Structure and Function of Sodium-Proton Antiporters

Chiara Natasha Lee

Division of Molecular Bioscience

Imperial College London

Thesis submitted for the Degree of Doctor of Philosophy

2013

Abstract

Sodium-proton (Na^+/H^+) antiporters are secondary membrane protein transporters present in all living cells and are critical for sodium, pH and cell volume homeostasis. Their deregulation of transport activity has been linked to human diseases, such as, hypertension, heart failure and epilepsy and consequently they may be targets for drugs. In 2005, the first crystal structure of a Na^+/H^+ antiporter, NhaA from *Escherichia coli*, was solved at 3.45 Å resolution in an inward-facing conformation at pH 4 where the protein is inactive. Like many Na^+/H^+ antiporters, the activity of NhaA is regulated by pH and is active above pH 6.5.

The main goal of this thesis was to solve a crystal structure of a Na^+ /H^+ antiporter in an active state. One of the problems in producing milligram quantities of purified membrane protein for crystallography is poor overexpression. For this reason, we first sought to improve membrane protein overexpression by developing a new expression platform, which we have called "MemStar" in *E. coli* using a test-case of control proteins, overall showing a boost in expression levels to at least 12 $mg.L^{-1}$.

This thesis describes the crystal structure of NapA from *Thermus thermophilus*, an NhaA homologue, which was solved to 3 Å in an outward-facing conformation at pH 7.8 in an active state. This NhaA homologue was selected as purified protein could grow better diffracting crystals than NhaA. The stability of NapA was also more suitable for purification in a small micelle detergent to improve diffracting resolution. Although NhaA crystals did not form above pH 6.5, a stabilised mutant was useful to confirm the position of a critical residue important in the mechanism.

Structural comparisons with the NapA structure show the core domain moving relative to the dimerisation domain, similar to a rocking bundle model observed in other structures of different secondary active transporters sharing conserved structural features in their membrane protein folds also present in NhaA and NapA. This work has provided us with a fresh insight into the mechanism of Na^+/H^+ antiporters.

Declaration

I declare that the work presented in this thesis has been carried out by myself, unless otherwise stated, and that all sources of information have been specifically acknowledged.

Acknowledgements

To my work colleagues;

First of all I would like to thank So Iwata, who was the first person to really introduce me into the world of membrane proteins.

I would like to thank my supervisor, David, for his constant support, guidance, valued advice and, above all, friendship throughout my PhD. My thanks unquestionably extends to my co-supervisor, Alex, who kept my spirits up in those rare but dark hours, such as on a beamline that is down at 3am or when better high resolution crystals are still needed.

A big thank you must go to all the past and present MPC members at Imperial for making me enjoy coming into the lab every day. I particularly would like to thank Hae Joo, who I have really enjoyed working next to. To Sam, Manami and Mutsuko, I owe you all a big thank you for helping me with the tasks I cannot stand; such as orders and expense claim forms! I would like to thank everyone from MPL at Diamond for making me feel very welcome. My thanks in particular to Nien-Jen, Indran and Matt who amazingly were always happy, no matter how many times, to check I was running the Shimadzu correctly.

I must also thank all the members of Jan Willem de Gier's team at Stockholm University; I really enjoyed my short but fun and productive time spent with you all.

With regards to the work of this thesis, I would like to thank Shoko Yashiro for providing me with the NhaA thermostable mutant, Oliver Beckstein for collaborative MD simulations, Christoph von Ballmoos for the NapA transport assays, Anna Hjelm for helping me with flow cytometry analysis and Hassanul Choudhury for purifying and crystallising TehA.

To my family;

I have to thank my brother, Joshua, whose superior knowledge of computers had come to my rescue on many occasions as well as your endless humour!

I owe so much to my grandmother, who has been on this three year journey as much as I have and I could not have done without her. You have always been there for me.

To my father, who has given me such enormous support in every way throughout my life.

This thesis is dedicated to my late grandfather,

Dr. Desmond Mulcahy MD (Cantab), FRCP,

who guided me throughout my education.

Table of Contents

Abbreviations

β-OG	β-octyl-glucoside
CMC	Critical micelle concentration
CPA	Cation proton antiporter family
CPM	7-diethylamino-3-(4'-maleimidylphenyl)-4-methylcoumarin
CV	Column volume
DDM	*n*-dodecyl β-D-maltoside
EM	Electron microscopy
FACS	Fluorescence-activated cell sorting
FSEC	Fluorescent size exclusion chromatography analysis
GFP	Green fluorescent protein
GPCR	G-coupled protein receptor
HA	Heavy atom
IMAC	Immobilised metal affinity chromatography
IMP	Integral membrane protein
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Luria Broth
LDAO	*N*,*N*-Dimethyldodecylamine *N*-oxide
MBP	Maltose binding protein
MFS	Major facilitator superfamily
MIRAS	Molecular isomorphous replacement and anomalous scattering
MR	Molecular replacement
NCS	Non-crystallographic symmetry
NM	Nonyl-maltoside
PDB	Protein Data Bank
PMF	Proton motive force
RFU	Relative fluorescence unit
RNC	Ribosome nascent chain
SAD	Single wavelength anomalous dispersion
SEC	Size exclusion chromatography
SEC-LS	Size exclusion chromatography-light scattering
SeMet	Selenomethionine
SLC	Solute carrier transporter family
SMF	Sodium motive force
SRP	Signal recognition particle
$T_{1/2}$	Half-life
TEV	Tobacco Etch Virus
TM	Transmembrane

Figure List

Table List

Thesis Objectives

The main objective of this thesis was to understand how sodium-proton antiporters function at the atomic level. X-ray crystallography was the method of choice for structural understanding and this information was combined with collaborative work from Dr. Christoph von Ballmoos (transport assays) and Dr. Oliver Beckstein (molecular dynamic simulations).

Chapter 3 first describes an expression protocol that was developed to boost membrane protein production yields in *E. coli* to facilitate structural studies. This general expression protocol, MemStar, was later applied to structural studies of sodium-proton antiporters, as outlined in Chapters 4 and 5.

Chapter 4 focuses on the bacterial sodium-proton antiporter NhaA from *E. coli*, which is the best characterised model system for sodium-proton antiport with both EM and crystal structures available. However, as only one crystal structure was solved at low pH where the protein is inactive, this has made it difficult to put together a transport mechanism. The practical objective was to try and improve the stability of NhaA to enable the growth of well ordered crystals above pH 6.5, where the protein is active. A random mutagenesis strategy was employed that improved the thermostability of NhaA. Although we did not manage to obtain an active state, nonetheless this approach did further improve the efficiency of growing well ordered crystals at low pH. This became an important issue as we uncovered that there had been an incorrect assignment of helix 10 on the original structure.

Chapter 5 describes the crystallographic work on NapA from *Thermus thermophilus*, a bacterial NhaA homologue, for which we were able to obtain well diffracting crystals grown in pH conditions where the protein is active.

Structural comparisons were made against the inward-facing NhaA structure. Chapter 6 gives a final discussion of what we have learnt from these structures with regards to the transport cycle of Na^+/H^+ antiport.

1.0 -

Introduction

Introduction

All living cells must exist in an isolated non-equilibrium system, which is maintained by a plasma membrane acting as a barrier to the external environment. The basic architecture of a membrane is a lipid bilayer composed from phospholipids. Each phospholipid contains two non-polar hydrophobic tails composed of acyl carbon chains situated in the interior with a polar hydrophilic phosphate head group facing the external aqueous environment. This arrangement creates a hydrophobic interior of the membrane that is selectively permeable to small lipophilic solutes or uncharged molecules (such as oxygen or carbon dioxide), but retards the diffusion of hydrophilic solutes and ions, which are essential for maintaining cellular homeostasis and thus cell survival. Facilitated transport of these polar molecules and ions across the selectively permeable membrane is carried out by integral membrane proteins (IMPs), which can span the membrane to provide a continuous protein-lined pathway through the lipid bilayer.

1.1- Integral membrane proteins

There are two major structural folds of IMPs present in bacteria and eukaryotes, namely the β-barrel or the α-helical bundle (figure 1.1). The β-barrel fold consists of a series of β-strands which hydrogen bond with neighbouring strands forming a closed pore-like structure (Schulz, 2000). This β-barrel structural fold is most common in the porin family of gram-negative bacteria for the transport of hydrophobic and hydrophilic molecules and extrusion of drugs by passive diffusion across the outer membrane (Nikaido, 2003; Hong *et al*., 2006). β-barrel IMPs are also found in the outer membranes of mitochondria in eukaryotes, reflecting their bacterial

origin. Here they are required for the transport of hydrophilic metabolites across the mitochondria outer membrane to and from the cytoplasm (Ujwal *et al*., 2008; Walther and Rapaport, 2009).

The majority of membrane proteins are made up of α-helices with an average length of 25 amino acids and sufficient hydrophobicity to partition into the lipid bilayer and, in most cases, cross the membrane. These α-helical stretches are often referred to as transmembrane (TM) segments. They interact and associate with other TM segments, lipid and the bulk solvent to create many different types of structures, e.g., ion-channels, G-coupled protein receptors (GPCRs) and transporters (Alberts, 2008). α-helical IMPs are mostly present in the inner membrane of bacterial cells and plasma membrane of eukaryotes and make up 20-30 % of all sequenced genomes to date (Wallin and von Heijne, 1998). Human α-helical membrane proteins also represent 40% of all pharmaceutical drug targets confirming their level of importance in all living cells (Overington *et al*., 2006). Henceforth, I use the term membrane proteins to denote α-helical membrane proteins.

(a) **(b)**

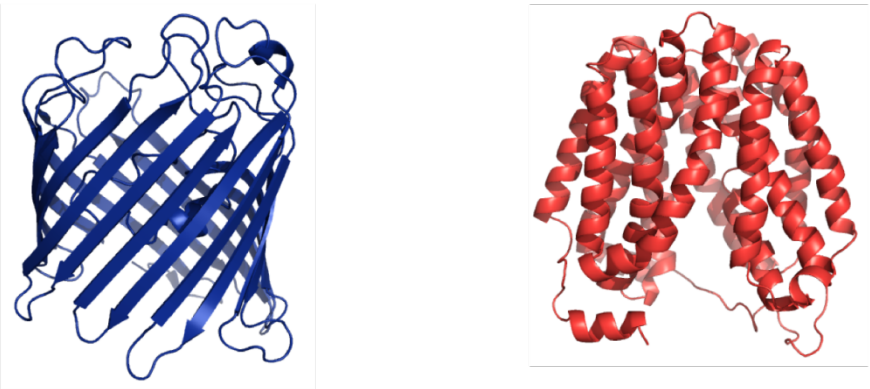


Figure 1.1 – Illustration of two types of integral membrane proteins. (a) An example of a β-barrel membrane protein (PDB: 1PHO, (Cowan *et al*., 1992)) and (b) an example of an α-helical membrane protein (PDB: 1PV6, (Abramson *et al*., 2003)). This and subsequent molecular structures are generated using Pymol (DeLano, 2002).

1.2 - Membrane protein biogenesis

Membrane proteins need to be transcribed, translated and translocated into the membrane bilayer. This is a multi-step process that must be managed to avoid protein misfolding and subsequent aggregation in the cytoplasm (figure 1.2) (Dalbey *et al*., 2011). In *Escherichia coli,* membrane protein biogenesis begins when a hydrophobic α-helical segment is exposed from the exit tunnel of the ribosome and is recognised by the signal recognition particle (SRP) (Batey *et al*., 2000). SRP binding targets the ribosome nascent chain (RNC) complex to the inner membrane and docks onto the membrane bound SRP receptor, FtsY (Luirink and Sinning, 2004; Dalbey *et al*., 2011). In eukaryotic cells, translation is halted when the SRP-RNC complex is formed and does not resume until it makes contact with the Sec61 translocon. Sufficiently hydrophilic α-helical TM segments pass laterally through the Sec translocon, whereas more polar ones pass through, as is the case for signal sequences

of secretory proteins. For membrane proteins the first TM segment is the signal sequence, which is generally not cleaved. -helix partitioning into the membrane is mostly driven by favourable helix-lipid interactions (Hessa *et al.*, 2007). In some cases very polar TM segments cannot insert by themselves, but may either be pulled into the membrane by neighbouring TMs, which are much more hydrophobic (Kauko *et al.*, 2010) or by a cluster of downstream charged residues in keeping with the positive-inside rule where positively charged residues face the cytoplasmic side of the membrane (Heijne, 1986). Membrane protein topology predictions in *E. coli* and *Saccharomyces cerevisiae* show 80% of -helical membrane proteins to have the C-terminus facing the cytoplasm (C_{IN}) (Rapp *et al.*, 2004; Kim *et al.*, 2006).

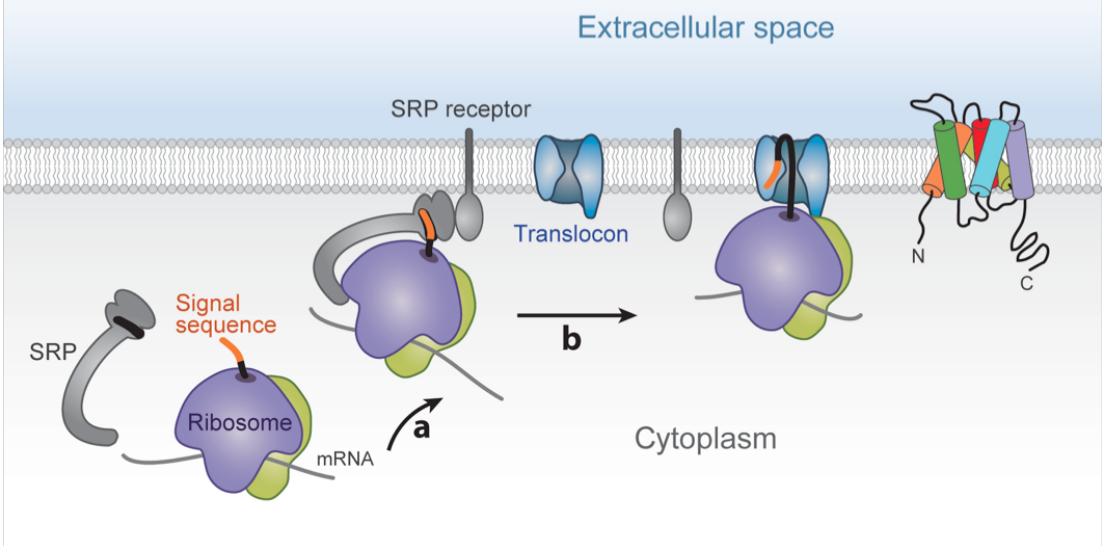


Figure 1.2 – Illustration of membrane protein biogenesis. Membrane proteins fold co-translationally. After the signal sequence (shown in orange) exits the ribosome it is recognised by the SRP (step a). Binding of SRP halts further translocation in eukaryotes and the SRP-ribosome nascent chain complex is targeted to the SRP receptor, which is closely associated with the Sec translocon (step b). Translation resumes and TM segments pass laterally through an opening in the Sec translocon and into the membrane. This image is taken with permission from (Kang *et al.*, 2013).

1.3 - Classes of membrane transport proteins

Membrane transport proteins are essential for facilitated transport of hydrophilic molecules and ions that cannot otherwise cross the membrane passively. This is due to either their hydrophilic nature and/or because they are moving against a concentration gradient. There are two major classes of membrane transport proteins, namely channels and transporters.

Channels are understood to simply act as a selective "passageway" through the membrane. Ion channels carry out passive transport by opening a pore through the middle of the protein for ions to diffuse down their concentration gradient in keeping with the Nernst potential (Gouaux and Mackinnon, 2005). Channel pores act as selective filters that will only allow certain types of ions to pass through depending on their size, shape, and charge (Zhou *et al*., 2001). Different types of stimuli can also modulate a channel to open and close, such as ligands (neurotransmitters) or changes in membrane potential.

Membrane transporters, in contrast, specifically bind to a given molecule (substrate) and physically undergo a "bind and release" conformational change to transport the substrate across the membrane. Active transport requires an energy source to transport molecules and ions against their concentration gradients across the membrane, which is critical for all living cells, as they exist in a non-equilibrium state. Primary active transporters use a direct energy source to drive conformational changes to transport molecules across the membrane, such as ATP hydrolysis. For example the sodium/potassium (Na^+/K^+) ATPase exchanges $3Na^+$ (out) for $2K^+$ (in) to maintain the electrochemical potential used to drive action potentials and consumes 20% of ATP

production at rest (Silver and Erecinska, 1997). Secondary active transporters utilise the energy stored in ion gradients. Movement of ions down their electrochemical gradients is used to transport substrates against their concentration gradient across the membrane (figure 1.3). Membrane symporters transport the ion and substrate in the same direction across the membrane and membrane antiporters transport the components in opposite directions. Membrane uniporters share similar kinetic characteristics to symporters and antiporters, but instead undergo passive transport of one substrate species down its concentration gradient, as shown from the family of glucose transporters (GLUT family), which are present on the plasma membranes of human erythrocytes for concentration dependent glucose uptake (Gould and Holman, 1993).

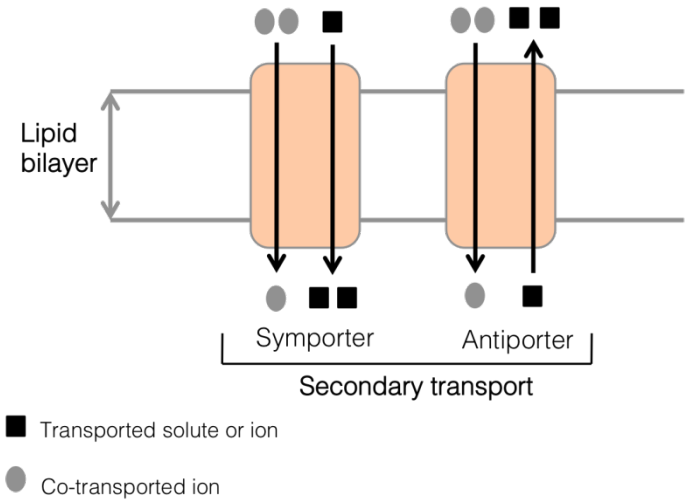


Figure 1.3 – Schematic illustration of two types of secondary active transporters. From left to right, a membrane symporter simultaneously transports its substrate across the membrane in the same direction as the electrochemical gradient of its co-transported ion used as an energy source. A membrane antiporter transports its substrate across the membrane in the opposite direction of the electrochemical gradient of the counter-ion.

Secondary active transporters have been proposed to function by an alternating access mechanism, involving structural movements to take place during transport to form an ‘outward-facing' state when the substrate binding site is accessible and open to the external medium and an ‘inward-facing' state where the protein is open towards the cytoplasm (Jardetzky, 1966). In the last ten years, a number of high resolution crystal structures from bacterial homologues of different Na^+- and H^+-coupled transporters have been solved in the inward and outward states as well as intermediate states where the substrate binding site is not accessible to bulk solvent on either side of the membrane; referred as the occluded state (figure 1.4).

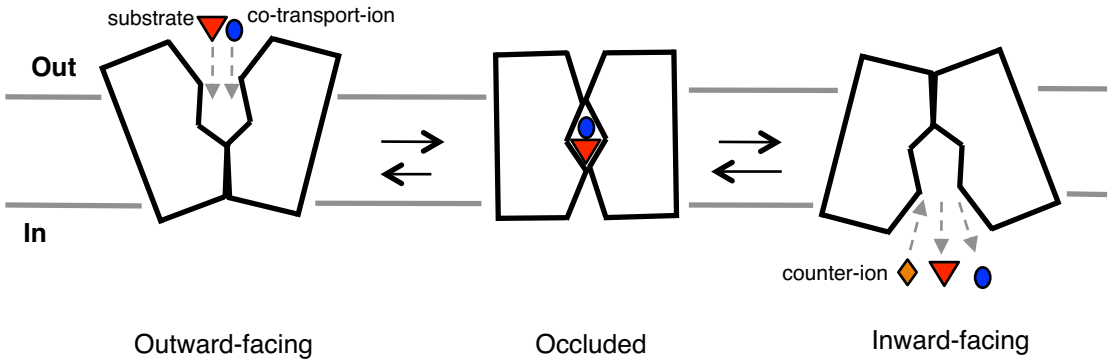


Figure 1.4 – Schematic illustration of the alternating access mechanism. From left to right, an outward-facing conformation is present for the substrate and co-transported ion to enter a central binding site from the extracellular space. An occluded state is followed with conformational changes taking place to first seal off access of the binding site to both sides of the membrane. The protein switches to the inward-facing conformation with the binding site solvent exposed to the cytoplasm to release the substrate and co-transported ion. With membrane antiporters, the counter-ion is co-transported in the opposite direction to the substrate.

The solute carrier transporter (SLC) superfamily comprises over 300 identified secondary active transporters, classified into 55 SLC gene families based on their sequence and function in the co-transport of essential nutrients, such as sugars, amino acids, nucleosides, neurotransmitters, cofactors (vitamins and metal ions), drugs and ions (Saier *et al*., 2006). In the last ten years a cumulative increase in the number of high resolution crystal structures of secondary transporters has helped to narrow the gap between the concepts of secondary transport and understanding how the alternating access mechanism is established at the molecular level.

Different structural themes are also emerging as more structures are being solved from secondary transporters in different SLC families. Firstly, a number of structures with diverse functions have been observed to show similarities in their structural fold. This especially true for the LeuT fold superfamily containing multiple structures solved in the same/different conformations from 7 different SLC families, all classified to share the LeuT fold (Yamashita *et al*., 2005; Faham *et al*., 2008; Weyand *et al*., 2008; Fang *et al*., 2009; Ressl *et al*., 2009; Shaffer *et al*., 2009; Tang *et al*., 2010). Secondly, all secondary transporter structures solved to date show topology-repeat units that can form separate structural domains parallel to the membrane, as observed with transporters in the major facilitator superfamily (MFS) (Abramson *et al*., 2003; Huang *et al*., 2003) or inverted domains related by pseudosymmetry perpendicular to the membrane, as observed with transporters of the LeuT, Glt_{Ph} (Yernool *et al*., 2004) and NhaA (Hunte *et al*., 2005) folds as well as within the separate domains of MFS transporters (Radestock and Forrest, 2011). The pseudosymmetry observed between these repeat units has been suggested to facilitate switching between the outward and inward-facing states. Through structural

comparison of the same or similar proteins solved in multiple conformations, crystal structures play a significant part in helping us to better understand the molecular basis of the alternating access mechanism. A common structural theme seen among a number of secondary transporters is the movement of two structural domains that rock against one another to promote alternate access of a central substrate binding site between the inward and outward states (Forrest *et al*., 2011).

This thesis focuses on the SLC9 Sodium-proton (Na^+/H^+) exchanger family.

2.0 –

Sodium-Proton Antiporter family

2.1 - Physiological roles of sodium-proton antiporters

The presence of electrochemical gradients of sodium ions (Na^+) and protons (H^+) across the cell membrane is critical for all living cells to provide an energy source for secondary active transport of cellular solute and ion uptake as well as ATP synthesis (Skulachev, 1991). Sodium-proton (Na^+/H^+) antiporters are essential for regulating sodium and pH homeostasis and are subsequently driven by these electrochemical gradients depending on the cell type (figure 2.1). Consequently their deregulation can cause the intracellular concentration of both ions to become potent, leading to cellular stress and eventually cell death.

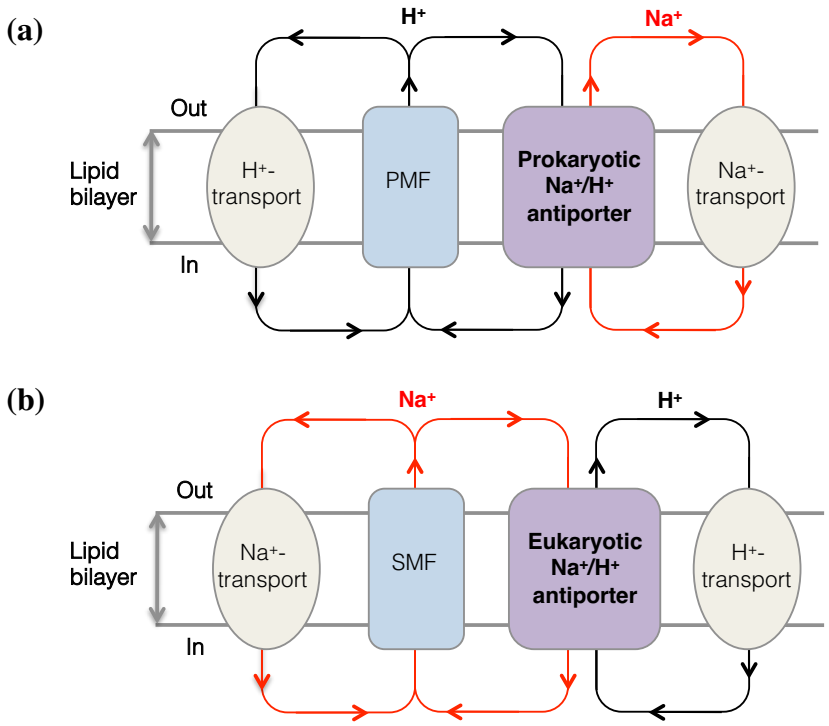


Figure 2.1 – Schematic illustration of Na^+/H^+ antiport activity and movement of H^+ and Na^+ ions across the cell membrane. Na^+/H^+ antiporters play an important role in cell homeostasis in all living cells by converting a generated proton motive force (PMF) into a sodium motive force (SMF) and vice versa to maintain the function of H^+-coupled and Na^+-coupled transporters for the transport of essential solutes and ions across cellular and organelle membranes. (a) Prokaryotic Na^+/H^+ antiporters are driven by a PMF generated from oxidative phosphorylation. (b) Eukaryotic Na^+/H^+ antiporters are mostly driven by a SMF generated from Na^+/K^+ ATPase activity. This figure is adapted from (Hase *et al.*, 2001).

The transport activity of Na^+/H^+ antiporters was first discovered by Peter Mitchell and Ian West in 1974, who showed in *E. coli* that after the addition of Na^+ ions to the bacterial cells, an efflux of Na^+ ions is driven by the proton motive force (PMF) (West and Mitchell, 1974). Since then, their transport activity has been identified in the cell membranes of bacteria, yeast, plants and animals as well as organelle membranes, such as mitochondria and plant vacuoles (Brett *et al*., 2005). Although the antiport exchange of Na^+ and H^+ ions is conserved across all species, their physiological roles show more variation depending on the cell type and surrounding environment.

Bacterial Na^+/H^+ antiporters

In bacteria, Na^+/H^+ antiporters present on the inner membrane play central roles in the regulation of internal Na^+ and H^+ concentrations necessary for cell survival, by exporting excess intracellular Na^+ ions using the PMF generated by oxidative phosphorylation (figure 2.1) (Padan and Schuldiner, 1994). Their function has been shown to be particularly important in preventing cellular alkalidosis under environmental conditions of extreme salt stress and high pH. Bacterial Na^+/H^+ antiporters are also shown to have an important constructive role in using the free energy from the PMF to generate a sodium motive force (SMF) to drive Na^+-coupled secondary transporters for the uptake of sugars, amino acids, nucleosides and ions as well as drive Na^+ -coupled drug efflux systems (Padan and Schuldiner, 1994; Padan *et al*., 2005). *E. coli* NhaA has been the most functionally characterised bacterial Na^+/H^+ antiporter and carries out electrogenic transport in a $2H^+$:$1Na^+$ ratio, also observed from other bacterial NhaA homologues (Padan *et al*., 1989; Inoue *et al*., 1999; Furrer *et al*., 2007). *E. coli* NhaA was also the first Na^+/H^+ antiporter to be solved by X-ray crystallography (Hunte *et al*., 2005) and therefore will be further discussed.

Mammalian Na^+/H^+ antiporters

Mammalian Na^+/H^+ antiporters (often referred to as NHEs) consist of a family of gene-related isoforms (NHE1-9), sharing ~25-70% amino acid identity overall, but differing in their physiological functions, depending on their location in the plasma and/or intracellular membrane and cell type (Orlowski and Grinstein, 2004).

Plasma-membrane type NHEs are important for regulating intracellular pH, cell volume and preventing cellular acidification by exporting excess H^+ ions from the cytoplasm using the SMF generated by Na^+/K^+ ATPase (figure 2.1) (Demaurex and Grinstein, 1994; Counillon and Pouyssegur, 2000). By controlling cell volume and maintaining cellular and organellar pH, mammalian NHEs have been proposed to play pivotal roles in numerous physiological processes, including transepithelial ion transport, cell proliferation and apoptosis, cell migration and vesicle trafficking (Orlowski and Grinstein, 2004). Consequently their loss of regulated activity has been shown in different human diseases, such as cancer, diabetes, cardiovascular disease, hypertension and organ ischemia (Putney *et al*., 2002; Bobulescu *et al*., 2005). Human NHE1 in particular has been the most extensively studied isoform expressed on all cell types and regulates cell proliferation and apoptosis (Slepkov *et al*., 2007), making it a major drug target in human diseases, such as cardiovascular disease and breast cancer (Lee *et al*., 2012; Amith and Fliegel, 2013). In contrast, the transport activity of the NHE3 isoform expressed on the apical surface of intestinal epithelial cells is critical for generating a PMF, driven by a SMF generated from the basolateral Na^+/K^+ ATPase, for the absorption amino acids, oligopeptides, metal ions and drug compounds by apical H^+-coupled secondary transporters (Thwaites and Anderson, 2007). This dual ion transport activity network between the apical NHE3 and

basolateral Na^+/K^+ ATPase is also present in the kidneys for the reabsorption of filtered solutes and ions by H^+-coupled transporters (Bobulescu and Moe, 2009). The NHE6 isoform located on the inner mitochondrial membrane also uses a PMF generated by oxidative phosphorylation to export excess matrix Na^+ ions to regulate the mitochondrial volume and pH (Numata *et al*., 1998). NHE6 is also involved in maintaining mitochondrial calcium homeostasis, which is very important for metabolic regulation, with excess Ca^{2+} ions exported by Na^+/Ca^{2+} antiporters utilising a SMF generated from the NHE6 activity (Bernardi, 1999). Unlike their bacterial orthologues, mammalian NHEs are understood to carry out electroneutral antiport exchange ($1H^+$:$1Na^+$)(Demaurex *et al*., 1995).

Plant Na^+/H^+ antiporters

Plant vacuolar Na^+/H^+ antiporters (NHX) are very important for plant agriculture to prevent salt stress and confer halotolerance by compartmentation of excess Na^+ from the cytoplasm into vacuoles using a PMF generated by the vacuolar H^+-ATPase (Rodriguez-Rosales *et al*., 2009). Subsequently, transgenic Arabidposis plants overexpressing the NHX1 gene from *Arabidopsis thaliana* have enhanced salt tolerance (Apse *et al*., 1999). The SOS1 family is also another important class of plant Na^+/H^+ antiporters located on the plasma membrane for sodium extrusion (Nunez-Ramirez *et al*., 2012).

2.2 - Cation proton antiporter superfamily

All sequenced Na^+/H^+ antiporter genes have been classified into the monovalent cation proton antiporter (CPA) superfamily; containing other secondary transporters involved in cation regulation and pH homeostasis by the exchange of Na^+, Li^+ or K^+ for H^+ ions (Brett *et al*., 2005; Saier *et al*., 2006). Based on phylogenetic analysis, Na^+/H^+ antiporters from bacteria, plants and humans in the CPA superfamily are divided into two main branches, the CPA1 and CPA2 families; shown in figure 2.2 (Brett *et al*., 2005).

CPA1 family

The CPA1 family evolved from bacterial NhaP genes, which encode for a family of Na^+/H^+ antiporters that carry out pH-dependent electroneutral transport of Na^+ or Li^+ ions using the PMF (Hellmer *et al*., 2002). These ancestral bacterial NhaP genes are shown to have evolved and diverge into two gene clusters; the NhaP-SOS1 clade, containing multiple paralogues of the bacterial NhaP family sharing homology with the plant SOS1 Na^+/H^+ antiporter family and the eukaryotic NHE clade containing the mammalian NHE isoforms and plant vacuolar NHX family (Brett *et al*., 2005).

Sequence homology between the mammalian NHEs also reflects their subcellular localisation in the plasma membrane (isoforms 1-5) and intracellular membrane (isoforms 6-9) sharing 25–70% amino acid identity overall (Orlowski and Grinstein, 2007). All Na^+/H^+ antiporters in the CPA1 family are understood to carry out electroneutral exchange according to the transport classification database (http://www.tcdb.org/)(Saier *et al*., 2006).

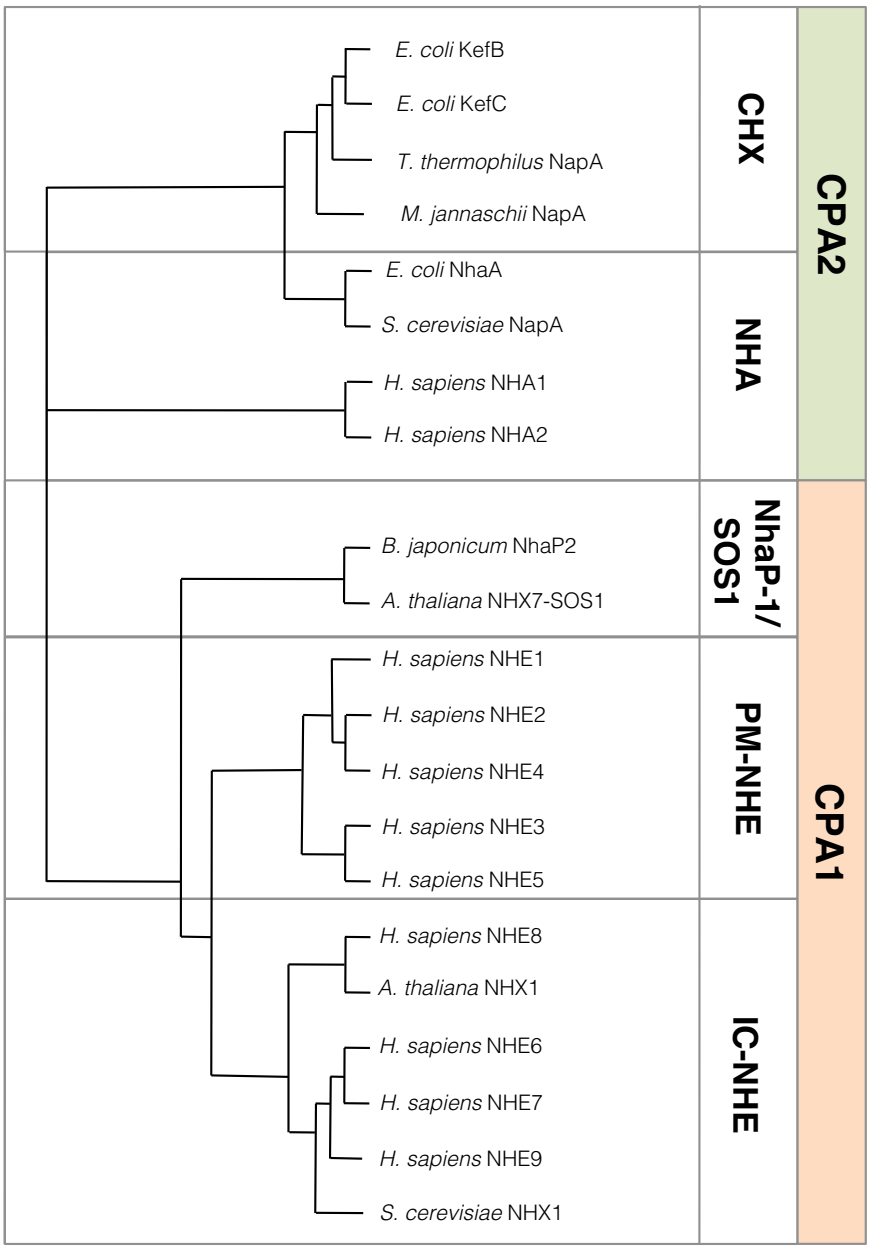


Figure 2.2 – Phylogenetic distribution of bacterial and eukaryotic members from the CPA superfamily. The unrooted dendrogram shows phylogenetic relationships between 22 selected CPA sequences from sequenced genomes according to the NCBI Taxonomy database (Wheeler *et al*., 2000). The coloured bars on the right indicate classification into the CPA1 and CPA2 families, which can be further subclassified: CPA2 is divided into the CHX and NHA clades and CPA1 is divided into the NhaP-1/SOS1, plasma-membrane (PM)-NHE and intracellular (IC)-NHE clades. Phylogenetic analysis was carried out in ClustalW Phylogeny and displayed using the ATV software (Zmasek and Eddy, 2001).

CPA2 family

The CPA2 family is smaller than CPA1 and contains both Na^+/H^+ and K^+/H^+ antiporter genes from bacteria, fungi, plants and animals (Brett *et al*., 2005). It is understood that gene origin of this family is shared between two bacterial ancestral Na^+/H^+ antiporters genes, NhaA and NapA, and the bacterial K^+/H^+ antiporters KefB and KefC required for K^+ ion export to regulate intracellular pH during metabolite detoxification (Goldberg *et al*., 1987; Waser *et al*., 1992; Ferguson *et al*., 1993; Furrer *et al*., 2007). Bacterial homologues of NhaA and NapA are both shown to conserve pH-dependent electrogenic transport (Taglicht *et al*., 1993; Furrer *et al*., 2007). The CPA2 family is divided into two gene clusters; the CHX clade branching from the bacterial NapA and Kef genes and the NHA clade branching from only the bacterial NhaA genes. Unlike the CPA1 family, there are fewer CPA2 animal Na^+/H^+ antiporters with members not functionally characterised (Brett *et al*., 2005). In the NHA clade, the bacterial NhaA genes share homology with the fungal NHA family (Na^+, K^+/H^+), and a more recently identified animal NHA Na^+/H^+ antiporter family. To date only two human NHA genes have been identified, NHA1 and NHA2, compared to the more studied human NHE genes of the CPA1 family (Ye *et al*., 2006; Xiang *et al*., 2007), with the functional activity of the NHA2 isoform linked to pH and salt homeostasis in the kidneys and essential hypertension (Canessa *et al*., 1980; Fuster *et al*., 2008; Kondapalli *et al*., 2012).

2.3 - *Escherichia coli* NhaA

E. coli NhaA from the CPA2 family has been the most extensively studied Na^+/H^+ antiporter and has become a model-representative of a pH-regulated Na^+/H^+ antiporter, particularly after it was the first crystal structure of a Na^+/H^+ antiporter to be solved in 2005 (Padan *et al*., 2004; Hunte *et al*., 2005).

The transport activity of NhaA was first identified in the 1980s from an *E. coli* mutant strain, which showed increased Na^+/H^+ activity due to a mutation in a gene (*nha*A) that was sequenced to encode a Na^+/H^+ antiporter that extrudes Na^+ or Li^+ ions in exchange for a H^+ ion (Goldberg *et al*., 1987; Karpel *et al*., 1988; Taglicht *et al*., 1991). Cell expression of NhaA is regulated by a positive regulator, *nha*R, which is induced by intracellular Na^+ or Li^+ ions with its transport activity strictly regulated by intracellular pH and shown to increase by three orders of magnitude from pH 6.5 to 8.5 in the presence of both ions (Karpel *et al*., 1991). By using salt-sensitive *E. coli* mutant strains, NhaA was confirmed to be indispensible for cell growth under extreme alkaline and high salt conditions (0.6 M NaCl or 0.1 M LiCl both at pH 8.5), with increased expression and an exceptionally fast turnover rate of 100,000 ions per minute to prevent cellular alkalidosis (Taglicht *et al*., 1991, 1993).

2.4 – NhaA structure

Ten years after the *nhaA* gene was discovered, the first structural insight of the NhaA antiporter was made by electron microscopy (EM) (Williams *et al*., 1999). Cryo-EM analysis of 2D crystals showed NhaA as a dimer with two monomers containing 12 TMs (Williams *et al*., 1999; Williams, 2000). An NhaA dimeric state was later confirmed by biochemical cross-linking and genetic complementation experiments

using native membranes under the extreme salt stress conditions when NhaA is functionally most active (Gerchman *et al*., 2001).

Four years later, an NhaA crystal structure was solved at 3.45 Å in an inward-facing conformation from crystals grown at low pH where the protein is inactive (figure 2.3)(Williams *et al*., 1999; Hunte *et al*., 2005). Although the protein was not crystallised as a physiological dimer, alignment of the X-ray and cryo-EM electron density maps show the same helix positions, confirming the crystal structure to represent a native conformation (Screpanti *et al*., 2006).

The structure of the NhaA monomer shows an N_{IN} - C_{IN} topology with TMs 1 to 5 and TMs 8 to 12 related by a pseudo two-fold axis of symmetry parallel to the membrane plane and are separated by TMs 6 and 7 (figure 2.3, part a) (Hunte *et al*., 2005). In agreement with the EM and cryo-EM maps, two densely packed domains are formed by these two inverted topology repeats intertwining to form a dimerisation (interface) domain and a six-helical bundle core domain (figure 2.3, part b and c). After fitting the X-ray monomeric structure onto the cryo-EM map, it can be seen that the two adjacent □-hairpin motifs on the periplasmic side (between TMs 1 and 2) from both monomers form an antiparallel four-stranded □-sheet (figure 2.3, part c) (Appel *et al*., 2009).

This □-hairpin motif is critical for dimerisation but not for transport activity depending on the external cell growth conditions (Rimon *et al*., 2007). From complementation growth assays using salt-sensitive *E. coli* mutant strains, the wild type NhaA dimeric state has shown to complement cell growth better under extreme

salt stress conditions (0.1 M LiCl pH 7 or 0.6 M NaCl pH 8.5) and is also more thermostable compared to an NhaA monomer mutant with a deletion of only the □-hairpin motif (Herz *et al*., 2009).

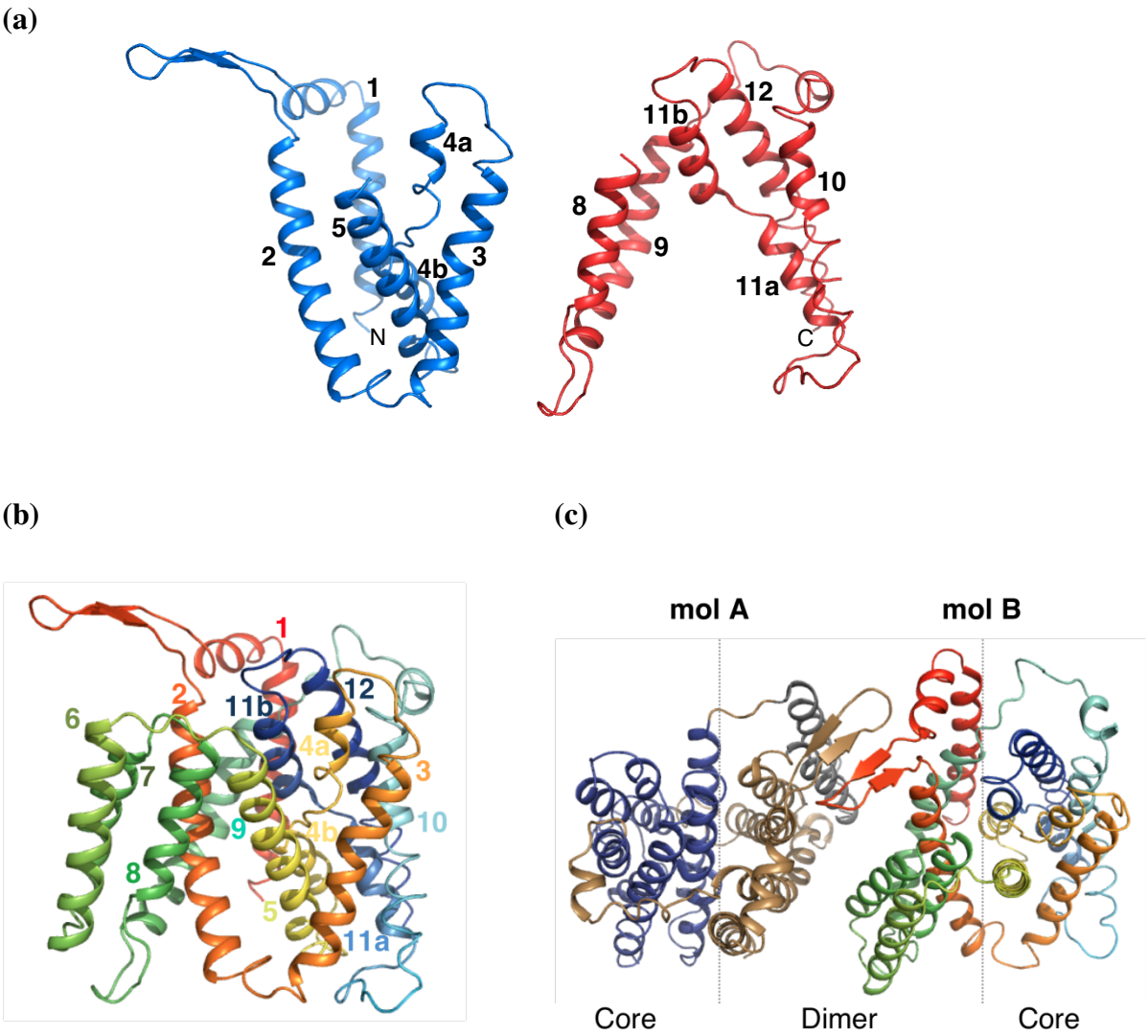
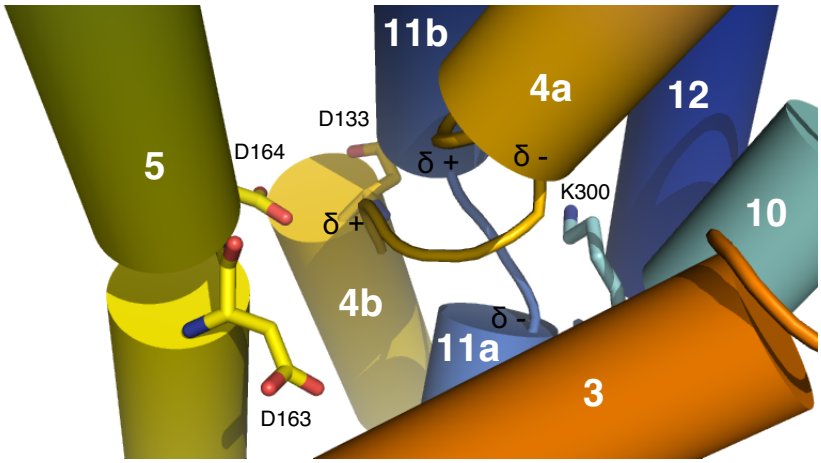


Figure 2.3 – Structure of *Escherichia coli* NhaA. (a) Two inverted topology repeats of TMs 1 to 5 (blue) and TMs 8 to 12 (red) related by a pseudo two-fold axis of symmetry parallel to the membrane plane. (b) Structure viewed in the membrane plane coloured as a rainbow from the N terminus (red) to the C terminus (blue). (c) Structure of NhaA dimeric model (Appel *et al*., 2009) viewed from the extracellular side. Mol A shows the core domain (TMs 3,4,5,10,11 and 12) coloured dark blue and the dimer domain (TMs 1,2,8 and 9) coloured beige which are connected by TMs 6 and 7 coloured grey. Mol B shows the same helix colour scheme as part (b). All structures are shown as a cartoon representation. Unless otherwise stated the colour scheme of NhaA in part (b) is applied henceforth.

NhaA fold

The NhaA structure introduced a new structural fold of a secondary transporter based on the arrangement of TMs 4 and 11 in the core domain (figure 2.4, part a) (Hunte *et al*., 2005). These helices are antiparallel and often termed "discontinuous" due to the presence of extended peptide chain regions in the middle where they cross over with each other in the centre of the core domain. Close to these extended peptide chains are two highly conserved aspartic acid (Asp) residues, Asp163 and Asp164, from TM 5, of which transport activity is abolished when their carboxylate groups are removed and therefore likely to be involved in sodium binding (Inoue *et al*., 1995). A deep narrow cytoplasmic funnel-like cavity and an oppositely facing shallow periplasmic cavity is also present showing these two aspartate residues located at the bottom of both cavities in the middle of the membrane, further supporting this extended peptide region to be the location of the ion binding site. The cytoplasmic-facing cavity is lined with negatively charged residues likely to attract hydrated sodium ions to move down towards the ion binding site for transport. These two aspartate residues are strictly conserved across all CPA2 bacterial and human NHA homologues with a similar loss of transport activity if either residue is removed (figure 2.4 part b) (Tsuboi *et al*., 2003; Furrer *et al*., 2007; Xiang *et al*., 2007). Similarly bacterial and human NHE homologues from the CPA1 family strictly conserve one aspartate residue equivalent to Asp164 from NhaA, which has also been confirmed from the NHE1 isoform to be essential for transport activity (Murtazina *et al*., 2001; Ding *et al*., 2006).

(a)



(b)

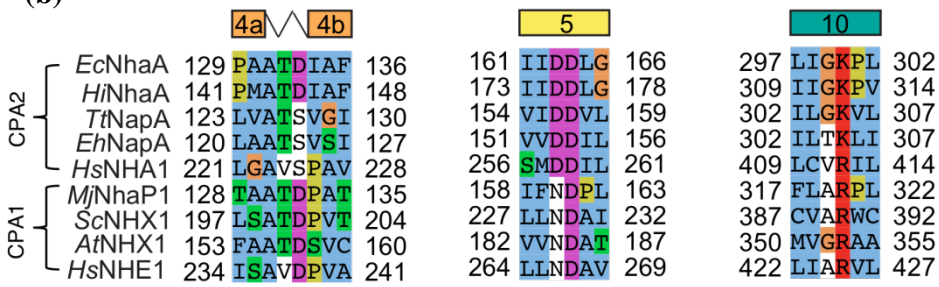


Figure 2.4 – Core domain of NhaA. (a) Cartoon representation of the core domain with the same coloured helices (as shown in Figure 2.3, part b) viewed from the periplasmic-facing side of the membrane. Essential ion binding residues Asp163 (D163) and Asp164 (D164) from TM 5 are located in the ion binding site with critical residues Asp133 (D133) from the peptide region to charge-compensate the opposing positive N-termini of TMs 4b and 11b and similarly Lys300 (K300) from TM10 for the opposing negative C-termini of TMs 4a and 11a. (b) Sequence alignment of bacterial and eukaryotic Na^+/H^+ exchangers from the CPA1 and CPA2 families. For clarity only parts of TMs 4, 5 and 10 are shown. The ion binding residue equivalent to Asp164 in NhaA on TM 5 is strictly conserved across both families as well as a positively charged residue aligning with Lys300 on TM 10 of NhaA. Asp133 on TM 4 of NhaA is only conserved across bacterial and eukaryotic Na^+/H^+ antiporters of the CPA1 family. Multiple sequence alignment was carried out in ClustalW and the MAFFT server in JalView with the sequences coloured in ClustalW format.

The ion binding site contains regions of opposing positive and negative charges from the antiparallel N- and C-termini of TMs 4a-b and 11a-b, which from the structure are proposed to be stabilised by Asp133 from TM 4 and lysine (Lys) 300 from TM 10 respectively. Both residues are confirmed to be important for transport in the ion

binding site and conserved across bacterial and human homologues (figure 2.4, part b)(Inoue *et al*., 1995; Kozachkov *et al*., 2007).

Based on the NhaA structure in combination with functional analyses, it has been proposed that when *E. coli* cells are exposed to extreme conditions of high salinity and pH, electrogenic transport exchange in NhaA is up-regulated by the nhaR regulator upon high intracellular Na^+ concentrations with the deprotonation of Asp163 and Asp164 in the ion binding site to export one sodium ion into the periplasm with the release of their protons into the cytoplasm. Other experimental techniques have indirectly confirmed the presence of a pH-induced conformational change of NhaA when it is likely to be active, such as the use of monoclonal antibodies (Venturi *et al*., 2000), accessibility to trypsin (Rothman *et al*., 1997) and fluorescent probes (Tzubery *et al*., 2004). Difference cryo-EM maps have also more recently observed helical movements in the core domain when higher pH buffers (with/without Na^+ ions) are added to 2D NhaA crystals grown at low pH (Appel *et al*., 2009).

However, the single NhaA structure to date is still the only structural model of a Na^+/H^+ antiporter in the Protein Data Bank (PDB), which makes it difficult to identify the specific movements of the helices at the molecular level during pH activation and transport activity in Na^+/H^+ antiporters. Given that NhaA has the fastest reported turnover rate of a secondary transporter, it has been proposed that the transport mechanism will involve small helical movements in the core domain rather than large movements across the whole protein (Padan *et al*., 2009).

To try and answer these questions, the main aim of this thesis was firstly to solve an NhaA structure under conditions where the protein is active and, ideally, in a different conformation to the inward-facing inactive state (Padan *et al*., 2004). The most likely reason why an NhaA active structure has still not been solved is due to reduced stability in a detergent solution. The most common approach to solving a membrane protein structure by crystallography has been through the use of detergents, as a substitute for the lipid bilayer during protein purification, but this can affect the native folded state; this is discussed further in the following section.

At the start of my PhD, two approaches had been taken to address this issue. Firstly the detergent stability of NhaA was investigated to determine the best possible solution for optimising its stability for crystallisation in an active pH condition. The second approach was to screen more bacterial homologues of NhaA from thermophilic bacteria, which are likely to be more naturally stable than NhaA and thus may be more stable under the unnatural and harsh conditions used for protein crystallisation.

2.5 - Working with membrane proteins

X-ray crystallography has been the most successful technique for high resolution structural determination of IMPs, of which 387 unique membrane protein structures have to date (March 2013) been reported in the database "Membrane proteins of known structure" (White, 2009). The general approach to solve a membrane protein structure using X-ray crystallography is to first overexpress the target protein to milligram quantities for purification using detergents to substitute the lipid bilayer and keep the membrane protein in its native folded state. From the exposed hydrophilic loop regions protein-protein interactions can be made to initiate crystal growth. As this treatment forces the transfer of a folded membrane protein from its natural lipidic membrane environment into an unnatural environment during crystallisation, it is crucial that the membrane protein is not partially unfolded which will result in protein aggregation. This is the major reason why membrane proteins represent only 0.005% of all protein structures in the PDB, reflecting the technical difficulties associated with structural studies of membrane proteins (Carpenter *et al*., 2008).

Membrane protein overexpression is commonly known to produce inclusion bodies due to oversaturation of the Sec-translocation machinery resulting in cellular toxicity. Unlike β-barrel membrane proteins, α-helical membrane proteins are most often overexpressed and extracted from the cell membranes. To prevent the membrane protein unfolding during membrane extraction, the detergent *n*-dodecyl β-D-maltoside (DDM) is often the first detergent choice for purification as it contains a long hydrophobic CH_2 chain with a large uncharged maltoside head group; therefore considered mild and less likely to destabilise the protein. Detergents stabilise the folded membrane protein as a protein-detergent micelle, provided the detergent

concentration is above its critical micelle concentration (CMC). The problem with using the relatively large DDM detergent is that it can mask the hydrophilic surface of membrane proteins needed for protein-protein interactions for crystal growth, especially with secondary transporters that do not contain intracellular or extracellular domains. Small micelle detergents containing shorter hydrophobic chains and/or smaller head groups are an alternative choice to optimise crystal packing. The head group can also remain uncharged such as β-octyl-glucoside (β-OG), or charged such as *N,N*-Dimethyldodecylamine *N*-oxide (LDAO); considered to be a harsh detergent. The use of small micelle detergents is statistically more likely to result in membrane crystals diffracting to a higher resolution compared to DDM (Sonoda *et al*., 2010). However if the membrane protein cannot remain stable in a small micelle detergent, it will unfold and aggregate.

One of the best approaches to take at the first stages of a membrane protein crystallography project is to screen for the most stable membrane protein candidates or homologues, which are more likely to produce high resolution diffracting crystals using a small micelle detergent. Stability screens have since been introduced to monitor intrinsic stability in a detergent micelle (Sonoda *et al*., 2011), but is still dependent on the membrane protein overexpression yields to produce milligram amounts of high quality pure protein for crystallisation. In other words if a membrane protein target is known to remain stable in a small micelle detergent, low expressions levels can result in over-concentrating free detergent micelles during purification and hinder crystallisation. For this reason, method developments to optimise membrane protein expression were also carried out in this thesis.

3.0 -

MemStar: A new strategy for membrane protein production in *Escherichia coli*

3.1 - Introduction

As outlined in chapter 1, one of the important parameters for obtaining well-diffracting crystals for solving membrane protein structures is their stability in detergent solution (Serrano-Vega *et al*., 2008; Sonoda *et al*., 2010; Bill *et al*., 2011). In general, the better the stability of the membrane protein in a small micelle detergent, the easier it is to optimise crystals as it increases the number of potential crystal contacts. However, the problem encountered with screening membrane proteins based on stability, such as GPCRs, is that it does not correlate with overexpression levels (Shibata *et al*., 2009; Sonoda *et al*., 2011). As such, many of the most detergent stable membrane proteins may only be produced to low levels and the extra financial cost associated with purification in small micelle detergents can make them practically infeasible to work with.

Our goal was to determine an active-state structure of a Na^+/H^+ antiporter. As will later be discussed in Chapter 5, bacterial NhaA homologues were screened to try and identify a naturally more stable homologue at active pH. Unfortunately, however, the expression levels were quite poor and we needed to improve overexpression levels of the best target to facilitate cost-effective purifications in small micelle detergents. Despite published papers already describing new methodologies shown to improve membrane protein expression yields, it is still considered to be a matter of "trial and error" to find a good combination of different parameters to obtain high membrane protein overexpression levels, e.g., such as the type of promoter, strain, fusion tag, media and culture conditions (Wang *et al*., 2003). Our approach was to firstly identify an optimum type of medium, as this is the easiest parameter to first adjust.

One common feature observed with optimised expression conditions for membrane protein production is the use of an auto-induction medium (Deacon *et al*., 2008). For this reason we started our approach by testing different auto-induction media.

In general, the *E. coli* Walker strains C43(DE3) and C41(DE3) are considered the best suited for membrane protein production in *E. coli*. The Walker strains have mutations in the *lac*UV5 promotor site, which reduce the levels of T7 RNA polymerase activity and this minimises over-saturation of the Sec-translocon (Miroux and Walker, 1996; Wagner *et al*., 2008). As such, membrane protein overexpression is generally less toxic in comparison to BL21(DE3) strains. Both strains have been successful in optimising membrane protein expression from large-scale expression screening studies (Korepanova *et al*., 2005; Hammon *et al*., 2009). More recently the Lemo21(DE3) strain was developed to regulate the T7 RNA polymerase activity by controlling the expression of its natural inhibitor T7 lysozyme (Wagner *et al*., 2008; Schlegel *et al*., 2012). In the method outlined in this chapter, both C43(DE3) and Lemo21(DE3) strains were tested in combination with the different types of media.

3.2 - Materials and methods

A detailed step-by-step protocol of the GFP-based overexpression and purification pipeline used in *E. coli* is given in Drew *et al.,* 2006.

3.2.1- Expression strains and plasmid

E. coli expression strains C43(DE3) and Lemo21(DE3) were used in this chapter (Miroux and Walker, 1996; Wagner *et al*., 2008). All control proteins were cloned in a pWaldo-GFPe vector (Drew *et al*., 2006). This expression plasmid was derived from the pET28(+) vector with the expression cassette under the transcriptional control of the T7 promotor using kanamycin selection. This vector is modified with a C-terminal green fluorescent protein (GFP) fusion containing an 8 x Histidine tag. A TEV protease recognition site is situated between the membrane protein template and GFP-His_8 tag for removal of the latter tag (Drew *et al*., 2006). Transformations and expression culture conditions were supplemented with 50 μg/mL kanamycin and 30 μg/mL chloramphenicol when using the Lemo21(DE3) strain.

3.2.2 – Types of media

Luria Broth (LB) and two auto-induction media, ZYM-5052 and PASM-5052, were used for expression screening given in table 1.

Table 1 – Ingredient preparation of the media used.

Type of media	Name of media	Media composition
Rich	Luria-Broth (LB)	10 g/L Sodium chloride 10 g/L Tryptone 5 g/L yeast extract
Auto-Induction	ZYM-5052*	1% N-Z- amine AS 0.5 % yeast extract 25 mM Na_2HPO_4 25 mM KH_2PO_4 50 mM NH_4Cl 5 mM Na_2SO_4 2 mM $MgSO_4$ 0.2x metals (optional) 0.5 % glycerol 0.05 % glucose 0.2 % α-lactose
Auto-Induction	PASM-5052* (optional SeMet labelling)	50 mM Na_2HPO_4 50 mM KH_2PO_4 25 mM $(NH_4)_2SO_4$ 2 mM $MgSO_4$ 0.2x metals 0.5 % glycerol 0.05 % glucose 0.2 % α-lactose 200 µg/mL each of 17 amino acids (no C,Y,M) 10 µg/mL Met 125 µg/mL SeMet (optional) 100nM vitamin B_{12}

* Further details of medium preparation are given in Studier, 2005.

3.2.3 - Small-scale expression

Expression screening was carried out from 5 mL cultures volumes using 24-well culture plates (Qiagen) with aerated plate sheet covers. From a fresh colony, overnight LB cultures were diluted 1:50 in the tested medium and incubated at 37°C to an OD_{600} of 0.5 followed by a further 15 hours at 25°C. An IPTG-induction step at an OD_{600} of 0.5 was carried out with 0.4 mM IPTG where specified. With the Lemo21(DE3) strain, media were supplemented with 0, 0.1, 0.25, 0.5, 0.75 and 1.0 mM L-rhamnose before inoculation.

3.2.4 - Flow cytometry analysis

Membrane protein expression was monitored on a cell population level, from fresh small-scale cultures using a FACSCalibur instrument (BD Biosciences) (Wagner *et al*., 2008; Schlegel *et al*., 2012). Cells were diluted to a final concentration of ~10^6 in 2 mL of ice-cold 1 x PBS buffer. Diluted cells were incubated on ice for 30 min in the dark with 0.2 μM of membrane staining fluorophore FM4-64 (Invitrogen). For data analysis, cell gating was based on FM4-64 staining and GFP-fluorescence was detected and presented logarithmically using the FlowJo software (Treestar). The FACS experiments and data analysis were carried out by Anna Hjelm (University of Stockholm).

3.2.5 - Large-scale expression

Large-scale expression followed the same protocol as small-scale using 1 L baffled shaker flasks. With the Lemo21(DE3) strain, autoclaved media were supplemented with optimum L-rhamnose concentration. Cells were harvested at 5,000*g* for 15 minutes and re-suspended in 50 mL 1 x PBS buffer for storage at -80°C.

3.2.6 - Membrane preparation

From large-scale expression, thawed cells were lysed by a pressure of 25 kpsi (~ 1.7 x 10^3 atm) (twice) through a cell disruptor (Constant Systems) at 4°C. Cell debris and unbroken cells were removed by centrifugation at 12,000*g* for 10 min at 4°C. From the supernatant, crude membranes were isolated as a pellet by ultracentrifugation at 140,000*g* for 2 hours at 4°C. Membrane pellets were resuspended in 15 mL 1 x PBS buffer and flash-frozen in liquid nitrogen for storage at -80°C.

3.2.7 - Quantification of membrane protein expression

From small- and large-scale expression, a 1 mL aliquot of cells were spun down at 16,000*g* for 10 min followed by resuspension in 100 μL 1 x PBS buffer. Cell suspension was transferred to a 96-well black Nunc plate with GFP fluorescence measured as relative fluorescence units (RFUs) using a SpectraMax 340PC384 Absorbance Microplate Reader (Molecular Devices, CA). GFP emission was measured at 512 nm with an excitation wavelength of 488 nm using the Softmax Pro software suite (Molecular devices). Fluorescence was converted to $mg.L^{-1}$ based the ratio of 1:1 of membrane protein expressed relative to its GFP tag (Drew *et al*., 2008).

3.2.8 – Fluorescent-detection size exclusion chromatography analysis

For each control protein, crude membranes from a 1 L culture (from standard conditions) were diluted to 3.5 mg/mL total protein concentration in 1 mL of solubilisation buffer; 1 x PBS, 150 mM NaCl and 1% (w/v) low-α DDM ($DDM_{L\alpha}$) (Generon). The total protein concentration of crude membranes was measured by a BCA assay (Thermo Scientific) with bovine serum albumin as a protein standard. The same dilution factor was applied to crude membranes of the same control protein

expressed in the MemStar condition. Membranes were solubilised for 1 hour at 4°C and isolated by ultracentrifugation at 140,000*g* for 1 hour. The detergent solubilisation efficiencies were calculated based on RFU readings taken before and after ultracentrifugation. From the supernatent, 100 µL of DDM-solubilised GFP-fusion was loaded onto a Superose 6 column (10/30, GE-healthcare) pre-equilibrated with 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.03% DDM at a flow rate of 0.4 mL/min. On-line GFP fluorescence was measured at each protein fraction (emission wavelength of 512 nm and excitation wavelength of 488 nm). GFP fluorescence was plotted against retention volume (mL) using GraphPad Prism software (San Diego, CA).

3.2.9 - Membrane protein solubilisation and purification

Buffers

Solubilisation buffer	1 x PBS, 150 mM NaCl and 1% $DDM_{L\alpha}$
Wash buffer	1 x PBS, 150 mM NaCl and 0.1% $DDM_{L\alpha}$, 10 mM Imidazole
Dialysis buffer	20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.03% $DDM_{L\alpha}$

Solubilisation

Crude membranes were diluted in solubilisation buffer to a final protein concentration of 3.5 mg/mL. After 1 hour stirring at 4°C, solubilised membranes were isolated by ultracentrifugation at 140,000*g* for 1 hour at 4°C to remove unsolubilised material in the pellet. Detergent solubilisation efficiencies were calculated from RFU readings taken before and after ultracentrifugation.

Purification

Solubilised membranes were stirred with 1 mL of Ni-NTA slurry (Qiagen) per 1 mg GFP pre-equilibrated with wash buffer for 2 hours at 4°C. The slurry was then transferred to a glass Econo-Column (BioRad) and washed with 20 column volumes (CVs) of wash buffer containing increasing amounts of imidazole: 20, 30, and 40 mM. The membrane protein-GFP fusion was eluted in 50 mL of wash buffer containing 250 mM imidazole. Cleavage of the GFP-His_8-tag was followed with the addition of equimolar His_6-TEV protease (section 3.2.10) during overnight dialysis in 3 L of dialysis buffer at 4°C. Reverse immobilised metal affinity chromatography (IMAC) was performed using a 5 mL Ni-NTA HisTrap HP column (GE-healthcare) pre-equilibrated in dialysis buffer to remove the cleaved GFP-His_8 fusion tag and His_6-TEV protease. Flowthrough containing the target protein was concentrated to 500 μL using 50 kDa MWCO Amicon Ultra-15 centrifugal concentrators (Millipore) at 3,000*g* at 4°C with 10 min intervals to minimise precipitation during concentrating. Size exclusion chromatography (SEC) was performed using a Superdex 200 10/30 gel filtration column pre-equilibrated in dialysis buffer. The monodisperse protein peak was collected and concentrated using 100 kDa molecular weight cut-off concentrators (Vivaspin) to 10 mg/mL. Final protein concentrations were measured by a BCA assay and purity assessed from 10 μg of protein by SDS-PAGE analysis. Pure protein was used for crystallisation or flash-frozen in liquid nitrogen for storage at -80°C.

3.2.10 – Expression and purification of TEV protease

The plasmid construct of His_6-TEV protease was a personal gift from Dr. Simon Newstead (University of Oxford), cloned into the pMal-C2 vector and containing an N-terminal maltose binding protein tag (MBP) to improve soluble expression yields.

A TEV protease cleavage site is present between the MBP and TEV protease for removal of the MBP during expression by self-cleavage.

Buffers

Cell resuspension buffer	1 x PBS, 300 mM NaCl
Purification buffer	20 mM Tris-HCl pH 7.5, 300 mM NaCl, 3 mM DTT and 20 % glycerol
Dialysis buffer	20 mM Tris-HCl pH 7.5, 300 mM NaCl, 5 mM DTT and 30 % glycerol
Storage buffer	20 mM Tris-HCl pH 7.5, 300 mM NaCl, 5 mM DTT and 50 % glycerol

Transformations and expression was carried out in the *E. coli* BL21-CodonPlus(DE3)-RP strain (Agilent Technologies) supplemented with 100 μg/mL carbenicillin. From a fresh colony, overnight LB culture was diluted 1:50 in 5 L of autoclaved LB prepared using 1 L baffled shaker flasks. Large-scale cultures were incubated at 37°C to an OD_{600} of 0.5-0.6 followed by induction with 0.4 mM IPTG and continued incubation for a further 15 hours at 25°C. Cells were harvested at 5,000*g* for 15 min and the cell pellet resuspended in 200 mL cell resuspension buffer and either stored at -80°C or immediately lysed by a pressure of 25 kpsi (twice) through a cell disruptor (Constant systems) at 4°C. Cell debris and unbroken cells were removed by centrifugation at 12,000*g* for 10 min at 4°C, with the supernatent extracted and ultracentrifuged at 140,000*g* for 45 min at 4°C. The supernatent, after ultracentrifugation, was isolated from the pellet of crude membranes and supplemented with 20 mM imidazole (final concentration) before being loaded onto two 5 mL Ni-NTA HisTrap HP columns (GE-healthcare) connected in series to an

ÄKTAprime™ (GE-healthcare) pre-equilibrated with purification buffer containing 20 mM imidazole. The supernatent was run at 2 mL. min^{-1} followed by purification buffer containing 50 mM imidazole for 20 CVs or until the monitored UV absorbance at 280 nm had decreased to zero. Bound TEV protease protein was eluted in 100 mL purification buffer containing 250 mM imidazole and dialysed overnight in 3 L of dialysis buffer at 4°C. The dialysed sample was re-loaded onto the Ni-NTA HisTrap columns pre-equilibrated with dialysis buffer containing 20 mM imidazole followed by 20 CVs of dialysis buffer containing 50 mM imidazole. Elution of the bound TEV protease was carried out using 100 mL of dialysis buffer containing 250 mM imidazole, which was added to 350-400 mL storage buffer with a final protein concentration of 1 $mg.mL^{-1}$. Purified TEV protease was finally aliquoted as 5 mL volumes into 15 mL falcon tubes for storage at -80ºC.

3.2.11 – Fluorescent based CPM thermostability assay

The thermostability assay was carried out according to the protocol described by Stevens and co-workers (Alexandrov *et al*., 2008). Stocks of CPM (7-diethylamino-3-(4'-maleimidylphenyl)-4-methylcoumarin) dye (Sigma) dissolved in DMSO to 4 mg/mL were prepared as 5 µL aliquots and stored at -80°C. 1 µL of 10 mg/mL purified protein was added to 150 µL of buffer containing 20 mM Tris-HCl pH 7.5, 150 mM NaCl and 0.1% LDAO in a 96-well black Nunc plate. A thawed aliquot of CPM dye was diluted to 40 µg/mL in a dilution buffer (20 mM Tris-HCl pH 7.5, 150 mM, 0.03% DDM) followed by 3 µL added quickly to the protein in the dark. The plate was covered with a clear cover plate with fluorescence monitored at an emission wavelength of 463 nm with excitation of 387 nm on SpectraMax plate reader pre-incubated at 40°C. Fluorescent readings were measured every 5 minutes for 3 hours

with a 15 second shaking interval. The fraction of folded protein at each time point was calculated by the quotient of raw fluorescence measured at each time point divided by the maximal fluorescence measured. The $T_{1/2}$ values were calculated from a single exponential decay curve plotted using Graphpad Prism software (San Diego, CA).

3.2.12 - Crystallisation

Crystallisation screens used the MemGold™ 96-well conditions from Hampton Research (Appendix I) on 96-well sitting drop plates with 200 nL of purified protein mixed with 200 nL of reservoir solution. Crystallisation plates were prepared using a Mosquito robot (TTP labtech) and stored at 20°C. Crystallisation trials of the control protein TehA from *Haemophilus influenza* were carried out by Hassanul Choudhury (Imperial College London).

3.3 - Results

3.3.1 – Control membrane proteins

The objective of this chapter was to develop a generic expression screen to optimise membrane protein expression in *E. coli* in a high-throughput manner for structural biology studies. Expression screening was first conducted from a bacterial library of 10 "control" integral membrane proteins with published crystal structures, of which their crystal forms had successfully been reproduced in-house using a GFP-based overexpression and purification pipeline (table 2)(Drew *et al*., 2001, 2005).

Table 2 – List of 10 control membrane proteins

Membrane Protein	**Function**	**Species**	**Reference**
Mhp1	Benzyl-hydantoin transporter	*Microbacterium liquefaciens*	(Weyand *et al*., 2008)
NhaA*	Sodium/proton antiporter	*Escherichia coli*	(Hunte *et al*., 2005)
Yiip	Zinc transporter	*Escherichia coli*	(Lu and Fu, 2007)
EmrD	Multidrug transporter	*Escherichia coli*	(Yin *et al*., 2006)
GlpG	Intramembrane protease	*Escherichia coli*	(Wang *et al*., 2006)
AsbT	Sodium/bile acid transporter	*Neisseria meningitidis*	(Hu *et al*., 2011)
GlpT	Glycerol-3-phosphate transporter	*Escherichia coli*	(Huang *et al*., 2003)
AmtB	Ammonium channel	*Escherichia coli*	(Zheng *et al*., 2004)
LacY	Lactose permease	*Escherichia coli*	(Abramson *et al*., 2003)
TehA	Tellurite channel	*Haemophilus influenza*	(Chen *et al*., 2010)

* An NhaA mutant (A109T, Q277G and L296M) was used as it produced higher resolution diffracting crystals than the wild type, described in Chapter 4.

3.3.2 - Effect of media type

The type of medium was first tested, as this is the easiest parameter to adjust. Membrane protein expression yields in C43(DE3) cells were first monitored in LB and a routinely used auto-induction medium ZYM-5052. The expression yields in each medium condition were compared by converting the measured GFP relative fluorescence units (RFUs) to $mg.L^{-1}$ and divided by the final OD_{600} to give the final overexpression level normalised to cell density ($mg.L^{-1}.OD_{600}^{-1}$). To simplify the comparisons between the different overexpression conditions, the average overexpression levels normalised to cell density is only indicated for each expression condition.

The ZYM-5052 auto-induction medium produced a slightly lower average overexpression level (2.8 $mg.L^{-1}.OD_{600}^{-1}$) compared to cells grown in LB (3.0 $mg.L^{-1}.OD_{600}^{-1}$) (figure 3.1), indicating that the increased yields reported in auto-induction media are most likely due to the increased biomass (Deacon *et al*., 2008). It is understood, however, that recombinant membrane protein expression is typically poor if induced late in the *E. coli* growth phase (Drew *et al*., 2005). With reference to this, expression screening was repeated in the ZYM-5052 medium with a standard IPTG-induction step at mid-log phase as carried out in the LB medium. In other words, we were using the ZYM-5052 as a "non" auto-induction medium.

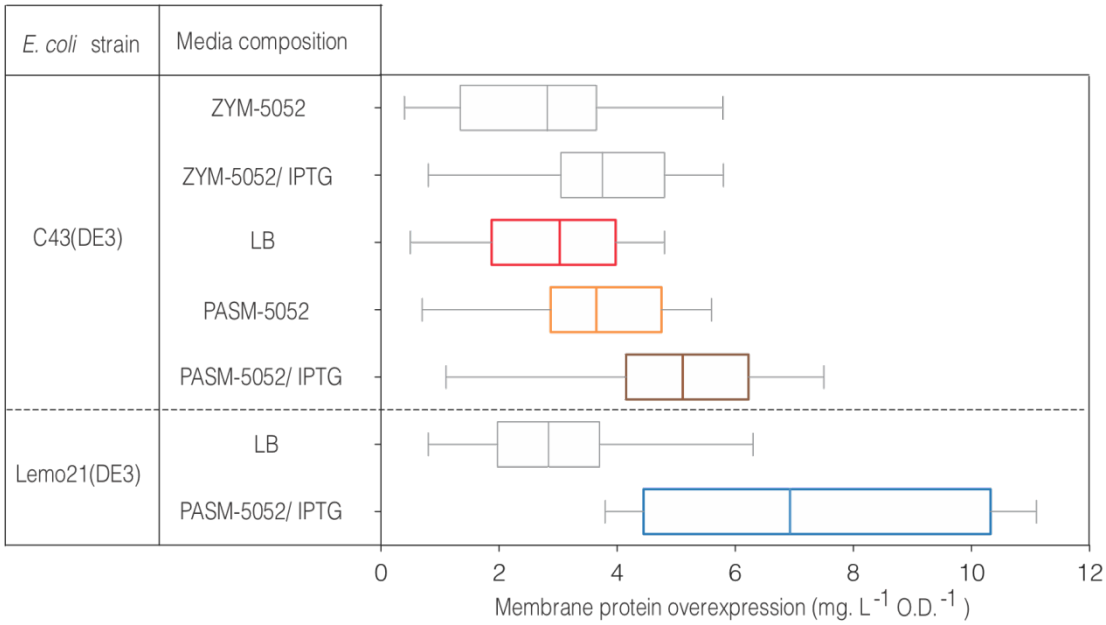


Figure 3.1 – Expression screening in different medias and culture conditions. Box-and-whisker plot illustrating the distribution of membrane protein expression (mg. L^{-1}) normalised to cell density (OD_{600}) for the 10 control membrane proteins in each of the different culture conditions as labelled. The bottom and top of each box represents the 25th and 75th percentile with the whisker lines indicating maximum and minimum expression levels in each condition with the vertical line representing the mean normalised expression level (mg. $L^{-1}.OD_{600}^{-1}$). The coloured boxes represent the culture conditions further tested.

The addition of IPTG-induction in the ZYM-5052 medium further increased average overexpression levels from 2.8 mg.$L^{-1}.OD_{600}^{-1}$ to 3.7 mg.$L^{-1}.OD_{600}^{-1}$ (figure 3.1). To investigate this combination further, another auto-induction medium, PASM-5052, was next tested. PASM-5052 auto-induction medium is generally used for selenomethionine (SeMet) protein labelling, a commonly used heavy atom derivative for crystallographic phasing (Morth *et al*., 2006), but can still be used as a standard medium without selenomethionine (Studier, 2005). By switching from ZYM-5052 to PASM-5052 with IPTG induction, overexpression levels increased even further from 3.7 mg.$L^{-1}.OD_{600}^{-1}$ to 5.1 mg.$L^{-1}.OD_{600}^{-1}$ (figure 3.1). So far, just by switching the medium, a ~ 1.6 fold higher average overexpression level normalised to cell density was obtainable using PASM-5052 with IPTG induction in comparison to LB. For this

reason, further expression optimisation was carried out on PASM-5052 with IPTG induction.

3.3.3 - Expression optimisation using Lemo21(DE3)

We next decided to test the Lemo21(DE3) strain in both LB and the PASM-5052 medium with IPTG induction to see if it could further improve overexpression levels compared to the C43(DE3) strain.

In Lemo21(DE3) the average membrane protein overexpression level increased from 5.1 to 6.9 mg.L^{-1}.OD_{600}^{-1}. As such, a further 35 % improvement was made compared to growth in C43(DE3) cells, as shown previously in figure 3.1. Furthermore, the presence of selenomethionine in the PASM-5052 medium did not drastically affect overexpression levels from Lemo21(DE3) cells (figure 3.2). In LB, however, there is no improvement in the use of Lemo21(DE3) compared to C43(DE3).

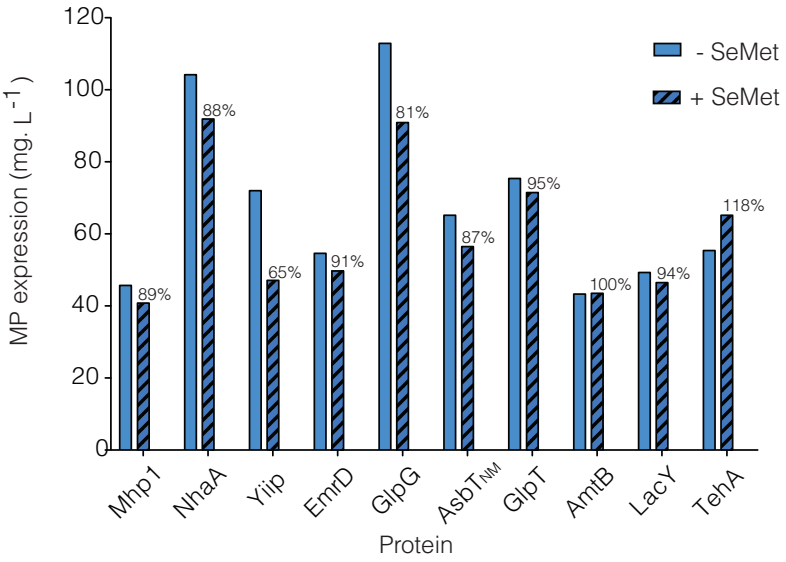


Figure 3.2 – Membrane protein expression yields with selenomethionine incorporation. Expression levels from the 10 control proteins using the Lemo21(DE3) strain in the PASM 5052 medium with IPTG induction with/without selenomethionine incorporation. Expression levels with selenomethionine are given as a percentage of the expression levels produced without selenomethionine present.

The optimal L-rhamnose concentration varied between the control proteins in both media types (figure 3.3). Mhp1 was the only membrane protein to have optimal overexpression without the addition of L-rhamnose. Interestingly for each control protein, the optimal L-rhamnose concentration with maximal overexpression was the same in both media types. After analysing the overexpression data more closely, we also found that the expression yields using 0.25 mM L-rhamnose were always at least 65% of the maximal overexpression level attainable, as shown in figure 3.3 and later in figure 3.7.

Using the final combination, the average overexpression levels normalised to cell density increased by 230%. That is, from 3.0 $mg.L^{-1}$. OD_{600}^{-1} for the standard conditions (C43(DE3) cells grown in LB) to 6.9 $mg.L^{-1}$. OD_{600}^{-1} for the new optimised conditions of Lemo21(DE3) cells grown in PASM-5052 with an IPTG induction step, as shown in figure 3.1. For sake of simplicity, it was decided to name this final overexpression combination as MemStar.

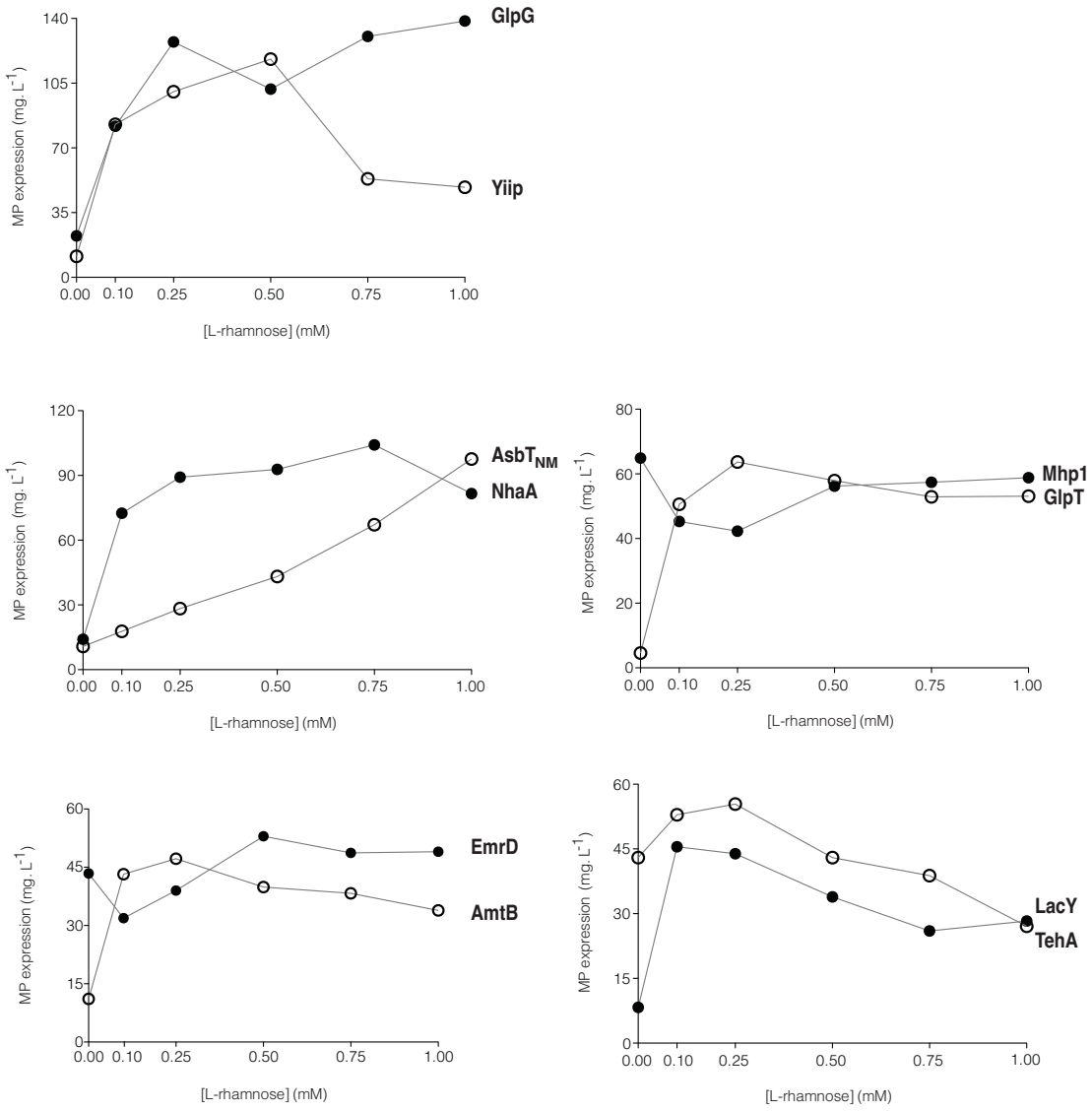


Figure 3.3 – Optimised expression levels using the Lemo21(DE3) strain. Expression levels from the 10-control proteins in the presence of the indicated amounts of L-rhamnose using PASM-5052 with IPTG induction. The L-rhamnose concentration producing the highest overexpression levels was selected for large-scale expression.

3.3.4 - Flow cytometry analysis

To understand how the PASM-5052 auto-induction medium and Lemo21(DE3) strain was improving membrane protein expression, we analysed the evolution of MemStar by fluorescence-activated cell sorting (FACS) analysis. Membrane protein

overexpression levels were monitored by GFP fluorescence and the total cell population by the use of a membrane binding fluorescent dye FM4-64. Instead of conducting FACS analysis on all 10-control proteins, we focused our efforts on the control protein $AsbT_{NM}$. This control was chosen since its overexpression levels was closest to the calculated average overexpression level normalised to cell density across all 10-control proteins.

With C43(DE3) cells grown in LB, a large fraction of cells did not express the target protein (non-expressing cells) (figure 3.4, red line). By switching to PASM-5052, however, this non-expressing cell population reduced (figure 3.4, orange line). By including an IPTG induction step, the fraction of cells expressing the control protein further increased as well as the amount produced from the expressing cells (figure 3.4, brown line). The most straightforward explanation is that by switching to the auto-induction medium, we have suppressed residual "leaky" expression. That is, since the overexpression of membrane proteins is generally "toxic" to cell growth, in LB the leaky expression may have been culling viable expressing cells. Interestingly, by including IPTG induction to the auto-induction media, we dramatically maximise the amount of membrane-integrated expression levels from an increased fraction of expressing cells. Although we do not have an explanation for this, it seems that cells reaching stationary phase are not as viable for membrane protein production as those at mid-log phase (unpublished data).

As expected, Lemo21(DE3) cells grown in PASM-5052 with IPTG induction produced the highest overexpression levels from expressing cells as well as a greater fraction of cells expressing the target protein (figure 3.4, blue line). The

Lemo21(DE3) strain has recently been shown to reduce the amount of inclusion body formation (Schlegel *et al*., 2012). In line with this, another expression screen study using the C43(DE3) strain was shown to express 70 out of 99 membrane proteins from *Mycobacterium tuberculosis* in *E. coli* with >94% of expressed proteins forming some degree of insoluble aggregation (Korepanova *et al*., 2005). Since cytoplasmic aggregates can deplete the cell of viable chaperones, such as blockage of the Sec translocon machinery, this is one reason why we think Lemo21(DE3) can increase membrane protein overexpression yields. With Lemo21(DE3) we can fine-tune the optimal strength of the T7 promoter to further maximise the amount of membrane-integrated expression with a reduction of the adverse effects leading to cell toxicity.

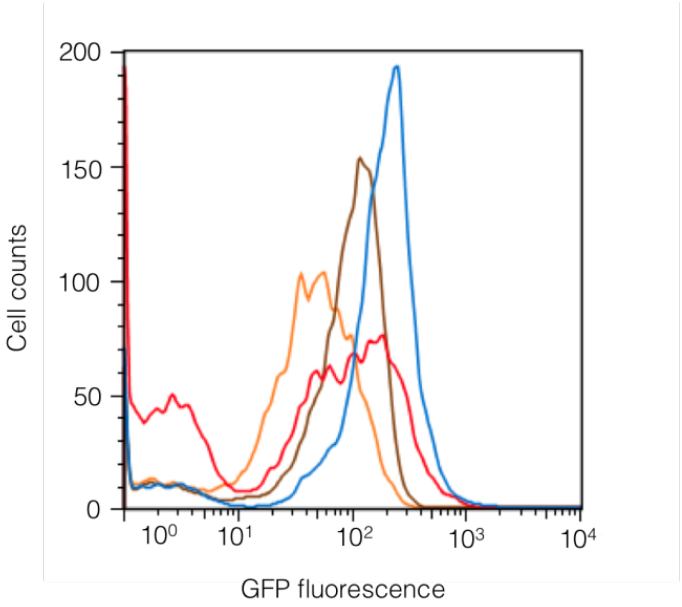


Figure 3.4 – Flow cytometry analysis of the control protein $AsbT_{NM}$. Overexpression of the $AsbT_{NM}$ was carried out in C43(DE3) cells grown in LB (red), PASM-5052 (orange), PASM-5052 with IPTG induction (brown) and Lemo21(DE3) cells grown in PASM-5052 with IPTG induction (blue). The coloured culture conditions correspond with the box-and-whisker plots in Figure 3.1. (note: $AsbT_{NM}$ was selected due to its close membrane protein expression levels ($mg.L^{-1}.OD_{600}^{-1}$) to the mean expression level of the 10-control proteins). FACS experiment and data analysis were carried out by Anna Hjelm (University of Stockholm).

3.3.5 - FSEC analysis

To examine the quality of a folded membrane protein from using MemStar, crude membranes from the 10-control proteins were isolated from 1 L cultures and solubilised in 1% DDM to monitor monodispersity of the detergent-solubilised extracts by FSEC analysis. In this case, the use of FSEC has a further practical advantage as we can directly compare the quality and amount of DDM-solubilised material isolated from the MemStar and standard conditions. All control proteins were confirmed to be stable in DDM showing a single, sharp monodisperse trace in both expression conditions (figure 3.5).

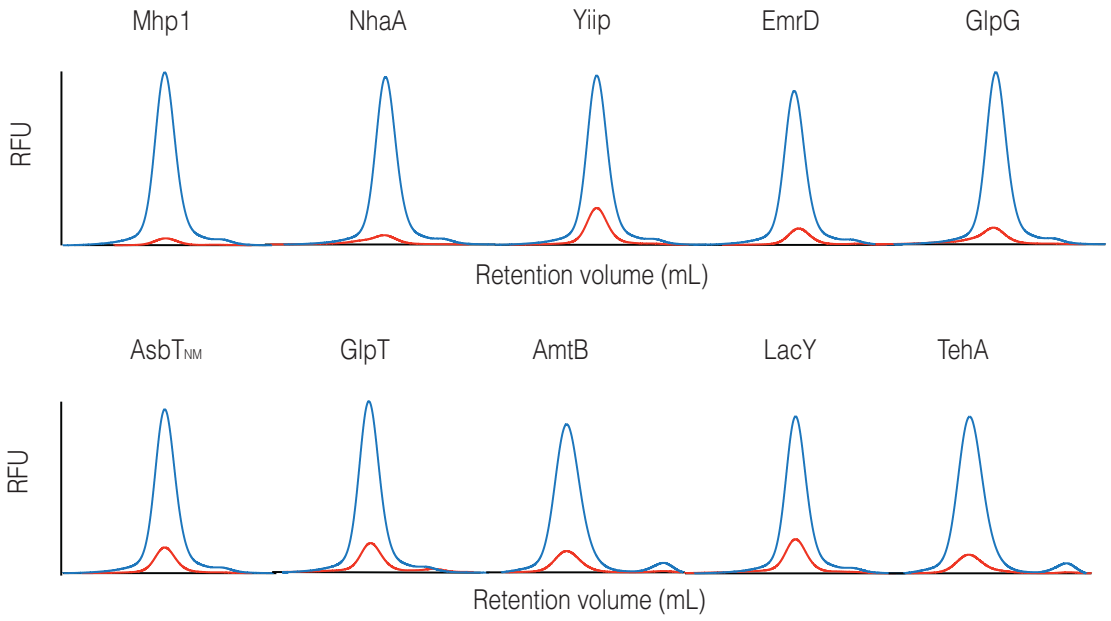


Figure 3.5 – Quality assessment of overexpressed control proteins by FSEC analysis. FSEC traces of DDM-solubilised membranes from the 10 control proteins from 1 L cultures grown in standard (red) or MemStar (blue) conditions.

From isolated crude membranes, the average peak heights from using MemStar is 9-fold higher than from using the standard conditions, which is significantly more than the average 2.3-fold increase based only on their normalised overexpression levels measured directly from overexpressed cells. Furthermore, the DDM-solubilisation

efficiencies were only 40-50% from the MemStar membranes compared to 100% from the standard membranes. Even though we could have increased the amount of detergent to extract more protein in the crude membranes isolated from MemStar, we decided not to, as one of the reasons for developing this overexpression platform in the first place was to keep down the financial costs of detergent usage.

3.3.6 - Purification and crystallisation

By using membranes isolated from our MemStar protocol, we were able to reproduce the crystals from all control proteins in a similar manner to the same protein produced previously from standard conditions (Sonoda *et al*., 2011). This analysis included the bacterial homolog TehA of the plant SLAC1 anion channel, where X-diffraction was observed up to 1.2 Å (figure 3.6). As one would expect, the final protein yields using MemStar were much higher for all proteins. Taken together, there is no apparent compromise in the quality of the protein produced for structural work by MemStar.

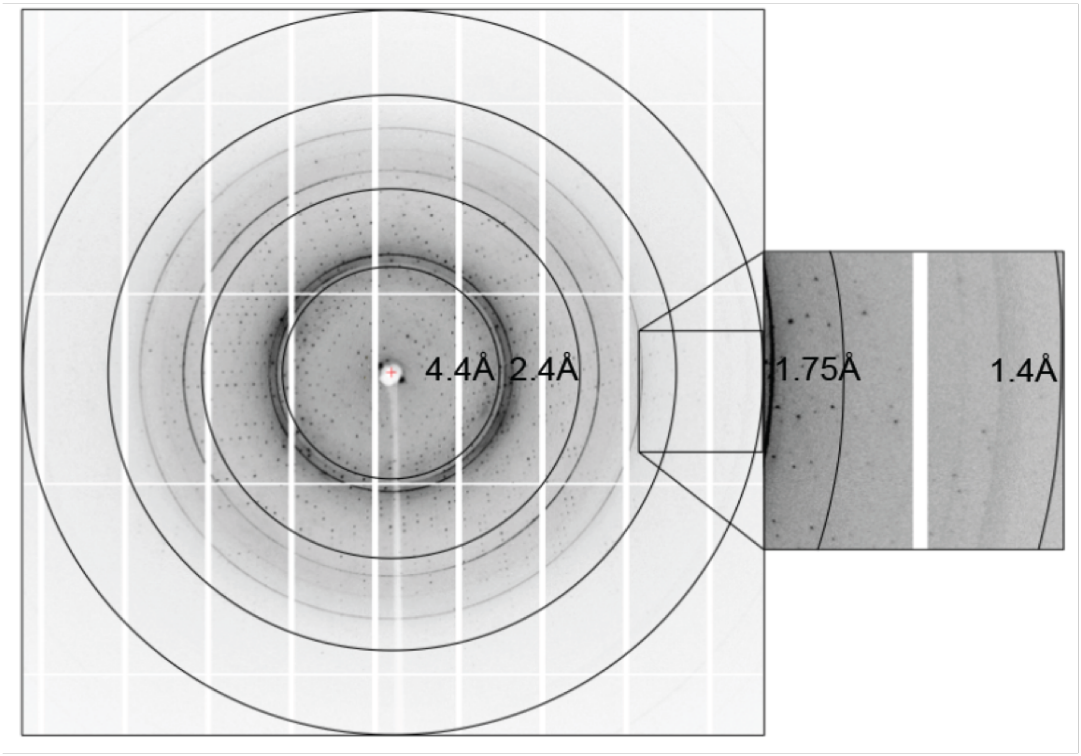


Figure 3.6 - X-ray diffraction of the control protein TehA. Crystals were grown in the MemGold 2.15 condition (refer to Appendix I) from protein expressed using MemStar. Crystal screening was carried out on the microfocus beamline I24 at the Diamond light source using a Pilatus 6M-F detector. TehA crystals grew in the R3 spacegroup with the following cell dimensions: a=b= 97 Å, c= 136 Å and □=□= 90° and □= 120°. TehA crystallisation and X-ray crystal screening was carried out by Hassanul Choudhury (Imperial College London).

3.3.7 - Large-scale expression screening

The next step was to test MemStar on a broader range of membrane proteins, in terms of their function and size. For a sample set, 24 membrane proteins from a previous study were chosen that are known to express from very low (< 1 mg) to high levels (> 5 mg) in *E. coli* using the same pWaldo GFP-fusion vector (Drew *et al*., 2005). MemStar should be useful for large-scale screening, however L-rhamnose titrations for every construct becomes impractical. As mentioned previously, we observed that with using 0.25 mM L-rhamnose the overexpression levels for the control proteins were at least 65% of the maximum expression level obtainable. To investigate this

trend more thoroughly, we also screened 11 different mutants of the control membrane protein NhaA. Interestingly, the variation in expression levels between point mutations of the same protein was as large as that across the 10 different proteins (figure 3.7). On a practical level again we observe 0.25 mM rhamnose results in at least 65% of the maximum expression level for each tested NhaA mutant. As such, we tested expression levels only with this concentration of L-rhamnose in the MemStar condition across the additional 24 membrane proteins as well as the standard conditions for consistency.

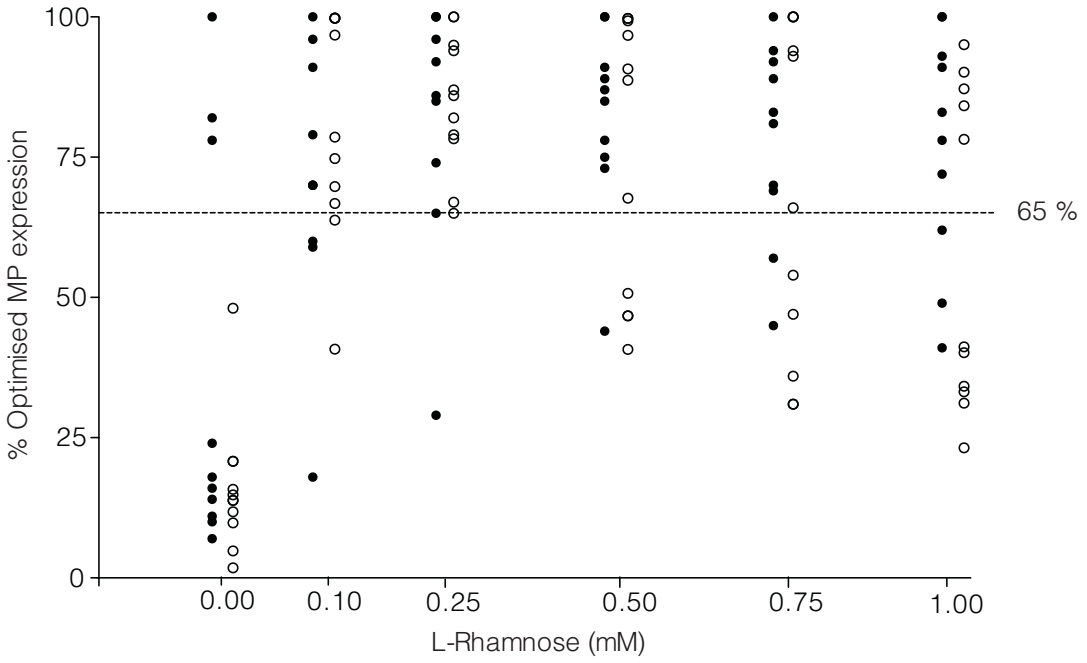
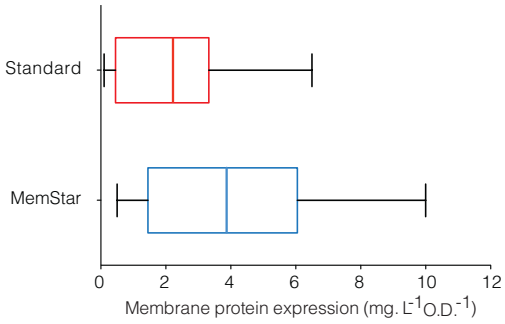


Figure 3.7 – Correlation between expression levels and L-rhamnose concentration using MemStar. Expression levels are represented as a percentage of the highest expression across the indicated L-rhamnose concentrations for the 10 control proteins (filled circles) and NhaA mutant library (unfilled circles). The dotted line on the *y* axis represents ≥ 65% maximal expression present in 20 out of 21 cases at 0.25 mM concentration of L-rhamnose.

Similarly to the first expression screen, the normalised average overexpression level from the standard to MemStar condition increased 2-fold, that is, from 2.2 to 3.9

mg.L^{-1}.OD_{600}^{-1} (figure 3.8, part a). After analysing the data in more detail, we observed 17 out of the 24 proteins expressing above 12.5 mg.L^{-1} in the MemStar condition compared to only 6 proteins in the standard condition (figure 3.8, part b).

(a)



(b)

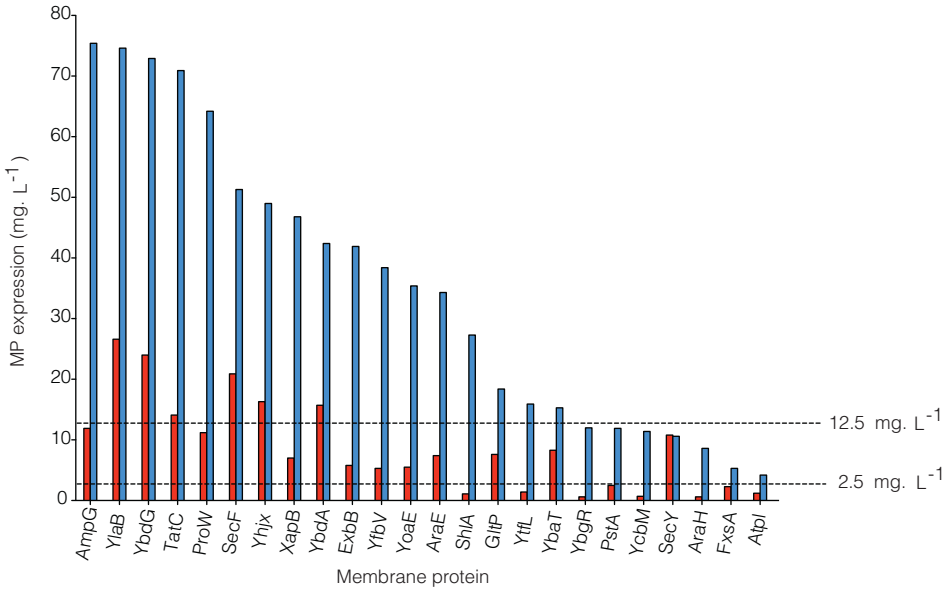


Figure 3.8 – Medium-scale analysis of the application of MemStar using an established 24-membrane protein-GFP fusion library (a) Box-and-whisker plot illustrating the distribution of membrane protein expression (mg.L^{-1}) normalised to cell density (OD_{600}) for the 24 control membrane proteins in standard and MemStar conditions. The bottom and top of each box represents the 25th and 75th percentile with the whisker lines indicating maximum and minimum expression levels in each condition with the vertical line representing the mean normalised expression level (mg. L^{-1}.OD_{600}^{-1}). (b) Overexpression levels from all membrane proteins in the standard and MemStar conditions. Expression levels of at least 2.5 mg.L^{-1} (lowest dotted line) from the standard conditions can be increased to above 12.5 mg.L^{-1} using MemStar (highest dotted line).

The average normalised overexpression level across these 17 membrane proteins increased slightly more from 2.8 to 5.0 $mg.L^{-1}.OD_{600}^{-1}$. Interestingly, 6 out of the 7 membrane proteins expressing less than 12.5 $mg.L^{-1}$ using MemStar were also the poorest expressing proteins in the standard condition showing expression yields less than 2.5 $mg.L^{-1}$. Based on the expression data from this larger test-case of membrane proteins, an *ad hoc* cut-off expression level of ~ 2.5 $mg.L^{-1}$ can be a suggested minimal level of expression required to increase overexpression to above 12.5 $mg.L^{-1}$ using MemStar.

Given that the membrane proteins expressing below this *ad-hoc* level have very different functions, this can suggest that MemStar's failure not to optimise effectively their expression levels is not linked to protein-specific toxicity. The lowest expressing membrane protein is Atp1, which had its expression increased from 1.2 to 4.2 $mg.L^{-1}$ using MemStar. This target is an accessory subunit of the F_0F_1 ATP synthase and documented to poorly translate due to many rare codons (Walker *et al*., 1984, Norholm *et al*., 2012) and thus its expression levels may be independent of the MemStar conditions.

Translation inefficiency was investigated further by codon optimisation from one control protein, AraH, which similarly expressed below the *ad hoc* cut-off level of ~ 2.5 $mg.L^{-1}$ by the lab of Professor Daniel Daley (University of Stockholm). Using the standard conditions, AraH expression levels increased from 1.1 to more than 2.5 $mg.L^{-1}$ after codon optimisation, which was satisfactorily boosted to more than 17 $mg.L^{-1}$ using MemStar.

3.3.8 – Homologue screening

To investigate further the expression limits of MemStar, we decided to screen homologues of two control proteins; EmrD and GlpG (table 3). In this way we can minimise the individual membrane protein effects on expression levels; in other words homologues with the same membrane protein function are more likely to have similar toxicity effects. For EmrD the poorest expressing homologue (*Bacillus substilis*) in standard conditions of 2.4 $mg.L^{-1}$ was boosted to 35 $mg.L^{-1}$ using MemStar. For homologues of GlpG, however, the three that expressed to levels lower than 2.3 $mg.L^{-1}$ in the standard condition (*Shewanella oneidensis, Haemophilus influenzae* and *Pyrococcus horikoshii*) could not be optimised to expression levels above 12.5 $mg.L^{-1}$ using MemStar. As a result, this work has shown that the *ad hoc* cut-off expression level of 2.5 $mg.L^{-1}$ is applicable to homologues of the same protein as well as for completely different membrane proteins.

Table 3 – List of selected EmrD and GlpG homologues

Membrane protein	Species	Standard (mg. L^{-1})	MemStar (mg. L^{-1})
EmrD	*Bacillus cereus*	15.5	55.8
	Pseudomonas putida	11.8	39.6
	Haemophilus influenzae	7.3	25.5
	Shewanella oneidensis	5.5	23.4
	Escherichia coli	4.4	53.0
	Shigella flexneri	3.9	36.7
	Bacillus subtilis	2.4	35.4
GlpG	*Escherichia coli*	17.8	127.3
	Bacillus subtilis	8.4	37.5
	Shigella flexneri	5.9	31.9
	Thermus thermophilus	2.3	34.8
	Shewanella oneidensis	1.8	8.0
	Haemophilus influenzae	1.1	5.3
	Pyrococcus horkoshii	1.2	4.2

3.3.9 – Correlation between optimised expression using MemStar and thermostability

One of the original arguments for the development of this system was to improve the expression of membrane proteins, which can only be produced to low levels even though they may be more stable and therefore more likely to produce well-ordered crystals using a small micelle detergent. Our final analysis of the MemStar condition was therefore conducted from 9 control proteins, from the first test-case, to compare their optimised expression levels with thermostability monitored by their unfolding rate in the LDAO detergent; known to produce harsh membrane protein unfolding conditions due to its small hydrophobic chain and small charged head group. Based on cysteine accessibility, membrane protein unfolding rates were monitored using the thiol-specific fluorochrome CPM dye incubated with pure protein in LDAO at 40ºC, producing fluorescence upon reacting with free sulfhydryl groups exposed during membrane protein unfolding. The calculated LDAO half-life ($T_{1/2}$) was used to compare membrane protein stability with optimised expression. Across all 9 membrane proteins, the data did not show an obvious linear correlation (R^2 = 0.1) between optimised expression levels using MemStar and protein stability, as shown from figure 3.9. A correlation was more visible with the exclusion of the control protein AmtB (R^2 = 0.6), however this trend would need to be further verified by testing more control proteins.

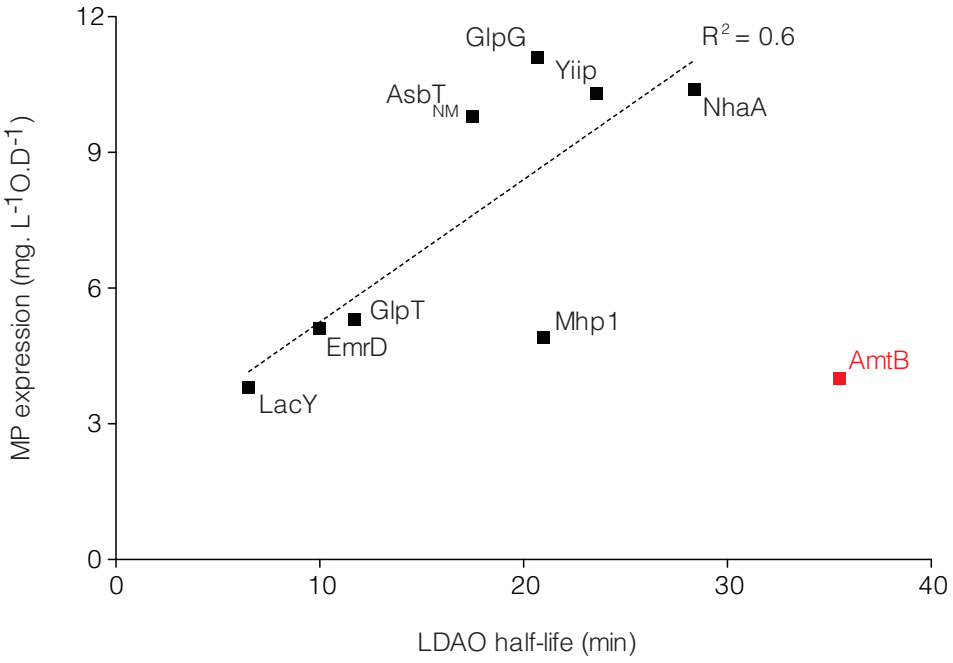


Figure 3.9 – Correlation between membrane protein thermostability and optimised expression levels using MemStar. The data across 8 membrane proteins (excluding AmtB; shown in red) shows a reasonable correlation (R^2 = 0.6) between membrane protein stability monitored by the unfolding half-life rates in LDAO at 40ºC, and the optimised expression levels using MemStar.

3.4 - Discussion

To resolve a high resolution crystal structure of a membrane protein, one important parameter is the production of sufficient isolated quantities of monodisperse and stable membrane protein. Most often membrane proteins are first tested for crystallisation in a protein–detergent micelle. However if expression levels are low (<10 $mg.L^{-1}$), free detergent molecules are likely to be overconcentrated and hinder the growth of well-ordered crystals. Ideally, if expression levels could be optimised in a high throughput manner, this could justify using small micelle detergents routinely in a membrane protein crystallography laboratory and also help hasten the time-consuming process of crystal optimisation.

Previously in our laboratory, routine protein production of bacterial membrane proteins were overexpressed using LB medium and the *E. coli* expression strain C43(DE3); referred to as standard conditions (Miroux and Walker, 1996). By combining the PASM-5052 auto-induction medium and the Lemo21(DE3) expression strain with IPTG induction, this produced a 2-fold average increase in the normalised membrane protein overexpression levels relative to expression levels using the standard conditions. This 2-fold increase was established from 10 control membrane proteins with solved crystal structures and a library of 24 different bacterial membrane proteins known to express from less than 1 $mg.L^{-1}$ to 5 $mg.L^{-1}$ in a similar setup (Drew *et al*., 2005).

Expression studies in *E. coli* have shown auto-induction to produce higher expression yields for membrane proteins (Deacon *et al*., 2008; Gordon *et al*., 2008). However, in some media types, such as Terrific Broth and auto-induction media, the cell biomass

increases as well as the expression levels. This was not the case with MemStar. We were able to improve the expression level per OD_{600}. Flow cytometry showed that we had increased not only the amount of expressed protein produced from an expressing cell, but also the population of expressing cells. FSEC analysis confirmed a substantial increase in the amount of membrane protein extracted from crude membranes, with no apparent loss of quality. The DDM-solubilisation efficiencies from MemStar are less than 100% due to their higher total protein concentrations in the crude membranes (10-16 $mg.mL^{-1}$) compared to the standard membranes (4-5 $mg.mL^{-1}$). However, given the high expression levels produced from MemStar, one can afford to solubilise less protein in order to obtain a higher protein: detergent ratio, particularly when using expensive small micelle detergents.

The conditions of MemStar also hold two very practical advantages, making it even more suitable for membrane protein crystallography projects. Firstly the PASM-5052 medium is designed to express selenomethionine-labelled proteins for phase estimation in crystallography (Studier, 2005). The expression levels from the 10-control proteins were virtually independent of selenomethionine incorporation, which can provide more flexibility as well as making MemStar, financially, more cost-effective without the requirement of using a methionine auxotrophic strain. Secondly, the Lemo21(DE3) strain can be applied in a high-throughput manner to optimise membrane protein expression levels by only altering the concentration of L-rhamnose rather than testing a number of different expression strains, which is considered to be a more "trial and error" approach (Wagner *et al*., 2008). Interestingly, 0.25 mM L-rhamnose produced at least 65% of maximum expression levels of the optimum concentration from 10 different membrane proteins and 11 mutants from the same

protein. Consequently, 70% of a further 24 different membrane proteins produced expression levels above 12.5 mg.L^{-1} from using only 0.25 mM L-rhamnose. This practical approach can therefore be applied for large-scale high-throughput expression screening of multiple constructs or homologues, at the beginning of a structural project, followed if necessary by further expression screening using more L-rhamnose concentrations.

One unexpected result from screening the 24 control membrane proteins was that overall the MemStar condition could not optimise expression levels above 12.5 mg.L^{-1} if the control proteins were poorly expressing below 2.5 mg.L^{-1} from the standard conditions. The most likely reason for these membrane proteins still expressing poorly is either due to poor translation efficiencies or difficulties in membrane protein folding when overexpressed. One control protein, AraH, previously expressed to less than 2.5 mg.L^{-1} in the standard conditions was increased to more than 17 mg.L^{-1} by combining codon optimisation with the use of MemStar.

Based on this data, an *ad hoc* cut-off expression level of ~ 2.5 mg.L^{-1} is proposed to be the minimum level of expression required to increase expression levels above 12 mg.L^{-1} using MemStar. However this cut-off value, itself, should not be taken too strictly as an expression benchmark given that one EmrD homologue did show a significant increase in expression from 2.4 mg.L^{-1} to 35 mg.L^{-1} from using MemStar. Nonetheless, out of all 59 membrane proteins tested in this chapter, 83% expressed above 12.5 mg.L^{-1} using MemStar compared to 29% from the previous standard conditions. Altogether, this work can justify MemStar to be a potential candidate system for high throughput expression for membrane protein production in *E. coli*.

A reasonable correlation is shown between the optimised expresssion levels using MemStar from 8 control membrane proteins and their protein stability, which is an important factor for obtaining well-ordered crystals (Sonoda *et al.*, 2011). Although, one must not ignore the possibility that more stable membrane proteins suitable for crystallisation may still show impractically low levels of expression, which require the availability of large-scale fermenters to produce enough material compared to the more common and easy to use shaker flasks.

MemStar has been confirmed to eliminate this obstacle based on the 3.0 Å crystal structure of the sodium: proton antiporter NapA from *Thermus thermophilus*, which was solved after expression screening for more thermostable cysteine mutants for phasing purposes. With these much higher expression levels, crystal optimisation took less time as routine purifications using a small micelle detergent could be carried out from 5 L MemStar cultures. This work is described in Chapter 5.

4.0 -

Structural studies of

Escherichia coli NhaA

4.1 - Introduction

Even though the first crystal structure of NhaA was solved in an inactive state, it still gave us the first insight into the structural arrangement of the antiporter. However, since then, an NhaA structure solved in an active state from crystals grown above pH 6.5 has still not been obtained, which has made it very difficult to understand its transport mechanism. Membrane protein instability in a detergent solution at high pH is the most likely reason. In this Chapter, our initial approach was to purify and crystallise NhaA in a high pH condition where the protein is active.

James Bowie and co-workers first demonstrated that the likelihood of uncovering a stabilising point mutation for a membrane protein was significantly higher than for globular proteins (Zhou and Bowie, 2000; Bowie, 2001). However, this observation was not taken advantage of until the work of Christopher Tate and colleagues, who carried out alanine scanning to design a thermostable mutant of an eukaryotic GPCR (turkey β1-adrenergic receptor) that would remain in its functional folded state using a small micelle detergent (octyl-thiol-glucoside) for crystallisation (Serrano-Vega *et al*., 2008; Warne *et al*., 2008). The structure was subsequently solved to 2.9 Å.

The limitation of the above approach is that it requires a high-affinity ligand to confirm functionality as part of the screening and thermostabilisation process. Prior to the start of my PhD, as part of a small project by a former Master student, NhaA was selected to create a more general stabilisation strategy. Stabilising point mutants were generated by random mutagenesis and selected based on expression levels (> 1 $mg.L^{-1}$) and those that had better than 50% extraction efficiency in 1% β-OG from total

membranes, at a concentration of 3 $mg.mL^{-1}$, after one-hour incubation at 4°C. This procedure was carried out because it was apparent that there was a correlation between the stability of membrane proteins, as judged by their $T_{1/2}$ unfolding rates in LDAO at 40°C using the CPM-based assay (described in Chapter 3) and their β-OG solubilisation efficiency (personal communication with Dr. David Drew). From this work, a thermostabilised NhaA mutant was generated that harboured two random mutations (Ala109Thr and Gln277His) that produced better crystal growth at high pH (above pH 6.5) compared to the wild type. When I joined this project, the first task was to see if we could use this thermostabilised NhaA mutant to obtain an active state structure.

At the same time, NhaA wild type crystals in the inactive state were also repeated (by a former Master student) from purified protein using the GFP-based pipeline. These crystals grown at pH 4 were crystallised as a dimer instead of the monomer observed in the initial crystal structure (Hunte *et al*., 2005). A 3.7 Å structural model of the NhaA dimer was obtained by molecular replacement using the monomeric structure to calculate the electron density map.

The crystal structure of the NhaA dimer was overall very similar to the dimer modelled from the cryo-EM maps using the monomeric X-ray structure (Appel *et al*., 2009). However, the crystal structure showed the four-stranded β-sheet protruding along the membrane plane with a hydrogen bonding network present in the two adjacent β-hairpins, compared to the model showing a more curved β-sheet with the tips of the β-hairpins pointing towards the periplasm (figure 4.1). During refinement of this structure, it was noticed in the core domain that TM 10 showed a misalignment

between the electron density and the residue positions. It appeared that the sequence of TM 10 extended from residue 287 to 313, rather than from 290 to 316 in the published structure (Hunte *et al*., 2005). Realignment of TM 10 was, therefore, also important to confirm as it contains the highly conserved Lys300 residue in the middle of the helix, which is critical for ion binding (Kozachkov *et al*., 2007). For this side project, it was decided to introduce a methionine residue on TM 10 for selenomethionine protein labelling using MemStar, to use the anomalous signal from the selenium atom as a structural parameter to confirm the reassignment of TM 10 in the refined NhaA dimeric structure.

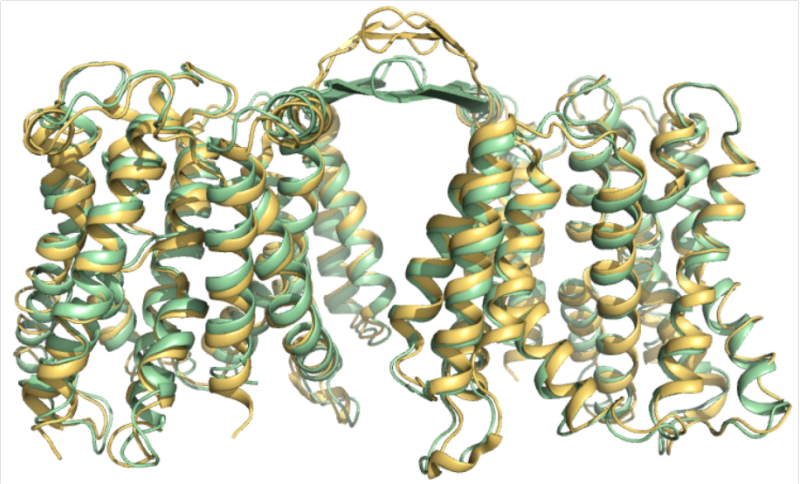


Figure 4.1 – Structure of the NhaA dimer. Cartoon representation of the NhaA dimeric model (PDB: 1ZCD, (Appel *et al*., 2009)) (beige) superposed onto a crystal structure of NhaA (green) solved as a dimer. The crystal structure shows the orientation of the four-stranded β-sheet on the periplasmic surface more flat compared to the model.

4.2 - Materials and methods

Detailed descriptions of the protocols used for expression and FSEC analysis have been given in Chapter 3. Only further experiments or modifications are described in this chapter.

4.2.1 - Site-directed mutagenesis

The gene cloned in the pWaldo-GFPe plasmid was used as the template for site-directed mutagenesis. All point mutations were made using the Quikchange Lightning Site-Directed Mutagenesis Kit (Agilent Technologies) with forward and reverse primers carrying the complementary codon mutation. PCR reactions were carried out according to the protocol given with the kit. Point mutations were confirmed by DNA sequencing (Source Bioscience) and the resulting DNA sequences analysed using EnzymeX.

4.2.2 - β-Octyl-glucoside solubilisation efficiency

Crude membranes from a 1 L culture were diluted to 3.0 mg/mL total protein concentration in 1 mL of solubilisation buffer; 1 x PBS, 150 mM NaCl and 1% (w/v) β-OG (Anatrace). The total protein concentration of crude membranes was measured by a BCA assay (Thermo Scientific) with bovine serum albumin as a protein standard. The membranes were incubated in the solubilisation buffer for 1 hour at 4°C with mild agitation, followed by ultracentrifugation at 140,000*g* at 4°C for 1 hour. Detergent solubilisation efficiencies were calculated based on RFU readings taken before and after ultracentrifucation.

4.2.3 - Heat-FSEC analysis

The protocol was identical to FSEC analysis with a heat induction step. Detergent-solubilised membranes (after ultracentrifugation) were incubated at 40°C for 20 min followed by centrifugation at 13,000*g* for 10 min to remove any precipitation formed before being loaded into the Superose 6 column.

4.2.4 – NhaA solubilisation and purification

Buffers

Solubilisation buffer	1 x PBS, 150 mM NaCl and 1% $DDM_{L\alpha}$
Wash buffer	1 x PBS, 150 mM NaCl and 0.1% $DDM_{L\alpha}$, 10 mM imidazole
Dialysis buffer	30 mM MES pH 6.5, 150 mM NaCl, 0.03% $DDM_{L\alpha}$
Crystallisation buffer	20 mM sodium citrate pH 4.5, 150 mM NaCl, 0.03% high-α DDM ($DDM_{H\alpha}$)

Crude membranes were diluted in solubilisation buffer to a final protein concentration of 6.5 mg/mL. After 1 hour stirring at 4°C, solubilised membranes were isolated by ultracentrifugation at 140,000*g* at 4°C for 1 hour to remove unsolubilised material in the pellet. Detergent solubilisation efficiencies were calculated from RFU readings taken before and after ultracentrifugation. Solubilised membranes were stirred at 4°C with 1 mL of Ni-NTA slurry per 1 mg GFP pre-equilibrated with wash buffer. After 2 hours the slurry was transferred to a glass Econo-Column (BioRad) and washed with 20 column volumes (CVs) of wash buffer containing increasing amounts of imidazole: 20, 30, 40 and 50 mM. The membrane protein-GFP fusion was eluted in

50 mL of wash buffer containing 250 mM imidazole. Cleavage of the GFP-His_8-tag was followed by addition of equimolar His_6-TEV protease during overnight dialysis in 3 L of dialysis buffer. Reverse IMAC was carried out using a 5 mL Ni-NTA HisTrap HP column pre-equilibrated in dialysis buffer. NhaA contains 2 histidine residues within the first 5 residues of the N-terminus, which resulted in the protein binding to the column and not being eluted in the flowthrough. The column was washed with 5 CVs of dialysis buffer and then washed with 3 CVs of buffer containing 250 mM imidazole with the flowthrough collected containing eluted protein, free GFP-His_6 tag and His_6-TEV protease. The flowthrough was concentrated to 500 μL using 100 kDa MWCO Amicon Ultra-15 centrifugal concentrators (Millipore) at 3,000*g* with 10 min intervals to remove any precipitation during concentrating. The concentrated sample was diluted 1:60 into crystallisation buffer followed by reconcentration to 500 μL with free GFP precipitating out due to the low pH. SEC was followed using a Superdex 200 10/30 gel filtration column pre-equilibrated in crystallisation buffer. The monodisperse protein peak was collected and concentrated with 100 kDa molecular weight cut-off concentrators (Vivaspin) to 10 mg/mL. Final protein concentrations were measured by a BCA assay and purity assessed from 10 μg of protein by SDS-PAGE analysis. Pure protein was used for crystallisation trials or flash-frozen in liquid nitrogen for storage at -80°C.

4.2.5 - Crystallisation

Crystallisation trials were set up by the hanging drop vapour method manually using 24-well VDXm plates (Hampton Research) using the following optimised crystallisation conditions previously determined:

Low pH (inactive NhaA state) -	0.1 M sodium citrate pH 3.5 (and 4.0),
	0.1 M lithium sulphate and 22 – 36% PEG 400.
High pH (active NhaA state) -	0.1M MES pH 5.5 – 6.5 or
	0.1M Tris-HCl pH 7.5,
	0.1 M lithium sulphate and 22 – 36% PEG 400.

1 μL of purified protein was mixed with 1 μL of mother liquor on a siliconised glass cover slide (Hampton Research) and equilibrated over 500 μL of reservoir (mother liquor) solution at 20 °C.

4.2.6 - Additive and Detergent screens

Crystal optimisation was carried out using an additive screen HR2-428 (Hampton Research), and detergent screen 1 (Hampton Research) manually on 24-well hanging drop plates. A 10-fold dilution of each condition was added to purified protein prior to being mixed with reservoir solution.

4.2.7 - Amphiphiles

Pure protein was incubated for 5 min with 1% Facade-EM (Avanti Polar Lipids) before being mixed with reservoir solution. This was also carried out in conjunction with the additive and detergent screens.

4.2.8 - Crystal dehydration

Cover slides containing equilibrated crystal drops were transferred, sequentially overnight, to wells containing the same reservoir solution with a 2% increment of

PEG 400 up to a final concentration of 34%. Crystals were flash-frozen from different PEG concentrations.

4.2.9 - Crystal freezing

Crystals were mounted in a cryo-loop (Hampton Research) and transferred to 1 μL of cryoprotectant solution; crystallisation condition containing 40% PEG 400 with 1% crystallisation detergent. After a few seconds, crystals were flash-frozen in liquid nitrogen.

4.2.10 - Data collection, processing and structure determination

X-ray crystal screening and data collection were carried out on synchrotron beamlines at Diamond Light Source (Didcot, United Kingdom) and at the ESRF (Grenoble, France). The best data set was collected on Diamond beamline I03 from an NhaA triple mutant (Ala109Thr, Gln277Gly and Leu296Met), which was designed to confirm the proposed repositioning of TM 10 from a refined model of the NhaA wild type dimeric structure.

The purpose of growing crystals of this mutant with selenomethionine labelling was to locate the position of the selenium atom site from Met296 on TM 10 based on its anomalous scattering signal (*f''*). This occurs when energy from the emitted X-rays is absorbed by electrons in the heavy atom, which changes the phase of the diffracted X-ray waves and as a result breaks Friedel's law between pairs of X-ray reflections (*hkl* and *–h-k-l*), needed for structure determination. The reflections are collected as intensities from diffraction images, of which the number of images required for a complete data set is dependent on the crystal lattice symmetry, defined by its

spacegroup in the asymmetric unit cell. The contribution of anomalous scattering in the crystal can be estimated from differences in the measured X-ray intensities between these Friedel pairs. This can be used to locate the positions of the selenium atoms in the structure.

The maximum anomalous scattering signal from selenium occurs close to its absorption edge wavelength, which was determined by a fluorescent scan of the crystal, using the program CHOOCH (Evans and Pettifer, 2001) to show the X-ray energy (keV) required to produce a peak *f''* signal during data collection. Data collection was carried out at a fixed X-ray wavelength of 0.98 Å (12.659 keV). The triple mutant crystallised in the same $P2_1$ spacegroup with similar cell dimensions (shown in table 4 in the results section) to the wild type structure. In this spacegroup, at least 180° of diffraction images are required for a complete data set due to the anomalous scattering breaking Friedel's Law. To increase the redundancy of the anomalous signal, whilst minimising radiation damage to the crystal during data collection, 90° of diffraction images (0.5° per image) were collected from six separate crystal wedges across the crystal, showing the same diffraction quality and resolution. A total of 450° of data from the first 5 wedges were used in the final data set, which were processed and merged together to 3.5 Å using the Xia2 pipeline (Winter, 2010) to XDS (Kabsh, 2010) with further processing using the CCP4 suite of programs (Collaborative Computational Project Number 4, 1994).

The initial phases from the triple mutant data were estimated by molecular replacement (MR) in the program PHASER (McCoy *et al*., 2007), using the published monomeric structure as the search model. In this case, MR was considered the most

convenient phasing method by using the atomic positions from the published structure to estimate the phases of the triple mutant. An anomalous difference fourier map was calculated in COOT (Emsley and Cowtan, 2004) using the estimated phases with the differences in the observed structure factor amplitudes between the Friedel pairs. The structure factor amplitudes are proportional to the square-root of the intensities, collected from the diffraction images. The anomalous difference map showed the anomalous peak signals consistent with selenium atoms, of which their positions were compared with their corresponding methionine residues from a calculated electron density map of the triple mutant. MR was also repeated using one monomer from the wild type dimeric structure after TM 10 had been reassigned as the search model. Using COOT, an anomalous peak observed at 4.1σ consistent with a selenium atom from Met296 was at the same position as Lys300 from the electron density map calculated using the published structure for MR. In contrast, the same anomalous peak aligned with Leu296 from the electron density map using the wild type structure with the reassigned TM 10 for MR. After this leucine residue was substituted with a methionine, the selenium peak was exactly on top of the sulphur atom; confirming its position in the new wild type structure to be correct. The anomalous peak positions from the remaining selenium atoms were also consistent with their corresponding methionine residues from both electron density maps, confirming the repositioning of only TM 10 in the NhaA structure.

Model building of the NhaA dimer in the program O (Jones and Kjeldgaard, 1997) was interspersed with structural refinement using the PHENIX package (Adams, 2010) to improve the refined model. Since the data from the triple mutant was isomorphous with the wild type, the previously refined model of the wild type dimer

was used for model building and refinement. During map calculation, B-factor sharpening (DeLaBarre and Brunger, 2006) was applied, followed by real space electron-density averaging using RAVE (Kleywegt, 2001).

Structural refinement was conducted to minimise the differences between the calculated structure factor amplitudes from the refined NhaA model with the observed amplitudes from the triple mutant data set. After a refinement cycle, comparisons between the calculated and observed amplitudes were made from the R_{work} and R_{free} values, of which the latter uses 5% of reflections not used in refinement to give an unbiased estimate of the improvement of the model. Consequently, a decrease in both R values would indicate fewer differences between the model and the observed data. Due to the 3.5 Å resolution, structural restraints were used to help improve refinement of the model by adding more observational data to the model. Non-crystallographic symmetry (NCS) restraints were first applied across all four monomers (one monomer per NCS restraint group). During the following refinement cycles, secondary structure restraints were also applied to add hydrogen bonding distance restraints in the helices. The refined model was also improved with the addition of group B-factor restraints to refine the B-factors from multiple atoms per monomer group with also TLS (translation/liberation/screw) rigid body restraints present per monomer (Winn *et al*., 2001). Refinement was improved slightly further in the last few cycles using the anomalous scattering data from the triple mutant as an experimental phase restraint. In the final refinement cycles, the NCS group restraints were only applied across the two dimers in the asymmetric unit cell (one dimer per NCS group), to provide less restraint between the two monomers to adopt slightly different conformations but still preserve the non-crystallographic symmetry between the two dimers.

4.3 - Results

4.3.1 - Improving NhaA stability at high pH

The provided thermostable NhaA mutant contained two mutations at Ala109Thr positioned on TM 3 and Gln277His on the loop proceeding TM 9 (figure 4.2). This mutant had previously been shown to produce the best diffraction to 5 Å from crystals grown at pH 6.5. To improve the crystal growth quality and resolution, it was decided to test more point mutations on residue 277 to see if thermostability could be further improved. Using MemStar all tested mutants (given in Table 7, Appendix II) expressed to 130 mg.L^{-1} compared to 14 mg.L^{-1} from the standard conditions, previously described in Chapter 3.

Figure 4.2 – An NhaA thermostable mutant. Cartoon representation of NhaA showing the position of two thermostabilising point mutations (spheres) generated by random mutagenesis (by a former Master student): Ala109Thr (orange) on TM3 and Gln277His (turquoise) on loop 9.

Protein stability was monitored by the extraction efficiencies in □-OG from crude membranes and LDAO stability using the CPM assay from purified protein. As mentioned previously, this procedure was carried out due to a correlation between the stability of membrane proteins judged by their LDAO $T_{1/2}$ unfolding rates at 40°C and their solubilisation efficiencies in □-OG (personal communication with Dr. David Drew). One NhaA mutant now containing Gln277Gly showed the most promising improvement in crystal growth between pH 6.5 – 7.5 with a further small increase in protein stability (figure 4.3). However the crystal diffraction did not significantly improve further to a resolution better than 4-5 Å.

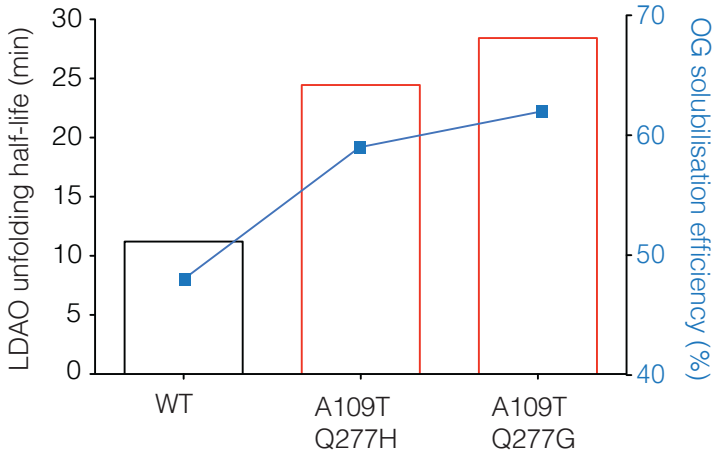


Figure 4.3 –Protein stability of NhaA wild type and thermostable mutants. Protein thermostability, shown as a bar chart, was monitored by the LDAO unfolding $T_{1/2}$ at 40°C from DDM-purified protein of the wild type (black) and two thermostable mutants (red). A correlation is also observed between their OG solubilisation efficiencies, shown in blue, measured from OG-solubilised crude membranes and protein thermostability.

However, this new thermostable mutant showed a better quality of crystal growth and X-ray diffaction at the inactive pH range compared to what had been previously seen from the wild type. This was found to be extremely useful for the work in the following section.

4.3.2 - Structural realignment of helix 10 using selenomethionine labelling

To confirm the realignment of TM 10 from the refined NhaA dimeric structure was correct, selenomethionine-labelling was chosen as a structural parameter. By collecting data at the absorption edge wavelength of selenium, the selenomethionine residue position could be located based on its anomalous peak signal. Since we were most interested in confirming the position of Lys300, which is important in the ion binding site, leucine-296 was first chosen for methionine substitution based on its close distance to Lys300. At the same time though, two more single methionine mutation sites were made at isoleucine- 293 and leucine-302 in case crystals from the Leu296Met mutant did not diffract well (figure 4.4)

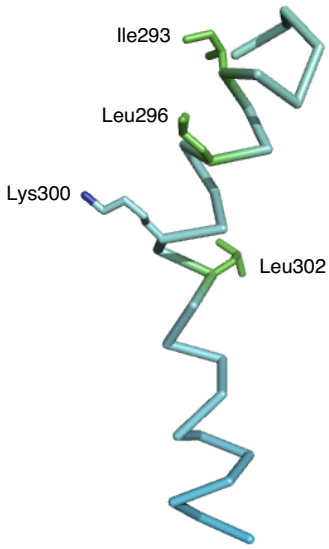


Figure 4.4 – Position of three methionine substitution sites on helix 10 of NhaA. Ribbon representation of helix 10 showing the position of three chosen residues (green sticks) point mutated to methionine with their position close to the strictly conserved Lys300 residue (turquoise).

With the MemStar expression system, expression screening of these mutants was efficient, producing yields of 20-40 $mg.L^{-1}$ which made it very convenient for large-scale expression from as little as 2 L cultures. Using total membranes, FSEC analysis

was carried out from DDM- and LDAO-solubilised membranes, with all three mutants showing the same equal and overlapping FSEC traces in these two detergents; shown to be a good indicator of membrane protein stability (figure 4.5, part a)(Sonoda *et al*., 2011). Solubilisation and purification was carried out in DDM to grow the same crystal form of the wild type dimer. All mutants showed the same symmetrical SEC trace with concentrated protein of a pure quality suitable for crystallisation (figure 4.5, part b).

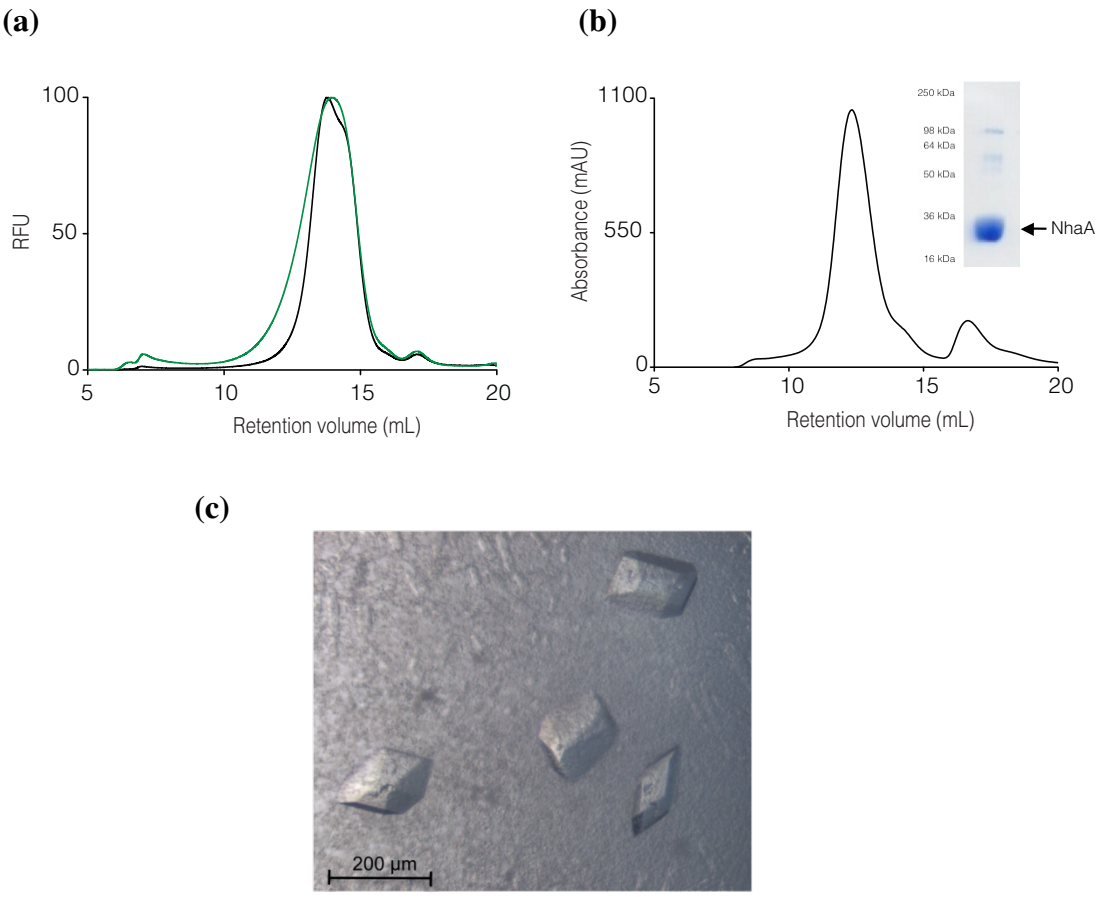


Figure 4.5 – Characterisation of an NhaA methionine mutant. All three single methionine NhaA mutants showed the same results as shown here from Leu296Met. (a) FSEC analysis from detergent solubilised membranes in DDM (black) and LDAO (green) that were extracted from crude membranes. (b) SEC analysis from DDM-purified protein with protein quality confirmed by SDS-PAGE analysis (in inset). (c) Crystals grown to full size after 3-4 days at 20°C in 0.1 M sodium citrate pH 3.5, 0.1 M lithium sulphate and 28% PEG 400 with 1% (w/v) heptyl-thiol-□-D-glucoside and 1% hexanediol.

Crystallisation was carried out using the optimised crystallisation conditions of the wild type; 0.1 M sodium citrate pH 3.5 (and 4.0), 0.1 M lithium sulphate and 22 – 36% PEG 400. The same wild type crystal form grew from all the mutants (figure 4.5, part c) with the most promising X-ray diffracting crystals from Leu296Met with a resolution of ~ 4-5 Å. These crystals grew in 32% PEG 400 at pH 3.5 with 1% (w/v) heptyl-thiol-β-D-glucoside and 1% hexanediol. However, the resolution could not be improved further despite numerous crystallisation trials. Secondly, radiation damage was a problem during data collection due to the low crystal symmetry spacegroup ($P2_1$), meaning that at least 180°-360° of diffraction patterns were needed to get a complete data set due to the breakage of Friedel's Law from the selenium atom. In this case we were also more interested in getting an accurate anomalous signal from the selenium atom, in which one then wants to collect even more data to get a high redundancy.

4.3.3 - Crystal optimisation using a thermostable NhaA mutant

To improve the efficiency of crystal optimisation of the Leu296Met mutant, this mutation was added onto the thermostable NhaA mutant that had grown consistently better crystals than the wild type at low pH (section 4.3.1). For clarity this mutant with now three point mutations is referred to as the triple mutant. Encouragingly, this triple mutant showed a further increase in its β-OG-extraction efficiency and LDAO stability (figure 4.6, part a). Protein thermostability was also monitored by heat-FSEC analysis from LDAO-solubilised membranes, with the wild type showing a slightly a broader trace compared to the other two tested thermostable mutants (figure 4.6, part b). Applying a heat step before FSEC analysis has become a more efficient way to screen thermostable membrane proteins earlier from isolated crude membranes

without the process of protein purification (Hattori *et al*., 2012; Leung *et al*., 2012). Using MemStar, expression levels increased significantly to 100 mg.L^{-1} with selenomethionine labelling, which was sufficient to grow 2 L large-scale cultures. Protein of a pure quality was concentrated to 10 mg.mL^{-1} with a final volume of 1 mL, which significantly reduced the time taken for crystal optimisation as numerous crystallisation plates could be set up after one purification prep (figure 4.6, part c).

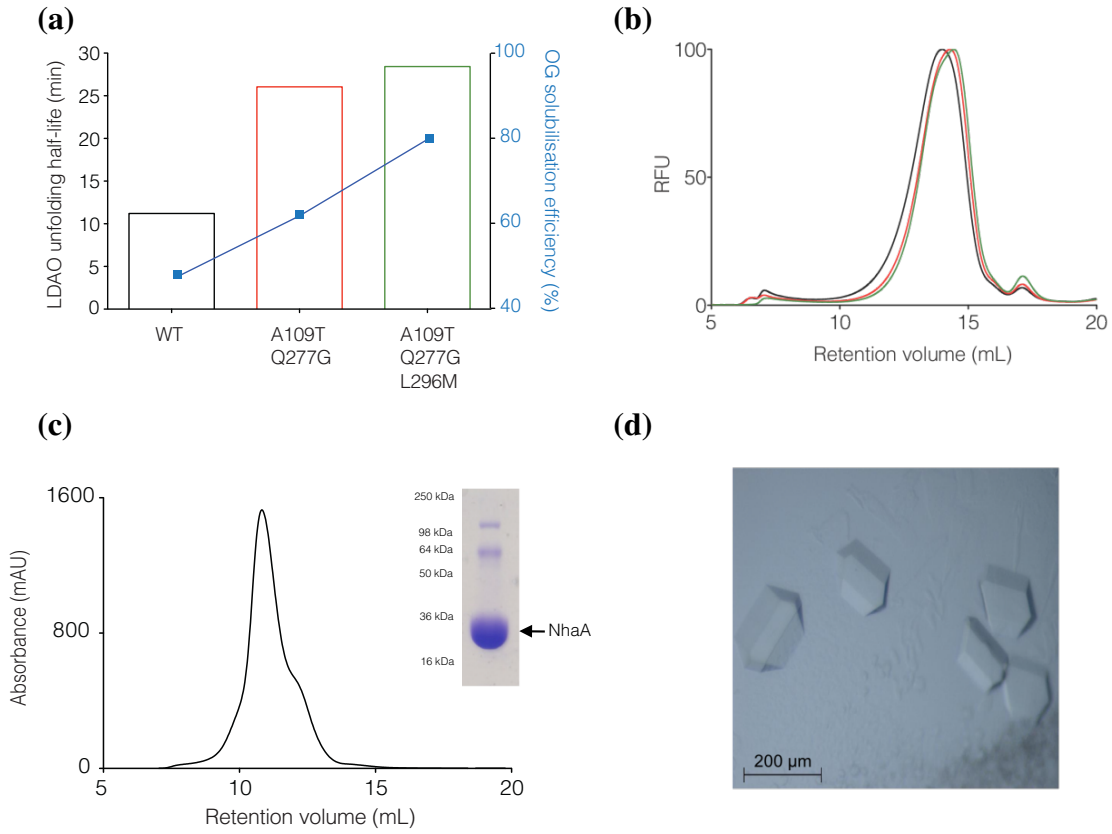


Figure 4.6 – Protein stability and characterisation of an NhaA triple mutant. This mutant was derived by adding the Leu296Met mutation onto the previously made NhaA thermostable mutant containing two point mutations (Ala109Thr and Gln277Gly). (a) The triple mutant (green) shows a further improvement in thermostability based on its LDAO unfolding $T_{1/2}$, shown as a bar chart, than the wild type (black) and thermostable mutant before the methionine mutation was added (red) and OG solubilisation efficiency, shown in blue. (b) Heat-FSEC analysis involved incubation LDAO-solubilised membranes for 20 mins at 40°C before FSEC. Coloured traces are the same as part (a). (c) SEC analysis from DDM-purified protein with protein quality confirmed by SDS-PAGE analysis (in inset). (d) Crystals grown to full size after 3-4 days at 20°C in 0.1 M sodium citrate pH 3.5, 0.1 M lithium sulphate and 28% PEG 400 with 1% (w/v) heptyl-thiol-β-D-glucoside, 1% hexanediol and 1% Facade-EM.

The same crystal form was grown from the triple mutant using the crystallisation conditions previously described for the Leu296Met crystals (figure 4.6, part d). From this particular prep, the amphiphile detergent Facade-EM (Avanti Polar lipids) was also tested to optimise crystal growth. Due to the largest crystals growing in 22% PEG 400, crystal dehydration was carried out to minimise the formation of ice crystals when freezing the crystals under liquid nitrogen conditions before X-ray analysis.

One crystal dehydrated to 32% PEG 400 showed the best diffraction to 3.5 Å with the same unit cell dimensions as the wild type (table 4). Unlike the crystals only containing the methionine mutation, much more data (>360^{o}) was able to be collected from this triple mutant from different sections of the crystal showing uniform diffraction and resolution to help maximise the anomalous signal: noise ratio.

4.3.4 - Realignment of helix 10

The aim of collecting a high redundancy data set was to obtain a convincing anomalous peak signal from the selenium atom to confirm if its position aligned with residue 296 from the refined NhaA dimeric structure with the proposed realignment of TM 10. Molecular replacement was, therefore, carried out by using only the monomer as a search model from the refined dimeric structure, before and after the realignment of TM 10, as well as the published structure for consistency. The resulting electron density maps using the triple mutant data all showed the same overall size and shape as the previous wild type dimeric structure with the anomalous difference map of the selenium atom aligned with Met296 from TM 10 after realignment (figure 4.7). The nine other selenium atoms aligned with the methionine

residues from both the wild type monomer and dimeric structures, suggesting no other major structural changes in the other 11 helices apart from TM 10. This was further verified after structure refinement with the refined 3.5 Å wild type structure showing no visible differences in the helix positions, apart from TM 10 (table 4).

Table 4 – Data collection and refinement statistics of the NhaA triple mutant

Crystal	**Triple-mutant**
Beamline	I03
Detector	Pilatus 6M-F
Wavelength (Å)	0.9793
Space group	$P2_1$
Resolution (Å)	56.5-3.5 (3.54-3.50)[a]
Cell dimensions (Å)	***a***=115.8 ***b***=99.4, ***c***=140.2 □=90º, □= 97º, □=90º
Number of measured reflections	313,049
Number of unique reflections	37,951
Completeness (%)	94.6 (69.5)
Redundancy	8.2 (7.2)
I/□(I)	22.0 (1.4)
R_{merge} (%)	4.4 (111.4)
Refinement	
Number of reflections used	34,855
Number of atoms	11,331
R-factor (%)	27.2
R-free[b] (%)	30.4
R.m.s.d. from ideal values	
Bond lengths (Å)	0.010
Bond angles (°)	1.48
Ramachandran plot outliers[c] (%)	2.4

[a] Values in parentheses refer to data in the highest resolution shell
[b] Based on 5% of the reflections
[c] From Molprobity (Chen *et al.*, 2010)

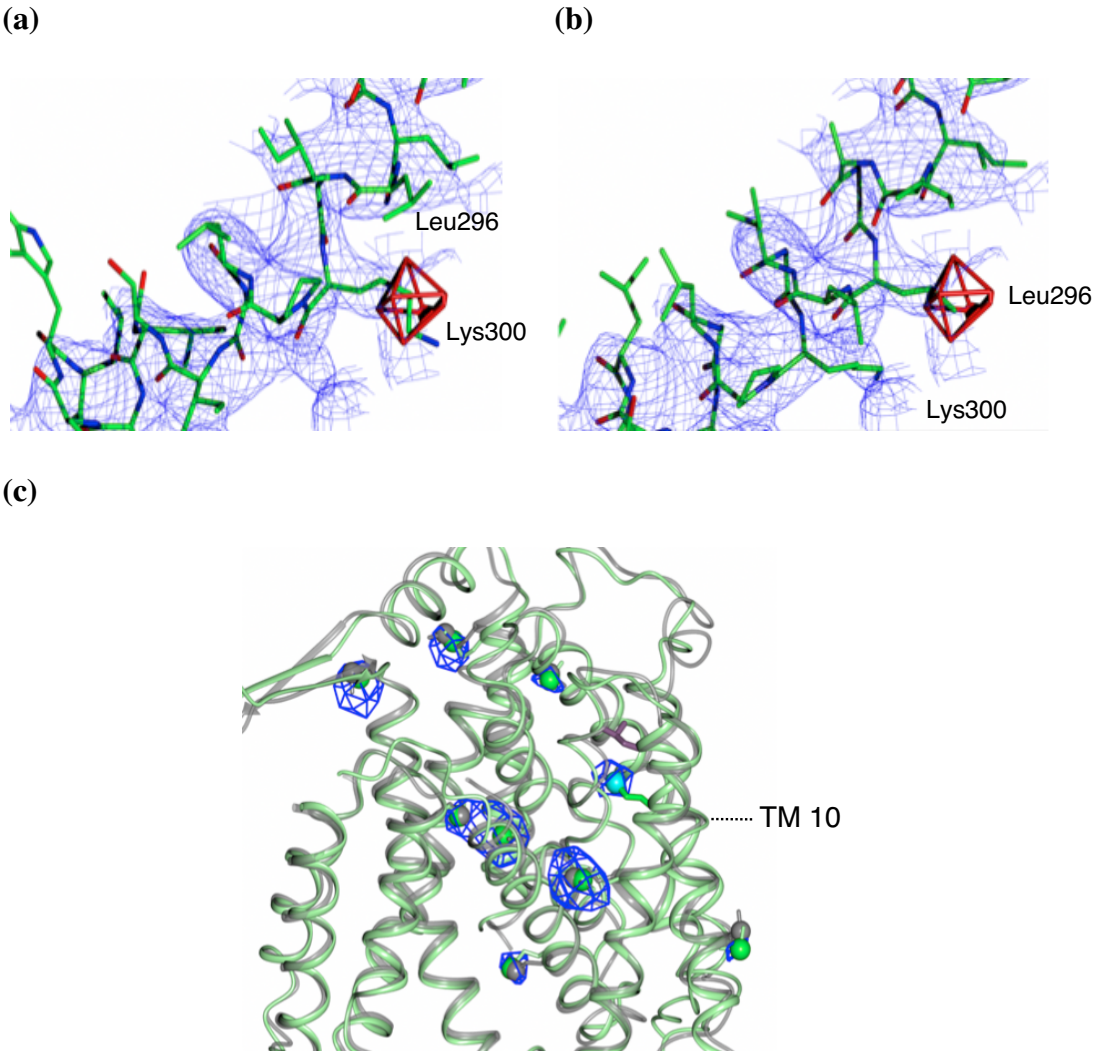


Figure 4.7 – Helix 10 realignment of NhaA. (a) Electron density of TM 10 with the 2mFo-DFc map shown in blue with the phases estimated from the NhaA wild type dimer structure before realignment of TM 10 (contoured at 1.5 σ). The anomalous difference map shown in red has been calculated from the SeMet-labelled triple mutant (contoured at 3.6σ) showing alignment of the anomalous selenium peak from Met296 with Lys300 (left) instead of Leu296. (b) The same electron density maps displayed as part (a) but with the phases estimated from the wild type structure with the predicted realignment of TM 10 starting from 287 to 313 (compared to 290 to 316 in published structure) with Leu296 aligned with an anomalous selenium peak. (c) Cartoon representation of the the final refined structure of the NhaA triple mutant (green) superposed onto the published structure (grey) showing a good overall alignment of the position of the helices (excluding TM 10) with anomalous difference map (same as part a) showing all methionine residues (green spheres) in the same corresponding positions as their anomalous selenium peaks (blue) including Met296 (cyan sphere) from the refined NhaA structure, which has moved 7 Å lower towards the cytoplasm compared to the original position of Leu296 (purple stick) in the published structure. Figures were prepared by Dr. Alex Cameron using ccp4mg (Potterton *et al*., 2004).

4.3.5 - New observed salt-bridge pair in the ion binding site

The realignment of TM 10 did not change the arrangement of the core domain showing a root mean standard deviation (r.m.s.d.) of 0.9 Å from 173 out of 179 Cα atoms between the NhaA dimer and published monomer. However, the position of Lys300 is now shown to be 9 Å closer to the ion binding site and within a salt bridge distance (2.7 Å) with the putative ion binding residue Asp163 (figure 4.8). We think that this Lys300-Asp163 salt bridge pair observed at low pH may be important to the transport mechanism, as mutagenesis of Asp163 is known to abolish transport activity (Inoue *et al.*, 1995). This is discussed in more detail in Chapters 5 and 6.

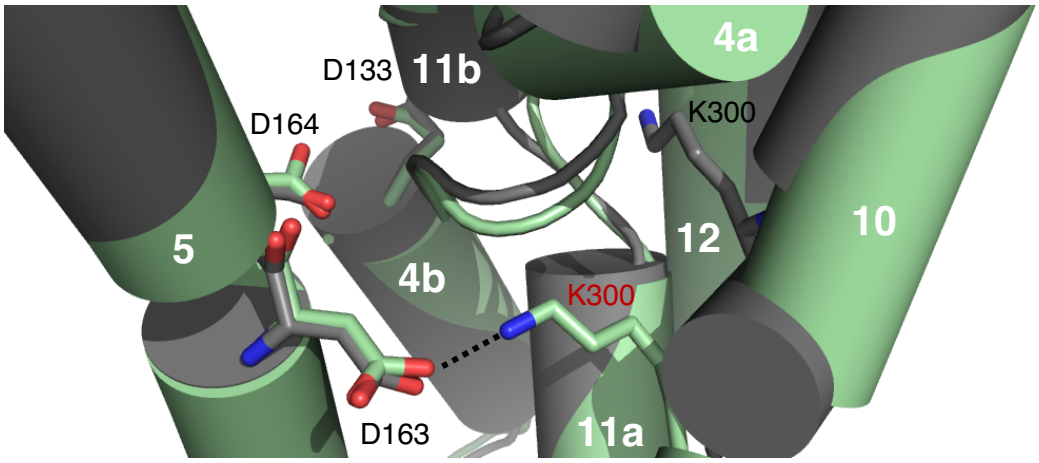


Figure 4.8 – Position of Lys300 after realignment of helix 10. Cartoon representation of the NhaA structure showing only the core domain viewed from the periplasmic facing side of the membrane with the published structure (grey) superposed onto the new NhaA structure (green). Important charged residues in the ion binding site are shown as sticks with D133 (TM 4), D163 and D164 (TM 5) in the same positions from both structures apart from K300 (TM 10). The new position of K300 in the new structure (labelled in red text) is proposed to form a salt-bridge interaction with D163. TM 3 has been omitted for clarity.

5.0 -

Structural studies of

Thermus thermophilus NapA

5.1 - Introduction

The work from this Chapter was carried out to see if a more naturally thermostable bacterial NhaA homologue could grow a better quality of crystal in an active pH condition compared to the thermostabilised NhaA mutant described in the previous chapter.

Prior to starting my PhD, NapA from *Thermus thermophilus* was selected as a suitable candidate for crystallisation. This CPA2 bacterial Na^+/H^+ antiporter was confirmed to be stable in a wide range of detergents, including LDAO, from FSEC analysis with a higher LDAO unfolding $T_{1/2}$ compared to NhaA wild type. Although the sequence identity between both proteins is less than 15%, the bacteria NapA genes, as expected, strictly conserve the same important residues localised in the ion binding site of NhaA (previously shown in figure 2.4), with similar mutant phenotypes confirming both to function using the same pH-dependent transport mechanism (Furrer *et al*., 2007). Interestingly, NapA from *T. thermophilus* shows 21% sequence identity to the human NHA2 isoform in the CPA2 family compared to 15% with *E. coli* NhaA (Xiang *et al.*, 2007). Thus, this justified crystallographic studies on NapA to complement the structural and biochemical work of NhaA.

Much better X-ray diffraction crystals were obtained from NapA purified in DDM compared to the thermostable NhaA mutant, which were grown at pH 9 where the protein is most likely to be in an active state. Before joining this project, a 3.5 Å data set had previously been collected using SAD from a selenomethionine-labelled

crystal. Unfortunately the anomalous signal from the selenium atom was not good enough to estimate the phases, which were needed to calculate an electron density map of the NapA structure using the electron density equation. Molecular replacement was also tried to determine the phases from this data set using the NhaA monomeric structure as a search model, however, this approach did not work most likely due to the low sequence identity between NapA and NhaA. Crystal optimisation was also continued for native NapA crystals to try and collect a data set of the same quality as the collected selenomethionine data set, which could both be used to estimate the phases by isomorphous replacement. However, NapA crystals were often fragile due to the high solvent contents, making them more prone to radiation damage and resulting in poor X-ray diffracting resolution. Another problem was that DDM-purified NapA protein crystallised in the lowest crystal symmetry spacegroup (P1), which meant that when collecting data from a crystal, it had to be rotated at least 360^{o} to obtain all the structural information needed from the diffraction patterns to solve the structure. Crystal damage during data collection therefore became an ongoing problem.

The main task was to improve the NapA data to estimate phases from primarily NapA crystals derivatised with a different type of heavy atom (HA) bound ideally on the membrane protein surface. In this work we also investigated the use of a small micelle detergent to try and further improve crystal quality and X-ray diffracting resolution.

5.2 - Materials and methods

Experiment protocols not described in this chapter are previously described in both Chapters 3 and 4 with only further experiments or modifications given in this Chapter.

5.2.1 – NapA solubilisation and purification

Buffers

Solubilisation buffer	1 x PBS, 150 mM NaCl and 1% $DDM_{L\alpha}$
Wash buffer	1 x PBS, 150 mM NaCl, 10 mM imidazole and 0.1% $DDM_{L\alpha}$,
Detergent exchange	1x PBS, 150 mM NaCl, 250 mM imidazole and 0.6% buffer NM
Dialysis buffer	20 mM Tris-HCl pH 7.5, 150 mM NaCl and 0.5% NM
Crystallisation buffer	20 mM Tris-HCl pH 7.5, 150 mM NaCl and 0.45% NM

Crude membranes were diluted in solubilisation buffer to a final protein concentration of 3.5 mg/mL. After 1 hour stirring at 4°C, solubilised membranes were isolated by ultracentrifugation at 140,000*g* for 1 hour to remove unsolubilised material in the pellet. Solubilisation efficiencies were calculated from RFU readings taken before and after ultracentrifugation. Solubilised membranes were stirred with 1 mL of Ni-NTA slurry per 1 mg GFP pre-equilibrated with wash buffer. After 2 hours the slurry was

transferred to a glass Econo-Column (BioRad) and washed with 20 column volumes (CVs) of wash buffer containing increasing amounts of imidazole: 20, 30,40 and 50 mM. The membrane protein-GFP fusion was eluted using 50 mL of detergent containing 250 mM imidazole. Cleavage of the GFP-His_8-tag was followed by addition of equimolar His_6-TEV protease during overnight dialysis in 1.5 L of dialysis buffer. Reverse IMAC was carried out using a 5 mL Ni-NTA HisTrap HP column pre-equilibrated in dialysis buffer with eluted protein collected in the flowthrough. Protein was concentrated to 500 μL using 100 kDa MWCO Amicon Ultra-15 centrifugal concentrators (Millipore) at 3,000*g* with 10 min intervals to remove any precipitation during concentrating. SEC was followed using a Superdex 200 10/30 gel filtration column pre-equilibrated in crystallisation buffer. The monodisperse protein peak was collected and concentrated with 100 kDa molecular weight cut-off concentrators (Vivaspin) to 10 mg/mL. Final protein concentrations were measured by a BCA assay and purity assessed from 10 μg of protein by SDS-PAGE analysis. Pure protein was used for crystallisation trials or flash-frozen in liquid nitrogen for storage at -80°C.

5.2.2 – Mercury derivatisation

Mercury (Hg) derivatisation was carried out to provide phasing information during structure determination. Prior to SEC, 500 μL of concentrated protein was incubated on ice with 0.5 to 2.5 mM mercury acetate at different time periods (30 min to 2 hours). Free unbound Hg was removed during SEC. Hg-derivatisation was confirmed either by the CPM-based unfolding assay or by in-gel fluorescence analysis by mixing 4 μg of CPM dye (Sigma) to 1 μg of native and Hg-derivatised pure protein of the same concentration (10 mg/mL) with the SDS-PAGE gel cassette run in the dark. In-

gel fluorescence analysis was carried out using an LAS-1000-3000 charged-coupled device (CCD) imaging system (Fujiflm) at a fluorescence emission/excitation spectrum of 463/387 nm. Absence of a fluorescent band confirmed Hg-derivatisation.

5.2.3 – Size exclusion chromatography light scattering analysis

Purified NapA protein was loaded onto a Superdex 200 10/300 size exclusion column (GE Healthcare) coupled to a Viscotek TDAmax tetra detector array (Malvern) with GPCmax solvent pump and integrated auto-sampler, using the OmniSEC software for data analysis. The SEC-UV/LS/refractive index (RI) system was equilibrated in 20 mM Tris-HCl pH 7.5, 150 mM NaCl and 0.03% DDM at a flow rate of 0.3 mL/min. Standard gel filtration molecular weight markers were used for calibration and all proteins were analysed under the same experimental conditions. Data was collected from the RI, Right Angle LS (RALS) and UV280 detectors. The oligomeric state of NapA from the NapA-detergent micelle was calculated using methods described previously (Slotboom *et al*., 2008).

5.2.4 – Crystallisation screening and optimisation

The MemGold™ 96-well crystallisation conditions provided by Hampton Research (Appendix I) were used by mixing 200 nL of protein with 200 nL of reservoir solution onto 96-well sitting drop plates using a Mosquito robot (TTP labtech). Plates were incubated at 4°C and 20°C.

Crystal optimisation trials were set up by the hanging drop vapour method manually using 24-well VDXm plates (Hampton Research) mixing 1 μL of purified protein

with 1 μL of mother liquor on a siliconised glass cover slide and equilibrated over 500 μL of equilibrated reservoir solution at 20°C.

The optimised crystallisation conditions for NapA purified in DDM (personal communication with Dr. David Drew) were 0.05 M magnesium acetate, 0.05 M glycine pH 9, 22- 36% PEG 400. From the initial crystallisation screening plate, the optimised crystallisation conditions for NapA purified in NM were 0.001 M zinc sulphate, 0.05 M HEPES pH 7.8, 24% PEG 600 (MemGold condition- 2.36; shown in Appendix I).

5.2.5 - Crystal dehydration

Cover slides containing equilibrated crystal drops were transferred, sequentially overnight, to wells containing the same reservoir solution with a 2% increment of PEG 400 up to a final concentration of 34%. Crystals were flash-frozen from different PEG concentrations. Crystal annealing was also carried out on one native crystal of the NapA triple mutant on the ID23_2 beamline at the ESRF synchrotron.

5.2.6 - Data processing and structure determination

Data collection

X-ray crystal screening and data collection were carried out on synchrotron beamlines at Diamond Light Source (Didcot, United Kingdom) and at the ESRF (Grenoble, France). As our main aim was to estimate the phases for structure determination of NapA, data collections of mercury- and selenomethionine- derivatised NapA crystals were conducted close to their absorption edge wavelengths, to collect anomalous

scattering data from the two different heavy atoms for phase estimation. Two data sets were also collected from native NapA crystals for phasing purposes. All data sets used to solve the NapA structure were collected from a number of synchrotron visits at different beamlines, with their data collection statistics summarised in table 5; presented in the results section.

The NapA crystals used during data collection, were grown in three different spacegroups. Thus, one can collect fewer reflections for a complete data set if crystal symmetry is present. However similarly to the NhaA project, all data sets were collected at high redundancy to get as accurate a measure of the observed intensities as possible as well as the anomalous scattering from the heavy atoms. The data sets collected from the triclinic crystals (P1 spacegroup symmetry) were used to estimate the phases. The monoclinic ($P2_1$ spacegroup) and orthorhombic ($C222_1$ spacegroup) crystals were used for multi-crystal averaging during density modification.

Estimation of Phases

The first good quality data set was collected from a mercury-derivatised crystal of the NapA cysteine double mutant. The data set was processed using the Xia2 pipeline to XDS to a resolution of 4 Å, with the same P1 spacegroup and similar cell dimensions to the previously collected NapA wild type data set (given in table 5, section 5.3.2). The position of the two mercury sites bound to the two mutated cysteine residues in the double mutant were located based on their anomalous scattering determined from an anomalous difference patterson map using the program RSPS (Knight, 2000). Heavy-atom refinement of the two mercury sites was followed using the program SHARP (Bricogne *et al*., 2003), of which initial phases were estimated using the

single anomalous dispersion (SAD) function, to calculate an electron density map using the electron density equation. The quality of the electron density regions in the maps were improved with solvent flattening using SOLOMON (Abrahams and Leslie, 1996) to reduce the noise signal from the solvent content, representing 71% of the asymmetric unit cell.

A native and mercury-derivatised data set collected from crystals of the NapA triple mutant were later merged and scaled using SCALEIT in the CCP4 software package with the previous double mutant data set as they were all isomorphous; sharing the same P1 spacegroup and unit cell dimensions. With the phasing from the double mutant, the third mercury atom site from the triple mutant was located from an anomalous difference fourier map. In SHARP, heavy atom refinement was repeated to refine the positions of all three mercury atom sites using the observed structure factor amplitudes of the native and two mercury-derivatised data sets (double and triple mutant), where differences would be due to the presence of the heavy atoms. All three data sets further improved the phase estimates by multiple isomorphous replacement with anomalous scattering (MIRAS) based on the positions of the refined mercury atom sites. An isomorphous data set from a selenomethionine-derivatised crystal of the NapA single mutant, processed to 3.5 Å was also merged and scaled together with the native and mercury-derivatised data sets. The selenium sites were located from anomalous difference fourier maps, which further improved the phase estimates.

After incorporating all these data sets containing two different heavy atom derivatives, the updated electron density map after solvent flattening showed an improvement compared to the previous maps from only using the double mutant data

set. Based on the number of mercury atom sites located in the asymmetric unit cell (8 from the double mutant and 12 from the triple mutant), four NapA monomers were present.

Density modification

To improve interpretability of the rod-shaped density regions from the helices for model building, density modification was the next step. Density modification of the electron density map was carried out in DM (Cowtan, 1994) with solvent flattening, histogram mapping and non-crystallographic symmetry (NCS) averaging, of which the latter was particularly useful in improving the quality of electron density by averaging the density across each of the four NCS-related monomers. Four NCS operators (one per monomer) were determined in the program O (Jones and Kjeldgaard, 1997) based on the mercury atom site positions, which were needed to locate the positions of three of the monomers relative to the position of one reference monomer. Additionally, an averaging NCS mask was also used to define a set electron density volume of each monomer to enclose density averaging between the four NCS operators. For this, a mask was made also in O based on the NhaA monomer structure. The electron density was becoming more interpretable with iterative cycles in DM whilst simultaneously refining the shape of the mask.

Density modification with multi-crystal averaging

Two more data sets were also collected from two crystals processed in two different spacegroups; $P2_1$ from a mercury-derivatised double mutant crystal and $C222_1$ from a native triple mutant crystal. These two additional data sets were used for multi-crystal NCS averaging in DMMULTI (Cowtan, 1994) using the previously determined NCS

operators from the triclinic crystals (P1) with new NCS operators determined for the monoclinic ($P2_1$) and orthorhombic ($C222_1$) crystals, using the most recently refined mask of the monomer. The NCS operators for the $P2_1$ and $C222_1$ data sets were determined by carrying out MR in PHASER using the refined electron density of one monomer as a search model to give the orientation and positions of the search model to move between the monomer units in the asymmetric unit cell. Similarly, both the electron density maps and shape of the mask were improved by iterative cycles of multi-crystal averaging. Once no visible improvements could be made to the maps, model building was followed.

Model building and refinement

Model building was carried out in both O and COOT and was facilitated by the anomalous peak positions of mercury and selenomethionine sites from anomalous difference maps. The electron density of each monomer subunit contained 13 rod-shaped helices, showing four monomers crystallised as two side by side dimers in the same orientation, as if parallel to the membrane plane, which were operated by a crystallographic two-fold symmetry axis. The overall size and shape of the density from the NapA dimer was similar to the density from the wild type NhaA dimer. Model building and structural refinement were interspersed to improve the final refined structural model.

Refinement of the model was carried out in PHENIX using the observed structure factor amplitudes from the 3.0 Å $C222_1$ data set, whilst monitoring the R_{free} and R_{work} values. The refinement cycles were carried out in a similar manner to the NhaA project, with restraints first put on the atomic XYZ coordinates and individual B-

factors. Secondary structure restraints also showed an improvement to the refined model. The final refinement statistics are shown in the results section of Chapter 5. Since the data sets used to solve the final NapA structural model were all from cysteine mutant crystals, to check that the mutations did not introduce any structural effects, MR was carried out using the refined structure as a search model with the previously collected selenomethionine data set of the wild type. The electron density map produced from the wild type data set was overall very similar to the final density map from the refined structure, of which its anomalous difference map also showed anomalous peaks from the selenium atom sites in the same positions as their corresponding methionine residues.

5.2.7 – Structural analysis with NhaA

Sequence alignment of NapA and NhaA with their structures solved in this thesis were carried out using the PROMALS3D server with manual adjustment in JalView. Superpositions were carried out from only the TMs of NapA and NhaA using the DaliLite server (Holm and Park, 2000). Lsqman was also used when the core and dimer domains of the two structures were superposed separately such that all matching Cα pairs were less than 3.8 Å apart after the superposition (Kleywegt and and Jones, 1994).

5.3 - Results

5.3.1 – Mercury derivatisation of NapA

To help with the phasing of NapA, more focus was chosen to be put on improving the X-ray diffraction of HA-derivatised crystals using mercury, which is commonly used for phasing by binding to the sulfhydryl group of cysteine residues (Morth *et al.*, 2006). However, NapA contains no cysteine residues. Using the NhaA structure as a model and pair-wise sequence alignment with NapA, two cysteine substitution sites were chosen at methionine-20 and valine-166, predicted to be accessible to a free mercury atom in a detergent micelle based on their corresponding residue positions in NhaA (figure 5.1). To help improve the accuracy of the mercury anomalous signal, a double mutant was also made combining Met20Cys and Val166Cys and a triple mutant with the same two mutation sites with a third at valine-326.

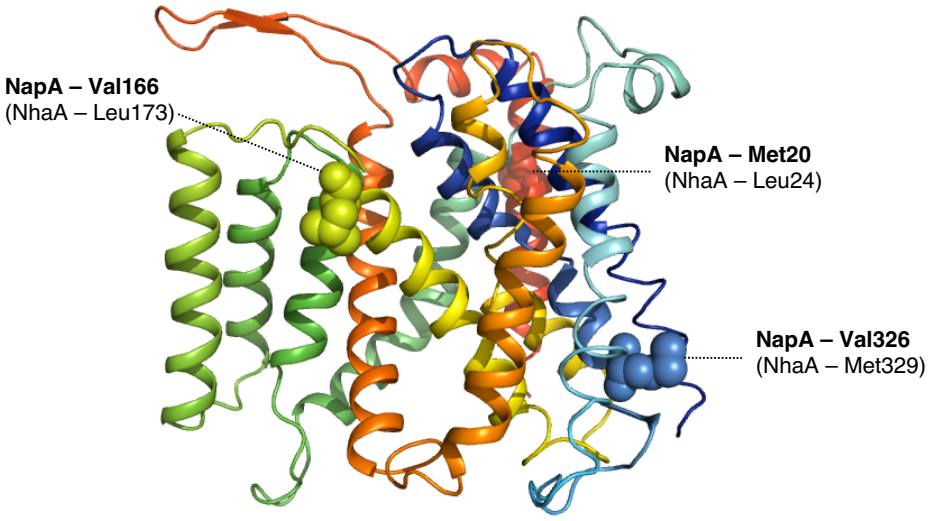


Figure 5.1 – Predicted structural position of three cysteine substitution sites of NapA based on NhaA. NhaA structure viewed in the membrane plane showing three residues as spheres, of which their corresponding residues in NapA are chosen to be mutated to cysteine for mercury-derivatisation. With pair-wise alignment using JalView, Leu24 from TM 1 in NhaA aligns with Met20 in NapA (red), Leu173 from TM 5 with Val166 (yellow) and Met329 from TM 11a with Val326 (blue).

The NapA wild type and cysteine mutants expressed to less than 3 mg.L^{-1} using the standard conditions; described in Chapter 3. The MemStar system, however, boosted expression levels to ~ 10-20 mg.L^{-1}, which were much more suitable levels for large-scale expression and purification.

FSEC analysis was carried out from DDM and LDAO-solubilised extracts from crude membranes to screen detergent stability. All mutants were monodisperse in both detergents and suitable for purification. Even though the Met20Cys showed a slightly broader LDAO trace with a small aggregation peak, it was still considered stable enough for purification in the mild DDM detergent (figure 5.2).

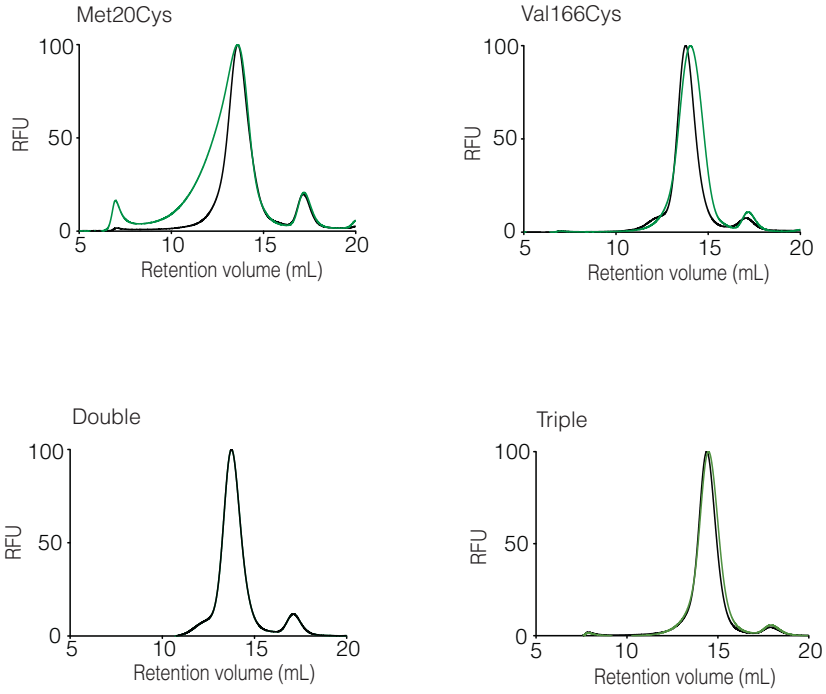


Figure 5.2 – FSEC analysis of the NapA cysteine mutants. FSEC traces of detergent-solubilised membranes in DDM (black) and LDAO (green) from Met20Cys, Val166Cys, double (Met20Cys and Val166Cys) and triple (Met20Cys, Val166Cys and Val326Cys) mutants.

It was decided to first purify and crystallise Met20Cys and Val166Cys in DDM with/without mercury derivatisation. Pure protein of a stable quality was obtained from both mutants with bound mercury confirmed by both in-gel fluorescence and the CPM-based unfolding assay (figure 5.3) as well as by mass spectrometry (St. Andrews University). As expected, both mutants grew the wild type crystal form using the optimised crystallisation conditions previously determined for the wild type at pH 9.

After crystal optimisation trials, the resolution of crystal diffraction from both mutants could not be improved higher than 5 Å. Furthermore, the mercury-derivatised crystals were more fragile, which made it extremely difficult to collect a complete data set with a high redundancy without radiation damage. Crystal optimisation trials were also attempted on NapA wild type to obtain a high resolution native data set. However, the resolution similarly could not be improved higher than 4 Å due primarily to radiation damage during data collection.

The next approach chosen was to test how well the crystal quality and resolution of X-ray diffraction of NapA could be improved when purified using a small micelle detergent.

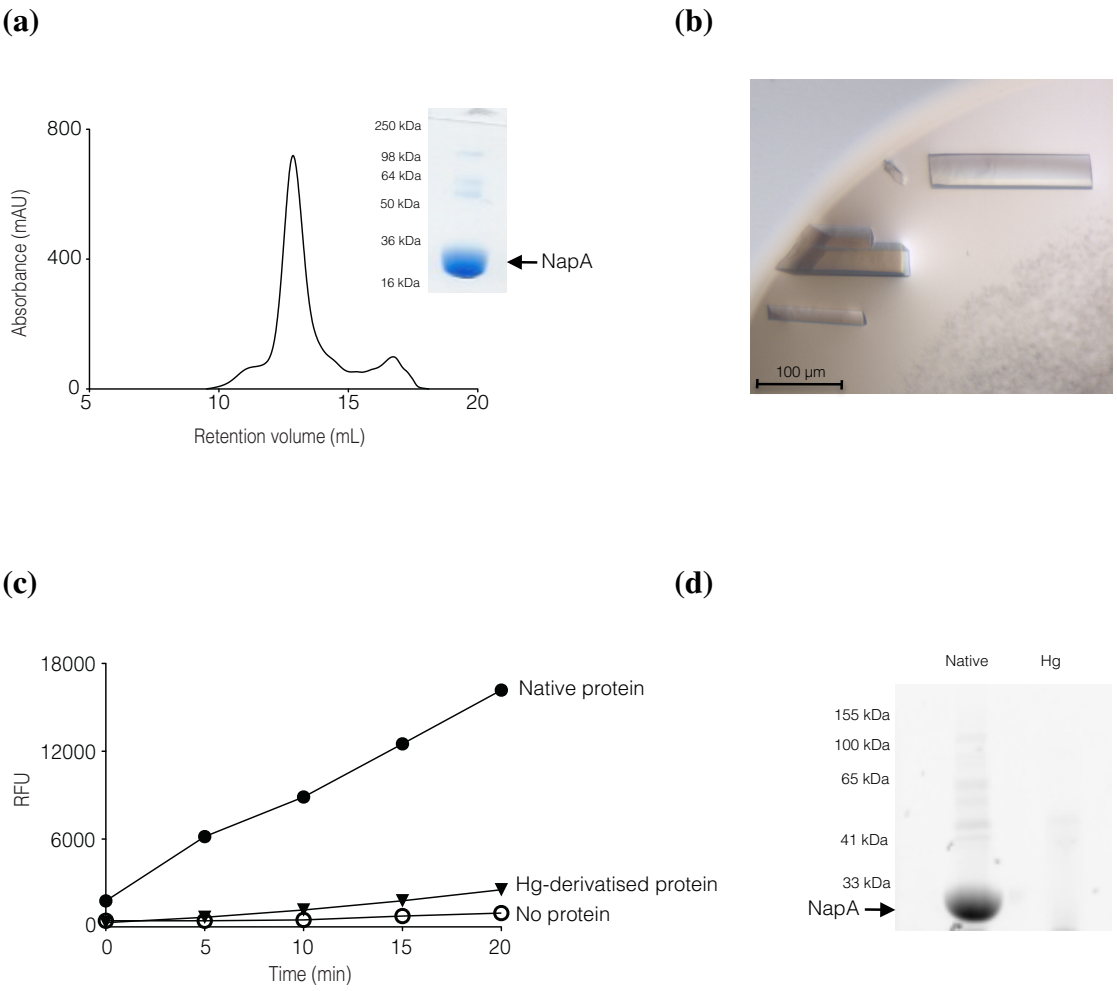


Figure 5.3 – Characterisation of NapA. All cysteine mutants produced the same results as shown here with Met20Cys. (a) SEC profile of purified NapA protein in DDM with the purity of 10 mg.L^{-1} protein used for crystallisation illustrated by the Coomassie-stained SDS-PAGE gel (inset). (b) Crystal grown after 3-4 days at 20°C using the wild type crystallisation conditions. Two approaches were taken to monitor Hg-derivatisation from purified protein using the thiol-specific CPM dye: (c) the CPM-based unfolding assay; filled circle, native protein with no Hg added; filled triangle, Hg-derivatised NapA protein; open circle, no protein addition (only CPM dye) and (d) In-gel fluorescence analysis by mixing 1 µg protein with 4 µg CPM dye before loading onto the SDS-PAGE gel. Both methods produced a fluorescent signal from only the native NapA protein.

5.3.2 - Crystal optimisation using a small micelle detergent

The detergent nonyl-maltoside (NM) was chosen that only differed from DDM by the length of its hydrophobic chain consisting of nine CH_2 groups compared to twelve of DDM. NapA wild type was first purified to confirm if this protein could be crystallised in NM. From 5 L cultures, isolated crude membranes were though solubilised in DDM with a detergent exchange into NM beginning from the first purification step on the Ni-NTA column. Purification in NM showed the same monodisperse SEC peak as in DDM with the same purity of concentrated protein.

Initial crystal hits were made from the MemGold crystallisation screen set up at 20°C, with 4 Å X-ray diffraction from crystals grown in 0.001 M zinc sulphate, 0.05 M HEPES pH 7.8, 24% PEG 600. This result was encouraging since these first NapA crystals already showed an improvement in the resolution before crystal optimisation compared to the NapA crystals purified in DDM. While crystal optimisation trials were carried out on the wild type, the cysteine mutants were also to be purified in NM.

Heat-FSEC analysis was first carried out from crude membranes of the cysteine mutants to confirm the mutations did not affect protein stability. As expected, NapA wild type showed the same monodispersity traces in LDAO and DDM from FSEC and heat-FSEC analysis; therefore was used as a control (figure 5.4). Met20Cys was the only mutant showing an aggregation peak in both detergents, therefore was excluded from future experiments.

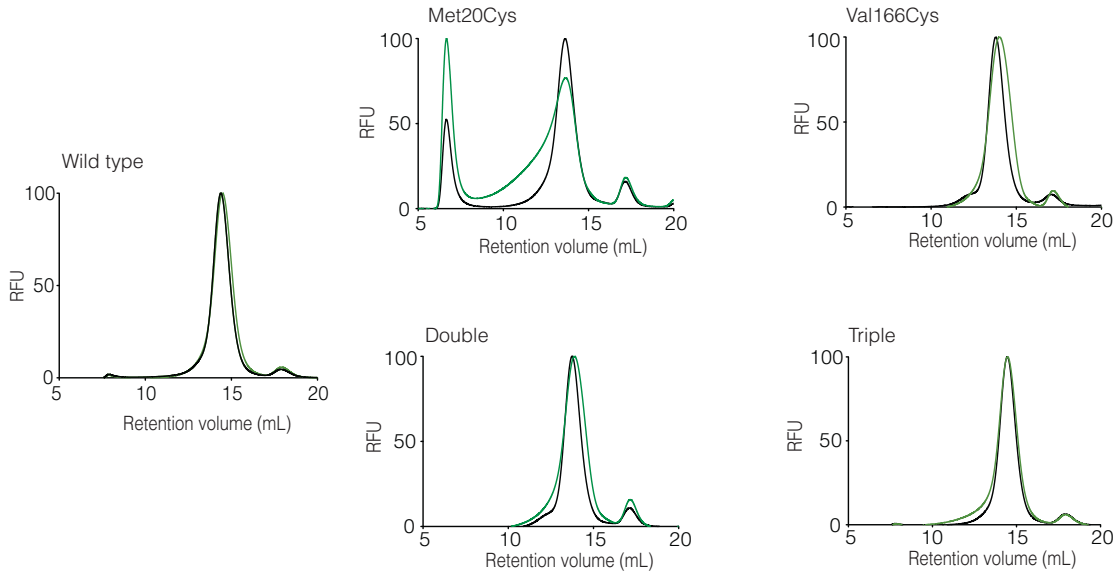


Figure 5.4. – Heat-FSEC analysis of NapA wild type and cysteine mutants. DDM- (black) and LDAO- (green) solubilised membranes were incubated for 20 min at 40°C before FSEC analysis. Only the Met20Cys mutant showed an aggregation peak in both detergents suggesting it to be less thermostable compared to the wild type and other mutants.

Purification of the single Val166Cys, double and triple mutants were carried out in NM as the wild type. To help reduce financial detergent costs, 10 L cultures were routinely prepared for purification of each mutant, with one half of the purified protein sample remaining native and the second half for mercury-derivatisation. The same crystal form was grown from all mutants using the same wild type crystallisation conditions. During crystal optimisation trials, crystals appeared overnight with maximum size reached after 3–4 days (figure 5.5). Interestingly, a second new crystal form was also obtained only from the native triple mutant with 0.025% dichloromethane added as an additive. Selenomethionine-labelled protein of Val166Cys was also purified in NM for crystallisation to try and obtain more anomalous data from a second HA to improve the phasing information. Crystal dehydration was finally carried out before flash-freezing the crystals, primarily to

minimise the formation of ice crystals in the cryo-loop but also to test if the quality of crystal diffraction could be improved (Heras and Martin, 2005).

(a) **(b)**

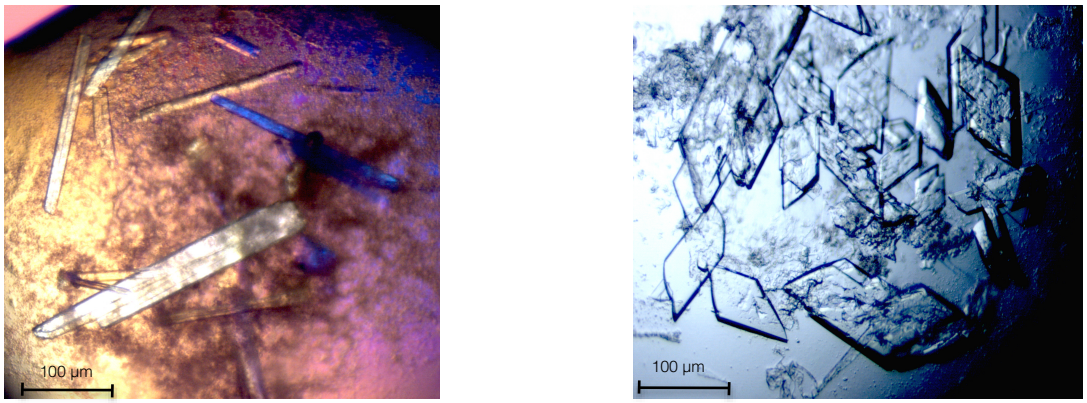


Figure 5.5 – NapA crystal forms grown from protein purified in NM. (a) Crystal form grown from the wild type and all cysteine mutants (native and Hg-derivatised) in the wild type crystallisation conditions. (b) A second crystal form grown from only protein of the native triple mutant with 0.025% dichloromethane present in the crystal drop.

The cysteine mutants produced much better X-ray diffraction than the wild type, which were mostly triclinic, but very occasionally crystals with different space groups were observed (table 5). The best quality data sets were collected from crystals of each cysteine mutant which were dehydrated between 26-32% PEG 400. The highest resolution data set was collected from the second native crystal form of the triple mutant which was orthorhombic. A further improvement in the diffraction spots from this crystal was also observed after reannealing on the beamline.

Table 5 – Data sets collected from NapA wild type and cysteine mutants

SPACEGROUP	$C222_1$	P1					$P2_1$
	Native	Native	Hg	Hg	Se	Se	Hg
	Triple	Triple	Double	Triple	Single	Wild Type	Triple
Beamline	ID23_2	ID23_2	I03	ID23_2	ID23_1	I24	I03
Detector	CCD	CCD	Pilatus 6M-F	CCD	CCD	Pilatus 6M-F	Pilatus 6M-F
Wavelength (Å)	0.873	0.873	1.008	0.873	0.979	0.978	1.005
Resolution (Å)	63.9 – 3.0 (3.06-2.98)[a]	76.6 - 4.0 (4.09-3.99)	99.0 - 4.0 (4.15-4.02)	92.0 - 4.9 (5.0-4.87)	32.0 - 3.5 (3.62-3.52)	30 - 3.7 (3.76-3.70)	102.0 - 4.4 (4.56-4.44)
Cell dimensions a, b, c (Å) □, □, □(°)	73.8 82.2 191.8 □= 90.0 □ = 90.0 □= 90.0	79.4 95.5 105.2 □= 77.9 □ = 76.0 □= 80.9	78.8 95.3 103.8 □= 77.5 □ = 75.7 □= 80.3	77.2 95.3 101.7 □= 77.1 □ = 77.8 □= 79.5	79.5 95.1 104.2 □= 77.7 □ = 76.3 □= 80.8	a = 79.5 b = 93.7 c = 106.0 □= 77.6 □ = 76.2 □= 81.0	55.0 204.1 64.2 □= 90.0 □ = 98.4 □= 90
Number of unique reflections	12265 (918)	24327 (1759)	22725 (1653)	12636 (925)	34726 (2613)	30528 (1594)	8456 (662)
Completeness (%)	99.9 (100)	98.3 (97.3)	97.8 (95.3)	98.8 (98.1)	97.7 (97.8)	99.3 (99.1)	96.9 (96.2)
Redundancy	11.0 (9.3)	4.6 (4.6)	7.0 (6.8)	8.6 (8.6)	15.3 (15.7)	3.5 (3.5)	3.3 (3.4)
I/□(I)	18.8 (2.8)	13.6 (2.1)	11.2 (2.3)	12.1 (3.6)	19.8 (3.0)	21.7 (1.4)	5.9 (2.9)
R_{merge}	8.3 (95.0)	6.6 (85.8)	7.5 (98.4)	10.2 (67.8)	7.5 (115.6)	11.3 (99.7)	7.8 (98.2)
Phasing power[b] isomorphous anomalous			 0.59 0.75	 0.63 0.76	 2.29 2.11		

[a] Values in parentheses refer to data in the highest resolution shell
[b] Calculated to 3 Å; Phasing power = rms ($|F_H|/|F_{PH} - |F_P + F_H||$)

5.3.3 - Structure determination of NapA

As shown from table 5, more data sets were from the triclinic NapA crystals, which were native, mercury and selenomethionine-derivatised. These were first chosen to estimate the phases using MIRAS. Due to the timing of when these data sets were collected, the mercury-derivatised double mutant was first used to estimate the phases based on its anomalous signal from the two mercury atoms. Four NapA monomers were predicted in the asymmetric unit cell where 8 mercury atom sites were located in SHARP. A NapA dimer was later confirmed from the triple (both native and Hg-derivatised) and single (selenomethionine-derivatised) mutants, which were incorporated in SHARP to further improve the phases by both anomalous scattering and isomorphous replacement.

Due to a 71% solvent content in the unit cell calculated by the Matthews coefficient in the CCP4 software package, solvent flattening was first used to improve the initial electron density maps in SOLOMON. Density modification was then applied by firstly NCS averaging in DM using the four NapA monomeric molecules as operators, based on the HA positions using an initial mask made by overlaying the NhaA monomer onto the solvent flattened electron density maps using O. Both the electron density maps and the shape of the mask were improved during cycle rounds of density modification still only using the triclinic data of the cysteine mutants. Once the electron density within the refined mask could not be improved further, the maps were used as a search model for MR against the mercury-derivatised $P2_1$ data set and the native $C222_1$ data set, to improve the quality and particularly resolution of the maps given the latter data set was collected at the highest resolution of 3.0 Å. Further density modification cycles were repeated using DMMULTI to carry out multi-crystal

NCS averaging using the data sets from all the cysteine mutants with operators calculated from the refined mask across the three different spacegroups (P1, $P2_1$ and $C222_1$). The resolution quality of the final refined electron density maps were good enough to see two NapA monomers crystallised as a side by side dimer in the same orientation parallel to the membrane plane, which are operated by a crystallographic two-fold symmetry axis.

Model building of the NapA monomeric structure into the electron density map was helped from the anomalous difference maps of the 3 mercury and 5 selenium atom sites in the monomer. To confirm the structural model agreed with the observed structure factors, the 3.0 Å native $C222_1$ data set was used during structural refinement (table 6).

Table 6: Refinement statistics of NapA

Spacegroup	**$C222_1$**
Number of reflections used (non-anomalous)	22,911
Number of atoms	2,824
R-factor (%)	22.3
R-free[a] (%)	24.8
R.m.s.d. from ideal values	
Bond lengths (Å)	0.008
Bond angles (°)	1.151
Ramachandran plot outliers[b] (%)	0.8

[a] Based on 5% of the reflections
[b] From Molprobity (Chen *et al.*, 2010)

As a final confirmation that the final refined NapA structure did represent a physiological conformation in a pH-activated state, the previously collected 3.5 Å wild type data set collected from selenomethionine crystals grown at pH 9 was molecular replaced with the NapA model. At a resolution of 3.7 Å, we could see no clear structural differences between the positions of the selenium peaks in the anomalous difference maps from the wild type and the methionine positions from the refined NapA structure.

5.3.4 - Outward-facing state of NapA

The size and shape of the NapA dimer is similar to the solved structure of the NhaA dimer. To confirm this state was physiological in NapA, SEC-LS analysis was carried out from purified protein of the wild type and triple mutant to calculate the protein molecular weight in a detergent micelle (figure 5.6). From the wild type and triple mutant, the molecular weight was calculated to be 80 kDa, confirming the presence of a NapA dimer.

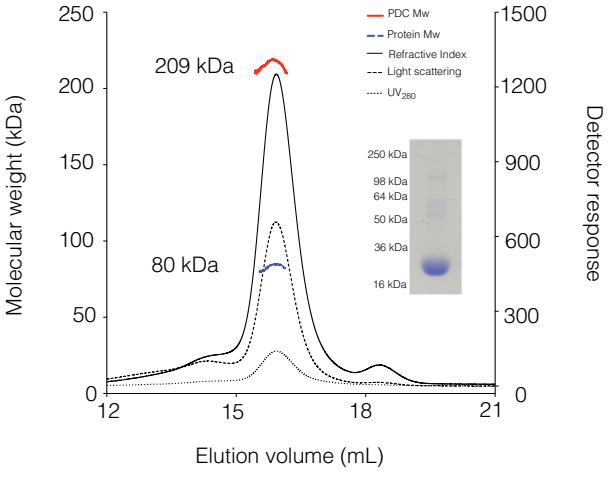
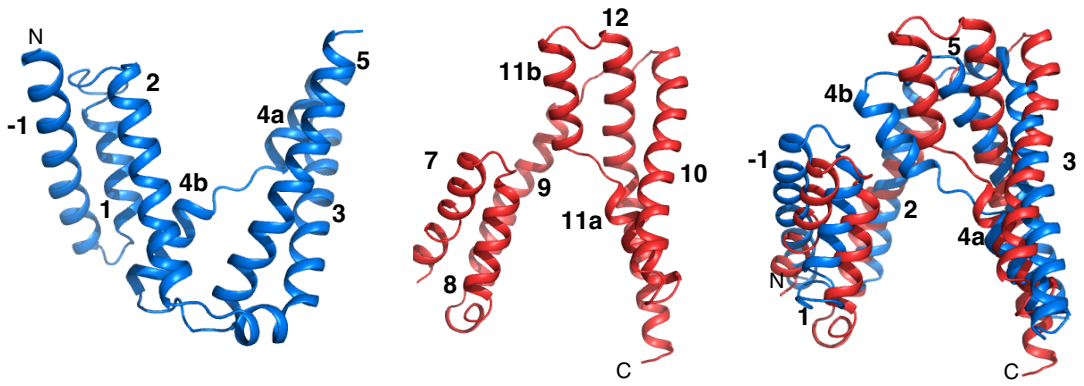


Figure 5.6 – SEC-LS analysis of NapA wild type. The molecular weight corresponding to the NapA-detergent complex (red line) and NapA only (blue line) is as shown (the predicted Mw of NapA is 40.8 kDa) and the purity of the protein used for analysis is illustrated by the Coomassie-stained SDS-gel (inset). The triple NapA mutant showed the same result as the wild type.

The crystal structure shows two NapA monomers positioned side by side with a similar arrangement of the helices compared to the NhaA dimer. Structure based alignment shows the NapA monomer to have an extra helix at the N-terminus, which gives the membrane protein an N_{OUT} - C_{IN} topology. To facilitate structural comparisons between the 13 helices of NapA with the 12 from NhaA, this extra helix from NapA is referred to as TM -1 (figure 5.7). The structure of the NapA monomer shows two pseudosymmetry-related inverted repeats formed by TMs 1 to 5 and TMs 7 to 12, which intertwine to form a separate core domain and a dimer (interface) domain linked together by TM 6 (figure 5.8).

Figure 5.7 – Structure based alignment of NapA and NhaA. The coloured helices are depicted as coloured filled boxes from red at the N-terminus to blue at the C-terminus as in Figure 5.8. Aspartate residues, D156/157 (NapA) and D163/164 (NhaA), likely to coordinate a sodium ion are labelled with a red square. Ionisable residues, E333/K305 (NapA) corresponding to D133/K300 (NhaA) are labelled with a black square. Sequence alignment was carried out using the PROMALS3D server with the NhaA and NapA structures and manual adjustment in JalView. Sequences are displayed in ClustalW format.

(a)



(b)

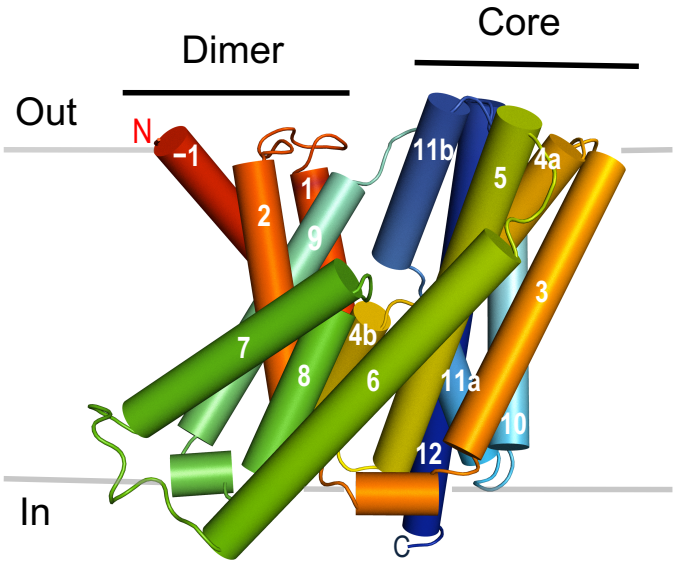


Figure 5.8 – NapA topology and structure. (a) From left to right, two inverted topology repeat units of TMs -1 to 5 (blue) and TMs 7 to 12 (red) are related by a pseudo two-fold axis of symmetry shown by the first repeat overlayed on the second after being rotated 90° perpendicular to the membrane. (b) Cartoon representation of the NapA structure as viewed in the plane of the membrane depicted in grey.

The dimer domains of the two monomers show a buried extensive surface area of 1800 Å^2 with tight hydrophobic helix-helix interactions shown mainly between TM −1 from one monomer and TM 7 from the other monomer (figure 5.9).

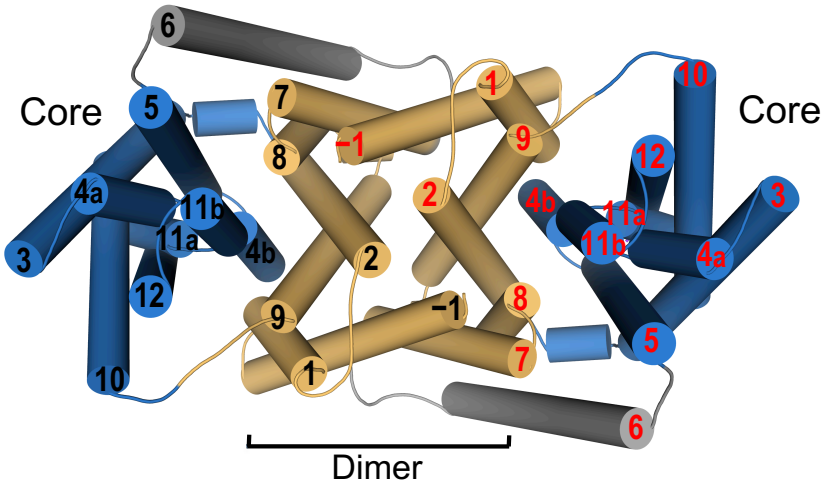


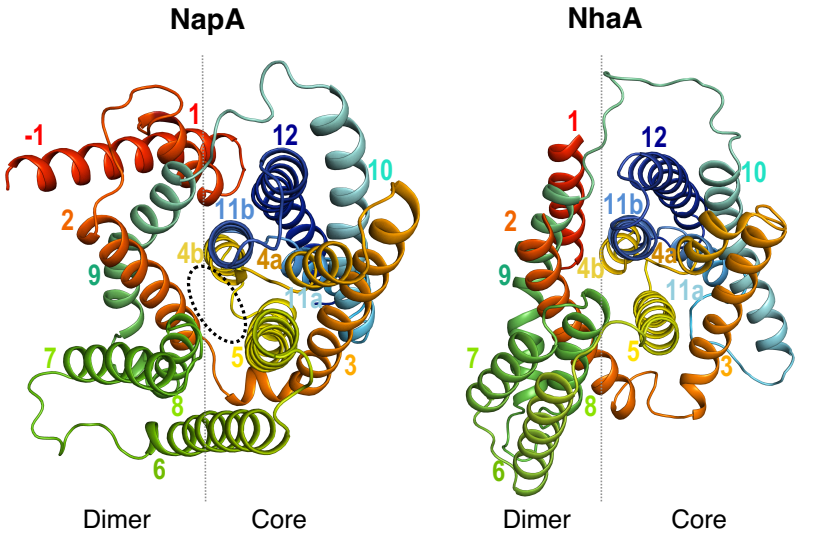
Figure 5.9 – Structure of the NapA dimer. Cartoon representation of the NapA dimeric structure viewed from the extracellular side with TMs from one monomer labelled black and the other red. The TMs of the dimerisation domains are coloured in pale orange (TMs -1, 1, 2, 7, 8 and 9) and the core ion-translocation domains in sky blue (TMs 3, 4a-4b, 5, 10, 11a-11b and 12), connected together by TM6 shown in grey.

5.3.5 - The inward-facing state of NhaA state and outward-facing state of NapA

Structural comparisons were first made between the inward-facing state of NhaA and the outward-facing state of NapA based on their internal asymmetry of the inverted topology repeats. Even though the repeats from both proteins are overall structurally arranged in a similar manner, NapA contains two 6 TM inverted repeats linked together by TM 6 compared to the two 5 TM inverted repeats in NhaA linked together by both TM 6 and 7 (figure 5.10). This is due to the extra TM -1 from the first repeat in NapA, which is pseudosymmetry-related to TM 7 on the second. To facilitate

structural comparisons between the two structures, TM 6 was excluded from both with also the removal of TM -1 from NapA and the interfacial TM with the preceeding β-hairpin from NhaA.

(a)



(b)

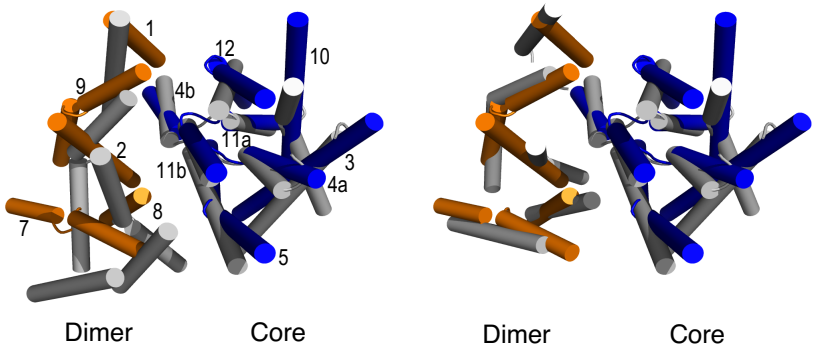


Figure 5.10 – Structural comparisons of NapA with NhaA. (a) Outward-facing NapA structure (left) and inward-facing NhaA structure (right) as viewed from the extracellular side with the removal of the β-hairpin domain located between TMs 1 and 2 from NhaA to facilitate visual comparison. The core and dimer domain boundaries are represented by a grey line. The NapA structure shows an open extracellular-facing cavity between the two domains, illustrated as an oval black circle. (b) Superposing the helices of the core and dimer domains of NhaA (grey) with the same domains of NapA (blue for the Core domain and orange for the dimer domain) together (left) and separately (right). TMs -1 and 6 have been omitted for clarity.

The r.m.s.d. value was 4.7 Å for 215 out of 247 pairs of Cα atoms from both structures when the core and dimer domains were superposed together (figure 5.10, part a). However this value decreased when the core and dimer domains were superposed separately with an r.m.s.d. of 1.8 Å for 134 out of 148 pairs of Cα atoms from the TMs in the core domains and 1.9 Å for 62 out of 88 pairs of Cα atoms from the TMs in the dimer domains (Figure 5.10, part b). This suggests that the whole movement of the core domain relative to the dimer domain may promote alternate access between the inward and outward-facing states.

On the cytoplasmic side of NapA, interactions between TMs 3,4 and 5 from the core domain and TMs 2 from the dimer domain close the cytoplasmic-facing cavity, which opens an outward-facing periplasmic cavity lined with negatively charged glutamic acid residues. At the bottom of the outward-facing funnel-like cavity is the location of the two ion binding aspartate residues Asp156 and Asp157, which align with the corresponding aspartate residues in NhaA also in the same location at the bottom of the inward-facing cavity (figure 5.11 and figure 2.4). In the inward-facing state of NhaA, both aspartate residues point towards the core domain compared to the outward-facing state of NapA, which shows Asp156 pointing towards the dimer domain with its side chain positioned at the base of the outward facing cavity (figure 5.11). Interestingly, the salt bridge pair previously seen between Asp164 and Lys300 from the NhaA structure after realignment of TM 10 (described in Chapter 4) is also observed in the NapA structure between the equivalent Asp157 and Lys305 from TM 10.

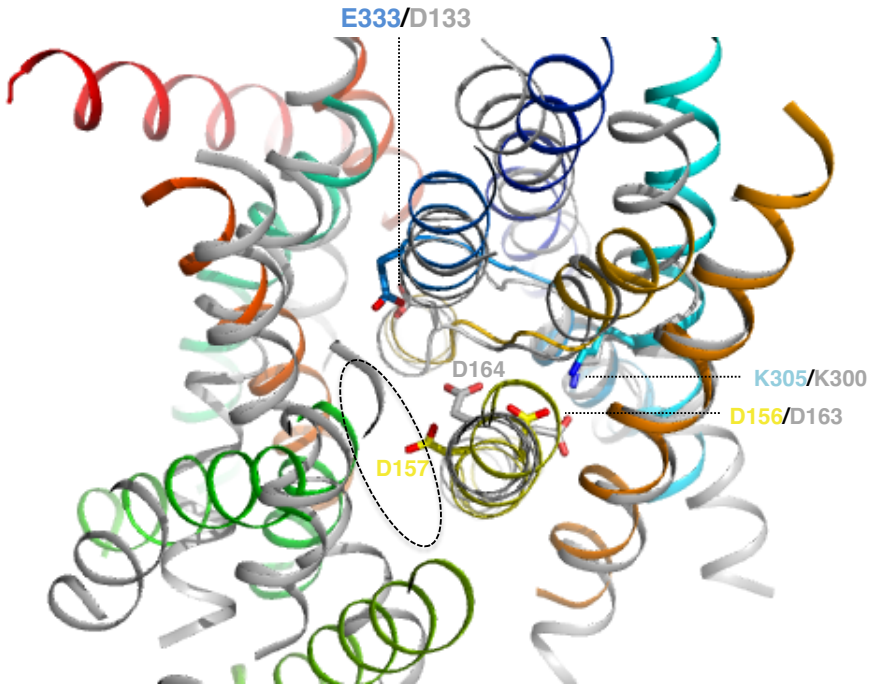


Figure 5.11 – Structure comparison of important residues in the ion binding site between NapA and NhaA. Cartoon representation of NapA (coloured as in Figure 5.8) and NhaA (grey) viewed from the extracellular side of the membrane with the residues understood to be involved in ion-translocation shown as sticks. The corresponding positions of the charged neutralising dipole residues (E333/D133, K305/K300) and the ion-binding aspartate (D156/D163) are the same apart from the position of the ion-binding aspartate (D157/D164), with D157 in NapA pointing towards the base of the open extracellular cavity, illustrated as a black oval circle. A sodium ion is proposed to be bound to the solvent-exposed D157 residue for release into the periplasm. A salt-bridge interaction is also suggested to be present between K305/K300 and D156/D163 in the inward and outward-facing states.

The Lys305 residue from TM 10 in NapA is equivalent to Lys300 in NhaA (figure 5.7), which both have shown to abolish transport activity when the positive charge at this position is removed (Furrer *et al*., 2007; Maes *et al*., 2012). This new position of the lysine residue from both structures is also likely to be involved in neutralising the negative dipoles from the opposing C-termini. However, the aspartate residue (Asp133) of NhaA positioned to neutralise the positive dipoles of the opposing N-termini is not conserved in NapA based on its structural alignment to Ser127 (figure

5.7). Instead from the structure, we can see the side-chain of Glu333 from TM 11b taking a similar position to the carboxylate of Asp133 in NhaA, which may be an alternative charge-compensating residue (figure 5.11). Alanine substitution of both residues have shown the same significant increase in the sodium binding affinity (Noumi *et al*., 1997), implying that these two pseudosymmery-related residues may also play a similar role in the ion binding site as well as not being the only negatively charged residue involved in charge compensation.

6.0 -

Final Discussion

Na^+/H^+ antiporters are located on the plasma membrane in all living cells and play a vital role in cellular sodium and pH homeostasis (Padan, 2008). Since 2005, the inward-facing structure of the well-characterised *E. coli* NhaA has been the only representative crystal structure of a Na^+/H^+ antiporter. This was crystallised at low pH where the protein is inactive (Hunte *et al*., 2005). With the very fast transport rate of NhaA, the first proposal of the transport mechanism was suggested to involve local rearrangements of the finely electrostatically balanced discontinuous helices (TMs 4a-b and 11a-b) in the core domain (Padan, 2008).

However, in the last ten years a number of bacterial crystal structures from different secondary transporter families have been determined at high resolution, revealing the presence of topology repeats to be a common structural theme (Boudker and Verdon, 2010). Sodium-coupled transporters in the LeuT and Glt_{Ph} family folds are shown to contain inverted structural repeats that intertwine to form two domains as well as discontinuous helices important for ion binding, similar to NhaA (Yernool *et al*., 2004; Yamashita *et al*., 2005; Screpanti and Hunte, 2007). The structure of the Na^+-coupled bile acid symporter $AsbT_{NM}$ has also been shown to have the NhaA fold, which first suggested it to be conserved across other secondary transporters (Hu *et al*., 2011). Extensive structural comparisons between multiple structures from the same or different secondary transporters of the same structural fold have overall given a general model of secondary transport involving a two-domain rocking bundle movement around a central substrate binding site (Forrest *et al*., 2011). Consequently we cannot rule out the presence of more global domain movements taking place in the Na^+/H^+ antiporter mechanism.

The initial goal of this thesis was to try and solve a crystal structure of NhaA in an active conformational state, preferably in an outward-facing state to compare with the inward-facing NhaA structure. Due to the reduced NhaA stability in a detergent solution above pH 6.5 where the protein is active, more progress was made in growing crystals at pH 7.8 and 9 from the NhaA homologue, NapA from *T. thermophilus*. Crystallisation studies of NapA were justified to complement the structural and extensive functional work of NhaA, since it also strictly conserved residues important for transport with the same electrogenic antiport mechanism as NhaA (Furrer *et al*., 2007). Structural progress was significantly made with NapA purified in a small micelle detergent, which could not have been a feasible option without the use of the MemStar optimised expression system.

The main purpose for developing the MemStar system was to reduce the number of membrane proteins being filtered out during the first stages of expression screening, which may in fact be more stable and therefore more likely to grow crystals using a small micelle detergent. Consequently the use of MemStar helped to obtain two crystal structures; an inward-facing state of the NhaA physiological dimer solved at pH 3.5 and a new outward-facing state of the NapA dimer solved at pH 7.8.

Bacterial and mammalian Na^+/H^+ antiporters have been shown to purify as dimers, suggesting this oligomeric state to be conserved in the SLC9 family (Williams, 2000; Hisamitsu *et al*., 2006; Rimon *et al*., 2007). The dimer interface region of NhaA is more localised compared to NapA, primarily consisting of the four-stranded β-sheet on the periplasmic side rather than the tight hydrophobic packing between helices in the inner dimer domains seen in NapA. As yet we do not know the importance of the

interactions between TM -1 and TM 7 in NapA for maintaining dimerisation and transport activity. Recently, a 7 Å cryo-EM structure of the CPA1 bacterial Na^+/H^+ antiporter, NhaP1 from *Methanococcus jannaschii*, was solved also containing 13 TM helices with two 6 TM topology inverted repeat units forming a dimer and core domain similar to NapA (Goswami *et al*., 2011). Deletion of the TM -1 of NhaP1 abolished transport activity but did not affect dimerisation, suggesting this helix in both proteins to possibly have a direct/indirect role in the transport mechanism (Goswami *et al*., 2011). Interestingly, a new NHE1 topology model has recently been proposed based on the observation of 13 helices present in the NhaP1 structure (Lee *et al*., 2012).

The physiological reason for a dimeric state is unclear, although one can suggest that a dimer would form a more rigid body and may therefore be important for protein stability in the membrane. The NhaA transport activity is not dependent on dimerisation although removal of the □-sheet motif is shown to reduce complementation growth under extreme salt stress conditions in addition to protein stability (Rimon *et al*., 2007; Herz *et al*., 2009). The human NHE1 exists as a dimer with the two monomers proposed to allosterically regulate each other (Moncoq *et al*., 2008). Indeed allosteric regulation of NhaP1 antiport activity has also been recently proposed through the action of the additional TM -1 (Goswami *et al*., 2011). In NapA, the beginning of TM -1 points closely towards the outward-facing cavity of its neighbouring monomer, therefore such a role of TM -1 involved transport regulation cannot be ruled out.

The NapA structure is similar to NhaA, which is characterised by the antiparallel discontinuous helices, TMs 4a-b and TMs 11a-11b that cross over each other. At the base of the outward and inward-facing cavities of the NapA and NhaA structures, respectively, is the location of the two strictly conserved aspartate residues likely to coordinate a sodium ion based on their position, conservation with mammalian Na^+/H^+ antiporters (Brett *et al*., 2005), phenotypes of mutants (Kuwabara *et al*., 2004; Padan, 2008), isothermal titration calorimetry (ITC) experiments (Maes *et al*., 2012) and MD simulations (Arkin *et al*., 2007). Mutation of either residue to an alanine or asparagine in NapA completely abolishes transport activity (Furrer *et al*., 2007). In the inward-facing NhaA structure, both Asp163 and Asp164 point towards the centre of the core domain. In the outward-facing NapA structure, Asp156 (equivalent to Asp163 in NhaA) is similarly orientated towards the centre of the core domain but Asp157 (equivalent to Asp164 in NhaA) points more towards the dimer domain, with its side chain carboxylate group exposed at the base of the outward-facing periplasmic cavity. Given that the NapA crystals were grown at pH 7.8 where the protein is active, we predict from our structure and the pKa of aspartic acid that Asp157 is likely to be deprotonated.

It was recently shown using solid-state membrane electrophysiology that the transport activity profile of NhaA fits a simple H^+ *vs* Na^+ kinetic binding model to a single common site and that they work equally well in either direction (Mager *et al*., 2011). In other words, the protons and sodium ions compete for binding to the aspartate residues. In the NapA and NhaA structures, the location of Asp157/Asp164 (NapA/NhaA) is shown to be better suited for ion binding. MD simulations of NapA have also found that Asp157 can rapidly bind sodium. Looking at this in a broader

perspective, Asp157/Asp164 is strictly conserved in all Na^+/H^+ antiporters from bacteria to man in the CPA superfamily as well as in bacterial, yeast and plant K^+/H^+ antiporters from the CPA2 family (Wiebe *et al.*, 2001; Aranda-Sicilia *et al.*, 2012). Mechanically, one can consider a simple situation where Na^+ and H^+ ions compete for binding to Asp157/Asp164 that becomes solvent exposed when alternating between the outward to inward-facing states. However, how can this explain electrogenic exchange from CPA2 Na^+/H^+ antiporters?

Interestingly, a salt-bridge is observed in both structures between the neighbouring aspartate, Asp156 of NapA and Asp163 of NhaA, to the strictly conserved lysine residue on TM 10 (Lys305/Lys300 in NapA/NhaA). Lysine residues located in the ion binding site have been implicated as proton acceptors in other proton-coupled transport mechanisms, such as in the Na^+/H^+ antiporter-like Nuo subunits of the respiratory complex I (Efremov and Sazanov, 2011) and the H^+-coupled transporter ApcT (Shaffer *et al*., 2009). Structures of sodium-coupled symporters such as LeuT (Yamashita *et al*., 2005), Mhp1 (Weyand *et al*., 2008), vSGLT (Faham *et al*., 2008) and BetP (Ressl *et al*., 2009) also show a co-transported sodium ion in the ion binding site at the same position as the proton-accepting lysine residue in the ApcT transporter (Shaffer *et al*., 2009). The same analysis of NhaA with the repositioned Lys300 and NapA has similarly shown its amine group in the same position as the co-transported sodium ion in the structural homologue $AsbT_{NM}$ (figure 6.1)(Hu *et al*., 2011).

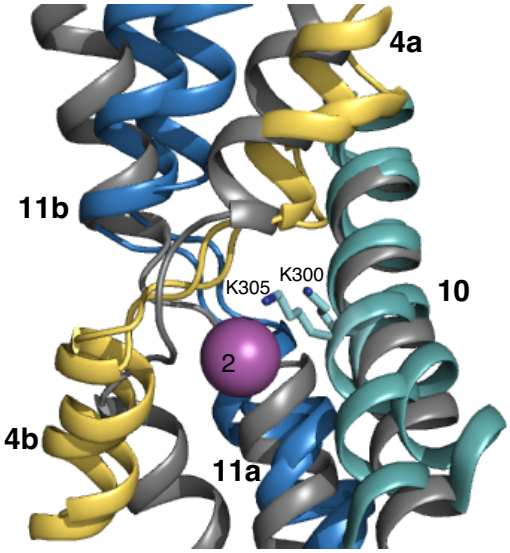


Figure 6.1 – Structural comparison of the core domains of NapA, NhaA and $AsbT_{NM}$. Cartoon representation of TMs 4a-b, 10 and 11a-b of NapA/NhaA superposed on TMs 4,8 and 9 of $AsbT_{NM}$ (PDB: 3ZUX; (Hu *et al*., 2011)) respectively and are coloured grey. The sodium Na2 site of $AsbT_{NM}$ is depicted as a purple sphere with K305 from NapA and K300 from NhaA (after realignment of TM 10) shown as a stick model in a similar position.

Given that the active NapA structure was solved at pH 7.8, the lysine residue is more likely to be protonated than the aspartate residue based on its higher pKa value. However, their pKa values may be reduced by their local environment in the ion binding site, causing the aspartate to possibly become deprotonated and consequently interact with the protonated lysine residue. With the new NhaA structure also showing this salt bridge pair, it is possible that this polar interaction between these two strictly conserved residues may be an important component of the transport mechanism, which offers a new reason why transport activity is lost when either the carboxylate or amine group is mutated (Inoue *et al*., 1995; Furrer *et al*., 2007). A recent NhaA study has also shown an increase in the pH to activate transport when Lys300 is replaced with arginine, implying that the strength of the observed salt-bridge may also affect activation of the transport mechanism (Maes *et al*., 2012).

From these two structures it appears that this salt-bridge between the lysine and aspartate residue is central to the transport mechanism. In collaboration with Dr. Oliver Beckstein, during MD simulations of NhaA and NapA the presence of the salt-bridge was only maintained in the absence of sodium. With NhaA, sodium binding to Asp164 eventually caused the salt-bridge to be broken with some simulations showing Asp163 coordinating a sodium ion after breakage of the salt-bridge.

Based on the NapA structure and the MD simulations, we suggest that in the inward-facing state the protonated Asp157/164 residue may first lose a proton (at a higher pH of the cytoplasm) to bind to a Na^+ ion (figure 6.2). This will subsequently cause the Asp156/163 – Lys305/300 salt bridge to break, with the second aspartate also coordinating the sodium ion and deprotonation of the lysine residue to release the second proton into the cytoplasm. Although Asp133 in NhaA and Glu333 in NapA are not essential for transport activity, these residues could also be protonated or deprotonated and therefore be involved in sodium coordination when switching from the inward to outward-facing state.

Once this salt-bridge is broken, the inward-facing state will switch conformation to expose the sodium ion to the periplasm in the outward-facing state for its release in the large water-filled cavity. Based on the NapA structure, we propose that after the sodium ion is released, Asp156/163 may possibly first accept a proton but is immediately taken up by Lys305/Lys300 for reprotonation and consequently reform the salt-bridge. Protonation of the Asp157/164 residue may then trigger the outward facing state to switch back to the inward-facing state.

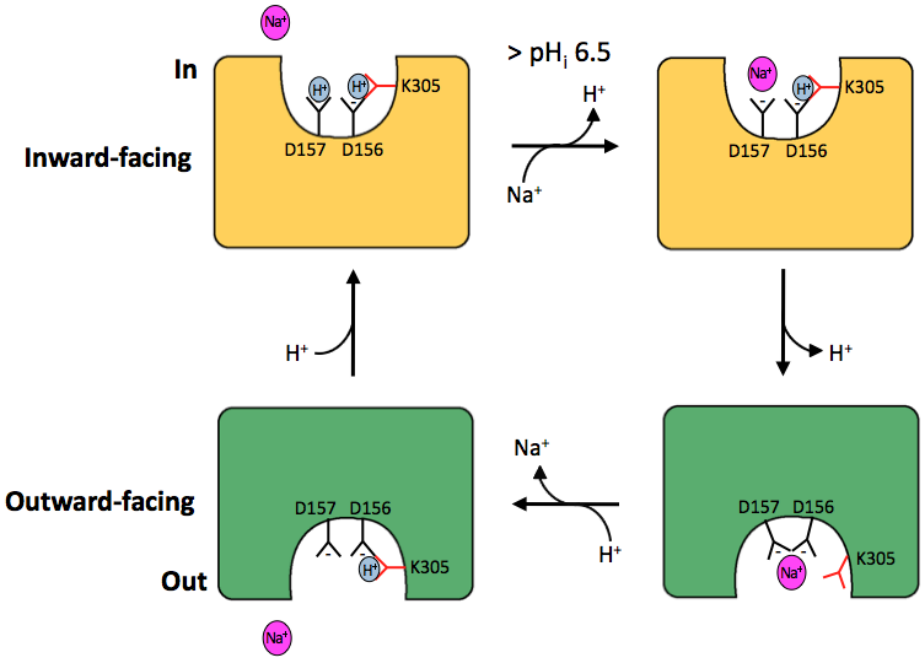


Figure 6.2 – Schematic model of the NapA transport cycle in the ion binding site. From the inward-facing state, D157 will deprotonate when intracellular pH is above 6.5 and coordinate a sodium ion (magenta sphere) that has entered the ion binding site, causing the D156-K305 salt-bridge to break. The subsequent reorientation of D156 to coordinate the sodium ion and deprotonation of K305 may then trigger conformational changes to switch from the inward-facing to outward-facing state. Based on the NapA structure, it is proposed that after sodium ion release into the periplasm, the D156-K305 salt-bridge is reformed with protonation of D157 to switch back to the inward-facing state (NhaA structure).

Conformational changes during Na^+/H^+ antiport were previously proposed to involve local rearrangements of the finely electrostatically balanced discontinuous helices in the core domain (Hunte *et al*., 2005; Padan, 2008). However, we have seen major structural differences in the position of the core domain relative to the dimer domain between NapA and NhaA. With reference to the dimerisation interface, which is shown from cryo-EM studies not to change appreciably in response to Na^+ binding (Vinothkumar *et al*., 2005; Appel *et al*., 2009), the core domain in the NhaA structure can be rotated by 21° relative to the core of NapA. This large rotation of the core domain closes the cavity seen on the outside of NapA and opens the cytoplasmic

funnel on the inside, as observed in NhaA. During this process the proposed single cation binding aspartate (Asp157/Asp164) is shifted 10 Å towards the cytoplasmic surface of the transporter (figure 6.3).

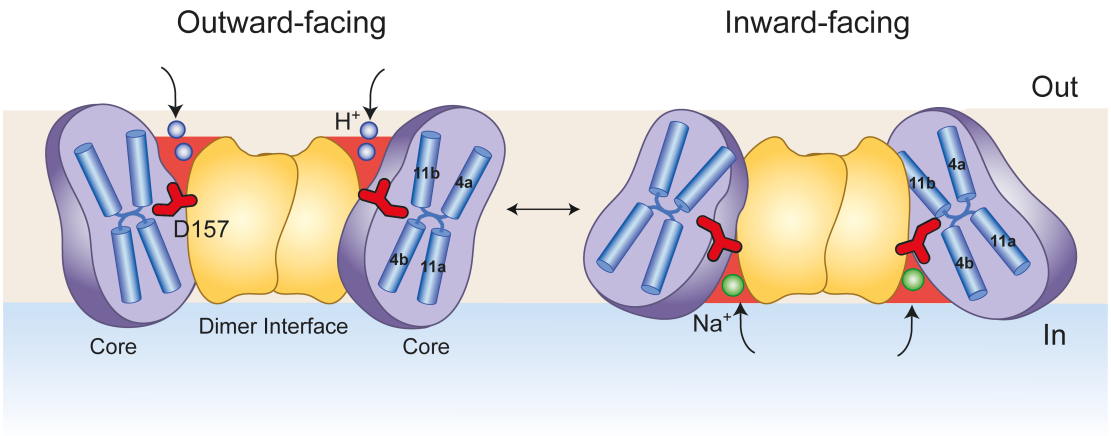
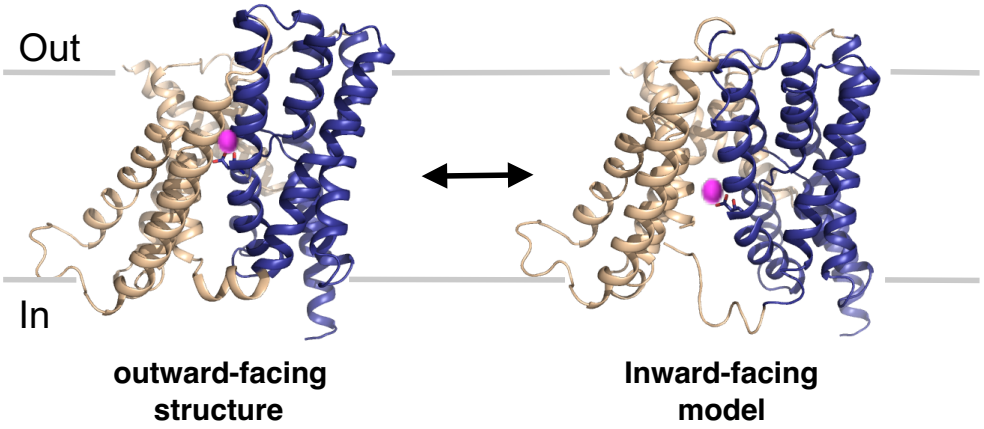


Figure 6.3 – Alternating access model of Na^+/H^+ antiport. Schematic representation of the proposed mechanism illustrating the conformational changes, that of the core moving against the dimerisation domain, showing the transition from the outward-open (left) to the inward-open state (right). During this process the location of ion binding aspartate residue (Asp157/Asp164), shown in red, moves some 10 Å towards the cytoplasmic surface.

This elevator movement of a substrate-binding core domain, in this case to carry Na^+ ions from one side of the membrane to the other, resembles that of the two-domain rocking bundle movements seen in the transport mechanism of the Na^+-coupled glutamate transporter Glt_{Ph} (Reyes *et al*., 2009)(figure 6.4) and has also been proposed for the structural NhaA homologue $AsbT_{NM}$ (Hu *et al*., 2011). Interestingly, a two domain transport mechanism was also recently predicted in NhaA, based on a consideration of the two symmetry related inverted repeats as well as elastic network models and biochemical cross-linking (Schushan *et al*., 2012). It was also noticed that the position of Asp164 from this NhaA model points towards the outward-facing

cavity, similar to the equivalent Asp157 residue in the outward-facing NapA structure.

(a)



(b)

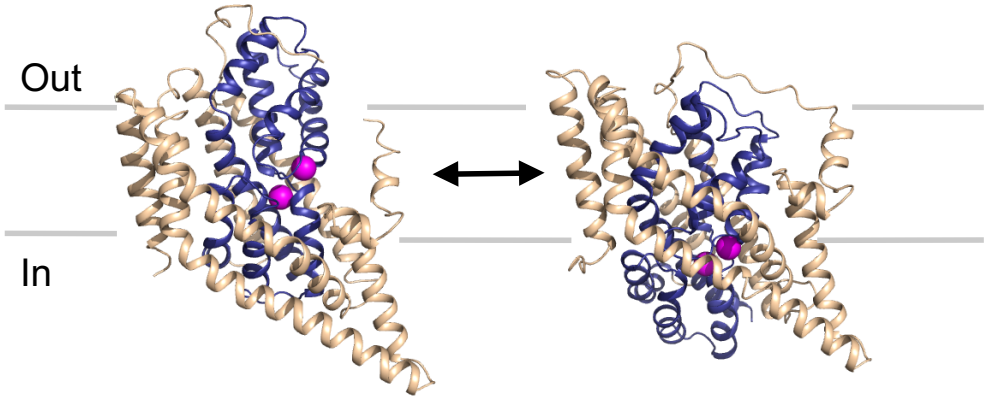


Figure 6.4 – Two-domain rocking bundle model of NapA and Glt_{Ph}. (a) Cartoon representation of the outward-facing NapA structure (left) and inward-facing NapA model (right), which was constructed by reorientation of the core domain (blue) to the same position as the core domain in the inward-facing NhaA structure relative to the fixed position of the dimer domain (beige). Asp157 is shown as a stick and modelled to coordinate a sodium ion (pink sphere) between the two states. (b). The crystal structures of Glt_{Ph}, solved in the outward (PDB: 2NWX; (Boudker *et al.*, 2007)) and inward-occluded (PDB: 3KBC; (Reyes *et al.*, 2009)) states, shown in the same format as NapA with two coordinated sodium ions (pink spheres). Position of the membrane is shown by a grey line.

7.0 - Concluding Remarks

The work from this thesis has enabled us to describe global conformational changes that make up the alternating access model of Na^+/H^+ antiport. Surprisingly, the conformational change between the inward NhaA state and outward NapA state are very large. It therefore seems that the rate of secondary active transport may not necessarily correlate with the size of the conformational change or the substrate itself, as previously assumed. Although we have not been able to resolve sodium binding, the NhaA and NapA structures are consistent with a single ion-translocation site mechanism, of which a strictly conserved aspartate residue (Asp157/Asp164 in NapA/NhaA) is solvent accessible in both opposing conformations. Further structures and methods that capture membrane dynamics are however needed to clarify how the local ion binding and release are coupled to these global changes now outlined in this thesis.

8.0 – Bibliography

Abrahams JP, Leslie AGW (1996) Methods used in the structure determination of bovine mitochondrial F_1 ATPase. *Acta Cryst*, **D52:** 30–42

Abramson J, Smirnova I, Kasho V, Verner G, Kaback HR, Iwata S (2003) Structure and mechanism of the lactose permease of Escherichia coli. *Science*, **301:** 610–615

Adams PD (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D*, **66:** 213–221

Alberts B. (2008) *Molecular biology of the cell*. Reference edition. Garland Pub.

Alexandrov AI, Mileni M, Chien EY, Hanson MA, Stevens RC (2008) Microscale fluorescent thermal stability assay for membrane proteins. *Structure*, **16:** 351–359

Amith SR, Fliegel L (2013) Regulation of the Na+/H+ Exchanger (NHE1) in Breast Cancer Metastasis. *Cancer Res*, **73:** 1259–1264

Appel M, Hizlan D, Vinothkumar KR, Ziegler C, Kuhlbrandt W (2009) Conformations of NhaA, the Na/H exchanger from Escherichia coli, in the pH-activated and ion-translocating states. *J Mol Biol*, **386:** 351–365

Apse MP, Aharon GS, Snedden WA, Blumwald E (1999) Salt tolerance conferred by overexpression of a vacuolar Na+/H+ antiport in Arabidopsis. *Science*, **285:** 1256–1258

Aranda-Sicilia MN, Cagnac O, Chanroj S, Sze H, Rodriguez-Rosales MP, Venema K (2012) Arabidopsis KEA2, a homolog of bacterial KefC, encodes a K(+)/H(+) antiporter with a chloroplast transit peptide. *Biochim Biophys Acta*, **1818:** 2362–2371

Arkin IT, Xu H, Jensen MO, Arbely E, Bennett ER, Bowers KJ, Chow E, Dror RO, Eastwood MP, Flitman-Tene R, Gregersen BA, Klepeis JL, Kolossvary I, Shan Y, Shaw DE (2007) Mechanism of Na+/H+ antiporting. *Science*, **317:** 799–803

Batey RT, Rambo RP, Lucast L, Rha B, Doudna JA (2000) Crystal structure of the ribonucleoprotein core of the signal recognition particle. *Science*, **287:** 1232–1239

Bernardi P (1999) Mitochondrial transport of cations: channels, exchangers, and permeability transition. *Physiol Rev*, **79:** 1127–1155

Bill RM, Henderson PJ, Iwata S, Kunji ER, Michel H, Neutze R, Newstead S, Poolman B, Tate CG, Vogel H (2011) Overcoming barriers to membrane protein structure determination. *Nat Biotechnol*, **29:** 335–340

Bobulescu IA, Di Sole F, Moe OW (2005) Na+/H+ exchangers: physiology and link to hypertension and organ ischemia. *Curr Opin Nephrol Hypertens*, **14:** 485–494

Bobulescu IA, Moe OW (2009) Luminal Na(+)/H (+) exchange in the proximal tubule. *Pflugers Arch*, **458:** 5–21

Boudker O, Ryan RM, Yernool D, Shimamoto K, Gouaux E (2007) Coupling substrate and ion binding to extracellular gate of a sodium-dependent aspartate transporter. *Nature*, **445:** 387–393

Boudker O, Verdon G (2010) Structural perspectives on secondary active transporters. *Trends Pharmacol Sci*, **31:** 418–426

Bowie JU (2001) Stabilizing membrane proteins. *Curr Opin Struct Biol*, **11:** 397–402
Brett CL, Donowitz M, Rao R (2005) Evolutionary origins of eukaryotic sodium/proton exchangers. *Am J Physiol Cell Physiol*, **288:** C223–39
Bricogne G, Vonrhein G, Flensburg C, Schiltz M, and Paciorek W (2003) Generation, representation and flow of phase informtion in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallogr D Biol Crystallogr*, **59:** 2023–2030
Canessa M, Adragna N, Solomon HS, Connolly TM, Tosteson DC (1980) Increased sodium-lithium countertransport in red cells of patients with essential hypertension. *N Engl J Med*, **302:** 772–776
Carpenter EP, Beis K, Cameron AD, Iwata S (2008) Overcoming the challenges of membrane protein crystallography. *Curr Opin Struct Biol*, **18:** 581–586
Chen VB, Arendall WBr, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC (2010a) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr*, **66:** 12–21
Chen YH, Hu L, Punta M, Bruni R, Hillerich B, Kloss B, Rost B, Love J, Siegelbaum SA, Hendrickson WA (2010b) Homologue structure of the SLAC1 anion channel for closing stomata in leaves. *Nature*, **467:** 1074–1080
Collaborative Computational Project Number 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D,* **50**, 760-763
Counillon L, Pouyssegur J (2000) The expanding family of eucaryotic Na(+)/H(+) exchangers. *J Biol Chem*, **275:** 1–4
Cowan SW, Schirmer T, Rummel G, Steiert M, Ghosh R, Pauptit RA, Jansonius JN, Rosenbusch JP (1992) Crystal structures explain functional properties of two E. coli porins. *Nature*, **358:** 727–733
Cowtan K (1994) 'dm': An automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography*, **31:** 34–38
Dalbey RE, Wang P, Kuhn A (2011) Assembly of bacterial inner membrane proteins. *Annu Rev Biochem*, **80:** 161–187
Deacon SE, Roach PC, Postis VL, Wright GS, Xia X, Phillips SE, Knox JP, Henderson PJ, McPherson MJ, Baldwin SA (2008) Reliable scale-up of membrane protein over-expression by bacterial auto-induction: from microwell plates to pilot scale fermentations. *Mol Membr Biol*, **25:** 588–598
DeLaBarre B, Brunger AT (2006) Considerations for the refinement of low-resolution crystal structures. *Acta Crystallogr D Biol Crystallogr*, **62:** 923–932
DeLano WL (2002) The PyMOL molecular graphics system. *Delano Scientific, Palo Alto, CA, USA*
Demaurex N, Grinstein S (1994) Na+/H+ antiport: modulation by ATP and role in cell volume regulation. *J Exp Biol*, **196:** 389–404
Demaurex N, Orlowski J, Brisseau G, Woodside M, Grinstein S (1995) The mammalian Na+/H+ antiporters NHE-1, NHE-2, and NHE-3 are electroneutral and voltage independent, but can couple to an H+ conductance. *J Gen Physiol*, **106:** 85–111
Ding J, Rainey JK, Xu C, Sykes BD, Fliegel L (2006) Structural and functional characterization of transmembrane segment VII of the Na+/H+ exchanger isoform 1. *J Biol Chem*, **281:** 29817–29829

Drew DE, von Heijne G, Nordlund P, de Gier JW (2001) Green fluorescent protein as an indicator to monitor membrane protein overexpression in Escherichia coli. *FEBS Lett*, **507:** 220–224

Drew D, Slotboom DJ, Friso G, Reda T, Genevaux P, Rapp M, Meindl-Beinker NM, Lambert W, Lerch M, Daley DO, Van Wijk KJ, Hirst J, Kunji E, De Gier JW (2005) A scalable, GFP-based pipeline for membrane protein overexpression screening and purification. *Protein Sci*, **14:** 2011–2017

Drew D, Lerch M, Kunji E, Slotboom DJ, de Gier JW (2006) Optimization of membrane protein overexpression and purification using GFP fusions. *Nat Methods*, **3:** 303–313

Drew D, Newstead S, Sonoda Y, Kim H, von Heijne G, Iwata S (2008) GFP-based optimization scheme for the overexpression and purification of eukaryotic membrane proteins in Saccharomyces cerevisiae. *Nat Protoc*, **3:** 784–798

Efremov RG, Sazanov LA (2011) Structure of the membrane domain of respiratory complex I. *Nature*, **476:** 414–420

Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*, **60:** 2126–2132

Evans G, Pettifer RF (2001) CHOOCH: a program for deriving anomalous-scattering factors from X-ray fluorescence spectra. *J Appl Cryst*, **34:** 82–86

Faham S, Watanabe A, Besserer GM, Cascio D, Specht A, Hirayama BA, Wright EM, Abramson J (2008) The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na+/sugar symport. *Science*, **321:** 810–814

Fang Y, Jayaram H, Shane T, Kolmakova-Partensky L, Wu F, Williams C, Xiong Y, Miller C (2009) Structure of a prokaryotic virtual proton pump at 3.2 A resolution. *Nature*, **460:** 1040–1043

Ferguson GP, Munro AW, Douglas RM, McLaggan D, Booth IR (1993) Activation of potassium channels during metabolite detoxification in Escherichia coli. *Mol Microbiol*, **9:** 1297–1303

Forrest LR, Kramer R, Ziegler C (2011) The structural basis of secondary active transport mechanisms. *Biochim Biophys Acta*, **1807:** 167–188

Furrer EM, Ronchetti MF, Verrey F, Pos KM (2007) Functional characterization of a NapA Na(+)/H(+) antiporter from Thermus thermophilus. *FEBS Lett*, **581:** 572–578

Fuster DG, Zhang J, Shi M, Bobulescu IA, Andersson S, Moe OW (2008) Characterization of the sodium/hydrogen exchanger NHA2. *J Am Soc Nephrol*, **19:** 1547–1556

Gerchman Y, Rimon A, Venturi M, Padan E (2001) Oligomerization of NhaA, the Na+/H+ antiporter of Escherichia coli in the membrane and its functional and structural consequences. *Biochemistry*, **40:** 3403–3412

Goldberg EB, Arbel T, Chen J, Karpel R, Mackie GA, Schuldiner S, Padan E (1987) Characterization of a Na+/H+ antiporter gene of Escherichia coli. *Proc Natl Acad Sci U S A*, **84:** 2615–2619

Gordon E, Horsefield R, Swarts HG, de Pont JJ, Neutze R, Snijder A (2008) Effective high-throughput overproduction of membrane proteins in Escherichia coli. *Protein Expr Purif*, **62:** 1–8

Goswami P, Paulino C, Hizlan D, Vonck J, Yildiz O, Kuhlbrandt W (2011) Structure of the archaeal Na+/H+ antiporter NhaP1 and functional role of transmembrane helix 1. *EMBO J*, **30:** 439–449

Gouaux E, Mackinnon R (2005) Principles of selective ion transport in channels and pumps. *Science*, **310:** 1461–1465

Gould GW, Holman GD (1993) The glucose transporter family: structure, function and tissue-specific expression. *Biochem J*, **295:** 329–341

Hammon J, Palanivelu DV, Chen J, Patel C, Minor DL, Jr. (2009) A green fluorescent protein screen for identification of well-expressed membrane proteins from a cohort of extremophilic organisms. *Protein Sci*, **18:** 121–133

Hase CC, Fedorova ND, Galperin MY, Dibrov PA (2001) Sodium ion cycle in bacterial pathogens: evidence from cross-genome comparisons. *Microbiol Mol Biol Rev*, **65:** 353–70

Hattori M, Hibbs RE, Gouaux E (2012) A fluorescence-detection size-exclusion chromatography-based thermostability assay for membrane protein precrystallization screening. *Structure*, **20:** 1293–1299

Heijne G (1986) The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology. *EMBO J*, **5:** 3021–3027

Hellmer J, Patzold R, Zeilinger C (2002) Identification of a pH regulated Na(+)/H(+) antiporter of Methanococcus jannaschii. *FEBS Lett*, **527:** 245–249

Heras B, Martin JL (2005) Post-crystallization treatments for improving diffraction quality of protein crystals. *Acta Crystallogr D Biol Crystallogr*, **61:** 1173–1180

Herz K, Rimon A, Jeschke G, Padan E (2009) Beta-sheet-dependent dimerization is essential for the stability of NhaA Na+/H+ antiporter. *J Biol Chem*, **284:** 6337–6347

Hessa T, Meindl-Beinker NM, Bernsel A, Kim H, Sato Y, Lerch-Bader M, Nilsson I, White SH, von Heijne G (2007) Molecular code for transmembrane-helix recognition by the Sec61 translocon. *Nature*, **450:** 1026–1030

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

Holm L, Park J (2000) DaliLite workbench for protein structure comparison. *Bioinformatics*, **16:** 566–567

Hong H, Patel DR, Tamm LK, van den Berg B (2006) The outer membrane protein OmpW forms an eight-stranded beta-barrel with a hydrophobic channel. *J Biol Chem*, **281:** 7568–7577

Hu NJ, Iwata S, Cameron AD, Drew D (2011) Crystal structure of a bacterial homologue of the bile acid sodium symporter ASBT. *Nature*, **478:** 408–411

Huang Y, Lemieux MJ, Song J, Auer M, Wang DN (2003) Structure and mechanism of the glycerol-3-phosphate transporter from Escherichia coli. *Science*, **301:** 616–620

Hunte C, Screpanti E, Venturi M, Rimon A, Padan E, Michel H (2005) Structure of a Na+/H+ antiporter and insights into mechanism of action and regulation by pH. *Nature*, **435:** 1197–1202

Inoue H, Noumi T, Tsuchiya T, Kanazawa H (1995) Essential aspartic acid residues, Asp-133, Asp-163 and Asp-164, in the transmembrane helices of a Na+/H+ antiporter (NhaA) from Escherichia coli. *FEBS Lett*, **363:** 264–268

Inoue H, Sakurai T, Ujike S, Tsuchiya T, Murakami H, Kanazawa H (1999) Expression of functional Na+/H+ antiporters of Helicobacter pylori in antiporter-deficient Escherichia coli mutants. *FEBS Lett*, **443:** 11–16

Jardetzky O (1966) Simple allosteric model for membrane pumps. *Nature*, **211:** 969–970

Jones TA, Kjeldgaard M (1997) Electron-density map interpretation. *Methods in Enzymology*, **277:** 173–208

Kabsh W (2010) XDS. *Acta Crystallogr D*, **66:** 125–132
Kang HJ, Lee C, Drew D (2013) Breaking the barriers in membrane protein crystallography. *Int J Biochem Cell Biol*, **45:** 636–644
Karpel R, Olami Y, Taglicht D, Schuldiner S, Padan E (1988) Sequencing of the gene ant which affects the Na+/H+ antiporter activity in Escherichia coli. *J Biol Chem*, **263:** 10408–10414
Karpel R, Alon T, Glaser G, Schuldiner S, Padan E (1991) Expression of a sodium proton antiporter (NhaA) in Escherichia coli is induced by Na+ and Li+ ions. *J Biol Chem*, **266:** 21753–21759
Kauko A, Hedin LE, Thebaud E, Cristobal S, Elofsson A, von Heijne G (2010) Repositioning of transmembrane alpha-helices during membrane protein folding. *J Mol Biol*, **397:** 190–201
Kim H, Melen K, Osterberg M, von Heijne G (2006) A global topology map of the Saccharomyces cerevisiae membrane proteome. *Proc Natl Acad Sci U S A*, **103:** 11142–11147
Kleywegt GJ, and Jones TA (1994) A super position. *ESF/CCp4 Newsletter*, **31:** 9
Kleywegt GJ, Zou, J.Y., Kjeldgaard, M., and Jones, T.A (2001) Around O. In *International Tables for Crystallography, Volume F. Crystallography of Biological Macromolecules,*M.G. Rossmann and Arnold E (eds.), pp 497–506, 526-528.
Knight SD (2000) RSPS version 4.0: a semi-interactive vector-search program for solving heavy-atom derivatives. *Acta Crystallogr D Biol Crystallogr*, **56:** 42–47
Kondapalli KC, Kallay LM, 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
Korepanova A, Gao FP, Hua Y, Qin H, Nakamoto RK, Cross TA (2005) Cloning and expression of multiple integral membrane proteins from Mycobacterium tuberculosis in Escherichia coli. *Protein Sci*, **14:** 148–158
Kozachkov L, Herz K, Padan E (2007) Functional and structural interactions of the transmembrane domain X of NhaA, Na+/H+ antiporter of Escherichia coli, at physiological pH. *Biochemistry*, **46:** 2419–2430
Kuwabara N, Inoue H, Tsuboi Y, Nakamura N, Kanazawa H (2004) The fourth transmembrane domain of the Helicobacter pylori Na+/H+ antiporter NhaA faces a water-filled channel required for ion transport. *J Biol Chem*, **279:** 40567–40575
Lee BL, Sykes BD, Fliegel L (2012) Structural and functional insights into the cardiac Na(+)/H(+) exchanger. *J Mol Cell Cardiol*, **89:** 189–199
Leung J, Cameron AD, Diallinas G, Byrne B (2012) Stabilizing the heterologously expressed uric acid-xanthine transporter UapA from the lower eukaryote Aspergillus nidulans. *Mol Membr Biol*, **30:** 32–42
Lu M, Fu D (2007) Structure of the zinc transporter YiiP. *Science*, **317:** 1746–1748
Luirink J, Sinning I (2004) SRP-mediated protein targeting: structure and function revisited. *Biochim Biophys Acta*, **1694:** 17–35
Maes M, Rimon A, Kozachkov-Magrisso L, Friedler A, Padan E (2012) Revealing the Ligand Binding Site of NhaA Na+/H+ Antiporter and its pH Dependence. *J Biol Chem*, **287:** 38150–38157
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

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr*, **40:** 658–674
Miroux B, Walker JE (1996) Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol*, **260:** 289–298
Moncoq K, Kemp G, Li X, Fliegel L, Young HS (2008) Dimeric structure of human Na+/H+ exchanger isoform 1 overproduced in Saccharomyces cerevisiae. *J Biol Chem*, **283:** 4145–4154
Morth JP, Sorensen TL, Nissen P (2006) Membrane's Eleven: heavy-atom derivatives of membrane-protein crystals. *Acta Crystallogr D Biol Crystallogr*, **62:** 877–882
Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood method. *Acta Cryst*, **D53:** 240–255
Murtazina R, Booth BJ, Bullis BL, Singh DN, 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
Nikaido H (2003) Molecular basis of bacterial outer membrane permeability revisited. *Microbiol Mol Biol Rev*, **67:** 593–656
Norholm MH, Light S, Virkki MT, Elofsson A, von Heijne G, Daley DO (2012) Manipulating the genetic code for membrane protein production: What have we learnt so far? *Biochim Biophys Acta*, **1818:** 1091–1096
Noumi T, Inoue H, Sakurai T, Tsuchiya T, Kanazawa H (1997) Identification and characterization of functional residues in a Na+/H+ antiporter (NhaA) from Escherichia coli by random mutagenesis. *J Biochem*, **121:** 661–670
Numata M, Petrecca K, Lake N, Orlowski J (1998) Identification of a mitochondrial Na+/H+ exchanger. *J Biol Chem*, **273:** 6951–6959
Nunez-Ramirez R, Sanchez-Barrena MJ, Villalta I, Vega JF, Pardo JM, Quintero FJ, Martinez-Salazar J, Albert A (2012) Structural Insights on the Plant Salt-Overly-Sensitive 1 (SOS1) Na(+)/H(+) Antiporter. *J Mol Biol*, **424:** 283–294
Orlowski J, Grinstein S (2004) Diversity of the mammalian sodium/proton exchanger SLC9 gene family. *Pflugers Arch*, **447:** 549–565
Orlowski J, Grinstein S (2007) Emerging roles of alkali cation/proton exchangers in organellar homeostasis. *Curr Opin Cell Biol*, **19:** 483–492
Overington JP, Al-Lazikani B, Hopkins AL (2006) How many drug targets are there? *Nat Rev Drug Discov*, **5:** 993–996
Padan E, Maisler N, Taglicht D, Karpel R, Schuldiner S (1989) Deletion of ant in Escherichia coli reveals its function in adaptation to high salinity and an alternative Na+/H+ antiporter system(s). *J Biol Chem*, **264:** 20297–20302
Padan E, Schuldiner S (1994) Molecular physiology of Na+/H+ antiporters, key transporters in circulation of Na+ and H+ in cells. *Biochim Biophys Acta*, **1185:** 129–151
Padan E, Tzubery T, Herz K, Kozachkov L, Rimon A, Galili L (2004) NhaA of Escherichia coli, as a model of a pH-regulated Na+/H+antiporter. *Biochim Biophys Acta*, **1658:** 2–13
Padan E, Bibi E, Ito M, Krulwich TA (2005) Alkaline pH homeostasis in bacteria: new insights. *Biochim Biophys Acta*, **1717:** 67–88
Padan E (2008) The enlightening encounter between structure and function in the NhaA Na+-H+ antiporter. *Trends Biochem Sci*, **33:** 435–443
Padan E, Kozachkov L, Herz K, Rimon A (2009) NhaA crystal structure: functional-structural insights. *J Exp Biol*, **212:** 1593–1603

Potterton L, McNicholas S, Krissinel E, Gruber J, Cowtan K, Emsley P, Murshudov GN, Cohen S, Perrakis A, Noble M (2004) Developments in the CCP4 molecular-graphics project. *Acta Crystallogr D Biol Crystallogr*, **60:** 2288–2294

Putney LK, Denker SP, Barber DL (2002) The changing face of the Na+/H+ exchanger, NHE1: structure, regulation, and cellular actions. *Annu Rev Pharmacol Toxicol*, **42:** 527–552

Radestock S, Forrest LR (2011) The alternating-access mechanism of MFS transporters arises from inverted-topology repeats. *J Mol Biol*, **407:** 698–715

Rapp M, Drew D, Daley DO, Nilsson J, Carvalho T, Melen K, De Gier JW, Von Heijne G (2004) Experimentally based topology models for E. coli inner membrane proteins. *Protein Sci*, **13:** 937–945

Ressl S, Terwisscha van Scheltinga AC, Vonrhein C, Ott V, Ziegler C (2009) Molecular basis of transport and regulation in the Na(+)/betaine symporter BetP. *Nature*, **458:** 47–52

Reyes N, Ginter C, Boudker O (2009) Transport mechanism of a bacterial homologue of glutamate transporters. *Nature*, **462:** 880–885

Rimon A, Tzubery T, Padan E (2007) Monomers of the NhaA Na+/H+ antiporter of Escherichia coli are fully functional yet dimers are beneficial under extreme stress conditions at alkaline pH in the presence of Na+ or Li+. *J Biol Chem*, **282:** 26810–26821

Rodriguez-Rosales MP, Galvez FJ, Huertas R, Aranda MN, Baghour M, Cagnac O, Venema K (2009) Plant NHX cation/proton antiporters. *Plant Signal Behav*, **4:** 265–276

Rothman A, Gerchman Y, Padan E, Schuldiner S (1997) Probing the conformation of NhaA, a Na+/H+ antiporter from Escherichia coli, with trypsin. *Biochemistry*, **36:** 14572–14576

Saier MHJ, Tran CV, Barabote RD (2006) TCDB: the Transporter Classification Database for membrane transport protein analyses and information. *Nucleic Acids Res*, **34:** D181–6

Schlegel S, Lofblom J, Lee C, Hjelm A, Klepsch M, Strous M, Drew D, Slotboom DJ, de Gier JW (2012) Optimizing Membrane Protein Overexpression in the Escherichia coli strain Lemo21(DE3). *J Mol Biol*, **423:** 648–659

Schomaker V, Trueblood KN (1968) On the rigid-body motion of molecules in crystals. *Acta Cryst*, **B24:** 63–76

Schulz GE (2000) beta-Barrel membrane proteins. *Curr Opin Struct Biol*, **10:** 443–447

Schushan M, Rimon A, Haliloglu T, Forrest LR, Padan E, Ben-Tal N (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

Screpanti E, Padan E, Rimon A, Michel H, Hunte C (2006) Crucial steps in the structure determination of the Na+/H+ antiporter NhaA in its native conformation. *J Mol Biol*, **362:** 192–202

Screpanti E, Hunte C (2007) Discontinuous membrane helices in transport proteins and their correlation with function. *J Struct Biol*, **159:** 261–267

Serrano-Vega MJ, Magnani F, Shibata Y, Tate CG (2008) Conformational thermostabilization of the beta1-adrenergic receptor in a detergent-resistant form. *Proc Natl Acad Sci U S A*, **105:** 877–882

Shaffer PL, Goehring A, Shankaranarayanan A, Gouaux E (2009) Structure and mechanism of a Na+-independent amino acid transporter. *Science*, **325:** 1010–1014

Shibata Y, White JF, Serrano-Vega MJ, Magnani F, Aloia AL, Grisshammer R, Tate CG (2009) Thermostabilization of the neurotensin receptor NTS1. *J Mol Biol*, **390:** 262–277

Silver IA, Erecinska M (1997) Energetic demands of the Na+/K+ ATPase in mammalian astrocytes. *Glia*, **21:** 35–45

Skulachev VP (1991) Chemiosmotic systems in bioenergetics: H(+)-cycles and Na(+)-cycles. *Biosci Rep*, **11:** 387–441

Slepkov ER, Rainey JK, Sykes BD, Fliegel L (2007) Structural and functional analysis of the Na+/H+ exchanger. *Biochem J*, **401:** 623–633

Slotboom DJ, Duurkens RH, Olieman K, Erkens GB (2008) Static light scattering to characterize membrane proteins in detergent solution. *Methods*, **46:** 73–82

Sonoda Y, Cameron A, Newstead S, Omote H, Moriyama Y, Kasahara M, Iwata S, Drew D (2010) Tricks of the trade used to accelerate high-resolution structure determination of membrane proteins. *FEBS Lett*, **584:** 2539–2547

Sonoda Y, Newstead S, Hu NJ, Alguel Y, Nji E, Beis K, Yashiro S, Lee C, Leung J, Cameron AD, Byrne B, Iwata S, Drew D (2011) Benchmarking membrane protein detergent stability for improving throughput of high-resolution X-ray structures. *Structure*, **19:** 17–25

Studier FW (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif*, **41:** 207–234

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

Taglicht D, Padan E, Schuldiner S (1993) Proton-sodium stoichiometry of NhaA, an electrogenic antiporter from Escherichia coli. *J Biol Chem*, **268:** 5382–5387

Tang L, Bai L, Wang WH, Jiang T (2010) Crystal structure of the carnitine transporter and insights into the antiport mechanism. *Nat Struct Mol Biol*, **17:** 492–496

Tsuboi Y, Inoue H, Nakamura N, Kanazawa H (2003) Identification of membrane domains of the Na+/H+ antiporter (NhaA) protein from Helicobacter pylori required for ion transport and pH sensing. *J Biol Chem*, **278:** 21467–21473

Tzubery T, Rimon A, Padan E (2004) Mutation E252C increases drastically the Km value for Na+ and causes an alkaline shift of the pH dependence of NhaA Na+/H+ antiporter of Escherichia coli. *J Biol Chem*, **279:** 3265–3272

Ujwal R, Cascio D, Colletier JP, Faham S, Zhang J, Toro L, Ping P, Abramson J (2008) The crystal structure of mouse VDAC1 at 2.3 A resolution reveals mechanistic insights into metabolite gating. *Proc Natl Acad Sci U S A*, **105:** 17742–17747

Venturi M, Rimon A, Gerchman Y, Hunte C, Padan E, Michel H (2000) The monoclonal antibody 1F6 identifies a pH-dependent conformational change in the hydrophilic NH(2) terminus of NhaA Na(+)/H(+) antiporter of Escherichia coli. *J Biol Chem*, **275:** 4734–4742

Vinothkumar KR, Smits SH, Kuhlbrandt W (2005) pH-induced structural change in a sodium/proton antiporter from Methanococcus jannaschii. *EMBO J*, **24:** 2720–2729

Wagner S, Klepsch MM, Schlegel S, Appel A, Draheim R, Tarry M, Hogbom M, van Wijk KJ, Slotboom DJ, Persson JO, de Gier JW (2008) Tuning Escherichia coli

for membrane protein overexpression. *Proc Natl Acad Sci U S A*, **105:** 14371–14376
Walker JE, Saraste M, Gay NJ (1984) The unc operon. Nucleotide sequence, regulation and structure of ATP-synthase. *Biochim Biophys Acta*, **768:** 164–200
Wallin E, von Heijne G (1998) Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci*, **7:** 1029–1038
Walther DM, Rapaport D (2009) Biogenesis of mitochondrial outer membrane proteins. *Biochim Biophys Acta*, **1793:** 42–51
Wang DN, Safferling M, Lemieux MJ, Griffith H, Chen Y, Li XD (2003) Practical aspects of overexpressing bacterial secondary membrane transporters for structural studies. *Biochim Biophys Acta*, **1610:** 23–36
Wang Y, Zhang Y, Ha Y (2006) Crystal structure of a rhomboid family intramembrane protease. *Nature*, **444:** 179–180
Warne T, Serrano-Vega MJ, Baker JG, Moukhametzianov R, Edwards PC, Henderson R, Leslie AG, Tate CG, Schertler GF (2008) Structure of a beta(1)-adrenergic G-protein-coupled receptor. *Nature*, **454:** 486–491
Waser M, Hess-Bienz D, Davies K, Solioz M (1992) Cloning and disruption of a putative NaH-antiporter gene of Enterococcus hirae. *J Biol Chem*, **267:** 5396–5400
West IC, Mitchell P (1974) Proton/sodium ion antiport in Escherichia coli. *Biochem J*, **144:** 87–90
Weyand S, Shimamura T, Yajima S, Suzuki S, Mirza O, Krusong K, Carpenter EP, Rutherford NG, Hadden JM, O'Reilly J, Ma P, Saidijam M, Patching SG, Hope RJ, Norbertczak HT, Roach PC, Iwata S, Henderson PJ, Cameron AD (2008) Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter. *Science*, **322:** 709–713
Wheeler DL, Chappey C, Lash AE, Leipe DD, Madden TL, Schuler GD, Tatusova TA, Rapp BA (2000) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res*, **28:** 10–14
White SH (2009) Biophysical dissection of membrane proteins. *Nature*, **459:** 344–346
Wiebe CA, Dibattista ER, Fliegel L (2001) Functional role of polar amino acid residues in Na+/H+ exchangers. *Biochem J*, **357:** 1–10
Williams KA, Geldmacher-Kaufer U, Padan E, Schuldiner S, Kuhlbrandt W (1999) Projection structure of NhaA, a secondary transporter from Escherichia coli, at 4.0 A resolution. *EMBO J*, **18:** 3558–3563
Williams KA (2000) Three-dimensional structure of the ion-coupled transport protein NhaA. *Nature*, **403:** 112–115
Winn MD, Isupov MN, Murshudov GN (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr D Biol Crystallogr*, **57:** 122–133
Winter G (2010) xia2: an expert system for macromolecular crystallography data reduction. *Journal of Appl Crystallogr*, **43:** 186–190
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
Yamashita A, Singh SK, Kawate T, Jin Y, Gouaux E (2005) Crystal structure of a bacterial homologue of Na+/Cl--dependent neurotransmitter transporters. *Nature*, **437:** 215–223

Ye G, Chen C, Han D, Xiong X, Kong Y, Wan B, Yu L (2006) Cloning of a novel human NHEDC1 (Na+/H+ exchanger like domain containing 1) gene expressed specifically in testis. *Mol Biol Rep*, **33:** 175–180

Yernool D, Boudker O, Jin Y, Gouaux E (2004) Structure of a glutamate transporter homologue from Pyrococcus horikoshii. *Nature*, **431:** 811–818

Yin Y, He X, Szewczyk P, Nguyen T, Chang G (2006) Structure of the multidrug transporter EmrD from Escherichia coli. *Science*, **312:** 741–744

Zheng L, Kostrewa D, Berneche S, Winkler FK, Li XD (2004) The mechanism of ammonia transport based on the crystal structure of AmtB of Escherichia coli. *Proc Natl Acad Sci U S A*, **101:** 17090–17095

Zhou Y, Bowie JU (2000) Building a thermostable membrane protein. *J Biol Chem*, **275:** 6975–6979

Zhou Y, Morais-Cabral JH, Kaufman A, MacKinnon R (2001) Chemistry of ion coordination and hydration revealed by a K+ channel-Fab complex at 2.0 A resolution. *Nature*, **414:** 43–48

Zmasek CM, Eddy SR (2001) ATV: display and manipulation of annotated phylogenetic trees. *Bioinformatics*, **17:** 383–384

Appendix I

MemGold™ screen (Hamptons Research)

moleculardimensions.com

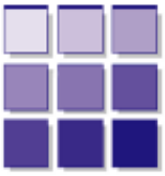
MemGold **Box 1 - Tubes 1-24** **MD1-39**

Tube #	Salt	Buffer	pH	Precipitant
1.1	None	0.08 M sodium citrate	5.2	2.2 M ammonium sulfate
1.2	None	0.01 M Tris	8.0	1.2 M tri-sodium citrate
1.3	None	0.015 M tricine	8.5	24 % w/v PEG 4000
1.4	0.36 M sodium chloride/0.1% w/v sodium azide	0.015 M sodium phosphate	7.0	9.9 % w/v PEG 4000
1.5	0.3 M sodium chloride	0.01 M Tris	8.0	27.5 % w/v PEG 4000
1.6	None	0.225 M MES/bis-tris	6.6	6.6 % w/v PEG 6000
1.7	0.1 M ammonium sulfate	0.1 M HEPES	7.5	12.0 % w/v PEG 4000/ 22 % v/v glycerol
1.8	0.02 M calcium chloride/0.01 M magnesium sulfate/0.02 M sodium chloride	0.02 M MES	6.5	7.7 % w/v PEG 1500
1.9	None	0.05 M HEPES	7.5	2.5 M ammonium sulfate
1.10	None	0.0665 M HEPES	7.5	1.1 M tri-sodium citrate
1.11	None	0.15 M potassium phosphate	6.5	3.3 M ammonium sulfate
1.12	0.1 M magnesium acetate	0.1 M sodium citrate	5.8	14 % w/v PEG 5000 MME
1.13	0.1 M sodium chloride	0.02 M sodium citrate	5.6	11 % w/v PEG 3350
1.14	0.1 M sodium chloride	0.02 M sodium citrate	5.6	5.5 % w/v PEG 3350
1.15	0.05 M calcium chloride/0.05 M barium chloride	0.1 M Tris	8.2	32 % v/v PEG 400
1.16	0.05 M sodium chloride	0.1 M sodium phosphate	6.2	16 % w/v PEG 4000
1.17*	0.1 M magnesium chloride	0.03 M Tris-hydrochloride	8.2	19 % w/v PEG 4000
1.18	0.2 M sodium chloride	0.025 M HEPES	7.5	13 % w/v PEG 4000
1.19	None	0.1 M HEPES	7.5	11 % w/v PEG 3350
1.20	0.1 M sodium chloride	0.02 M KMES	6.7	6.6 % w/v PEG 4000
1.21	0.1 M potassium chloride	0.02 M Tris	7.0	20 % w/v PEG 4000
1.22	0.05 M magnesium chloride/0.1% w/v sodium azide	0.1 M sodium cacodylate	6.7	6.6 % w/v PEG 3350
1.23*	0.2 M potassium chloride	0.1 M sodium citrate	5.5	37 % v/v pentaerythritol propoxylate (5/4 PO/OH)
1.24	None	0.1 M Tris	8.0	5.5 % w/v PEG 4000

moleculardimensions.com

MemGold **Box 1 - Tubes 25-48** **MD1-39**

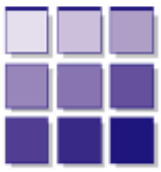
Tube #	Salt	Buffer	pH	Precipitant
1.25	0.1 M sodium chloride	0.02 M Tris	7.0	7.7 % w/v PEG 4000
1.26	0.1 M magnesium chloride	0.1 M Tris	7.5	22 % v/v PEG 400
1.27	0.04 M sodium chloride	0.04 M Tris	8.0	27 % v/v PEG 350 MME
1.28	0.05 M sodium chloride/0.02 M magnesium chloride	0.1 M sodium citrate	6.0	22 % v/v PEG 400
1.29	None	0.1 M sodium acetate	5.5	8.8 % w/v PEG 2000 MME
1.30	None	0.4 M ammonium acetate	8.0	13 % w/v PEG 2000 MME
1.31	None	0.02 M bis Tris	7.0	15 % w/v PEG 2000
1.32	0.1 M sodium chloride/0.1 M magnesium chloride	0.02 M Tris	7.5	11 % w/v PEG 1500
1.33	0.1 M sodium chloride/0.1 M magnesium chloride	0.1 M HEPES	8.0	11 % w/v PEG 1500
1.34	0.2 M sodium acetate/0.2 M Potassium Chloride	0.1 M HEPES	7.0	22 % w/v PEG 3000
1.35	0.02 M nickel sulfate	0.01 M HEPES	7.0	33 % v/v Jeffamine-M600
1.36	0.15 M sodium chloride	0.1 M Tris	8.0	13 % w/v PEG 6000
1.37	0.2 M calcium chloride	0.1 M HEPES	7.5	53 % v/v PEG 400
1.38	0.05 M magnesium acetate	0.05 M sodium acetate	5.0	28 % v/v PEG 400
1.39	None	0.05 M HEPES	7.5	22 % v/v PEG 4000
1.40	0.2 M calcium chloride	0.1 M Tris hydrochloride	8.0	44 % v/v PEG 400
1.41	0.05 M magnesium acetate	0.05 M sodium acetate	5.4	24 % v/v PEG 400
1.42	0.2 M calcium chloride	0.1 M MES	6.5	26 % v/v PEG 350 MME
1.43	0.1 M potassium chloride	0.1 M Tris	8.5	39 % v/v PEG 400
1.44	0.05 M magnesium chloride	0.1 M glycine	9.0	22 % v/v PEG 400
1.45	0.1 M ammonium sulfate	0.1 M glycine	3.8	28 % w/v tri-ethylene glycol
1.46	0.15 M sodium formate	0.1 M HEPES	7.2	18 % w/v PEG 3350
1.47	None	0.2 M sodium acetate	6.8	8.8 % w/v PEG 6000
1.48	0.2 M potassium chloride	0.1 M MES	6.5	18 % w/v PEG 6000



moleculardimensions.com

MemGold **Box 2 - Tubes 1-24** **MD1-39**

Tube #	Salt	Buffer	pH	Precipitant
2.1	0.22 M sodium citrate	0.1 M Tris	8.0	35 % v/v PEG 400
2.2	None	0.1 M sodium acetate	4.5	17 % v/v PEG 400
2.3	None	0.02 M Tris	8.5	1.0 M lithium sulfate/1.8 % w/v PEG 8000
2.4	None	0.02 M Tris	7.5	22 % v/v PEG 550 MME
2.5	0.05 M sodium chloride	0.02 M glycine	10.0	33 % w/v PEG 1000
2.6	0.2 M magnesium chloride	0.1 M Tris	8.5	25 % w/v PEG 4000
2.7	0.2 M magnesium chloride	0.1 M sodium cacodylate	6.5	31 % w/v PEG 2000
2.8	None	0.64 M sodium acetate	4.6	18 % w/v PEG 3350
2.9	0.1 M sodium chloride/0.1 M cadmium chloride	0.1 M Tris hydrochloride	8.0	33 % v/v PEG 400
2.10	None	0.1 M Bicine	8.9	31 % w/v PEG 2000
2.11	0.05 M sodium sulfate/0.05 M lithium sulfate	0.05 M Tris	8.5	35 % v/v PEG 400
2.12	0.1 M sodium chloride	0.05 M glycine	9.5	33 % v/v PEG 300
2.13	0.3 M magnesium nitrate	0.1 M Tris	8.0	23 % w/v PEG 2000
2.14	0.12 M lithium sulfate	0.02 M Tris/0.1 M sodium citrate	7.5/ 5.0	20 % v/v PEG 300
2.15	0.1 M sodium chloride	0.12 M Tris	9.4	20 % v/v PEG 400
2.16	0.2 M sodium chloride	0.1 M HEPES	7.0	22 % v/v PEG 550 MME
2.17	0.1 M sodium chloride/0.325 M sodium acetate	0.1 M Tris	8.0	21 % v/v PEG 400
2.18	0.02 M sodium citrate	0.08 M sodium phosphate	6.2	18 % w/v PEG 2000
2.19	0.02 M potassium nitrate	0.03 M potassium citrate	6.5	7.7 % w/v PEG 4000
2.20	0.1 M sodium chloride/0.005 M magnesium chloride	0.1 M Tris	8.5	30 % w/v PEG 2000 MME
2.21	0.2 M calcium chloride	0.1 M HEPES	7.0	33 % v/v PEG 400
2.22	0.1 M calcium chloride	0.1 M Tris	6.5	13 % w/v PEG 2000 MME
2.23	0.2 M ammonium sulfate/0.02 M sodium chloride	0.02 M sodium acetate	4.0	33 % v/v PEG 200
2.24	0.07 M sodium chloride	0.05 M sodium citrate	4.5	22 % v/v PEG 400



moleculardimensions.com

MemGold Box 2 - Tubes 25-48 MD1-39

Tube #	Salt	Buffer	pH	Precipitant
2.25	0.2 M ammonium sulfate	0.1 M sodium acetate	4.6	28 % v/v PEG 550 MME
2.26	None	0.05 M glycine	9.0	55 % v/v PEG 400
2.27	0.1 M magnesium chloride/0.1M sodium chloride	0.1 M Tris	8.5	33 % v/v PEG 400
2.28	0.1 M lithium sulfate/0.05 M di-sodium hydrogen phosphate	0.05 M citric acid	None	19 % w/v PEG 1000
2.29	0.2 M magnesium chloride/ 0.1 M potassium chloride	0.025 M sodium citrate	4.0	33 % v/v PEG 400
2.30	0.05 M zinc acetate	0.05 M MES	6.1	11 % w/v PEG 8000
2.31	0.3 M magnesium nitrate	0.1 M Tris	8.0	22 % w/v PEG 8000
2.32	0.1 M sodium chloride/4% v/v ethylene glycol	0.1 M MES	6.5	33 % v/v PEG 400
2.33	0.05 M sodium chloride	0.1 M sodium citrate	5.5	26 % v/v PEG 400
2.34	0.1 M lithium sulfate	0.1 M glycine	9.3	30 % v/v PEG 400
2.35*	0.15 M potassium citrate/ 0.05 M lithium citrate	0.1 M sodium phosphate	-	22 % w/v PEG 6000
2.36	0.001 M zinc sulfate	0.05 M HEPES	7.8	28 % v/v PEG 600
2.37	0.1 M sodium chloride	0.1 M sodium phosphate	7.0	33 % v/v PEG 300
2.38	0.1 M sodium chloride	0.05 M Bicine	9.0	33 % v/v PEG 300
2.39	0.05 M zinc acetate/6% v/v ethylene glycol	0.1 M sodium cacodylate	6.0	6.6 % w/v PEG 8000
2.40	0.2 M lithium sulfate	0.1 M sodium citrate	3.5	28 % v/v PEG 400
2.41	0.1 M sodium chloride	0.1 M Tris	7.5	11 % w/v PEG 4000
2.42*	0.05 M lithium sulfate	0.1 M tricine	7.4	7 % w/v PEG 3000
2.43*	0.2 M calcium chloride	0.1 M MES	6.5	33% v/v PEG 400
2.44*	1 M sodium chloride	0.1 M sodium citrate	6.0	28% w/v PEG 4000
2.45*	None	0.1 M HEPES	7.5	11% w/v PEG 4000
2.46*	0.002 M zinc sulfate	0.08 M HEPES	7.0	25 % v/v Jeffamine ED2001
2.47*	0.001 M cadmium chloride/0.03 M magnesium chloride	0.1 M MES	6.5	30 % v/v PEG 400
2.48*	None	0.1 M bis-tris-propane	7.0	3.0 M sodium chloride

- ***These conditions have been changed from the pre-release (prior to June 2007) beta version of MemGold. The pre-release conditions have been moved to the sister screen MemPlus (MD1-44), a new screen for Outer Membrane protein crystallisation."**

Abbreviations:
ADA; N-(2-Acetamido)iminodiacetic Acid, **Bicine**; N,N-Bis(2-hydroxyethyl)glycine, **CHES**; 2-(N-Cyclohexylamino)ethane sulfonic Acid, **HEPES**; N-(2-hydroxyethyl)-piperazine-N'-2-ethanesulfonic acid, **KMES**; 2-(N-morpholino)ethanesulfonic acid potassium salt, **MES**; 2-(N-morpholino)ethanesulfonic acid, **MME**; Monomethylether, **PEG**; Polyethylene glycol, **Tricine**; N-[Tris(hydroxymethyl)methyl]glycine, **Tris**; 2-Amino-2-(hydroxymethyl)propane-1,3-diol, **Tris HCl**; 2-Amino-2-(hydroxymethyl)propane-1,3-diol, hydrochloride.]

Appendix II

Table 7 – Expression screening of NhaA mutants using MemStar. All mutants contain the Ala109Thr mutation.

NhaA mutant	Standard (mg. L^{-1})	MemStar (mg. L^{-1})
Gln277His	14	135
Gln277Ala	13	123
Gln277Ser	14	137
Gln277Gly	11	126
Gln277Leu	16	111
Gln277Asp	13	131

Appendix III

The experimental work of this thesis is presented in the following manuscripts;

Chapter 3

Lee C, Kang HJ, Hjelm A, Choudhury H, Beis K, de Gier JW, Drew D.
MemStar: A new strategy for membrane protein production in *E. coli*. (Manuscript in preparation).

Chapter 4

Drew D, Beckstein O, Lee C, Yashiro S, Sansom M, Iwata S, Cameron A.D.
A salt-bridge in the sodium-proton antiporter NhaA acts as a sodium-sensitive electrostatic switch. (Manuscript in preparation).

Chapter 5

Lee C, Kang HJ, von Balmoos C, Newstead S, Uzdavinys P, Iwata S, Beckstein O, Cameron A.D., Drew D (2013) A two-domain elevator mechanism for sodium/proton antiport. *Nature,* (Manuscript under revision).