectrofluorometer (λ_{ex}=463 nm and λ_{em}=510 nm[4]). The relative hNHE1 activities were measured by regression fitting of a 20 s linear region of the fluorescence intensity curves. For instance, the fluorescence curve in Figure 2.14 was fit from 40 to 60 s yielding a change in fluorescence intensity of 7400 per second (set to 100% relative hNHE1 activity). Empty liposomes (prepared in the same fashion except without the addition of hNHE1) were assayed in the same way as a control for vesicle leakiness. To determine the specificity of the Na^+/H^+ exchange, proteoliposomes were treated with the specific hNHE1 inhibitor EMD87580 (Figure 2.8-B) (32).

Single Particle Electron Microscopy

Completed by Dr. Karine Moncoq.

Purified hNHE1 was diluted with 1% trehalose, 3% ammonium molybdate (pH 7.0) from an initial concentration of 0.2 mg/mL to a concentration of 0.01 mg/mL. Immediately following dilution, five microlitres were pipetted onto a glow-discharged, carbon-coated grid for 30 s. The grid was washed with one drop of 2% uranyl acetate, and then allowed to sit with a drop of 2% uranyl acetate for 1 min. The excess stain was blotted with filter paper, and the grid was allowed to air dry. Data were collected on a Tecnai F20 (FEI Company) located in the Microscopy and Imaging Facility at the University of Calgary (Calgary, Alberta, Canada). The microscope was operated at 200 keV and images were recorded on Kodak SO-163 film under low-dose conditions at a magnification of 50,000 with a defocus ranging from -2.0 µm to -2.5 µm. Micrographs

[4] λ_{ex} is the excitation wavelength and λ_{em} is the emission wavelength

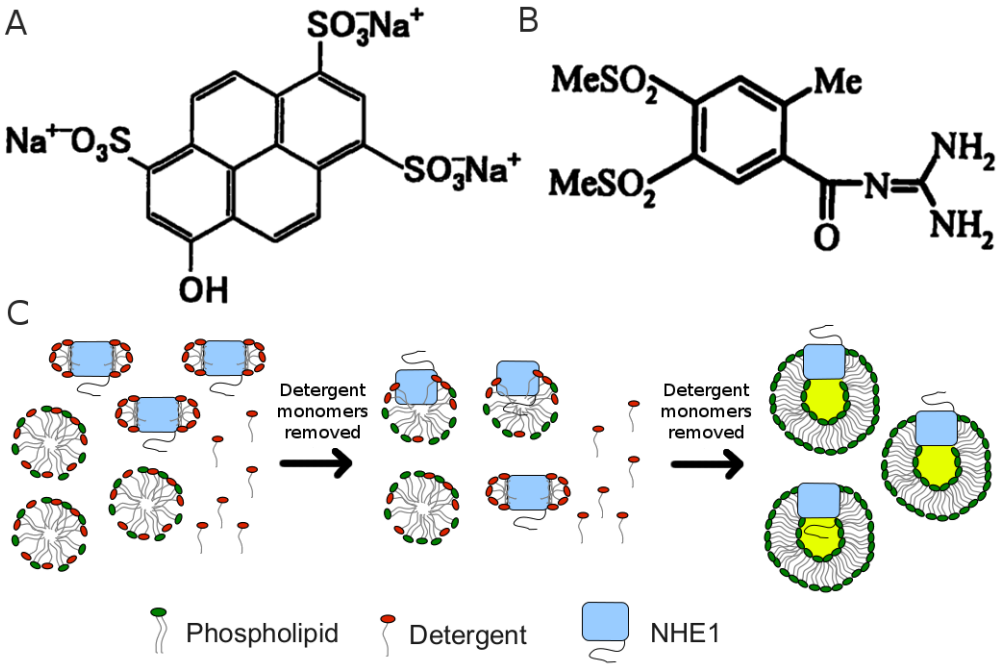

Figure 2.8: Reconstitution schematic and structures of pyranine and EMD87580. ***A***. Structure of pH sensitive fluorophore pyranine (trisodium 8-hydroxypyrene-1,3,6-trisulfonate), containing three protonatable sulphonic acid groups. When fully protonated (low pH) pyranine fluorescence is low. As each sulphonic acid becomes deprotonated the fluorescence increases. ***B***. Structure of EMD87580, belonging to the 6-membered heteroaroylguanidine class (amiloride family). ***C***. Schematic representation of proteoliposome reconstitution. Beginning on the left: purified hNHE1 in detergent is mixed with detergent solubilized EYPC lipid; during gel filtration or biobead treatment, detergent monomers are removed leading to association of lipid and hNHE1; following complete detergent removal hNHE1 has become incorporated into proteoliposomes. If performed in the presence of pyranine, the sealed proteoliposomes retain trapped pyranine (represented by a yellow interior). External pyranine is removed by gel fitration or centrifugation.

were digitized with a Nikon Super Coolscan 9000 with a scanning resolution of 6.35 μm/pixel, followed by pixel averaging to achieve a final resolution of 5.08 Å/pixel.

Image processing and reconstruction were carried out using both EMAN (33) and SPIDER (34) software packages. Using EMAN's boxer, 7930 particles were selected semi-automatically from 24 micrographs with a box size of 50 x 50 pixels. The boxed images were corrected for the CTF using ctfit (EMAN). Reference-free classification into 135 groups proceeded using startnr-classes (EMAN) and AP CA (SPIDER). The class averages generated by EMAN and SPIDER were combined into a common set of 44 groups using classesbymra (EMAN). A set of Euler angles was then assigned to these class averages (startAny command in EMAN; OP command in SPIDER), and initial three-dimensional models were built using common lines in Fourier space. These models were low pass filtered to 20 Å resolution, aligned and averaged (align3d and avg3d commands in EMAN). The average was taken as a preliminary model for five iterations of refinement in EMAN and SPIDER. A resolution of 22 Å of the final reconstruction was determined by calculating the Fourier shell correlation between two independent half datasets (eotest command in EMAN; 0.5 FSC criterion). Comparison of the final reconstruction with the predicted molecular mass for recombinant hNHE1 (92 kDa) indicated that the structure represented an hNHE1 dimer. Twofold symmetry was apparent and was applied to the final reconstruction (align3dsym and proc3d commands in EMAN).

SDS-PAGE, Western blotting and Protein Assay

Completed by Dr. Karine Moncoq and Grant Kemp.

Protein samples were heated at 65 °C for 3 min and were separated on 10% SDS-polyacrylamide gels as described (35). For Western blot analysis, SDS-PAGE gels were transferred onto nitrocellulose membranes, and detection of recombinant hNHE1 was with HisProbe-HRP using the Supersignal West HisProbe Kit (Pierce) according to the manufacturer's instructions. Alternatively, a monoclonal anti-NHE1 antibody was used as the primary antibody as described earlier (35) and peroxidase-conjugated goat anti-mouse antibody (Pierce) was used as a secondary antibody. X-ray films were digitized with an Epson (Toronto, ON) Perfection 3200 densitometer and bands were quantified using ImageQuant software (GE Healthcare Life Sciences).

Protein concentrations were determined for yeast membrane preparations using a Lowry assay and for purified protein using an Amido Black assay (36, 37) with bovine serum albumin as

standard.

Results

Expression of NHE1 constructs in yeast

To obtain suitable amounts of functional protein for structural studies we expressed full-length hNHE1 using *S. cerevisiae*, a proven and efficient heterologous expression system. To avoid potential complications that may arise during crystallization due to protein glycosylation, the known N-linked glycosylation site at Asn75 (38) was mutated to aspartate. This glycosylation site, along with O-linked glycosylation occurring within the first extracellular loop, have been shown previously to be non-essential for either cation exchange activity or hNHE1 biosynthesis (38, 39). In addition, a His-tag with a flexible Gly-Gly linker (Gly_2His_{10}) was introduced at the C terminus of hNHE1 to facilitate protein purification via immobilized metal affinity chromatography (IMAC). Thus, a His-tagged hNHE1-N75D was inserted into the expression vector pYeDP60 (40) under the control of a galactose inducible promoter. This expression system was originally developed for the expression of cytochrome P450s (40, 41) and has been shown to be suitable for the heterologous expression of rabbit SERCA1a Ca^{2+}-ATPase (42), and its subsequent crystallization and structure determination (43), and the expression and crystallization of plant plasma membrane H^+ ATPase (44).

To improve hNHE1 expression, purity, stability, and possibly crystallizability, a human hNHE1 transmembrane domain construct (hTMD) was developed (Figure 2.1). Although optimizing the protein sequence of a target in order to increase stability and crystallizability is an important step, a major bottleneck for obtaining large eukaryotic membrane proteins structure is expression levels (45). While a low expression level can be countered to some extent by making larger preparations or pooling multiple preparations of protein, this is ultimately limited by size and cost and may lead to heterogeneity in detergent concentration, lipid content, and protein aggregation. By increasing the initial protein expression level we aimed to minimize these effects. Additionally, the carboxy-terminal regulatory domain of hNHE1 is known to contain regions of inherent flexibility (46, 47) that lead to *in vitro* degradation (*unpublished observation*). It has even been postulated that some regions are intrinsically disordered and that this flexibility is required for

proper regulation of transport activity (48). Employing this hTMD construct was hypothesized to allow us to probe the structure of the membrane domain while avoiding these caveats.

There has long been evidence that the codon composition of highly expressed genes is biased (49, 50). In fact an early survey of several highly expressed proteins showed that only ⅓ of the possible codons were present in these proteins whereas proteins expressed at lower levels contained all the possible codons (51). This bias was linked to the relative cellular abundance of various tRNAs (52), thus allowing the organism to efficiently translate highly expressed proteins during rapid growth or environmental challenge without needing large amounts of all tRNAs (53). Immediately researchers became interested in how to take advantage of this bias to increase the expression of heterologous proteins in yeast (20, 54, 55) with some success (56, 57). We applied genome-wide codon usage information (Table 2.1) to make an *optimized* hTMD construct. Additionally, very recently, a novel way of thinking about codon bias was developed by the laboratory of Dr. Ina Urbatsch that has been coined *harmonization* (18). While *optimization* involves using codons that appear more often in the genome as a whole, *harmonization* identifies codons that are most often used in highly expressed proteins. Interestingly several rare codons are actually preferred in highly expressed proteins and a few more commonly used codons are nearly excluded (highlighted in Table 2.2). Applying this codon usage information (very generously shared with us by Dr. Ina Urbatsch) we designed the mNHE1 construct.

Since the final protein used for crystallization must by very pure and homogeneous, additional features were incorporated into these new constructs that may facilitate these aims. A tobacco etch virus protease (TEV) site and a high-affinity calmodulin (CaM) tag were incorporated upstream of the polyhistidine tag on mNHE1 that would allow both tag removal—to decrease flexibility—and CaM-Sepharose affinity purification, if required. The sequence chosen for CaM binding was from the Strategene pCal vector:

KRRWKKNFIAVSAANRFKKISSSGAL

In hTMD an alternative strategy using a *tag cloning site* was engineered between a TEV and polyhistidine tag. This included two unique restriction enzyme sites, *SalI* and *MluI*, that would allow the insertion of novel purification tags if required.

Alongside the optimization and harmonization of hTMD and mNHE1, we also assayed different expression hosts and vectors for increased expression levels. Yeast is a useful heterologous

expression system for eukaryotic membrane proteins because it has the machinery required for proper folding, membrane insertion and secretion. In the past decade *P. pastoris* has become a very popular system for the production of recombinant proteins (reviewed (56, 58)) due to its relative ease and low cost. Several studies have shown that many eukaryotic proteins expressed in bacteria become trapped in inclusion bodies and require delicate refolding techniques that can be laborious to optimize. An excellent example of this comes from Lueking *et al.* (59) who took 29 cDNA clones from a human fetal brain library and expressed them in a dual *E. coli* and *P. pastoris* expression vector. In *E. coli* five of the proteins did not express at all and 15 were detected in inclusion bodies, whereas all the proteins were expressed as soluble proteins in yeast. As *S. cerevisiae* was able to successfully produce functional hNHE1, and even greater success has been reported in the literature with *P. pastoris*, we cloned mNHE1 into a compatible vector for expression.

The hTMD construct, cloned into pYeDP60, was assayed for protein expression in a similar manner to hNHE1. While yeast transformants could be attained, although with more difficulty, no significant amount of protein expression could be detected by Western blot.

mNHE1, cloned into pPICZ-A, was recombined into GS115 and X33 strains of *P. pastoris*. They were then screened for copy multiplicity by plating on increasing concentrations of Zeocin (Figure 2.9). It was noted in the EasySelect manual (Invitrogen) that these colonies should have the highest expression levels. Expression tests for seven colonies that showed a high degree of insertion (based on growth in high concentrations of Zeocin) were carried out. Various types of media were assayed for growth and induction conditions. No significant expression could be detected by Western blot. Due to time constraints of my PhD thesis, these trials were not investigated further.

Concurrently to *P. pastoris* expression tests, we assayed other *S. cereviae* plasmids that have been successful for membrane protein expression. The p423-GAL1 system, developed by Robert Stroud's laboratory (60–62) based on previous work on the GAL1 promoter (63), has resulted in two eukaryotic membrane protein crystal structures (64, 65). Additionally, pYES-DEST52 from the Gateway® Cloning system (Invitrogen) has been employed to produce membrane proteins in *S. cerevisiae* that could be purified and studied *in vitro* (66, 67). Although successful clones were obtained, limited advancements in expression tests and time limitations prevented further exploration of these systems.

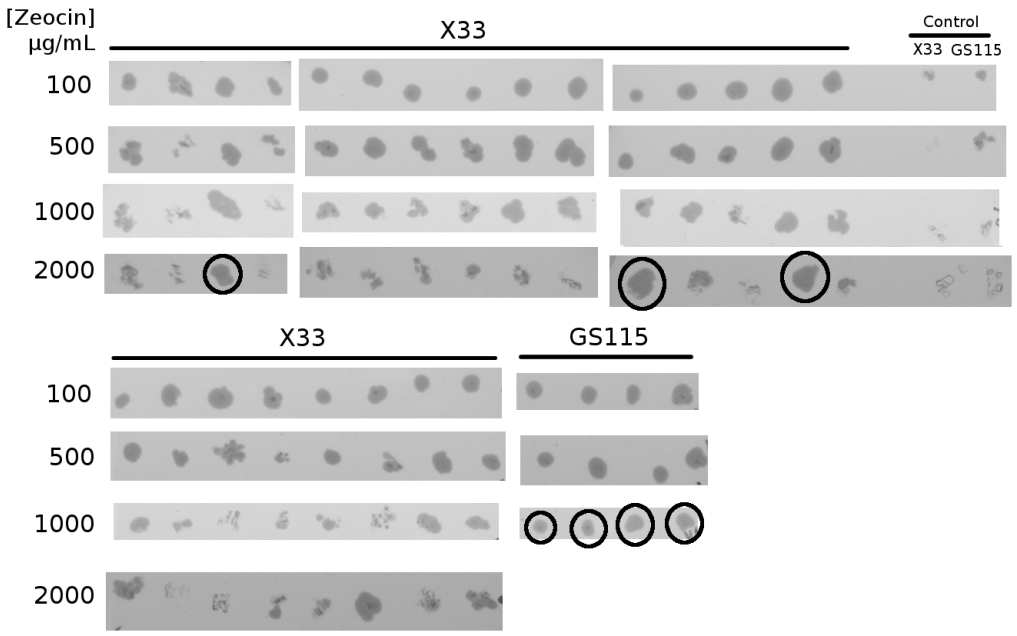

Figure 2.9: Measuring recombination multiplicity of successful *P. pastoris* transformants. Strips from YPDS plates spotted with positive GS115 and X33 transformants (NB: fewer positive GS115 colonies were obtained during transformation) and a single untransformed control colony. At low concentration of Zeocin (100 µg/mL) all transformants, and to a lesser degree the control colonies, are able to grow. As the concentration of Zeocin is increased fewer transformants remain. The transformants that were selected for protein expression trials are circled.

While many of the above strategies were ineffective, some success was achieved with other modifications to the hNHE1 expression and purification scheme. First, expression in *S. cerevisiae* was improved by using the strain DSY864 (also known as NKY879 (16)). This improved expression by roughly 10 fold (Figure 2.10-A). It was also observed that the expression began soon after induction and was sustained for at least 22 h (Figure 2.10-B).

To further increase the cell mass obtained, we used a fermentor. Following standard protocols using rich (YPDE-2) media, large quantities of yeast were easily obtained. However, while cells continued to grow to OD_{600} >50 the plasmid pYeDP60 was lost as indicated by the pinkish colour of the pellet, which is a phenotype of cells lacking *ade2*. This was confirmed as no expression was visible by Western blot. In an attempt to circumvent this, selective media was also assayed. Similar to flask grown cultures, where no hNHE1 expression was found in SDaa, this yielded poor cell mass and no expression. This is a common problem observed for plasmid-born expression of proteins in *S. cerevisiae* fermentor cultures. The key to increased cell density in fermentor culture is aeration. In standard Erlenmyer flasks a ratio of liquid to air of at least 1:4 is recommended during growth in a shaker. This prevents anaerobic growth by ensuring the culture remains in good contact with air during growth. In a fermentor there is mechanical mixing and air infusion during growth increasing the available oxygen. As a compromise we began using Tunair® flasks. They are baffled plastic flasks that have an open top that protects the culture from contamination using a filter. Under normal flask conditions cultures may reach a maximun OD_{600} of 15-20 (compared to >50 in the fermentor). In the Tunair® we were routinely able to achieve densities of 30-35 without losing expression; a reasonable compromise.

Membrane Isolation

A differential centrifugation procedure was used to prepare membrane fractions that were enriched in hNHE1 (Figure 2.7). Western blot analysis of the fractionated membranes illustrated the profile of hNHE1 yield in the various fractions (Figure 2.10-D). The expressed recombinant hNHE1 migrated on SDS-PAGE slightly lower than the 100 kDa standard, which is in good agreement with the predicted molecular weight of 92 kDa for the unglycosylated protein. The P3 light membranes (corresponding to endoplasmic reticulum and secretion vesicles) contained most of the hNHE1 protein, while the heavy P2 membranes (plasma membranes and mitochondria) contained much less hNHE1. From 1 L of yeast culture, 20 mL of P3 membrane fraction

were obtained at a protein concentration of 15 mg/mL. We used the P3 membrane fraction for all subsequent solubilization and purification steps.

To further improve our protein yield, we investigated whether the amount of hNHE1 lost immediately following lysis (P1) could be decreased. As the protein lost at this step is primarily contained in unlysed cells the goal was to increase the efficiency of lysis. Unlike bacteria, yeast have a tough cell wall that is not easily broken during lysis. Originally yeast were lysed mechanically by vortexing with fine glass beads, which was only modestly effective. This was improved by using a high-pressure lysis method (Emulsiflex), which more effectively pulverized the hardy cell wall (data not shown). However, there was still a significant portion of cells that remained unlysed. Crotti *et al.* (27) demonstrated that yeast could be effectively permeabilized using an osmotic shock step, greatly abating the integrity of the cell wall. Our application of this method led to a two-fold increase in the amount of membrane recovered following cell lysis.

Membrane Stripping and hNHE1 Solubilization

Prior to solubilization we tested the effect of high KCl treatment to remove peripheral membrane proteins and to enrich the membrane fractions for the hNHE1 protein. The incubation of P3 membranes (B_{KCl}) with 0.7 M KCl (or NaCl) removed up to 70% of the protein content (Figure 2.12-A) while not more than 20% of the hNHE1 was lost to the supernatant during this step (Figure 2.12-B). This led to an approximate 2-fold enrichment of hNHE1, and this step proved to be invaluable in improving the subsequent purification by immobilized metal ion affinity chromatography (Figure 2.12-D). When this step was omitted, we observed that detergent solubilized-NHE1 eluted from the Ni-NTA agarose at low concentrations of imidazole (50 mM). With the inclusion of this step, hNHE1 was retained on the column after low concentration imidazole washes (50 mM) and eluted at high imidazole concentration (250 mM). Therefore, P3 membranes were routinely treated with either KCl or NaCl before solubilization and purification.

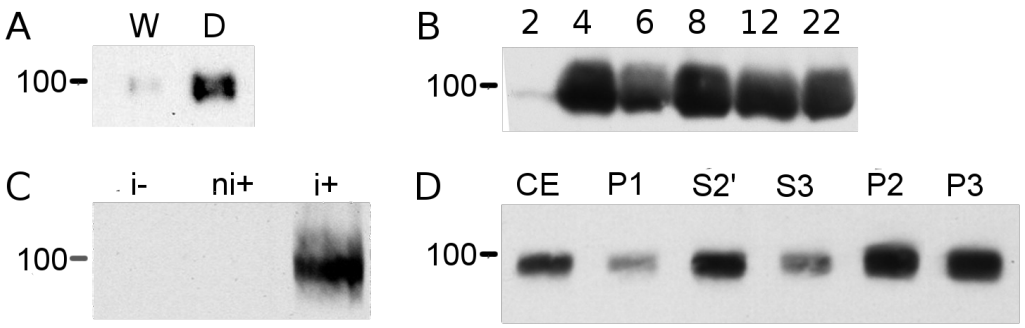

Figure 2.10: Western Blots showing hNHE1 expression and membrane isolation. Proteins were transferred from 10% SDS-PAGE gels onto nitrocellulose membranes and detected with His-Probe HRP. ***A***. Western blot showing the comparison of expression level of *S. cerevisiae* strains W303.1b (W) to DSY864 (D). ***B***. Time course of hNHE1 expression in DSY864 (lanes are labelled with hours following induction) indicating a fast and sustained response to galactose. ***C***. Western blot indicating the specific inducible expression of hNHE1 in W303.1b: (i-) W303.1b transformed with empty pYeDP60 plasmid and induced with 2% galactose; (ni+) W303.1b transformed with pYeDP60-NHE1 but not induced; (i+) W303.1b transformed with pYeDP60-NHE1 and induced. ***D***. Western blot of several of the fractions outlined in Figure 2.7. Each lane contains 1.5 μg of total protein.

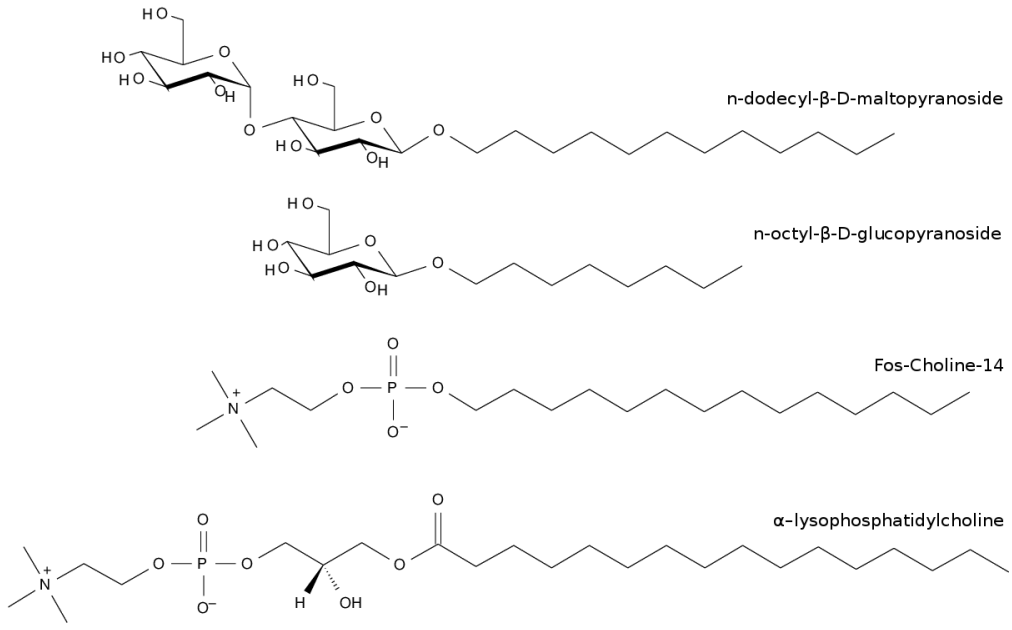

Figure 2.11: Structures of detergents used for hNHE1 solubilization, purification and reconstitution. n-dodecyl-β-D-maltopyranoside (DDM) was used for affinity chromatography purification; n-octyl-β-D-glucopyranoside (OG) was used to solubilize lipid for hNHE1 proteoliposome reconstitution; Fos-Choline-14 (FC-14) was used to solubilize hNHE1 from yeast membranes; α-lysophosphatidylcholine was used as an alternative detergent for yeast membrane solubilization.

Detergents are central to the isolation and solubilization of membrane proteins, yet the selection process remains somewhat empirical (68). The efficacy of detergents varies with individual proteins and selection of a suitable detergent is important in determining protein yield and in preserving protein function. Consequently a variety of detergents (LPC, $C_{12}E_8$, DDM, FC-10, FC-12, and FC-14) were tested for their ability to solubilize hNHE1 from stripped P3 membranes (Figure 2.12-C). To increase the ease of membrane solubilization we began with homogenized membrane resuspensions at 4 mg/mL that were diluted two-fold in solubilization buffer containing 1.2% detergent—the final detergent:protein (w/w) ratio was 3:1. Beginning with higher protein concentrations or using a lower detergent:protein ratio decreased hNHE1 recovery whereas increasing the detergent:protein ratio did not significantly improve recovery. The recovery of hNHE1 from the membrane was estimated by Western blot analysis from aliquots withdrawn before and after centrifugation, detected with HisProbe-HRP. The most efficient solubilization of the hNHE1 protein occurred with the detergents LPC, FC-10, FC-12, and FC-14. The maximum efficiency was 61 and 92% with LPC and FC-14, respectively, and increasing the detergent/protein ratio did not enhance solubilization. Interestingly, increasing the acyl chain length of Fos-choline detergents enhanced solubilization, suggesting that hydrophobic matching may be important in the solubilization of hNHE1. Thus, LPC and FC-14 were suitable detergents that could be used interchangeably in the solubilization and functional preservation of recombinant hNHE1 (Figure 2.11).

Purification of Recombinant hNHE1

Detergent-solubilized membranes were applied to Ni-NTA agarose, followed by SDS-PAGE and Western blotting to monitor the purification process (Figure 2.13-A & B). The majority of hNHE1 protein was retained on the column and washing the column with either 20 or 50 mM imidazole only removed trace amounts of the hNHE1. Fractions enriched in hNHE1 eluted at 250 mM imidazole with a purity of ~70% (Figure 2.13-A, lane E). The purification achieved approximately a 100-fold enrichment for hNHE1 following a single cycle of binding and elution from the Ni-NTA agarose. The protein yield after this chromatography step was 0.5 mg/L (~0.4 mg/L hNHE1) of culture according to quantification by Amido Black assay and densitometry of Coomassie Blue-stained gels. Further purification of hNHE1 was achieved via calmodulin affinity chromatography. This was designed to take advantage of the presence of the endogenous

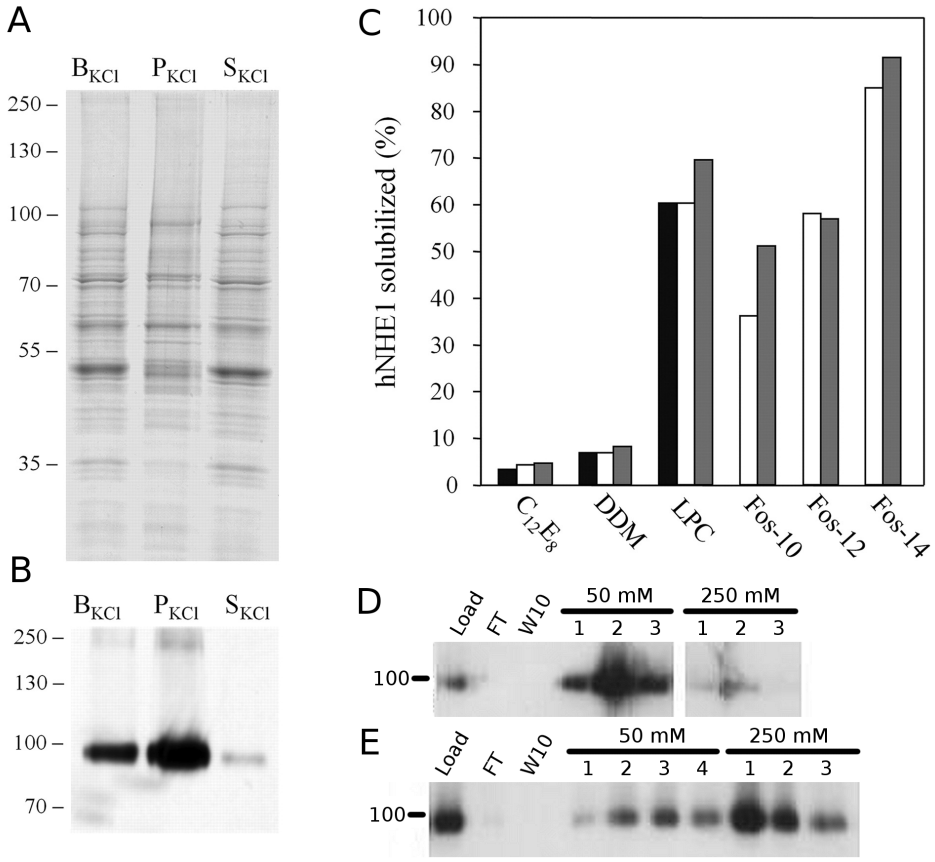

Figure 2.12: Analysis of hNHE1 recovery following membrane stripping and solubilization. All Western blots were from gels transferred to nitrocellulose and hNHE1 was detected with HisProbe-HRP. ***A***. 7.5% SDS-PAGE gel of samples taken during the membrane stripping process. B_{KCl}, total membrane fraction before centrifugation; P_{KCl}, high speed centrifugation pellet resuspended in an equivolume of Resuspension Buffer; S_{KCl}, high speed centrifugation supernatant. ***B***. Western blot of the same fractions as in ***A***. ***C***. Graph depicting the effectiveness of hNHE1 recovery to the supernatant following membrane solubilization (measured by Western blot quantification using GelQuant). ***D***. Western blot showing the Ni-NTA resin affinity of hNHE1 solubilized from unstripped membrane. Load, unstripped solubilized membranes; FT, Ni-NTA unbound fraction; W10, 10 mM imidazole wash fraction; 50 mM, 50 mM imidazole elution fractions; 250 mM, 250 mM imidazole elution fractions. NB: 10 mM imidazole washes were used during protocol development and was increased to 20 mM with optimization. ***E***. Replication of ***D*** using hNHE1 solubilized from stripped membranes.

calmodulin binding site present on the C-terminal tail of hNHE1 (69). In this procedure, the Ni-NTA fractions most enriched in hNHE1 were pooled and incubated with calmodulin-agarose in the presence of calcium, followed by elution with calcium-free buffer. This second purification resulted in removal of many contaminants and produced a more homogeneous protein (≥90%; Figure 2.13-C, lane E), based on quantification of Coomassie Blue-stained gels. However, the purification procedure was not routinely used because relatively low amounts of purified protein (~0.1 mg/L) were recovered.

Finally, we wished to examine the glycosylation state of the yeast-expressed hNHE1 by comparing it with NHE1 expressed in mammalian AP1 cells. Our recombinant hNHE1 possessed a similar molecular mass to the un- or partially glycosylated form of NHE1 in AP1 cells (lower bands, Figure 2.13-D) (1). This is consistent with the N75D mutation removing the site of N-linked glycosylation from the yeast-expressed hNHE1.

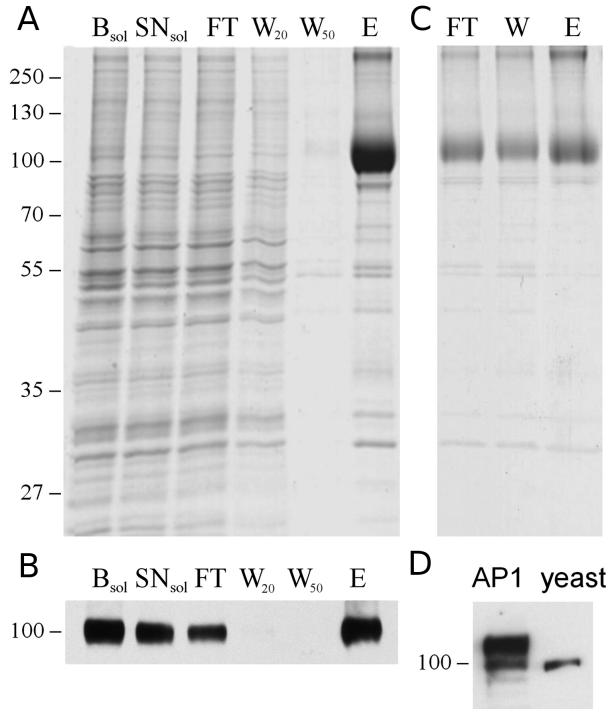

Figure 2.13: SDS-PAGE and Western blot analysis of purified hNHE1. ***A***. 7.5% SDS-PAGE gel showing fractions from Ni-NTA affinity chromatography. B_{sol}, stripped membrane fraction before solubilization; SN_{sol}, supernatant following solubilization; FT, Ni-NTA unbound fraction; W_{20}, 20 mM imidazole wash fraction; W_{50}, 50 mM imidazole wash fraction; E, pooled elution fraction. ***B***. Western blot of the same fractions as in ***A***. ***C***. 7.5% SDS-PAGE gel showing fractions from CaM affinity chromatography. FT, unbound fraction; W, wash fraction; E, pooled elution fraction. ***D***. Western blot for hNHE1 expressed in either mammalian AP1 cells (AP1) (1) or w303.1b *S. cerevisiae* (yeast). Two bands in AP1 cells indicate a mature glycosylated form (upper band) and an immature unglycosylated form (lower band). Yeast expressed hNHE1 (N75D) exists only as the unglycosylated form.

Na^+/H^+ Exchange Activity

To confirm that recombinant hNHE1 was functional and could conduct cation exchange, purified hNHE1 was reconstituted into phosphatidylcholine vesicles in the presence of ammonium (NH_4^+) ions and the fluorescent pH indicator pyranine. Pyranine fluorescence has been shown to directly reflect the intravesicular pH (70, 71). Dilution of the proteoliposomes into ammonium-free buffer resulted in acid loading of the vesicles due to outward diffusion of ammonia (NH_3). Efflux of intra-vesicular H^+ in exchange for extravesicular Na^+ was monitored by the increase of pyranine fluorescence upon addition of NaCl. The activity assays showed that hNHE1-containing proteoliposomes mediated rapid cation exchange, compared with a low background of cation exchange by vesicles lacking NHE1 (Figure 2.14-A). The addition of $(NH_4)_2SO_4$ completely collapsed the cation gradient across the vesicle membranes and resulted in full recovery of pyranine fluorescence. When vesicles were preincubated with the amiloride analog EMD87580, a potent specific NHE1 inhibitor (2), cation exchange activity of the hNHE1 proteoliposomes was inhibited in a concentration-dependent manner (Figure 2.14-B). The measured IC_{50} value for hNHE1 inhibition by EMD87580 was 5 μM (Figure 2.14-C), which is similar to the value reported for wild-type NHE1 expressed in mammalian cells (2).

We also measured the cation dependence of transport to further confirm the identity of our recombinant protein. NHE activity was measured over a range of cation concentrations (5–150 mM NaCl, LiCl, and KCl) in the absence and presence of the specific inhibitor EMD87580. The NHE1-specific transport was then reflected in the difference between transport in the absence and presence of the inhibitor. NHE1 carries out sodium- and lithium-stimulated proton transport, but potassium does not compete with sodium for transport and is not an appropriate substrate for NHE1 (72, 73). Our recombinant hNHE1 exhibited sodium-dependent transport that increased rapidly with NaCl concentration and saturated at 50–75 mM NaCl (no observable transport at 5 mM NaCl; 30% transport activity at 25 mM NaCl; 100% transport activity at 75 mM NaCl). Lithium also stimulated transport activity, but the concentration dependence was not characterized. By comparison, there was no measurable potassium-stimulated transport. These results indicate that our recombinant hNHE1 is fully functional, with an inhibitor sensitivity and cation-dependence that is characteristic of this exchanger.

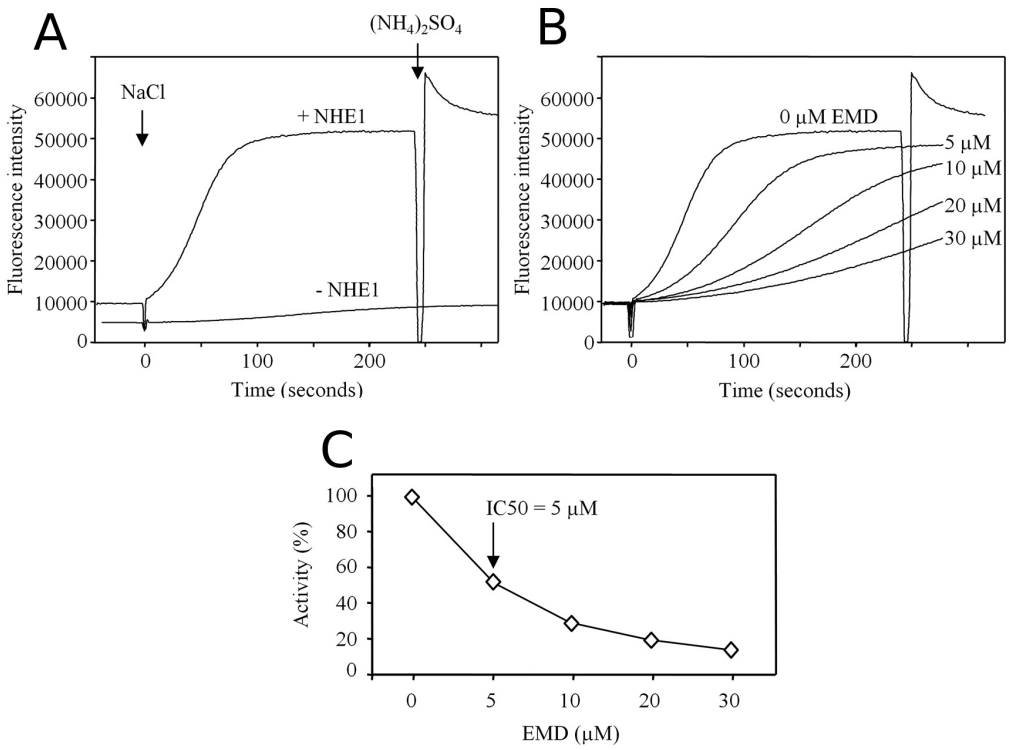

Figure 2.14: hNHE1 activity assay curves showing pyranine fluorescence change over time upon hNHE1 activity stimulation. ***A***. Representative curves showing pyranine fluorescence ($[H^+]$ over time). Proteoliposomes containing hNHE1 (+NHE1) and empty liposomes (-NHE1) were diluted into Reaction Buffer inducing an acid load (low pyranine fluorescence). After a stable signal was acquired NaCl was added to initiate Na^+/H^+ exchange (t=0, indicated). Proton efflux increased trapped pyranine fluorescence. Following maximum fluorescence the remaining proton gradient was collapsed by $(NH_4)_2SO_4$ addition (indicated). ***B***. hNHE1 activity was confirmed by repeating the ***A*** protocol following the pre-incubation of proteoliposomes with increasing concentrations of NHE1-specific inhibitor EMD87580. ***C***. The [EMD87580] at which 50% of hNHE1 is inhibited (IC_{50}) was determined graphically (comparing the relative NHE activity versus the [EMD87580]) to be 5 μM.

Biophysical and Biochemical Characterization

The oligomeric state of purified hNHE1 was analyzed by size-exclusion chromatography on a Sephadex 200 column (Figure 2.15). Purified hNHE1 in detergent (FC-14) was applied to a column pre-equilibrated with the same detergent. The elution profile consisted of two main peaks: a small peak running with the void volume and a large peak with a retention time of 23.2 min indicating a molecular mass in slight excess of 200 kDa. Each peak contained hNHE1 as determined by Western blot analysis (data not shown). Including the FC-14 detergent micelle (108 molecules; aggregation number of FC-14 obtained from Anatrace, Maumee, OH), the approximate molecular mass for an hNHE1 dimer is 233 kDa, consistent with the larger second peak observed in the chromatogram (arrow in Figure 2.15).

To gain insights into the secondary structure of purified hNHE1 we used CD spectroscopy. The CD spectral profile (Figure 2.16) showed two local minima at 208 and 222 nm characteristic of a high α-helical content (74). Analysis of the spectrum indicated that the recombinant hNHE1 contains 41% α-helix, 23% β-sheet, and 36% random coil. These values are consistent with topological predictions of NHE1 that suggest 12 transmembrane helices (13, 14).

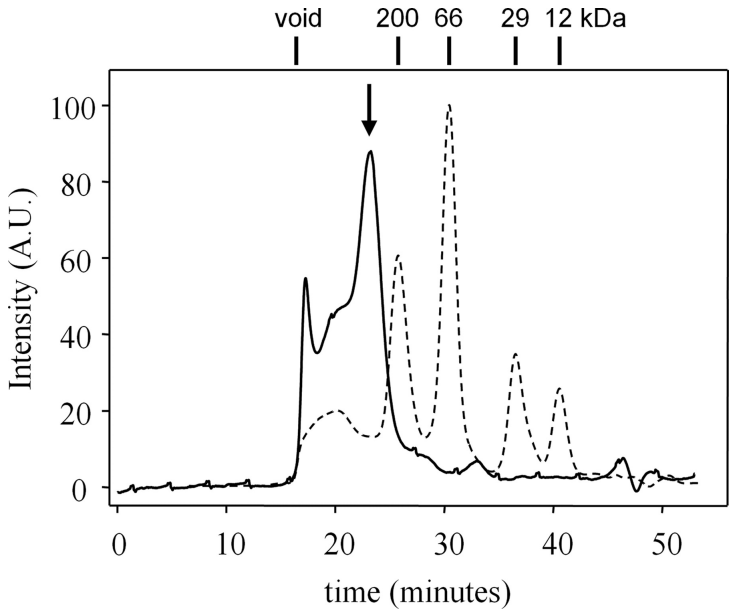

Figure 2.15: Size excusion chromatography of Ni-NTA purified hNHE1. Following affinity purification and protein concentration (see Methods), hNHE1 was loaded onto a Superdex 200 HR 10/30 gel filtration column to determine it's hydrodynamic volume. The solid line shows the A_{280} (y-axis) of the loaded hNHE1 sample over time (x-axis). An arrow indicates the retardation time of the hNHE1 dimer. The dashed line follows the A_{280} of various protein standards, weights (kDa) are indicated above the chromatogram.

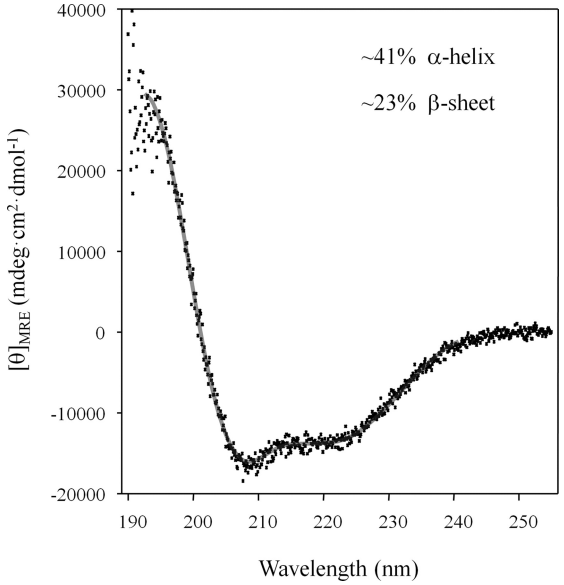

Figure 2.16: Circular dichroism (CD) spectroscopy of Calmodulin-affinity purified hNHE1. The far-UV CD spectrum of hNHE1 following Ni-NTA and calmodulin affinity purification and protein dialysis (see Methods).

Electron Microscopy and Single Particle Reconstruction of hNHE1

Electron micrographs of negatively stained hNHE1 revealed a homogeneous and uniform distribution of particles that were ~100 Å in diameter (Figure 2.17-A). Particle selection yielded 7,930 projections that were grouped by reference-free alignment and classification. Some class averages exhibited 2-fold symmetry (Figure 2.17-B). Fourier common lines approaches implemented in both EMAN and SPIDER were utilized to determine the relative orientations of 44 class averages, which were then combined to generate an initial three-dimensional model. This initial model was used as a starting reference for subsequent rounds of alignment and classification, followed by reference-free orientation determination in EMAN and SPIDER. The final three-dimensional model after imposing 2-fold symmetry (Figure 2.18) indicated a resolution of 22 Å, with a fairly even distribution of particle orientations used in the reconstruction (data not shown). The volume shown in Figure 2.18 (4σ density cutoff) corresponds to a 184-kDa protein, which is exactly twice the mass expected for an hNHE1 monomer. Thus, the reconstruction reveals that recombinant hNHE1 is a dimer. This size is consistent with our predictions based on size-exclusion chromatography (Figure 2.15), as well as previous studies suggesting that NHE1 is a dimer (46, 75, 76).

The hNHE1 reconstruction can be divided into two regions: a compact, globular domain that we assign to the transmembrane domain of hNHE1 and an apical ridge that we assign to the cytoplasmic domain of hNHE1. The relative size of these regions is consistent with their expected molecular mass (58 and 34 kDa, respectively). At higher density thresholds (4.8σ), the globular region of the reconstruction splits into two subdomains (Figure 2.18). The size and shape of each of these subdomains compares well with the distribution of 12 transmembrane helices expected from NHE1 topological predictions. The cytoplasmic domain of the hNHE1 reconstruction forms an elongated ridge that sits atop the transmembrane domain. The size of this ridge is consistent with a dimer formed by hNHE1 cytoplasmic domains, each consisting of ~300 amino acids. Unlike the transmembrane domain, which subdivides at higher density thresholds (Figure 2.18-A), the apical ridge remains a single continuous domain (Figure 2.18-B). This observation suggests extensive contact between cytoplasmic domains in the hNHE1 dimer. The extensive contact between cytoplasmic domains in the hNHE1 reconstruction is in good agreement with reported data that the proximal region of the cytoplasmic tail (amino acids

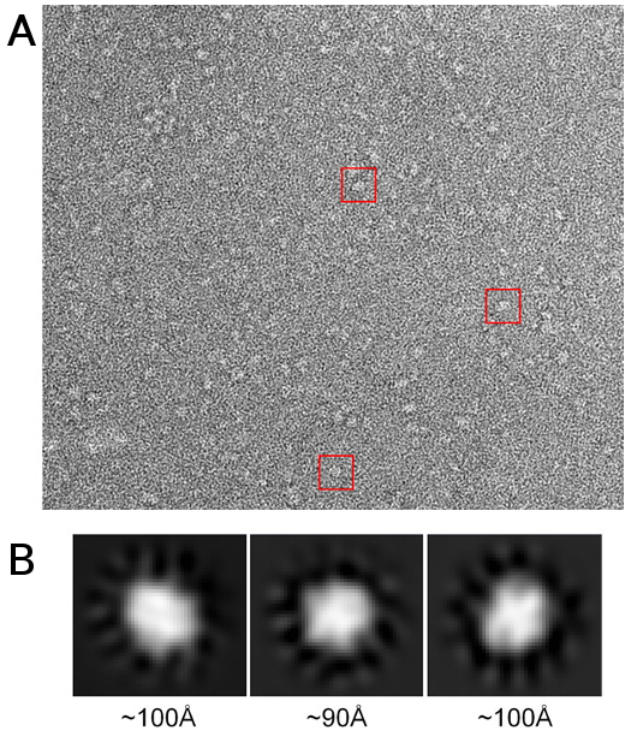

Figure 2.17: hNHE1 single particle reconstruction. ***A***. Electron micrograph image showing single particles randomly situated on a grid square. For reference three individual particles (lighter colour) are boxed in red. ***B***. Example of three class averages of single particles representing different surfaces of hNHE1. The dimensions of the molecular envelope measured from these classes are shown.

530–580) of NHE1 has a strong propensity to self-associate (46).

Discussion

Human NHE1 is an important target for the clinical improvement of heart disease, and a number of inhibitors of NHE1 have been developed for this purpose (77). However, clinical trials with NHE inhibitors have been unsuccessful or have resulted in negative side effects that overshadow any beneficial outcome (78–81). The molecular structure of NHE1 would aid the development of novel therapeutics, yet this structure remains unknown. While the Na^+/H^+ exchangers possess high activity, they are present in very low abundance in mammalian tissues. Therefore, it has not been possible to purify significant amounts of these proteins from native sources for structural characterization. While we have expressed single transmembrane segments of NHE1 in *E.*

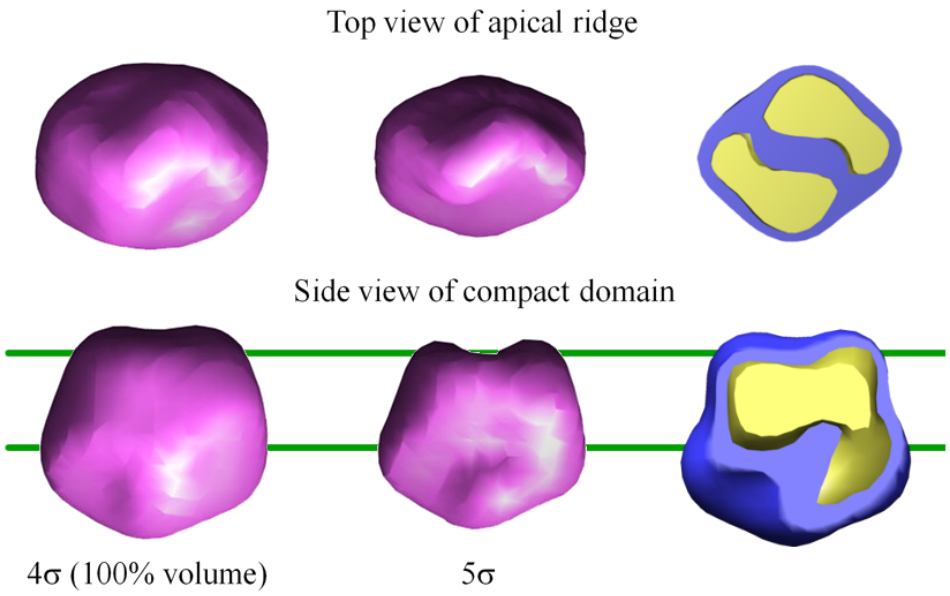

Figure 2.18: hNHE1 molecular envelope. In magenta the calculated molecular envelope of the hNHE1 dimer shown from two different angles at two levels of electron density (4 and 5 σ). The top view highlights the *apical ridge* feature that is likely the cytoplasmic tail domain. The side view highlights the membrane or *compact domain*. The putative membrane boundaries are represented as green lines. Overlaying the two magenta envelopes, now represented in blue and yellow at 4 σ and 5 σ, respectively, and taking a thick cross-section indicates how the *compact domain* is composed of two monomers (top view) and how the cytoplasmic tail protrudes below the membrane plane (side view).

coli and characterized them by NMR spectroscopy (2–5, 82–84), only very recently have larger segments containing two or more transmembrane helices been produced by the laboratory and their characterization is ongoing (see *Chapter 5*). In addition, we have not been successful in the overproduction of a yeast Na^+/H^+ exchanger, sod2, in *E. coli* (6). In an attempt to mitigate these limitations, we have developed a yeast expression system to overproduce the hNHE1 protein in amounts suitable for biophysical characterization of the protein. *S. cerevisiae* as an expression system has proven successful with transport proteins of similar overall function and origin (12, 42, 44). We successfully expressed hNHE1 in *S. cerevisiae* in amounts large enough for purification and characterization. Approximately 400 μg of purified protein could be obtained per litre of culture, consistent with the yields obtained for the heterologous expression of other eukaryotic membrane proteins in *S. cerevisiae* (43, 44, 64, 65, 85–87). Our data demonstrate that hNHE1 expressed in yeast retains functionality after purification in detergent solution and reconstitution into proteoliposomes. The reconstituted hNHE1 is capable of Na^+-dependent H^+ transport, and can be inhibited by the specific inhibitor EMD87580 (2). Because we expressed hNHE1 with the N-linked glycosylation site removed, we directly confirmed our earlier suggestion (38, 39) that N-linked glycosylation is not essential for cation transport activity.

In these studies, we used detergent-solubilized, purified hNHE1 for biophysical characterization of secondary structure content, oligomeric state, and low resolution molecular structure. Analysis of purified hNHE1 by CD spectroscopy demonstrated significant α-helical content, consistent with the predicted topology of 12 transmembrane segments. Based on this analysis and the transport measurements, we were confident that recombinant hNHE1 was properly folded and fully functional. Since there were persistent reports in the literature that NHE1 exists as a dimer (75, 76), we utilized gel filtration chromatography to address the oligomeric state of the protein in detergent solution. The molecular mass estimated by this technique was in slight excess of 200 kDa, while the predicted molecular weight of our recombinant hNHE1 was 92 kDa. The calculated molecular mass for an hNHE1 dimer was 233 kDa, when including 108 detergent molecules in the micelle. It therefore seems likely that the majority of our detergent-solubilized hNHE1 eluted as a dimer by gel filtration chromatography; however, a minor fraction of the protein eluted in the void volume and may form a higher order aggregate or larger micelle.

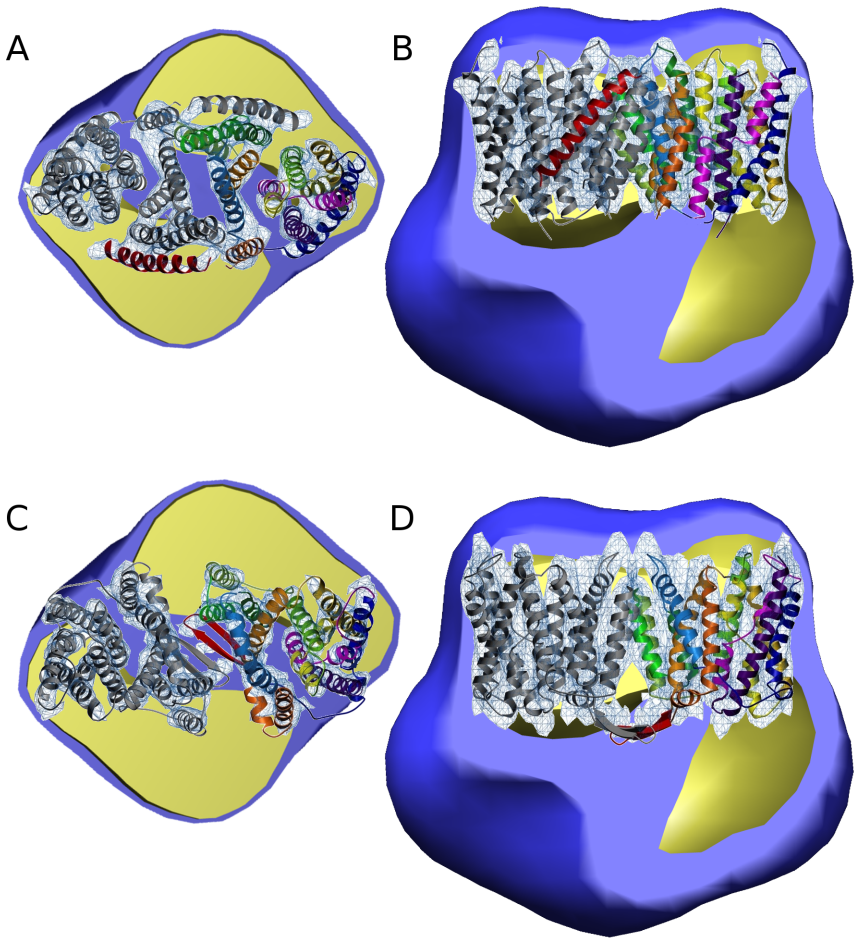

Figure 2.19: hNHE1 molecular envelope compared for *Methanococcus jannaschii* NhaP1 and *E. coli* NhaA. ***A***. View from the cytoplasm of the hNHE1 molecular envelope, from Figure 2.18, overlaid with: the projection map (white/blue mesh) of *M. jannaschii* NhaP1 calculated from 2D crystals using electron cryomicroscopy (88); and an homology model of NhaP1 (cartoon) calculated from using the *E. coli* NhaA structure (88). ***B***. A side view of ***A***. ***C***. Similar to ***A*** except using the 2D projection map of NhaA (89) and the crystal structure of NhaA (90). ***D***. A side view of ***C***. The NhaP1 and NhaA images in this figure were modified then reprinted with permission from Macmillan Publishers Ltd: *The EMBO Journal* (2011) **30**, 439–449, ©2010 (88).

Lastly, we used single particle electron microscopy of recombinant hNHE1 to confirm the oligomeric state and to determine the molecular shape of the exchanger (Figure 2.18). The three-dimensional reconstruction of hNHE1 at 22 Å resolution revealed an elongated cytoplasmic domain closely associated with a compact transmembrane domain. The overall shape of the transmembrane domain in each hNHE1 monomer is comparable to the known structural data for Na^+/H^+ exchangers: the crystal structure of NhaA from *E. coli* (90), as well as electron crystallographic analyses of NhaA (89, 91) and the archeal *Methanococcus jannaschi* NhaP1 (88, 92) (Figure 2.19). A more detailed comparison of the structural data and a final working model of NHE1 is presented in *Chapter 6*.

The lack of success with the alternative constructs (hTMD and mNHE1) could have many possible explanations. For hTMD it is quite possible that we have not created a stable construct that is amenable to expression. Our work with hNHE1 and the work of others (46) strongly suggests that the cytoplasmic domain is required for dimerization. This dimerization may be important to the overall stability of the protein. If our design of hTMD prevents dimerization, the protein may become quickly degraded *in vivo* preventing its purification. This may be surmounted by either extending the length of the cytoplasmic tail to allow dimerization, adding an engineered dimerization motif (e.g. coiled-coil) or finding a suitable *S. cerevisiae* host that does not significantly degrade the protein. However, there is also evidence that there may be constitutive regulators of NHE1 that bind to the cytoplasmic domain (93, 94) and if this region is removed protein stability may become impossible. Additionally, we intend to explore the possibilities for mNHE1 expression and purification for structure determination. While expression in *P. pastoris* did not yield positive results, the similarity of mNHE1 to hNHE1 means that it is highly likely to be expressed in the pYeDP60 system and purified in the same manner as hNHE1. Although this similarity may also result in difficulties during crystallization there is evidence that even small changes in amino acid sequence can significantly improve crystal packing interfaces (95).

This research has made many positive developments towards understanding NHE1 structure and function. We have developed a reproducible system for the expression and purification of functional full length human NHE1 in the yeast *Saccharomyces cerevisiae*. While crystallization trials were attempted, no significant progress was made towards achieving crystals (data not shown). To address this we developed alternative strategies including: a human transmembrane domain construct with yeast optimized codons and a full-length murine NHE1 with yeast harmo-

nized codons. These constructs can also be employed in a *Pichia pastoris* system. Although time limitations have prevented the progress of these possible projects, they provide a good base to continue pursuing the structure of NHE1. In addition, significant advances in the expression and purification of hNHE1 should be applicable to future projects: DSY864 strain, Tunair® baffled flasks, osmotic shock preceding lysis, and the inclusion of stripping during membrane isolation. Additional knowledge gained by trial and error is also useful including: concentrations for freezing membrane and purified protein, freeze-thaw stability, convenient stopping points, and optimal solubilization and resin binding volumes. With further optimization and modification the techniques outlined above may help lead to the determination of the full structure of NHE1.

References

1. Slepkov, E. R. *et al.* (2005). Structural and functional characterization of transmembrane segment IV of the NHE1 isoform of the Na^+/H^+ exchanger. *J Biol Chem* **280**, 17863–17872.

2. Ding, J., Rainey, J. K., Xu, C., Sykes, B. D. & Fliegel, L. (2006). Structural and functional characterization of transmembrane segment VII of the Na^+/H^+ exchanger isoform 1. *J Biol Chem* **281**, 29817–29829.

3. Reddy, T. *et al.* (2008). Structural and functional characterization of transmembrane segment IX of the NHE1 isoform of the Na^+/H^+ exchanger. *J Biol Chem* **283**, 22018–22030.

4. Lee, B. L., Li, X., Liu, Y., Sykes, B. D. & Fliegel, L. (2009). Structural and functional analysis of transmembrane XI of the NHE1 isoform of the Na^+/H^+ exchanger. *J Biol Chem* **284**, 11546–11556.

5. Lee, B. L., Li, X., Liu, Y., Sykes, B. D. & Fliegel, L. (2009). Structural and functional analysis of extracellular loop 2 of the Na^+/H^+ exchanger. *Biochim Biophys Acta* **1788**, 2481–2488.

6. Dibrov, P., Young, P. G. & Fliegel, L. (1998). Functional analysis of amino acid residues essential for activity in the Na^+/H^+ exchanger of fission yeast. *Biochemistry* **37**, 8282–8288.

7. Parcej, D. N. & Eckhardt-Strelau, L. (2003). Structural characterisation of neuronal voltage-sensitive K^+ channels heterologously expressed in *Pichia pastoris*. *J Mol Biol* **333**, 103–116.

8. Long, S. B., Campbell, E. B. & MacKinnon, R. (2005). Crystal structure of a mammalian voltage-dependent shaker family K^+ channel. *Science* **309**, 897–903.

9. Flegelova, H., Haguenauer-Tsapis, R. & Sychrova, H. (2006). Heterologous expression of mammalian Na/H antiporters in saccharomyces cerevisiae. *Biochim Biophys Acta* **1760**, 504–516.

10. Flegelova, H. & Sychrova, H. (2005). Mammalian NHE2 Na^+/H^+ exchanger mediates efflux of potassium upon heterologous expression in yeast. *FEBS letters* **579**, 4733–4738.

11. Montero-Lomelí, M. & Okorokova Façanha, A. L. (1999). Expression of a mammalian Na^+/H^+ antiporter in *Saccharomyces cerevisiae*. *Biochem Cell Biol* **77**, 25–31.

12. Sekler, I., Kopito, R. & Casey, J. R. (1995). High level expression, partial purification, and functional reconstitution of the human AE1 anion exchanger in *Saccharomyces cerevisiae*. *J Biol Chem* **270**, 21028–21034.

13. Wakabayashi, S., Pang, T., Su, X. & Shigekawa, M. (2000). A novel topology model of the human Na^+/H^+ exchanger isoform 1. *J Biol Chem* **275**, 7942–7949.

14. Landau, M., Herz, K., Padan, E. & Ben-Tal, N. (2007). Model structure of the Na^+/H^+ exchanger 1 (NHE1): functional and clinical implications. *J Biol Chem* **282**, 37854–37863.

15. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *J Mol Biol* **305**, 567–580.

16. Worland, S. T. & Wang, J. C. (1989). Inducible overexpression, purification, and active site mapping of DNA topoisomerase II from the yeast *Saccharomyces cerevisiae*. *J Biol Chem* **264**, 4412–4416.

17. Lindsley, J. E. & Wang, J. C. (1993). On the coupling between ATP usage and DNA transport by yeast DNA topoisomerase II. *J Biol Chem* **268**, 8096–8104.

18. Bai, J. *et al.* (2011). A gene optimization strategy that enhances production of fully functional p-glycoprotein in *Pichia pastoris*. *PLoS ONE* **6**, e22577.

19. De Schutter, K. *et al.* (2009). Genome sequence of the recombinant protein production host *Pichia* pastoris. *Nat Biotechnol* **27**, 561–566.

20. Hani, J. & Feldmann, H. (1998). tRNA genes and retroelements in the yeast genome. *Nucleic Acids Res* **26**, 689–696.

21. Dragosits, M. *et al.* (2009). The effect of temperature on the proteome of recombinant *Pichia pastoris*. *J Proteome Res* **8**, 1380–1392.

22. Dragosits, M. *et al.* (2010). The response to unfolded protein is involved in osmotolerance of *Pichia pastoris*. *BMC Genomics* **11**, 207.

23. Baumann, K. *et al.* (2010). A multi-level study of recombinant *Pichia pastoris* in different oxygen conditions. *BMC Systems Biology* **4**, 141.

24. Schiestl, R. H. & Gietz, R. D. (1989). High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Curr Genet* **16**, 339–346.

25. Clare, J. J., Rayment, F. B., Ballantine, S. P., Sreekrishna, K. & Romanos, M. A. (1991). High-level expression of tetanus toxin fragment C in *Pichia pastoris* strains containing multiple tandem integrations of the gene. *Nat Biotechnol* **9**, 455–460.

26. Clare, J. J. *et al.* (1991). Production of mouse epidermal growth factor in yeast: high-level secretion using *Pichia pastoris* strains containing multiple gene copies. *Gene* **105**, 205–212.

27. Crotti, L. B., Drgon, T. & Cabib, E. (2001). Yeast cell permeabilization by osmotic shock allows determination of enzymatic activities in situ. *Anal Biochem* **292**, 8–16.

28. Gebreselassie, D., Rajarathnam, K. & Fliegel, L. (1998). Expression, purification, and characterization of the carboxyl-terminal region of the Na^+/H^+ exchanger. *Biochem Cell Biol* **76**, 837–842.

29. Sreerama, N. & Woody, R. W. (2004). On the analysis of membrane protein circular dichroism spectra. *Protein Sci* **13**, 100–112.

30. Sreerama, N. & Woody, R. W. (2000). Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem* **287**, 252–260.

31. Nakamura, N., Tanaka, S., Teko, Y., Mitsui, K. & Kanazawa, H. (2005). Four Na^+/H^+ exchanger isoforms are distributed to golgi and post-golgi compartments and are involved in organelle pH regulation. *J Biol Chem* **280**, 1561–1572.

32. Karmazyn, M., Avkiran, M. & Fliegel, L. *The sodium-hydrogen exchanger: from molecule to its role in disease* (Kluwer Academic Publishers, 2003).

33. Ludtke, S. J., Baldwin, P. R. & Chiu, W. (1999). EMAN: semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol* **128**, 82–97.

34. Frank, J. *et al.* (1996). SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol* **116**, 190–199.

35. Rieder, C. V. & Fliegel, L. (2002). Developmental regulation of Na^+/H^+exchanger expression in fetal and neonatal mice. *Am J Physiol* **283**, H273–H283.

36. Kaplan, R. S. & Pedersen, P. L. (1985). Determination of microgram quantities of protein in the presence of milligram levels of lipid with amido black 10B. *Anal Biochem* **150**, 97–104.

37. Schaffner, W. & Weissmann, C. (1973). A rapid, sensitive, and specific method for the determination of protein in dilute solution. *Anal Biochem* **56**, 502–514.

38. Counillon, L., Pouysségur, J. & Reithmeier, R. A. (1994). The Na^+/H^+ exchanger NHE-1 possesses N- and O-linked glycosylation restricted to the first N-terminal extracellular domain. *Biochemistry* **33**, 10463–10469.

39. Haworth, R. S., Fröhlich, O. & Fliegel, L. (1993). Multiple carbohydrate moieties on the Na^+/H^+ exchanger. *Biochem J* **289**, 637–640.

40. Pompon, D., Louerat, B., Bronine, A. & Urban, P. (1996). Yeast expression of animal and plant P450s in optimized redox environments. *Methods Enzymol* **272**, 51–64.

41. Urban, P., Cullin, C. & Pompon, D. (1990). Maximizing the expression of mammalian cytochrome P-450 monooxygenase activities in yeast cells. *Biochimie* **72**, 463–472.

42. Lenoir, G. *et al.* (2002). Overproduction in yeast and rapid and efficient purification of the rabbit SERCA1a Ca^{2+}-ATPase. *Biochim Biophys Acta* **1560**, 67–83.

43. Jidenko, M. *et al.* (2005). Crystallization of a mammalian membrane protein overexpressed in saccharomyces cerevisiae. *Proc Natl Acad Sci U S A* **102**, 11687–11691.

44. Jahn, T. *et al.* (2001). Large scale expression, purification and 2D crystallization of recombinant plant plasma membrane H^+-ATPase. *J Mol Biol* **309**, 465–476.

45. Midgett, C. R. & Madden, D. R. (2007). Breaking the bottleneck: Eukaryotic membrane protein expression for high-resolution structural studies. *J Struct Biol* **160**, 265–274.

46. Hisamitsu, T., Pang, T., Shigekawa, M. & Wakabayashi, S. (2004). Dimeric interaction between the cytoplasmic domains of the Na^+/H^+ exchanger NHE1 revealed by symmetrical intermolecular cross-linking and selective co-immunoprecipitation. *Biochemistry* **43**, 11135–11143.

47. Li, X., Liu, Y., Kay, C. M., Müller-Esterl, W. & Fliegel, L. (2003). The Na^+/H^+ exchanger cytoplasmic tail: structure, function, and interactions with tescalcin. *Biochemistry* **42**, 74487456.

48. Nørholm, A.-B. *et al.* (2011). The intracellular distal tail of the Na^+/H^+ exchanger NHE1 is intrinsically disordered: implications for NHE1 trafficking. *Biochemistry* **50**, 3469–3480.

49. Holland, J. P. & Holland, M. J. (1979). The primary structure of a glyceraldehyde-3-phosphate dehydrogenase gene from *Saccharomyces cerevisiae*. *J Biol Chem* **254**, 9839–9845.

50. Bennetzen, J. L. & Hall, B. D. (1982). Codon selection in yeast. *J Biol Chem* **257**, 3026–3031.

51. Ikemura, T. (1982). Correlation between the abundance of yeast transfer RNAs and the occurrence of the respective codons in protein genes: Differences in synonymous codon choice patterns of yeast and *Escherichia coli* with reference to the abundance of isoaccepting transfer RNAs. *J Mol Biol* **158**, 573–597.

52. Ikemura, T. (1985). Codon usage and tRNA content in unicellular and multicellular organisms. *Mol Biol Evol* **2**, 13–34.

53. Sharp, P. M. & Li, W. H. (1986). An evolutionary perspective on synonymous codon usage in unicellular organisms. *J Mol Evol* **24**, 28–38.

54. Sharp, P. M., Tuohy, T. M. F. & Mosurski, K. R. (1986). Codon usage in yeast: cluster analysis clearly differentiates highly and lowly expressed genes. *Nucleic Acids Res* **14**, 5125–5143.

55. Romanos, M. A., Scorer, C. A. & Clare, J. J. (1992). Foreign gene expression in yeast: a review. *Yeast* **8**, 423–488.

56. Daly, R. & Hearn, M. T. W. (2005). Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. *J Mol Recognit* **18**, 119–138.

57. Huang, H. *et al.* (2008). High-level expression of a truncated 1,3–1,4-β-D-glucanase from *Fibrobacter succinogenes* in *Pichia pastoris* by optimization of codons and fermentation. *Appl Microbiol Biotechnol* **78**, 95–103.

58. Cregg, J. M. *et al.* (2009). Expression in the yeast *Pichia pastoris*. *Methods Enzymol* **463**, 169–189.

59. Lueking, A., Holz, C., Gotthold, C., Lehrach, H. & Cahill, D. (2000). A system for dual protein expression in *Pichia pastoris* and *Escherichia coli*. *Protein Expression Purif* **20**, 372–378.

60. Li, M. *et al.* (2009). Selecting optimum eukaryotic integral membrane proteins for structure determination by rapid expression and solubilization screening. *J Mol Biol* **385**, 820–830.

61. Hays, F. A. *et al.* (2009). Ratiocinative screen of eukaryotic integral membrane protein expression and solubilization for structure determination. *J Struct Funct Genomics* **10**, 9–16.

62. Hays, F. A., Roe-Zurz, Z. & Stroud, R. M. (2010). Overexpression and purification of integral membrane proteins in yeast. *Methods Enzymol* **470**, 695–707.

63. Mumberg, D., Muller, R. & Funk, M. (1994). Regulatable promoters of *Saccharomyces cerevisiae*: comparison of transcriptional activity and their use for heterologous expression. *Nucleic Acids Res* **22**, 5767–5768.

64. Waight, A. B. *et al.* (2013). Structural basis for alternating access of a eukaryotic calcium/proton exchanger. *Nature* .

65. Pedersen, B. P. *et al.* (2013). Crystal structure of a eukaryotic phosphate transporter. *Nature* **496**, 533–536.

66. Rollwitz, I., Santaella, M., Hille, D., Flgge, U.-I. & Fischer, K. (2006). Characterization of AtNST-KT1, a novel UDP-galactose transporter from arabidopsis thaliana. *FEBS Letters* **580**, 4246–4251.

67. Otto, B. *et al.* (2010). Aquaporin tetramer composition modifies the function of tobacco aquaporins. *J Biol Chem* **285**, 31253–31260.

68. White, M. A., Clark, K. M., Grayhack, E. J. & Dumont, M. E. (2007). Characteristics affecting expression and solubilization of yeast membrane proteins. *J Mol Biol* **365**, 621–636.

69. Bertrand, B., Wakabayashi, S., Ikeda, T., Pouysségur, J. & Shigekawa, M. (1994). The Na^+/H^+ exchanger isoform 1 (NHE1) is a novel member of the calmodulin-binding proteins. identification and characterization of calmodulin-binding sites. *J Biol Chem* **269**, 13703–13709.

70. Venema, K., Gibrat, R., Grouzis, J. P. & Grignon, C. (1993). Quantitative measurement of cationic fluxes, selectivity and membrane potential using liposomes multilabelled with fluorescent probes. *Biochim Biophys Acta* **1146**, 87–96.

71. Venema, K., Quintero, F. J., Pardo, J. M. & Donaire, J. P. (2002). The arabidopsis Na^+/H^+exchanger AtNHX1 catalyzes low affinity Na^+ and K^+ transport in reconstituted liposomes. *J Biol Chem* **277**, 2413–2418.

72. Kleinman, J. G., Harig, J. M., Barry, J. A. & Ramaswamy, K. (1988). Na^+ and H^+ transport in human jejunal brush-border membrane vesicles. *Am J Physiol* **255**, G206–211.

73. Jean, T., Frelin, C., Vigne, P., Barbry, P. & Lazdunski, M. (1985). Biochemical properties of the Na^+/H^+ exchange system in rat brain synaptosomes. interdependence of internal and external pH control of the exchange activity. *J Biol Chem* **260**, 9678–9684.

74. Saxena, V. P. & Wetlaufer, D. B. (1971). A new basis for interpreting the circular dichroic spectra of proteins. *Proc Natl Acad Sci U S A* **68**, 969–972.

75. Fliegel, L., Haworth, R. S. & Dyck, J. R. (1993). Characterization of the placental brush border membrane Na^+/H^+ exchanger: identification of thiol-dependent transitions in apparent molecular size. *Biochem J* **289 (Pt 1)**, 101–107.

76. Fafournoux, P., Noël, J. & Pouysségur, J. (1994). Evidence that Na^+/H^+ exchanger isoforms NHE1 and NHE3 exist as stable dimers in membranes with a high degree of specificity for homodimers. *J Biol Chem* **269**, 2589–2596.

77. Avkiran, M. & Marber, M. S. (2002). Na^+/H^+ exchange inhibitors for cardioprotective therapy: progress, problems and prospects. *J Am Coll Cardiol* **39**, 747–753.

78. Theroux, P. *et al.* (2000). Design of a trial evaluating myocardial cell protection with cariporide, an inhibitor of the transmembrane sodium-hydrogen exchanger: the Guard during ischemia against necrosis (GUARDIAN) trial. *Current controlled trials in cardiovascular medicine* **1**, 59–67.

79. Zeymer, U. *et al.* (2001). The Na^+/H^+ exchange inhibitor eniporide as an adjunct to early reperfusion therapy for acute myocardial infarction. results of the evaluation of the safety and cardioprotective effects of eniporide in acute myocardial infarction (ESCAMI) trial. *J Am Coll Cardiol* **38**, 1644–1650.

80. Mentzer Jr., R. M. (2003). Effects of Na^+/H^+ exchange inhibition by cariporide on death and nonfatal myocardial infarction in patients undergoing coronary artery bypass graft surgery: The expedition study. *Circulation* **108**, 2723–2723.

81. Chaitman, B. R. (2003). A review of the GUARDIAN trial results: clinical implications and the significance of elevated perioperative CK-MB on 6-month survival. *J Card Surg* **18**, 13–20.

82. Slepkov, E. R., Rainey, J. K., Sykes, B. D. & Fliegel, L. (2007). Structural and functional analysis of the Na^+/H^+ exchanger. *Biochem J* **401**, 623–633.

83. Tzeng, J., Lee, B. L., Sykes, B. D. & Fliegel, L. (2010). Structural and functional analysis of transmembrane segment VI of the NHE1 isoform of the Na^+/H^+ exchanger. *J Biol Chem* **285**, 36656–36665.

84. Lee, B. L., Sykes, B. D. & Fliegel, L. (2011). Structural analysis of the Na^+/H^+ exchanger isoform 1 (NHE1) using the divide and conquer approach. *Biochem Cell Biol* **89**, 189–199.

85. Sarramegna, V., Talmont, F., Demange, P. & Milon, A. (2003). Heterologous expression of G-protein-coupled receptors: comparison of expression systems from the standpoint of large-scale production and purification. *Cell Mol Life Sci* **60**, 1529–1546.

86. Pedersen, B. P., Buch-Pedersen, M. J., Morth, J. P., Palmgren, M. G. & Nissen, P. (2007). Crystal structure of the plasma membrane proton pump. *Nature* **450**, 1111–1114.

87. Pryor, E. E. *et al.* (2013). Structure of the integral membrane protein CAAX protease Ste24p. *Science* **339**, 1600–1604.

88. Goswami, P. *et al.* (2011). Structure of the archaeal Na^+/H^+ antiporter NhaP1 and functional role of transmembrane helix 1. *EMBO J* **30**, 439–449.

89. Williams, K. A. (2000). Three-dimensional structure of the ion-coupled transport protein NhaA. *Nature* **403**, 112–115.

90. Hunte, C. *et al.* (2005). Structure of a Na^+/H^+ antiporter and insights into mechanism of action and regulation by pH. *Nature* **435**, 1197–1202.

91. Williams, K. A., Geldmacher-Kaufer, U., Padan, E., Schuldiner, S. & Kühlbrandt, W. (1999). Projection structure of NhaA, a secondary transporter from *Escherichia coli*, at 4.0 å resolution. *EMBO J* **18**, 3558–3563.

92. Vinothkumar, K. R., Smits, S. H. J. & Kühlbrandt, W. (2005). pH-induced structural change in a sodium/proton antiporter from *Methanococcus jannaschii*. *EMBO J* **24**, 2720–2729.

93. Pang, T., Su, X., Wakabayashi, S. & Shigekawa, M. (2001). Calcineurin homologous protein as an essential cofactor for Na^+/H^+ exchangers. *J Biol Chem* **276**, 17367–17372.

94. Zaun, H. C., Shrier, A. & Orlowski, J. (2008). Calcineurin b homologous protein 3 promotes the biosynthetic maturation, cell surface stability, and optimal transport of the Na^+/H^+ exchanger NHE1 isoform. *J Biol Chem* **283**, 12456–12467.

95. Derewenda, Z. S. (2004). Rational protein crystallization by mutational surface engineering. *Structure* **12**, 529–535.

Chapter 3

Expression and Purification of Multi-transmembrane Segment Peptide for Structural Studies

The techniques and protocols presented in this *chapter* were published in *Current Protocols in Protein Science* by John Wiley & Sons Inc., Hoboken, NJ, USA (1). They are used here with permission of the publisher. The unique format of this publication has been largely maintained including detailed numbered protocol steps.

Introduction

Structural biology has proven to be an important tool in the progression of modern medicine. From understanding the chemistry of relatively simple molecules such as drugs and antibiotics to being able to visualize large proteins and protein complexes, structural biology has changed the way we understand the human body and its environment. However, our understanding of membrane proteins lags behind that of soluble proteins, and determining a membrane structure remains a challenging endeavour. If one considers that approximately one-third of the human genome encodes membrane proteins (2), many of which are potential drug targets, membrane proteins only account for ~1% of structures deposited in the Protein Data Bank (3). One reason

for this lies in the relative difficulty in expressing, purifying and studying membrane proteins *in vitro*. Not only must all the considerations required for soluble proteins be optimized, unique challenges associated with maintaining the appropriate membrane or membrane-like environment must also be included. Before this optimization can begin, the first hurdle comes with the fact that most membrane proteins are not naturally abundant, necessitating their heterologous expression. It is typically more difficult to express and purify membrane proteins, especially if one is interested in human or mammalian variants (4). This is compounded when producing large quantities of polytopic membrane proteins for structural studies.

This unit will discuss one strategy in the *divide-and-conquer* approach (5) to design, heterologously express and purify to homogeneity transmembrane peptides that can be used for structural biology. The method uses a Maltose Binding Protein (MBP) fusion that allows difficult hydrophobic peptides to be expressed and purified as a soluble protein construct (6). To illustrate we use the human sodium proton exchanger isoform 1 as an example. The challenges that may be encountered, alternative approaches and the physiological relevance are discussed below.

BASIC PROTOCOL 1: Designing and cloning a transmembrane peptide for expression

The recipes for all media and solutions are found at the end of the chapter.

Design of the peptide

1. Acquire the sequence of your protein of interest.
   (a) UniProt and NCBI are good online sources.
2. Hydropathy analysis and secondary structure prediction are recommended to help determine the appropriate choice of transmembrane segment.
   (a) Recommended for transmembrane prediction: TMHMM2 and HMMTop2.
   (b) Recommended for secondary structure prediction: JPred3.
   (c) Use the above analyses to select a helical sequence that also crosses the membrane then ensure that you include >3 extramembrane residues on either side of the predicted segment.
   (d) A thorough literature review for experimental evidence (mutagenesis, labelling studies, etc.) will also be useful for supporting or refuting the limits of the predicted transmembrane segment.

If solubility problems are encountered in the initial stages of peptide purification (see below), it may be desirable to include either additional residues from extra-membranous loops or capping lysine residues that flank the transmembrane segment.

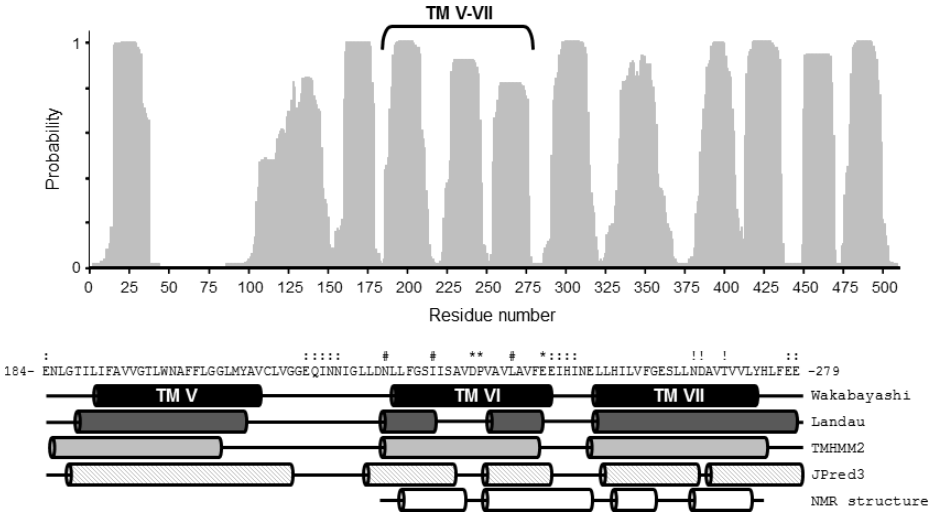

Figure 3.1: Transmembrane peptide design. The upper panel depicts the output of TMHMM2 prediction of membrane topology (7) for the human sodium proton exchanger isoform 1 (NHE1) residues 1–510. The probability of each amino acid (x-axis) being located in the membrane is shown on the y-axis. The lower panel displays the sequence of NHE1 between residues 184 and 279 with various transmembrane segment predictions, the regions of helical content determined by NMR, and known functional mutations. The black cylinders labelled *Wakabayashi* represent transmembrane helices predicted by Wakabayashi *et al.* (8). The dark grey cylinders labelled *Landau* represent transmembrane helices predicted by Landau *et al.* (9). The light grey cylinders labelled *TMHMM* correspond to the transmembrane regions from the upper panel. The hatched cylinders labelled *JPred3* represent helices predicted by the JPred3 algorithm (10). The white cylinders labelled *NMR structure* represent the solution structures of two purified peptides presented by Tzeng *et al.* (11) and Ding *et al.* (12). Above the sequence functionally important residues are labelled: (*) represents Cys mutants that inactivate NHE1 (12); (#) represents Cys mutants that are affected by cysteine modification reagent (12); (!) represents Ala mutants that greatly reduce NHE1 activity (11); (:) represents Cys mutants that are accessible from the extra- or intracellular space (8).

Cloning peptide for expression as a Maltose Binding Protein fusion

Expression vectors for MBP fusion proteins are commercially available (New England Biolabs Inc., Ipswich, MA, USA). We have previously described how pMal-c2x was adapted to allow more efficient screening of multiple constructs of phospholamban and sarcolipin (13). Below we adapt this procedure for the currently available, pMal-c5X. Briefly, a polymerase chain reaction (PCR) product containing a tobacco etch virus (TEV) protease site and the restriction sites *BamHI* and *EcoRI* is created allowing a single protocol for the cloning, expression, purification and characterization of various peptide constructs. Selection of the TEV protease was based on its specificity and efficiency in our preparation conditions and because recombinant protein can be readily obtained. Factor Xa, included in the pMal-c5X vector, is a popular alternative (14).

Constructing the pMal-c5X+TEV vector

This need only be done for the first construct. Subsequent cloning can use the *BamHI* and *EcoRI* sites that are incorporated during this protocol. Initially, PCR is used to create a product with 5'-TEV-*BamHI*-gene of interest-STOP-*EcoRI*-*SalI*-3'. It is important that the initiating methionine is not included so that a C-terminal fusion protein is created. Also, care must be taken to ensure that the TEV site will be in frame with MBP.

Perform PCR and prepare the vector for cloning

1. Before beginning PCR, phosphorylate the forward primer using T4 kinase kit.
   (a) Performing this step on the primer rather than the PCR product reduces PCR product loss.
2. In a 200 μL PCR tube prepare the PCR reaction on ice using the phosphorylated forward primer.
   (a) PCR reaction conditions should be obtained from the polymerase manufacturer.
   (b) Primer design was modified from Douglas *et al.* (13) and primers are as follows: (forward) 5'-GAGAACCTGTACTTCCAGGGATCCNNN-3', TEV site (ENLYFQ) is underlined, *BamHI* site is in bold, gene of interest (starting at the second codon!) is shown as NNN; (reverse) 5'-ACTGGAATTCTCANNN-3', STOP codon is bold, *EcoRI* site is italicized, gene of interest is represented by NNN. ACTG at the 5' end of the reverse primer allows efficient restriction digestion of the PCR product.
3. Run the PCR reaction in a thermal cycler.
   (a) Cycle conditions should be obtained from the polymerase manufacturer.

(b) Annealing temperature (Ta) should be the melting temperature of the primers minus 5 °C.

4. Gel purify the PCR product (e.g. using a QIAquick method).

5. Digest the purified PCR product with *EcoRI* and re-purify.

6. Digest 4 mg of pMal-c5X vector with PdmI and *EcoRI* and purify by gel extraction (QIAquick).

   (a) Check double digest enzyme compatibility with your supplier.

Ligation of the PCR product into pMal-c5X and transformation into *E. coli*

7. Ligate together the cut PCR product and cut vector using a DNA ligation kit (Fermentas).

   (a) Typical conditions are: 13 μL cut PCR product, 2 μL cut vector, 5 μL 5x ligation buffer, 1 μL T4 ligase. Incubate at room temperature for 2 h or overnight at 4 °C.

8. Transform the ligation into DH5α competent cells for maintenance and DNA production.

   (a) Another similar cloning strain of *E. coli* is also suitable.

   (b) LB+amp plates are used to select for the pMal vector.

9. Individual transformants, typically 4–10, are screened for correct PCR product ligation by PCR.

   (a) A single colony is selected using a sterile loop and suspended in 100 μL of sterile water in a sterile microcentrifuge tube by vigorous stirring of the loop. This loop is then streaked onto a fresh LB+amp plate (and placed at 37 °C overnight) to maintain the colony if it is positive. 4–10 colonies are selected this way.

   (b) To create PCR template, resuspended colonies are boiled for 5 min.

10. Repeat PCR reaction above using 2 μL of the boiled colony resuspension as template.

11. If the PCR yields a product, the cloning has been successful.

12. Grow a positive streaked transformant, from step 9a, in 2 mL of LB+amp for 8 h in a 37 °C shaking incubator.

13. Transfer this culture to 50 mL of LB+amp and grow overnight in a 37 °C shaking incubator

14. The next day make a glycerol stock: take 800 μL of the overnight culture and mix it with 200 μL of sterile 80% glycerol in a 1.5 mL cryotube and store at -80 °C.

15. Use the remainder of the culture for plasmid purification (QIAGEN plamid MidiPrep kit).

16. Send a sample of the plasmid DNA for sequencing to confirm construct insertion and check for any mutations arising during PCR.

The resultant plasmid contains the following elements: MBP-linker-FacXa-TEV-*BamHI*-yourgene-*EcoRI*. The above procedure can be modified to include other compatible restriction sites, alternate protease cleavage sites, or additional affinity tags added before or after your target sequence (15). By digesting this new vector with *BamHI* and *EcoRI* a different PCR product can be added after the TEV protease site. This plasmid can now be transformed into different bacterial strains for expression testing.

BASIC PROTOCOL 2: Expression and purification of fusion protein

After successfully cloning your gene into the modified pMal vector and transformation into a bacterial strain, the transformants are screened for expression. We have found that expression levels do not vary widely between transformants, yet occasionally a particular colony shows better growth and expression. Thus, it is advisable to choose several colonies for expression testing.

Screening transformants for expression

1. A single transformant is picked from a plate and added to 1.5 mL of LB+amp media and grown overnight at 37 °C in a rotary shaker at high speed (>200 rpm).
   (a) Large scale: single colony added to 10 mL LB+amp
2. In the morning 100 µL of the overnight culture is inoculated into 10 mL of media + amp (1/100 dilution of stock recipe) and grown at 37 °C in a rotary shaker (150–200 rpm) until OD_{600} of 0.4–0.6 is reached (typically 3–6 h).
   (a) Media = LB, Terrific, or M9
   (b) Large scale: the entire 10 mL overnight culture is added to 1 L of media + amp.
   (c) Large scale: the culture can be moved to the lower expression temperature for at least 30 min prior to induction to ensure induction at the lower temperature. This is particularly important for difficult constructs.
   (d) A glycerol stock should also be made to preserve this overnight culture.
3. A 1 mL non-induced sample is taken for gel analysis.
4. The appropriate amount of inducing agent (0.1 to 1 mM IPTG) is added and the culture is placed in a rotary shaker incubator (100–150 rpm) at the appropriate temperature for the appropriate time.

(a) Typically 18–22 °C is used for 12–48 h. During a pilot experiment several samples should be taken at convenient intervals (4, 8, 12, 24, 48 h) and analysed by SDS-PAGE.

Fusion protein purification

The general steps for purification include centrifugation, chromatographic purification of the fusion protein, protease cleavage to liberate the target peptide, and peptide purification. If the fusion protein is found in the soluble fraction, as is the case for our MBP fusions, it may precipitate at later stages of purification. In this case, the chromatography buffers should be supplemented with 10–20% glycerol and the salt concentration (PSE base) should be decreased.

Maltose affinity chromatography

Samples should be taken at each step of the fusion protein purification for SDS-PAGE analysis (Figure 3.2).

1. Bacteria are harvested (6000g for 15 min at 4 °C) and resuspended completely in cold lysis buffer.
   (a) For sonication resuspend the cells at ~0.1–2 g/mL. For high pressure methods use ~0.5–0.6 g/mL.
   (b) Steps 1–3 should be carried out at 4 °C. All other steps can be carried out at room temperature, though handling at 4 °C may increase protein stability.
2. Cells are lysed using standard procedures (sonication, Emulsiflex etc.).
   (a) Ensure the cells remain at 4 °C during lysis.
3. The lysate is clarified by centrifugation to remove unbroken cells and insoluble material (50,000g for 25 min at 4 °C).
4. Prepared maltose affinity resin is incubated in batch mode with the clarified lysate with gentle agitation for 1 hour.
   (a) The resin is prepared by washing with at least 3 column volumes of purification buffer. If the resin is in 20% ethanol, wash with 1–2 column volumes of water followed by 3 column volumes of purification buffer to avoid salt precipitation.
   (b) Gently agitate the resin on a rotary platform device (e.g. a gel rocker or Nutator® is recommended; a stir bar may damage the resin).
   (c) 25 mL of amylose resin will bind at least 150 mg of fusion protein. This resin can be washed as recommended by the manufacturer and reused at least 10 times (until binding efficiency is decreased).
5. Pour the resin into a column and allow it to pack by gravity while collecting the flow through.

6. The column is washed with at least 3 column volumes of purification buffer to remove any non-specifically bound protein.
   (a) To ensure complete washing, a quick qualitative protein assay can be used. Using a micropipette, take 2 µL of liquid from a drop hanging from the tip of the column and mix it with 3 drops of 1x Bradford reagent (Bio-Rad, Hercules, CA, USA), initially brown in colour. If protein is washing off of the resin, the mixture will turn blue. When all the contaminants have been washed off the column, the mixture will remain brown.
7. Add 0.2–0.5 column volumes of elution buffer to the column, mix well (the column is capped at both ends and placed on a rotary platform device), and allow it to incubate for 20 min.
   (a) This step helps remove the protein in a smaller volume and reduces the later requirement for concentration.
8. Collect the eluent and continue to add elution buffer until all the protein has been eluted (1–2 column volumes).
   (a) The same Bradford test described above can be used here as well.
9. Concentrate the protein to >5 mg/mL using a filter driven concentrator (10,000 MWCO).
   (a) This improves TEV cleavage as the enzyme is more effective at higher concentrations.
   (b) The eluent may be stored at 4 °C before or after concentration.

Protease digestion

10 U of TEV per 100 mg of fusion protein and 1 mM (final) of DTT is added to the concentrated eluent and allowed to digest at 16 °C until completion. This is typically 24–72 h but it can be longer. If longer times are required, add 1 mM fresh DTT every 2 or 3 days. Alternatively, more TEV protease can be added if poor digestion (<30%) is observed after 3 or more days of digestion. The effectiveness of digestion is observed by a mobility-shift from fusion (>42 kDa) to free MBP (~42 kDa) by SDS-PAGE (8% acrylamide should be sufficient to observe a mobility shift of >2 kDa). See Figure 3.2.

BASIC PROTOCOL 3: Peptide recovery

At this stage, there are many possible ways to remove MBP and purify the peptide. We will focus on methods to purify from a protein pellet, though not every one will work for a particular peptide.

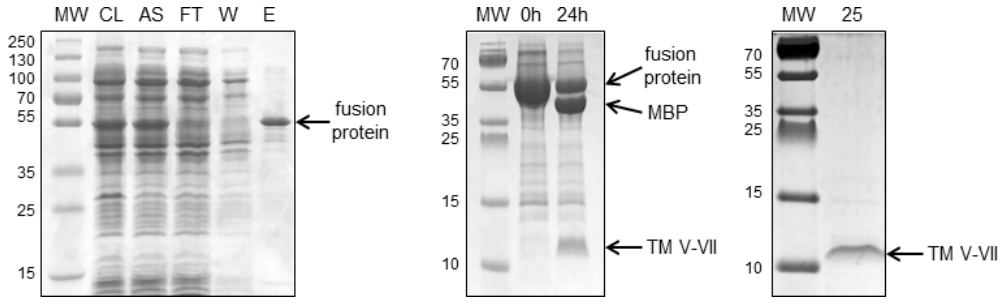

Figure 3.2: SDS-PAGE analysis of TM V-VII purification. Molecular weight ladder (MW) as marked, in kDa. Left panel is a 12% Tris-Glycine SDS-PAGE gel showing fractions of the purification: CL, crude lysate (5 μL); AS, after lysate ultracentrifugation (5 μL); FT, amylose column flow through (5 μL); W, amylose column wash (10 μL); E, amylose column elution (2 μL, ~10 μg). The position of the fusion protein (MBP-TM V-VII) is marked. Center panel is a 16% Tris-Tricine SDS-PAGE gel showing TEV protease digestion progress: 0 h, concentrated elution before TEV addition; 24 h, sample after 24 h digestion at 16 °C. The position of the fusion protein (MBP-TM V-VII), cleaved MBP and the peptide (TM V-VII) are marked. Right panel is a 16% Tris-Tricine SDS-PAGE gel showing 25 μL sample of the organic phase of the organic extraction, dried then resuspended in sample loading buffer. The position of the peptide (TM V-VII) is marked. The gels were stained using Coomassie blue.

If the target peptide precipitates following protease cleavage, high-speed centrifugation (100,000g for 45 min) may be used to collect a pellet enriched in the peptide. This is ideal for quantities of fusion proteins in excess of 200 mg without glycerol in the buffer. The peptide pellet may then be purified by HPLC (see protocol below) or organic extraction (see step 5 in protocol below).

If the peptide does not precipitate on its own, trichloroacetic acid (TCA) is used to precipitate all the protein (see step 1 in protocol below). The peptide is then extracted from this pellet by organic extraction (see protocols below).

Additionally a combination of techniques may be required to isolate the peptide depending on the sample. Below is a flowchart depicting these possible combinations (Figure 3.3). Initially, the sample may be split into aliquots for testing.

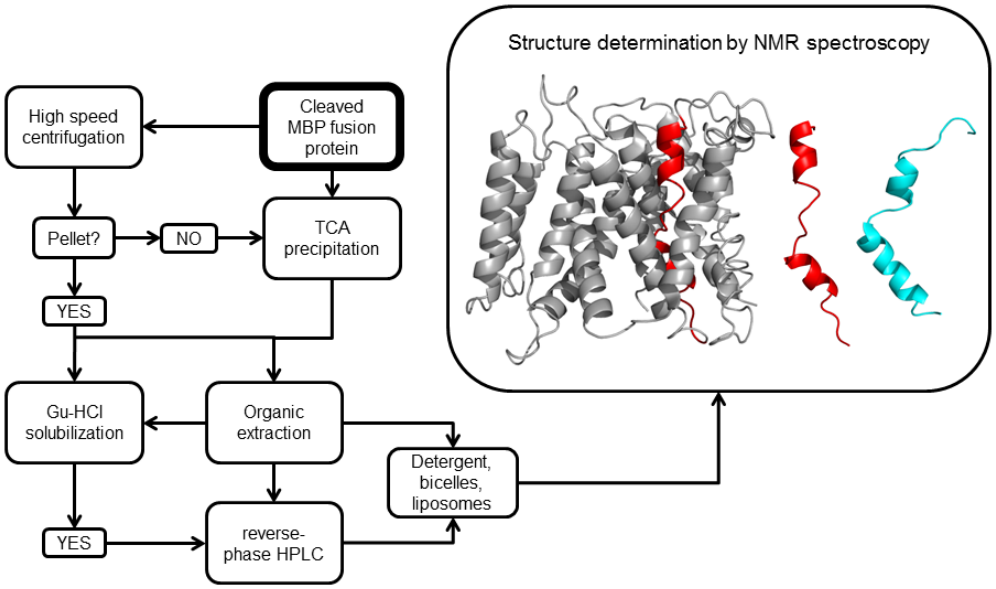

Figure 3.3: Flowchart depicting peptide isolation for further experiments. Begin with *Cleaved MBP fusion protein* and follow the arrows. The final box of the flowchart shows, on the left, a cartoon representation of a three dimensional model of the human sodium proton exchanger isoform 1 (9) with the sequence of TM VI highlighted in red. In the centre, the model structure of TM VI coloured red. On the right, the solution NMR structure of TM VI (11) for comparison.

Organic extraction

The following steps should be done using ONLY chloroform insensitive materials like glass, metal and Teflon. Many plastics dissolve in chloroform and will contaminate the sample.

1. After protease digestion is complete, the reaction mixture is transferred to a glass centrifuge tube and 2.5 mL of 60% TCA per 100 mg of fusion protein is SLOWLY added while swirling to precipitate the protein.
   (a) Slow addition of concentrated TCA helps prevent the formation of large aggregates of precipitate aiding in the later extraction step.
2. This mixture is incubated on ice for 30 min to complete the precipitation.
3. The precipitate is collected by centrifugation (6000g for 30 min at 4 °C)
   (a) Glass centrifuge tubes may break at RCF values over 9000g , therefore lower speeds for longer times are used.
4. The pellet is gently rinsed 3 times with purified water, being careful not to disturb the pellet.
   (a) If proceeding with option 2 (see above), go to HPLC sample preparation protocol below.
5. Prepare fresh organic solvent mixture composed of equal parts chloroform and isopropanol. A total volume of 32 mL of solvent mixture (16 mL of each) is used for every 100 mg of fusion protein.

(a) The final extraction solution will contain 5:5:1 chloroform:isopropanol:water (see below).

(b) NMR experiments of the peptide directly in the organic extraction mixture may be possible for some samples. However costly deuterated solvents must be used to avoid a strong solvent signal from confounding the peptide signal. To reduce the cost, the volume of solvent mixture can be reduced by up to 5-fold. Although this reduces the efficiency of extraction and the total yield of peptide, it increases the concentration of peptide in the organic solvent and consumes less solvent.

6. Add 25 mL of the solvent mixture to the precipitated protein and let it sit on ice for 15 min.

7. Using a metal spatula, mince the pellet and gently scrape the walls of the tube until the entire pellet has been resuspended.

   (a) Achieving a fine resuspension aids in the later steps.

8. Pour the mixture into an appropriately sized all-glass Dounce homogenizer and homogenize until the solution is uniform in appearance.

   (a) Care should be taken to avoid splashing the solvent during homogenization. A fume hood is recommended to avoid chloroform fume inhalation.

   (b) This can be done over the course of 30 min or more allowing the homogenizer to rest on ice in between homogenization sessions.

   (c) A fibrous pellet of insoluble denatured protein will appear at the bottom of the homogenizer (whitish in appearance).

9. Add 3.2 mL of pure water and continue to homogenize.

   (a) The pellet may become more translucent and fibrous.

   (b) This may be done over the course of 30 min or more allowing the homogenizer to rest on ice in between homogenization sessions.

10. Pour the entire contents of the homogenizer into a separatory funnel. Rinse the centrifuge tube and the homogenizer with the remaining solvent mixture (~7 mL) and pour into the funnel.

11. The funnel is shaken and vented every 30 min for 2 h or more and the layers are allowed to separate for >12 h.

    (a) If possible the mixture may be shaken and vented intermittently for the remainder of the day and left to separate overnight.

12. After the layers have separated completely the organic (bottom) layer is collected and added to a clean separatory funnel. One volume of purified water is added and step 11 is repeated.

    (a) The majority of remaining insoluble material and the aqueous layer should remain in the funnel but as the extraction is being repeated ensure that all the organic layer is collected even if a small amount of aqueous layer is also collected.

13. Collect the bottom layer, avoiding the collection of any aqueous phase, and store it in a glass container with a Teflon-lined or ground glass closure.

(a) This sample may be used directly for some experiments such as mass spectrometry or NMR.

(b) Step 12 can also be repeated to increase sample purity.

14. The organic solvent is removed from the sample by drying under nitrogen or argon gas, or by lyophilization. The sample should be analysed for purity by Tris-Tricine SDS-PAGE (16) (Figure 3.2). At this point, the sample may be used directly or further purification may be required (proceed to HPLC sample preparation protocol). It is recommended that the solubility of the peptide be tested in detergent solution or more polar solvent systems. It may also be possible to reconstitute the dried peptide into liposomes by various methods (17).

    (a) For solubility in detergent, dry 100 μg of peptide to a thin film on the wall of a glass test tube. This can be done by gently blowing a stream of nitrogen gas over the solvent while vortexing.

    (b) Add water (100–200 μL) and heat the solution (in the range of 37–50 °C) for 30 min. The peptide should become flocculent.

    (c) Bring solution to room temperature and add buffer components of choice (buffer, pH, salts) followed by detergent (e.g. 0.7 mg dodecylphosphocholine (DPC)) and vortex vigorously for 3 min. At this point, the clarity of the solution is a good indicator of peptide solubility.

The starting material for the following process can be (see options listed above): a 100,000g pellet, a TCA precipitated pellet, or the dried extracted sample from option 3. The pellet is solubilized in Guanidinium chloride (GuHCl) as follows.

HPLC sample preparation

1. Resuspend the pellet in 20 mL of 1 M buffered GuHCl per 100 mg of fusion protein using a spatula, vigorous vortexing and a Dounce homogenizer

2. Centrifuge the sample at 10,000g for 20 min at 4 °C to pellet insoluble material. Discard the supernatant.

    (a) This step should remove any remaining more soluble impurities like MBP and TEV from the sample.

    (b) The supernatant should be analysed by Tris-Tricine SDS-PAGE to ensure no significant peptide loss (16).

3. Resuspend the pellet in ~5 mL of 7 M buffered GuHCl per 100 mg of fusion protein.

    (a) Keeping the sample volume low will aid in HPLC column loading.

4. Centrifuge the sample at 10,000g for 20 min at 4 °C. Retain the supernatant.

5. Subject the supernatant to reverse-phase HPLC purification.

    (a) Hydrophobic peptide samples in GuHCl can be directly injected onto a reverse-phase HPLC semi-preparative column (e.g. a Zorbax SB300 C8 column, Agilent Technologies).

(b) Very hydrophobic peptides may be difficult or impossible to elute from the column, thereby reducing the life of the HPLC column. We find water-isopropanol-TFA gradients to be far superior to water-acetonitrile-TFA.

(c) Extensive washing of the column with high concentrations of isopropanol-TFA is required to remove unwanted contamination from peptides and GuHCl between runs.

6. Reverse-phase HPLC is a good polishing method for getting a highly pure peptide that can be solubilised in detergent or reconstituted into liposomes. However, very hydrophobic peptides may not separate or elute efficiently using this method. Organic extraction may serve as an alternative purification option, or it may be necessary to modify the peptide design and begin anew.

Commentary

Background Information

Unfortunately there is no *silver bullet* protocol that can be used for the expression and purification of any protein, and this is even more evident when considering membrane proteins. Many great resources for membrane protein expression in heterologous systems have already been published. A few recent examples include bacteria (18, 19), yeast (20–22), insect cells (23), mammalian cells (24), and cell-free systems (25). Herein, we present an application of the *divide and conquer* approach (26) using, as an example, a human membrane transport protein that is a recognized clinical target in heart disease and cancer, the human sodium proton exchanger isoform 1 (27). The basic tenet of our approach is that well-defined pieces of a membrane protein retain their native structure, yet they are easier to express, purify and analyse. Several groups including ours have successfully applied this method (28–33).

One of the central weaknesses of this approach is the correct selection of a transmembrane segment or subdomain in the absence of a validated topology model for the target membrane protein. Fortunately, there have been huge improvements in predictive algorithms that use both empirical and bioinformatic information to predict protein topology, secondary structure and even tertiary fold (many are listed on ExPASy). These tools combined with the wealth of biochemical data on many important membrane protein targets can be used as a base for the design of hypothetical transmembrane segments and subdomains that can be more readily studied *in vitro* (34). Of course, dividing any protein into pieces brings to question whether the structure of this

piece in isolation is relevant to the full-length physiological structure. While the inability to collect functional data on a membrane protein fragment prevents any surety of proper folding, there is ample support for the idea that isolated membrane segments and loops can remain properly folded. This is perhaps best exemplified by the numerous pioneering studies of Yeagle and Albert of the G-protein coupled receptor rhodopsin. Using chemically synthesized peptides and nuclear magnetic resonance (NMR) spectroscopy, the structure of the carboxy-terminal domain was determined (5). This success was followed up with the structures of all four individual cytoplasmic loops (35–37) and then the sixth transmembrane segment (33). As more crystal structures of membrane proteins were published various other studies demonstrated that structures of individual TM segments and loops are very similar to those in the full length structure (32, 38–40). This demonstration that individual TM segments in isolation may provide relevant structural information has led many researchers to begin studying individual transmembrane (TM) segments.

Critical Parameters and Troubleshooting

Peptide Design

There are many considerations and pitfalls in designing a peptide fragment when the membrane protein of interest has a complex topology. Unless the membrane protein target is sufficiently small (e.g. phospholamban (41) or sarcolipin (42)) or has a well-defined topology (e.g. GPCRs (43)), the ends of the putative membrane spanning regions are ambiguous. Transmembrane α-helices are generally considered to be 20 residues in length, yet the available membrane protein structures reveal that transmembrane regions can be tilted, kinked, unwound, or buried shorter or elongated helices. They can also cause local compaction or stretching of the membrane bilayer leading to a shorter or longer than expected transmembrane segment (44, 45). Nonetheless, many accurate predictive tools exist for the *in silico* study of membrane proteins owing to quickly expanding bioinformatic databases and improved sequence and structural alignment algorithms (46). No algorithm is able to correctly predict every transmembrane segment and it is therefore recommended to run a few different algorithms and then manually analyse the sequence to pick out the most likely candidate (see the example below). At this point, one needs to consider all biochemical information available for the membrane protein target. For example, cysteine-scanning mutagenesis combined with *in vivo* labelling by membrane permeant and

impermeant reactive compounds can help to define the limits of transmembrane segments and extra-membranous loops (8, 47, 48). In addition, transmembrane segments containing highly conserved residues that are critical for membrane protein function are often chosen first (49). Based on these analyses, a putative transmembrane segment is chosen including at least three extramembrane residues on the N- and C-termini. Often limited solubility of the transmembrane peptide in detergents or organic solvents is encountered, in which case it is advantageous to either lengthen the extra-membranous portions of the peptide or include two lysine residues on the N- and C-termini of the peptide. In the latter case, this has been shown to retain function (50, 51), as well as increase peptide solubility and increase the likelihood of correct insertion in membrane mimetics such as detergent micelles, bicelles, and liposomes (52). Many other considerations of peptide design reviewed by others (e.g. Cunningham and Deber (53)) may also be applicable.

Figure 3.1 depicts an example of peptide design using the human sodium proton exchanger isoform 1. Functional and structural data from mutagenesis and NMR studies (8, 11, 12) as well as two conflicting three dimensional models of the membrane domain of the protein (9, 54), made us interested in the region around amino acids 180–280. TMHMM analysis (7) of the sequence predicts 12 transmembrane segments. The sequence of residues 184–279 is displayed along with known functional and structural data, and three separate transmembrane predictions: TMHMM2, JPred3 (55), and evolutionary conservation/fold alignment (9). This demonstrates the importance of using the available data coupled with different prediction methods when designing a transmembrane construct.

Peptide production

Once the peptide is designed, there are many choices on how to obtain suitable quantities for structural analyses. Peptide synthesis is the simplest approach, though the strong hydrophobicity of many transmembrane peptides present challenges for their chemical synthesis and subsequent purification. Consequently, our personal experience has been that several peptide synthesis companies have either completely failed or refused to synthesize some of our constructs. If one chooses to use an expression system, as is our preference, there are many choices for the expression of small hydrophobic proteins. By using a simple affinity tag such as hexahistidine (His6), aggregated peptides, which tend to accumulate in inclusion bodies (56), can be solubilized and purified under denaturing conditions, then subsequently refolded by dilution or dialysis into de-

tergent or lipid suspensions (reviewed by Singh and Panda (57)). Although many peptides can be purified and properly refolded in this manner, their hydrophobic nature can promote aggregation and low peptide yields (58). Compounding this issue, hydrophobic peptides are generally toxic to the cell, which leads to down-regulation of expression or targeted degradation. An attractive alternative is to use a large soluble tag linked to the hydrophobic transmembrane peptide. The advantage is that a naturally abundant and highly soluble protein can be used both to increase expression and to maintain the transmembrane segment in a state suitable for purification. Ideally, when fused to the hydrophobic peptide, the large soluble protein tag would allow purification as a soluble protein, avoiding all of the caveats and complications associated with inclusion bodies.

Several solubility enhancing fusion protein systems are commercially available with the most popular being maltose binding protein (MBP) (6), glutathione-S-transferase (GST) (59), Mistic (60), NusA and small ubiquitin-like modifier (SUMO) (61). GST is a smaller tag, about 26 kDa, that modestly increases the solubility of the target protein, but allows facile and specific affinity purification. SUMO and Mistic are relatively recently characterized tags that have both been shown to be effective in increasing solubility of membrane proteins for expression and purification. MBP (42 kDa) has proven to be a robust system for transmembrane peptide expression (13, 29, 31, 62). This is because MBP has the remarkable ability to resist the denaturation forces imposed by fusion to an insoluble peptide sequence, thereby allowing high cellular expression levels of a soluble protein construct (6, 63). In our early efforts to purify phospholamban, a short single transmembrane protein, an MBP fusion protein maintained a soluble state, while a GST fusion was found in inclusion bodies (13).

Expression and purification of fusion protein

The next major consideration is bacterial strain and growth conditions. Due to its favourable expression, *Escherichia coli* strains like DH5α, normally reserved for cloning, are sometimes sufficient for expressing large amounts of fusion protein. Unfortunately this step is largely empirical depending on the construct. We have had success with DH5α, Rosetta DE3, XL1-Blue, BL21-pLysS and TB1. TB1 (JM83) is recommended by New England Biolabs (the manufacturer of the pMal vector system) for expression. Another important consideration is the type of media the cells are grown in. For a strain that expresses well, commonly used rich complex broths like LB and Terrific are low cost and typically result in high protein yields. However, if structural

studies are being pursued, M9 minimal media is recommended because it can be conveniently modified to allow isotopic labelling. Commonly, more difficult to express constructs seem to express better in M9 minimal media than in Lysogeny Broth (LB). Other media considerations to account for biosynthetic deficiencies may also be required, such as TB1 supplementation with proline when grown in minimal media. Generally minimal media is supplemented with building blocks, such as vitamins and amino acids, to improve bacterial growth. For specific labelling strategies, such as the use of ^{15}N for collecting two dimensional NMR spectra of a target peptide, minimal media can be generated with ^{15}N-labelled ammonium chloride as the sole source of nitrogen.

Once fusion protein expression is achieved with a construct-strain-media combination, several other variables can be optimized to improve the quality and quantity of the expressed material. Although typically bacteria are grown at 37 °C, inducible expression of exogenous proteins at this temperature often results in little or no fusion protein. Lowering the temperature during induction changes the cell state to favour protein synthesis (64). Our usual protocol is as follows. The cell culture is incubated at 37 °C until the beginning of logarithmic growth, corresponding to an OD_{600} of ~0.6–0.8 (for difficult constructs an earlier induction point, OD_{600} of 0.4–0.5, may be preferable). The culture is shifted to 22 °C and equilibrated for 30 min. Once the culture reaches 22 °C, the inducing agent is added (in our case isopropyl β-D-1-thiogalactopyranoside (IPTG)) and the cells are incubated for 12–72 h. The concentration of inducing agent can also affect fusion protein quality and quantity. For IPTG a range of 0.1 to 1 mM may be tested to optimize protein yield. Of course, the goal here is to optimize the quality and the level of protein expression (per cell) as well as the total cell mass such that large quantities of material suitable for structural analyses are obtained. Induction of expression at lower cell density, lower temperatures and with lower IPTG concentrations will require longer induction times (up to 72 h), while higher cell densities, temperatures and IPTG concentrations will generally require shorter induction times (e.g. >1 mM IPTG at 37 °C usually maximizes expression within 2–4 h). Some constructs are stable and express well with overnight induction at 37 °C, whereas other constructs degrade at 37 °C and require longer induction times at lower temperature. It is difficult to predict the expression conditions for a particular fusion protein construct, so a range of conditions should be empirically tested. Once a working expression protocol is determined, further optimization for large scale cultures may be performed. It is recommended that this is done using one litre

culture volumes, such that scale up simply involves multiple 1 L flasks and variables like aeration and growth times remain consistent between cultures.

Following over-expression, the purification scheme invariably starts with cell lysis, regardless of whether the target protein is found in inclusion bodies, cell membranes or the soluble cytosolic fraction. The purification scheme is: cell lysis, isolation of cytosolic fraction, maltose affinity chromatography, protease cleavage, and peptide purification (via organic extraction, and/or HPLC purification). Since MBP is such a robust system, little optimization of the fusion protein purification should be required. However, some constructs may express well as soluble proteins in the cytosol but become unstable during purification. If the fusion protein precipitates, supplement the affinity chromatography buffers with 10–20% glycerol and decrease the salt concentration by using 0.5x PSE (see recipe). Additionally all purification steps can be carried out at 4 °C. Following fusion protein purification we have occasionally experienced difficulty achieving complete cleavage of some constructs. Since the fusion protein and peptide should be stable in the elution buffer, samples can be left to digest for many weeks if required. Additional enzyme and fresh DTT every 3 days may also help. If cleavage is still poor, the linker between MBP and the peptide may be lengthened.

Peptide recovery

This is the most challenging step. After removal of MBP the peptide is no longer stable in solution. This can be used to our advantage through centrifugation, denaturants and organic solvents to collect and purify the peptide. If denaturants are used, not all peptides will spontaneously refold properly, especially multiple transmembrane segments. This makes HPLC purification challenging as the peptide is unlikely to be soluble in any solvent system that is also compatible with peptide retention on the column. If guanidinium chloride is not suitably solubilizing the peptide, urea may be used as an alternative. There is also choice in selecting the organic solvent for the extraction procedure. Although the peptide may be soluble in many different organic solvents, there are relatively few suitable two phase three component systems for the organic extraction. The other short chain alcohols, methanol and ethanol, have been used successfully in place of isopropanol, but longer chain alcohols like butanol are not effective. 1,2-dichloroethane, which has a similar polarity as chloroform, was also successful in extracting some peptides. While organic extraction yields highly pure peptide in some cases, not all peptides will be extracted

efficiently and many may require further purification by HPLC before structural and functional characterization is possible. Although the peptide may be purified to relative homogeneity, difficulties in moving the peptide into an environment suitable for further experiments may mean that different purification techniques may be required. More challenging constructs, those with lower expression or poor fusion protein purity or solubility, may also require further purification before cleavage (e.g. ion exchange or size exclusion chromatography) to increase the efficacy of peptide isolation.

Time considerations

Once optimized, the entire process of expression and purification takes four to ten days. There are many stopping points where the sample can be stored and the process continued later. For a typical construct, your experience may differ: the harvested bacterial pellet can be stored at -20 °C for months; the amylose column eluent (before or after concentration) can be stored at 4 °C for up to 3 days; the TEV digestion reaction is stable for at least one week at 16 °C; the organic extraction, if sealed, is stable for months at room temperature; and a lyophilized peptide, from HPLC or dried from organic extraction is stable for months under vacuum or at -80 °C. The optimization process for expression may take several weeks or months depending on the number of bacterial strains screened but adaptation of the fusion protein purification and protease cleavage should be simpler. Purifying the peptide following cleavage may take many months of optimization while some peptides may be impossible to successful isolate in a usable form for further characterization.

Acknowledgements

The authors would like to acknowledge several people who were instrumental in the optimization of the methods described above. JL Douglas and CA Trieber for early design and optimization of the expression and purification system; JP Glaves and PA Gorski for the optimization of expression in M9 media and the organic extraction technique; DK Ceholski and PA Gorski for the optimization of HPLC preparation techniques.

Reagents and Solutions

All the solutions described below, unless noted, are stable for at least 1 year and can be stored at room temperature. The media components and maltose are stable as long as they remain sterile. Pure water refers to Milli-Q-purified water or equivalent.

Miller Lysogeny Broth (LB)

In 800 mL of water dissolve: Tryptone, 10 g; Yeast Extract, 5 g; NaCl 10 g. Adjust the pH to 7.5 and dilute to a final volume of 1000 mL with pure water.

LB+amp plates

Prepare LB as directed above except add 1 g agar per 100 mL LB before autoclaving. After autoclaving, allow liquid to cool to ~50 °C. Add 1 mL of ampicillin stock (100 mg/mL, see below) mix gently (to prevent bubble formation), then pour into sterile petri plates. Allow agar to completely set for at least 60 min. For storage, turn the plates upside down and allow them to dry overnight on the bench. The following day stack the plates upside down into their original bag and store at 4 °C. Stable for at least 30 days.

M9 minimal medium

For 1 L mix together: 1g $(NH_4)_2SO_4$; 100mL 10x M9 salts; 673mL water. Autoclave and allow to cool. Using sterile technique add: 200 mL 5x phosphate solution; 1 mL 1000x Metal Mix; 1 mL 1% thiamine; 25 mL 20% glucose.

10x M9 salts

Dissolve the following in 1 L (final volume) of water and filter (0.45 μm): $Na_2HPO_4{\cdot}7\,H_2O$, 128 g (or 67.8g of anhydrous); KH_2PO_4, 30 g; NaCl, 5 g.

5x phosphates

Dissolve the following in 800 mL of water: K_2HPO_4, 53 g; KH_2PO_4, 24.7 g. Adjust pH to 7.5, dilute to 1 L, and autoclave.

1000x metal mix

Dissolve the following in 100 mL (final volume) of water and autoclave: $MnSO_4$, 500 mg; $FeSO_4{\cdot}7\,H_2O$, 92.5 mg; $MgSO_4{\cdot}7\,H_2O$, 5 g; $CaCl{\cdot}2\,H_2O$, 50 mg.

1% (w/v) thiamine (Vitamin B1)

Dissolve 1 g in 100 mL of pure water and autoclave.

20% (w/v) glucose

Dissolve 200 g of glucose in 1 L of pure water and filter sterilize or autoclave.

Cassamino acids (0.7% final) can also be added for non-labelled protein preparations. Make a 10% (w/v) stock solution, autoclave and add sterilely to M9 medium.

Ampicillin stock

Dissolve 1 g of ampicillin in 10 mL of pure water, filter sterilize into a sterile container and aliquot into 1 mL volumes in sterile microcentrifuge tubes. Store at -20 °C.

80% sterile glycerol

Mix 80 mL of glycerol with 20 mL of pure water and autoclave. Store at room temperature.

0.5 M IPTG

Dissolve 1.19 g of isopropyl β-D-1-thiogalactopyranoside in 10 mL of pure water, filter sterilize into a sterile container and aliquot into 1 mL volumes in sterile microcentrifuge tubes. Store at -20 °C.

10x PSE (Phosphate-Sodium-EDTA)

Dissolve the following in 1 L (final volume) of water: Na_2HPO_4, 1.462 g; NaH_2PO_4, 26.88 g; NaCl, 70.2 g; EDTA, 0.372g.

50% (v/v) glycerol

Dilute 500 mL glycerol to 1 L with water.

5% (w/v) NaN_3

Dissolve 5 g of $NNaN_3$ in 100 mL (final volume) of water.

500 mM maltose

Dissolve 90 g of maltose in 500 mL (final volume) of water, autoclave.

Purification buffer

Mix together the following and dilute to 1L with water: 100 or 50 mL 10x PSE[1]; 4 mL 5% NaN_3; 400 mL 50% glycerol[2].

Lysis buffer

Add 25 µL HALT Protease Cocktail and 100 µL of 1 M DTT (see below) to 100 mL of purification buffer. Make fresh daily.

Elution buffer

Dilute 12 mL 500 mM maltose to 100 mL with purification buffer (60 mM final). Make fresh daily.

1 M DTT

Dissolve 1.54 g of dithiothreitol in 10 mL of pure water. Aliquot into 1 mL volumes in microcentrifuge tubes. Store at -20 °C.

60% (w/v) trichloroacetic acid (TCA)

Dissolve 60 g of TCA in 100 mL (final volume) of water.

1 M buffered GuHCl

Mix together: 8 M GuHCl, 3.252 mL; 0.1 M Na-phosphate buffer, pH 8, 12.5 mL and dilute to 25 mL with water.

7M buffered GuHCl

Mix together: 8 M GuHCl, 21.875 mL; 0.1 M Na2HPO4, 2.5 mL[3].

8M Guanadinium chloride (GuHCl)

Dissolve 76.424 g of guanidinium chloride in 100 mL (final volume) of water.

1.1 M phosphate buffer, pH 8

[1] Using less PSE may increase protein solubility

[2] optional, increases protein stability

[3] not buffered, the final pH will be around 7.15

Mix together: 0.2 M NaH_2PO_4, 8.5 mL; 0.2 M Na_2HPO_4, 91.5 mL.

0.2 M NaH_2PO_4
Dissolve 5.52 g of NaH_2PO_4 in 1 L (final volume) of water.

0.2 M Na_2HPO_4
Dissolve 5.68 g of Na_2HPO_4 in 1 L (final volume) of water.

References

1. Kemp, G., Larry, F. & Young, H. S. Membrane transport piece by piece: Production of transmembrane peptides for structural and functional studies. In *Current Protocols in Protein Science* (John Wiley & Sons, Inc., 2001).

2. Wallin, E. & von Heijne, G. (1998). Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci* **7**, 1029–1038.

3. Raman, P., Cherezov, V. & Caffrey, M. (2006). The membrane protein data bank. *Cell Mol Life Sci* **63**, 36–51.

4. Junge, F. *et al.* (2008). Large-scale production of functional membrane proteins. *Cell Mol Life Sci* **65**, 1729–1755.

5. Yeagle, P. L., Alderfer, J. L. & Albert, A. D. (1995). Structure of the carboxy-terminal domain of bovine rhodopsin. *Nat Struct Biol* **2**, 832–834.

6. Kapust, R. B. & Waugh, D. S. (1999). *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci* **8**, 1668–1674.

7. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *J Mol Biol* **305**, 567–580.

8. Wakabayashi, S., Pang, T., Su, X. & Shigekawa, M. (2000). A novel topology model of the human Na^+/H^+ exchanger isoform 1. *J Biol Chem* **275**, 7942–7949.

9. Landau, M., Herz, K., Padan, E. & Ben-Tal, N. (2007). Model structure of the Na^+/H^+ exchanger 1 (NHE1): functional and clinical implications. *J Biol Chem* **282**, 37854–37863.

10. Cole, C., Barber, J. D. & Barton, G. J. (2008). The Jpred 3 secondary structure prediction server. *Nucleic Acids Res* **36**, W197–W201.

11. Tzeng, J., Lee, B. L., Sykes, B. D. & Fliegel, L. (2010). Structural and functional analysis of transmembrane segment VI of the NHE1 isoform of the Na^+/H^+ exchanger. *J Biol Chem* **285**, 36656–36665.

12. Ding, J., Rainey, J. K., Xu, C., Sykes, B. D. & Fliegel, L. (2006). Structural and functional characterization of transmembrane segment VII of the Na^+/H^+ exchanger isoform 1. *J Biol Chem* **281**, 29817–29829.

13. Douglas, J. L., Trieber, C. A., Afara, M. & Young, H. S. (2005). Rapid, high-yield expression and purification of Ca^{2+}-ATPase regulatory proteins for high-resolution structural studies. *Protein Expression Purif* **40**, 118–125.

14. Nagai, K., Perutz, M. F. & Poyart, C. (1985). Oxygen binding properties of human mutant hemoglobins synthesized in *Escherichia coli*. *Proc Natl Acad Sci U S A* **82**, 7252–7255.

15. Routzahn, K. M. & Waugh, D. S. (2002). Differential effects of supplementary affinity tags on the solubility of MBP fusion proteins. *J Struct Funct Genomics* **2**, 83–92.

16. Schägger, H. (2006). Tricine-SDS-PAGE. *Nat Protoc* **1**, 16–22.

17. Reddy, L. *et al.* (1995). Functional reconstitution of recombinant phospholamban with rabbit skeletal Ca^{2+}-ATPase. *J Biol Chem* **270**, 9390–9397.

18. Newby, Z. E. R. *et al.* (2009). A general protocol for the crystallization of membrane proteins for x-ray structural investigation. *Nat Protoc* **4**, 619–637.

19. Miroux, B. & Walker, J. E. (1996). Over-production of proteins *Escherichia coli*: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol* **260**, 289–298.

20. Jidenko, M. *et al.* (2005). Crystallization of a mammalian membrane protein overexpressed in saccharomyces cerevisiae. *Proc Natl Acad Sci U S A* **102**, 11687–11691.

21. Cregg, J. M. *et al.* (2009). Expression in the yeast *Pichia pastoris*. *Methods Enzymol* **463**, 169–189.

22. Daly, R. & Hearn, M. T. W. (2005). Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. *J Mol Recognit* **18**, 119–138.

23. Trometer, C. & Falson, P. (2010). Mammalian membrane protein expression in baculovirus-infected insect cells. *Methods Mol Biol* **601**, 105–117.

24. Lundstrom, K. (2010). Expression of mammalian membrane proteins in mammalian cells using semliki forest virus vectors. *Methods Mol Biol* **601**, 149–163.

25. Klammt, C., Schwarz, D., Dötsch, V. & Bernhard, F. (2007). Cell-free production of integral membrane proteins on a preparative scale. *Methods Mol Biol* **375**, 57–78.

26. White, S. H., Ladokhin, A. S., Jayasinghe, S. & Hristova, K. (2001). How membranes shape protein structure. *J Biol Chem* **276**, 32395–32398.

27. Karmazyn, M., Avkiran, M. & Fliegel, L. *The sodium-hydrogen exchanger: from molecule to its role in disease* (Kluwer Academic Publishers, 2003).

28. Kocherla, H. *et al.* (2012). Biosynthesis and spectroscopic characterization of 2-TM fragments encompassing the sequence of a human GPCR, the y4 receptor. *ChemBioChem* **13**, 818–828.

29. Lee, B. L., Sykes, B. D. & Fliegel, L. (2011). Structural analysis of the Na^+/H^+ exchanger isoform 1 (NHE1) using the divide and conquer approach. *Biochem Cell Biol* **89**, 189–199.

30. Bordag, N. & Keller, S. (2010). α-helical transmembrane peptides: A Divide and Conquer approach to membrane proteins. *Chem Phys Lipids* **163**, 1–26.

31. Hu, J. *et al.* (2007). Structural biology of transmembrane domains: Efficient production and characterization of transmembrane peptides by NMR. *Protein Sci* **16**, 2153–2165.

32. Katragadda, M., Alderfer, J. L. & Yeagle, P. L. (2001). Assembly of a polytopic membrane protein structure from the solution structures of overlapping peptide fragments of bacteriorhodopsin. *Biophys J* **81**, 1029–1036.

33. Chopra, A., Yeagle, P. L., Alderfer, J. A. & Albert, A. D. (2000). Solution structure of the sixth transmembrane helix of the G-protein-coupled receptor, rhodopsin. *Biochim Biophys Acta, Biomembr* **1463**, 1–5.

34. Rost, B., Sander, C., Casadio, R. & Fariselli, P. (1995). Transmembrane helices predicted at 95% accuracy. *Protein Sci* **4**, 521–533.

35. Yeagle, P. L., Alderfer, J. L. & Albert, A. D. (1995). Structure of the third cytoplasmic loop of bovine rhodopsin. *Biochemistry* **34**, 14621–14625.

36. Yeagle, P. L., Alderfer, J. L. & Albert, A. D. (1996). Structure determination of the fourth cytoplasmic loop and carboxyl terminal domain of bovine rhodopsin. *Mol Vis* **2**, 12.

37. Yeagle, P. L., Alderfer, J. L., Salloum, A. C., Ali, L. & Albert, A. D. (1997). The first and second cytoplasmic loops of the G-protein receptor, rhodopsin, independently form beta-turns. *Biochemistry* **36**, 3864–3869.

38. Duff, K. C. & Ashley, R. H. (1992). The transmembrane domain of influenza A M2 protein forms amantadine-sensitive proton channels in planar lipid bilayers. *Virology* **190**, 485–489.

39. Reddy, G. L., Iwamoto, T., Tomich, J. M. & Montal, M. (1993). Synthetic peptides and four-helix bundle proteins as model systems for the pore-forming structure of channel proteins. II. transmembrane segment M2 of the brain glycine receptor is a plausible candidate for the pore-lining structure. *J Biol Chem* **268**, 14608–14615.

40. Katragadda, M., Alderfer, J. L. & Yeagle, P. L. (2000). Solution structure of the loops of bacteriorhodopsin closely resembles the crystal structure. *Biochim Biophys Acta* **1466**, 1–6.

41. Simmerman, H. K., Collins, J. H., Theibert, J. L., Wegener, A. D. & Jones, L. R. (1986). Sequence analysis of phospholamban. identification of phosphorylation sites and two major structural domains. *J Biol Chem* **261**, 13333–13341.

42. Wawrzynow, A. *et al.* (1992). Sarcolipin, the "proteolipid" of skeletal muscle sarcoplasmic reticulum, is a unique, amphipathic, 31-residue peptide. *Arch Biochem Biophys* **298**, 620–623.

43. Bhave, G. *et al.* (2003). Membrane topology of a metabotropic glutamate receptor. *J Biol Chem* **278**, 30294–30301.

44. de Planque, M. R. R. & Killian, J. A. (2003). Protein-lipid interactions studied with designed transmembrane peptides: role of hydrophobic matching and interfacial anchoring. *Mol Membr Biol* **20**, 271–284.

45. Cybulski, L. E. & de Mendoza, D. (2011). Bilayer hydrophobic thickness and integral membrane protein function. *Current protein & peptide science* **12**, 760–766.

46. Punta, M. *et al.* (2007). Membrane protein prediction methods. *Methods (San Diego, Calif.)* **41**, 460–474.

47. Akabas, M. H., Stauffer, D. A., Xu, M. & Karlin, A. (1992). Acetylcholine receptor channel structure probed in cysteine-substitution mutants. *Science* **258**, 307–310.

48. Tang, X.-B., Fujinaga, J., Kopito, R. & Casey, J. R. (1998). Topology of the region surrounding Glu681 of human AE1 protein, the erythrocyte anion exchanger. *J Biol Chem* **273**, 22545–22553.

49. Slepkov, E. R. *et al.* (2005). Structural and functional characterization of transmembrane segment IV of the NHE1 isoform of the Na^+/H^+ exchanger. *J Biol Chem* **280**, 17863–17872.

50. Afara, M. R., Trieber, C. A., Glaves, J. P. & Young, H. S. (2006). Rational design of peptide inhibitors of the sarcoplasmic reticulum calcium pump. *Biochemistry* **45**, 8617–8627.

51. Afara, M. R., Trieber, C. A., Ceholski, D. K. & Young, H. S. (2008). Peptide inhibitors use two related mechanisms to alter the apparent calcium affinity of the sarcoplasmic reticulum calcium pump. *Biochemistry* **47**, 9522–9530.

52. Melnyk, R. A. *et al.* (2003). Polar residue tagging of transmembrane peptides. *Pept Sci* **71**, 675–685.

53. Cunningham, F. & Deber, C. M. (2007). Optimizing synthesis and expression of transmembrane peptides and proteins. *Methods (San Diego, Calif.)* **41**, 370–380.

54. Nygaard, E. B. *et al.* (2011). Structural modeling and electron paramagnetic resonance spectroscopy of the human Na^+/H^+ exchanger isoform 1, NHE1. *J Biol Chem* **286**, 634–648.

55. Cuff, J. A. & Barton, G. J. (1999). Evaluation and improvement of multiple sequence methods for protein secondary structure prediction. *Proteins: Struct , Funct , Bioinf* **34**, 508–519.

56. Kane, J. F. & Hartley, D. L. (1988). Formation of recombinant protein inclusion bodies in *Escherichia coli*. *Trends Biotechnol* **6**, 95–101.

57. Singh, S. M. & Panda, A. K. (2005). Solubilization and refolding of bacterial inclusion body proteins. *J Biosci Bioeng* **99**, 303–310.

58. De Bernardez Clark, E. (1998). Refolding of recombinant proteins. *Curr Opin Biotechnol* **9**, 157.

59. Bichet, P., Mollat, P., Capdevila, C. & Sarubbi, E. (2000). Endogenous glutathione-binding proteins of insect cell lines: characterization and removal from glutathione S-transferase (GST) fusion proteins. *Protein Expression Purif* **19**, 197–201.

60. Roosild, T. P. *et al.* (2005). NMR structure of mistic, a membrane-integrating protein for membrane protein expression. *Science* **307**, 1317–1321.

61. Zuo, X. *et al.* (2005). Enhanced expression and purification of membrane proteins by SUMO fusion in *Escherichia coli*. *J Struct Funct Genomics* **6**, 103–111.

62. Ceholski, D. K., Trieber, C. A., Holmes, C. F. B. & Young, H. S. (2012). Lethal, hereditary mutants of phospholamban elude phosphorylation by protein kinase A. *J Biol Chem* **287**, 26596–26605.

63. Nallamsetty, S. & Waugh, D. S. (2006). Solubility-enhancing proteins MBP and NusA play a passive role in the folding of their fusion partners. *Protein Expression Purif* **45**, 175–182.

64. Jones, P. G. & Inouye, M. (1994). The cold-shock response–a hot topic. *Mol Microbiol* **11**, 811–818.

Chapter 4

Structural and Functional Analysis of Transmembrane Segment IV of the Salt Tolerance Protein Sod2

This research was originally published in The Journal of Biological Chemistry. Asad Ullah, Grant Kemp, Brian Lee, Claudia Alves, Howard S. Young, Brian D. Sykes and Larry Fliegel. Structural and Functional Analysis of Transmembrane Segment IV of the Salt Tolerance Protein Sod2. *J. Biol. Chem.* 2013; **288**(34):24609–24624. As second author this chapter represents my contributions to this research, with the exception of data that was included for clarity. Notably, the NMR data collection and analysis was done by Brian Lee while the majority of the NMR sample preparation and structural modelling was completed by myself in close collaboration with Brian Lee. The functional data was collected and analyzed by Dr. Asad Ullah.

Introduction

Under normal physiological conditions, plants, yeast and mammalian cells have relatively low Na^+ concentrations in their cytosol. Since the external Na^+ concentration is much higher than internal Na^+ concentration, an accumulation of intracellular Na^+ ions results. Organisms respond

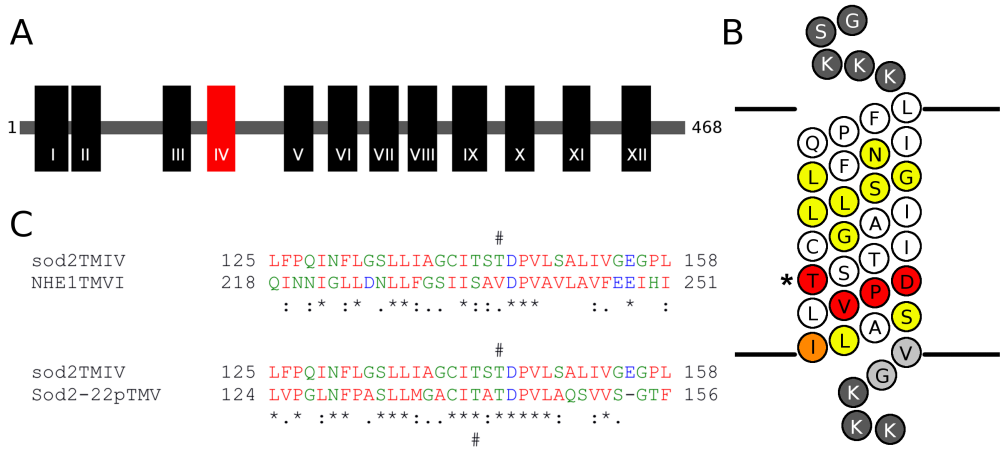

Figure 4.1: sod2 topology and TM IV alignment. ***A***. Topology of full length *S. pombe* sod2 protein modified from Wiebe *et al.* (5). Length of each labelled transmembrane segment (black rectangles) and their relative placement along the sequence (grey bar) are to scale. The red box is TM IV. ***B***. A 2D representation of TM IV mapped with known mutations (see Table 4.1 for details): group I mutations, white; group II, yellow; group III, red; group V, orange; and group IV (T144S) is labelled with an (*). sod2 amino acids that were not mutated are labelled in light grey and exogenous amino acids incorporated into the peptide used for NMR experiments are in dark grey. ***C***. Labelled Clustal Ω alignment of TM IV with the putatively analogous human NHE1 TM VI and *Zygosaccharomyces rouxii* Sod2-22p TM V. Residues are coloured by type: red, hydrophobic; blue, acidic; green, polar; magenta, basic; and labelled with Gonnet PAM 250 amino acid conservation scores: (*****) identical; (**:**) strongly similar; (**.**) weakly similar.

to this salt stress in several ways. Plants and yeast deal with these excess *toxic* levels of intracellular Na^+ primarily by either extruding it, or by sequestering it into vacuoles, thereby reducing the cytosolic concentration. This process is mediated by transporters and other regulatory proteins. In the fission yeast *Schizosaccharomyces pombe*, the Na^+/H^+ antiporter (sod2) is responsible for most of the salt removal from the cytosol (1). This protein functions by using the external proton gradient to pump out internal sodium ions. Disruption of this gene results in a reduced extrusion of cytoplasmic Na^+ and a decreased tolerance of external Na^+ (1). Sod2 removes both Na^+ and Li^+ from the cytosol using the proton gradient created by the plasma membrane ATPase (1, 2). We have previously (3, 4) used *S. pombe* with a knockout of the sod2 gene to study the effects of mutation of amino acids in this protein. Because of the limited number of other salt tolerance mechanisms in this species, removal of this gene results in a severe salt tolerance phenotype (1, 3, 4). This makes *S. pombe* a very useful organism for the study of salt tolerance proteins.

Sod2 belongs to the cation proton antiporter 2 family, which shares its origins with prokaryotic NhaA. The mechanism of transport of Na^+/H^+ exchangers is of great interest both because

of the potential to improve salt tolerance in plants and make salt resistant phenotypes, but also as a fundamental scientific problem. Significant progress has been made in the understanding of bacterial transport by NhaA (6), however the eukaryotic transporters are not as well understood. They have a different exchange stoichiometry and are activated by different physiological conditions. Although we have earlier examined several key residues of sod2 (3, 4, 7, 8), the mammalian NHE1 is the most well characterized eukaryotic Na^+/H^+ exchanger. We have determined the structure of several membrane associated fragments (9–15) as well as elucidated many essential residues. One of these critical TM segments in NHE1 is TM IV, which has been compared to the important TM IV of *E. coli* NhaA (10, 16). The conclusions from this study, in combination with earlier findings, suggest that TM IV of the yeast salt tolerance protein sod2 is analogous to NHE1 TM VI and NhaA TM IV. Our results demonstrate that this segment is critical to sod2 function with amino acids 144-147 comprising part of a region critical for transport. This study is the first structural and functional characterization of an entire transmembrane segment of a yeast salt tolerance membrane protein and demonstrates that this region is critical for cation selectivity and salt tolerance.

Experimental Procedures

Materials, Strains and Media

Restriction enzymes were obtained from New England Biolabs, Inc. (Mississauga ON, Canada). PWO DNA polymerase was obtained from Roche Applied Science (Roche Molecular Biochemicals, Mannheim, Germany).

S. pombe bearing the sod2 gene disruption (sod2::ura4) was used for all transformations and as a control where indicated (3). It was maintained on low sodium minimal KMA medium or yeast extract adenine (YEA) using methods described earlier (1, 3). KMA medium was used where indicated and contains (per 1 liter): potassium hydrogen phthalate, 3 g; K_2HPO_4, 3 g; yeast nitrogen base without amino acids, 7 g; glucose, 20 g; and adenine, 200 mg. Leucine at 200 mg/l was added to maintain the sod2::ura4 *leu1-32* strain where indicated and all media was buffered using 50 mM MES/Citrate and adjusted to pH 5.0 with KOH. *S. pombe* containing the pREP-41sod2GFP plasmid (and mutant derivatives) were routinely grown in medium in the

absence of thiamine. Cultures were grown at 30 °C with constant agitation using a rotary shaker. The plasmid pREP-41sod2GFP has been described earlier (17) and contains the entire sod2 gene plus a C-terminal GFP tag separated by a nine amino acid Gly-Ala spacer (Figure 4.2).

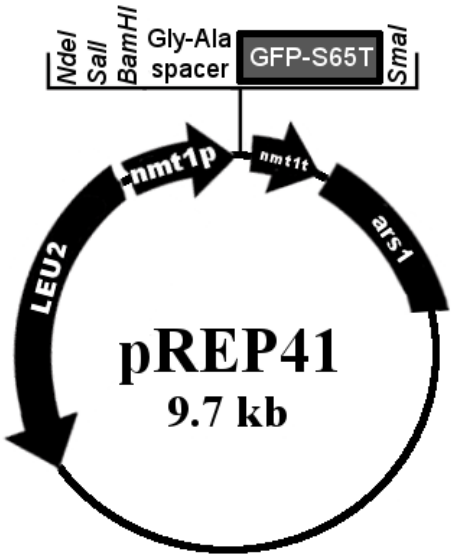

Figure 4.2: Plasmid map of *S. pombe* expression vector, pREP-41GFP. Nutritional selection of auxotrophic yeast in KMA media can be obtained via leucine limitation (LEU2). sod2 is cloned into the multiple cloning site (indicated by the large square bracket at the top of the map) between the *NdeI* and *BamHI* restriction sites. This allows the expression of a C-terminal GFP(S65T) fusion protein from the constitutive nmt1 promoter (nmt1p) with the complementary terminator (nmt1t). Nmt1 is a thiamine biosynthetic promoter and allows for consistent expression. The other element is the alanine tRNA ligase (ars1) which is not employed in this research.

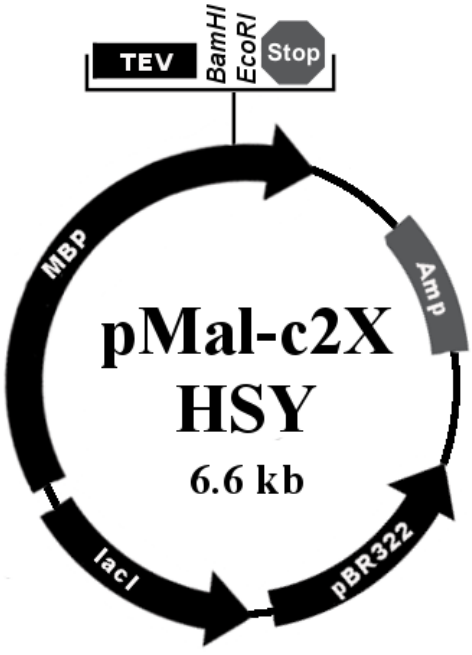

Figure 4.3: Plasmid map of maltose binding protein fusion expression vector, pMal-c2X-HSY. Designed for *E. coli* expression of target protein with C-terminal TEV cleavable maltose binding protein. Protein of interest is cloned between the *BamHI* and *EcoRI* sites and the plasmid can be selected in *E. coli* using ampicillin (Amp). Replication of the plasmid is through the pBR322 origin. The LacI element is not used in this research.

sod2 TM IV was expressed in XL1-Blue cells as a maltose binding protein fusion using a modified pMAL-c2x containing a tobacco etch virus (TEV) protease cleavage site between maltose binding protein and the peptide of interest (Figure 4.3). Expression was induced with IPTG in either LB (for unlabelled peptide) or M9 (for labelled samples) media. Expression was initially optimized with: LB, 0.6 mM IPTG, 37 °C, 24 h; and M9, 1 mM IPTG, 22 °C, 48 h. LB media contains: tryptone, 1% w/v; yeast extract, 0.5% w/v; NaCl, 1% w/v. M9 media contains: $(^{15}NH_4)_2SO_4$, 7.5 mM; glucose, 28 mM; thiamine, 30 μM; [Na_2HPO_4, 47 mM; KH_2PO_4, 22 mM; NaCl, 8.5 mM]; [$MnSO_4 \cdot 1H_2O$, 30 μM; $FeSO_4 \cdot 7H_2O$, 3.3 μM; $MgSO_4 \cdot 7H_2O$, 203 μM; $CaCl_2 \cdot 2H_2O$, 3.4 μM]; [K_2HPO_4, 60 mM; KH_2PO_4, 36 mM; pH adjusted to 7.5 with KOH]; where the individual components or mixtures contained within the [ ] were autoclaved separately and combined afterwards.

Trypsin Treatment of Microsomal Membranes

Yeast microsomal membranes were prepared as follows. Typically membranes were isolated from a 50 ml yeast culture transformed with either wild type or mutant pREP-41sod2GFP. Yeast cells were grown in KMA to an OD_{600} of 2 at 30 °C. Cells were harvested (3500g , 10 min), washed with double distilled water and resuspended in lysis buffer (50 mM Tris-HCl, pH 8.0; 5 mM EDTA; protease inhibitor cocktail (18); 1 mM dithiothreitol). Cells were lysed in a Bullet Blender® using 0.5 mm zirconium oxide beads (speed setting: 10) for 40 minutes. Alternatively they were passed through an Emulsiflex homogenizer at a pressure of 25,000 psi. Unbroken cells were harvested by centrifugation and this supernatant was enriched for microsomal membranes by centrifugation at 14,000g for 10 min. The membranes were then pelleted from this supernatant at 100,000g for 1 h and resuspended at 2 mg/ml in 1 mM EDTA, adjusted to pH 7.4 with Tris HCl. For storage, 15% glycerol (v/v) was added and small aliquots made for flash freezing in liquid N_2. For digestion trypsin (phenylalanyl chloromethyl ketone-trypsin, Sigma, St., Louis, MO) was added to an aliquot to yield a trypsin:protein ratio (1:200) and was incubated at 30 °C, for the times indicated. The reaction was terminated by the addition of SDS-PAGE sample buffer and samples were resolved on 12% SDS-PAGE gels. Nitrocellulose transfers were immunostained using a primary polyclonal anti-GFP antibody (generously provided by Dr. Luc Berthiaume, Dept. of Cell Biology, University of Alberta), labelled with an IRDye 680 -conjugated goat anti-rabbit polyclonal antibody (Bio/Can, Mississauga, ON, Canada) and detected using the Odyssey

scanning system (LI-COR Biosciences, USA) (4, 19). Western blot bands were quantified using ImageQuant TL software (GE Healthcare Life Sciences).

Sod2 TM IV peptide purification

Amino acids ^{125}LFPQINFLGSLLIAGCITSTDPVLSALIVG^{154} of sod2 were expressed as a fusion protein with maltose binding protein. The design was such that we also introduced 3 additional N- and C- terminal lysines to the peptide. This was done to aid in the solubility of the peptide, as described earlier (11, 14). The primers:

SodMBPf: 5'–CATGGGATCCAAAAAAAAATTGTTTCCACAAATTAACTTTTTAGG–3'

SodMBPr: 5'–CCGGGAATTCTCATTTCTTTTTTCCTACAATCAATGCTGATAG–3'

were used to amplify the DNA of sod2 while adding the terminal lysines. They were designed to allow for in frame expression with MBP in a modified pMal-c2X plasmid (20) (pMal-c2X-HSY, Figure 4.3) which also has a tobacco etch virus (TEV) protease site upstream of the insert to allow for cleavage from the MBP. PCR was performed using the pREP-41sod2GFP plasmid as a template and the PCR product contained the TM IV sequence with the modified termini flanked by a 5' *BamHI* restriction endonuclease site and a 3' *EcoRI* site. After cloning into the modified pMal-c2X vector mentioned above, the plasmid construction was confirmed by DNA sequencing.

The *E. coli* strain XL1 blue was used for expression. The purification protocol is detailed in *Chapter 2*. The salient optimizations are as follows. The purification buffer (PB) contained (per 1 liter): 10 mM sodium phosphate buffer, pH 7; 60 mM NaCl; 0.5 mM EDTA; 20% v/v glycerol; 0.02% NaN_3 plus added HALT protease inhibitor cocktail (EDTA-free, Thermo Scientific, Rockford, IL, USA). The cells were lysed using ultrasonication (Branson Sonifier, Emerson Industrial Automation, Danbury, CT, USA) or high-pressure homogenization (Emulsiflex-C3, Avestin Inc., Ottawa, Canada). The sod2 fragment was cleaved free of the MBP using TEV protease at 10 U/mg of fusion protein for 48–72 h at 16 °C. The protein was precipitated with trichloroacetic acid, pelleted by centrifugation and subjected to two rounds of organic extraction with $CHCl_3$:isopropanol:water (ratio 5:5:1). Most complete extraction of the hydrophobic peptide was achieved at 1 ml = 1 vol per 30 mg fusion protein (i.e. [5 ml $CHCl_3$:5 ml isopropanol:1 ml water]/30 mg fusion protein. For studies of unlabelled peptide, deuterated solvents (Cam-

bridge Isotope Laboratories, Andover, MA) were used and the extraction was carried out at higher protein concentrations (i.e. 1 vol per 100 mg fusion protein). For ^{15}N labelled samples, $CDCl_3$ and undeuterated isopropanol were used. Purification by HPLC was not required. The identity of the purified peptide was confirmed by MALDI-TOF mass spectrometry. The purity of the sod2-TM IV peptide was estimated at over 95%.

NMR spectroscopy

Peptide samples obtained from liquid-liquid extraction were used for high-resolution NMR spectroscopy. After testing with various membrane mimetics, examination of 1D ^{1}H NMR and 2D ^{15}N HSQC NMR spectra indicated that the best solvent system for sod2 TM IV was $CDCl_3$/IPA-d8. By drying the organic extract under a gentle stream of nitrogen gas an adequately concentrated sample for structure determination was obtained. $CDCl_3$ in the extract was used as a lock solvent and spectra were referenced to tetramethylsilane added to the solvent.

All NMR spectra were acquired at 500 MHz and 30 °C. 2D ^{15}N HSQC (21), 3D ^{15}N NOESY-HSQC (150 ms mixing time) (22), TOCSY-HSQC (50 ms mixing time) (22), and HNHA (23, 24) spectra were collected with VnmrJ (Varian Inc.) and processed in NMRPipe (25). Resonance assignment and analysis was performed in NMRViewJ (26). Xplor-NIH (27) was used to model the secondary structure of the peptide.

Homology modelling

A homology model of sod2 was created with the program Modeller (28) using the crystal structure of *Escherichia coli* NhaA (6) (PDB entry: 1ZCD). *Arabidopsis thaliana* SOS1 was used as a mediator to align sod2 and NhaA using Clustal Ω (29) on the EMBL-EBI server (30). The alignment was supported and refined by using the TMHMM2 α–helical transmembrane segment prediction algorithm (31) and the JPred3 secondary structure prediction algorithm (32). These data, along with the known position of transmembrane helices in NhaA, were used to adjust the alignment to best reflect the most likely position of transmembrane segments in sod2. Short helical restraints (<10 residues) were given to Modeller, in regions where longer helices were predicted. The validity of the model was analyzed using a protein validation software suite (33) and the ConSurf method (34, 35) on the online server (36).

Results

Selection of TM IV

We examined amino acids of a transmembrane segment of the yeast *S. pombe* salt tolerance protein sod2 that were thought to be important for activity. While the transmembrane segment assignments are not certain, amino acids 126–151 were assigned as TM IV based on a previously published (5) model of the protein (Figure 4.1-A) using hydrophobicity analysis. This segment (FPQINFLGSLLIAGCITSTDPVLSAL) is mainly comprised of very hydrophobic residues, with a few polar residues interspersed (Figure 4.1-B). As the borders of the membrane lipid interface are not assured 26 amino acids were chosen for purposes of mutational analysis—longer than required for a typical α–helical transmembrane segment to cross a lipid bilayer (37).

We compared the amino acid sequence of sod2 with several other related Na^+/H^+ exchanger proteins using the program Clustal Ω (29) (multiple sequence alignment tool from EMBL-EBI (30)). Sod2 did not align well directly with the complete sequences of human NHE1 or *E. coli* NhaA although isolated regions did show some similarity (not shown). The region around the putative TM IV did align well with NHE1 TM VI and the related *Zygosaccharomyces rouxii* Sod2-22p TM V (Figure 4.1-C). We have recently (14) shown that TM VI of NHE1, rather than TM IV, is structurally related to TM IV of NhaA which may account for this finding (see 6). However the closely associated salt tolerance protein SOS1 from *Arabidopsis thaliana* (38) showed significant areas of identity and similarity especially in the sequence proposed to be TM IV. Fortunately SOS1 also closely resembles NhaA permitting a good three-way alignment of sod2, SOS1 and NHE1 TMVI (Figure 4.4). This provided a good starting point for homology modelling with Modeller.

Functional effect of TM IV mutagenesis

Since sod2 is the major salt exporter of *S. pombe* (17), we measured the ability of sod2 TM IV mutants to rescue salt tolerance in the salt sensitive sod2::ura4 deletion strain. The results of these experiments, performed by Dr. Ullah, including the growth phenotype and transport ability of sod2 mutants, have been summarized in Table 4.1. Further data can be found in the published work (39).

```
sod2        --------MGWRQLDI---------------DKVHLALIVAGGFITFFCYFSEVFRKKL 36
SOS1        MTTVIDATMAYRFLEEATDSSSSSSSSKLESSPVDAVLFVGMS------LVLGIASRHL 53
NhaA        -----------MKHLH---------------RFFSSDASGGIILIIAAILAMIMAN--- 30

sod2        LVGEAVLGSIT----GLIFGPHAAKLV--------DPFSWGDHGDYLTVEICRIVLDVR 83
SOS1        LRGTRVPYTVALLVIGIALGSLEYGAKHNLGKIGHGIRIWNEIDPELL---------- 101
NhaA        -SGATSGWYHD----FLETPVQLRVGS--------LE-----INKNMLLWINDALMAVF 71

sod2        VFASAIELPGAYFQHNFR----------SIIVM-----LLPVMAYGWLVTAGFAYALF 126
SOS1        ---LAVFLPALLFESSFSMEVHQIKRCLGQMVLLAVPGVLISTACLGSLVKVTFPY--- 154
NhaA        FLLVGLEVKRELMQGSLA----------SLRQA-----AFPVIAAIGGMIVPALLYLA 114
                                                                    *
                         TMIV
sod2        PQINFLGSLLIAGCITSTDPVLSALIVGEGPLAKKTPERI--RSLLIAESGCNDGMAVP 183
SOS1        -EWDWKTSLLLGGLLSATDPVAVVALLKELGAS----KKL--STIIEGESLMNDGTAIV 206
NhaA        FNYADPITREGWAIPAATDIAFALGVLALLGSR----VPLALKIFLMALAIIDDLGAII 169
             :     :    .  .:** .::  ::.

sod2        FFYFAIKLLTVKPSRNAGRDWVLLV------VLYECAFGIFFGCVIGYLLSFILKHAQ- 235
SOS1        VFQLFLKM-----AMGQNSDWSSIIKFLLKVALGAVGIGLAFGIASVIWLKFIFNDTVI 260
NhaA        IIALFYTN-DLSMASLGVAAVAIAV------LAVLNLCGARRTGVYILVGVVLWTAVL- 220

sod2        KYRLIDAISYYSLPLAIPLLCSGIGTIIGVDDLLM--SFFAGILFNWNDLFSKNIS-AC 291
SOS1        EITLTIAVSYFAYYTAQE--WAGASGVLTVMTLGMFYAAFARTAFKGDS--QKSLHHFW 315
NhaA        KS----------GVHATLAGVIVGFFIPLKE------KHGRS---------PAKR-LE 252

sod2        SVPAFIDQTFSLLFFTYYGTIIPWNNFNW--SVEGLPVWRLIVFSILTLVCRRLPVVFS 348
SOS1        EMVAYIANT---LIFILSGVVIAEGILDSDKIAYQGNSWRFLFLLYVYIQLSRVVVVGV 371
NhaA        HVLHPWVAYLILPLFAFANAGVSLQGVTL--DGLTSILPLGIIAGLLIGKPLGISLFCW 309

sod2        VKPLV-----PDIKTWKEALFVGHFGPIGVCAVYMAFLAKLLLSPDEIEKSIYESTTVF 402
SOS1        LYPLLCRFGYGLDWKESIILVWSGLRGAVALALSLSVKQSSGNSHISK---ETGTLFLF 427
NhaA        LALRLKLAHLPEGTTYQQIMVVGILCGIGFTMSIFIASLAFGSV----------DPEL 357

sod2        STLNEIIWPIISFVILSSIIVHGFSIH 429
SOS1        FT-------GGIVFLTLIVNGSTTQFV 447
NhaA        INWAKLGILVGSISSAVIGYSWLRVRL 388
```

Figure 4.4: Alignment of sod2 with NhaA using SOS1 as an intermediate as described in the *Experimental procedures*. Sod2 TM IV and aligned residues in NhaA are colored (red, hydrophobic; blue, acidic; green, polar; magenta, basic) and labeled with Gonnet PAM 250 amino acid conservation scores between sod2 and NhaA TM IV (* identical; : strongly similar; . weakly similar). The transmembrane segments, from the crystal structure or the homology model of NhaA and sod2, respectively, are shaded in grey.

Mutation	Liquid LiCl	Liquid NaCl	Solid LiCl	Solid NaCl	Group	Str NMR	Str Model
sod2 wt	+++	+++	+++	+++			
F126A	+++	+++	+++	+++	I	D	L
P127A	+++	+++	+++	+++	I	D	L
Q128A	+++	++	+++	+++	I	D	L
I129A	+++	+++	++	+++	I	H	L
N130A	+++	+++	++	+++	II	H	L
F131A	+++	+++	+++	+++	I	H	H
L132A	++	++	++	+++	II	H	H
G133A	++	++	++	++	II	H	H
S134A	++	++	++	++	II	H	H
L135A	++	++	++	++	II	H	H
L136A	+++	+++	++	++	II	H	H
I137A	+++	+++	+++	+++	I	H	H
G139A	+++	+++	++	++	II	H	H
C140A	+++	+++	+++	+++	I	H	H
I141A	+++	+++	+++	+++	I	H	H
T142A	+++	+++	+++	+++	I	H	H
S143A	+++	+++	+++	+++	I	E	E
T144A	-	-	+	+	III	E	E
T144D	-	-	+	+	III	E	E
T144K	-	-	+	+	III	E	E
T144S	+++	-	+++	+	IV	E	E
D145	N/A	N/A	N/A	N/A	N/A	E	E
P146	N/A	N/A	N/A	N/A	N/A	E	H
V147A	-	-	+	+	III	E	H
V147L	-	-	-	-	III	E	H
L148A	+++	+++	+++	+++	I	H	H
S149A	++	+++	+++	+++	II	H	H
L151A	+++	+++	+++	+++	I	H	H
L151S	++	+++	+++	+++	II	H	H
I152A	+	++	++	++	V	H	H

Table 4.1: Summary of sod2 TM IV structure and function data. All of the data is taken from (39) except for the functional data (columns 2-5) for D145 and P146 which were derived from other collected data, (3) and (4), respectively. The mutational groups (column 6) are defined by their effect on the ability of sod2 to grow in the presence of salt. They are as follows: I, no effect; II, mild effect; III, marked effect; IV, ion selective; V, intermediate effect. Columns 7 and 8 define the secondary structure determined by either NMR or homology modelling where: H, α-helical; L, loop; D, disordered; E, extended.

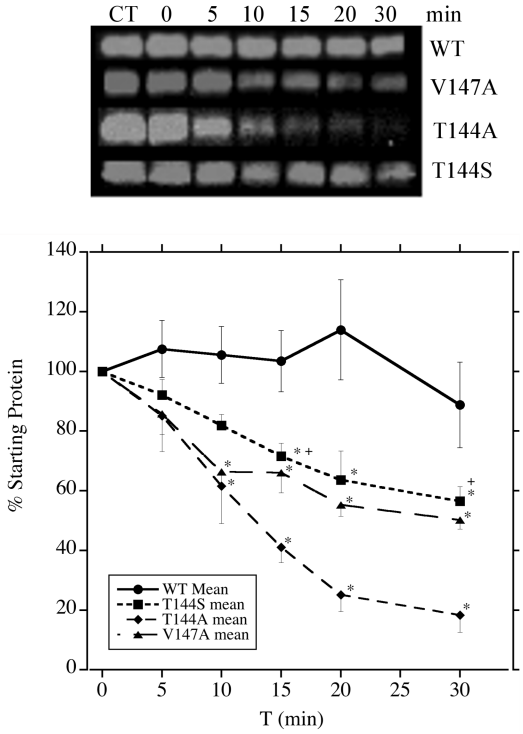

Figure 4.5: Analysis of susceptibility of mutant and wild type sod2 protein to trypsinolysis. Yeast cell membrane fractions were incubated with a 1:200 trypsin:protein ratio for 0–30 min at 30 °C as described in the *Experimental procedures*. Samples were then analyzed by SDS-PAGE and Western blotting using anti-GFP antibody. ***A***, Western blot of wild type (WT) and sod2 mutant proteins. ***B***, Summary of the effect of trypsin on percentage of sod2 protein remaining over time. Results are mean ± SE of at least 4 experiments. * indicates significantly different from WT at $P < 0.05$. + indicates significantly different from T144A at $P < 0.05$.

Trypsinolysis

To determine whether the mutant sod2 protein was properly folded we carried out limited digestion with trypsin. This method examines the proteolytic attack on accessible Arg and Lys residues, and has been used earlier to examine changes in the structure of sod2 (4) and of the mammalian Na^+/H^+ exchanger (40). We compared the digestion of wild type sod2 with that of the V147A, T144A and T144S mutants (Figure 4.5). We found that the T144A mutant sod2 protein was digested significantly more rapidly than the wild type with the T144S and the V147A showing intermediate patterns of digestion.

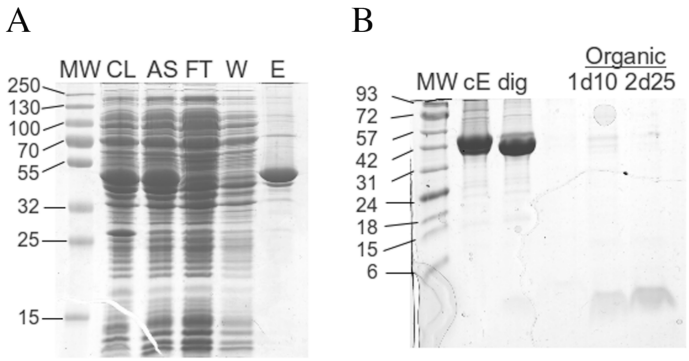

Figure 4.6: SDS-PAGE (12%) analysis of purification of MBP-sod2 fusion protein. ***A***, sod2 TM IV fusion protein was produced and purified as described in the *Experimental procedures* and samples from the purification procedure are illustrated. CL, crude bacterial lysate; AS, clarified lysate supernatant after ultracentrifugation; FT, flow through, maltose affinity column unbound fraction; W, maltose affinity column wash fraction; E, maltose affinity column eluted fraction (~2 mg protein/ml). MW, molecular weight ladder (kDa). ***B***, Tris-Tricine (16%) analysis of TEV cleavage of sod2 fusion protein followed by organic extraction. cE, 10x concentrated elution fraction; dig, concentrated elution fraction following 3 day cleavage with TEV; Organic, samples of the organic layer following extraction; 1d10, 10 μL of organic layer following one round of extraction; 2d25, 25 μL sample of organic layer following two rounds of extraction; MW, molecular weight ladder (kDa).

TM IV peptide purification

To gain molecular insights into how these functionally important residues may be arranged in the structure of sod2 we expressed and purified a peptide representing TM IV for NMR experiments (see Figure 4.1-B for a 2D representation).

$$KKKGS^{125}LFPQINFLGSLLIAGCITSTDPVLSALIVG^{154}KKK$$

Following purification of a maltose binding protein fusion protein and cleavage with TEV, an organic liquid extraction technique was used to isolate the hydrophobic TM IV peptide from any remaining aqueous contaminants (Figure 4.6). After two rounds of extraction the organic layer contained relatively pure peptide with only trace amounts of contaminants remaining and we proceeded directly with NMR experiments using the sample in organic solvent.

Nuclear magnetic resonance spectroscopy

Initially, 1D ^{1}H and then 2D ^{15}N HSQC spectra were used to judge the quality of peptide samples for structure determination while varying the solvent/detergent, pH, and temperature. The

peptide was poorly soluble in a $CHCl_3$/MeOH/H_2O mixture, which has been used successfully in previous work for a transmembrane segment of the mammalian Na^+/H^+ exchanger (16). It was also poorly soluble in trifluoroethanol/H_2O and although the peptide could be solubilized in dodecylphosphocholine, SDS, or DMSO, these samples only gave moderate quality 2D ^{15}N HSQC spectra (Figure 4.12). The peaks were broader and less well resolved than a sample in $CHCl_3$/IPA. The best spectra were obtained in the solvent mixture used for the organic extraction, $CHCl_3$/IPA, and this solvent mixture was used for further NMR analysis. The 1D spectrum showed good dispersion of the backbone amide peaks, however there was an additional broad intensity underneath the amide region which may have been peptide aggregates or impurities. Peaks were well resolved in a 2D ^{15}N HSQC spectrum suggesting that the solvent mixture would be suitable for further experiments (Figure 4.7). With the exception of proline residues and the N-terminal glycine, complete backbone resonance and partial side chain resonance assignments were obtained from the 3D NMR spectra. A labeled 2D HSQC spectrum is shown in Figure 4.7. A few additional peaks were present in the HSQC that could not be assigned that could result from impurities in the sample or minor conformations from cis-trans isomerization from one or both prolines in the TM IV peptide.

A comparison of the Hα chemical shifts with random coil chemical shifts (Figure 4.8) (41) suggests that there are two alpha-helical regions: residues 128–142 form a longer helix on the N-terminal half of the peptide and residues 147–153 form a shorter helix on the C-terminal half. The residues between the helices, 143-146, have shifts close to random coil values, and are likely dynamic or extended in structure. The termini are also likely unstructured, with Hα shifts close to random coil. The chemical shift data is supported by the $^3J_{HNH\alpha}$ coupling constants, as the helical regions have coupling constants around 4–6 Hz, consistent with helical structure, and with termini with coupling constants at around 7 Hz, indicating dynamic or unstructured residues. Critical residue T144 has a coupling constant of about 8 Hz, suggesting this region between the two helices is adopting an extended conformation. Values for some residues were not determined due to overlap of peaks in the HN region of the spectrum. Proton-proton distances from NOESY data also support the secondary structure prediction (Figure 4.8-C). Hα-HN (i, i+3) and (i, i+4) contacts can be observed within the two helical regions. Further evidence of the extended region between the two helices is provided by the presence of exchange NOE peaks at 4.8 ppm to the residual water in the sample, indicating interactions with water with this region.

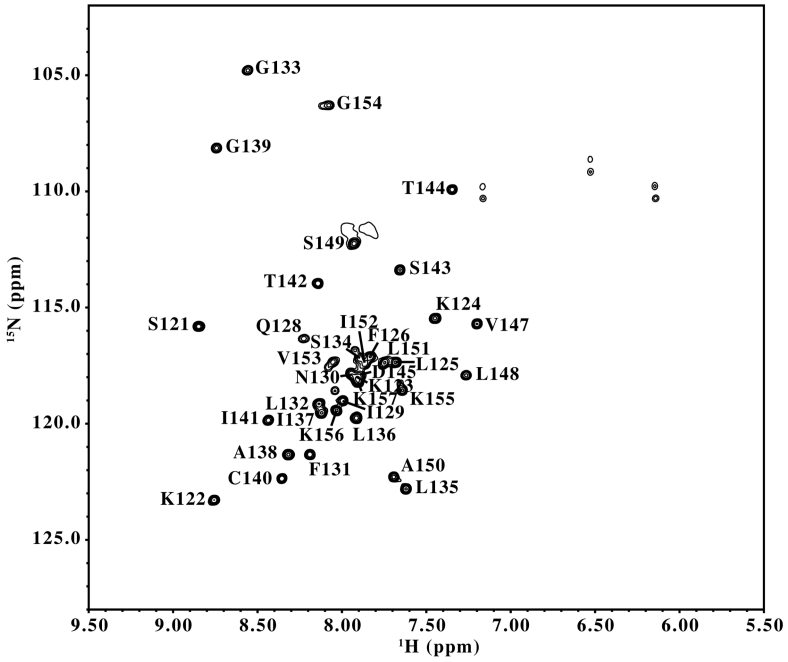

Figure 4.7: 2D ^{15}N HSQC spectra of sod2 TM IV peptide. 2D ^{15}N HSQC of sod2 TM IV in 50% CDCl3/50% isopropanol with assignments indicated. The sample was taken directly from the organic extraction procedure and partially concentrated by slow evaporation under $N_{2(g)}$.

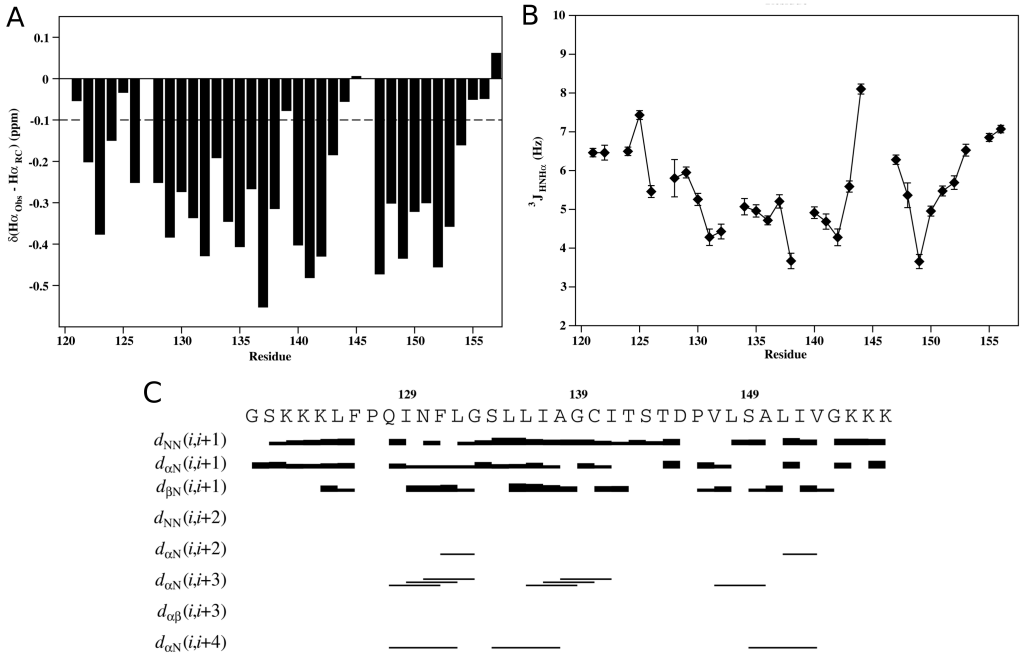

Figure 4.8: Secondary structure determination using Hα chemical shift, $^3J_{HNH\alpha}$ data, and Nuclear Overhauser Effect distances. ***A***, Chemical shift index prediction of sod2 TM IV secondary structure. Contiguous regions with deviations lower than -0.1 ppm (dotted line) from random coil chemical shifts indicate alpha-helical structure. Glycine chemical shifts were averaged if two Hα peaks are resolved. ***B***, $^3J_{HNH\alpha}$ coupling constants calculated for the peptide from a 3D HNHA NMR spectrum. Regions approximately <5 Hz suggest alpha helical structure, and >8 Hz extended structure. Missing values are glycines or were not calculated due to peak overlap. ***C***, Summary of Nuclear Overhauser Effect distances observed in a 3D ^{15}N NOESY-HSQC spectrum of sod2 TM IV. Regions containing (i, i+3) and (i, i+4) contacts suggest helical structure.

A model (Figure 4.9) representing the likely structure that sod2 TM IV could adopt was constructed with Xplor-NIH (27) using the $^{3}J_{HNH\alpha}$ and the Hα chemical shift data (Figure 4.8). The model structure calculation begins with a simple extended polypeptide upon which the measured NMR restraints are added and energy minimization calculations are performed. This process was followed several times analyzing violations resulting from conflicts between the data and the model. Residues that had large errors in the calculations, or where the $^{3}J_{HNH}\alpha$ and chemical shift data conflicted or were ambiguous, were not included in the modelling. The model displays disordered termini, a longer helical stretch of residues 129–142, an extended region from residues 143–147 and a shorter helix from residues 148–153. The NMR data and model have been deposited in the Protein Data Bank and Biological Magnetic Resonance Data Bank (PDB entry: 2M7X).

Homology modelling

To gain insight into the potential role of TM IV in the folded structure of sod2 and to provide evidence that the deduced structure of TM IV was correct, we constructed a homology based model of the entire sod2 protein. The model was based on the crystal structure of *E. coli* NhaA, the only Na^{+}/H^{+} exchanger with a high resolution crystal structure available. As noted above, sod2 did not align directly with NhaA, using the Clustal Omega alignment program (29). To overcome this problem we used SOS1 as an alignment template as it aligned with both NhaA and sod2 (Figure 4.4). These alignments were then used to construct a sequence comparison between NhaA and sod2. The initial model largely resembled the architecture of NhaA except for one large structural anomaly wherein amino acids 237–255 formed a long loop that traversed parallel through the membrane domain looping around two transmembrane segments. Closer analysis of this feature indicated that several of the residues were in steric conflict and that this structure was invalid. To resolve this we sought to improve our sequence alignment using the TMHMM2 secondary structure prediction algorithm (31) to predict the location of α-helical transmembrane segments in sod2. Significantly, the predicted helical regions aligned approximately with the actual locations of transmembrane segments in NhaA, giving an unbiased indication that our initial alignment was reasonable. The predictions were further corroborated using the JPred3 secondary structure prediction algorithm (32) which indicated very similar helical regions to TMHMM2. Using these secondary structure predictions, the alignment was adjusted so that predicted he-

lices of sod2 lined up with the known helices of NhaA. This alignment replaced the aberrant intramembrane loop at residues 237–255 with a transmembrane alpha helix that satisfied the steric conflicts. However closer inspection revealed that up to seven amino acids at the ends of several transmembrane helices dipped back into the membrane. To address this we looked at the predicted length of transmembrane helices in sod2 versus NhaA and noted that many were longer in sod2 than NhaA. Since Modeller does not account for significant differences in transmembrane segment length, short helical restraints were added to the ends of the sod2 transmembrane segments predicted to be longer than the aligned sequence in NhaA. To ensure no undue bias was added by secondary structure prediction the same TMHMM2 and JPred3 algorithms were used to predict helical regions of NhaA. The algorithm predicted helices that aligned nearly perfectly (less than 2 residues different) with the actual location of transmembrane helices in the crystal structure (with the exception of the two short helices, 7 and 8, which were predicted to be one long helix). The resulting model (Figure 4.10-A) no longer appeared to have any structural anomalies and was in good agreement with the NhaA structure (Figure 4.10-B). The validity of the model was confirmed using the protein structure validation software suite (33). This evaluation revealed only eleven residues with backbone angles outside of the allowable region of the Ramachandran plot (Figure 4.13). Ten of these residues are located in extramembrane loops where the backbone bonds would be flexible enough to adopt a more preferable angle. The seventh residue, V348, is located in a slightly unwound region of transmembrane segment X. As this residue is near the end of the helix, and to resolve the significance of this disallowed dihedral angle, sculpting was performed using PyMol (Schrodinger, LLC). Upon relaxation this region became more helical adopting an admissible dihedral angle.

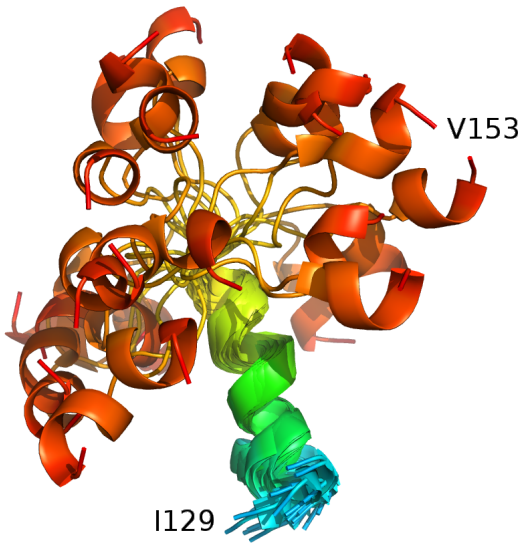

Figure 4.9: Ensemble of 25 models of sod2 TM IV created with Xplor-NIH. Side view of the ensemble with the structures superimposed along the N-terminal helix (PDB entry: 2M7X). Only residues I129-V153 are shown for clarity as the amino and carboxyl ends of the structures were disordered.

The model was further validated from a bioinformatics standpoint using the ConSurf method (34, 35). Using the ConSurf server (36) 91 protein sequences (of >35% identity and a BLAST score of >300) were selected and aligned with the sod2 sequence. The server then calculated a Bayesian conservation score (0–9), gave it a corresponding color and mapped it onto the model of sod2 (Figure 4.11). Regions of high sequence conservation (pink to purple) are, as expected, found in the core of the transporter while regions of less conservation (white to blue) are found on the outer surface. Significantly the critical amino acids ^{144}TDP^{146} all have the highest conservation score of 9 and V147 has a score of 8.

Comparison of the deduced and modelled TM segments of sod2, TM IV and TM IV of NhaA and TM VI of NHE1 (Figure 4.10-C) shows that all display a characteristic helix–extended region–helix conformation. The NMR structure of TM IV of sod2 is very similar, though not identical, to the structure of sod2 deduced by modelling. Analysis of the known biochemical data with respect to the model (Figure 4.10-D) illustrates the homology model of sod2 viewed perpendicular to the membrane plane from the extracellular side. The key residues of TM IV (T144-V147) are shown as orange spheres and have a central, potentially pore lining location. Figure 4.10-E illustrates the position of all known functional mutations in the model.

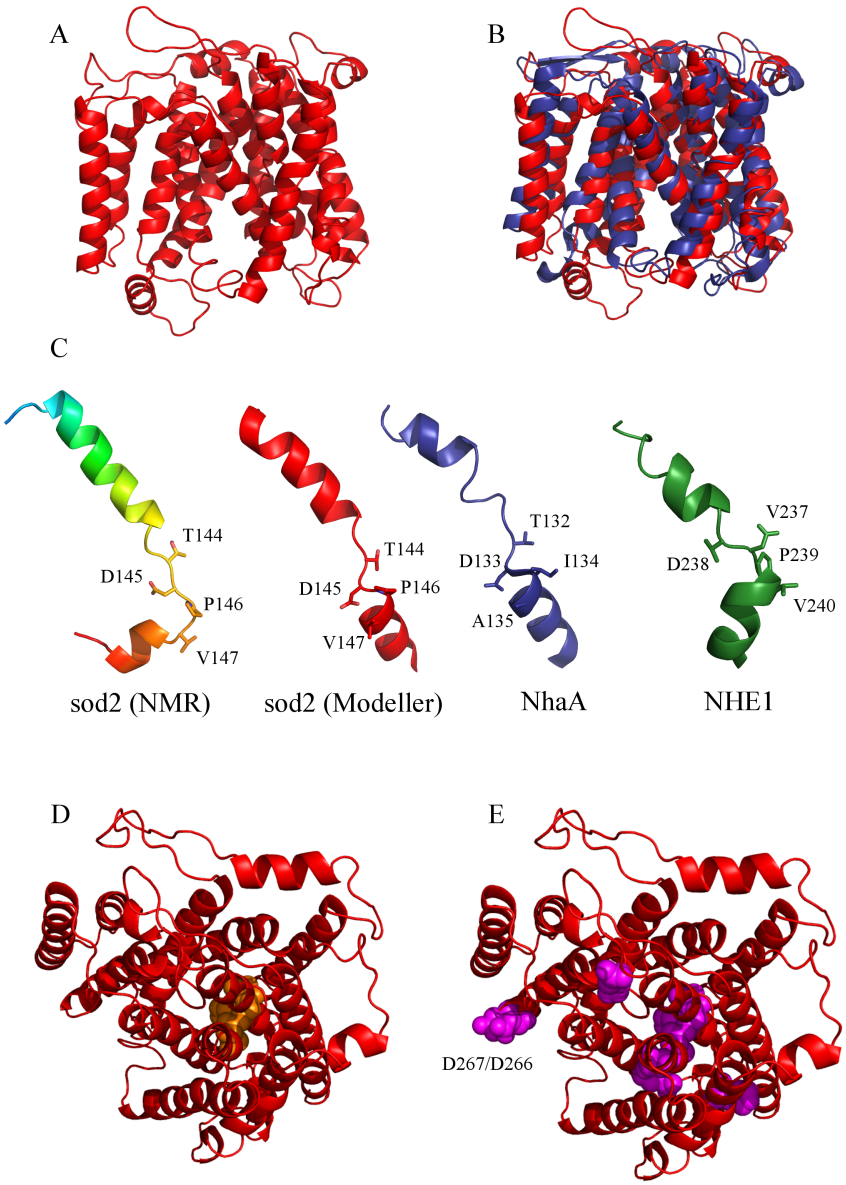

Figure 4.10: Homology model and TM segment structures compared. ***A***, Homology model of sod2 membrane domain (residues 12–429) based on the crystal structure of *E. coli* NhaA (PDB entry:1ZCD) using Modeller software (28). ***B***, Homology model (red) aligned with NhaA (blue). ***C***, Structures of analogous TM segments of NhaA TM IV (blue) and NHE1 TM VI (green) are shown next to the structural model of sod2 (NMR) made with Xplor-NIH (rainbow) and sod2 (Modeller) homology model TM IV(red). Key residues have been labelled for each peptide. ***D***, Homology model of sod2 viewed perpendicular to the membrane plane from the extracellular side. The key residues of TM IV (T144-V147) are shown as orange spheres. ***E***, Reproduction of panel ***D*** with all known functional mutations indicated in magenta spheres. D266 and D267 are labelled, indicating their distance from the transport pore.

Discussion

Removal of excess salt in plants and yeast is crucial in dealing with excess *toxic* levels of intracellular Na^+. In plants, improvements in salt tolerance have significant agricultural implications for crops. The yeast *Schizosaccharomyces pombe* is a useful model organism for the study of salt tolerance as the Na^+/H^+ antiporter (sod2) plays the major role in salt removal from the cytosol and its deletion results in a Na^+ and Li^+ sensitive phenotype (1). While significant advances have been made towards understanding the molecular mechanism of mammalian (42) and *E. coli* (43) Na^+/H^+ exchangers, there has been much less study on plant and yeast Na^+/H^+ exchangers. Here we present the first systematic examination of a transmembrane segment of this type of transporter by alanine scanning mutagenesis and NMR spectroscopy. Alanine scanning mutagenesis has been used earlier to characterize transmembrane segments of membrane proteins including the NHE1 isoform of the mammalian Na^+/H^+ exchanger (9, 44). The premise is that the small side chain of alanine can substitute for amino acids without disrupting the protein structure, while at the same time altering the nature of the side chain. All the amino acids of TM IV were mutated to alanine, with the exception of A138 and A150 which were already alanine. Also, in earlier studies we had already characterized D145 and P146 (4, 7).

Most of the alanine mutants had either no effect, or mild to intermediate effects on the ability of sod2 to restore salt tolerance in sod2::ura4 *S. pombe* (Table 4.1, Group I and II mutants). However, effects of mutation of amino acids T144 and V147 were notable and were explored in more detail. For T144, mutation to either Ala, Asp or Lys did not restore the ability of sod2 to confer salt tolerance. However, a serine substitution conferred tolerance to LiCl but not to NaCl, suggesting that both the size and the chemistry of this residue are important for activity. Amino acids 124–156 of the *Zygosaccharomyces rouxii* Na^+/H^+ exchanger, a related member of the yeast plasma membrane Na^+/H^+ antiporters, exhibit striking similarity to amino acids 125–158 of *S. pombe* sod2. In particular, T141 and T143 of *Zr*sod2-22p align with T142 and T144 of sod2 (Figure 4.1-C). Interestingly it has been demonstrated that mutation of T141 and S150 of *Zr*sod2-22p to serine and threonine, respectively, altered and broadened the cation selectivity of this antiporter and indicated the importance of the hydroxyl sidechain (45). Similarly, the nearby residue P145 was also shown to contribute to substrate specificity. While T142 of sod2, which aligns with T141 of *Zr*sod2-22p, was not critical for transport by sod2, we found that the

downstream T144 of sod2 appears to fill the role of *Zr*sod2-22p T141.

A simple explanation for the change in cation selectivity upon mutation of Ser to Thr may be that presentation of the hydroxyl by Ser is such that it can coordinate the smaller Li^+ ion (0.95 Å) but not the larger Na^+ (0.65 Å) (40). Alternatively, hydrogen bond formation between the side chain oxygen atom of a serine and threonine and the i-3 or i-4 peptide carbonyl oxygen may induce or stabilize a 3–4 degree bend in a helix relative to a helix with Ala in this position (46). Therefore T144A may have a different TM segment conformation that could affect cation coordination. Importantly, T132 of the *E. coli* Na^+/H^+ antiporter NhaA has been shown to participate in cation coordination (47) and other Ser or Thr residues have also been shown to be important in cation binding and transport in other membrane transporters (48–50). However, experiments on the mammalian NHE1 isoform of the Na^+/H^+ exchanger have shown that not all serine and threonine residues are important in this regard (9, 14).

Conversely V147A, a moderately conservative mutation, also produced a non-functional protein but it could not be rescued by substitution with Leu. This suggests that size or shape of Val is structurally important to sod2 activity.

While T144 and V147 mutants resulted in the most notable difference in activity, mutation of residues L132, G133, S134, L135, L136 and G139 (Group II, Table 4.1) exhibited phenotypes with slightly depressed growth in both NaCl and LiCl containing media. Although the effect is small, this region is conserved in both human NHE1 and *Zr*sod2-22p (45) (Figure 4.1-D) and it may be possible that changes in amino acid packing or helix dynamics lead to less efficient transport and therefore reduced resistance. All these amino acids were in a helical region of the protein (Figure 4.9).

Our experiments have demonstrated a severe functional defect upon mutation of T144 and V147, possibly affecting cation coordination directly (as suggested by T144S) or indirectly (by changing the conformational stability of the protein). These conclusions are supported by the fact that neither the expression level nor targeting of the mutant forms of sod2 were affected (39), indicating that global protein misfolding is not the culprit. However, limited tryptic digestion suggested that there was a local decrease in protein conformational stability upon mutation. This was especially noticeable in the T144A protein which was much more quickly digested than the wild type. However this effect was greatly reduced in the conservative substitution of T144S, implicating the hydroxyl side chain's role in protein stability. Sensitivity to trypsin was also

increased in V147A but only slightly more than T144S, suggesting a more subtle structural role.

In order expand to our understanding of the trypsinolysis results, we expressed and purified sod2 TM IV and analyzed the structure of the protein using NMR. Earlier studies have demonstrated that the amino acid sequence of TM segments of membrane proteins contain most of the required structural information needed to form their native structures (51, 52). For example, Katragadda *et al.* demonstrated that the structure of individual transmembrane helices of rhodopsin and bacteriorhodopsin corresponded very well to the structures obtained by x-ray crystallography (52, 53). We have also earlier used this approach and successfully produced transmembrane peptides of the mammalian Na^+/H^+ exchanger in *E. coli* or synthetically (14, 16). We therefore produced TM IV of sod2 as a MBP fusion protein. After cleavage free of MBP, we used an organic extraction procedure to purify the TM IV fragment which resulted in a sample of sufficient purity for structure determination by NMR. Hu *et al.* used a similar protocol and demonstrated that sections of the membrane domain of CorA, M2 and KdpF are properly folded (54).

NMR experiments revealed significant secondary structure information. Measuring Hα chemical shift deviations from random coil, we noted that there were contiguous regions of negative deviation which are indicative of helical secondary structure (55). Likewise we calculated $^3J_{HNH\alpha}$ coupling constants from the ratio of the intensities of the diagonal and crosspeaks for each residue in the HNHA spectrum (23, 24). These values are related to the ϕ dihedral angle. Although the amino terminal portion of the TM IV peptide showed largely alpha helical character Hα chemical shift deviations 133 and 139 are closer to random coil values, which could suggest that this N-terminal helix is broken, distorted or dynamic at these positions. Importantly the coupling constants show the same trends as the chemical shift deviations, corroborating our secondary structure observations. Although some supporting NOE data was collected, the relatively low number of helical distances observed could be a further indication of dynamic behaviour or distortion of the helices. This could simply be a result of the peptide being in organic solvent rather than in a membrane, or perhaps it indicates the importance of other transmembrane segments that might stabilize partially unwound helices in the tertiary structure. Additionally, although data for the proline at position 146 could not be observed, it is predicted to be a helix breaker (56). This is not unprecedented as the structure of hNHE1 TM IV shows two proline residues that disrupt the helix and form an extended section (19). Most interestingly, all the residues in this segment that have been found to be critical to sod2 function (T144 and V147 in

this work, and D145 and P146 shown previously) are found in the extended region or very close to it (Table 4.1). This is consistent with the results of other structural studies that have shown that discontinuous membrane helices are often involved in substrate coordination, an example being *E. coli* NhaA (47). The minor functional effects observed in the flanking helical regions may be connected to the structural dynamics suggested by the NMR data.

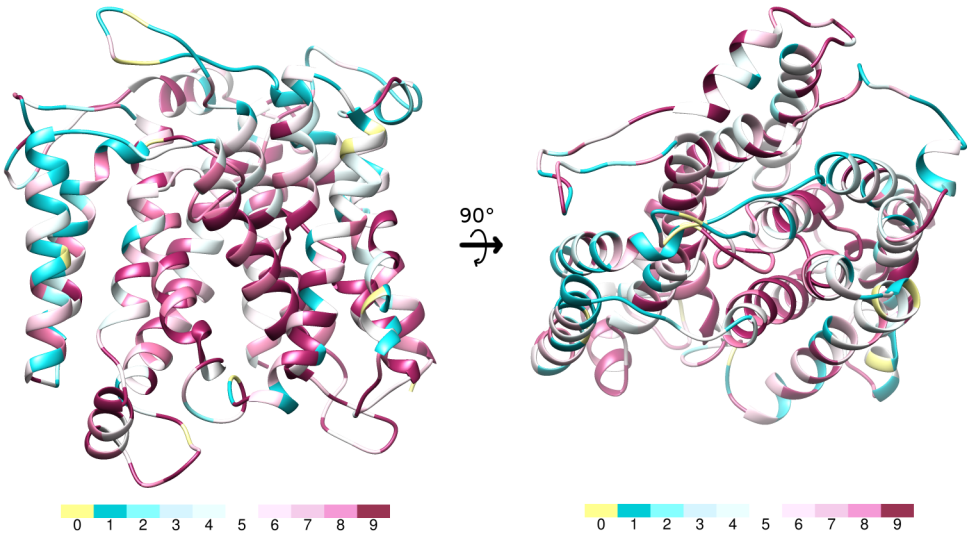

Figure 4.11: ConSurf conservation mapping on sod2 homology model. The ConSurf server (36) was used to create a multiple sequence alignment and calculate the overall conservation score for each residue in sod2. This conservation score (0–9) was mapped onto the homology model of sod2 using a colour scale. Internal regions of the protein important for transport show generally higher levels of conservation and external regions of the protein are less well conserved. The key residues in TM IV have either the highest conservation score, 9 (T144, D145, P146) or the second highest, 8 (V147).

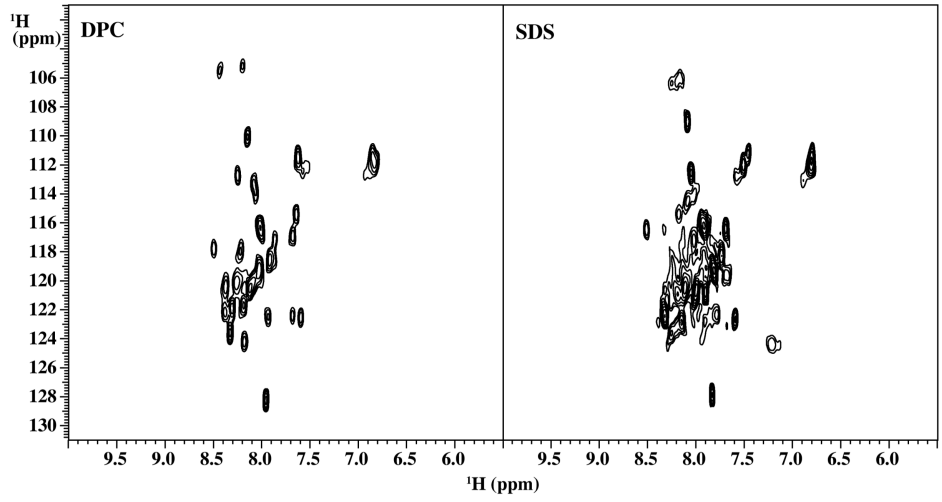

Figure 4.12: 2D ^{15}N HSQC spectra of sod2 TM IV peptide in detergent micelles. Samples in detergent micelles were prepared from peptide in organic solvent that was dried under argon gas and reconstituted in 1% DPC or SDS, 10 mM imidazole, 0.25 mM DSS, 95% H_2O, 5% D_2O, pH 6.0. Spectra were acquired at 500 MHz and 30 °C.

Although the data was not sufficient to calculate a precise structure, a model of sod2 TM IV was created using Xplor-NIH (27) to better visualize the secondary structure and to provide a basis for comparison to other published structures. We suggest that TM IV of sod2 may be functionally more similar to TM VI of NHE1 than to TM IV of NHE1. This is based on analysis of amino acid alignments, data examining critical amino acids in this region and overall structure of the region. The overall structure of human NHE1 TM VI is similar to that of TM IV of sod2 (Figure 4.10-C). Both contain two helices with an intervening unwound region that is critical for function (14). TM IV of human NHE1 has an unwound region, but this is flanked by only one helix (16).

Assessment of functional properties of aligned residues also supports the notion that TM IV of sod2 is similar to TM VI of NHE1. Mutation of residues D238 and P239 of human NHE1 results in a non-functional protein. These residues correspond to D145 and P146 in sod2 and we showed earlier that when D145 and P146 of sod2 are mutated that the protein is inactive (4, 7). Similarly, the mutation of V237 of NHE1 to Cys had reduced function in the presence of cysteine-modifying compounds, and this corresponds to T144 of sod2 which is critical to function.

We have earlier also noted that TM VI of NHE1 has more similarities to TM IV of NhaA than TM IV of NHE1 (14). Comparison of sod2 TM IV with *E. coli* NhaA TM IV also suggests that although they have different primary structures their secondary structures are similar, both with helix-unwound region-helix conformation (Figure 4.10-C). Interestingly, TM IV of *E. coli* NhaA also contains important residues in this segment as mutation of either T132 or D133 has large effects on the apparent Km indicating a role in the integrity of the transport mechanism (6, 57). The extended peptide region may help accommodate the charged ion substrate (47). Studies on other membrane proteins have confirmed that unwound helices participate in substrate coordination and transport of many membrane proteins aside from Na^+/H^+ antiporters (47, 58).

To assess the effect these mutations might have on the folded structure, and to provide evidence that the deduced structure of sod2 TM IV is valid, we constructed a homology model of the membrane domain of sod2. Mounting evidence suggests that a finite number of folds exist for membrane transporters (59) and that proteins with quite different primary amino acid sequence can have surprisingly similar structures. For example, the bile acid sodium symporter of *Neisseria meningitides* has a low primary sequence identity to *E. coli* NhaA but the structure

is surprisingly similar to that of NhaA (54).

With this evidence for justification, we decided to use the functionally related NhaA structure as a template for homology modelling of sod2. After several iterative rounds of modelling, described above, we obtained a model that appears to generally satisfy our NMR data, and the collected biochemical data (Figure 4.10). The model is largely free of steric conflicts as demonstrated by 98% of residues falling within the acceptable regions of the Ramachandran plot and remaining residues in flexible regions of the structure (Figure 4.13). Additionally, analysis using the ConSurf method (34, 35) indicates that functionally important regions in the interior of the transporter show the highest levels of conservation (Figure 4.11) including the unwound section of TM IV. This method was also used to validate a model structure of NHE1 (36, 60).

The overall architecture of the model of sod2 indicates the same fold as NhaA (6) with helices IV and XI having an unwound, crossed configuration in the membrane. Functionally critical residues H367, Asp241, (3), P146 (4), D145, and E173 (8) are all situated in or near the putative transport pore (Figure 4.10-E). The exceptions are residues D266 and D267. These residues were found to be critically important for transport activity (3) but are not located near the transport site. They are located on a short alpha helix within a long extracellular loop. A possible explanation is that this long loop may form a *cap-like* structure on the cytoplasmic side of the pore opening. The negative charges could help attract sodium ions for transport. Since this is not a feature found in NhaA it would be unlikely to be predicted by molecular modelling.

Significantly, the critical residues in the model of TM IV T144-P146 are in an extended conformation and face the hypothetical transport pore, in agreement with our deduced NMR structure. This offers further evidence that these residues are involved in ion translocation and supports the NMR structure we obtained. V147, however, is in a helical conformation and faces towards the outer helices in the homology model. From the homology model, one possible explanation of the functional defect of V147A is that mutation at this site affects helical packing, likely destabilizing the structure, leading to a functional defect. This is supported by our confocal data (see Ullah *et al.* (39)) that suggests that V147A is targeted the same as the wild type protein and by our trypsinolysis data (Figure 4.5) that shows that V147A is more susceptible to digestion suggesting increased flexibility. Alternatively, Val has a lower propensity to form helices than either Ala or Leu; therefore these mutations may lead to increased helical character perturbing the placement of TM IV in the membrane or decreasing the segments flexibility preventing proper

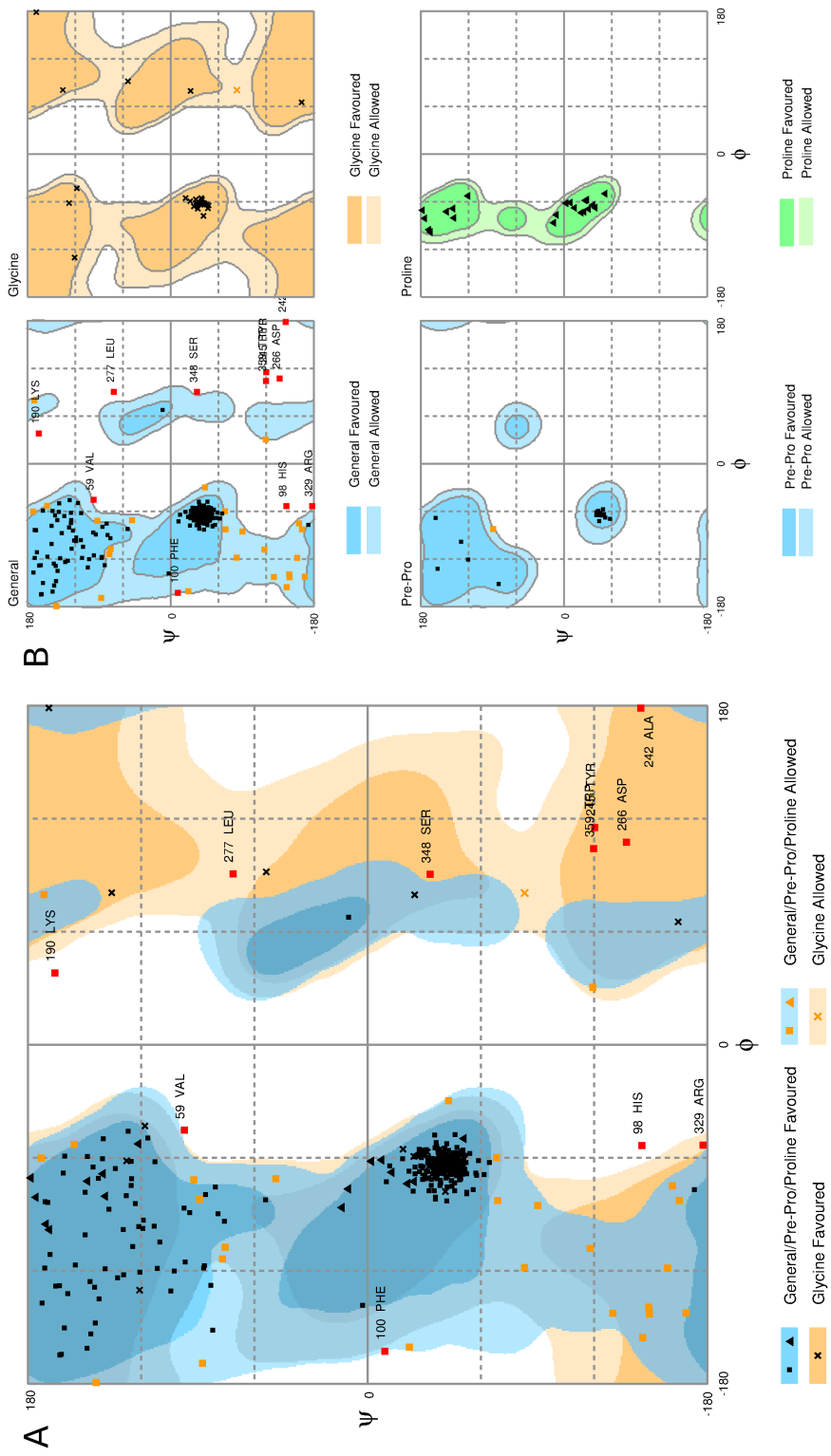

Figure 4.13: Ramachadran plot for sod2 homology model produced using RAMPAGE (61). ***A***, Ramachandran plot for every residue of sod2. Darker areas (and black symbols) indicate *favourable* bond angles while lighter shades (and brown symbols) indicate *allowed* angles. Residues falling outside of the *allowed* regions are shown as red symbols and labelled. ***B***, The plots of ***A*** broken down into specific amino acid classes for increased clarity.

function (62). This is still consistent with the trypsin digestion data as increased helical content may increase overall protein flexibility while decreasing the local flexibility of TM IV. In contrast V147 in the NMR structure is shown to be in an extended conformation. In fact the $^{3}J_{HNH\alpha}$ coupling constants and chemical shift data for V147 are just past the helical cut-off indicating weak helical character. This suggests that V147 may be more conformationally dynamic than the Ala and Leu mutants and that this flexibility is required for proper function. The fact that V147 is extended in the NMR structure and helical in the homology model (Table 4.1) further suggests that this position is likely locally flexible and requires the full membrane domain in order to form proper folded helical contacts.

In summary, the deduced secondary structure of the peptide based on NMR data consists of a slightly kinked N-terminal helix at residues 128–142, followed by a very flexible extended segment at residues 143–146, and a short C-terminal helix at residues 147–154. This secondary structure is supported by a homology model of TM IV of sod2, derived from *E. coli* NhaA. The functional data suggest that residues 144–146 appear to have analogous partners in *E. coli* NhaA, human NHE1 and *Zr*sod2-22p, indicating the importance of this region in transport. The salt tolerance protein sod2 has a partially unwound helix containing functionally important residues for ion transport, consistent with results obtained with human NHE1 and *E. coli* NhaA. This study extends our current molecular understanding of salt tolerance and salt tolerance proteins. Previously we demonstrated that mutation of D145 (7) and P146 (4) of sod2 resulted in a transport defect that was unable to confer salt tolerance. Our present results demonstrate that the flanking amino acids T144 and V147 are also critical for sod2 structure and function and that TM IV is likely to be directly involved in ion translocation. We have also demonstrated that the hydroxyl of T144 is required for activity and that the conservative Thr to Ser mutation results in the inability to transport Na^{+} while maintaining Li^{+} transport. We suggest that TM IV of sod2 serves a role similar to that of TM IV of *E. coli* NhaA and TM VI of human NHE1.

References

1. Jia, Z. P., McCullough, N., Martel, R., Hemmingsen, S. & Young, P. G. (1992). Gene amplification at a locus encoding a putative Na^+/H^+ antiporter confers sodium and lithium tolerance in fission yeast. *EMBO J* **11**, 1631–1640.

2. Haworth, R. S., Lemire, B. D., Crandall, D., Cragoe, J., E J & Fliegel, L. (1991). Characterisation of proton fluxes across the cytoplasmic membrane of the yeast *Saccharomyces cerevisiae*. *Biochim Biophys Acta* **1098**, 79–89.

3. Dibrov, P., Young, P. G. & Fliegel, L. (1998). Functional analysis of amino acid residues essential for activity in the Na^+/H^+ exchanger of fission yeast. *Biochemistry* **37**, 8282–8288.

4. Ndayizeye, M., Touret, N. & Fliegel, L. (2009). Proline 146 is critical to the structure, function and targeting of sod2, the Na^+/H^+ exchanger of *Schizosaccharomyces pombe*. *Biochim Biophys Acta* **1788**, 983–992.

5. Wiebe, C. A., Dibattista, E. R. & Fliegel, L. (2001). Functional role of polar amino acid residues in Na^+/H^+ exchangers. *Biochem J* **357**, 1–10.

6. Hunte, C. *et al.* (2005). Structure of a Na^+/H^+ antiporter and insights into mechanism of action and regulation by pH. *Nature* **435**, 1197–1202.

7. Wiebe, C. A., Rieder, C., Young, P. G., Dibrov, P. & Fliegel, L. (2003). Functional analysis of amino acids of the Na^+/H^+ exchanger that are important for proton translocation. *Mol Cell Biochem* **254**, 117–124.

8. Fliegel, L. (2005). Identification of conserved polar residues important for salt tolerance by the Na^+/H^+ exchanger of *Schizosaccharomyces pombe*. *Mol Cell Biochem* **268**, 83–92.

9. Ding, J., Rainey, J. K., Xu, C., Sykes, B. D. & Fliegel, L. (2006). Structural and functional characterization of transmembrane segment VII of the Na^+/H^+ exchanger isoform 1. *J Biol Chem* **281**, 29817–29829.

10. Slepkov, E. R., Rainey, J. K., Sykes, B. D. & Fliegel, L. (2007). Structural and functional analysis of the Na^+/H^+ exchanger. *Biochem J* **401**, 623–633.

11. Reddy, T. *et al.* (2008). Structural and functional characterization of transmembrane segment IX of the NHE1 isoform of the Na^+/H^+ exchanger. *J Biol Chem* **283**, 22018–22030.

12. Lee, B. L., Li, X., Liu, Y., Sykes, B. D. & Fliegel, L. (2009). Structural and functional analysis of transmembrane XI of the NHE1 isoform of the Na^+/H^+ exchanger. *J Biol Chem* **284**, 11546–11556.

13. Lee, B. L., Li, X., Liu, Y., Sykes, B. D. & Fliegel, L. (2009). Structural and functional analysis of extracellular loop 2 of the Na^+/H^+ exchanger. *Biochim Biophys Acta* **1788**, 2481–2488.

14. Tzeng, J., Lee, B. L., Sykes, B. D. & Fliegel, L. (2010). Structural and functional analysis of transmembrane segment VI of the NHE1 isoform of the Na^+/H^+ exchanger. *J Biol Chem* **285**, 36656–36665.

15. Lee, B. L., Sykes, B. D. & Fliegel, L. (2011). Structural analysis of the Na^+/H^+ exchanger isoform 1 (NHE1) using the divide and conquer approach. *Biochem Cell Biol* **89**, 189–199.

16. Slepkov, E. R. *et al.* (2005). Structural and functional characterization of transmembrane segment IV of the NHE1 isoform of the Na^+/H^+ exchanger. *J Biol Chem* **280**, 17863–17872.

17. Fliegel, L., Wiebe, C., Chua, G. & Young, P. G. (2005). Functional expression and cellular localization of the Na^+/H^+ exchanger sod2 of the fission yeast *Schizosaccharomyces pombe*. *Can J Physiol Pharmacol* **83**, 565–572.

18. Silva, N. L., Wang, H., Harris, C. V., Singh, D. & Fliegel, L. (1997). Characterization of the Na^+/H^+ exchanger in human choriocarcinoma (BeWo) cells. *Pflügers Arch* **433**, 792–802.

19. Slepkov, E. R., Chow, S., Lemieux, M. J. & Fliegel, L. (2004). Proline residues in transmembrane segment IV are critical for activity, expression and targeting of the Na^+/H^+ exchanger isoform 1. *Biochem J* **379**, 31–38.

20. Douglas, J. L., Trieber, C. A., Afara, M. & Young, H. S. (2005). Rapid, high-yield expression and purification of Ca^{2+}-ATPase regulatory proteins for high-resolution structural studies. *Protein Expression Purif* **40**, 118–125.

21. Kay, L., Keifer, P. & Saarinen, T. (1992). Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J Am Chem Soc* **114**, 10663–10665.

22. Zhang, O., Kay, L. E., Olivier, J. P. & Forman-Kay, J. D. (1994). Backbone ^1H and ^{15}N resonance assignments of the n-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. *J Biomol NMR* **4**, 845–858.

23. Vuister, G. W. & Bax, A. (1993). Quantitative J correlation: a new approach for measuring homonuclear three-bond $J_{HNH\alpha}$ coupling constants in ^{15}N-enriched proteins. *J Am Chem Soc* **115**, 7772–7777.

24. Grzesiek, S., Kuboniwa, H., Hinck, A. P. & Bax, A. (1995). Multiple-quantum line narrowing for measurement of hα-hβ J couplings in isotopically enriched proteins. *J Am Chem Soc* **117**, 5312–5315.

25. Delaglio, F. *et al.* (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277–293.

26. Johnson, B. A. & Blevins, R. A. (1994). NMR view: A computer program for the visualization and analysis of NMR data. *J Biomol NMR* **4**, 603–614.

27. Schwieters, C. D., Kuszewski, J. J., Tjandra, N. & Clore, G. M. (2003). The Xplor-NIH NMR molecular structure determination package. *J Magn Reson* **160**, 65–73.

28. Eswar, N. *et al.* (2006). Comparative protein structure modeling using modeller. *Current Protocols in Bioinformatics* **UNIT 5.6**.

29. Sievers, F. *et al.* (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**.

30. Goujon, M. *et al.* (2010). A new bioinformatics analysis tools framework at EMBL–EBI. *Nucleic Acids Res* **38**, W695–W699.

31. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *J Mol Biol* **305**, 567–580.

32. Cole, C., Barber, J. D. & Barton, G. J. (2008). The Jpred 3 secondary structure prediction server. *Nucleic Acids Res* **36**, W197–W201.

33. Bhattacharya, A., Tejero, R. & Montelione, G. T. (2007). Evaluating protein structures determined by structural genomics consortia. *Proteins: Struct., Funct., Bioinf.* **66**, 778–795.

34. Landau, M. *et al.* (2005). ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res* **33**, W299–W302.

35. Glaser, F. *et al.* (2003). ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. *Bioinformatics* **19**, 163–164.

36. Ashkenazy, H., Erez, E., Martz, E., Pupko, T. & Ben-Tal, N. (2010). ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res* **38**, W529–W533.

37. de Planque, M. R. R. & Killian, J. A. (2003). Protein-lipid interactions studied with designed transmembrane peptides: role of hydrophobic matching and interfacial anchoring. *Mol Membr Biol* **20**, 271–284.

38. Shi, H., Ishitani, M., Kim, C. & Zhu, J.-K. (2000). The arabidopsis thaliana salt tolerance gene SOS1 encodes a putative Na^+/H^+ antiporter. *Proc Natl Acad Sci U S A* **97**, 6896–6901.

39. Ullah, A. *et al.* (2013). Structural and functional analysis of transmembrane segment IV of the salt tolerance protein sod2. *J Biol Chem* **288**, 24609–24624.

40. Murtazina, R., Booth, B. J., Bullis, B. L., Singh, D. N. & Fliegel, L. (2001). Functional analysis of polar amino-acid residues in membrane associated regions of the NHE1 isoform of the mammalian Na^+/H^+ exchanger. *Eur J Biochem* **268**, 4674–4685.

41. Wishart, D. S. *et al.* (1995). ^1H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *J Biomol NMR* **6**, 135–140.

42. Kemp, G., Young, H. S. & Fliegel, L. (2008). Structure and function of the human Na^+/H^+ exchanger isoform 1. *Channels (Austin)* **2**, 329–336.

43. Padan, E. (2008). The enlightening encounter between structure and function in the NhaA Na^+/H^+ antiporter. *Trends Biochem Sci* **33**, 435–443.

44. Fleming, K. G. & Engelman, D. M. (2001). Specificity in transmembrane helix-helix interactions can define a hierarchy of stability for sequence variants. *Proc Natl Acad Sci U S A* **98**, 14340–14344.

45. Kinclova-Zimmermannova, O., Zavrel, M. & Sychrova, H. (2006). Importance of the seryl and threonyl residues of the fifth transmembrane domain to the substrate specificity of yeast plasma membrane Na^+/H^+ antiporters. *Mol Membr Biol* **23**, 349–361.

46. Ballesteros, J. A., Deupi, X., Olivella, M., Haaksma, E. E. & Pardo, L. (2000). Serine and threonine residues bend alpha-helices in the $\chi_1 = g^-$ conformation. *Biophys J* **79**, 2754–2760.

47. Padan, E., Kozachkov, L., Herz, K. & Rimon, A. (2009). NhaA crystal structure: functional-structural insights. *J Exp Biol* **212**, 1593–1603.

48. Díez-Sampedro, A., Wright, E. M. & Hirayama, B. A. (2001). Residue 457 controls sugar binding and transport in the Na^+/glucose cotransporter. *J Biol Chem* **276**, 49188–49194.

49. Kamiya, T. & Maeshima, M. (2004). Residues in internal repeats of the rice cation/H^+ exchanger are involved in the transport and selection of cations. *J Biol Chem* **279**, 812–819.

50. Li, H. & Pajor, A. M. (2003). Serines 260 and 288 are involved in sulfate transport by hNaSi-1. *J Biol Chem* **278**, 37204–37212.

51. Cunningham, F. & Deber, C. M. (2007). Optimizing synthesis and expression of transmembrane peptides and proteins. *Methods (San Diego, Calif.)* **41**, 370–380.

52. Katragadda, M., Alderfer, J. L. & Yeagle, P. L. (2001). Assembly of a polytopic membrane protein structure from the solution structures of overlapping peptide fragments of bacteriorhodopsin. *Biophys J* **81**, 1029–1036.

53. Katragadda, M. *et al.* (2001). Structures of the transmembrane helices of the g-protein coupled receptor, rhodopsin. *J Pept Res* **58**, 79–89.

54. Hu, N.-J., Iwata, S., Cameron, A. D. & Drew, D. (2011). Crystal structure of a bacterial homologue of the bile acid sodium symporter ASBT. *Nature* **478**, 408–411.

55. Wishart, D. S., Sykes, B. D. & Richards, F. M. (1991). Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J Mol Biol* **222**, 311–333.

56. Cordes, F. S., Bright, J. N. & Sansom, M. S. P. (2002). Proline-induced distortions of transmembrane helices. *J Mol Biol* **323**, 951–960.

57. Galili, L., Rothman, A., Kozachkov, L., Rimon, A. & Padan, E. (2002). Transmembrane domain IV is involved in ion transport activity and pH regulation of the NhaA Na^+/H^+ antiporter of *Escherichia coli*. *Biochemistry* **41**, 609–617.

58. Screpanti, E. & Hunte, C. (2007). Discontinuous membrane helices in transport proteins and their correlation with function. *J Struct Biol* **159**, 261–267.

59. Oberai, A., Ihm, Y., Kim, S. & Bowie, J. U. (2006). A limited universe of membrane protein families and folds. *Protein Sci* **15**, 1723–1734.

60. Landau, M., Herz, K., Padan, E. & Ben-Tal, N. (2007). Model structure of the Na^+/H^+ exchanger 1 (NHE1): functional and clinical implications. *J Biol Chem* **282**, 37854–37863.

61. Lovell, S. C. *et al.* (2003). Structure validation by c-α geometry: ϕ,ψ and c-β deviation. *Proteins* **50**, 437–450.

62. Pace, C. N. & Scholtz, J. M. (1998). A helix propensity scale based on experimental studies of peptides and proteins. *Biophys J* **75**, 422–427.

Chapter 5

Expression, purification and progress towards structure determination of a portion of the transmembrane domain of NHE1

The data presented in this chapter comprise a progress report of an unfinished story. Much of the supporting data is only qualitative and will be repeated more rigorously before publication. Nonetheless, it represents a nice application of the methods in *Chapter 3* and demonstrates that it is possible to study the membrane domain of proteins in larger sections, potentially offering a more insightful view of NHE1 structure.

This work was done in collaboration with Brian Lee and Dr. Brian Sykes (University of Alberta). All the NMR experiments were completed by Mr. Lee and advice on NMR sample preparations were taken from Mr. Lee and Dr. Sykes. The remainder of the work was completed by myself.

Multiple transmembrane segments of NHE1: a *slightly* divide-and-conquer approach

The purification techniques described in *Chapter 3*, particularly the high-performance liquid chromatography (HPLC) methods, have provided peptides for many published studies from our laboratory (1–5). Additionally, we have had much success expressing and purifying single transmembrane segments of NHE1 using various techniques (see *Chapter 1*). However, the resulting structures are still unable to confidently posit whether NHE has a NhaA-like fold (see *Chapter 6*). This, coupled with challenges in determining the high resolution structure of NHE1 (see *Chapter 2*), have pushed us to explore alternative options for understanding NHE1 structure. Therefore, we sought to expand our structural knowledge of NHE1 by expressing multiple contiguous transmembrane segments in one large construct to examine how these segments are arranged relative to each other. This technique, using maltose binding protein fusions, has already been successful (6).

Therefore we set about expressing larger sections of the membrane domain in the maltose binding protein fusion system. Initially, attempting two contiguous transmembrane segments (TMs), this effort was met with disappointing results. Most of the constructs would simply not express or resulted in only maltose binding protein without any fused peptide (unpublished observations). Only one construct, TM VI-VII, expressed well. When the chosen constructs were compared to the Landau model (7), it was noted that TM VI-VII was the only construct whose helices were folded together in three dimensional space (Figure 5.1). Intrigued by this anecdotal evidence supporting this model, we designed several new constructs centred around the *transport* bundle of helices. This is TM 3–5 (Landau) or TM V–VII (Wakabayashi) and TM 10–12 or TM X–XII. From this point I will be using Wakabayashi numbering as this is consistent with the published NMR solution structures.

Methods

Detailed methods are covered in *Chapter 3* and only a brief outline of the steps with the relevant modifications is included.

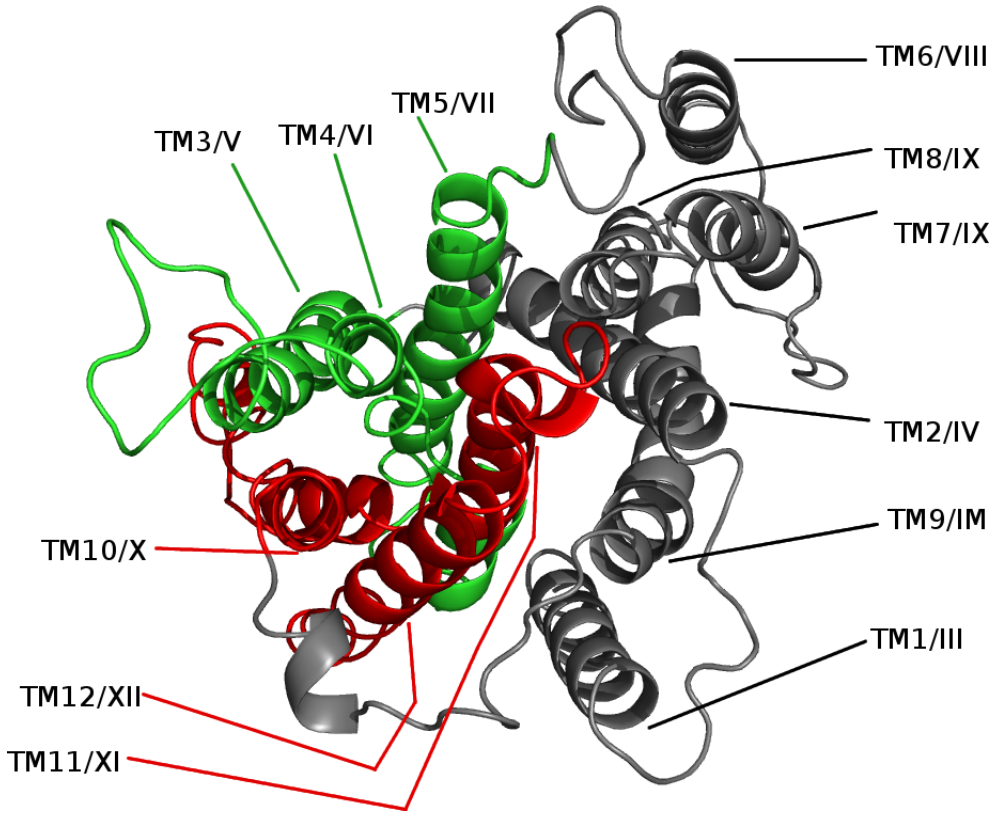

Figure 5.1: Directed selection of expression constructs using the Landau model. The transmembrane segments (TM) are labelled based in the Wakabayashi model (Roman numerals (8)) and the Landau model (Arabic numerals (7)). The long intramembrane loop between TM IX and X is labelled IM. The two contiguous three transmembrane segment sections, TM 3–5/V–VII and TM 10–12/X–XII, are coloured green and red, respectively.

Selection of constructs

Two three TM sections were chosen for expression screening: TM V-VII and TM X-XII. Owing to the controversial topology assignments of TM in NHE1, the sequences of these constructs were selected by consulting the existing models, structures, and functional data as well as doing *in-house* analysis of secondary structure and topology prediction (Figure 3.1 and 5.2). Based on both structural consensus as well as the presence of accessible residues at each terminus, the transmembrane segment sequences chosen are shown. Three lysines were added to each terminus to ease peptide purification by increasing its solubility and improving its behaviour in detergent micelles (9, 10).

Cloning and expression

TM V-VII and TM X-XII were cloned into pMal-c2X-HSY (Figure 4.3) and transformed into TB1 (JM83) *Escherichia coli* selected on LB+amp plates at 37°C. Glycerol stocks were made to preserve the transformants. For expression of the fusion protein, a glycerol stock is streaked onto a LB+amp plate and grown for 8–12 h at 37°C. A single colony is transferred using a sterile toothpick to a small overnight starter culture of LB+amp medium and grown overnight at 37°C. This starter culture is diluted at least 1/100 in 1 L of M9 medium and grown at 37°C until the A_{600} reaches 0.4–0.5. The culture is then transferred to 22°C for 30 min. Once the culture has cooled (A_{600} should be $\leq$0.6), expression is induced with 1 mM isopropyl β-D-1-thiogalactopyranoside at 22°C for 48 h. Gentle shaking is used throughout (<180 rpm). Initially an isotopically unlabelled peptide was produced to gauge its suitability for structure determination. Further experiments were done with either ^{15}N singly labelled or a $^{15}N/^{13}C$ doubly labelled peptide, which were produced by substituting the M9 medium with $^{15}NH_4$ and $^{13}C_6$–D–glucose (Cambridge Isotope Laboratories, Andover MA, USA).

Fusion protein purification

The induced cells were harvested and the fusion protein was purified essentially as described in *Chapter 3* using purification buffer containing 0.5x PSE and 20% glycerol. The peptide was cleaved from the maltose binding protein with tobacco etch virus protease (10 U/100 mg total protein) at 16°C for about 3 days (until cleavage was complete).

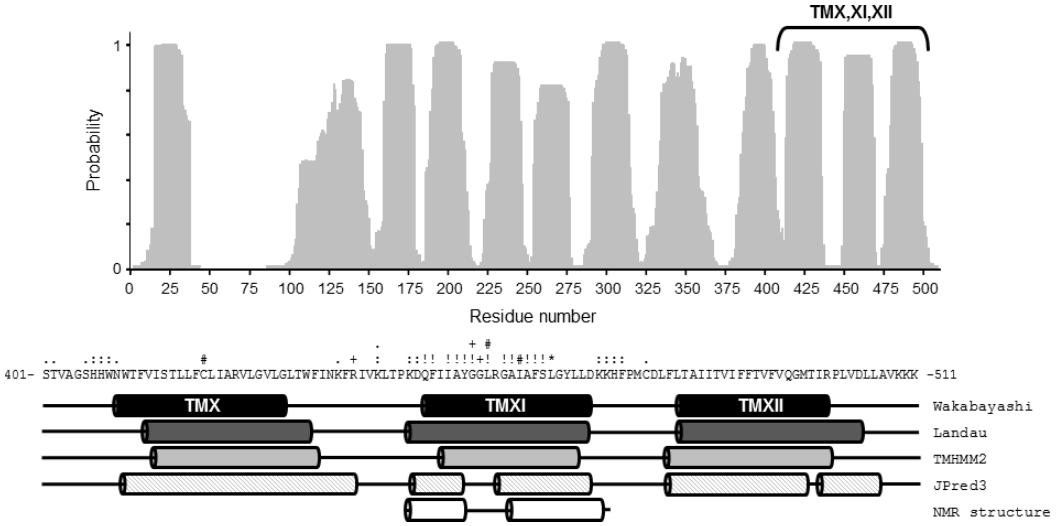

Figure 5.2: Transmembrane peptide design. Selecting an appropriate transmembrane section for expression and purification. The upper panel depicts the output of TMHMM2 prediction of membrane topology (11) for the human sodium proton exchanger isoform 1 (NHE1) residues 1–510. The probability of each amino acid (x-axis) being located in the membrane is shown on the y-axis. The lower panel displays the sequence of NHE1 between residues 401 and 511 with various transmembrane segment predictions, the regions of helical content determined by NMR, and known functional mutations. The black cylinders labelled *Wakabayashi* represent transmembrane helices predicted by Wakabayashi *et al.* (8). The dark grey cylinders labelled *Landau* represent transmembrane helices predicted by Landau *et al.* (7). The light grey cylinders labelled *TMHMM* correspond to the transmembrane regions from the upper panel. The hatched cylinders labelled *JPred3* represent helices predicted by the JPred3 algorithm (12). The white cylinders labelled *NMR structure* represent the solution structure of the purified peptide presented by Lee *et al.* (13). Above the sequence functionally important residues are labelled: (*) represents Cys mutants that inactivate NHE1 (13); (#) represents Cys mutants that are affected by cysteine modification reagent (13); (!) represents residues inaccessible to cysteine modifying agents (8); (:) represents Cys mutants that are accessible from the extra- or intracellular space (8). (+) represents mutants that had an effect on pH regulation (14); (.) represents mutants that were determined to be functionally non-essential (15–17).

Peptide purification

The entire cleavage reaction was precipitated with trichloroacetic acid and the pellet was subjected to organic extraction using chloroform-isopropanol-water (5:5:1). 32 mL of organic solvent (chloroform-isopropanol) mixture was used for every 100 mg of fusion protein that was cleaved. After two rounds of extraction the organic layer, containing the peptide, was carefully concentrated using a rotary evaporator ensuring the peptide did not come out of solution (≥5 ml). The extraction was exchanged into trifluoroethanol (TFE) by adding ~5 ml to the extraction then concentrating again. This was repeated until a negligible amount of chloroform and isopropanol remained. The peptide in TFE was injected onto a Zorbax SB300 C8 reverse phase high performance liquid chromatography column and the peptide was separated from the remaining contaminants and aggregates with a water-isopropanol gradient at 1 ml/min collecting 2 min fractions. The presence and purity of the peptide in each fraction was examined by SDS-PAGE and the pure peptide peak fractions were pooled. Sodium dodecyl sulphate (SDS) was added at ~ 1000 moles for every 1 mole of peptide and the mixture was concentrated in a rotary evaporator to remove all of the organic solvent. The peptide was lyophilized overnight and the dried peptide was resuspended at ~1 mM in 10 mM imidazole with 5% D_2O and 0.25 mM 4,4–dimethyl–4–silapentane–1–sulfonic acid added.

Size exclusion chromatography

A ^{15}N–labelled TM V-VII peptide was purified by organic extraction following cleavage from maltose binding protein. 5 mg of extracted peptide, without HPLC purification, was lyophilized and solubilized using 200 mg deuterated diphosphocholine (DPC) in 0.8 ml of pure water (as described above), for NMR experiments. However, the purity of this sample was insufficient to collect suitable heteronuclear single quantum coherence (HSQC) spectrum, so to prevent wasting the sample it was analyzed by size exclusion chromatography. 0.2 ml was loaded onto a Sephadex 200 column and 0.1% DPC was run at 0.4 ml/min, collecting 1 min fractions.

Circular dichroism

A ^{15}N–labelled TM V-VII peptide, purified by HPLC and solubilized in SDS, at a concentration of ~0.1 mM was placed in a Jasco J720 spectropolarimeter and five qualitative scans between

190 nm and 250 nm were collected.

NMR

Initial quality assessment ^{1}H spectra of peptide in the organic extraction were taken to judge peptide quality. For this the organic extraction was done in fully deuterated chloroform and isopropanol at a ratio of 5 ml of organic solvent mixture per 100 mg of peptide. Then a two dimensional ^{15}N HSQC spectrum was collected at 30°C in a 600 MHz magentic field of a ^{15}N labelled TM V-VII peptide that had been HPLC purified and solubilized in SDS at ~1 mM as described above. Added D_2O was used for signal locking and 4,4–dimethyl–4–silapentane–1–sulfonic acid for referencing spectra.

Results and discussion

Expression and purification

The expression of both TM V-VII and X-XII was successful. However, this expression is very sensitive to media, temperature and induction timing. No observable expression was obtained when induction was attempted in LB at either 37°C or 22°C. Induction in M9 media was weak or non-existent if cell density rose too high ($A_{600}\geq0.6$). Likewise, higher expression was obtained from cultures that were cooled to 22°C before induction (i.e. also before reaching $A_{600}\geq0.6$). Expression levels rose modestly from 24–48 h but no significant increase was evident after this. Despite this carefully optimized system, fusion protein yields were typically ≤200 mg/l, which is much lower than single transmembrane peptide constructs studied previously. For example, typical yields of sod2 TM IV (*Chapter 4*) were 400–450 mg/L. Also, the additional measures of reducing the salt concentration and adding glycerol were required to ensure that the fusion protein remained stable during purification. This is likely an indication that these longer constructs are more difficult for the cells to produce or possibly more fusion protein is being degraded. Despite these challenges, the fusion protein binds well to maltose affinity resin allowing the efficient elution of pure fusion protein (Figure 5.3-A). The peptide is also easily released by TEV proteolysis (Figure 5.3-B). Following cleavage, the peptide must be purified further. One strategy is to precipitate the protein and resolubilize it in a strong denaturant like concentrated

guanidinium-hydrochloride. However, neither of these constructs were significantly soluble in denaturant. Therefore, a different approach was taken and the peptide was extracted into organic solvent (Figure 5.3-C).

Because TM V-VII expressed better than TM X-XII, we focused on further purifying this construct for NMR. For HPLC purification of TM V-VII, initial efforts involved drying the peptide after organic extraction and solubilizing it in denaturant, but again the peptide remained insoluble. As isopropanol was too non-polar to allow the peptide to bind to a reverse-phase column, exchanging the peptide into a more polar solvent that was compatible with both column binding and peptide solubility was attempted. Trifluoroethanol (TFE) is a solvent that is often used to study folding of peptides and proteins because it is known to stabilize the formation of α-helices (18). Additionally, it enhances hydrophobic contacts making it a very effective solvent for protein folding regardless of hydrophobicity (19). By carefully exchanging the peptide into TFE, peptide solubility was maintained while creating a compatible solvent for peptide binding to a Zorbax SB300 C8 reverse phase HPLC column. The peptide eluted with high purity at 60–70% isopropanol (Figure 5.4). The HPLC purification step was crucial to collect high quality spectra. Previously, NMR analysis of the peptide directly from the organic extraction was attempted, both in deuterated solvents and resolubilized in deuterated diphosphocholine, but the resulting two dimensional spectra were noisy and unusable. This is likely due to either aggregation, impurities or both. Analysis by SDS-PAGE indicated that the organically extracted peptide was already quite pure and while HPLC removed some impurities it also greatly reduced the amount of recovered peptide. This could have been an indication that a significant portion of the peptide was soluble but somewhat aggregated, leading to lower quality NMR spectra. Thus, the HPLC step could be viewed as a polishing step that increased the homogeneity and monodispersity of the sample.

Biophysical characterization

Previously, in our hands, this purification system had only been used to structurally and functionally characterize single transmembrane segments, which seem to spontaneously adopt their native structures (20). During cellular protein translation, newly synthesized transmembrane segments are inserted into the membrane co-translationally and then the folded segments assemble together in the membrane (21). Therefore, our hypothesis is that three transmembrane segments that form a folded structure together will adopt their native conformation when expressed as a

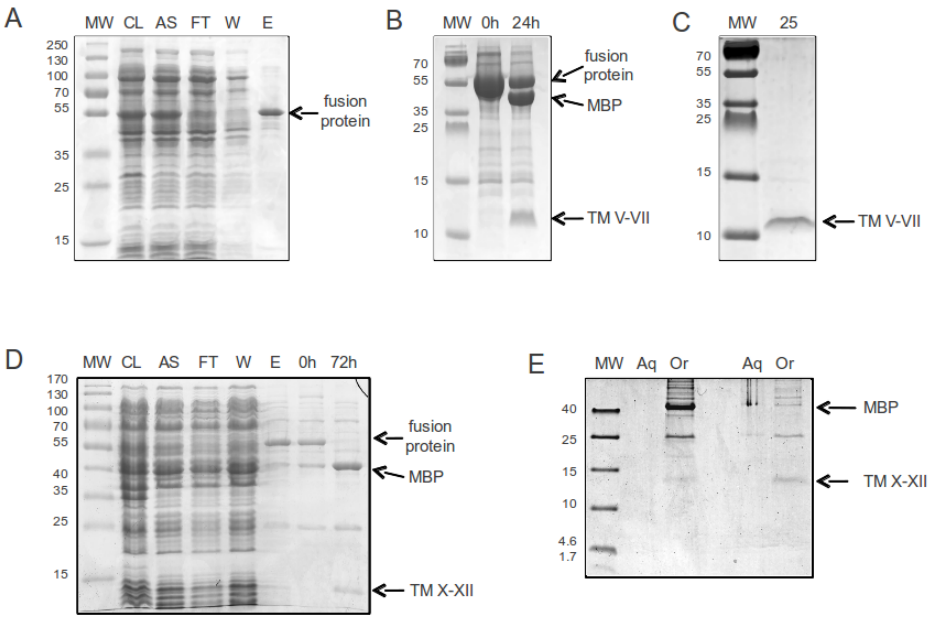

Figure 5.3: SDS-PAGE analysis of TM V-VII and TM X-XII purification. Molecular weight ladder (MW) as marked, in kDa. ***A*** 12% Tris-Glycine SDS-PAGE gel showing fractions of the purification of TM V-VII fusion protein from 2 L of TB1 cells: CL, crude lysate (5 µL); AS, after lysate ultracentrifugation (5 µL); FT, amylose column flow through (5 µL); W, amylose column wash (10 µL); E, amylose column elution (2 µL, ~10 µg). The position of the fusion protein is marked. ***B*** 16% Tris-Tricine SDS-PAGE gel showing TEV protease digestion progress of TM V-VII fusion protein: 0 h, concentrated elution before TEV addition; 24 h, sample after 24 h digestion at 16 °C. The position of the fusion protein, cleaved maltose binding protein and the peptide are marked. ***C*** 16% Tris-Tricine SDS-PAGE gel showing 25 µL sample of the organic phase of the organic extraction of TM V-VII, dried then resuspended in sample loading buffer. The position of the peptide is marked. ***D***, 12% Tris-Glycine SDS-PAGE gel showing fractions of the purification and proteolytic cleavage of TM X-XII fusion protein from 4 L of TB1 cells: CL, crude lysate (5 µL); AS, after lysate ultracentrifugation (5 µL); FT, amylose column flow through (5 µL); W, amylose column wash (10 µL); E, amylose column elution (2 µL, ~5 µg); 0h, concentrated elution before TEV addition; 72h, sample after 72 h digestion at 16 °C. The position of the fusion protein and free maltose binding protein (MBP) are marked. ***E*** 16% Tris-Tricine SDS-PAGE gel showing the result of the organic extraction of digested TM X-XII: Aq1, 25 µL sample of the aqueous phase after one round of extraction; Or1, 100 µL sample of the organic phase after one round of extraction; Aq2, 25 µL sample of the aqueous phase after two rounds of extraction; Or2, 100 µL sample of the organic phase after two rounds of extraction. The position of TM X-XII and maltose binding protein (MBP) are marked.

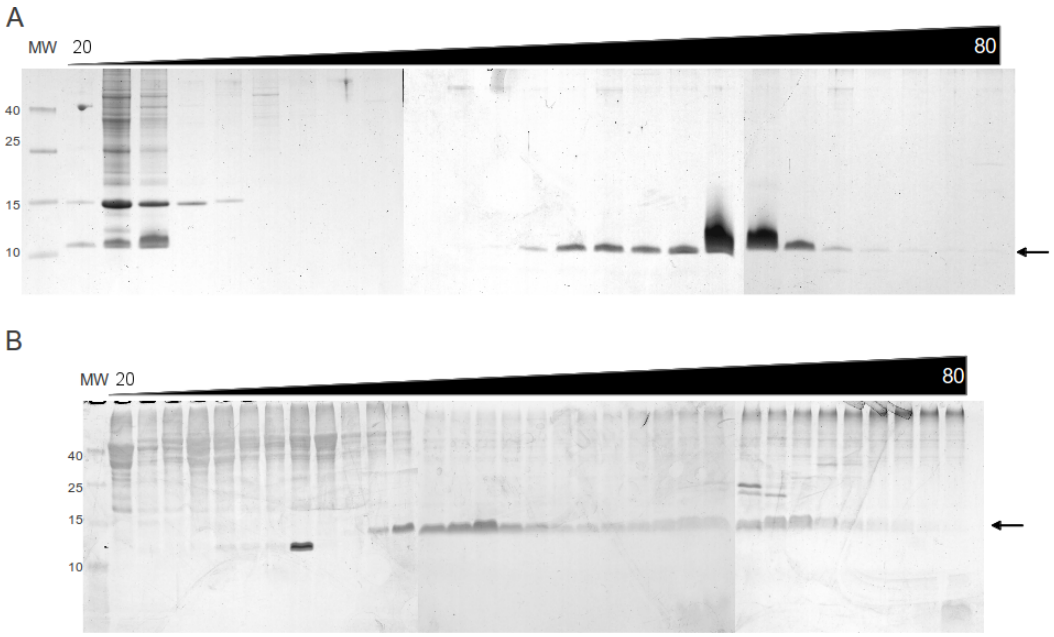

Figure 5.4: SDS-PAGE analysis of HPLC purification of TM V-VII and TM X-XII following organic extraction on a Zorbax SB300 C8 reverse phase column. Molecular weight ladder (MW) as marked, in kDa. Increasing concentration of isopropanol in the mobile phase (20–80%) is indicated, but it is not linear. ***A***, 16% Tris-Tricine SDS-PAGE gel stained with Serva blue showing the purification of TM V-VII. Early fractions contain most of the contaminants that do not bind to the column. Later fractions (60–70%) contain purified TM V-VII. The molecular weight of TM V-VII is marked by an arrow. ***B***, 16% Tris-Tricine SDS-PAGE gel stained with silver stain showing the purification of TM X-XII. Early fractions contain most of the contaminants that do not bind to the column. TM X-XII appears to *bleed* off the column slowly with an early pure peak (middle fractions, 60–65%) followed by a later less pure peak. The molecular weight of TM X-XII is marked by an arrow.

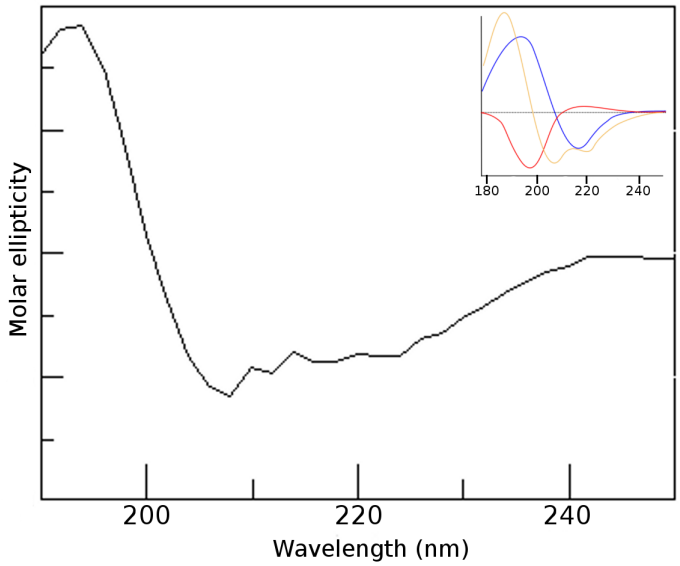

Figure 5.5: A qualitative circular dichroism spectrum of the TM V-VII peptide at ~0.1 mM in SDS solution between 190 nm and 250 nm. This curve is the average of five scans. Ideal reference spectra are shown in the top corner with α-helix in yellow, β-sheet in blue, and random coil in red.

fusion protein. As mentioned above, other constructs attempted in the lab that did not express also did not form a folded cluster of helices in the Landau molecular model (7). This anecdotal evidence suggests that indeed TM V-VII and TM X-XII are folded *in vivo*. However, does this purification scheme maintain this fold or, if they become partially unfolded during the process, are our experimental conditions suitable for spontaneous refolding? Although an NMR solution structure will answer this question, these are lengthy, costly and time-consuming experiments. To gauge whether NMR was worth pursuing, I conducted qualitative circular dichroism (CD) and size exclusion chromatography (SE) experiments using TM V-VII to qualitatively measure the secondary structural characteristics and determine if the peptide had a compact (folded) structure.

The shape of the CD spectrum indicated that the peptide has an α-helical structure, as expected, in SDS solution (Figure 5.5). This means that each individual transmembrane segment is folding into the appropriate secondary structure. To get a better measure of whether the construct was folded, we conducted a SE experiment (Figure 5.6-A). If each transmembrane segment had assumed the correct secondary structure, but had not folded into the correct tertiary structure, the peptide would appear less compact and more elongated as indicated by a larger hydrodynamic radius (Figure 5.6-B). However, if the peptide adopted its native tertiary structure, it would ap-

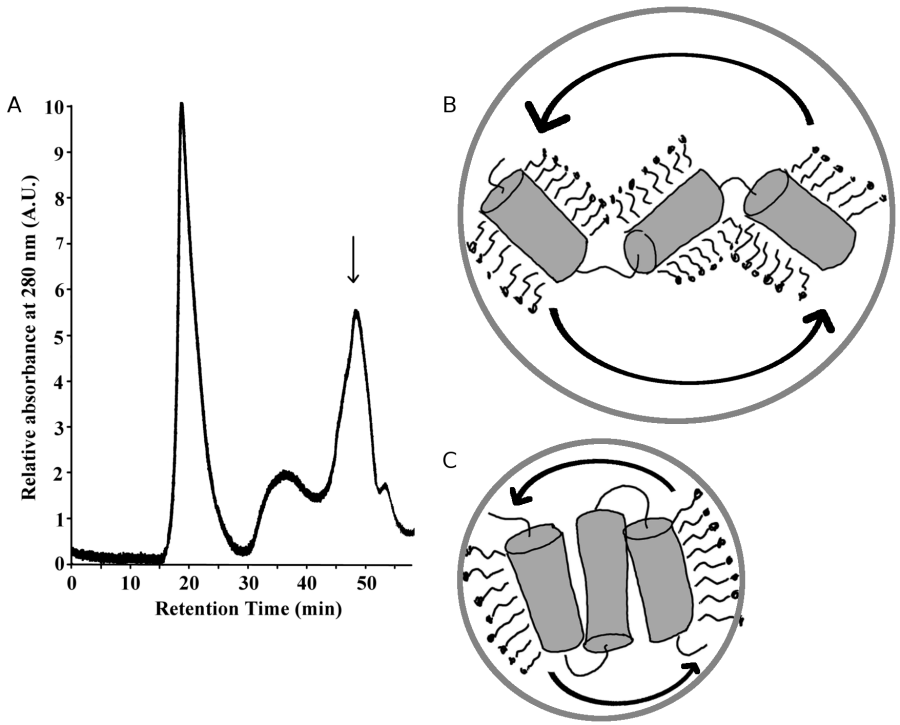

Figure 5.6: Size exclusion chromatogram of TM V-VII. ***A***, Size exclusion chromatogram (A_{280}) of TM V-VII run over a Sephadex 200 column. Three peaks were observed corresponding to an aggregate peak at the void volume, a ~44 kDa peak, and a ~17 kDa peak. The final peak, labelled with an arrow, was consistent with TM V-VII plus some associated detergent. ***B***, Cartoon depiction of the large hydrodynamic radius of an unfolded three transmembrane segment peptide. ***C***, Cartoon depiction of the compact hydrodynamic radius of a folded three transmembrane segment peptide.

pear more compact with a hydrodynamic radius closer to its molecular mass (Figure 5.6-C). Three observable peaks by SE chromatography indicated that there were three primary groups of peptide structures. The first peak corresponded to the void volume indicating that some of the peptide formed aggregates. The second peak at ~44 kDa could correspond to the unfolded structure shown in Figure 5.6-B. The final peak at ~17 kDa was appropriate for TM V-VII plus the associated detergent. This data indicated that at least some of the peptide was forming a folded conformation, but since the SE sample was prepared without HPLC purification a significant amount of aggregation was still present. Although SE was not repeated following HPLC, the large decrease in peptide yield and the vast improvement in the NMR spectra quality suggested that this additional purification step greatly increased homogeneity.

Following the optimization of the purification process and sample conditions for NMR we

collected a high quality two dimensional ^{15}N HSQC spectrum of TM V-VII (Figure 5.7). The dispersion, sharpness and number of peaks indicated that the peptide appeared to be structured and was suitable for further analysis. However, this project is not yet completed due to time restrictions in completing my doctoral program.

Future directions

Production of a $^{15}N/^{13}C$ doubly labelled peptide will enable the collection of more complicated three dimensional spectra that will allow the elucidation of the structure of TM V-VII (in collaboration with the laboratory of Dr. Brian Sykes). Furthermore, this optimized system could be used to produce TM X-XII. Lower fusion protein yields and time constraints led us to focus work on TM V-VII, but we now have a detailed experimental protocol for handling these larger membrane fragments.

Once experiments are complete, the structures of these two three transmembrane segment bundles should provide more evidence demonstrating whether the structure of NHE1 indeed has an NhaA-like fold. TM V-VII is particularly interesting because the precise functional role of this region has been debated in the literature (8, 22) (see *Chapter 6*). It will also demonstrate that this is another useful tool available to membrane protein structural biologists. Many eukaryotic membrane proteins are difficult to express and purify, hampering further progress in structure determination. This method can provide a meaningful way to elucidate some structural details about membrane proteins that have evaded significant progress.

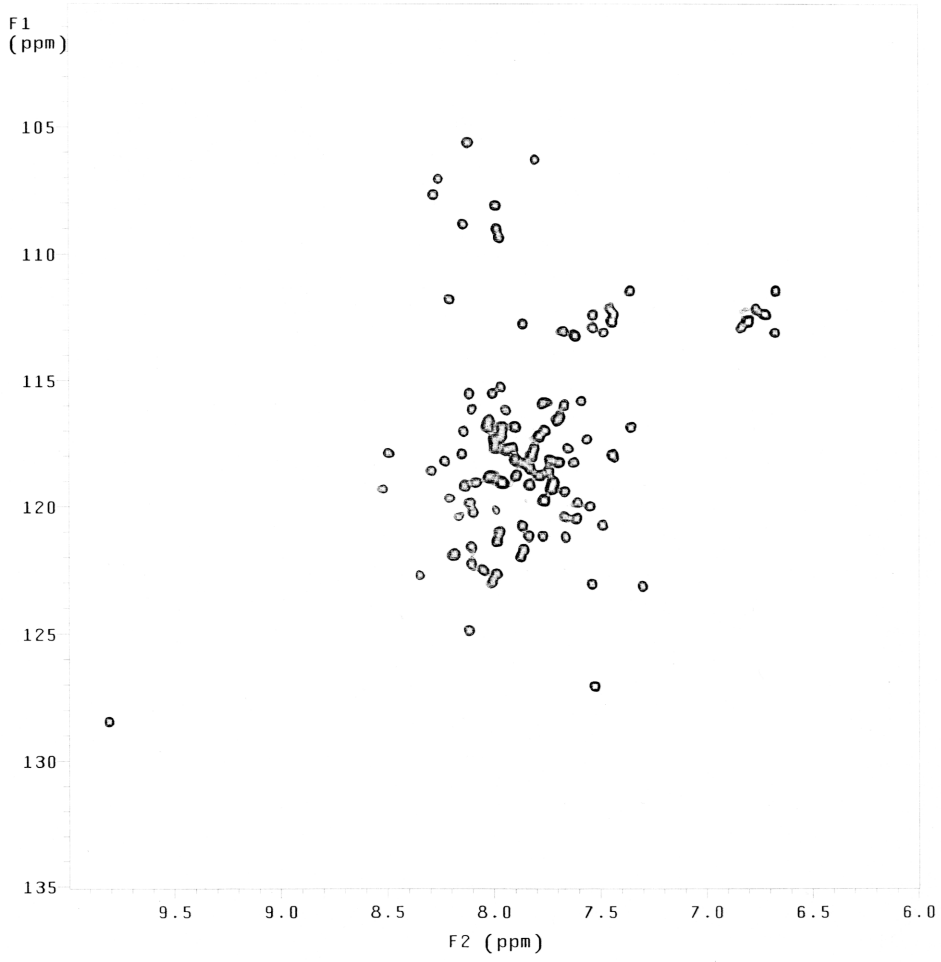

Figure 5.7: A two dimensional ^{15}N heteronuclear single quantum coherence spectrum of HPLC-purified TM V-VII in 0.1% SDS, 10 mM imidazole, 0.5% D_2O, 0.25 mM 4,4–dimethyl–4–silapentane–1–sulfonic acid. The spectrum was collected at 30°C and 600 MHz.

References

1. Glaves, J. P., Trieber, C. A., Ceholski, D. K., Stokes, D. L. & Young, H. S. (2011). Phosphorylation and mutation of phospholamban alter physical interactions with the sarcoplasmic reticulum calcium pump. *J Mol Biol* **405**, 707–723.

2. Ceholski, D. K., Trieber, C. A., Holmes, C. F. B. & Young, H. S. (2012). Lethal, hereditary mutants of phospholamban elude phosphorylation by protein kinase A. *J Biol Chem* **287**, 26596–26605.

3. Ceholski, D. K., Trieber, C. A. & Young, H. S. (2012). Hydrophobic imbalance in the cytoplasmic domain of phospholamban is a determinant for lethal dilated cardiomyopathy. *J Biol Chem* **287**, 16521–16529.

4. Gorski, P. A. *et al.* (2012). Transmembrane helix 11 is a genuine regulator of the endoplasmic reticulum ca2+ pump and acts as a functional parallel of -subunit on -Na+,K+-ATPase. *J Biol Chem* **287**, 19876–19885.

5. Gorski, P. A., Glaves, J. P., Vangheluwe, P. & Young, H. S. (2013). Sarco(endo)plasmic reticulum calcium ATPase (SERCA) inhibition by sarcolipin is encoded in its luminal tail. *J Biol Chem* **288**, 8456–8467.

6. Hu, J. *et al.* (2007). Structural biology of transmembrane domains: Efficient production and characterization of transmembrane peptides by NMR. *Protein Sci* **16**, 2153–2165.

7. Landau, M., Herz, K., Padan, E. & Ben-Tal, N. (2007). Model structure of the Na^+/H^+ exchanger 1 (NHE1): functional and clinical implications. *J Biol Chem* **282**, 37854–37863.

8. Wakabayashi, S., Pang, T., Su, X. & Shigekawa, M. (2000). A novel topology model of the human Na^+/H^+ exchanger isoform 1. *J Biol Chem* **275**, 7942–7949.

9. Melnyk, R. A. *et al.* (2003). Polar residue tagging of transmembrane peptides. *Pept Sci* **71**, 675–685.

10. Cunningham, F. & Deber, C. M. (2007). Optimizing synthesis and expression of transmembrane peptides and proteins. *Methods (San Diego, Calif.)* **41**, 370–380.

11. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *J Mol Biol* **305**, 567–580.

12. Cole, C., Barber, J. D. & Barton, G. J. (2008). The Jpred 3 secondary structure prediction server. *Nucleic Acids Res* **36**, W197–W201.

13. Lee, B. L., Li, X., Liu, Y., Sykes, B. D. & Fliegel, L. (2009). Structural and functional analysis of transmembrane XI of the NHE1 isoform of the Na^+/H^+ exchanger. *J Biol Chem* **284**, 11546–11556.

14. Wakabayashi, S., Hisamitsu, T., Pang, T. & Shigekawa, M. (2003). Kinetic dissection of two distinct proton binding sites in Na^+/H^+ exchangers by measurement of reverse mode reaction. *J Biol Chem* **278**, 43580–43585.

15. Counillon, L., Pouysségur, J. & Reithmeier, R. A. (1994). The Na^+/H^+ exchanger NHE-1 possesses N- and O-linked glycosylation restricted to the first N-terminal extracellular domain. *Biochemistry* **33**, 10463–10469.

16. Wang, H., Singh, D. & Fliegel, L. (1998). Functional role of cysteine residues in the Na^+/H^+ exchanger effects of mutation of cysteine residues on targeting and activity of the Na^+/H^+ exchanger. *Arch. Biochem. Biophys.* **358**, 116–124.

17. Murtazina, R., Booth, B. J., Bullis, B. L., Singh, D. N. & Fliegel, L. (2001). Functional analysis of polar amino-acid residues in membrane associated regions of the NHE1 isoform of the mammalian Na^+/H^+ exchanger. *Eur J Biochem* **268**, 4674–4685.

18. Myers, J. K., Pace, C. N. & Scholtz, J. M. (1998). Trifluoroethanol effects on helix propensity and electrostatic interactions in the helical peptide from ribonuclease T1. *Protein Sci* **7**, 383–388.

19. Albert, J. S. & Hamilton, A. D. (1995). Stabilization of helical domains in short peptides using hydrophobic interactions. *Biochemistry* **34**, 984–990.

20. Katragadda, M., Alderfer, J. L. & Yeagle, P. L. (2001). Assembly of a polytopic membrane protein structure from the solution structures of overlapping peptide fragments of bacteriorhodopsin. *Biophys J* **81**, 1029–1036.

21. White, S. H. & von Heijne, G. (2008). How translocons select transmembrane helices. *Annu Rev Biophys* **37**, 23–42.

22. Nygaard, E. B. *et al.* (2011). Structural modeling and electron paramagnetic resonance spectroscopy of the human Na^+/H^+ exchanger isoform 1, NHE1. *J Biol Chem* **286**, 634–648.

Chapter 6

Conclusion and Future Directions

I have outlined in this thesis the central role that Na^+/H^+ exchangers play in cell physiology. This is particularly true for human NHE1, which has been demonstrated to be a cellular *butler* acting as both a pH regulatory protein and as a cell signalling scaffold. It is able to respond to and propagate events that lead to regulation of cell volume, growth, proliferation, differentiation, survival and death. It is also evident that NHE1 dysregulation plays an essential role in the development of heart disease and cancer. I have presented a summary of the structural and functional data collected in the past quarter century since the first Na^+/H^+ exchanger was cloned, including my contributions and those of my close collaborators. Below, I apply this data to describe a new model of NHE1 structure and function.

The tale of two models

As previously introduced, two popular topology models of NHE1 have been published (1, 2). The methods used to construct these models and their features were covered in *Chapter 1*. Here I compare these models in more depth, weighing them against our current structure-function knowledge of NHE1. To avoid confusion, I will use Roman numerals to indicate Wakabayashi transmembrane segments (1) and Arabic numerals to indicate Landau segments (2).

In 2000, with fewer than 20 point mutations in the membrane domain of NHE1 previously described, no concrete structural information on NHE1 was available. To glean insight into where regions of the protein were located in relation to the plasma membrane, Wakabayashi *et*

al. constructed a topology model tested with 83 novel introduced cysteine mutations. Previously the nine native cysteines in NHE1 had been individually mutated and tested for expression, membrane targeting and activity, but here the authors were looking for potential disulfide bridges (3). This was the first systematic large scale mutational study of NHE1 providing information about the effect of these mutations on activity and where these residues were located in the mature protein on the cell surface. Seven years later Landau *et al.* published a novel topology model based on a three dimensional homology model of NHE1 constructed using the recently elucidated crystal structure of *Escherichia coli* NhaA (4). At this time, a further 50 point mutations in the membrane domain of NHE1 had been described giving the authors a much larger pool of data with which to analyze their model. Despite this, the models are quite similar with only two large discrepancies: the amino-terminal 120 amino acids (the first 2 transmembrane segments of the Wakabayashi model) and amino acids 330–410 (region around Wakabayashi's long intramembrane loop). Figure 6.1 shows a comparison of the two models.

Interestingly, these two regions have long been controversial. A previously published limited topological analysis had suggested that the first transmembrane segment of NHE1 may actually be a signal peptide with the authors showing that the first extracellular loop is sensitive to limited chymotryptic digestion and that digestion has no effect on pH regulation (5). As the first two transmembrane segments show a high degree of sequence variability within the SLC9A subgroup, and signal peptides were later identified on NHE3 (6) and NHE6 (7), this idea garnered some support in the community. However, no further evidence supporting the existence of an NHE1 signal peptide has been published, and the following studies have suggested that this region is indeed intact on the cell surface. Processing of NHE1 in the endoplasmic reticulum includes N- and O-linked glycosylation in the first extracellular loop and, while NHE1 deglycosylation does not affect activity, these sites have been shown to be at the cell surface (5, 8, 9). It has also been shown that this loop has a cell volume regulating mechanism, that is not present in NHE2 or NHE3, and a chimeric NHE1 containing the amino-terminus of NHE2 loses volume regulation (10). This unique regulatory mechanism that is present in the more recently evolved NHE1 could possibly explain the lack of sequence homology in this region. Furthermore, in the Wakabayashi model several introduced cysteines in the first extracellular loop were accessible in intact cells. These three experiments demonstrate that the first extracellular loop is indeed located on the outside of the cell, suggesting that the topology of Wakabayashi segment II is accurate.

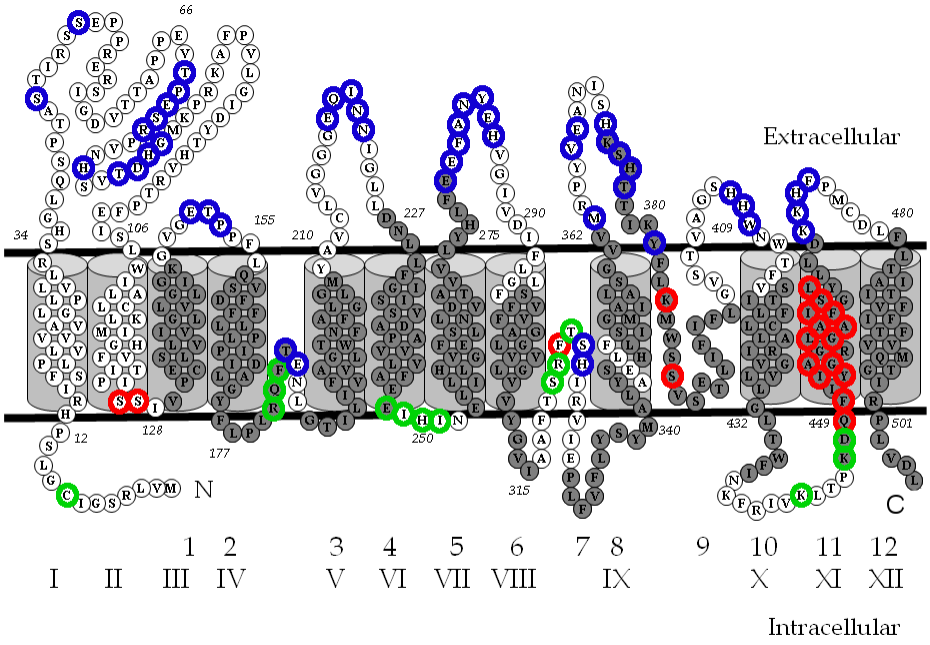

Figure 6.1: Comparison of the Wakabayashi and Landau topology models of NHE1. All amino acids in the putative membrane domain are labelled with single letter code, the membrane boundaries are marked with black lines, and Wakabayashi transmembrane segments are highlighted by grey cylinders. To compare the models, residues in Landau transmembrane helices are coloured in dark grey. For reference, introduced cysteines subjected to Substituted Cysteine Accessibility Method (SCAM) are surrounded by rings coloured as follows: blue, accessible on the extracellular side; green, accessible on the intracellular side; red, inaccessible.

Additionally, native Cys8 was labelled only in permeabilized cells indicating that the topology of this region is accurate. It has also been previously found that mutation of Cys8 to serine leads to improper processing of NHE1 (3), while the *cysteineless* protein was processed normally (1, 11). This behaviour has also been seen for some individual *S. pombe* sod2 cysteine mutants but not the cysteineless construct (Dr. Ullah, unpublished observations). This may indicate that specific unpaired cysteines in Na^+/H^+ exchangers lead to misfolding and that Cys8 of NHE1 is likely present in the folded protein and therefore unlikely to be removed in a signal peptide. These data together suggest that Wakabayashi segments I and II may not be essential to NHE1 function (5) but necessary for proper folding and targeting to the plasma membrane (1, 3) and may have a unique function in regulated cell volume control (10). Conversely, the first two TM segments of NhaA are important for activity (4), particularly segment 2 (12). This makes it conceivable that Wakabayashi segments I and II, which are absent in the Landau model, are not analogous to NhaA segments 1 and 2 and therefore not part of the canonical twelve transmembrane segments found in CPA transporters.

The next six transmembrane segments of both models agree relatively well (III–VIII and 1–6). The intracellular loops in the Wakabayashi model that have contiguous internally and externally accessible residues are also located in intracellular loops of the Landau model and therefore can be explained by the same reasoning. At this point (carboxy-terminus of segment VIII/6) the models diverge significantly. While both begin in the cytosol, Wakabayashi predicts a single transmembrane segment followed by a long intramembrane loop coming out on the extracellular side of the membrane whereas Landau predicts two short transmembrane helices (~15 residues each) followed by a longer helix (~24 residues) that yields three membrane passes and also comes out on the extracellular side. Before discussing the empirical support for each model, it is important to mention the technical biases associated with each method. Hydropathy modelling, the basis of the Wakabayashi model, predicts *standard* transmembrane helix lengths of 20-22 amino acids. This is sufficient to cross a *typical* phospholipid bilayer but it does not account for: protein-protein hydrophobic matching in a folded polytopic membrane protein; tilted, kinked or partially unwound helices; differences in lipid bilayer thicknesses between prokaryotes, eukuryotes, organelles and lipid microdomains. Indeed, the segment length in the Wakabayashi model is consistently ~20 residues long and in reality transmembrane segments can vary in length significantly (13, 14). Improved methods of transmembrane helix prediction

were used by Landau *et al.* (and myself in *Chapters 3* and *4*) in determining a reasonable sequence alignment. However, homology modelling, the basis for Landau's predictions, requires an aligned structural template that in and of itself may lead to unfairly biased transmembrane segment predictions. With those qualifications, Wakabayashi found their cysteine accessibility measurements indicated extracellular accessible residues on both ends of a transmembrane helix predicted by the Kyte-Doolittle hydropathy prediction. Specifically, the Kyte-Doolittle algorithm predicted that amino acids 373–377 and 381 were extracellular and amino acids 407–409 were intracellular. However, mutagenesis found that all these residues were accessible on the extracellular side. To explain this phenomenon they suggested that this segment was actually a long intramembrane loop. This topology was further supported by an independent group that used a cell-free expression system with functional rough microsomes to demonstrate that while both transmembrane segments IX and X cross the membrane, the loop between them is inserted into the membrane but does not cross it (15, 16). In contrast to this, the Landau model places amino acids 373–377 and 381 on the intracellular side with most of the residues actually being a part of the segment not in the loop. While this topology is consistent with their sequence alignment and the structure of NhaA, it is difficult to explain in light of the accessibility studies.

In addition to the two dimensional topology, Landau *et al.* have also produced a three dimensional homology model. The authors did an exhaustive job of examining their model in light of published mutagenesis experiments, showing that residues involved in ion translocation cluster in the putative transport bundle, while nonessential residues are primarily found towards the periphery (Figure 6.2). The mutants affecting pH regulation do appear to follow a particular cluster pattern but are all located at or very near the intracellular surface (except Gly455 and Gly456 which are in the middle of TM XI/11), but this position is identical in both models. Residues implicated in drug-binding cluster nicely together near the extracellular surface with three residues near the end of TM 2 (IV) and two in the loop between the short segments 7 and 8. Interestingly in the Wakbayashi model the second two residues are now in TM IX, however a small shift in how this segment sits in the membrane would cluster these residues with the other three on TM IV. In addition, Landau *et al.* justify the topology of TM 7 and 8 by suggesting that Ser351, which has functionally important analogous residues (17, 18), becomes a pore lining residue following a conformational change during transport. This prediction was later corroborated by both mutagenesis and structure determination (see below) (19).

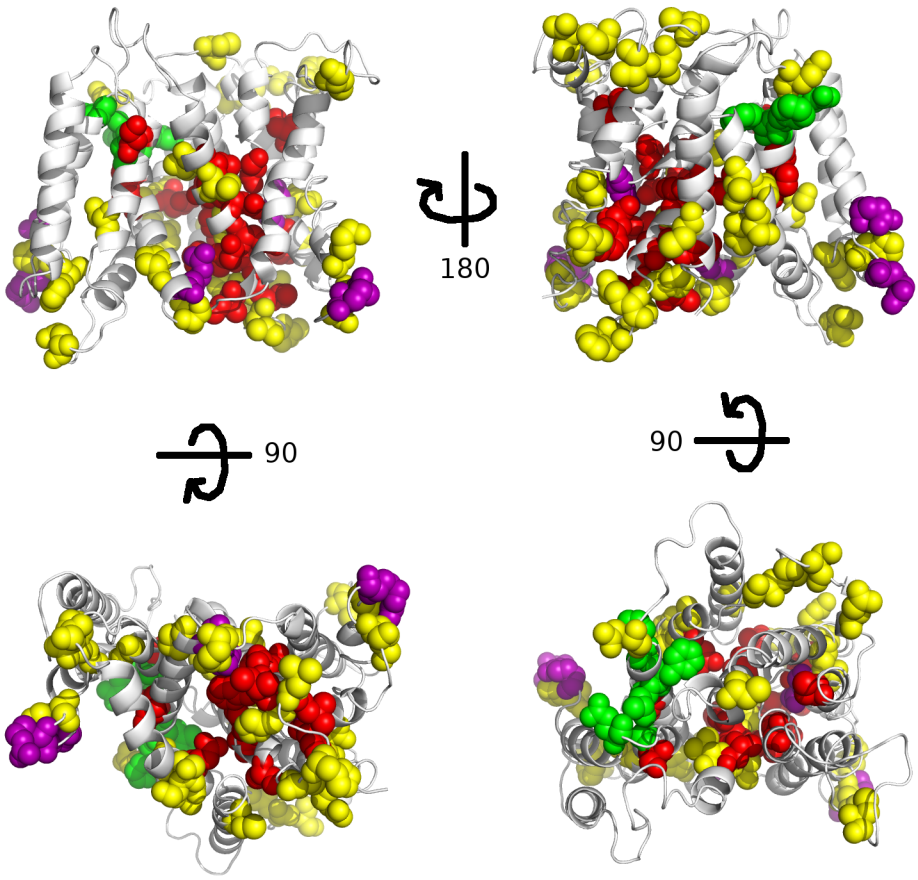

Figure 6.2: NHE1 mutagenesis data mapped onto the Landau homology model. This figure was reproduced based on the original figure by Landau et al. (2). The model is shown as a ribbon drawing in white from four angles. Residues that have been mutated are shown as spheres and coloured as follows: red, implicated in ion translocation; yellow, unessential; green, inhibitor-binding; purple, pH regulation.

Besides mutagenesis, the NMR structures of some short sequences of NHE1 can also be used to scrutinize the homology model (Figure 6.3). The first structure determined was TM IV (amino acids 155–180), which was shown to have an extended region in the middle of the segment and only residues 169–176 were α-helical (11). Immediately following this, the crystal structure of *E. coli* NhaA was published revealing that TM IV of NhaA also featured an extended region (4). Despite a lack of significant amino acid identity between the segments, this structural similarity and the fact that both contained residues important to transporter function, led to the conclusion that they were analogous (20). Therefore, when the Landau model was released, the alignment of TM IV (2) with NhaA TM II did not seem consistent with the NMR data (21). Also becoming available at the time was the NMR structure of TM VII (TM 5) (22). This transmembrane segment is the same in both models and the NMR structure agrees with the homology model. Following the Landau model, three more NMR structures have been elucidated, two of which favour this newer model of NHE1. The structure of amino acids 339–363 (TM IX or TM 7/8) shows a severely kinked peptide with helices at either end (19). This structure is more similar to the two short helices of the Landau model and it demonstrates that Ser351 faces the membrane lipids, as predicted in the Landau model. The NMR structure of TM XI (TM 11) was also elucidated, showing a kinked helix containing an extended region (23), precisely as predicted by the Landau model. Despite this information, it is still difficult to rule out either model. However, the recent structure-function analysis of TM VI (TM 4) was determined revealing another kinked helix with an extended region containing residues critical to ion translocation (24).

As stated above, the initial evidence (4, 11, 25, 26) suggested that NHE1 TM IV was analogous to TM IV of NhaA (20). However, the publication of the Landau model questioned this conclusion as transmembrane segment IV was not aligned in the three dimensional model with NhaA segment IV. Seeking clarification, other members of the Fliegel laboratory undertook mutagenesis and structure determination of transmembrane segment VI, which in Landau's model was analogous to NhaA segment IV (24). It was discovered that indeed transmembrane segment VI was structurally very similar to NhaA segment IV and that it contained several residues critical to NHE1 function (27). At about the same time, NhaA segment II was shown to be involved in conformational changes during ion transport (12) suggesting that, although it does not appear structurally similar to NHE1 segment IV (4, 11), they may share some functional equivalence. This strongly suggests that the Landau model is plausible. To solve the discrep-

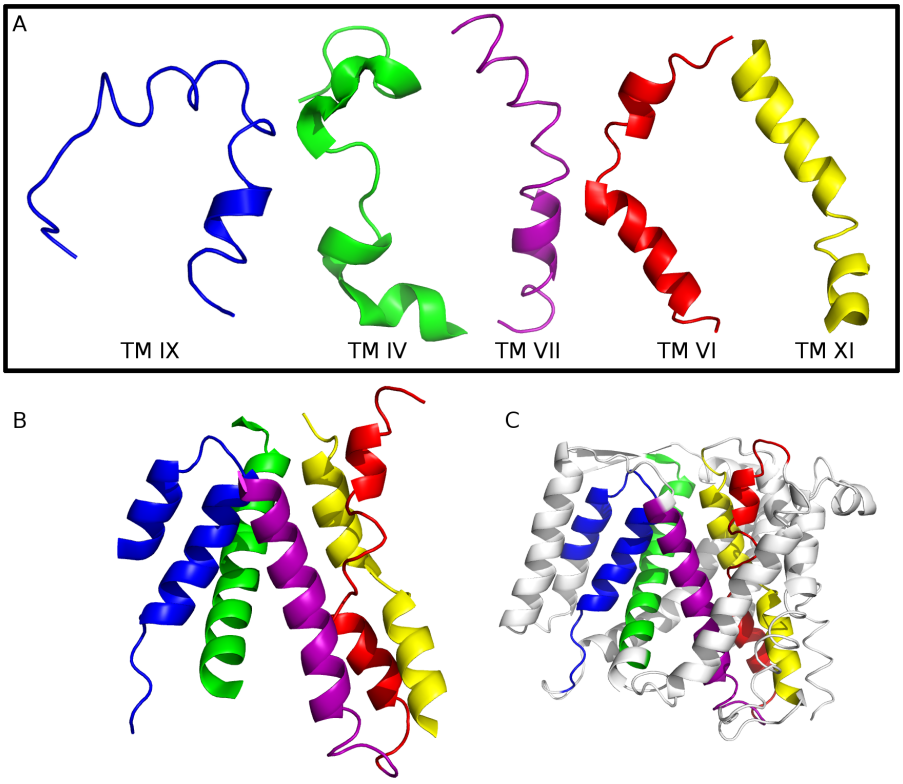

Figure 6.3: NMR structures of NHE1 transmembrane segments. ***A***, Cartoon representation of the five solved solution structures of NHE1 transmembrane segments. Each transmembrane segment (TM) is labelled with numbering from the Wakabayashi topology model (1): TM IX, blue (19); TM IV, green (11); TM VII, purple (22); TM VI, red (24); TM XI, yellow (23). ***B***, The analogous transmembrane helices of *Escherichia coli* NhaA as arranged in the crystal structure (4) based on alignment by Landau *et al.* (2): TM 7/8, blue; TM 2, green; TM 5, purple; TM 4, red; TM 11, yellow. ***C***, The complete crystal structure of NhaA (4) coloured as in ***B*** for reference.

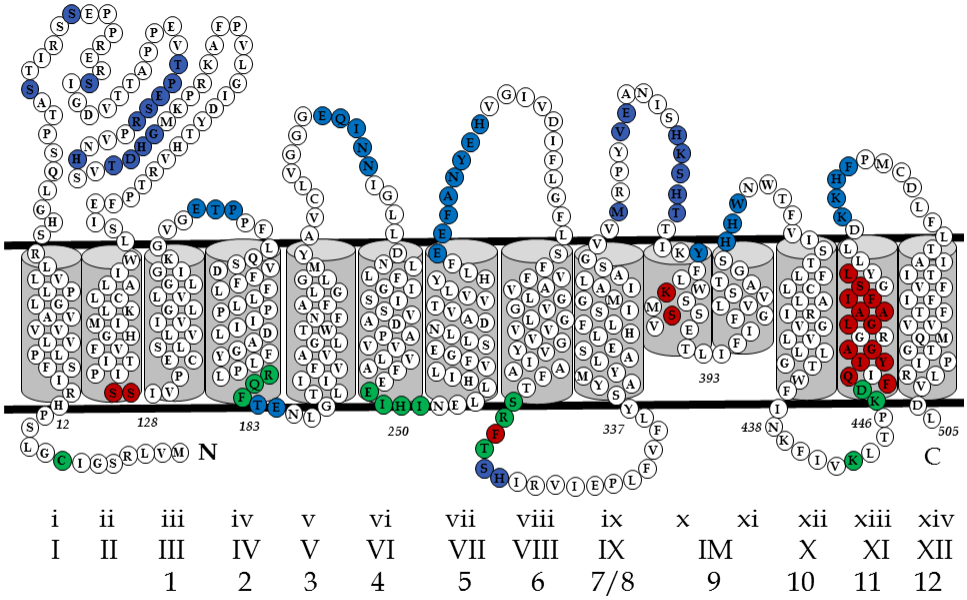

Figure 6.4: Novel *Alberta* topology model of NHE1. All amino acids in the putative membrane domain are labelled with single letter code, the membrane boundaries are marked with black lines, and transmembrane segments are highlighted by grey cylinders. For reference, introduced cysteines subjected to Substituted Cysteine Accessibility Method (SCAM) are coloured as follows: blue, accessible on the extracellular side; green, accessible on the intracellular side; red, inaccessible. The transmembrane helix numbers are labelled below: Alberta model, lower case Roman numerals; Wakabayashi model, upper case Roman numerals (except *intramembrane loop*, IM); Landau, Arabic numerals.

ancy in the accessibility of TM IX plus the intramembrane loop (TM 7/8/9), I propose a novel topology (Figure 6.4). Here TM IX (TM 7 and 8) remains a complete transmembrane helix and the intramembrane loop proposed by Wakabayashi, which is nearly the same length as TM 7 and 8 proposed by Landau, is modelled as two short helices reproducing this NhaA structural feature. This model satisfies the accessibility mutants and is still consistent with the mutagenesis mapping (Figure 6.2).

It should be noted that a third model of NHE1 also exists described by Nygaard *et al.* using electron paramagnetic resonance measurements and homology modelling (28). Like the Landau model, it is also presented in three dimensions but generally follows the topology of the Wakabayashi model. However, there is no evidence that supports the Nygaard model over the Landau model as the novel biophysical data presented by Nygaard *et al.* (29) are compatible

with both models (30). Most importantly, this model places residues that have been implicated in ion translocation (TM VI and VII) on the periphery of the protein, which is an unlikely structural arrangement.

The other unique feature of the novel topology model is that it contains fourteen transmembrane segments. Although CPA members, typified by NhaA, are defined as having 10–12 transmembrane helices by the TCDB, it has been demonstrated that *M. jannaschi* NhaP1, a closer relative to NHE1, has thirteen (31). Additionally, the yeast exchanger Nhx1 also has an odd number of transmembrane helices (32). Comparison of the structure to NhaA shows that amino-terminal helix 1 is the additional transmembrane segment (33, 34). An alignment of the SLC9A subgroup with NhaP1 shows that NhaP1 transmembrane segment 1 aligns with transmembrane segment II of the SLC9A members and only NHE1 has a significantly predicted transmembrane segment before this (Figure 6.5). Therefore, it is plausible that the SLC9A members typically have thirteen transmembrane segments, with some possibly acting as signal peptides. NHE1, being the most recently evolved member of the subgroup, has added a fourteenth helix that contributes to NHE1 trafficking and residence on the plasma membrane (3, 5, 9) or as a novel sensor for cell volume regulation (10).

Final model

I have described above functional and structural reasoning for comparing the NhaA structure with the NHE1 homology model. This model predicts that NHE1 will have an NhaA-like fold, which is bolstered by a great deal of mounting evidence that there are likely a limited number of membrane protein folds (36–38). Recent examples of unrelated proteins sharing the same fold include Mhp1 (39) and vSGLT (40), a nucleobase:cation and a Na^+:galactose transporter, respectively, which have the same fold as the amino acid transporter LeuT (41). Similarly ASBT, a Na^+/bile acid symporter, has a remarkably similar fold to NhaA despite having fewer transmembrane segments and transporting a much larger substrate (42). This gives us more confidence that functionally related proteins share a common architecture. Several models of Na^+/H^+ exchangers based on the crystal structure of NhaA have now been published. Two models of human NHE1 transmembrane domain (2, 28), one of SLC9BA or NHA2 (43), one of NhaA in the periplasmic-facing conformation (44), and one of sod2 (*Chapter 4* (45)).

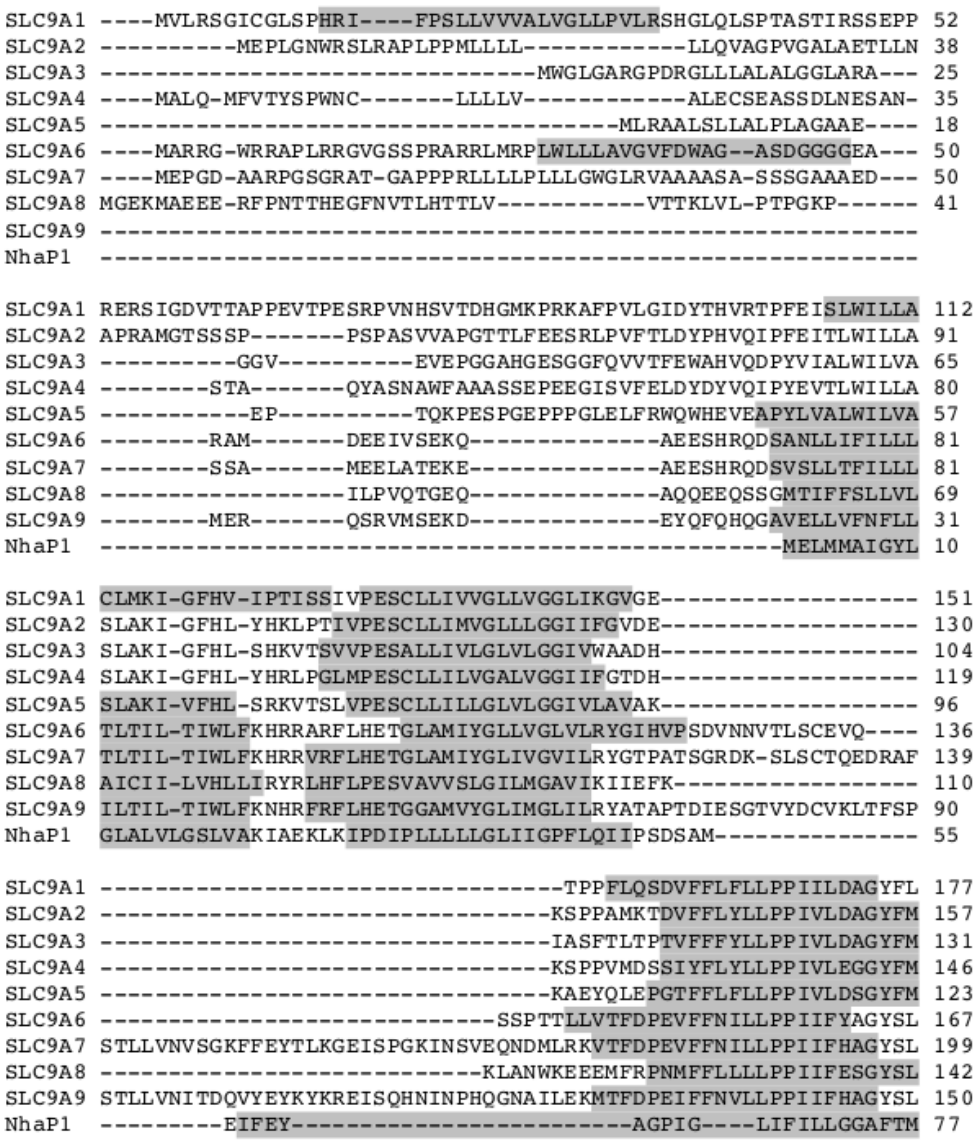

Figure 6.5: Sequence alignment of the amino-terminal region of *Methanococcus jannaschi* NhaP1 with the nine members of the SLC9A subfamily. Alignment was made using Clustal Omega (35). Transmembrane segment data for SLC9A and NhaP proteins were taken from the Uniprot database and are highlighted in grey.

To further justify both the Landau model and the new topology model proposed herein, I have fit the NhaA crystal structure and the NhaA projection structure in the molecular envelope of NHE1 (Figure 6.6). Using our interpretation of the *apical ridge* to be the cytoplasmic tail, I have oriented the envelope perpendicular to the membrane plane looking down into the cytoplasm. Using our interpretation of the membrane domain monomers at higher electron density, I located the approximate dimer interface. I then created a physiological NhaA dimer using the electron density from the NhaA dimer projection structure (46), the electron paramagnetic resonance measurements (47), and the high resolution structure of the NhaA monomer (4). I positioned and rotated the NhaA dimer so that it was in agreement with the NHE1 dimer interface (Figure 6.6-A). Importantly, this points the terminus of the final transmembrane segment of NhaA towards the NHE1 density representing the *apical ridge* (i.e. the beginning of the NHE1 cytoplasmic tail; Figure 6.6-B). With NhaA fit into the molecular envelope in this way there is significant unoccupied electron density. Figure 6.6-C shows the surface of NhaA, giving a better indication of the extra density. The location of the tops of each transmembrane helix in the Landau model have been superimposed on the surface labelled with Wakabayashi and Landau numbering. From this view, it is clear that the envelope could accommodate two more transmembrane segments at the amino terminus (indicated by the yellow circle).

Homology modelling and beyond: future directions for sod2

Very little information is available about the Na^+/H^+ exchanger, sod2. However, *S. pombe* and sod2 make a very interesting model with which to study eukaryotic exchangers. From a functional standpoint, sod2 appears to be a stable reporter for mutagenesis studies making it a useful tool to study salt tolerance in yeast and by extrapolation plants. Being in the CPA2 family of transporters, it may be a more suitable candidate for structural studies than human NHE1. At the very least, it offers a novel system to study the role of CPA2 exchangers in higher organisms and how they interrelate to CPA1 transporters. Indeed, any structure-function information that can be obtained from a unicellular eukaryotic Na^+/H^+ exchanger will have a positive impact on our understanding of how multicellular organisms regulate Na^+ and H^+ concentrations.

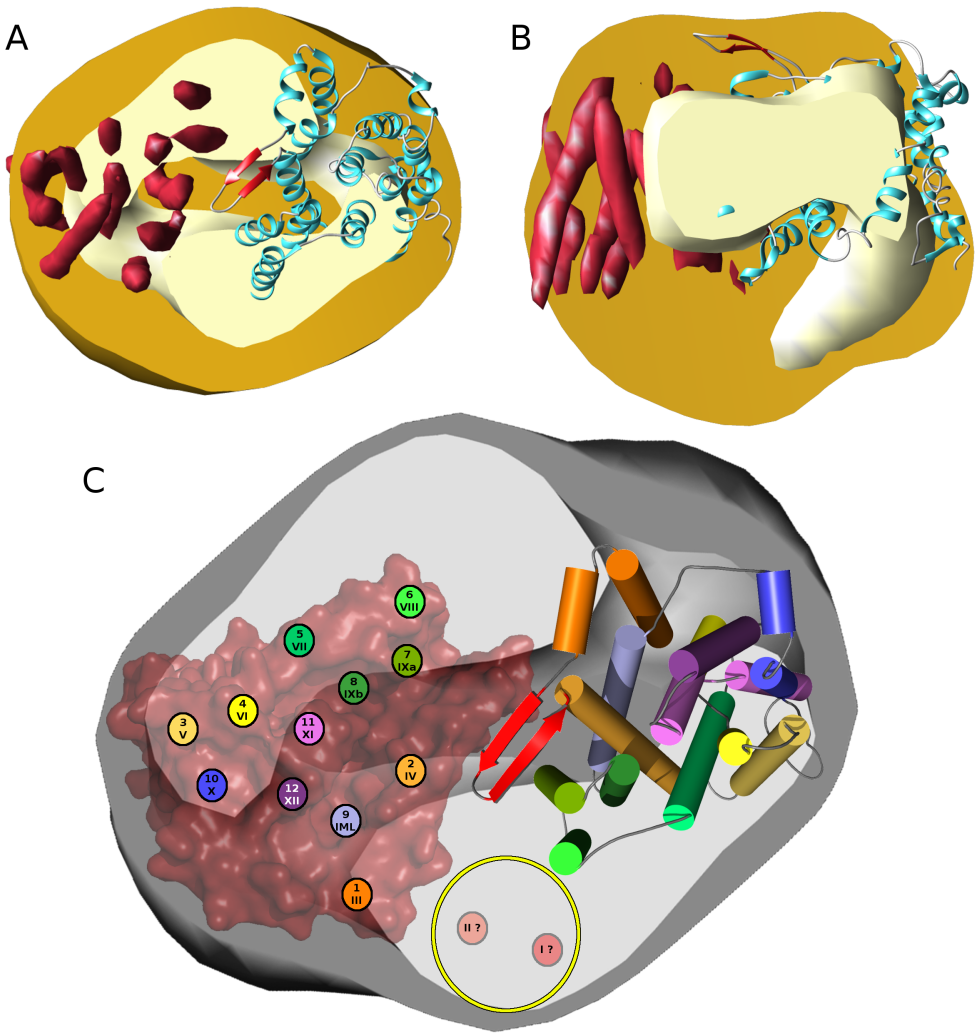

Figure 6.6: Detailed comparison of hNHE1 molecular envelope to NhaA. ***A***, The full volume of the molecular envelope of NHE1 (48), viewed from the extracellular space, shown cut in half (dark yellow) with a slab view of the envelope shown at higher electron density (light yellow). Superimposed on the left is the electron density of NhaA (49) and on the right a ribbon cartoon of the NhaA crystal structure (4). ***B***, An alternative view of ***A*** in the membrane plane. Here the envelope slab (light yellow) indicates density representing the the cytosolic regulatory domain (extending down). ***C***, A view similar to ***A*** with the molecular envelope now shown in shades of grey. Here NhaA monomers are shown as a surface (left, dark red) and as simple cartoon cylinders (multi-coloured, right, PDB:1ZCD (4)). Arranged on the NhaA surface are several circles representing the putative location and numbering of hNHE1 transmembrane helices based on the model by Landau *et al.* (2). The colouring matches the NhaA cylinders on the right and the segments are labelled with Arabic numerals (Landau *et al.*) and Roman numerals (Wakabayashi *et al.*, where IXa and b are a single helix and IML is an intramembrane loop in this model (1)). While both models have twelve transmembrane segments, the model by Landau *et al.* begins at amino acid 126 and excludes the first two segments predicted by Wakabayashi *et al.* Comparing this NhaA dimer orientation to the hNHE1 molecular envelope (in grey), there is significant *empty* density that contains no helices. By including the first two transmembrane segments predicted in the Wakabayahi *et al.* model (faded red circles I and II) and placing them next to segment III (large yellow ring), more of the molecular envelope is satisfied.

Final remarks

Na^+/H^+ exchangers represent an important family of membrane transport proteins present in all organisms. In humans, several isoforms carry out a variety of cellular functions through a diverse array of signalling events and dysregulation of NHE1 can lead to development of disease. Here I have presented my contributions towards understanding the structure of eukaryotic Na^+/H^+ exchangers. Chiefly, these include two optimized expression systems for studying eukaryotic exchangers *in vitro*, a low resolution molecular envelope of NHE1, and progress towards the high resolution structure of the transport bundle. From these small steps, it is evident that many questions remain about how NHE1 function and regulation are correlated to its structure and it remains as important as ever to determine the high resolution structure of a eukaryotic Na^+/H^+ exchanger. Only then will we be able to begin to elucidate mechanistic details that can bring potential insights into treating the dysregulation of NHE1 associated with heart disease and cancer.

References

1. Wakabayashi, S., Pang, T., Su, X. & Shigekawa, M. (2000). A novel topology model of the human Na^+/H^+ exchanger isoform 1. *J Biol Chem* **275**, 7942–7949.

2. Landau, M., Herz, K., Padan, E. & Ben-Tal, N. (2007). Model structure of the Na^+/H^+ exchanger 1 (NHE1): functional and clinical implications. *J Biol Chem* **282**, 37854–37863.

3. Wang, H., Singh, D. & Fliegel, L. (1998). Functional role of cysteine residues in the Na^+/H^+ exchanger effects of mutation of cysteine residues on targeting and activity of the Na^+/H^+ exchanger. *Arch. Biochem. Biophys.* **358**, 116–124.

4. Hunte, C. *et al.* (2005). Structure of a Na^+/H^+ antiporter and insights into mechanism of action and regulation by pH. *Nature* **435**, 1197–1202.

5. Shrode, L. D., Gan, B. S., D'Souza, S. J., Orlowski, J. & Grinstein, S. (1998). Topological analysis of NHE1, the ubiquitous Na^+/H^+ exchanger using chymotryptic cleavage. *Am J Physiol* **275**, C431–439.

6. Zizak, M. *et al.* (2000). Na^+/H^+ exchanger NHE3 has 11 membrane spanning domains and a cleaved signal peptide: topology analysis using in vitro transcription/translation. *Biochemistry* **39**, 8102–8112.

7. Miyazaki, E., Sakaguchi, M., Wakabayashi, S., Shigekawa, M. & Mihara, K. (2001). NHE6 protein possesses a signal peptide destined for endoplasmic reticulum membrane and localizes in secretory organelles of the cell. *J Biol Chem* **276**, 49221–49227.

8. Haworth, R. S., Fröhlich, O. & Fliegel, L. (1993). Multiple carbohydrate moieties on the Na^+/H^+ exchanger. *Biochem J* **289**, 637–640.

9. Counillon, L., Pouysségur, J. & Reithmeier, R. A. (1994). The Na^+/H^+ exchanger NHE-1 possesses N- and O-linked glycosylation restricted to the first N-terminal extracellular domain. *Biochemistry* **33**, 10463–10469.

10. Su, X., Pang, T., Wakabayashi, S. & Shigekawa, M. (2003). Evidence for involvement of the putative first extracellular loop in differential volume sensitivity of the Na^+/H^+ exchangers NHE1 and NHE2. *Biochemistry* **42**, 1086–1094.

11. Slepkov, E. R. *et al.* (2005). Structural and functional characterization of transmembrane segment IV of the NHE1 isoform of the Na^+/H^+ exchanger. *J Biol Chem* **280**, 17863–17872.

12. Herz, K., Rimon, A., Olkhova, E., Kozachkov, L. & Padan, E. (2010). Transmembrane segment II of NhaA Na^+/H^+ antiporter lines the cation passage, and asp65 is critical for pH activation of the antiporter. *J Biol Chem* **285**, 2211–2220.

13. de Planque, M. R. R. & Killian, J. A. (2003). Protein-lipid interactions studied with designed transmembrane peptides: role of hydrophobic matching and interfacial anchoring. *Mol Membr Biol* **20**, 271–284.

14. Holt, A. & Killian, J. A. (2010). Orientation and dynamics of transmembrane peptides: the power of simple models. *Eur Biophys J* **39**, 609–621.

15. Sato, Y., Ariyoshi, N., Mihara, K. & Sakaguchi, M. (2004). Topogenesis of NHE1: direct insertion of the membrane loop and sequestration of cryptic glycosylation and processing sites just after TM9. *Biochem Biophys Res Commun* **324**, 281–287.

16. Sato, Y. & Sakaguchi, M. (2005). Topogenic properties of transmembrane segments of *Arabidopsis thaliana* NHX1 reveal a common topology model of the Na^+/H^+ exchanger family. *J Biochem* **138**, 425–431.

17. Watanabe, Y., Shimono, Y., Tsuji, H. & Tamai, Y. (2002). Role of the glutamic and aspartic residues in Na^+-ATPase function in the ZrENA1 gene of *Zygosaccharomyces rouxii*. *FEMS microbiology letters* **209**, 39–43.

18. Gerchman, Y. *et al.* (1993). Histidine-226 is part of the pH sensor of NhaA, a Na^+/H^+ antiporter in *Escherichia coli*. *Proc Natl Acad Sci U S A* **90**, 1212–1216.

19. Reddy, T. *et al.* (2008). Structural and functional characterization of transmembrane segment IX of the NHE1 isoform of the Na^+/H^+ exchanger. *J Biol Chem* **283**, 22018–22030.

20. Slepkov, E. R., Rainey, J. K., Sykes, B. D. & Fliegel, L. (2007). Structural and functional analysis of the Na^+/H^+ exchanger. *Biochem J* **401**, 623–633.

21. Kemp, G., Young, H. S. & Fliegel, L. (2008). Structure and function of the human Na^+/H^+ exchanger isoform 1. *Channels (Austin)* **2**, 329–336.

22. Ding, J., Rainey, J. K., Xu, C., Sykes, B. D. & Fliegel, L. (2006). Structural and functional characterization of transmembrane segment VII of the Na^+/H^+ exchanger isoform 1. *J Biol Chem* **281**, 29817–29829.

23. Lee, B. L., Li, X., Liu, Y., Sykes, B. D. & Fliegel, L. (2009). Structural and functional analysis of transmembrane XI of the NHE1 isoform of the Na^+/H^+ exchanger. *J Biol Chem* **284**, 11546–11556.

24. Tzeng, J., Lee, B. L., Sykes, B. D. & Fliegel, L. (2010). Structural and functional analysis of transmembrane segment VI of the NHE1 isoform of the Na^+/H^+ exchanger. *J Biol Chem* **285**, 36656–36665.

25. Galili, L., Rothman, A., Kozachkov, L., Rimon, A. & Padan, E. (2002). Transmembrane domain IV is involved in ion transport activity and pH regulation of the NhaA Na^+/H^+ antiporter of *Escherichia coli*. *Biochemistry* **41**, 609–617.

26. Slepkov, E. R., Chow, S., Lemieux, M. J. & Fliegel, L. (2004). Proline residues in transmembrane segment IV are critical for activity, expression and targeting of the Na^+/H^+ exchanger isoform 1. *Biochem J* **379**, 31–38.

27. Tzeng, J., Lee, B. L., Sykes, B. D. & Fliegel, L. (2011). Structural and functional analysis of critical amino acids in TMVI of the NHE1 isoform of the Na^+/H^+ exchanger. *Biochim Biophys Acta, Biomembr* **1808**, 2327–2335.

28. Nygaard, E. B. *et al.* (2011). Structural modeling and electron paramagnetic resonance spectroscopy of the human Na^+/H^+ exchanger isoform 1, NHE1. *J Biol Chem* **286**, 634–648.

29. Pedersen, S. F., King, S. A., Nygaard, E. B., Rigor, R. R. & Cala, P. M. (2007). NHE1 inhibition by amiloride- and benzoylguanidine-type compounds. inhibitor binding loci deduced from chimeras of NHE1 homologues with endogenous differences in inhibitor sensitivity. *J Biol Chem* **282**, 19716–19727.

30. Schushan, M., Landau, M., Padan, E. & Ben-Tal, N. (2011). Two conflicting NHE1 model structures: compatibility with experimental data and implications for the transport mechanism. *J Biol Chem* **286**, le9; author reply Ie10.

31. Vinothkumar, K. R., Smits, S. H. J. & Kühlbrandt, W. (2005). pH-induced structural change in a sodium/proton antiporter from *Methanococcus jannaschii*. *EMBO J* **24**, 2720–2729.

32. Wells, K. M. & Rao, R. (2001). The yeast Na^+/H^+ exchanger Nhx1 is an N-linked glycoprotein. topological implications. *J Biol Chem* **276**, 3401–3407.

33. Goswami, P. *et al.* (2011). Structure of the archaeal Na^+/H^+ antiporter NhaP1 and functional role of transmembrane helix 1. *EMBO J* **30**, 439–449.

34. Lee, B. L., Sykes, B. D. & Fliegel, L. (2013). Structural and functional insights into the cardiac Na^+/H^+ exchanger. *J Mol Cell Cardiol* **61**, 60–67.

35. Sievers, F. *et al.* (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**.

36. Oberai, A., Ihm, Y., Kim, S. & Bowie, J. U. (2006). A limited universe of membrane protein families and folds. *Protein Sci* **15**, 1723–1734.

37. Screpanti, E. & Hunte, C. (2007). Discontinuous membrane helices in transport proteins and their correlation with function. *J Struct Biol* **159**, 261–267.

38. Shi, Y. (2013). Common folds and transport mechanisms of secondary active transporters. *Annu Rev Biophys* **42**, 51–72.

39. Weyand, S. *et al.* (2008). Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter. *Science* **322**, 709–713.

40. Faham, S. *et al.* (2008). The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na^+/sugar symport. *Science* **321**, 810–814.

41. Yamashita, A., Singh, S. K., Kawate, T., Jin, Y. & Gouaux, E. (2005). Crystal structure of a bacterial homologue of Na^+/Cl^--dependent neurotransmitter transporters. *Nature* **437**, 215–223.

42. Hu, N.-J., Iwata, S., Cameron, A. D. & Drew, D. (2011). Crystal structure of a bacterial homologue of the bile acid sodium symporter ASBT. *Nature* **478**, 408–411.

43. Schushan, M. *et al.* (2010). Model-guided mutagenesis drives functional studies of human NHA2, implicated in hypertension. *J Mol Biol* **396**, 1181–1196.

44. Schushan, M. *et al.* (2012). A model-structure of a periplasm-facing state of the NhaA antiporter suggests the molecular underpinnings of pH-induced conformational changes. *J Biol Chem* **287**, 18249–18261.

45. Ullah, A. *et al.* (2013). Structural and functional analysis of transmembrane segment IV of the salt tolerance protein sod2. *J Biol Chem* **288**, 24609–24624.

46. Williams, K. A. (2000). Three-dimensional structure of the ion-coupled transport protein NhaA. *Nature* **403**, 112–115.

47. Hilger, D., Polyhach, Y., Padan, E., Jung, H. & Jeschke, G. (2007). High-resolution structure of a Na^+/H^+ antiporter dimer obtained by pulsed electron paramagnetic resonance distance measurements. *Biophys J* **93**, 3675–3683.

48. Moncoq, K., Kemp, G., Li, X., Fliegel, L. & Young, H. S. (2008). Dimeric structure of human Na^+/H^+ exchanger isoform 1 overproduced in *Saccharomyces cerevisiae*. *J Biol Chem* **283**, 4145–4154.

49. Williams, K. A., Geldmacher-Kaufer, U., Padan, E., Schuldiner, S. & Kühlbrandt, W. (1999). Projection structure of NhaA, a secondary transporter from *Escherichia coli*, at 4.0 å resolution. *EMBO J* **18**, 3558–3563.