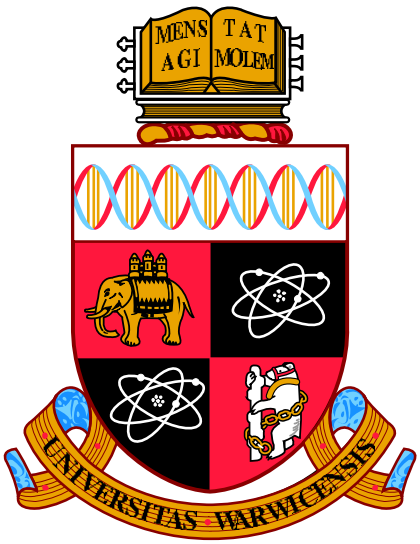
Structural and functional characterization of a bile acid sodium symporter from *Neisseria meningitidis*



by

Patrick Becker

School of Life Sciences
University of Warwick

A thesis submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Biological Sciences

United Kingdom
May 2021

Structural and functional characterization of a bile acid sodium symporter from *Neisseria meningitidis*

Patrick Becker

Department of Life Sciences, University of Warwick
Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Abstract

The apical sodium-dependent bile acid transporter from *Neisseria meningitidis* ($ASBT_{NM}$) shares 26% identity and 54% similarity to human ASBT which plays a key role in the reabsorption of bile acids in the terminal ileum and has gained interest as a pharmaceutical target. The structure of $ASBT_{NM}$ was solved at 2.2 Å (Hu et al. 2011) and can be used as a model to understand the transport mechanism of the human protein. Nevertheless, more structural and functional information are needed to elucidate the transport mechanism.
In this thesis, three different structures of the wild type and an E260A mutant are discussed which show different sodium-binding states in the inward-facing conformation. The structures showed that taurocholate-release and sodium-release do not induce major conformational changes in the protein. Furthermore, the E260A structure showed some differences in the positioning of the core and panel domain to each other indicating the beginning of a rotation movement. TM6 was shown to be able to unwind at the extracellular part and contains a potential hinge region allowing an outward movement in the outward-facing conformation to open the substrate binding site. Two liposome-based transport assays were performed. Both showed a time-dependent uptake of taurocholate by $ASBT_{NM}$. Sodium depletion did not abolish taurocholate transport which could indicate that the transporter does not work in a completely stringent sodium-dependent manner. Unfortunately, both assays have not been successful when the transport activity of different $ASBT_{NM}$ mutants was investigated. A wild type $ASBT_{NM}$ structure with pantoate bound was solved at 2.6 Å. Pantoate was shown to bind to the wild type protein with an estimated K_D of 124 μM.
The structural information from this thesis will help to further understand the transport mechanism of hASBT as it adds further elucidated conformations (inward-facing) to a complete transport cycle. The identification of the pantoate binding site in $ASBT_{NM}$ can help to understand the ligand recognition of the plant BASS transporters which share homology to $ASBT_{NM}$.

Declaration

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy.
I hereby declare that this thesis has been composed by myself under the supervision of Professor Alex Cameron. The thesis has not been previously submitted for a degree at another university. All sources of information are acknowledged in the form of references.

Patrick Becker

May 2021

Acknowledgements

First of all, I would like to express my deepest appreciation to my supervisor, Professor Alex Cameron, who gave me the opportunity to conduct my PhD research in his group. He was always there for guidance and advice. He always found some encouraging words if experiments have not worked as they were supposed to be.

I am also extremely grateful to Professor Oliver Beckstein and the NIH for funding the project.

I would also like to thank my advisory panel, Dr Corinne Smith and Professor Ioannis Nezis, who allowed me to discuss my research with them and gave me good advice. They also supported me to develop my transferable skills.

A special thanks to Dr Deborah Brotherton for her supervision in the laboratory and who have taught me the transport assay. She was always there in the laboratory for scientific advice and troubleshooting if something was not working.

Thanks also to my group members, Caitlin Hatton and Aurelien Grob.

Thanks should also go to all C10 lab members, especially to two former members, Mary and Raman.

Last, I want to thank my family and friends who always supported me and cheered me up if something was not working. They were always there for me either on phone or in person if I needed someone to talk to.

Contents

List of Figures

List of Tables

List of abbreviations

ABC	ATP-binding cassette
ASBT	Apical sodium-dependent bile acid transporter
ATP	Adenosine triphosphate
BASS	Bile Acid Sodium Symporter
BSEP	Bile salt export pump
CD	Circular dichroism
CMC	Critical micelle concentration
CPM	7-Diethylamino-3-(4'-Maleimidyl-phenyl)-4-Methylcoumarin
CV	Column volumes
DDM	n-Dodecyl β-D-maltoside
DM	n-Decyl β-D-maltoside
E. coli	*Escherichia coli*
EL	Extracellular loop
FT	flow through
GFP	Green fluorescent protein
GLUT	Glucose transporter family
IBABP	Ileal bile acid binding protein
IL	Intracellular loop
IMAC	Immobilized metal affinity chromatography
IPTG	Isopropyl-β-D-thiogalactopyranosid
ITC	Isothermal titration calorimetry

L-BABP	Liver bile acid binding protein
LB	Lysogeny broth
LCP	Lipidic cubic phase
LDAO	Lauryldimethylamine-N-Oxide
MFS	Major facility superfamily
MST	Microscale thermophoresis
MW	Molecular weight
Na1	Sodium binding site 1
Na2	Sodium binding site 2
NCS	Non crystallographic symmetry
NM	n-Nonyl β-D-maltoside
NSS	Neurotransmitter sodium symporter
NTCP	Na^+-taurocholate co-transporting polypeptide
OD	Optical density
OG	n-Octyl-β-D-Glucopyranoside
OST	Organic solute transporter
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
rIMAC	Reverse immobilized metal affinity chromatography
RMSD	Root-mean-square deviation
RT	Room temperature
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography
SLC	Solute carrier

SMALPs	Styrene maleic-acid lipid particles
SOAT	Sodium-dependent organic anion transporter
SOC	Super Optimal broth with Catabolite repression
TAE	Tris-acetate-EDTA
TEV	Tobacco etch virus
TM	Transmembrane helix
WT	Wild type

Chapter 1: Introduction

The foundation of protein crystallization reaches back more than 180 years with the discovery that proteins own the ability to form crystals by Hünefeld in 1840 (Hunefeld 1840, McPherson 1991). He obtained hemoglobin crystals when he left worm blood slowly drying between two glass slides. More than 100 years later, the first X-ray structure was obtained from a myoglobin crystal by Kendrew et al. who received the Nobel prize in Chemistry in 1962 together with Max Perutz for it (J. Kendrew et al. 1960, John C Kendrew et al. 1958). From then on, researchers became interested in the elucidation of protein structures by X-ray crystallography and in the idea that the protein structure is linked to its function and vice versa (McPherson 2004). Within the time of 20 years, around 50 new protein structures were solved which were all from soluble proteins (Jaskolski et al. 2014).

Surprisingly, no membrane protein structure has been reported up to then even though roughly 30% of the genes in most organisms are encoding for membrane proteins (Krogh et al. 2001). Located in the membrane at the interface between the outside of the cell and the inside, membrane proteins play crucial roles in solute transport across the membrane, cell signaling, and cell adhesion just to name a few. Furthermore, a lot of diseases are associated with the malfunction of membrane proteins. Therefore, it sounds surprising that only soluble proteins have been crystallized until the mid of 1980s (Jaskolski et al. 2014).

The reason for it is that it is more difficult to work with membrane proteins because of the following: 1. Membrane proteins usually show a lower expression which is related to their toxic effect on the membrane. When they are too highly overexpressed, they can make the membrane leaky and can lead to the cell collapsing. 2. Due to their membrane location, they require a more sophisticated purification process and the usage of detergent to avoid aggregation because of their hydrophobic parts. 3. The detergent micelle hinders tight crystal packing as it takes up a lot of space which usually results in small and fragile membrane protein crystals which do not diffract well (Carpenter et al. 2008). Nevertheless, because of their importance the structural characterization of them is fundamental and therefore, researchers started to work on them and have succeeded with the first membrane protein structure in 1980 of rhodopsin

from *Halobacterium halobium* by Hartmut Michel and Dieter Oesterhelt (Michel et al. 1980). From then on, more and more membrane protein structures have been elucidated helping to understand their function. Here, the focus will lie on membrane transporters which allow the transport of solutes across a biological membrane.

1.1 Transport across biological membranes

Biological membranes consist of a lipid bilayer and act as a semi-permeable enclosure allowing the creation of specific chemical environments in their inside. They allow the exchange of some small hydrophobic molecules and water but are impermeable to most other compounds. Therefore, membranes are crucial for the protection of the cell from any stimulus from the outside, but also for the creation of internal compartments which can have their own chemical environment such as the lysosome of eukaryotic cells which needs an acidic pH in the inside to perform its function (De Duve et al. 1966).

While acting as a protection, the membrane also causes problems for the cell. As it is impermeable to the majority of compounds, the cell needs specific transport systems to allow the exchange of solutes e.g. for metabolic and anabolic reasons or signaling reasons. Figure 1.1 shows the basic types of membrane transport (Stillwell 2013, Alberts et al. 2015).

Simple passive diffusion is the easiest type and describes the movement of a solute across the membrane because of its own concentration gradient. Molecules which can diffuse through the membrane by passive diffusion are water, small non-charged hydrophobic substances and gasses (Baldwin 2000).

For larger molecules or charged compounds like ions, a second type of diffusion can occur which is facilitated by an integral membrane protein (facilitated diffusion). Two types of facilitated diffusion are known: The first one is mediated by channel proteins which show a high specificity for different ions or other small compounds and they allow them to cross the membrane according to their concentration gradient by forming small pores through the membrane. Examples for channel proteins are aquaporins (Agre 2006) which facilitate the faster exchange

of water through the membrane or any type of ion channel like potassium channels (Lippiat 2008).

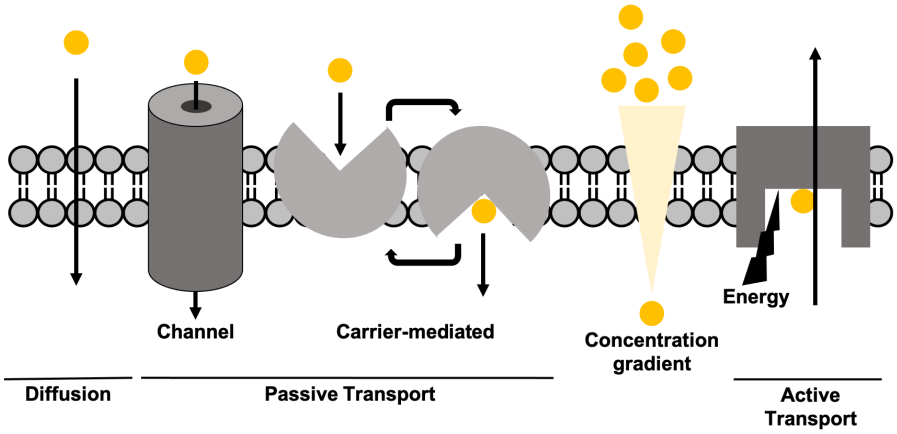


Figure 1.1: Schematic representation of transport types across a biological membrane: The different types are shown and labelled. Yellow circles represent the substrate. The different grey objects represent the possible integral membrane transport systems. The faint yellow triangle shows the concentration gradient of the substrate.

The second type of facilitated diffusion is carrier-mediated. In contrast to channels which open and close their pore to allow transport, carriers undergo more structural changes during the transport. They adopt at least two different conformations during their transport cycle with the substrate binding site only being accessible from one site of the membrane at a time. An example for the carrier-mediated diffusion is the glucose transporter family (GLUT) responsible for the glucose uptake into the cell (Mueckler et al. 2013).
While passive transport only occurs along the concentration gradient of the transported molecule, most often transport needs to occur against the concentration gradient leading to the accumulation of molecules on one site of the membrane. This is referred as active transport for which energy is necessary.

1.1.1 Active transport

Active transport can be divided into primary and secondary active transport (see figure 1.2) (Stillwell 2013, Nelson et al. 2008). Primary active transport systems

use the energy from ATP hydrolysis to drive the transport of their substrate against its concentration gradient. Examples for primary active transport systems are ATP-binding cassette transporters (ABC transporters) like McjD which facilitates self-immunity of the host against a toxin which is produced by itself. McjD pumps the self-produced toxin out the cell and therefore keeps its concentration to a low level in the cell that it is not toxic to the cell. (Choudhury et al. 2014). Another well-known example are the Na^+/K^+-ATPases which use the energy of ATP to create sodium or proton gradients across the membrane (Skou et al. 1992).

These produced ion gradients across the membrane are the energy source of secondary active transporters which transport their substrates across the membrane by co-transporting ions and using the electrochemical energy stored in their concentration gradients. It can be differentiated between symport and antiport. A symporter transports its substrate and the ion in the same direction, while an antiporter transports both in opposite directions (Nelson et al. 2008).

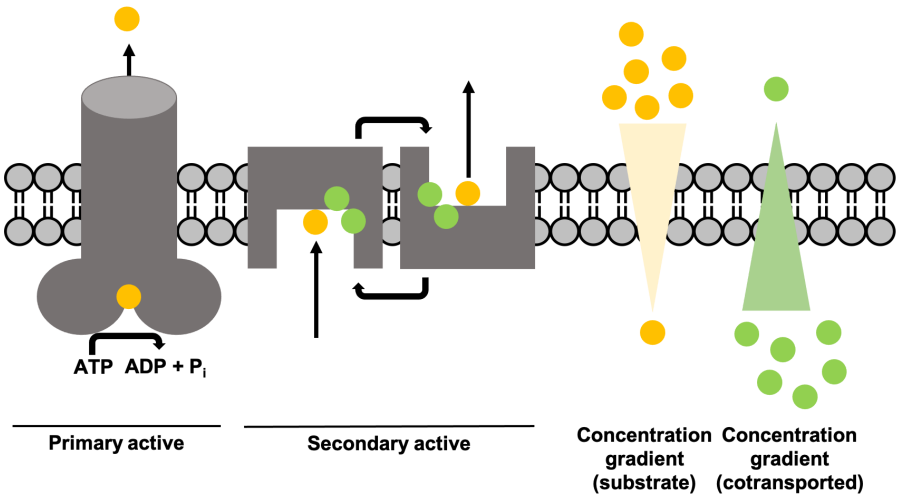


Figure 1.2: Schematic representation of active transport types across a biological membrane: The different types are shown and labelled. Yellow circles represent the substrate, while green circles indicate the co-transported ions. The different grey objects represent the possible integral membrane transport systems. The faint yellow triangle shows the concentration gradient of the substrate, while the faint green triangle shows the concentration gradient of the co-transported ion.

1.1.2 The alternating access mechanism

A transport model for secondary active transporters was developed by Jardeztky in 1966 (Jardetzky 1966). The alternating access model proposes that the transporter switches between an outward-open and an inward-open conformation in which it presents the binding site only to one site of the membrane at a time. For transport to occur, the transporter needs to undergo a series of different conformations. As an example, the different conformations of a sodium symporter will be discussed throughout a transport cycle (see figure 1.3).

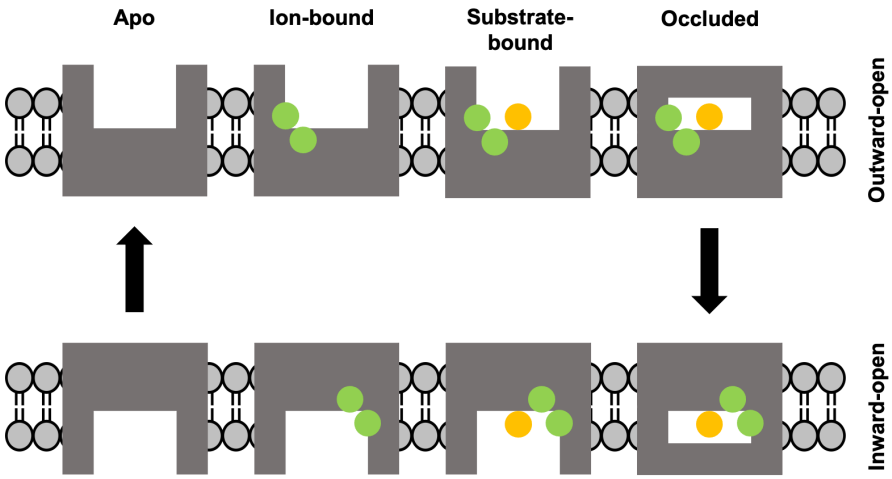


Figure 1.3: Model of the alternating-access mechanism: Different conformations for a whole transport cycle are shown for a symporter. The transporter is shown in grey. The substrate is represented as a yellow circle and its co-transported ions as green circles. The different conformations are labeled above for the outward-open conformation (upper membrane) and for the inward-open states below (lower membrane).

At the beginning, the transporter is in an outward-open apo state. The co-transported ions bind first and prime the binding site for the substrate by increasing the affinity of the transporter for its substrate. This is a common feature of transporters which have a lower affinity when the co-transported ions are not bound and exhibit a higher affinity when they are bound (Hänelt et al. 2015). Next, the substrate binds to the transporter (outward-open substrate bound state) which will lead to the closure of the substrate binding site (outward occluded state), often performed by small, local changes around the binding site.

Then, the transporter will undergo major conformational changes and swap into an inward-occluded state, followed by the opening of the substrate binding site (inward-open substrate bound state). Now, the substrate can be released. Finally, the co-transported ions leave the transporter (ion concentration is lower on this site) and the transporter can switch back again into the outward-open conformation (Jardetzky 1966, Tanford 1982, Forrest et al. 2011).

Depending on how the alternating access of the substrate binding site is facilitated, it can be differentiated between three different mechanisms (Ryan et al. 2016, Jeckelmann et al. 2020). Figure 1.4 shows the three potential transport mechanisms. In the rocker-switch mechanism, two homologous domains are moving around between a central substrate-binding site. This type is mainly seen in the major facilitator superfamily (MFS) with the best studied example being LacY, a lactose/proton symporter from archaea (Kaback et al. 2011, Yan 2015). In the rocking bundle mechanism, the transporter shows two heterologous domains, a core or bundle domain and a scaffold domain. The core domain undergoes a rigid-body movement around the scaffold domain to present the binding site to both sites of the membrane. An example of transporters which operate in this manner are the neurotransmitter sodium symporter (NSS) family with LeuT being the best studied one (Forrest et al. 2009, Loland 2015). Last, the core domain can undergo an up- and down movement in regards to the scaffold domain which is referred as the elevator mechanism. The glutamate transporter homolog Glt_{Ph} and the sodium proton antiporter from *Thermus thermophilus* (NapA) are examples of transporters which use the elevator mechanism (Reyes et al. 2009, Coincon et al. 2016, Lee et al. 2013). The core domain of Glt_{Ph} undergoes a vertical movement of about 16Å and rotates about 37° between the outward and inward facing conformation. For NapA, a similar elevator mechanism was observed with a smaller vertical movement of about 7-10 Å and it shows only a minimal rotation movement (Ryan et al. 2016).

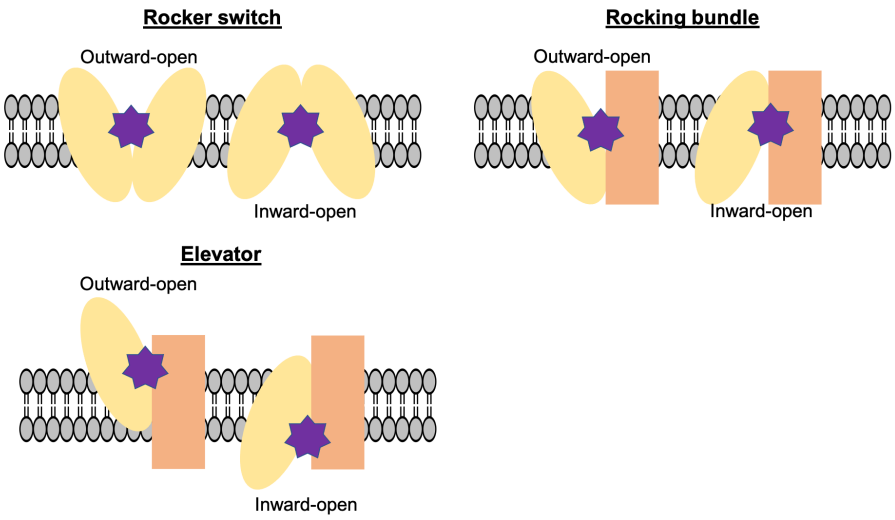


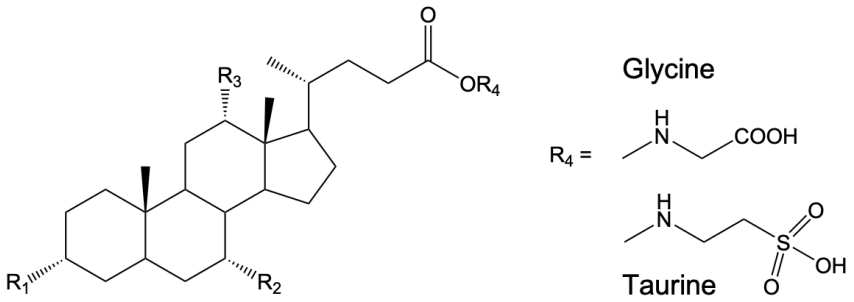
Figure 1.4: Schematic representation of different alternating-access mechanisms: Domains are colored in yellow and beige. The violet star indicates the substrate binding site. In the rocker switch mechanism, two homologues domains are moving. In the, rocking bundle and elevator mechanism, the core domain (yellow) moves around the scaffold domain (beige).

1.2 Bile acids and the SLC10 transporter family

Bile acids are amphipathic molecules consisting of a hydrophobic steroid backbone and a hydrophilic site which is made of several hydroxyl groups and a polar tail group (Alan F Hofmann 1999). The hydroxylation pattern and the tail group vary among different bile acids (see figure 1.5). Cholic acid and chenodeoxycholic acid are classified as primary bile acids as they are the first products of bile acid synthesis (see next section for more information). Chenodeoxycholic acid has one less hydroxyl group at position 12 in comparison to cholic acid. Both contain a free carboxyl group at the tail. Primary bile acids have low solubility in water and therefore they are mainly present as conjugates with glycine and taurine in the human bile which greatly increase their solubility and reduce the passive absorption of them (Stamp et al. 2008, Hofmann et al. 1992). These bile acids are referred as conjugated bile acids and are named with a prefix indicating their conjugated tail group, either with tauro- for taurine or glyco- for glycine. Other bile acids can be produced by bacteria in the intestine by removal of different hydroxyl groups and are called secondary bile acids with the most common being deoxycholic acid which misses the hydroxyl group at position 7 (Lewis et al.

1972).

Overall, these six conjugated bile acids make up for most of the human bile with glycocholic acid and glycochenodeoxycholate being the top in human, each making up of around 30% (Stamp et al. 2008, A. Hofmann et al. 1967).



Name	R1	R2	R3	R4	% in bile*
Cholic acid	-OH	-OH	-OH	-H	Trace
Chenodeoxycholic acid	-OH	-OH	-	-H	Trace
Deoxycholic acid	-OH	-	-OH	-H	Trace
Glyco-/taurocholic acid	-OH	-OH	-OH	Glycine/taurine	30/10
Glyco-/taurochenodeoxycholic acid	-OH	-OH	-	Glycine/taurine	30/5
Glyco-/taurodeoxycholic acid	-OH	-	-OH	Glycine/taurine	15/10

Figure 1.5: Chemical structure of bile acids and their names: The general chemical structure of bile acids is shown with the four alternative functional groups marked with R_1 to R_4. R_1 indicates position 3, R_2 - position 7, and R_3 - position 12. R_4 indicates the potential tail groups with the structure of taurine and glycine shown. The table shows the names of the bile acids according to their functional groups at the different positions. (* Stamp et al. 2008)

1.2.1 Bile acid synthesis

Bile acids are synthesized in the liver from cholesterol by a series of 17 enzymatic reactions (Chiang 2013, Chiang 1998, Šarenac et al. 2018, P. A. Dawson 2018) which are shown in figure 1.6. In a first rate-limiting step, the cholesterol is hydroxylated by cholesterol 7a-hxydroxylase (CYP7A1). This is followed by a series of modifications of the steroid backbone and the formation of two different products as the hydroxylation at position 12 is not mandatory for the pathway

to proceed. The two products show the typical bile acid steroid backbone with either two or three hydroxyl groups (see figure 1.6 for more information) but still contain the isooctane tail of the cholesterol.

In a next series of steps, the side chain undergoes an oxidative cleavage forming as end products the primary bile acids, cholic acid and chenodeoxycholic acid. In a final step, the primary bile acids are conjugated to taurine and glycine via an amide bond and are stored in the bile duct.

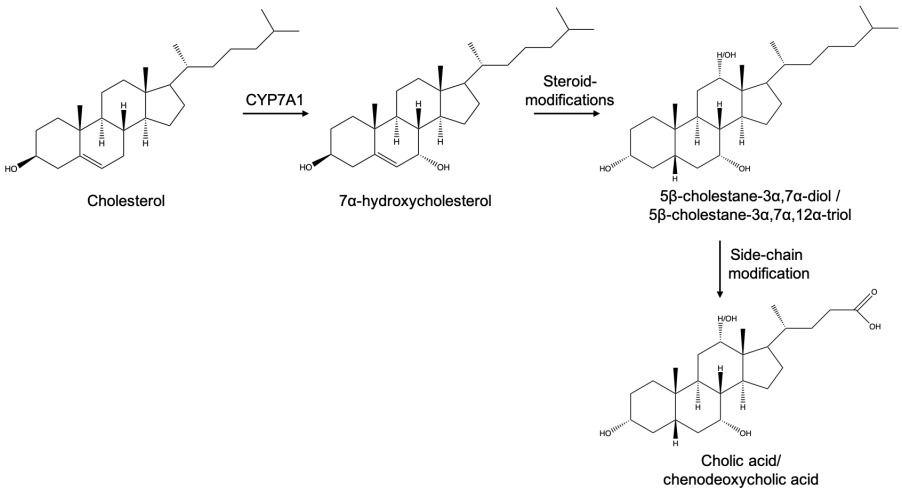


Figure 1.6: Simplified bile acid synthesis from cholesterol showing the key modifications: Cholesterol is hydroxylated by cholesterol 7a-hxydroxylase (CYP7A1). Next, the sterol backbone is modified caused by a series of enzymatic reactions leading to the formation of 5-beta-cholestane-3-alpha, 7-alpha-diol or 5-beta-cholestane-3-alpha, 7-alpha, 12-alpha -triol. These two intermediates are further modified at the tail to form the primary bile acids, cholic acid and chenodeoxycholic acid.

1.2.2 Bile acid function and the enterohepatic circulation

Upon food intake, bile acids are released into the intestine via the bile duct, where they enhance the solubilization and absorption of hydrophobic substances like lipids, cholesterol, and hydrophobic vitamins such as vitamin A, D, E, and K due to their amphipathic properties (Stamp et al. 2008). Due to their cost intensive synthesis, bile acids are recycled in humans once they have performed their function. This recycling pathway is known as the enterohepatic circulation

(Roberts et al. 2002, P. A. Dawson 2018) which is shown in figure 1.7.

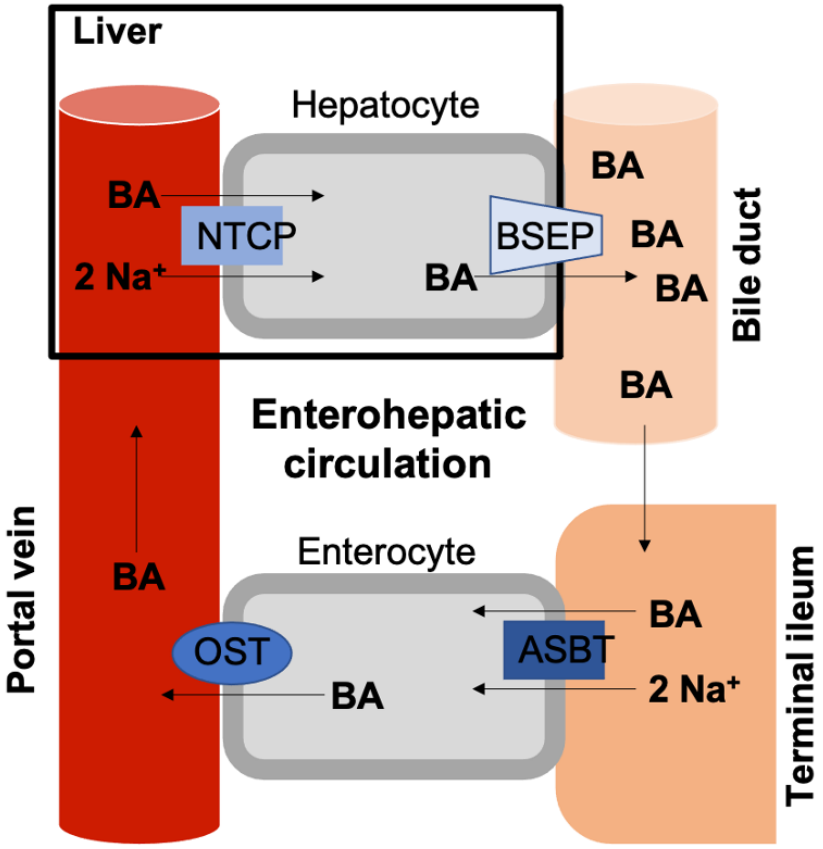


Figure 1.7: Schematic representation of the enterohepatic circulation: Bile acids (BA) are transported into enterocytes of the terminal ileum by ASBT in a sodium-dependent manner. On the basolateral membrane, the bile acids are transported into the blood stream by the OST complex. Through the blood stream, bile acids reach the liver and are transported into hepatocytes by NTCP in a sodium-dependent manner. BSEP is responsible for the active transport of bile acids into the bile duct by the usage of ATP. Bile acids are stored in the bile duct.

In the terminal ileum, bile acids are transported into the cell by the apical sodium dependent transporter (ASBT) which uses the energy of the sodium gradient across the membrane to transport one bile acid molecule together with two sodium ions into the cell (Craddock et al. 1998). The transporter will be discussed in more detail in the next section. Once in the cell, bile acids bind to the ileal bile acid binding protein (IBABP) (Kramer et al. 1993) and are transferred to the basolateral membrane where they are effluxed into the portal vein by the

organic solute transporter (OST) which consists of two different subunits (alpha and beta) and was shown to be sodium-independent (P. A. Dawson et al. 2010). Through the blood stream, bile acids reach the liver where they are transported into hepatocytes by the Na^+-taurocholate co-transporting polypeptide (NTCP) which is a close homologue of ASBT and similarly works in a sodium-dependent manner (Kouzuki et al. 1998). In the cell, they bind to the liver bile acid-binding protein (L-BABP) and are shuttled to the canicular membrane where they are actively transported into the bile duct by the bile salt export pump (BSEP) (Agellon et al. 2000, Monaco 2009, BBON Stieger et al. 1992). This transport is ATP-dependent and therefore allows to build up high concentrations of bile acids ($> 300\,\mathrm{mM}$) in the bile duct (Stamp et al. 2008). The bile acid pool of an adult human is approximately 3 g and this pool can turnover four to twelve times a day with a daily loss of less than 5% which indicates the high efficiency of the transporters involved, especially hASBT which performs the first step of salvage (John YL Chiang 2013).

1.2.3 The solute carrier 10 (SLC10) transporter family

NTCP and ASBT belong both to the solute carrier (SLC) family 10 which plays important roles in the transport of bile acids, drug, steroidal hormones, and other substrates (Da Silva et al. 2013, Geyer et al. 2006). Both were the first characterized members of the family and have been well studied (Hagenbuch et al. 1990, Wong et al. 1994). Therefore, the family is often referred as the ''sodium bile salt cotransport family" (Hagenbuch et al. 2004). But this name has been questioned recently due to the identification of SLC10A6 (Doring et al. 2006). The sodium-dependent organic anion transporter (SOAT) or SLC10A6 was shown to mainly transport sulfated steroids like estrone sulfate or estradiol sulfate and not bile acids despite the fact that it shares high similarity to ASBT (69.7%) and NTCP (62.3%) and was therefore assigned to this family (Da Silva et al. 2013). Beside these three members, the family contains four more members which are all orphan transporters with no known substrate. SLC10A3 or P3 is predicted to be ubiquitously expressed in humans due to its promoter region but its expression has yet to be verified (Alcaiay et al. 1988). Overall, it shares only 27%

identity with ASBT and NTCP (Da Silva et al. 2013, Hagenbuch et al. 2004). SLC10A4 or P4 is expressed in vesicles in neurons and is thought to play a role in vesicular storage but does not show transport of the typical substrates like bile acids or sulfate steroid (Geyer, C. Fernandes, et al. 2008). SLC10A5 or P5 is less characterized than the other members. It shows a high similarity with over 70% to other members and was shown to be expressed on the cell surface, mainly in the liver, indicating a role in solute transport similar to NTCP but no functional data is available for it (Carla F Fernandes et al. 2007). The last member is SLC10A7 which is expressed in the liver and testis. Interestingly, SLC10A7 shows a high sequence conservation among mammals and bacteria which might indicate a pivotal role (Godoy et al. 2007). Karakus et al described SLC10A7 as a regulator of calcium signaling, but it is not clear yet how it affects the calcium homeostasis and what it transports (Karakus et al. 2020).

1.2.4 NTCP and ASBT

hASBT (SLC10A2) and NTCP (SLC10A1) play important roles within the enterohepatic circulation (see figure 1.7). Both transport bile acids into the cell by using the energy of the sodium-gradient with a stoichiometry of two sodium ions per one substrate molecule (Hagenbuch et al. 1996, Weinman 1997). hASBT shows its highest expression in enterocytes of the terminal ileum where is potentially exposed to a variety of not naturally produced substances which can infiltrate the body via food uptake (Shneider et al. 1995). It is likely that this is the reason why hASBT shows a higher substrate specificity to avoid the uptake of any harmful substances. hASBT only accepts conjugated bile acids and primary bile acids as its substrates. On the other hand, NTCP is expressed in hepatocytes which are more protected from the environment. It shows a broader substrate specificity including bile acids, sulfated steroid hormones, bromosulphthalein, and xenobiotics (Bruno Stieger 2011).

1.2.5 Pharmaceutical interest of hASBT

Human ASBT has gained interest as a pharmaceutical target because of two reasons. First, as it is the entry point for bile acids into the enterohepatic circulation,

it was found that inhibition of hASBT leaves to the excretion of bile acids and a lower bile acid concentration in the blood (West et al. 2003). This causes the upregulation of bile acid synthesis in the liver from cholesterol which decreases the overall cholesterol concentration in the body. Hence, inhibitors for ASBT were investigated for treatment of hypercholesterolemia.
Secondly, ASBT has been discussed as a prodrug carrier (Balakrishnan et al. 2006). A lot of drugs show poor absorption when they are orally administrated. To increase the absorption and the bioavailability of a drug, they can be conjugated to a substance which is a substrate for an intestinal transporter. An example is acyclovir which is an antiviral medication. When it is conjugated to a valine (valacyclovir), the absorption could be enhanced by 2- to 3-fold as the valine-conjugate is a substrate for the intestinal peptide transporter. Similar results were obtained when acyclovir was conjugated to chenodeoxycholic acid. Acyclovir valylchenodeoxycholate was transported by hASBT with a K_I of around 36μM.

1.3 ASBT from *Neisseria meningitides*

The structure of ASBT from *Neisseria meningitides* ($ASBT_{NM}$) was solved at a resolution of 2.2 Å (Hu et al. 2011). The protein shares 26% identity and 54% similarity with its human homologue. All functional important residues are conserved among $ASBT_{NM}$, hASBT, and NTCP (see figure 1.8) and therefore the protein was considered as a good model to study the human protein. Additionally, the protein was shown to have similar transport kinetics with an estimated K_M of around 50μM for taurocholate in a cell-based transport assay.

The structure showed a similar fold than NhaA and NapA with the protein consisting of ten transmembrane helices (TM) with the last five being an 'inverted repeat" of the first five helices (see figure 1.9b) (Hu et al. 2011, Hunte et al. 2005, Lee et al. 2013). The protein can be divided into a smaller panel domain and a larger core domain (see figure 1.9a). The panel domain is made of TM 1/2 and TM 6/7 which each of the pair form a V-motif. The core domain consists

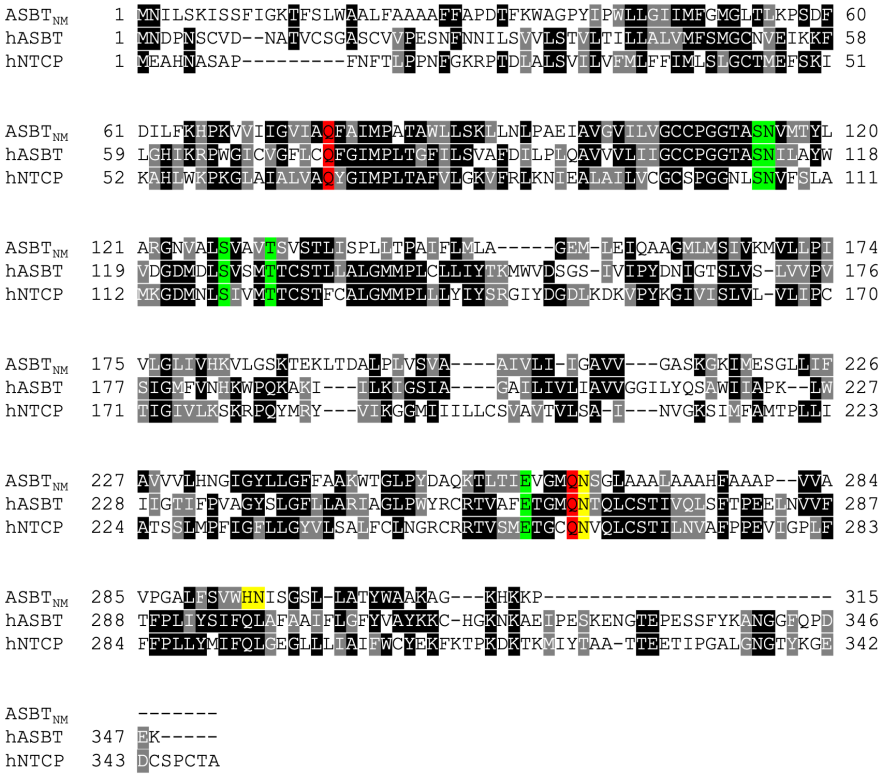


Figure 1.8: Sequence alignment of $ASBT_{NM}$, hASBT, and NTCP: Sequence alignment of $ASBT_{NM}$ (Uniprot: Q9K0A9) to hASBT (Uniprot: Q12908), and NTCP (Uniprot: Q14973) using T-coffee (*https* : *//www.ebi.ac.uk/Tools/msa/tcoffee/*) and Boxshade (*https* : *//embnet.vital − it.ch/software/BOX_form.html*). Identical residues are shown in black boxes while functionally conserved residues are shown in grey. Residues of the sodium sites are labelled in red (Na2) and green (Na1). Taurocholate interacting residues are highlighted in yellow.

of six transmembrane helices with TM4 and 9 being discontinuous and crossing each other approximately at the middle of the membrane. The panel and core domain form a cavity at the intracellular site between each other while they are closely interacting on the extracellular site. Hence, the structure was assigned as an inward-facing substrate-bound conformation.

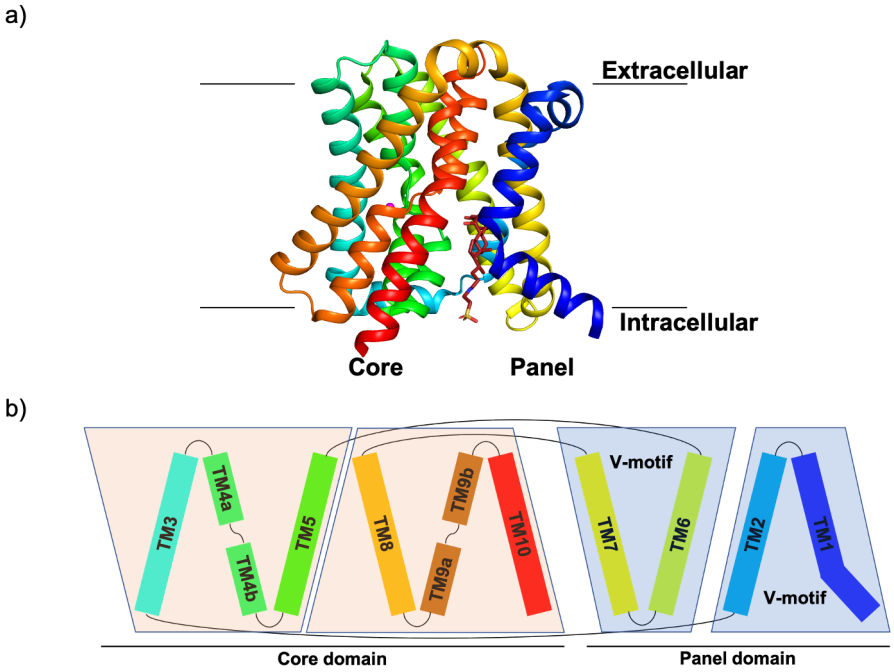


Figure 1.9: Cartoon representation and topology map of $ASBT_{NM}$: a) Side view of the $ASBT_{NM}$ wild type structure (3zuy) in cartoon representation. Helices are colored differently starting with TM1 in dark blue to TM10 in red. Sodium ions are shown as pink spheres. The taurocholate molecule is shown as sticks (firebrick). The membrane is shown as lines with the extracellular and intracellular site labeled. The core and panel domains are indicated. b) Topology map of $ASBT_{NM}$ showing all helices and the connecting loops. The core domain is shown on the left. The panel domain is shown on the right with the two V-motifs.

1.3.1 Sodium binding sites

Two sodium ion sites were identified and named Na1 and Na2 respectively (see figure 1.10). Both sites are located in the core domain. Na1 lies closer to the cytosol compared to Na2 and shows a hexa-coordination with the sodium ion made of the side chains of Ser114, Asn115, Ser128, Thr132, and Glu260 together with the carbonyl oxygen of Ser128. Na2 is located above Na1 approximately in the

middle of the protein where TM4 and 9 cross each other. Here, the sodium ion is penta-coordinated with only two side chains of Gln77 and Gln264 mediating contact and the other contacts are made by backbone carbonyl oxygens of Glu260, Val261, and Met263. Glu260 is the only residue which is involved in both sodium sites and might play a role of connecting the sites to each other.

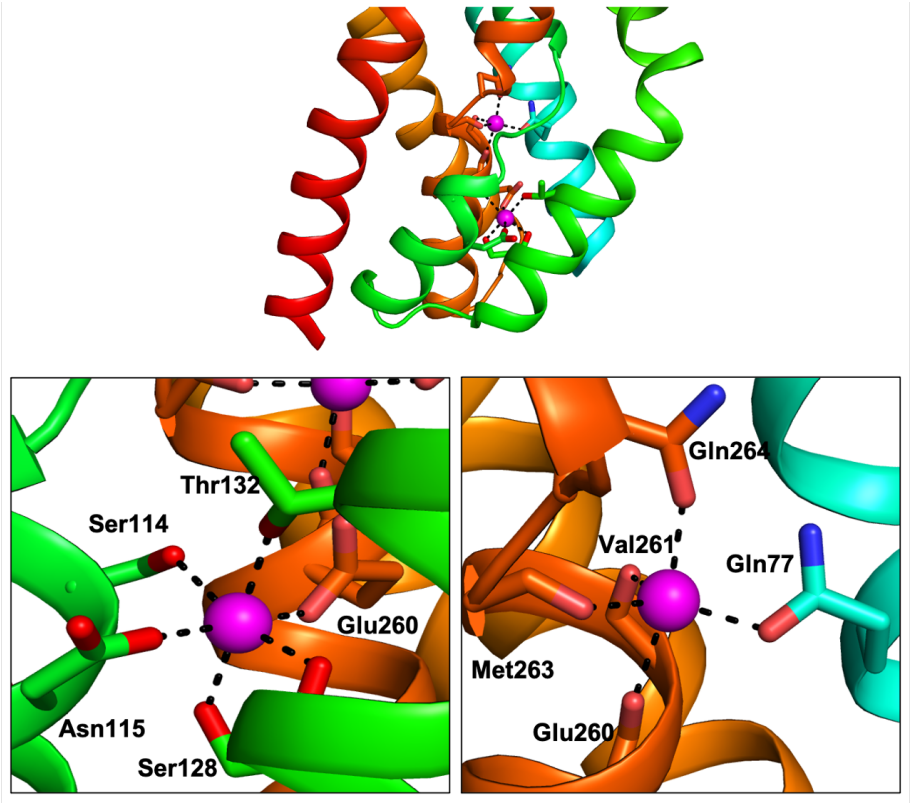


Figure 1.10: Sodium binding sites in $ASBT_{NM}$: A part of the core domain is shown (top) with the two sodium ions shown as pink spheres. Enlargement of Na1 (left) and Na2 (right) with residues interacting with the sodium ions are shown as sticks and labelled with their name and position.

1.3.2 Substrate binding pocket

One taurocholate molecule was found in the intracellular cavity between the panel and core domain (see figure 1.11). The cavity is lined by a lot of hydrophobic residues favoring the interaction with the hydrophobic cholesterol backbone of the taurocholate molecule and the hydrophilic taurine tail pointing towards the cytosolic site. Four residues in the binding pocket are interacting with the three

hydroxyl groups of the molecule. Gln295 and Thr112 form direct hydrogen bonds with the 7-hydroxyl and 3-hydroxyl group while His294 and Gln265 mediating the contact over a water molecule with the 7- and 3-hydroxyl group respectively. No interaction was observed for the hydroxyl group at position 12. All residues which are interacting with the taurocholate molecule are located in the core domain close to the sodium binding sites. Thr112 and Asn265 are located close to Na1 (Asn115) and Na2 (Gln264) which could indicate a cooperative binding between sodium and the substrate.

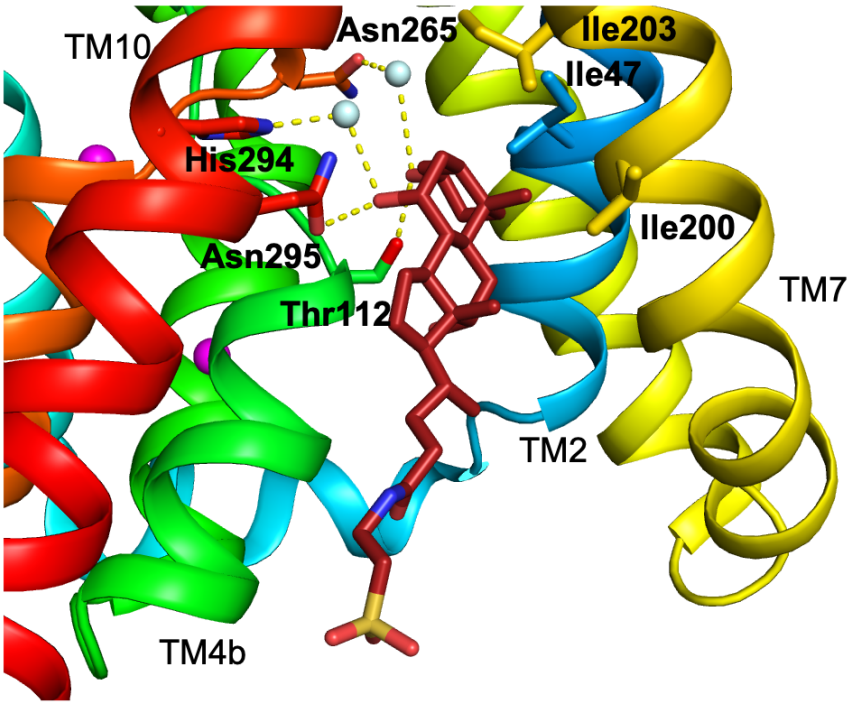


Figure 1.11: Taurocholate binding site in $ASBT_{NM}$: The bound taurocholate molecule is shown as sticks (firebrick red). Residues interacting with the taurocholate molecule are shown as sticks and labelled with their name and position. Hydrophobic residues lining the substrate binding site are shown as well. Residues on TM7 are shown in yellow and residues on TM2 are shown in light blue.

1.3.3 *Neisseria meningitides* and bile acids

Even though the source of the protein did not play a role for the structural characterization of $ASBT_{NM}$ and it was only chosen because of its beneficial properties such as expression level and stability, the host bacterium will be described here

briefly as it will be important for later purposes (Hu et al. 2011).
Neisseria meningitides is a gram-negative bacteria which was first identified by Anton Weichselbaum in the cerebrospinal fluid (CSF) as the cause of meningitis (Weichselbaum 1887, Stephens et al. 2007). The bacterium colonizes the human nasopharynx under normal condition and does not harm the host with approximately 10% of the population carrying it (Broom, Cartwright). Only occasionally the bacterium enters the blood stream and causes meningitis (Yazdankhah et al. 2004).

While no information could be found about the connection of *Neisseria meningitides* and bile acids, it is known for other bacteria that bile acids are toxic (Velkinburgh et al. 1999) mainly because of their detergent character. They were shown to cause disruption of the cell membrane (Fujisawa et al. 1996) and can induce DNA damage (Prieto et al. 2004) at high concentrations. Therefore, it is not surprising that bacteria developed resistance strategies to overcome the toxic effects of bile acids. A lot of bacteria possess efflux systems which actively transport bile acids out of the cell (Thanassi et al. 1997). Two efflux pump systems are described for *Neisseria meningitides* which provide resistance to antibacterial hydrophobic and cationic agents (C. E. Rouquette-Loughlin et al. 2004, C. Rouquette-Loughlin et al. 2003). While both studies did not mention bile acids as potential substrates, it is likely that the systems possess the ability to transport bile acids outside the cell as these multidrug efflux pumps are known to exhibit a huge substrate variety like for example AcrB from *E. coli* (Elkins et al. 2002).

1.4 ASBT from *Yersinia fredericksen*

In 2013, two new structures of another ASBT homologue were published by Zhou et. al (X. Zhou et al. 2014). $ASBT_{YF}$ is from the bacterium *Yersinia fredericksen* and shares around 40% identity with $ASBT_{NM}$. The wild type protein was crystallized in an inward-facing conformation similar to $ASBT_{NM}$. They also crystallized a Glu254Ala mutant (Glu260 in $ASBT_{NM}$) in an outward-facing conformation. Both crystal structures were obtained using the lipidic cubic phase

crystallization method (LCP).

1.4.1 Inward-facing conformation

The inward-facing conformation is quite similar with the $ASBT_{NM}$ structure showing only minor differences (see figure 1.12). The main differences are in TM1 and TM4b. TM1 does not show the same kink than in $ASBT_{NM}$ (see figure 1.12b) and therefore the intracellular half is closer to the core domain than in $ASBT_{NM}$. TM4b is partly unwound at the cross overpoint and shows one turn less than in $ASBT_{NM}$. Additionally, the helix is a bit more tilted towards the substrate binding pocket. Interestingly, even though the crystallization condition contained 100mM sodium chloride, no density for the sodium ions was visible. Overall, the sodium binding sites are well conserved between the two proteins with only two differences. Gln77 is a histidine in $ASBT_{YF}$ and Thr132 a serine. The unwinding occurs close to Na1 which is depicted in figure 1.12c. The unwound of TM4b causes a movement of Ser 108 and Asn109 (Ser114 and Asn115 in $ASBT_{NM}$). Overall, the structure shows an apo inward-facing conformation with no sodium.

1.4.2 Outward-facing conformation

The outward-facing conformation was obtained with an Glu254Ala mutant which was shown to interact with both sodium ions as previously discussed. Figure 1.13 shows the outward-facing conformation on the same side as all other structures with TM1 of the panel domain and TM10 formimg the core at the front.

As seen in the figure 1.13, the panel and core domain have moved towards each other leading to a closure of the intracellular binding pocket. A rotation of the molecule of 180° reveals a new binding pocket on the extracellular site between the core and panel domain. When comparing the extracellular sites of the $ASBT_{YF}$ outward-facing structure and the inward-facing structure, the extracellular binding site is caused by an outward movement of TM6 and a movement of TM9b and 10 towards the core domain. A side view of the core domain shows that the core domain, mainly TM4b and 9b, moves upwards with the extracellular part flipping backwards (see figure 1.14). The structure did not show any density

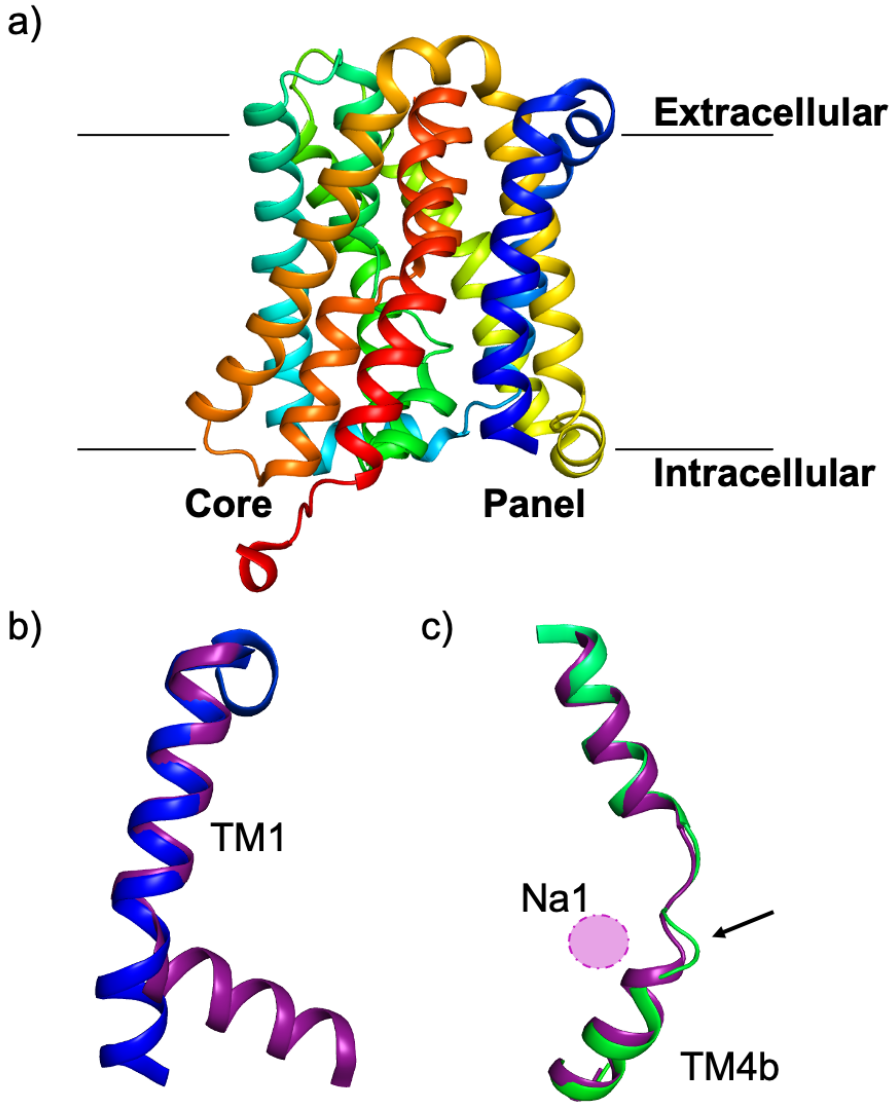


Figure 1.12: Cartoon representation of $ASBT_{YF}$ in the inward-facing conformation: a) Side view of the $ASBT_{YF}$ structure (4n7w) in cartoon representation. Helices are colored differently starting with TM1 in dark blue to TM10 in red. The membrane is shown as lines with the extracellular and intracellular site labeled. The core and panel domains are indicated. b) Overlay of TM1 from $ASBT_{YF}$ (dark blue) and $ASBT_{NM}$ (violet) c) Overlay of TM4 of $ASBT_{YF}$ (green) and $ASBT_{NM}$ (violet). The sodium 1 site is indicated as a sphere but was empty in $ASBT_{YF}$.

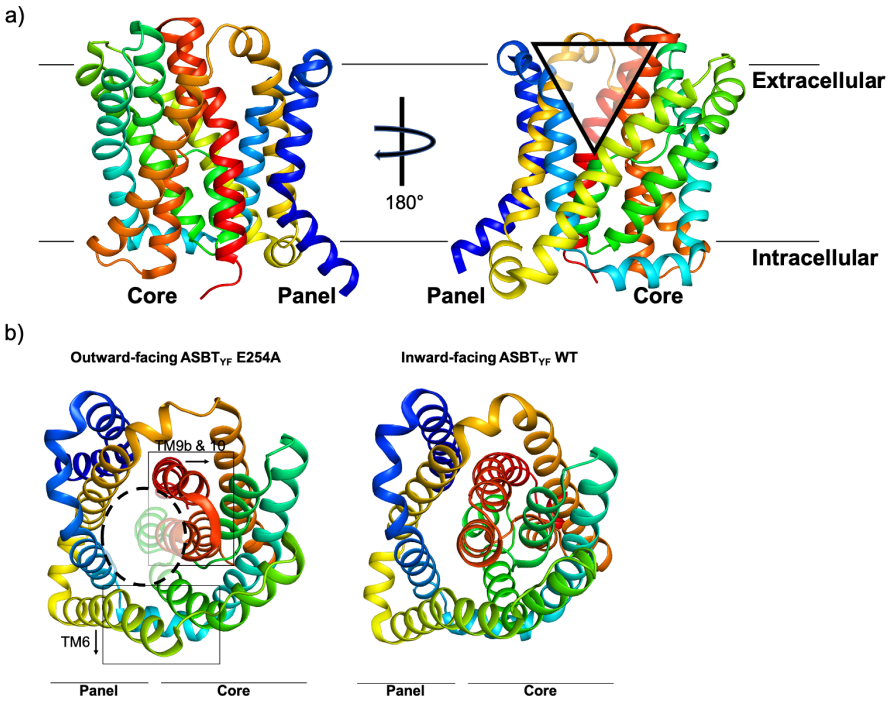


Figure 1.13: Cartoon representation of the E254A structure of $ASBT_{YF}$ in the outward facing conformation (4n7x) and comparison of the extracellular sites of the outward and inward facing conformations of $ASBT_{YF}$: a) Side view of the $ASBT_{YF}$ structure (4n7x) in cartoon representation. Helices are colored differently starting with TM1 in dark blue to TM10 in red. The membrane is shown as lines with the extracellular and intracellular site labeled. The core and panel domains are indicated. The right site structure shows the protein after a 180° rotation b) Comparison of the extracellular sites of the inward-facing and outward-facing conformation of $ASBT_{YF}$. The substrate binding site in the outward-facing conformation is indicated as a dotted circle (left side). Helices which showed movements between the outward and inward-facing conformation are highlighted and the movement direction for opening of the extracellular cavity is shown with arrows.

for the sodium ions nor for the taurocholate which was soaked with the crystal before freezing. Therefore, the structure can be assigned as an outward-facing apo no sodium conformation.

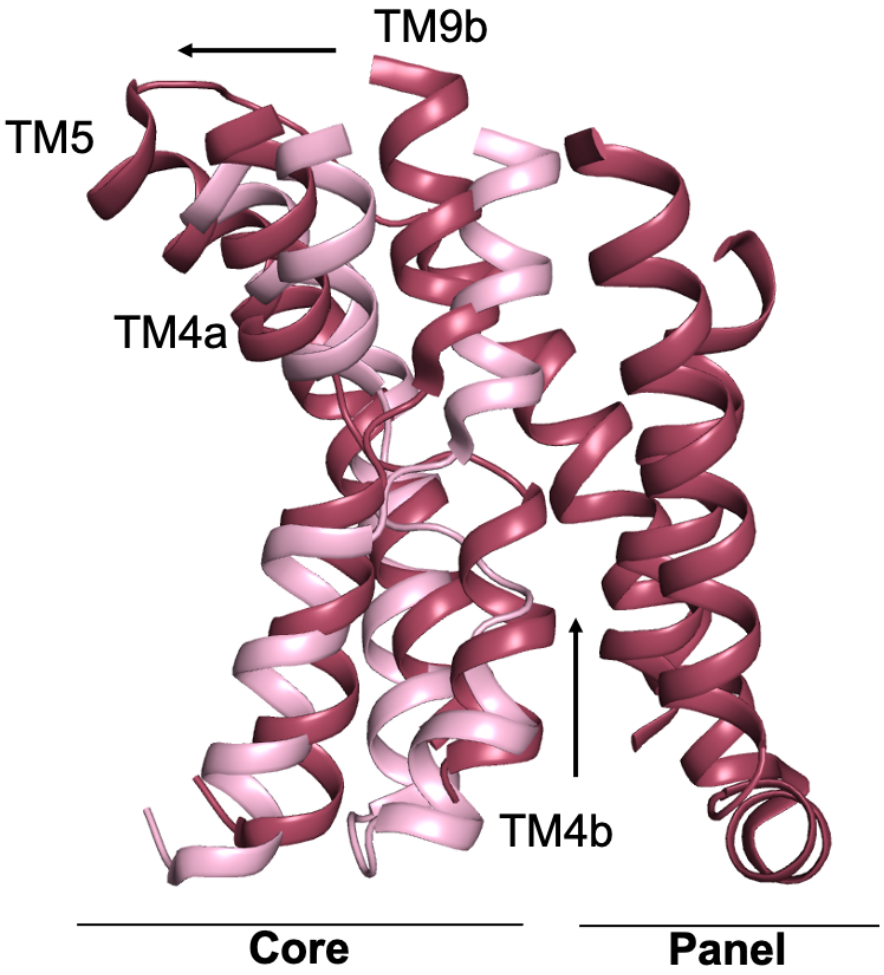


Figure 1.14: Core domain movement between the inward- and outward-facing conformation of $ASBT_{YF}$: The upwards movement of the core domain in the outward-facing conformation (red) is shown in comparison to the position of the core domain in the inward-facing conformation (pink). Helices are labelled and the movement direction for the transition between inward and outward conformation are shown as arrows.

1.5 The aim of this project

The aim of this project was to further structurally and functionally characterize $ASBT_{NM}$. From a structural point, there are in total three different conformations available of $ASBT_{NM}$ and $ASBT_{YF}$. Figure 1.15 shows a proposed transport cycle of ASBT with conformations which are likely to occur during the transport cycle. Regarding the inward-facing state, the previously discussed substrate-bound state ($ASBT_{NM}$) and the apo no sodium state ($ASBT_{YF}$) are known. During the transport cycle, the substrate is probably released first and leaves the protein in an apo sodium state. Even though it is unlikely that the substrate-release triggers a major conformational change, small differences could occur, maybe close to the sodium binding sites, indicating some connection between the substrate and sodium binding. For example, in LeuT the substrate directly interacts with one sodium ion and therefore, release of the substrate reduces the affinity of one sodium site and triggers the sodium release (Quick, Winther, et al. 2009). Although taurocholate does not directly participate in sodium binding in $ASBT_{NM}$, taurocholate release could indirectly affect the sodium sites.

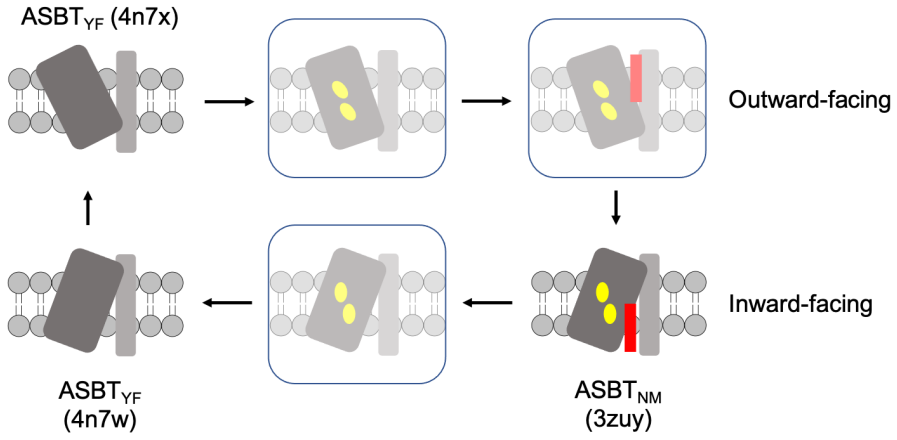


Figure 1.15: Schematic transport cycle of $ASBT_{NM}$: Different conformations of $ASBT_{NM}$ are shown which occur during one complete transport cycle. The transporter is shown in grey with the core domain in dark grey and the panel domain in light grey. The two sodium ions are depicted as yellow spheres and the taurocholate is shown as a red bar. Conformations for which structural information are known are labelled with the name and the PDB code in brackets. Conformations which are unknown are transparent.

Next, the sodium ions are released leading to an apo no sodium conformation ($ASBT_{YF}$). It is not known in which order the sodium ions leave the binding sites and again sodium binding could trigger conformational changes like the unwinding of TM4b which have been observed in $ASBT_{YF}$. Nevertheless, the sodium sites are slightly different among $ASBT_{NM}$ and $ASBT_{YF}$ and therefore conformational changes do not need to be similar in both proteins.

For the outward-facing states, only the structure of $ASBT_{YF}$ Glu254Ala is known. This structure is an outward-facing apo no sodium state and more structural information are needed to understand substrate and sodium binding in the outward-facing state. As mentioned before, the alternating-access mechanism proposes that sodium binding occurs first (outward-facing apo with sodium) and would prime the binding site for taurocholate. The structure of an outward-facing conformation with taurocholate bound would help to understand the binding of it and to identify important residues for substrate binding. This information could shed light on the differences in substrate recognition of hASBT and hNTCP and is probably the most interesting structure.

To sum up, there is still a lot of structural information missing for $ASBT_{NM}$ which can be really helpful to understand the transport mechanism and substrate recognition by it. Knowledge of the substrate binding site in the outward facing conformation could help to design more potent inhibitors for hASBT. Additionally, with more structural information the transport mechanism can be investigated in more detail which would allow to understand the size and volume of the substrate binding site in each conformation and during the transport in which the binding site might be blocked from the extracellular and intracellular site. This information could be beneficial for the design of more potent bile acid drug conjugates regarding the minimal requirements for transport and would allow to say where hASBT allows conjugation of the drug regarding the position in the bile acid backbone and what size allowance the protein has. Therefore, the wild type protein and mutants were prepared for crystallization using the lipidic cubic phase method which has been successfully used for $ASBT_{YF}$ and $ASBT_{NM}$ by a former PhD student (X. Zhou et al. 2014, Sekiguchi 2014).

The second part of the project was focused on the functional characterization of

$ASBT_{NM}$ wild type and mutants. 25 different mutants have been designed and prepared. The mutations are located in the sodium binding sites, in the substrate binding site, or involving side chains which were shown to be functionally important in hASBT. Attempts were made to characterize all mutants according to their transport behavior by using a liposome-based transport assay. Changes in the transport behavior of mutants can help to identify residues which are important for function and can be combined with the structural data to identify the role of mechanistically important residues and what their function is during a transport cycle. Additionally, the effect of these mutations on substrate binding was investigated by using Microscale thermophoresis (MST). By being able to measure the effect of mutations on substrate affinity, it could help to identify residues which are important for substrate binding.

Last, attempts were made to identify a new potential substrate for $ASBT_{NM}$. It was mentioned in the introduction that it is unlikely that ASBT has bile acids as its substrates because they are toxic to the bacteria. In regards to previous work, taurocholate has been a difficult substrate mainly in functional assays because of its detergent character. Conventional methods for affinity measurements like ITC or CD were not successful with the taurocholate as it potentially interferes with the detergent micelle of the protein (Sekiguchi 2014). Therefore, it would be useful to have a different substrate which might be more applicable for experiments and would allow to further characterize the protein in regards to functional important residues.

To sum up, three main aims were tried to address: 1. Crystallization trials of $ASBT_{NM}$ to obtain new structural information. 2. Functional characterization of important residues in substrate-binding and transport 3. Identification of a new substrate for $ASBT_{NM}$.

Chapter 2: Materials and methods

2.1 Materials

All chemicals and consumables were purchased from one of following companies unless otherwise stated.

Table 2.1: List of suppliers

Name	Location
Fisher Scientific^{TM} UK	Loughborough, UK
Sigma Aldrich	Darmstadt, Germany
Generon	Slough, UK
Avanti Polar Lipids, Inc.	Alabaster, USA
Stratech	Ely, UK
New England Biolabs (UK) Ltd.	Hitchin, UK
VWR International Ltd	Lutterworth, UK
STARLAB (UK), Ltd	Milton Keynes, UK
PerkinElmer	Waltham, USA

Detergents

All detergents were purchased either from Anatrace or Generon. Detergents were used at a stock concentration of 10% in ddH_2O. Stock solutions were stored at −20 °C and thawed at room temperature (RT) before usage.

Table 2.2: Detergents

Name	Critical micelle concentration (CMC)	Used concentration
n-Dodecyl β-D-maltoside (DDM)	0.0076%	0.03%
Decyl β-D-maltopyranoside (DM)	0.087%	0.25%
n-Nonyl-β-D-Maltopyranoside (NM)	0.28%	0.84%
n-Octyl-β-D-Glucopyranoside (OG)	0.53%	1.75%
Lauryldimethylamine-N-Oxide (LDAO)	0.023%	0.06 %

Antibiotics

Table 2.3: Antibiotics

Name	Stock concentration (mg/ml) (solvent)	Used concentration (μg/ml)
Kanamycin	50 (ddH_2O)	50
Chloramphenicol	35 (Ethanol)	35

Bacterial cultivation media

Table 2.4: Bacterial cultivation media

Name	Composition
Lysogeny broth (LB) Miller	1% (w/v) tryptone
	0.5% (w/v) yeast extract
	0.17 M NaCl
LB agar plates	1% (w/v) tryptone
	0.5% (w/v) yeast extract
	0.17 M NaCl
	1.5% agar
Super Optimal broth with	2% (w/v) tryptone
Catabolite repression (SOC) media	0.5% (w/v) yeast extract
	10 mM NaCl
	2 mM KCl
	10 mM $MgCl_2$
	10 mM $MgSO_4$
	20 mM glucose

General solutions and buffers

Table 2.5: General solutions and buffers

Name	Composition
10x Phosphate buffered saline (PBS)	1.37 M NaCl 27 mM KCl 100 mM Na_2PO_4 20 mM KH_2PO_4
50x TAE buffer (Tris-acetate-EDTA)	0.04 M Tris base 0.02 M acetic acid 0.002 M EDTA
10x Sodium Dodecyl Sulfate (SDS) running buffer	0.2501 M Tris base 1.924 M glycine 0.0347 M SDS

Primers

Table 2.6: Sequencing Primers

Name	Sequence (5' to 3'-end)
T7 forward	TAA TAC GAC TCA CTA TAG GG
Waldo GFP reverse	TTC ACC CTC TCC ACT GAC AG

Bacterial strains

Table 2.7: Bacterial strains

Name	Resistance	Used for
Lemo21(DE3) Competent *E. coli*	Chloramphenicol	Protein expression
NEB 5-alpha Competent *E. coli*	none	DNA amplification

2.2 Methods

2.2.1 Preparation of chemically competent cells

Desired cells were freshly streaked out on an agar plate supplemented with the appropriate antibiotics and left growing overnight at 37 °C. On the next morning, 100 ml of LB media supplemented with the appropriate antibiotics were inoculated with a single colony and cells were grown at 37 °C with shaking at 180 rpm until an optical density at 600 nm (OD_{600}) of 0.5. Cell solution was split into two 50 ml falcon tubes and cooled down on ice for 15 min. All buffers and equipment were prechilled on ice. Cells were spun down at 3000 g and 4 °C for 10 min and the supernatant was discarded. The pellets were resuspended in 10 ml of 0.1 M $CaCl_2$ each and left for another 10 min on ice. Cells were spun down again at 3000 g and 4 °C for 10 min and pellets were resuspended in 2 ml of 0.1 M $CaCl_2$, 10% glycerol and flash frozen in 50 μl aliquots.

2.2.2 Transformation of chemically competent cells

Desired cells were thawed on ice for 10 min. 1 μl of plasmid DNA (100 ng/μl to 200 ng/μl) was added to the cells and mixed by gently flipping the tube. Cells were incubated on ice for 30 min and heat shocked at 42 °C for the appropriate time. After an incubation for 5 min on ice, 950 μl of SOC media was added to the cells and cells were incubated at 37 °C for 60 min at 900 rpm. 150 μl of the cell suspension was plated on a LB agar plate supplemented with the appropriate antibiotics and incubated at 37 °C overnight. For cloning, the rest of the cell suspension was spun down for 5 min at 5000 rpm at RT. The supernatant was discarded and the pellet was resuspended in 150 μl SOC media and streaked out.

2.2.3 Amplification and Isolation of plasmid DNA

5 ml of LB media were supplemented with the appropriate antibiotics and inoculated with a single colony of NEB5α containing the desired plasmid. Cells were grown at 37 °C with shaking at 180 rpm overnight. Cells were pelleted on the next morning by centrifugation at 4000 rpm at RT. The supernatant was

discarded and DNA extraction was performed according to manufacturer's instruction (Monarch Plasmid Miniprep Kit, NEB). DNA was eluted in 30 μl of elution buffer and stored at $-20\,^{\circ}$C.

2.2.4 Sequencing

Sequencing was performed using Eurofins genomics. Samples were prepared according to their guidelines. Briefly, 5 μl of DNA template (80 ng/μl to 100 ng/μl) and 5 μl of the desired sequencing primer (final concentration: 5 μM) were transferred into a 1.5 ml Eppendorf tube and sent off for sequencing. Sequencing results were analyzed using the Expasy translate server and Blast server (Gasteiger, Gattiker, et al. 2003, Altschul et al. 1990).

2.2.5 Cloning into the 3C-cleavable vector

All enzymes and kits used for cloning were purchased from New England Biolabs (UK) Ltd.

Restriction enzyme digestion

The desired amount of DNA was transferred into a Polymerase chain reaction (PCR) tube together with 5 μl CutSmart buffer, 1 μl BamH1-HF, 1 μl Xho1 and the reaction was topped up to 50 μl with ddH_2O. The restriction digest was performed at 37 °C for 4 h to ensure full cleavage. For vector cleavage, an additional step was performed at the end of the incubation time (see next section).

Dephosphorylation of DNA

To prevent religation of the cleaved vector backbone, 1 μl of Alkaline Phosphatase was added only to the vector mixture 30min before the end of the incubation.

Agarose gel electrophoresis

Agarose gels were prepared by heating 1.5 g agarose in 100 ml of 1x TAE buffer until the agarose was completely dissolved. The solution was cooled down to around 50 °C and GelRed staining was added to a final concentration of 1x.

The gel was casted using the BioRad Mini-Sub Cell GT Systems. Agarose gel samples were mixed with 6x loading dye to reach a final concentration of 1x. For gel purifications, the whole sample was loaded by combining lanes. 1 μl of DNA ladder was mixed with 4 μl ddH_20 and 1 μl of 6x loading dye. All samples were loaded onto the gel and gel was ran at 120 V for 30 min to 1 h until the desired separation was observed.

Gelpurification

Gel extraction was performed according to manufacturer's instruction with the Monarch® DNA Gel Extraction Kit.

Ligation

For ligation, the desired backbone was mixed with the desired insert in different ratios. Usually, a 1 to 3 ratio (vector to insert) was chosen to ensure the incorporation of the insert. 2 μl of T4 DNA Ligase Buffer was added to the sample and the mixture was topped up to 19 μl with ddH_20. 1 μl of T4 DNA ligase was added and ligation was performed at 16 °C overnight. The ligation mix was transformed into NEB5alpha cells on the next day. Usually, 5 μl and 15 μl were used for transformations.

2.2.6 Expression

PASM -5052 media preparation

PASM-5052 media was freshly prepared in the morning of expression under sterile conditions in a laminar flow cabinet. All components were beforehand sterilized by autoclaving or sterile filtering. For small scale expression, the desired volume of PASM-5052 media was prepared in a sterile bottle. For large scale expression, the media was directly prepared in the autoclaved growth flasks (Tunair TM, Sigma Aldrich). The components were added in the order stated. After each addition, the solution was mixed by gently shaking the vessel to ensure a homogenous dispersion and to avoid precipitation upon addition of the next component.

Table 2.8: PASM media pipetting scheme

Name	For 1l
ddH_20	900 ml
1 M $MgSO_4$	1 ml
1000x trace metals	1 ml
50x 5052	20 ml
20x NPS	50 ml
100 μM vitamin B12	1 ml
17 amino acid solution	20 ml
25 mg/ml methionine	8 ml

Expressiontrials

Expression was performed using the MemStar system (Lee, Kang, Hjelm, et al. 2014). The desired volume of LB supplemented with the desired antibiotics was inoculated with a freshly transformed colony and the starter culture was grown overnight at 37 °C with shaking at 180 rpm. On the next morning, PASM-5052 media was freshly prepared according to the recipe. The media was supplemented with the desired antibiotics and growth was initiated by the 1:50 dilution of the overnight culture into PASM-5052 media. Growth was performed at 37 °C with shaking at 200 rpm. Once the culture reached an OD_{600} of 0.5, the culture was divided into 5 ml fractions and different amounts of isopropyl-β-D-thiogalactopyranosid (IPTG) (0 , 0.4 mM) and L-rhamnose (0 , 0.1, 0.25, 0.5, 0.75, 1 mM) were added to the subcultures. Subcultures were incubated at 25 °C at 200 rpm overnight. On the next morning, OD_{600} was measured for all subcultures by diluting cells 1:20 in 1x PBS. For Green Fluorescent Protein (GFP) readings, 1 ml of cultures were spun down at maximum speed for 10 min. Supernatant was carefully removed and the pellet was resuspended in 100 μl of 1x PBS. Cell suspension was transferred into a black-walled, 96-well plate and GFP fluorescence was measured.

Large scale expression

Expression was performed using the MemStar system (Lee, Kang, Hjelm, et al. 2014). The desired volume of LB supplemented with the desired antibiotics was inoculated with a freshly transformed colony and the starter culture was grown over night at 37 °C with shaking at 180 rpm. On the next morning, PASM-5052 media was freshly prepared according to the recipe. The media was supplemented with the desired antibiotics and growth was initiated by the 1:50 dilution of the overnight culture into PASM-5052 media. Growth was performed at 37 °C with shaking at 200 rpm. Expression was induced at an OD_{600} of 0.5 by the addition of 0.4 mM IPTG and 0.25 mM L-rhamnose. The temperature was decreased to 25 °C and expression was performed overnight. On the next day, cultures were harvested by centrifugation at 5000 g for 15 min at 4 °C. The supernatant was discarded and the pellet was transferred into a 50 ml falcon tube using a cell scraper. Centrifuge bottles were washed with a small amount of 1x PBS to remove residual cells and added to the falcon tube. Cells were flash frozen in liquid nitrogen and stored at −80 °C.

2.2.7 Protein purification

Membrane preparation

Cell pellet was thawed by incubating the falcon tube in luke warm water. All following steps were performed at 4 °C or on ice. Cells were transferred into lysis buffer (1x PBS, 1 mM $MgCl_2$) supplemented with 0.5 mM Pefabloc and DNaseI and stirred for 30 min to ensure a homogenous suspension. The cell suspension was filtered through a filter to remove lumps and larger particle which could clog the cell disruptor. Cells were lysed by passing them through a cell disruptor at a pressure of 25 kpsi twice. After the second round, residual lysate was collected by passing a small amount of lysis buffer through the system. Unbroken cells and cell debris were pelleted by a first, low centrifugation step at 15 000 g for 13 min. The supernatant was applied to a second round of centrifugation at 45 krpm for 1 h to pellet the membranes. The supernatant was discarded and 5 ml of 1x PBS was added to each tube. Membrane pellets were transferred to

a glass homogenizer by gently scratching them off the tube wall with a round-headed spatula. Membranes were homogenized by manually pulling the pestle up and down. Homogeneity was assessed visually by checking that no lumps or membrane fragments were left. The solution was transferred into a 50 ml falcon tube. Centrifugation tubes and the spatula were washed with a small amount of 1x PBS to remove residual membranes and the solution was homogenized in a second round. Both solutions were pooled together, topped up to 60 ml with 1x PBS and divided into two 50 ml falcon tubes. Membranes were flash frozen in liquid nitrogen and stored at −80 °C.

Large scale purification

Membranes were thawed by incubating the falcon tubes in luke warm water. From here onwards, all steps were performed at 4 °C or on ice. Membranes were transferred into solubilization buffer (1x PBS, 150 mM NaCl, 10 mM imidazole) supplemented with 0.5 mM Pefabloc. The solution was left stirring until the membranes were fully dissolved. Membrane proteins were solubilized by the addition of 1 % (w/v) DDM and gently stirring for 2 h. Unsolubilized material was removed by centrifugation at 45000 rpm for 45 min. Meanwhile, the needed amount of beads (Ni-NTA Agarose, Qiagen) was washed with 10 column volumes (CV) of dd$H_2$0, followed by 10 CV of solubilization buffer. The supernatant was transferred into a bottle and the desired amount of beads was added to the solution. 1 ml of resin was used for 1 mg of protein. The slurry was gently stirred for 3 h to allow binding and loaded onto a glass Econo-Column® (Bio-Rad Laboratories Ltd., Hemel Hempstead, UK) and the flow through (FT) was collected. The column was washed twice with 10 CV of wash buffer containing 20 mM (1st round) and 30 mM (2nd round) imidazole (1x PBS, 150 mM NaCl). 3C protease was added and mixed with the resin. Cleavage was performed overnight. On the next day, the flow through was collected and the column was washed with size exclusion chromatography (SEC) buffer (20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.03% DDM). A 5-ml HisTrap HP column (GE Healthcare, Chalfont, UK) was washed with dd$H_2$0 and equilibrated with SEC buffer. The cleaved protein was passed through the HisTrap column using a peristaltic pump to remove residual

3C protease, uncleaved protein, and unspecific bound proteins. The flow thorough was collected and concentrated using a 100 kDa molecular weight (MW) cutoff concentrator. The protein purity was assed by SDS polyacrylamide gel electrophoresis (PAGE).

Size exclusion chromatography

Size exclusion was performed using an ÄKTA purification system and a Superdex 200 Increase 10/300 GL column. The column was washed with ddH_20 and equilibrated in SEC buffer (20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.03% DDM). Protein was concentrated to 6 - 10 mg/ml and loaded into a 200 μl superloop. Run was performed according to a prepared program and fractions were collected in 200 μl fractions in a 96 deep well plate. Fractions containing protein were pooled together as assessed by the absorption at 280 nm.

2.2.8 SDS gel electrophoresis

Table 2.9: Recipe for 14% resolving gels

Solution	1 gel	2 gels	4 gels
1.5 M Tris pH 8.8	1.41 ml	2.82 ml	5.64 ml
ddH_20	1.51 ml	3.02 ml	6.04 ml
30% polyacrylamide (29:1)	2.66 ml	5.32 ml	10.64 ml
10% SDS	60 μl	120 μl	240 μl
10% AMPS	60 μl	120 μl	240 μl
TEMED	5 μl	10 μl	20 μl

Table 2.10: Recipe for stacking gels

Solution	1 gel	2 gels	4 gels
0.5 M Tris pH 6.8	630 μl	1260 μl	2520 μl
ddH_20	1.385 ml	2.77 ml	5.54 ml
30% polyacrylamide (29:1)	335 μl	670 μl	1340 μl
10% SDS	25 μl	50 μl	100 μl
10% AMPS	25 μl	50 μl	100 μl
TEMED	2.5 μl	5 μl	10 μl

SDS PAGE gels were casted using the Mini-PROTEAN system from BioRad. The desired volume of resolving gel solution was prepared and poured into the assembled glass plates. The resolving gel solution was covered with 100% ethanol and left polymerizing for 30 min. The ethanol was discarded and residual ethanol was removed with a filter paper. The stacking gel solution was prepared and poured on top of the resolving gel. The desired comb was placed into the gel trying to avoid any air bubbles. The stacking gel was allowed to polymerize for another 30 min. Gels were either used directly or were stored at 4 °C for one or two days wrapped in wet paper towel to avoid drying-out. Samples were diluted with 2x Laemmli sample buffer to obtain a final concentration of 1x. Gel samples were incubated at 42 °C for 5 min before loading. Gels were run using the Mini-Protean Tetra cell system. The system was assembled according to manufacturer's instructions and filled with 1x SDS running buffer. The comb was removed and gel samples were loaded (10 μl to 20 μl) including 2 μl of the markers. Gels were run at 200 V for 55 min. For in-gel fluorescence, the gel was carefully transferred into ddH_20 and imaged directly. For staining, Instant blue was added to the gel and heated in a microwave at full power for 20 s. The gel was left staining for a couple of hours with gentle agitation. The gel was destained by transferring it into ddH_20 and it was left destaining overnight with gentle agitation.

2.2.9 Protein concentration determination

Protein concentration was determined using a NanoDrop spectrophotometer by measuring the absorbance at 280 nm. Protein buffer only was used as a blank and 1 μl of protein solution was loaded onto the NanoDrop to measure the absorbance. The molar extinction coefficient and molecular mass were used to determine the protein concentration. Both values were obtained by using the ExpasyProtParam online tool (Gasteiger, Hoogland, et al. 2005).

Table 2.11: Molecular weight and extinction coefficient of $ASBT_{NM}$

Molecular weight	32936.78 g/mol
Molecular extinction coefficient	45950 $M^{-1}cm^{-1}l$

2.2.10 Lipidic cubic phase crystallization

The protein was concentrated to around 25 mg/ml and was filled into a 50 μl Hamilton syringe. The monoolein volume was calculated and was added to a 100 μl Hamilton syringe. The ratio of protein solution to monoolein is 2 to 3 (volume per volume). The two syringes were attached to each other using an adaptor and the two solutions were mixed until the solution cleared up indicating that the lipidic cubic phase (LCP) has formed. The LCP was dispensed in 50 nl drops on glass plates and covered in 800 nl crystallization solution using a Mosquito crystallization robot. The glass plates were covered with another glass plate and they were put into a 20 °C incubator. LCP plates were checked frequently under a microscope for crystal formation.

For initial screens, the LCP drop was covered with buffer from the following screens: MemGold, MemGold2, MemMeso and MemGoldMeso. All crystallization screens were purchased from Molecular Dimensions (Sheffield, United Kingdom).

Optimization screens were manually set up in 96-well plates by varying one or more concentrations of the crystallization condition which showed some crystal formation in the initial screens.

2.2.11 Crystal harvesting, Data collection, Structure determination and Structure analysis

The glass plates were opened using a glass cutter and the crystals were mounted using MicroMounts MiTeGen loops. The loops were flash frozen in liquid nitrogen and stored until usage. Data collection was performed on beamline I24 at the Diamond Light Source. The data collection strategy differed slightly for each data set and is shown in the statistics table for each data set. Diffraction images were integrated using DIALS and scaled in Aimless (CCP4i2) (Project et al. 1994). The initial structure was solved by Molecular replacement using the vapor diffusion structure as a search model (3zuy). The first 15 residues were deleted for that purpose. For later data sets, the wild type apo sodium structure was used as a model. Refinement was performed in Phenix (Adams et al. 2010) after each round of refinement the electron density and the model were investigated in Coot (Emsley et al. 2004). Structure building was also performed in Coot.
Structure analysis was performed in Coot and Pymol (Emsley et al. 2004, Schrödinger, LLC 2015). Each structure was visually investigated and checked for distinctive features. Superimpositions were performed in Pymol or Chimera (Pettersen et al. 2004). Images were further edited and labelled in Microsoft PowerPoint. Electron density maps were prepared by Phenix and uploaded into Pymol.

2.2.12 Soybean lipid-based transport assay

Liposome preparation

L-α-Phosphatidylcholine from soybean (P5638) was purchased from Sigma Aldrich. L-α-Phosphatidylcholine was dissolved at 20 mg/ml in MMK buffer (10 mM MOPS (pH 7), 5 mM $MgCl_2$, 100 mM KCl) by vortexing the solution for 30 min at RT at maximum speed. The solution containing multilamellar vesicles was applied to 6 cycles of sonication at 40% output for 30 s with 0.5 s on and 0.5 s off on ice. The solution was left on ice for 1 min between each cycle. Undissolved lipids and debris were removed by centrifugation at 15000 g at RT for 3 min in a JA25.50 rotor. The supernatant containing small unilamellar vesicles was transferred into clean 15 ml falcon tubes and the liposomes were applied to

8 rounds of freeze-thaw cycles using liquid nitrogen and lukewarm water. After the 8th freeze cycle, the liposome solution could be stored at −80 °C or could be used immediately for reconstitution.

Extrusion and reconstitution

The liposome extruder was assembled according to manufacturer's instruction. Briefly, the four filter supports and the two membranes (diameter 0.4 μm and 0.2 μm) were soaked first in dd$H_2$0 and second in MMK buffer. Two supports were inserted and the 0.4 μm membrane was added on the top. The other two supports were added to the other internal membrane support and the extruder was assembled. 1 ml of liposome solution was transferred into one of the syringes and the syringes were stuck into both sides of the extruder. The solution was passed 20 times through the extruder and the solution was transferred into a new tube and the step was repeated until the whole liposome solution was extruded. The extruder was disassembled and the 0.2 μm membrane was inserted between the filter supports. The whole extrusion process was repeated with the 0.2 μm membrane and the liposome solution was collected in a new tube. 1 ml of the extruded liposome solution (0.2 μm diameter) was destabilized by the addition of 90 μl of 20% sodium cholate solution. The solution was gently mixed and should clear up upon mixing. 40 μl of 2 mg/ml protein solution was added and incubated at RT for 30 min at 900 rpm in a thermomixer. Meanwhile, the needed number of PD-columns were washed in dd$H_2$0 and equilibrated with MMK buffer. 1.1 ml of the liposome solution was added to the top of the column and it was waited until the solution was completely absorbed onto the column. The column was then washed with 0.3 ml and with 1.4 ml of MMK buffer and the flow-through was discarded after each washing step. The proteoliposomes were eluted from the column with 2.3 ml of MMK buffer and collected into a 15 ml falcon tube. Liposomes were stored overnight at 4 °C or were applied to dialysis overnight against MMK buffer using a 3 kDa dialysis membrane.

Transport assay

150 μl of liposomes were transferred into a fresh Eppendorf tube. 5.2 μl of 4 M NaCl and 1 μl of valinomycin were added to the liposomes to obtain a final sodium concentration of 132 mM NaCl and 666 nM valinomycin. Liposomes were incubated at 37 °C at 900 rpm for 5 min. Meanwhile, the radioactive label was prepared. 0.5 μl of the ^{3}H-taurocholate stock solution (final concentration 0.209 μM) was mixed with 0.5 μl of 600 μM taurocholate solution in MMK buffer (final 0.8 μM) and 1 μl of MMK buffer. After the 5 min incubation, the 2 μl of substrate solution was added to the proteoliposome samples and transport was performed for 5 min at 37 °C. The transport was stopped by the addition of ice cold MMK buffer supplemented with 1 mM taurocholate (stop buffer). The solution was applied to a G-25 column. The column was washed with another 150 μl of stop buffer. Proteoliposomes were eluted into scintillation vials by the addition of 700 μl stop buffer to the column and the internalized radioactive was measured using a liquid scintillation counter. Data was analyzed using GraphPad Prism (Kairam et al. 2012).

2.2.13 *E. coli* polar lipid-based transport assay

The protocol was adapted from Zhou et al and Quick et al (X. Zhou et al. 2014,Quick and Javitch 2007). *E. coli* polar lipid extract was purchased from Avanti®Polar Lipids (Alabama, US).

Liposome preparation

The evaporator flask was rinsed with chloroform to remove impurities. Chloroform was discarded and residual chloroform was evaporated by applying a nitrogen stream. Meanwhile, the cooling system was switched on and left running to reach −20 °C. The empty evaporator flask was weighed before the lipid solution (at 25 mg/ml in chloroform) was transferred into the evaporator flasks using a glass pipette and the chloroform was evaporated by using a rotary evaporator. The systems was left evacuated for another 10 min to ensure the complete removal of chloroform. The vacuum was released and the flask was weighed again.

The mass of lipids were calculated and lipid resuspension buffer was added to a final concentration of 10 mg/ml (w/v) lipids. Lipid were resuspended in 100 mM KPi (pH 7.5), 2 mM beta-mercaptoethanol, 1.5% Octyl glucoside by rotating them in the flask using the rotary evaporator without applying a vacuum for 30 min at RT. Dialysis tubing was prepared by thoroughly washing the membranes with ddH_20 and liposome dialysis buffer (100 mM KPi (pH 7.5), 2 mM beta-mercaptoethanol). The lipid solution was transferred into the dialysis tubing and the solution was dialyzed overnight at 4 °C against 5 l of dialysis buffer to slowly reduce the concentration of OG and to allow the formation of liposomes. On the next day, liposomes were aliquoted at 10 mg/ml and flash frozen in liquid nitrogen. Liposomes were stored at −80 °C.

Reconstitution into liposomes

The liposome extruder was assembled according to manufacturer's instructions. Liposomes were thawed at 23 °C and passed 13x times through a 0.4 μm membrane. Liposomes were diluted to a concentration of 5 mg/ml and destabilized by the addition of 0.12% (v/v) Triton. The protein was added to the destabilized liposomes at a lipid to protein ratio of 100:1 and incubated for 10 min at RT with gentle agitation. BioBeads were added in a 1:5 ratio of detergent to beads (wet mass) followed by a 1 h incubation at RT with agitation. The old biobeads were removed and new biobeads were added (same ratio). This was repeated one more time resulting in three 1 h incubations with the biobeads at RT. The fourth incubation was performed overnight at 4 °C. On the next day, proteoliposomes were harvested at 65000 rpm for 1 h at 4 °C and the pellet was resuspended in 100 mM KPi pH 7.5, 2 mM beta-mercaptoethanol to a final lipid concentration of 100 mg/ml. Proteoliposomes were flash frozen in 10 μl aliquots and stored at −80 °C.

Transport assay

Proteoliposomes were thawed at RT and diluted into an appropriate volume of buffer (100 mM KPi pH 7.5, 2 mM beta-mercaptoethanol). Proteoliposomes were extruded through a 0.4 μm membrane (final diameter of proteoliposomes)

and were harvested at 65000 rpm for 1 h at 18 °C. The pellet was resuspended to a final lipid concentration of 100 mg/ml and directly used in the transport assay. 10 μl of proteoliposomes were diluted with 190 μl of buffer containing 10 mM HEPES (7.4), 100 mM NaCl, 1 mM $CaCl_2$, 1 mM $MgCl_2$ supplemented with 0.209 μM of ^{3}H-taurocholate and 0.8 μM of cold taurocholate. The transport was performed at 37 °C for 5 min. Meanwhile, membranes were washed with 3.5 ml of stop buffer (100 mM KPi (pH 6), 100 mM LiCl, 1 mM taurocholate). The wells were filled with 3.5 ml of stop buffer and the proteoliposome solution was added to the buffer when the incubation time was over. The vacuum was applied and the solution was pushed through the filter. Membranes were washed three times with 3.5 ml of stop buffer. The remaining radioactivity of the membranes was measured using a liquid scintillation counter. Data was analyzed using GraphPad Prism (Kairam et al. 2012).

2.2.14 CPM assay

7-Diethylamino-3-(4'-Maleimidylphenyl)-4-Methylcoumarin (CPM) was purchased from ThermoFisher and was dissolved in DMSO to a final concentration of 4 mg/ml. The dissolved dye was stored at −80 °C in small aliquots. The CPM assay was performed in 0.2 ml non-skirted low profile 96-well PCR plates (ThermoFisher). 50 μl of protein (in 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.03% DDM) was added to each well. The protein amount can be varied but was kept at 2.5 μg for the final experiments. 1 μl of substrate can be added at this point and the protein was incubated with the substrate for 30 min at RT in the plate. The dye was diluted 1:100 in 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.03% DDM and 2.5 μl of the diluted dye was added to each well. The assay was performed using a Stratagene Mx3005P Real-Time PCR machine (Strategene®, California, US) and samples were heated from 25 °C to 95 °C in 1 °C/min steps. Data was analyzed using GraphPad Prism (Kairam et al. 2012). For initial experiments, different dye dilutions and dye volumes were tried which were obtained from different publications (Sonoda et al. 2011, T. Shimamura, Shiroishi, et al. 2011, Alexandrov et al. 2008, Qin et al. 2018, Majd et al. 2018). Protocol 1: 1 μl of a 1:100 dilution was used. Protocol 2: 1.25 μl of a 1:40 dilution was used. Protocol

3: 4.16 μl of a 1:40 dilution was used. Protocol 4: 2.5 μl of a 1:100 dilution was used. Protocol 5: 5 μl of a 1:40 dilution was used.

2.2.15 Circular dichroism

The protein solution was dialyzed overnight against buffer containing 20 mM Tris-H_2SO_4 (pH 7.5), 150 mM Na_2SO_4, and 0.03% DDM. On the next morning, the dialyzed sample was spun down for 15 min at maximum speed at 4 °C. The protein was diluted to a final concentration of 0.4 mg/ml (initial trials were performed at different concentrations) and the protein sample was transferred into a glass cuvette. Circular dichroism (CD) measurements were performed on a Jasco J-815 Circular Dichroism Spectropolarimeter (JASCO UK Ltd., UK). Measurements in the far UV-region (190 nm to 260 nm) were performed at 25 °C and 8 scans. For the thermo stability, a far UV-region spectrum was measured at each temperature (6 scans). The temperature was increased in 5 °C intervals from 20 °C to 95 °C. The ellipticity at 222 nm was used to follow the unfolding of the protein. The data was analyzed using the provided software and GraphPad Prism (Kairam et al. 2012).

2.2.16 Isothermal calorimetry

The protein sample was dialyzed over night at 4 °C against 20 mM Tris-HCl (pH 7.5), 150 mM NaCl and 0.03% DDM. A 5 mM pantoate solution was prepared by resuspension of pantoate in the dialysis buffer. The protein solution was spun down at high speed at 4 °C for 30 min. Isothermal calorimetry (ITC) experiments were performed on a MicroCal PEAQ-ITC (Malvern Panalytical, UK). The protein solution (220 μM) was filled into the sample cell and the pantoate solution (5 mM) into the syringe. The cell temperature was set to 10 °C with a stirring speed of 750 rpm and a reference power of 10 μcal/sec. 20 injections were performed with an initial delay of 250 s. The initial injection was performed for 0.8 s with an injection volume of 0.4 μl. The later injections were performed for 4 s with an injection volume of 2 μl. 180 s spacing was left between each injection. Data was analyzed using the provided software and GraphPad Prism (Kairam et al. 2012).

2.2.17 Microscale Thermophoresis

Microscale Thermophoresis (MST) experiments were performed on a Nanolith NT.115 (NanoTemper Technologies GmbH, Germany). The experiments were performed according to the software recommendations. The protein solution was prepared at 100 nM and was diluted with the substrate in a 1:1 ratio. The protein-substrate mixture was transferred into the glass capillaries (Premium Capillaries) and experiments were performed at 25 °C with the automated fluorescence detection and the medium MST setting. Data were analyzed using the provided software.

Chapter 3: Expression, purification and quality assessment of $ASBT_{NM}$ wild type and mutants

The first part of the project focused on the continuation of the work of a former PhD student and Postdoc. The major goal was to express and purify all necessary samples for the transport assay which consisted of the $ASBT_{NM}$ wild type protein, all mutants, and NapA as a negative control. The wild type protein was used in the beginning to become familiar with the expression system and purification procedure which were later applied for all the missing mutants and NapA.

3.1 The TEV-cleavable construct

3.1.1 Expression trials of $ASBT_{NM}$ TEV-cleavable wild type protein

First, expression trials were performed to confirm the reproducibility of expression and to validate the known expression conditions. Expression was performed using the MemStar system (Lee, Kang, Hjelm, et al. 2014). Briefly, the system is a combination of the *E. coli* strain Lemo21(DE3) and the auto-induction media PASM-5052. The strain contains an additional plasmid which allows the repression of expression by the addition of L-rhamnose while the media facilitates a high biomass production in comparison to conventional growth media like LB (Studier 2005).

Expression was performed as described in material and methods. At an OD_{600} of 0.5, different amounts of L-rhamnose were added to the cultures while the IPTG concentration, the temperature and the time of expression were kept constant. Expression levels were determined by measuring the GFP fluorescence of each culture. The GFP fluorescence was normalized to the OD_{600} of the culture and converted into the protein concentration in mg L^{-1} with reference to a GFP standard curve.

Figure 3.1 shows the expression level of $ASBT_{NM}$ wild type for different L-rhamnose concentrations. 0.1 mM and 0.25 mM L-rhamnose showed the best expressions with mean protein yields of 5.9 and 6.3 $mgL^{-1}OD^{-1}$ respectively. The addition of higher concentrations of L-rhamnose decreased the expression level by about 50% which is in accordance to the inhibition of expression by L-rhamnose.

The addition of no L-rhamnose yielded the lowest expression. Reasons for this could be the toxic side effects on the bacterial plasma membrane by excessive membrane protein production or the formation of inclusion bodies (Mathieu et al. 2019). Overall, the experiment was consistent with former expression regarding the L-rhamnose concentration and the protein yield. It was continued with 0.25 mM L-rhamnose which have been used before.

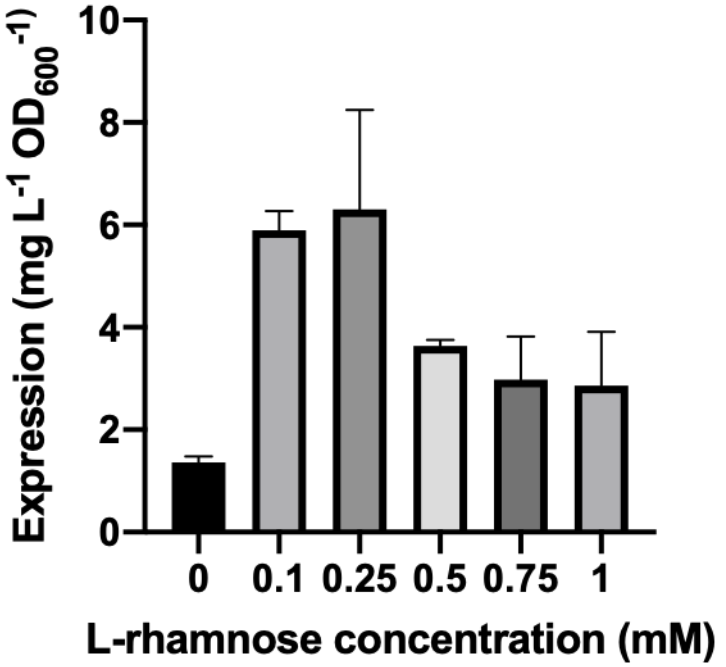


Figure 3.1: Expression trials of $ASBT_{NM}$ wild type TEV-cleavable construct: Effect of different L-rhamnose concentrations on the expression of the TEV-cleavable wild type construct. Expression levels were determined by measuring the GFP fluorescence of each culture. Expression levels were normalized to the OD_{600} of each individual culture and the fluorescence was converted into mgL^{-1}. Mean expression was calculated from three individual experiments.

3.1.2 Purification of $ASBT_{NM}$ TEV-cleavable wild type

After showing the reproducibility of expression and establishing the best expression condition, large scale growths were performed to obtained enough material for purifications. In a first step, the membrane fractions containing $ASBT_{NM}$ wild type were isolated from the bacteria. The membrane preparation was performed as described in material and methods. At different stages during the membrane

preparation, aliquots were taken and analyzed using SDS-PAGE. As the membrane preparation is a first crude purification to separate membrane proteins from soluble proteins, $ASBT_{NM}$ can only be visualized by in-gel fluorescence using the GFP tag.

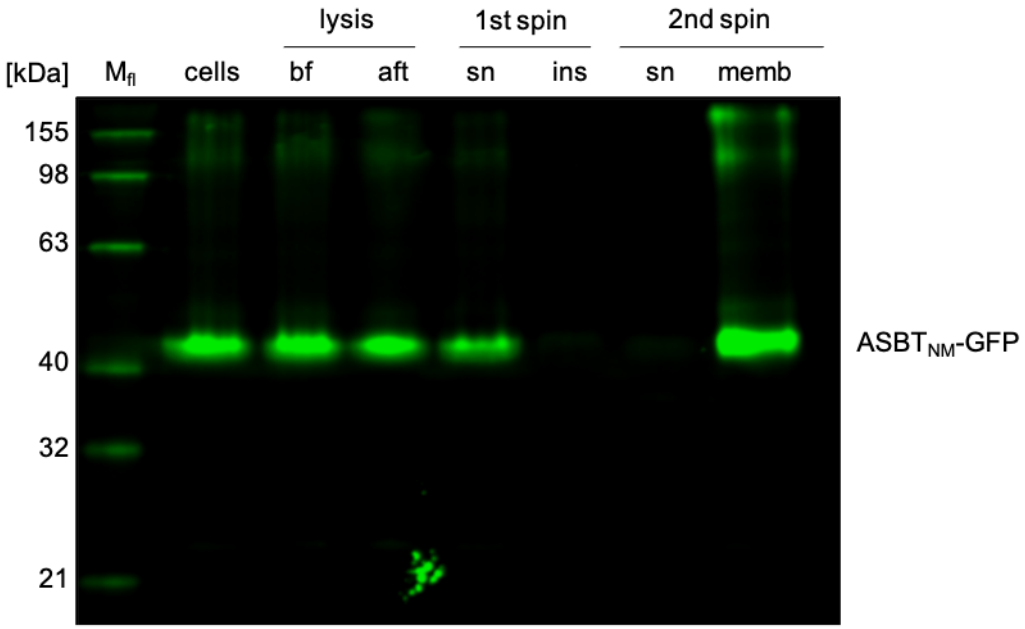


Figure 3.2: SDS-PAGE gel of $ASBT_{NM}$ membrane preparation: In-gel fluorescence of SDS-PAGE gel showing the $ASBT_{NM}$-GFP fusion protein (see right annotation) at different stages during the membrane preparation (cells – cell suspension, bf/aft lysis – before and after lysis, 1st spin sn/ ins – supernatant and insoluble fraction after first spin, 2nd spin sn/memb – supernatant and resuspended membranes after second spin). A fluorescence marker (M_{fl}) in the left lane was used to determine the molecular weight in kDa as indicated on the left.

The in-gel fluorescence gel is shown in figure 3.2 and shows a clear bright band for the ASBT-GFP fusion protein slightly above 40 kDa which is lower than the actual molecular weight of around 60 kDa (ASBT: 33 kDa/ GFP: 27 kDa). Nevertheless, it was proceeded with the experiment as this is a known phenomenon for membrane proteins, mainly due to the different ratio of bound SDS to protein ratio for membrane proteins (Rath et al. 2009). Before lysis, the cell suspension is filtered to remove impurities which could clog the cell disruptor. To ensure that the protein does not stick to the filter, an aliquot was taken and run on the same gel. $ASBT_{NM}$ did not stick to the filter, as the gel sample shows a similar

band intensity for the protein before and after filtration. Additionally, no fusion protein was lost during the lysis due to e.g. aggregation, as assessed by comparing the fluorescence intensities of the first three lanes. In a first low speed spin, unbroken cells, genomic DNA, and cell debris were pelleted. GFP fluorescence was only detected in the supernatant which contained membrane proteins and soluble proteins. In a second high speed spin, membranes were isolated from the solution by pelleting them. No fluorescence was detected in the supernatant which contained the soluble fraction, e.g. soluble proteins, metabolites, etc.. The resuspended membranes showed a clear band for the fusion protein. No lower molecular weight bands were observed indicating that no degradation products are present in the membrane faction. Some higher molecular weight bands were visible which could be potential SDS stable oligomers or aggregates. Nevertheless, it was continued as they were shown to not disturb the later purification process. The membranes could be stored at −80 °C for over two years with no impact on transport activity or crystallization behavior as assessed by former group members.

Purification was performed according to the established protocol in our group. The purification process was monitored using SDS-PAGE and is shown in figure 3.3. The crude membrane extract was thawed and diluted into solubilization buffer. Solubilization was performed by the addition of 1% DDM for 2 h. The in-gel fluorescence gel shows no differences in the amount of $ASBT_{NM}$-GFP fusion protein before and after solubilization indicating that no aggregation occurred. A third high speed ultracentrifugation step was performed to remove the insoluble material. The supernatant of the spin contained solubilized $ASBT_{NM}$-GFP (lane IMAC Inp) while the insoluble material usually showed no or only a faint band (not shown). The solubilized material was mixed with Ni-NTA beads and the slurry left stirring for 4 h to allow binding of the protein to the resin. The slurry was loaded onto a column and the flow through (FT) was collected. Interestingly, the flow through showed a fainter band at slightly lower molecular weight than the input. Nevertheless, the main GFP fusion protein was bound to the resin as assessed by the color of the resin.

To remove unspecific bound proteins, two wash steps with increasing imidazole

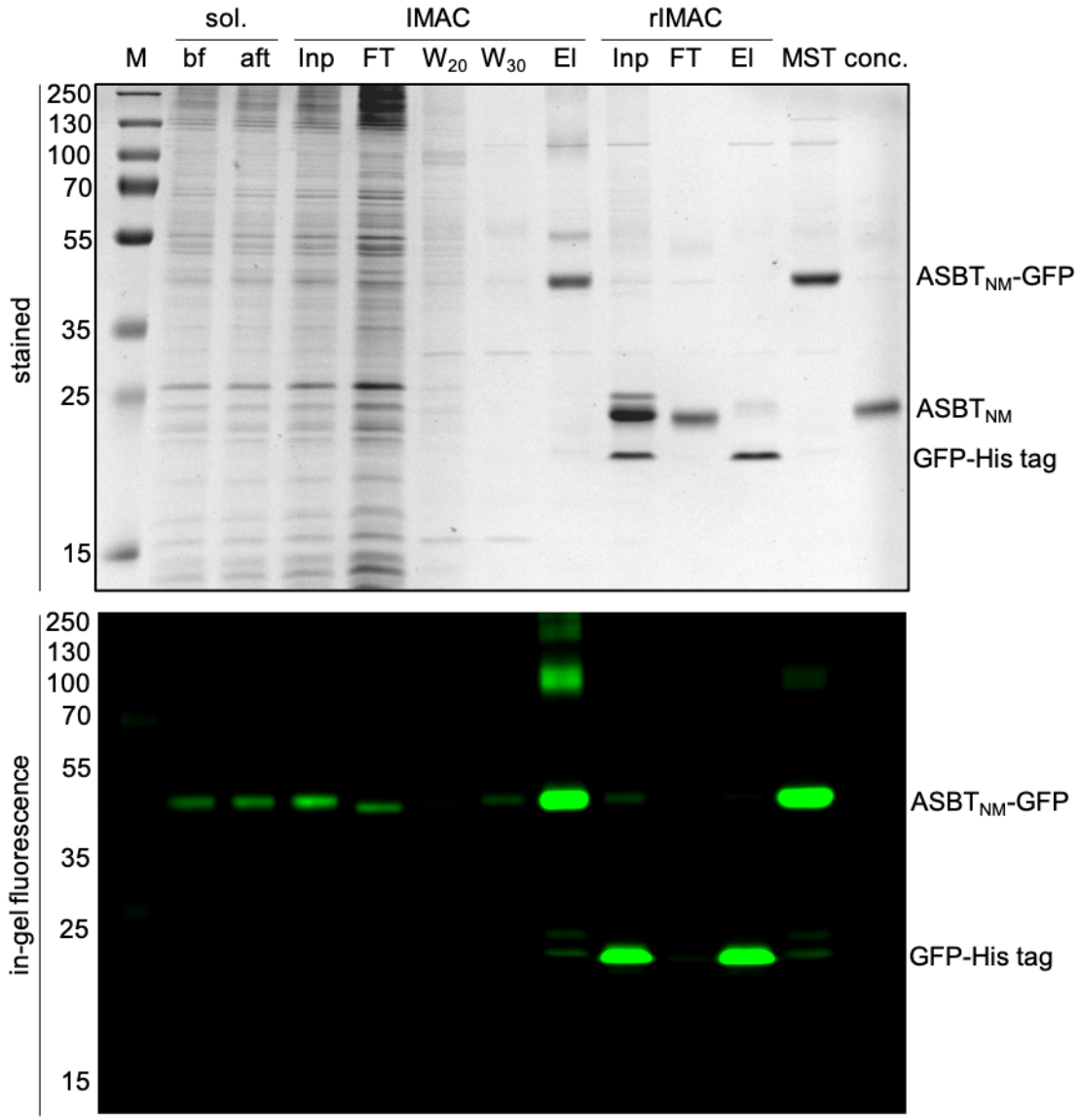


Figure 3.3: SDS-PAGE gel of $ASBT_{NM}$ TEV-cleavable wild type purification: Top image shows the Coomassie stained SDS-PAGE gel/ Bottom gel shows the in-gel fluorescence detected before staining. Lanes are showing the sample at different stages during the purification (bf/ aft sol. – before and after solubilization, IMAC Inp – supernatant of third spin, IMAC FT– Flow through after 4 h hours of binding, W20/W30 – wash fractions with 20 and 30 mM imidazole, IMAC el – elution with 250 mM imdazole, rIMAC Inp – Overnight dialyzed sample with TEV protease, rIMAC FT – FT after TEV digest containing the cleaved protein, rIMAC elution – Elution TEV protease and GFP-His tag with 250 mM imidazole, MST – ASBT-GFP fusion protein sample taken before TEV cleavage, dialyzed overnight for MST experiments, conc. – concentrated sample for transport assay). Prestained marker (M) is shown on the left with the corresponding molecular weights in kDa.

concentrations (20 and 30 mM) were performed. The stained gel showed multiple bands for the 20 mM imidazole wash step but no band in the in gel fluorescence showing that a lot of impurities were removed at this step. The 30 mM imidazole wash step contained some $ASBT_{NM}$-GFP but also showed some bands for unspecific bound proteins indicating that this step is necessary and increased the overall purity of the final protein sample even though slightly decreasing the yield. The protein was eluted into two column volumes of elution buffer containing 250 mM imidazole. The elution shows a clear, intense band for the $ASBT_{NM}$-GFP fusion protein with some minor impurities. The eluted protein was mixed with TEV protease and dialyzed overnight against SEC buffer (20 mM Tris-HCl (pH 7.4), 150 mM NaCl, and 0.03% DDM). After overnight incubation with TEV protease, the in-gel fluorescence gel showed only a weak band for the $ASBT_{NM}$-GFP fusion protein while a strong band for free GFP indicating that the cleavage was successful. A band for the cleaved $ASBT_{NM}$ was observed at around 22 kDa in the stained gel with a fainter band at a slightly higher molecular weight (around 25 kDa) corresponding to TEV protease. The dialyzed sample was applied to a Histrap (rIMAC) column to remove further impurities, the protease, and the fusion tag. The flow through containing $ASBT_{NM}$ showed a clear band with only minor impurities. The elution showed bands for TEV protease and free GFP. A fraction of the dialyzed $ASBT_{NM}$-GFP fusion protein was snap frozen and used for MST experiments. The purity of the fusion protein was considered as good enough for MST as small contaminations should not interfere with the method. The purity of the sample was good enough after the rIMAC and a further purification step (Size exclusion chromatography) was not necessary. For the transport assay, the protein was concentrated to 2 mgL^{-1} (gel lane conc.) and frozen until further usage. For LCP crystallization, the protein was further concentrated to a concentration between 25 and 35 mgL^{-1}.

3.1.3 Size exclusion chromatography

Size exclusion chromatography (SEC) was used to assess the quality of the purified protein. The protein was concentrated to 6 to 8 mgL^{-1} and loaded onto the column. Figure 3.4 shows the size exclusion chromatogram of $ASBT_{NM}$ WT in

20 mM Tris-HCl (pH 7.4), 150 mM NaCl, and 0.03% DDM. The chromatogram showed a single, monodisperse peak at 12.8 ml. Analysis of the peak by SDS PAGE showed that the peak corresponded to $ASBT_{NM}$ WT. The monodisperse peak suggests that $ASBT_{NM}$ is present in only one form, the monomeric form as suggested from previous experiments. Even though a small aggregation peak is visible between 8 and 9 ml, the majority of the protein seems to be properly folded. A small peak was also observed at 14 ml following the $ASBT_{NM}$ peak, but SDS PAGE analysis did not show any corresponding protein band.

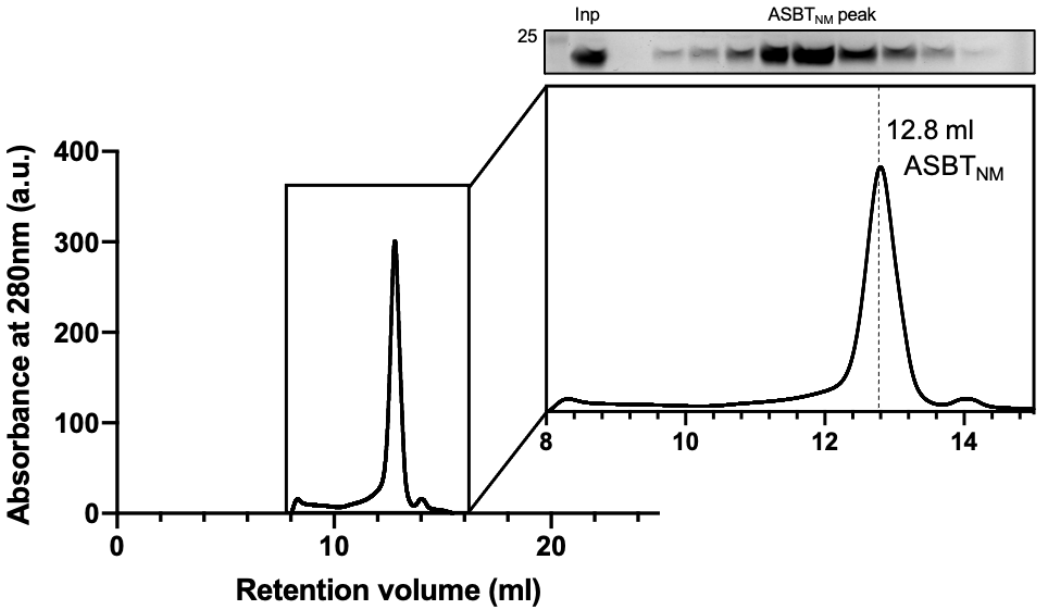


Figure 3.4: SEC chromatogram of $ASBT_{NM}$ wild type in DDM: Size exclusion chromatogram of $ASBT_{NM}$ wild type in 20 mM Tris-HCl (pH 7.4), 150 mM NaCl, and 0.03% DDM. SEC was performed using a Superdex 200 Increase 10/300 GL column. Top right image shows an enlargement of the area between 8 and 16 ml where aggregated protein and $ASBT_{NM}$ are supposed to elute. Almost no aggregation is observed as determined by the small peak between 8 and 9 ml. $ASBT_{NM}$ elutes at 12.8 ml as seen by the peak in absorbance at 280 nm. SDS-PAGE gel shows that the peak contains the $ASBT_{NM}$ protein.

3.2 3C-cleavable constructs

The conventional purification process involves a dialysis step of the eluted protein to reduce the amount of imidazole and to perform the TEV digest. The dialysis step can be considered as one of the most cost intensive steps in the purification

of membrane proteins because of the large amount of detergent needed to be added to the dialysis buffer to keep the protein in solution. Considering the large amount of detergent which will be needed for further purifications of wild type and mutants, a 3C protease-cleavable construct was designed and donated by Dr Deborah Brotherton to test if an on-column cleavage with 3C protease can be performed saving the dialysis step and detergent. 3C protease was chosen over TEV protease as the latter is known to be not as efficient if used for on-column cleavages.

3.2.1 Expression trials of the 3C-cleavable wild type protein

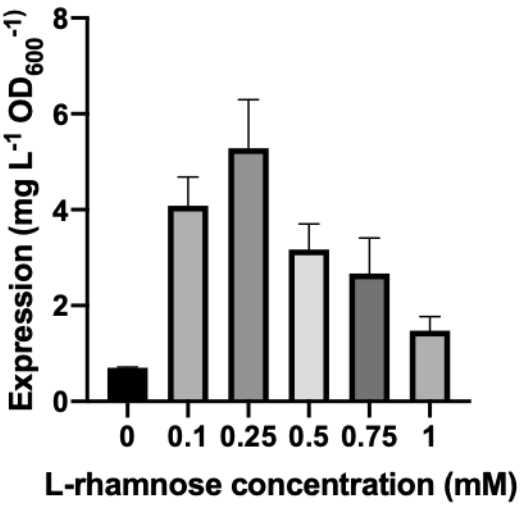


Figure 3.5: Expression trial of $ASBT_{NM}$ wild type 3C-cleavable construct: Effect of different L-rhamnose concentrations on the expression of the 3C-cleavable wild type construct. Expression levels were determined by measuring the GFP fluorescence of each culture. Expression levels were normalized to the OD_{600} of each individual culture and the fluorescence was converted into mg L^{-1}. Mean expression was calculated from three individual experiments.

First, expression trials were performed similar to the TEV construct (see figure 3.5). 0.25 mM rhamnose showed the highest expression followed by 0.1 mM. Overall, the 3C construct showed slightly less expression with a mean yield of 5.3 $mgL^{-1}OD^{-1}$ compared to the 6.3 $mgL^{-1}OD^{-1}$ of the TEV construct. The addition of higher concentrations of rhamnose yielded in expression but to a lower

level. Interestingly, there is a clear downward trend observable for the expression level with increasing rhamnose concentrations in contrast to the TEV construct. Overall, the best expression condition was similar to the TEV construct and it was proceeded with the large scale expression and purification.

3.2.2 Purification of the 3C-cleavable wild type protein

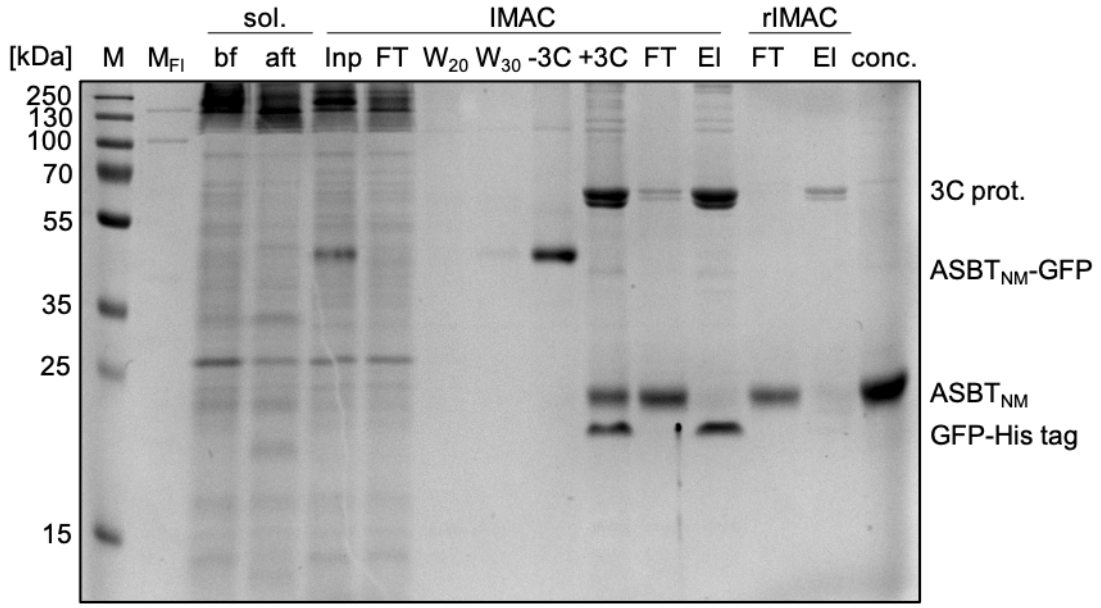


Figure 3.6: SDS-PAGE gel of $ASBT_{NM}$ 3C-cleavable wild type purification: Image shows the Coomassie stained SDS-PAGE gel showing the samples at different stages during the purification (bf/ aft sol. – before and after solubilization, IMAC Inp – supernatant of third spin, IMAC FT– Flowthrough after 4 h of binding, W_{20}/W_{30} – wash fractions with 20 and 30 mM imidazole, -3C – sample of beads before adding 3C protease, +3C – sample of beads after overnight incubation with 3C protease, IMAC FT2 – Flow through, IMAC elution – Elution of 3C protease and GFP-His tag with 250 mM imidazole, rIMAC FT – FT of reverse IMAC to remove residual 3C protease, rIMAC elution – Elution of 3C protease, conc. – concentrated sample for transport assay). Prestained marker is shown on the left with the corresponding molecular weights in kDa. In-gel fluorescence image not shown (see text for more information).

Membrane preparation was performed as previously described. The purification of the 3C cleavable construct was kept similar to the TEV cleavable construct until the IMAC elution step (see figure 3.6). Instead of eluting the GFP fusion protein with imidazole, the protein was kept on the resin after the wash step with 30 mM imidazole. An aliquot of the beads was run on the gel (-3C) and the sam-

ple showed a single band (-3C) corresponding to $ASBT_{NM}$-GFP with only minor impurities. 3C protease was added directly to the resin and gently mixed by pipetting up and down. Cleavage was performed by leaving the column overnight at 4 °C with no agitation. On the next morning, a small aliquot of the resin was taken (lane +3C). The 3C cleavage was successful and two bands appeared at lower molecular weight corresponding to cleaved $ASBT_{NM}$ and GFP. Additionally, a band at higher molecular weight is visible which is the 3C protease. The flow through was collected and the resin was washed with two column volumes of buffer containing 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, and 0.03% DDM to ensure that all cleaved $ASBT_{NM}$ was removed from the column. Elution with 250 mM imidazole showed that only 3C protease and free GFP remained bound to the column and that all $ASBT_{NM}$ was recovered in the flow through. The flow through was applied to a HisTrap column to remove residual 3C protease contamination. The rIMAC was able to separate the residual 3C protease from the pure protein. The flow through was concentrated and showed a high purity similar to the TEV cleavable $ASBT_{NM}$.

3.3 Purification of ASBT mutants and NapA

As mentioned at the beginning, once the expression conditions had been validated and the purification procedure had been performed a couple of times with the wild type protein, it was proceeded with the expression and purification of all missing samples for the transport assay and for MST experiments. In total, 25 $ASBT_{NM}$ mutants were designed by a former PhD student. The following mutants were missing and were purified accordingly: E260AQ264A, E260AL, Q264A, Q77K, Q77L, Q77AQ264A, H294A, T112V, N265A, E260A, M48A, N295A, H294Q. The mutants were stored in 50 μl aliquots at 2 mgL^{-1} for the transport assay. If the yield was good enough, some of the protein was concentrated further and used for crystallization experiments (see structure chapter).

NapA was expressed and purified similar to $ASBT_{NM}$. It was used as a negative control in the transport assay as it belongs to the same fold than $ASBT_{NM}$ but does not transport taurocholate and was therefore considered as a good negative

control for the assay. Expression and purification were performed similar to $ASBT_{NM}$. Overall, NapA showed a lower expression level than $ASBT_{NM}$ but it was enough for the transport assay.

3.4 Protein quality assessment

To assess the protein quality and structural integrity, initial experiments were performed with Circular dichroism spectroscopy and the CPM assay. Showing that the wild type and mutants are stable and properly folded were considered as critical controls for later experiments such as the transport assay.

3.4.1 Circular dichroism

Circular dichroism can be used to assess the secondary structure of a protein and therefore can be used to check the proper folding of the protein. At the beginning, experiments were performed using an in-house JASCO CD spectrometer. The protein was dialyzed overnight against 20 mM Tris-H_2SO_4 (pH 7.5), 150 mM Na_2SO_4, and 0.03% DDM to remove chloride ions which strongly absorb at wavelength less than 195 nm. Initial trials were performed with different concentrations of the protein to find the best and lowest possible concentration resulting in a good spectrum. Two factors were considered for this purpose. First, the CD spectra should show a smooth curve with no spikes. Second, the HT voltage should not exceed 700 V. All five different concentrations resulted in spectra typical for alpha helical proteins in the range between 260 and 200 nm which is in accordance with the solved structure of $ASBT_{NM}$ consisting of 10 helical transmembrane helices (see figure 3.7 a). For higher concentrations, the spectra showed spikes in the range between 200 and 190 nm. Additionally, the HT voltage exceeded 700 V for wavelength lower than 200 nm for higher concentrations (see figure 3.7 a) which probably caused the spikes in the spectra. The best spectrum was obtained with 0.4 mgL^{-1}. The spectra showed a smooth curve with the HT voltage not exceeding 700 V over the measured wavelength range. Lower concentrations showed noisy curves at wavelength below 200 nm similar to higher concentrations.

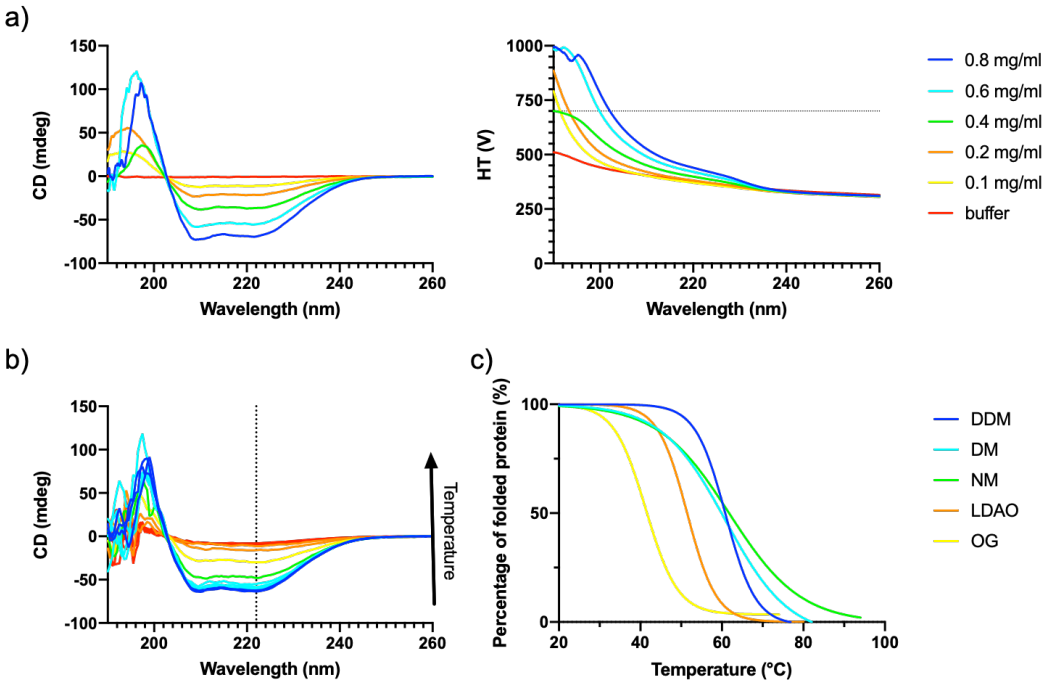


Figure 3.7: CD spectra of $ASBT_{NM}$ wild type at different concentrations and at different temperatures/ Melting curves of $ASBT_{NM}$ wild type in different detergents determined by CD spectroscopy: a) CD spectra of different concentrations of $ASBT_{NM}$ wild type in 20 mM Tris-H_2SO_4 pH 7.5, 150 mM Na_2SO_4, and 0.03% DDM (left hand side). Right diagram shows the dependency between the high tension voltage (HT) and the wavelength. Note that the voltage should not exceed 700 V (dashed line) as signal becomes noisy over this threshold. b) CD spectra of $ASBT_{NM}$ wild type with increasing temperatures (25 °C to 95 °C/ dark blue to red). CD spectra are changing upon heating from an alpha helical spectra to an unfolded protein spectra. Dashed line shows wavelength at 222 nm. Melting curves can be obtained by plotting the ellipticity as a function of temperature. d) Melting curves of wild type protein in different detergents (see material and methods for detergent final concentrations). Fraction of folded protein is shown as a function of the temperature.

The following experiments were carried out with a project student. The stability of $ASBT_{NM}$ in different detergents was checked. For this purpose, the samples were applied to increasing temperatures and a CD spectrum was measured at each temperature. As seen in figure 3.7 c, the protein shows a CD spectrum of a typical alpha helical protein at 20 °C. Upon heating, the protein starts unfolding and the ellipticity starts increasing in the range between 230 to 210 nm until the protein is completely unfolded at higher temperatures where the

Table 3.1: Melting temperatures of $ASBT_{NM}$ wild type protein in different detergents estimated by CD spectra: Melting temperatures of $ASBT_{NM}$ wild type in 20 mM Tris-HCl pH 7.5, 150 mM NaCl, and different detergents (0.03% DDM, 0.3% DM, 0.84% NM, 1.75% OG and 0.06% LDAO). The melting temperature (T_M) is the temperature at which 50% of the protein has unfolded and was obtained by curve fitting to a Boltzmann sigmoidal equation (see materials and methods for more information). T_M is shown as the mean and standard error from three individual repeats.

Detergent	Melting temperature/ T_M (°C)
DDM	60.92 ± 0.82
DM	60.89 ± 1.86
NM	61.84 ± 0.74
OG	41.34 ± 0.39
LDAO	51.47 ± 0.54

spectrum resembles a typical spectrum for unstructured proteins indicated the complete denaturation of the protein. By plotting the ellipticity at 222 nm at different temperatures, melting curves can be obtained which are useful to obtain information about the stability of the protein. Figure 3.7 d shows melting curves of $ASBT_{NM}$ in different detergents. The protein is completely folded in DDM over a long range of temperatures, approximately until 50 °C when the protein started to unfold rapidly and was completely unfolded at 70 °C. The melting temperature in DDM was around 61 °C. Interestingly, the melting temperatures for DM and NM were in the same range even these detergents are considered as harsher detergents. Nevertheless, the graph shows a different unfolding pattern. The protein starts to unfold earlier in DM and NM and unfolding happened upon a longer temperature range. LDAO and OG show the lowest melting temperatures with LDAO showing a melting temperature of 51.47 °C and OG of 41.34 °C degrees. Nevertheless, the protein is still completely folded at lower temperatures in these detergents.

3.4.2 CPM assay

To further check the stability of $ASBT_{NM}$ WT and to investigate the applicability of it in a thermal shift assay, initial experiments were performed using different published protocols. The CPM dye was used which becomes fluorescent upon re-

action with cysteines. If the protein is properly folded and only contains cysteines in the hydrophobic core, the cysteines are not accessible and no fluorescence signal is observed. Upon heating, the protein starts to unfold and cysteines will become accessible which leads to an increase in fluorescence. By measuring the fluorescence at different temperatures, similar melting curves as seen in the section before can be obtained. The different protocols differ in the added amount of CPM dye to the protein (see material and methods for more information). It was decided to perform the initial experiments with a relative high amount of protein (5 μg) to be able to see changes in the fluorescence in case the signal would be relatively low. All protocols worked and resulted in a typical S-shaped curve with the fluorescence increasing upon heating (see figure 3.8). Furthermore, the addition of 1 mM taurocholate stabilized the protein as indicated by the slightly right shifted curves (blue color) and the resulting higher melting temperatures. All protocols showed small differences in melting temperatures, possibly indicating an influence of the CPM concentration on the stability of $ASBT_{NM}$. All protocols except protocol 4 showed differences in the maximum fluorescence between the apo and the taurocholate samples which could be caused by an interaction of excessive dye with the taurocholate. Protocol 4 and 5 showed the smallest error bars indicating that these protocols showed a good reproducibility and it was decided to proceed with protocol 4 and 5.

In the next step, different amounts of protein were used to see which was the lowest amount of protein needed in the assay. As seen in figure 3.9, both protocols showing the typical curves in the range between 1 to 10 μg while the buffer did not show any increase in fluorescence indicating that the protein causes the rise in fluorescence. Comparing the two graphs, it is obvious that protocol 4 resulted in lower error bars and overall better S-shaped curves. Protocol 4 used a lower CPM concentration compared to protocol 5. The maximum fluorescence for 7.5 and 10 μg are the same for protocol 4 assuming that the dye concentration is the restricting factor here and that increasing the concentration of the protein will not lead to a higher signal. In contrast, protocol 5 shows an increase of the maximum fluorescence up to 10 μg indicating that the dye is added in excess allowing it to react with more protein molecules.

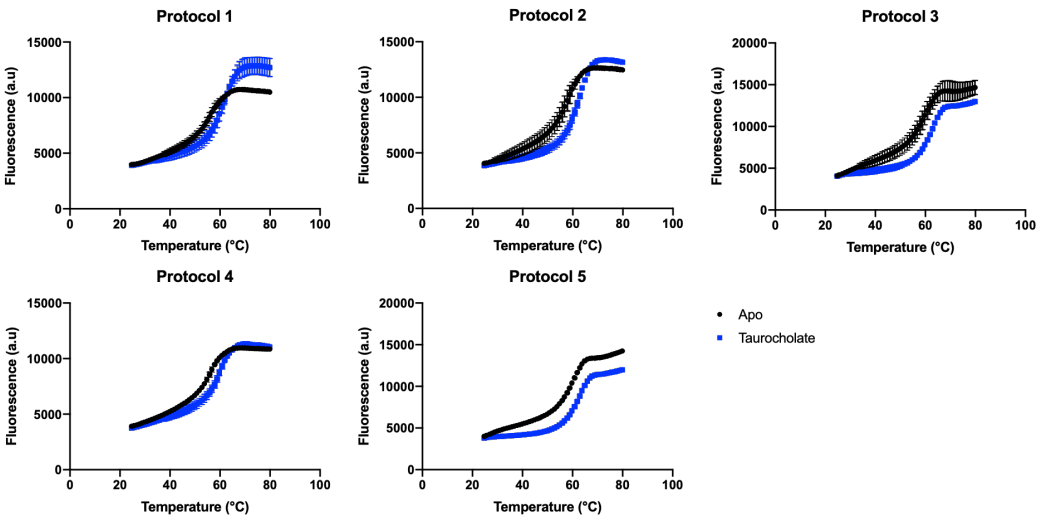


Figure 3.8: Initial melting curves of $ASBT_{NM}$ with and without taurocholate using different published CPM protocols: Melting curves of the apo protein (black) and with 1 mM taurocholate (blue) are shown for different CPM protocols (see material and methods for more information). 5 μg of protein was used in 20 mM Tris-HCl pH 7.5, 150 mM NaCl, and 0.03% DDM.

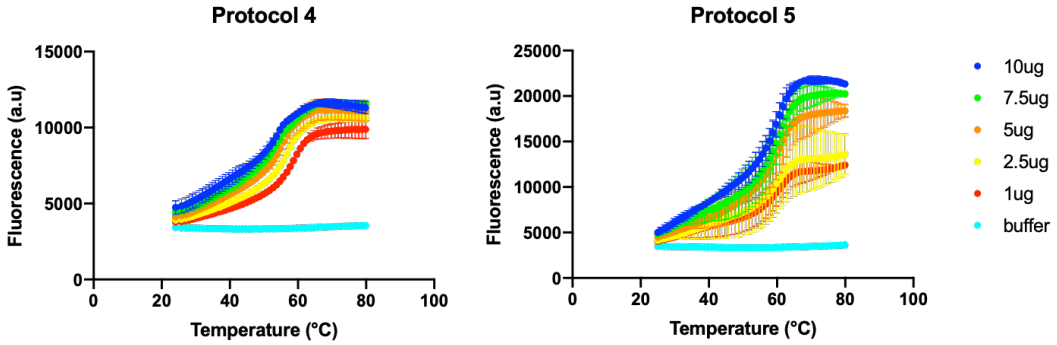


Figure 3.9: Melting curves of different amounts of $ASBT_{NM}$ wild type for protocol 4 and 5: Melting curves are shown for different amounts of protein (see color legend on the right). The assay was performed in 20 mM Tris-HCl pH 7.5, 150 mM NaCl, and 0.03% DDM. The buffer control is shown in cyan.

It was decided to use protocol 4 for further experiments in this thesis as it showed the best results even with low protein amounts like 2.5 μg. To put the CPM assay to the proof a similar experiment was performed than with the circular dichroism. The stability of the proteins was investigated in different detergents. Figure 3.10 shows the obtained melting curves with the fluorescence converted into the percentage of unfolded protein. DDM, DM, and NM show normal s-

Table 3.2: Melting temperatures of $ASBT_{NM}$ wild type protein in different detergents estimated by CPM assay: Melting temperatures of $ASBT_{NM}$ wild type in 20 mM Tris-HCl pH 7.5, 150 mM NaCl, and different detergents (0.03% DDM, 0.3% DM, 0.84% NM, 1.75% OG and 0.06% LDAO). The melting temperature (T_M) is the temperature at which 50% of the protein has unfolded and was obtained by curve fitting to a Boltzmann sigmoidal equation (see materials and methods for more information). T_M is shown as the mean and standard error from three individual repeats.

Detergent	Melting temperature/ T_M (°C)
DDM	55.91 ± 0.22
DM	52.09 ± 0.13
NM	53.20 ± 0.13
OG	37.03 ± 2.57
LDAO	44.12 ± 0.53

shaped curves. DDM shows the highest melting temperature of 55.91 °C followed by NM and DM with 53.20 and 52.09 °C respectively. OG and LDAO show much lower melting temperatures and OG shows an almost linear unfolding at higher temperatures. Nevertheless, these results are showing the same trend than the CD experiment and it was concluded that the CPM assay can be successfully used for further experiments to answer different experimental questions.

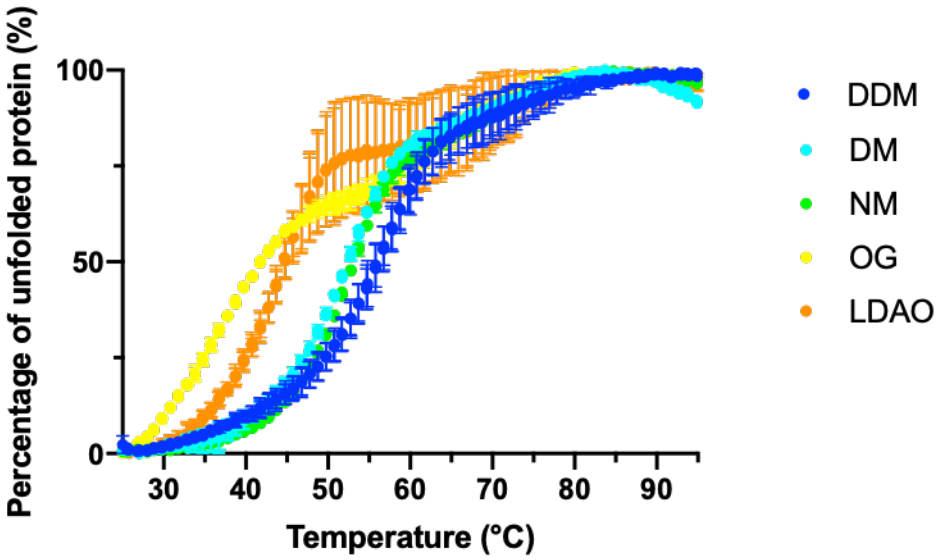


Figure 3.10: Melting curves of wild type protein in different detergents: Melting curves are depicted as the percentage of unfolded protein as a function of the temperature for different detergents (see material and methods for more details) while keeping the buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl) the same. Different detergents are shown in different colors (see legend on the right for more information).

3.5 Discussion

In this chapter, the expression and purification of $ASBT_{NM}$ wild type and mutants were discussed. The wild type and mutants were purified as TEV-cleavable constructs for later applications, mainly for the functional characterization which includes the planned transport assays and the MST experiments. They were purified as the TEV-construct because some mutants have been already purified by former group members and all of them were TEV-constructs.

One of the aims of the projects was to compare the transport activity of different mutants and therefore the construct was kept the same to allow a comparison between them in the transport assay. It needs to be noted here that it cannot be said for sure that the old mutants have not lost activity because of the storage and it was planned to perform a control experiment in which the activity of an old purified mutant and a newly purified mutant will be compared to see if longterm storage at $-80\,°C$ has an effect on the activity but this will be discussed more in detail in the transport assay chapter.

Another control experiment was to show that the wild type protein and mutants are properly folded before usage in any functional assay. This will assure that the differences in binding and affinity can be directly related to the mutation and not due to changes in the secondary structure or unfolding. Therefore, three methods were investigated in this chapter which can be used to show the proper folding of proteins.

Size exclusion chromatography is an indirect method to see if the protein is folded or unfolded by comparing the elution volume of the protein. As described before, the wild type protein was shown to be folded which is not surprising as the purification protocol is well established in our group. If the protein would be unfolded and aggregated, it would elute earlier due to the increase in size. An aggregation peak would be eluting at around 7 to 8 ml for the column which was used for this experiment. This elution volume is known in our group but could be easily verified by comparing the elution volume of an $ASBT_{NM}$ sample which has been denatured by heating or guanidine hydrochloride. Usually, for that purpose molecular weight markers are available which can be used as standards to

estimate the molecular weight of the protein of interest. Unfortunately, they are unsuitable for membrane proteins as they are surrounded by detergent micelles which increase their size and lead to an earlier elution of membrane proteins in comparison to soluble proteins of the same size. Overall, it was shown that size exclusion chromatography can be used for the mutants and that by comparing the peaks with the wild type protein it can be said if the mutants are folded or aggregated and if they can be used for the functional assays.

Next, Circular dichroism experiments were performed to determine the best experimental condition and to confirm the secondary structure of the wild type protein which will also indicate if the protein is folded or not. In agreement with the size exclusion chromatography, the wild type protein showed an alpha helical spectrum which is consistent with the published structure (Hu et al. 2011). Therefore, CD could also be used to confirm the proper folding of all mutants.

The CPM assay was tried to perform which was thought to be a good tool for later applications. By trying different published protocols and concentrations, the best experimental conditions were identified and a protocol has been successfully established.

To evaluate the significance of the results of the CPM assay, melting temperatures of the wild type protein in different detergents were investigated. The same experiment was performed with the CD. Both results showed the same trend with the protein being the stablest in DDM even though for the CD all maltopyranoside detergents (DDM, DM, and NM) showed similar melting temperatures with overlapping errors. LDAO and OG decreased the stability of the protein in both methods with OG showing the lowest stability. Overall, these results fit nicely with published data for the stability of membrane proteins in different detergents. Yang et al describe a general trend of detergent harshness which follows the trend nonionic > zwitterionic > anionic (Z. Yang et al. 2014). Additionally, within nonionic detergents there is a trend of increasing harshness with decreasing aliphatic chains. DDM, DM and NM are all nonionic detergents and in agreement show the highest stability. LDAO is a zwitterionic detergent and showed a reduced stability compared to the maltopyranoside detergents. OG is a nonionic detergent but with a really short aliphatic chain and therefore showed

a reduced stability as well.

Lastly, the introduction of a 3C-cleveable construct was shown in this chapter. Overall, the constructs are similar and they only differ in the protease cleavage sites (TEV: ENLYFQ/G and 3C: LEVLFQ/GP). Not surprisingly, these minimal changes in the protein sequence did not massively alter the expression level with the 3C-construct showing a slightly reduced expression level compared to the TEV-cleavable wild type protein. The protein did not show any changes in its stability during the purification and showed a comparable purity level to the TEV-cleavable constructs even though the purification protocol was slightly altered.

Overall, the 3C-cleavable construct is a good alternative to the TEV-cleavable construct and offers some advantages. The on-column cleavage saves the dialysis step which is necessary to reduce the imidazole concentration. The removal of imidazole is necessary for the rIMAC and imidazole can also affect the crystallization of proteins (Bhat et al. 2018). Therefore, the 3C-cleavable construct is useful as it avoids that the membrane protein will be in buffers with high concentrations of imidazole. Furthermore, the dialysis step uses quite a high volume of buffer depending on the volume of the elution fraction and therefore a reasonable amount of detergent is needed to keep the membrane protein in solution. For example, 0.3 g of DDM are needed for the preparation of 1 l of dialysis buffer to reach a final DDM concentration of 0.03%. Even though 0.3 g does not seem much at the beginning, if many purifications are performed, the used amount of DDM for dialysis will increase and can be considered as quite expensive. Therefore, the 3C-cleavable construct can help to save detergent and money if a lot purifications are performed, especially for purifications which require dialysis in a different, more expensive detergents. For example, the purification of Mhp1 includes a dialysis step in NM which is more expensive than DDM (T. Shimamura, Yajima, et al. 2008).

Chapter 4: Structural characterization of $ASBT_{NM}$ wild type and mutants

To obtain more structural information of $ASBT_{NM}$, crystallization trials using the purified wild type protein and mutants were carried out in the lipidic cubic phase (LCP) which was successfully used for the Q77A mutant by a former PhD student (Sekiguchi 2014). This chapter will discuss the successful crystallization of the wild type protein in different conditions and the E260A mutant.

4.1 Wild type structures and E260A mutant structure reveal differences in sodium binding in the inward-facing conformation

The original wild type structure was obtained using the vapor diffusion method with dehydration of the crystals. The structure was obtained in the presence of taurocholate and the presence of taurocholate was necessary to obtain crystals diffracting to high resolution. The LCP method had been successful in giving a structure of a Q77A mutation (Sekiguchi 2014) and consequently it was decided to focus on this method. For that purpose, the protein was directly used after the reverse IMAC as the purity was considered as good enough. The protein was concentrated to 25 mg/ml and an initial LCP crystallization experiment was set up using the MemMeso screen. After two to three days, small crystals were obtained in a few conditions. In particular, one crystallization condition was chosen at the beginning to proceed with as the condition (0.1 M Lithium sulfate, 0.1 M MES (6.5), 30% v/v PEG 400) did not contain any sodium ions. Initial crystals were small, but when tested for diffraction they diffracted to a resolution of about 4 Å (see figure 4.1 left image) and it was continued with optimizing the condition. Optimization screens were set up altering the salt/precipitant concentration and the pH. Unfortunately, despite several optimization attempts, it was not able to get a better data set of the wild type protein from this crystallization condition as the crystals were quite fragile and they never diffracted better than 4 Å. Only through the addition of either sodium or taurocholate to the crystallization condition, better diffracting crystals could be obtained and data sets were collected for both conditions. The single, rod-shaped crystals were visible after

three days and reached their full-size after three to four weeks. The optimized crystals diffracted to a resolution of around 3 Å (see figure 4.1 right image). First the protein without taurocholate was investigated.

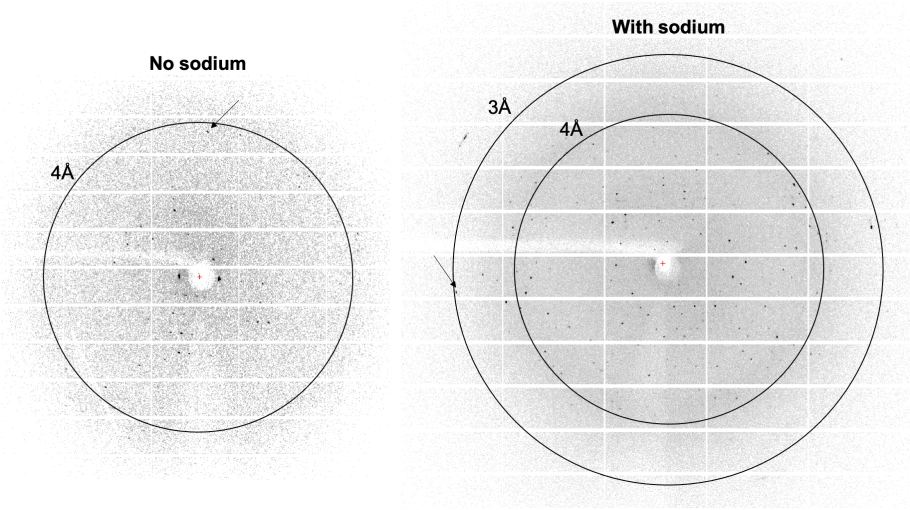


Figure 4.1: X-ray diffraction images of $ASBT_{NM}$ wild type crystals with and without sodium in the crystallization condition: Images show diffraction pattern of crystals grown in 0.1 M Lithium sulfate, 0.1 M MES (pH 6.5), 30% v/v PEG 400 (no sodium/ left site) and with the addition of 0.1 M NaCl (right site). Circles are indicating the resolution and arrows mark spots at the highest resolution.

4.1.1 A Wild type apo structure with sodium

The dataset which was collected from a single crystal was processed using DIALS and afterwards scaled using AIMLESS in CCP4i2 (Project et al. 1994). Diffraction images were inspected for radiation damage and R_{merge}-values were excluded from the analysis if necessary. The resolution was cutoff at 3.18 Å according to the $I/\sigma I$ value. The structure was solved by Molecular Replacement using the vapor diffusion structure (3zuy) as a model (see methods for more information). The crystal packing is shown in figure 4.2. The space group was $P2_12_12$ with the asymmetric unit containing two $ASBT_{NM}$ monomers which are interacting through their intracellular parts (see figure 4.2 b). The structure was refined in Phenix using XYZ coordinates refinement and individual B-factors (Adams et al. 2010). Additionally, the best results were obtained when NCS and secondary structure restraints were applied. After each round of refinement, the model was checked in Coot and further structure building, optimization and changes were

performed until the R-factors did not decrease anymore and the structure statistics were satisfying as estimated by MolProbity in Phenix (Adams et al. 2010). The structure was refined to a final R value of 22% and a R_{free} value of 27%. The final statistics can be seen in table 4.1.

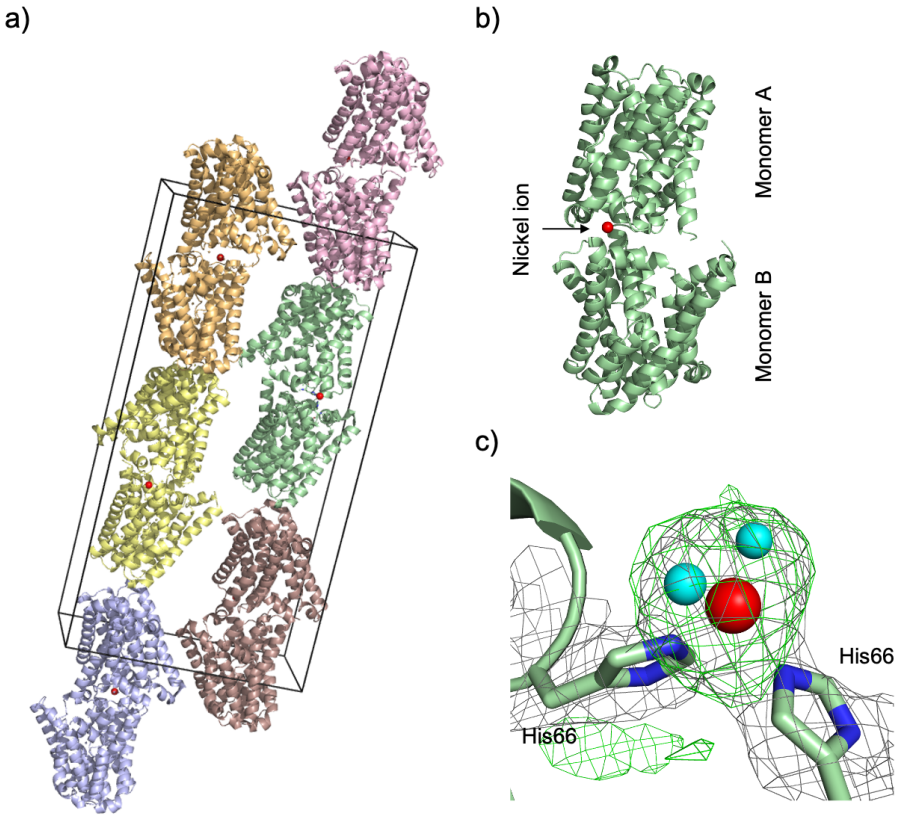


Figure 4.2: Crystal packing of WT apo sodium with asymmetric unit and Ni(II)-ion mediating crystal contact: a) Crystal packing is shown with monomers belonging to one asymmetric unit presented in the same color. Unit cell is indicated as black lines. The Nickel ions mediating crystal contact between monomers of the same asymmetric unit is depicted as red spheres. b) Closer view of the two monomers forming an asymmetric unit. Monomers are interacting through their intracellular sites with the arrow indicating the Nickel ion. c) Ni(II)-ion (red sphere) interaction with the two His66 residues of each monomer and two water molecules (cyan spheres). Electron density is shown as a grey mesh at 1.0 sigma according to the 2mFo-DFc map calculated after placing the Ni(II) ion and the two water molecules. mFo-DFc map is shown as green mesh at +3.0 sigma and was calculated when the ion and water molecules were deleted before refinement.

Interestingly, some strong positive electron density was observed at the interface of the two monomers in close proximity to histidine 66 of both monomers (see figure 4.2 c). The density was too strong to be from a water molecule or a sulfate

ion. The density was assigned as a metal ion, most likely a nickel ion which came from the purification process. In the structure, the Ni(II)-ion interacts with two histidines and two water molecules. Even though considered as a contamination from the purification process, the Ni(II)-ion actually seems to be beneficial for the crystallization by facilitating crystal contact.

Both monomers of the asymmetric unit are in the inward-facing conformation. Monomer B is shown as an example in figure 4.3. Strong density was observed in the sodium binding sites and two sodium ions (pink spheres) were modeled into the sites. Additionally, three monoolein molecules and one aliphatic carbon chain were built into the structure. The aliphatic carbon chain could belong to a monoolein or a copurified lipid but because there was no clear density for the headgroup it could not be unambiguously determined. One monoolein molecule was modeled into the binding pocket (see figure 4.3/ right site) but it could be also some other lipid molecule. The monoolein molecule does not show any polar interaction with the protein. The only residue close to the polar head group of the monoolein molecule is asparagine 295 but the hydrogens of the amide group are about 2.8 Å away from the nearest oxygen of the monoolein‘s hydroxyl group and therefore too far away to form a hydrogen bond. Overall, the binding pocket is quite hydrophobic and it is likely that the interaction is based on the hydrophobic properties of the binding pocket and the aliphatic tail of the monoolein. Even though the interaction seems to be rather unspecific the bound monoolein molecule could have benefited the crystallization process by decreasing the overall flexibility of the protein.

As mentioned before, both monomers crystallized in the inward-facing conformation with only minor structural differences showing an overall RMSD of 0.268 Å. If superimposed separately, the panel domain showed the main difference compared to the core domain with a RMSD of 0.384 Å and 0.165 Å respectively. A similar result was obtained when the structure was compared to the vapor diffusion structure of $ASBT_{NM}$ (3zuy) showing a more flexible panel domain (RMSD=0.763 Å) and a more rigid core domain (RMSD=0.258 Å) (aligned to monomer B). As seen in figure 4.4b, the main differences in the core domain are found in the extracellular and intercellular loops and adjacent helices but are

Table 4.1: Data collection and refinement statistics for the $ASBT_{NM}$ wild type apo structure with sodium: Crystals were grown in 0.1 M Lithium sulfate, 0.1 M MES (pH 6.5), 30% v/v PEG 400 and 0.1 M NaCl. Data was processed using DIALS and AIMLESS in CCP4i2 (Project et al. 1994). Structure was refined using Phenix (Adams et al. 2010).

Crystal	Wild type apo with sodium
Wavelength	0.9795Å
Beamline	I24/ DLS
Oscillation range	0.10°
Exposure time	0.034 s
Transmission	9.97 %
Beam size	20x20 μm
Space group	$P2_12_12$
Resolution	63.4 - 3.18Å(3.294 - 3.18)
CC_{50}	0.91 (0.652)
Cell dimensions	42.8917 92.3925 174.301 90 90 90
Number of measured reflections	67550 (6941)
Number of unique reflections	12283 (1192)
Completeness	96.84 % (97.87)
Multiplicity	5.5 (5.7)
$I/\sigma I$	5.63 (1.06)
R_{merge}	0.3257 (0.9171)
Refinement	
Number of reflections	11909 (1192)
R-factor	0.2156
R-free	0.2696
R.m.s.d from ideal values	
Number of non-hydrogen atoms	4696
Bond lengths	0.003
Bond angles	0.47
Ramachandran favored	98.21 %
Ramachandran allowed	1.63 %
Ramachandran outliers	0.16 %
Rotamer outliers	0.64 %

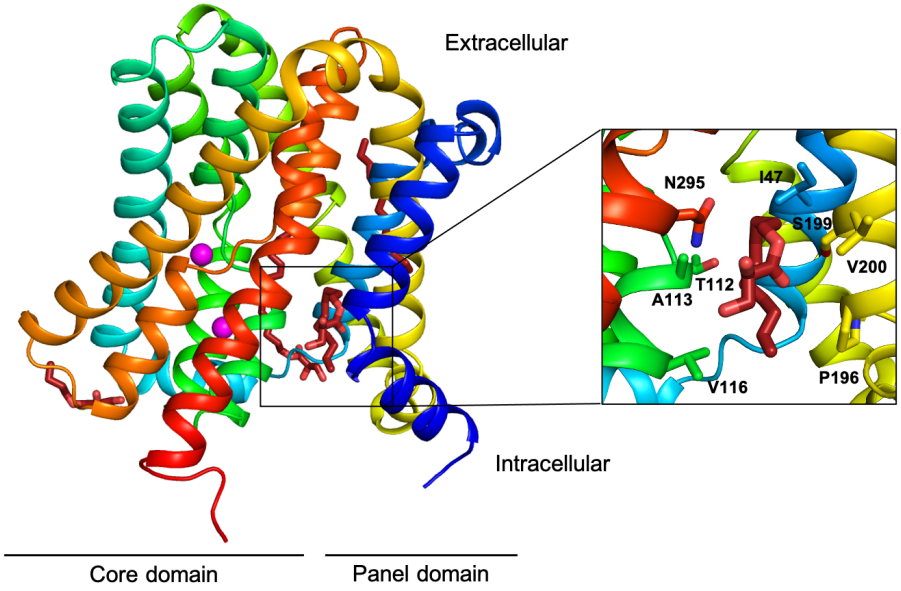


Figure 4.3: $ASBT_{NM}$ WT apo with sodium structure: Side view of $ASBT_{NM}$ WT apo with sodium in ribbon representation. Helices 1 to 10 are colored in different colors starting with TM1 in dark blue and TM10 in red. Sodium ions are shown as pink spheres and monoolein molecules/ aliphatic chains as sticks in dark red. The monoolein molecule found in the substrate binding pocket is enlarged on the right site with residues lining the pocket represented as sticks.

only small differences. The connecting loops between panel and core domain, intracellular loop 1 (IL1) and extracellular loop (EL4) (see figure 4.4a), were not included into the alignment of panel and core domain, but will be discussed together with the core domain. EL2/TM3 are slightly moved towards the center of the core domain compared to the vapor diffusion structure while the connecting loops/helices to the panel domain EL3/TM5 and EL4/TM8 are pushed towards the outside (see arrows in the figure for movement direction). On the intracellular site, IL1/TM3 have moved towards the outside to counteract the movement towards the center of TM3 on the extracellular site accompanied by the movement of IL2 towards IL1 and IL4/TM8 into the center to take up the freed space.

The panel domain shows more differences in comparison to the core domain (see figure 4.5). The intracellular half of TM1 has moved towards IL3 in the apo structure. Furthermore, the upper part of TM1 and IL1 show a slight upwards movement with TM2 moving towards the core domain in agreement with the movement of IL1/TM3 in the core domain. Even though EL3 moved towards the

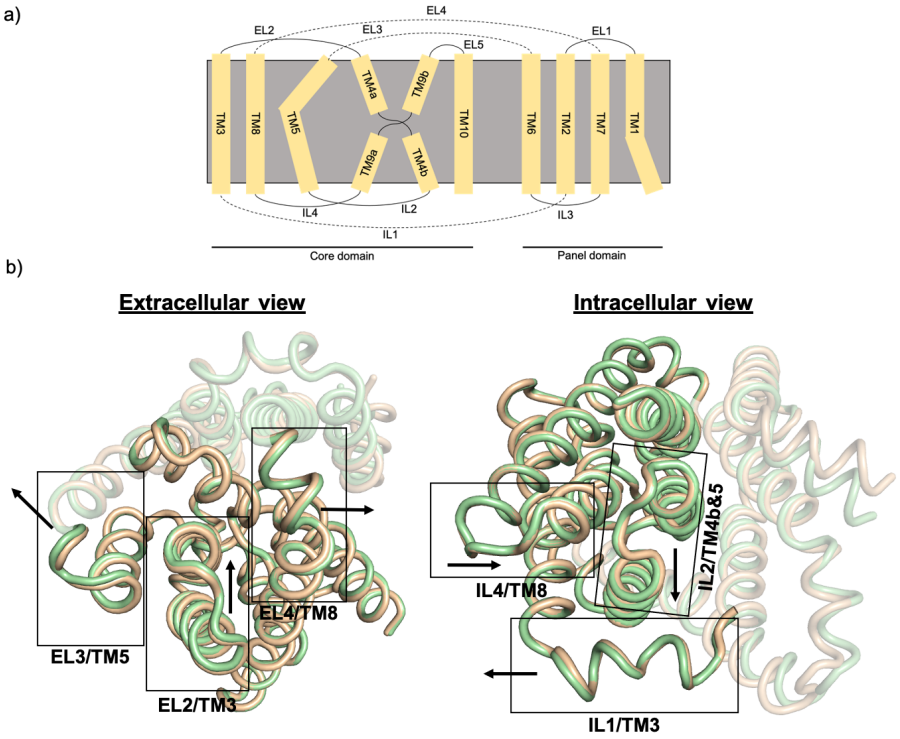


Figure 4.4: Schematic representation of $ASBT_{NM}$ topology and superimposition of the core domain of the WT apo structure with the taurocholate bound structure: a) Helices (TMs) are shown in yellow and the membrane plane is indicated in grey. Intracellular loops (IL) and extracellular loops (EL) are shown as lines. Loops connecting helices in the same domain are shown as solid lines, while loops connecting the different domains with each other are depicted as dotted lines. b) Cartoon representation of the core domain of the sodium structure (green) and taurocholate bound structure (wheat). The panel domain is transparent. Areas which show differences are highlighted with black boxes and boxes are labelled with the corresponding moving parts. The arrows indicate the direction of movement.

panel domain there is no difference visible in the upper part of TM6. In contrast, the bottom part of TM6 together with IL3 shows an upwards movement into the membrane plane which continues for TM7 but not to such an extent.

The structure was assigned as an inward-facing apo structure with sodium bound. As the vapor diffusion structure is a substrate-bound inward facing conformation, it was next investigated if the binding of the substrate might be responsible for the described differences between the structures. In the taurocholate bound structure, the cholesterol backbone is located between TM4b and TM7. In agreement with the relatively large substrate, it seems that in the substrate bound state, TM7 and TM2 have moved slightly away from the substrate bind-

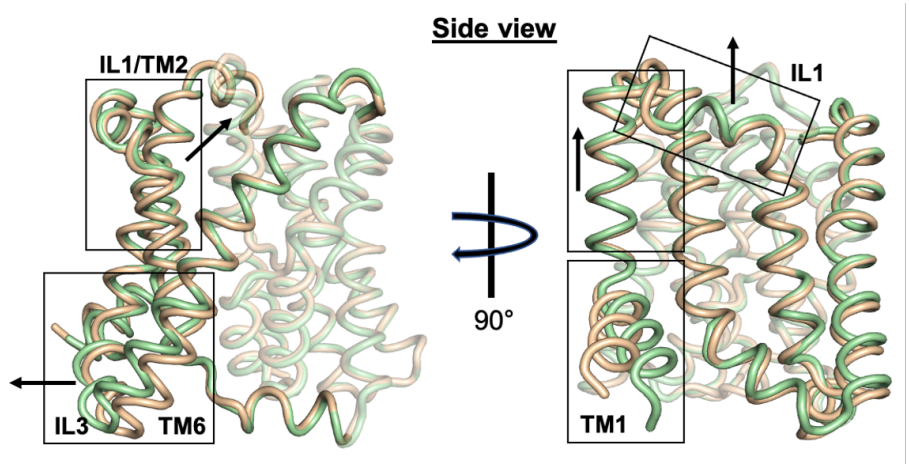


Figure 4.5: Comparison of the panel domains in the WT apo sodium structure and taurocholate bound structure: Side view of panel of the wild type sodium structure (green) and taurocholate bound structure (wheat). Right image shows the model 90° rotated. The core domain is transparent. Areas which show differences are highlighted with black boxes and boxes are labelled with the corresponding moving parts. The arrows indicate the direction of movement.

ing pocket to create enough space for the substrate (see figure 4.6a). The taurine moiety of the taurocholate projects into the outside and is located close to TM1. Additionally, a LDAO molecule is found at the kink of TM1 (see figure 4.6b). In contrast, TM1 is closer to the core domain in the substrate bound state even though both molecules are present. Last, it was checked if the different crystallization methods could have caused the structural differences. Indeed, there is a monoolein molecule located at the intracellular part of TM6 which could explain the movement of TM6 and IL3 in the apo structure (see figure 4.6c). In the vapor diffusion structure, nothing was seen at this position.

4.1.2 A WT structure in a low sodium condition

Next the structure of $ASBT_{NM}$ in the presence of taurocholate but under low sodium ion conditions was investigated. Similar to the addition of 100 mM NaCl, the addition of 1 mM taurocholate to the crystallization condition yielded better diffracting crystals and three data sets were collected from one single crystal and merged together. The resolution was cut off at 3.1 Å. The space group has changed to $P2_1$ (see table 4.2) with the asymmetric unit containing four

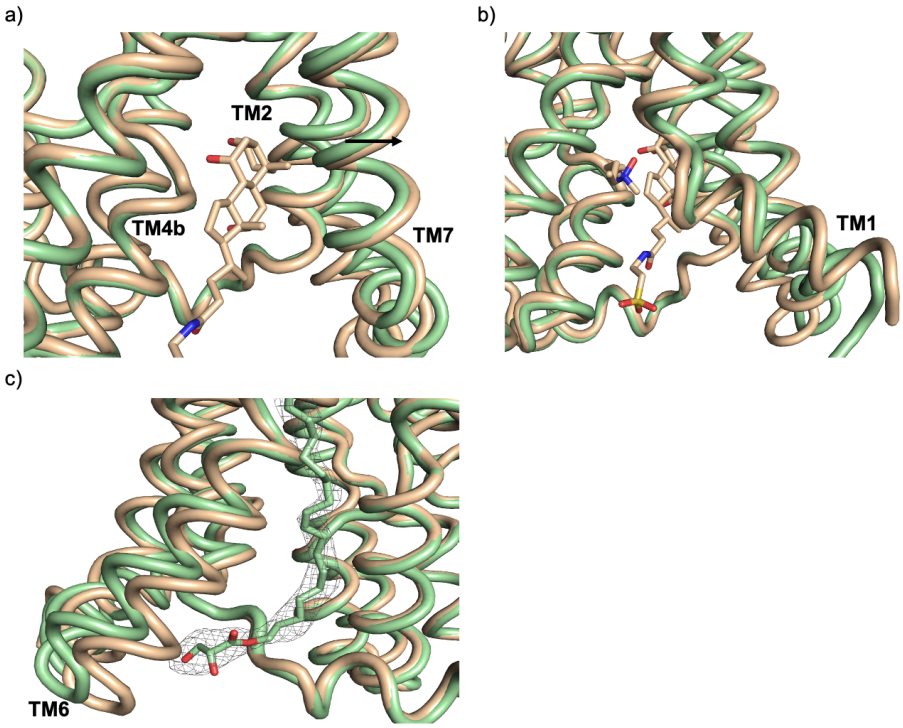


Figure 4.6: Potential substrate- and monoolein induced changes in the WT apo sodium structure: The wild type sodium structure is depicted in green and the taurocholate bound structure in wheat. a) Close-up view of the substrate binding pocket with taurocholate molecule shown as sticks. The arrow indicates the direction of movement of the taurocholate structure (3zuy). b) Close-up view of TM1 with taurocholate molecule and LDAO molecule shown as sticks. c) Monoolein molecule is shown which is bound close to TM6 in the WT apo sodium structure. Electron density is shown as grey mesh at +1 sigma.

monomers. Figure 4.7 shows the asymmetric unit and the labelling as how they will be referred to. All four monomers interact through their extracellular parts. Even though taurocholate was added to the crystallization condition, no taurocholate molecules could be found in the structure and therefore the structure will be referred as a WT apo low sodium structure as the only source of sodium comes from the addition of 1 mM sodium taurocholate. The sodium concentration can be calculated from the dilution during the LCP preparation and the additional sodium taurocholate. The final concentration was estimated to be around 4.5 mM which is in the range of the expected K_M for sodium for sodium-dependent transporters, e.g. the affinity of NapA for sodium is around 4 mM (Lee, Kang, Von Ballmoos, et al. 2013) and the EC_{50} for sodium in $ASBT_{YF}$ was estimated to be around 5 mM (X. Zhou et al. 2014).

Table 4.2: Data collection and refinement statistics for the $ASBT_{NM}$ wild type structure in a low sodium condition: Crystals were grown in 0.1 M Lithium sulfate, 0.1 M MES (pH 6.5), 30% v/v PEG 400 and 1 mM taurocholate. Data was processed using DIALS and AIMLESS in CCP4i2 (Project et al. 1994). Structure was refined using Phenix (Adams et al. 2010).

Crystal	WT low sodium
Wavelength	0.9795Å
Beamline	I24/ DLS
Oscillation range	0.15°
Exposure time	0.01 s
Transmission	23.11 %
Beam size	20x20 μm
Space group	$P12_11$
Resolution	77.1 - 3.1 Å(3.211 - 3.1)
CC_{50}	0.894 (0.68)
Cell dimensions	84.7999 86.5556 95.4162 90 114.604 90
Number of measured reflections	179058 (17864)
Number of unique reflections	22961 (2257)
Completeness	99.13 % (99.73)
Multiplicity	7.8 (7.9)
$I/\sigma I$	12.25 (1.76)
R_{merge}	0.4617 (0.7781)
Refinement	
Number of reflections	22773 (2256)
R-factor	0.2204
R-free	0.2866
R.m.s.d from ideal values	
Number of non-hydrogen atoms	9164
Bond lengths	0.003
Bond angles	0.74
Ramachandran favored	98.20 %
Ramachandran allowed	1.72 %
Ramachandran outliers	0.08 %
Rotamer outliers	0.75 %

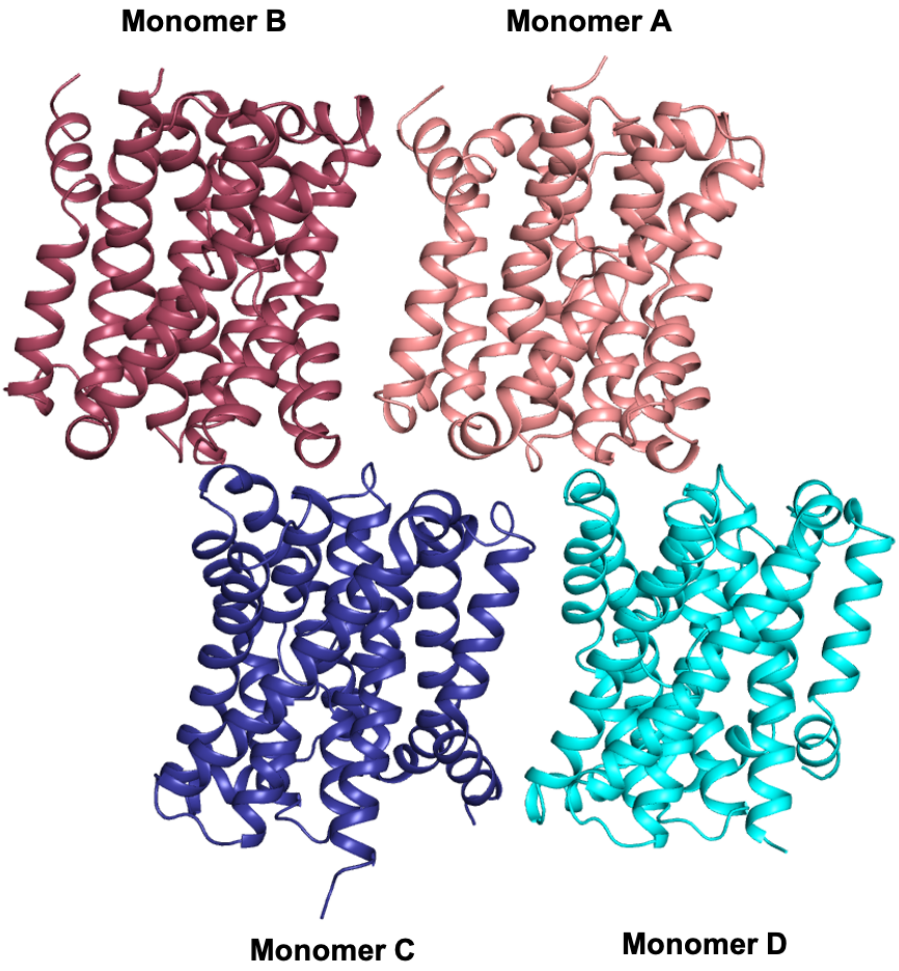


Figure 4.7: Asymmetric unit of the WT apo low sodium structure: Cartoon representation of all four monomers. Monomers are different colored (Monomer A – salmon, Monomer B- raspberry, Monomer C – dark blue, Monomer D – cyan)

When overlaid against each other and the WT apo structure, all monomers show a high similarity between each other ranging from 0.290 Å (MonB and D) and 0.381 Å of C and D. When compared separately, the panel domain shows a higher flexibility similar to the WT apo sodium structure discussed before. TM1 and EL1 adapt slightly different conformations in each monomer (see figure 4.8). TM2 stays almost the same in all monomers and does not seem to be affected by the movements of TM1 (Figure 4.8a). Similar to the WT apo structure, TM6 and IL3 show differences between the monomers indicating that these areas are more flexible.

The core domains are more similar to each other with overall lower RSMD to each other. The main differences are found on the linking loops between panel

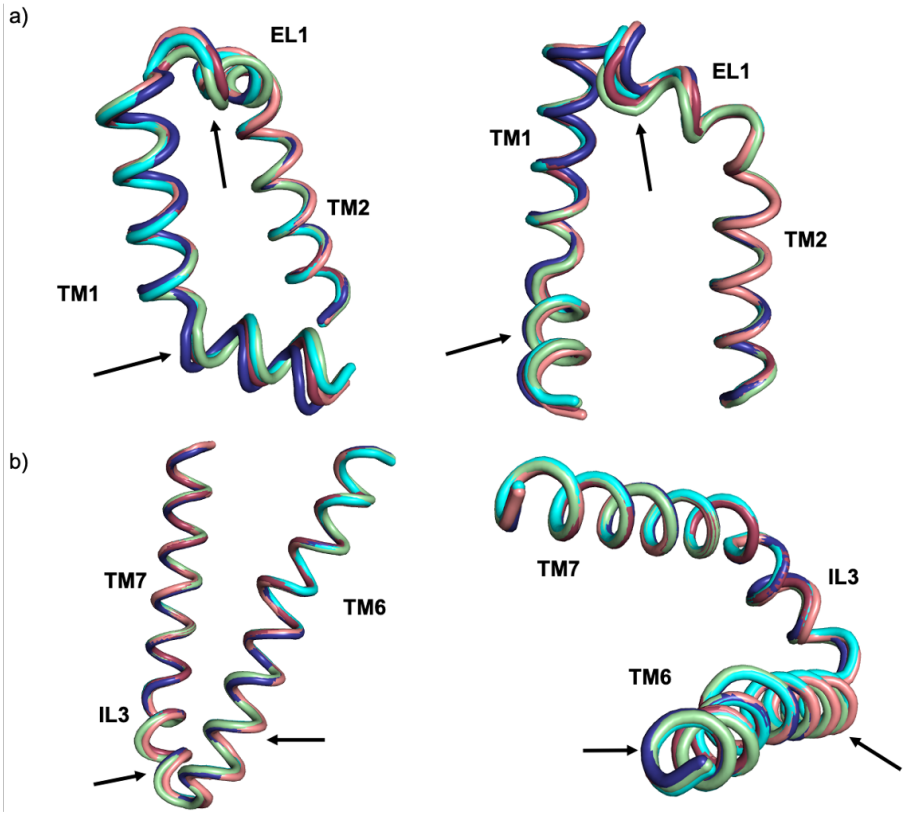


Figure 4.8: Differences in the panel domains of the WT apo low sodium structure: Helices are shown in ribbon presentation with the monomers represented in different colors (Monomer A – salmon, Monomer B- raspberry, Monomer C – dark blue, Monomer D – cyan). a) Differences in TM1, EL1, and TM2 from different angles. Structural parts showing higher flexibility are highlighted with arrows. b) Differences in TM6, IL3, and TM7 with areas showing differences highlighted with arrows.

and core domain and the adjacent helices in the core domain (see figure 4.9). IL1 adapts different conformation in all monomers together with the intracellular half of TM3. On the extracellular site, out of the both connecting loops, only EL3 shows some flexibility between the different monomers but the movement does not affect TM5 in the core domain. Furthermore, EL2 in monomer C shows a different conformation together with the extracellular half of TM3.

Interestingly, TM3 seems to show some flexibility between all monomers. During the refinement and structure building some differences in sodium density/ binding were observed between the monomers and will be discussed in the next paragraph in comparison to the WT apo sodium structure. In the WT apo sodium structure, clear positive density was observed in both sodium binding sites which is not surprisingly as the crystallization condition contained 100 mM

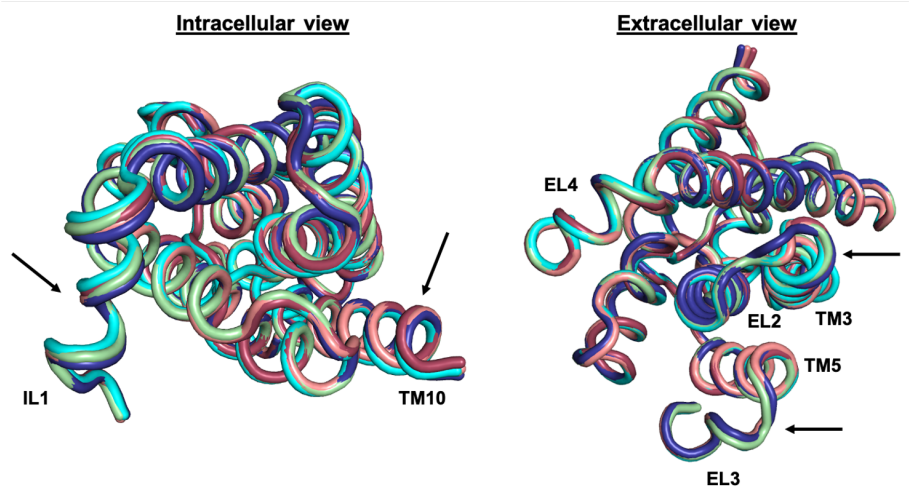


Figure 4.9: Differences in the core domain of the WT apo low sodium structure: Helices are shown in ribbon presentation with the monomers represented in different colors (Monomer A – salmon, Monomer B- raspberry, Monomer C – dark blue, Monomer D – cyan). Left image shows the core domain from the intracellular site and the right image from the extracellular site. Flexible areas are highlighted with arrows.

sodium chloride (see figure 4.10/ top left). The same applied to monomer A and B with density for both sodium ions in both sites, even though the crystallization condition contained less sodium (around 4.5 mM as estimated before). No major differences were observed in the sodium binding sites and in agreement, monomer A and B showed of all the monomers the highest similarity to the WT apo sodium structure with a RMSD of 0.345 Å.

For monomer C and D, density for sodium ions was only visible in one of the two sodium binding sites, but it is difficult to say at this resolution. Na1 is occupied in monomer C. Again, no major conformational changes were observed for the sodium binding residues with a RMSD of 0.208 Å and at this resolution, it is not possible to comment on minor movements. Gln77, a Na2 residue, lies on TM3 which seemed to act differently in monomer C in comparison to the other monomers. Although the intracellular half of TM3 including IL1 has shown flexibility in all the monomers, only in monomer C the intracellular half of the helix moved towards the core domain (see figure 4.11) while the extracellular part moved away from the core domain. It cannot be said if the movement is related to sodium release in Na2, but it was mentioned here and will be further investigated in later structures.

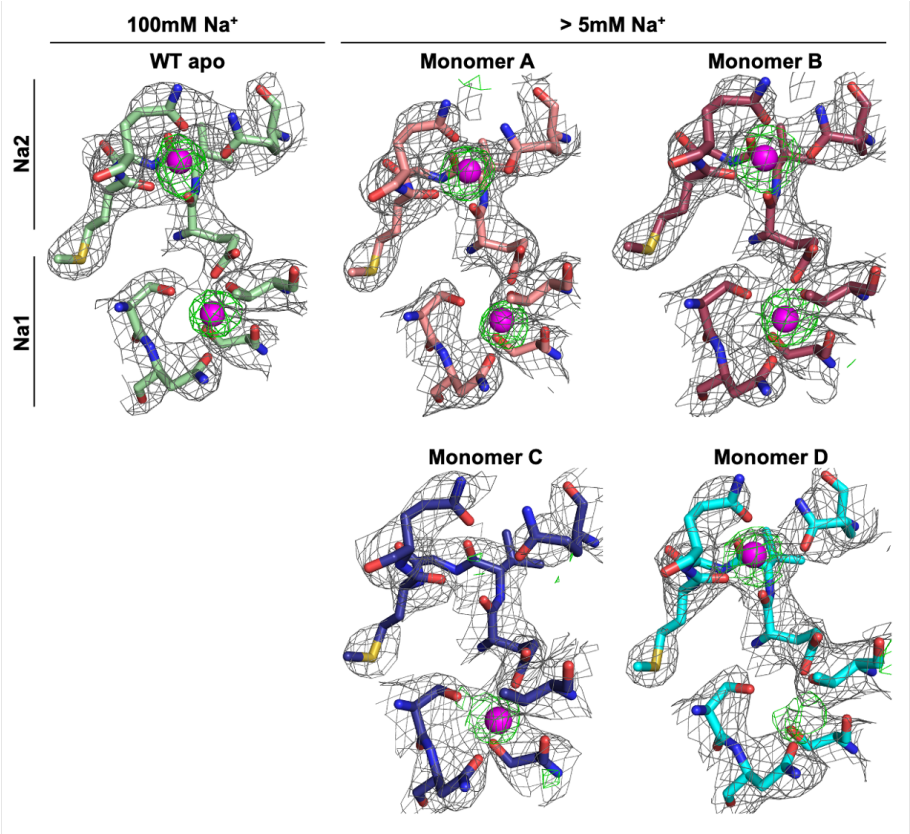


Figure 4.10: Comparison of sodium binding in the WT apo sodium structure and low sodium structure: Sodium binding sites are shown as stick representation (WT apo sodium- green, Monomer A – salmon, Monomer B- raspberry, Monomer C – dark blue, Monomer D – cyan). 2mFo-DFc map is shown at 1 sigma (grey mesh). The mFo-DFc map was calculated with no sodium ions present during the refinement and the resulting positive density is shown at +3 sigma (green mesh). Final sodium concentrations are indicated at the top.

For monomer D, it was trickier to say whether two sodium ions are bound or not. Na2 showed clear positive density without sodium which disappeared after modeling in the sodium ion. On the other hand, some positive density was visible in Na1, but the positive density did not disappear after another round of refinement with a sodium ion built in. This could be indicating that the density rather resembles noise instead of an actual atom at this position. Additionally, the B-factor of the modeled sodium ion differentiated greatly from the B-factors of the surrounding residues which also indicates that the Na1 is likely to be unoccupied in monomer D. Overall, no specific movements for monomer D could be identified when investigated against the other structures, especially around Na1. Some flexibility was observed in TM5 for all monomers which is connected

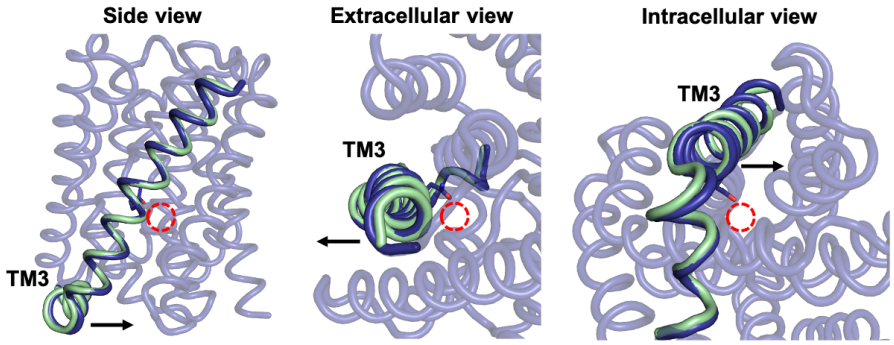


Figure 4.11: Comparison of TM3 in monomer C and WT sodium structure: Monomer C (dark blue) is shown in the ribbon representation with TM3 highlighted. TM3 of the WT apo sodium structure is shown in green for comparison. The location of Na2 is highlighted as a red circle and Gln77 is shown in stick representation. Images are shown movement of TM3 from different angles with the movement direction indicated as arrows.

to Na1 by residues Ser128 and Thr132, but in contrast to the described movement for monomer C, the movement of TM5 in monomer D was not in an opposite direction indicating a general flexibility of TM5 instead of possible difference because of sodium release in Na1.

4.1.3 A E260A structure with low sodium

Next, a E260A mutant structure was obtained using the LCP method as the equivalent E254A mutant of $ASBT_{YF}$ crystallized in an outward-facing conformation in LCP. A single crystal grew in an initial MemMeso screen in the A10 condition containing 0.1 M Lithium sulfate, 0.1 M HEPES 7, 40% v/v PEG 200. The crystal diffracted to a maximum resolution of about 3 Å and a data set was collected. The data set was indexed in C2 and the resolution was cutoff at 3.2 Å. Even though the R_{meas} and R_{merge} values are quite high for the outer shell at this resolution, the cutoff can be also justified from the quality of the electron density. The asymmetric unit contains two monomers in the same orientation which are tightly packed (see figure 4.12a). For one monomer, electron density for the whole C-terminus up to the 3C-cleavage site was visible and could be built in (see figure 4.12a). The C-terminus interacts with another C-terminus from a different layer via a Ni(II)-ion or different metal ion (not shown in figure) and therefore mediates crystal contact. The interaction seems to make the C-terminus more rigid

which potentially made it possible to see electron density for it and to model it in. As mentioned before, the two monomers crystallized in the same orientation regarding the cyto- and extrasolic part. Figure 4.12b shows the orientation of the two monomers from the extracellular site. Monomer B is slightly rotated of about 30° clockwise compared to monomer A as seen from the panel domains and EL3/4 with EL4 of monomer A and EL3 of monomer B being in close proximity. On the intracellular site, TM3 and IL1 of monomer A are moved to TM1 and the cavity between TM1 and the core domain of monomer B. Density for sodium in Na2 was found only in monomer A (pink spheres). No density was visible for Na1. Both will be discussed in more detail a bit later.

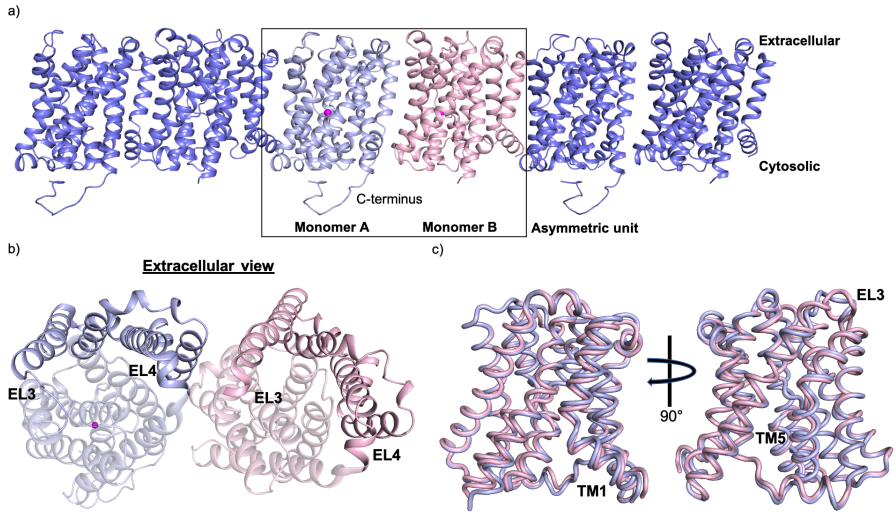


Figure 4.12: Crystal packing of E260A low sodium structure, extracellular view of the asymmetric unit and overlay of the two domains: a) Crystal packing of E260A is shown in one plane. Asymmetric unit is highlighted by a box and monomer labelling is shown. b) Asymmetric unit is shown from the extracellular site with the core domain (transparent) and the panel domain. Extracellular loop 3 and 4 (EL3/4) are labeled to see the rotation of the monomers to each other. c) Superimposition of the two monomers from two different angles showing some differences between the two monomers (monomer A – light blue, monomer B – pink)

Superimposition of the two monomers reveal that they show differences to each other with a RMSD of 0.635 Å (see figure 4.12c). Monomer B is more similar to the WT apo sodium structure with a RMSD of 0.462 Å. The main discrepancies are seen in TM1, TM3 and TM5/6 which are connecting the panel and core domain except for TM1 (see figure 4.12c). If overlaid separately, the

Table 4.3: Data collection and refinement statistics for the $ASBT_{NM}$ E260A low sodium structure: Crystals were grown in 0.1 M Lithium sulfate, 0.1 M HEPES 7 and 40% v/v PEG 200. Data was processed using DIALS and AIMLESS in CCP4i2 (Project et al. 1994). Structure was refined using Phenix (Adams et al. 2010).

Crystal	E260A low sodium
Wavelength	0.9686Å
Beamline	I24/ DLS
Oscillation range	0.10°
Exposure time	0.01 s
Transmission	10.09 %
Beam size	7x7 μm
Space group	C212
Resolution	69.8 - 3.2 Å(3.315 - 3.2)
CC_{50}	0.950 (0.471)
Cell dimensions	142.376 55.0133 98.841 90 101.344 90
Number of measured reflections	137481 (13805)
Number of unique reflections	12645 (1234)
Completeness	97.27 % (97.36)
Multiplicity	10.9 (11.0)
$I/\sigma I$	2.3 (0.6)
R_{merge}	0.5498 (2.417))
Refinement	
Number of reflections	12305 (1218)
R-factor	0.2320
R-free	0.2723
R.m.s.d from ideal values	
Number of non-hydrogen atoms	4870
Bond lengths	0.003
Bond angles	0.54
Ramachandran favored	95.21 %
Ramachandran allowed	4.63 %
Ramachandran outliers	0.16 %
Rotamer outliers	0.21 %

core domains show a RMSD of 0.318 Å and the panel domains 0.604 Å which agrees with the observation from before that the panel domain is always more flexible than the core domain. In the core domain, the intracellular part of TM10 of monomer A has moved closer to TM4b together with IL1 which has moved more into the substrate binding pocket (see figure 4.13). On the extracellular site, EL5/TM9b have moved into the same direction towards the panel domain in monomer A. The freed space on the extracellular site is taken up by TM8. IL1 has moved more into the substrate binding pocket making it slightly smaller.

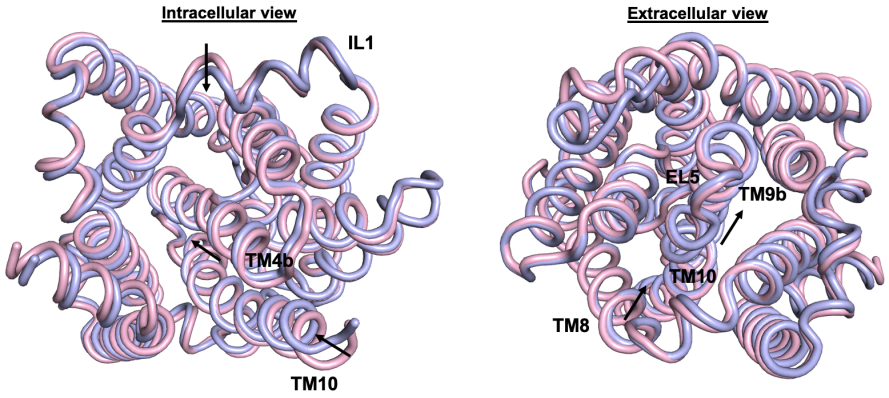


Figure 4.13: Superimposition of the two different monomers of the E260A low sodium structure: Intracellular view (left) and extracellular view (right) of the two monomers (monomer A – light blue, monomer B – pink) with parts showing differences highlighted. The arrows indicate the direction of movement of monomer A.

Overall, it seems that the intracellular cavity is smaller in monomer A in comparison to monomer B which is caused by a movement of the panel and core domain towards each other. The main helices which are responsible for the closure of the cavity are TM2/7 in the panel domain and TM4b in the core domain while the other helices in the core domain (TM8/9a/10) following the movement. The distances of the C-alpha atoms from residues lining the cavity on TM4b (core domain) and TM2/7 (panel domain) were measured and analyzed (see figure 4.14). The distances of the C-alpha atoms (backbone) were chosen as at this resolution the electron density does not allow to position the side chains unambiguously. Residues which were chosen are colored in green and red in the protein structure. Comparing the distances of Ala113 (TM4b/ core domain) to Met51 (TM2/ panel domain) and Pro196 (TM7/ panel domain) shows that

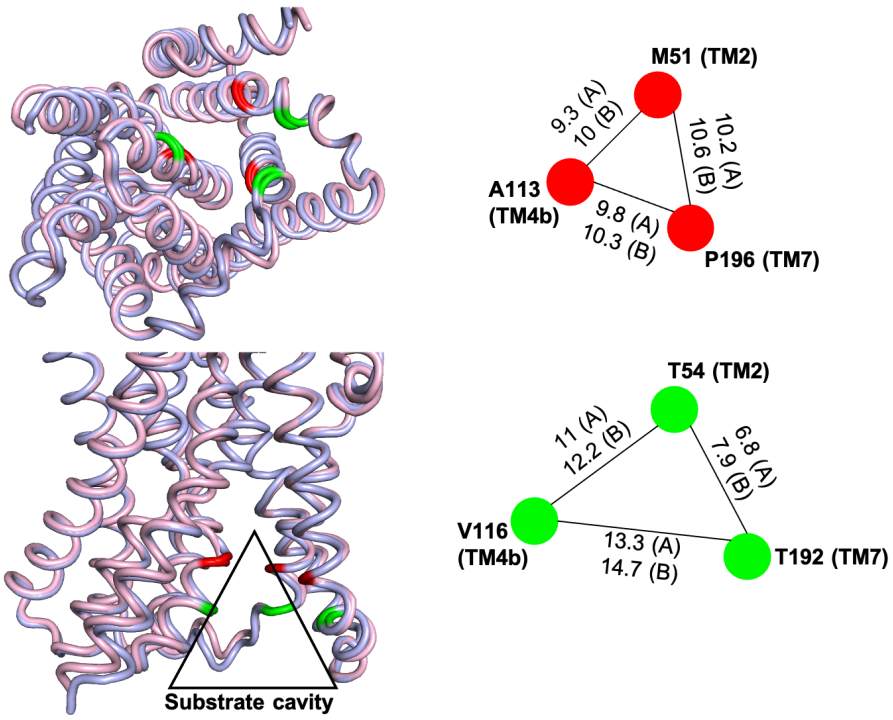


Figure 4.14: Differences in substrate binding pocket size between the two monomers of the E260A structure: Distance of six residues were measured to show the differences in the substrate binding pocket. Red color shows residues lining the substrate binding pocket closer to the core of the protein and green refers to residues closer to the cytosol. Substrate binding pocket is indicated as a triangle (bottom left). Distances (Å) between the residues are shown on the right side. Residue numbers and corresponding helices shown in bold. Distances are shown between the colored circles and the corresponding monomer is indicated in brackets.

the cavity in monomer A is slightly smaller than in monomer B as all distances are slightly smaller in monomer A. Ala113 is located almost at the center of the protein close to the crossover point of TM4 and TM9 and therefore these residues only show a small decrease in distance as the cavity is already narrower at this point as depicted in figure 4.14 (lower left image). To show the closure closer to the intracellular site, residues one turn down in the helices were included in the analysis. For TM2 and 4b, these residues were three residues away (Val116/TM4b and Thr54/TM2). For TM7, Thr192 was used as the helix is slightly bended and Asp193 points towards the outside (into the membrane plane). In agreement, the distances for monomer A are overall smaller compared to B with larger differences in the distance as the cavity is bigger at the intracellular site.

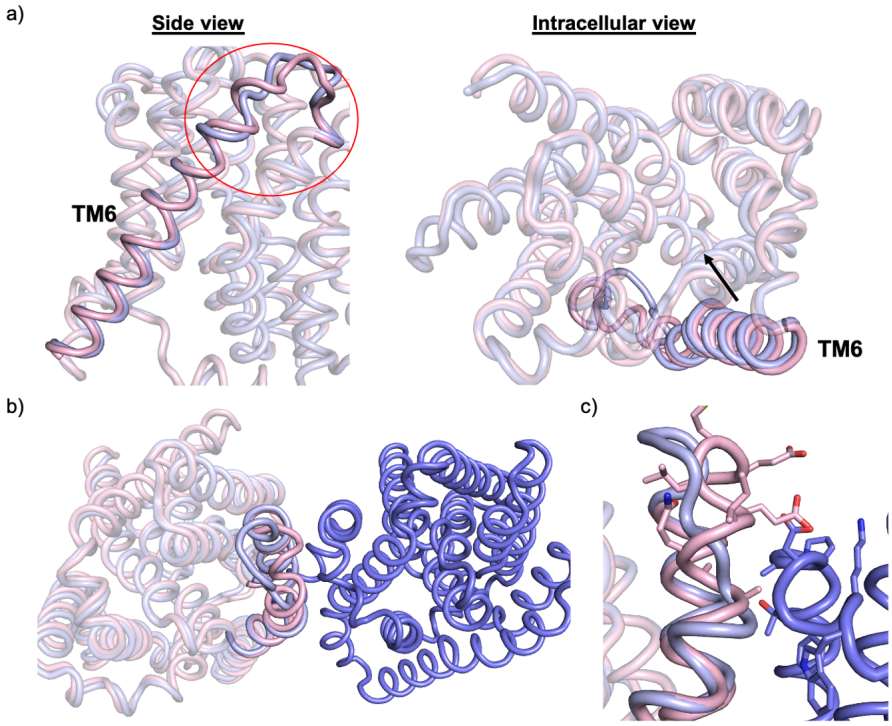


Figure 4.15: Differences in TM6 in the E260A structure: a) Side view of TM6 highlighting the difference between monomer A (light blue) and monomer B (pink). Extracellular part of TM6 in monomer A has unwound and became unstructured. Intracellular view of TM6 shows a movement towards the cavity with arrow indicating the direction of movement of monomer A. b) Extracellular view of the adjacent monomer (violet) of monomer A in the crystal. Monomer B is overlaid to show the difference in EL3/ TM6. c) Close-up view of the extracellular part of TM6 and EL3 of monomer A and EL1 of the adjacent monomer in the crystal (violet). Monomer A (pink) is overlaid and residues are shown as sticks to show potential clashes

While not involved in closure of the cavity, the intracellular part of TM6 has also moved towards the core domain while the extracellular part has unfolded earlier in monomer A on the extracellular site (see figure 4.15a) and became unstructured. EL3 connects the panel and core domain with each other and has moved towards the core domain and is elongated as it additionally contains the first two extracellular turns of TM6 . As seen in figure 15b, there is monomer from the next asymmetric unit close to EL3 of monomer A which could cause the difference in EL3. If TM6 would show the same secondary structure than in monomer B, residues on TM6 would clash with residues on EL1 on the adjacent molecule, mainly Glu156 and Asp31. Therefore, the observed structural difference could be a crystallographic artifact due to the crystal packaging. Nev-

ertheless, EL1 is known to be quite flexible, so it could potentially move around to accommodate for the elongated TM6 and therefore it cannot be precisely said what caused the differences, only that the extracellular part of TM6 is flexible and possess the ability to switch between structured and unstructured. Overall, the superimposition of the two monomers suggests a rocker-switch mechanism which is common for the MFS superfamily in which both domains move towards/apart each other to facilitate the opening and closing of the ligand binding site. But the MFS superfamily has two homologous domains unlike $ASBT_{NM}$. Therefore, it is more likely that $ASBT_{NM}$ follows a rocker-switch or elevator mechanism as proposed for NapA, NhaA, and $ASBT_{YF}$ in which one domain moves around the other domain. Therefore, it was decided to perform two separate superimpositions in which either the panel or core domain were kept fixed to investigate the movement of the other domain in regards to the fixed domain (see figure 4.16).

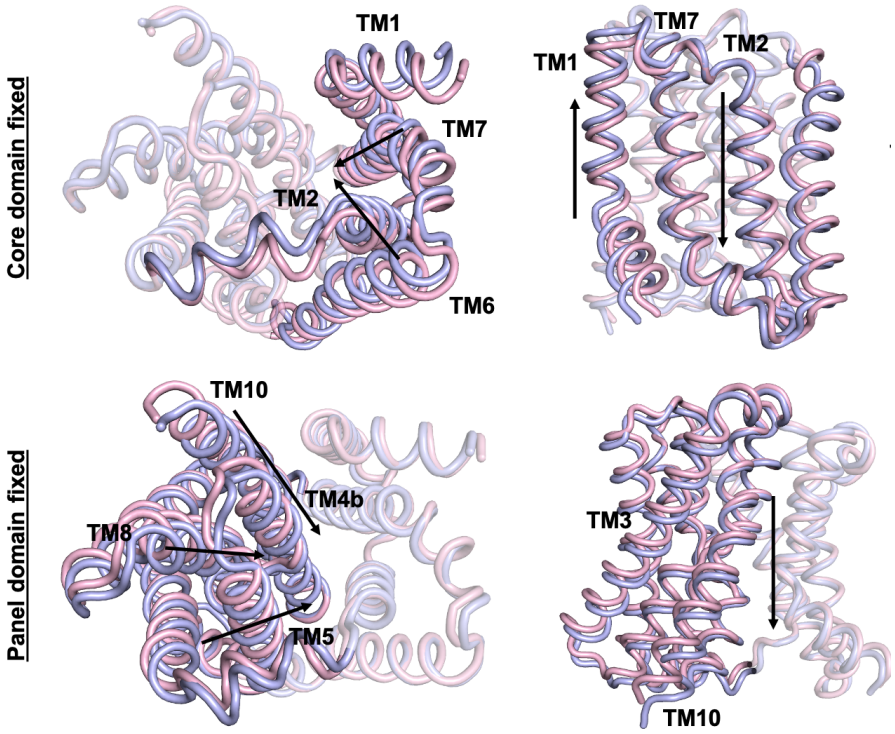


Figure 4.16: Superimposition of either panel or core domain of the two monomers in the E260A structure: Top images show the movement of the panel domain if the core domain kept fixed. Direction of movements of monomer A in regards to monomer B are indicated with arrows. Bottom images show the movement of the core domain if the panel domain is kept fixed.

Figure 4.16 shows the two superimpositions. If the core domain is overlaid and kept fixed, the panel domain shows an overall larger movement towards the substrate cavity compared to the movement seen before. From the side, this movement is accompanied by a downwards (intracellular) directed motion of TM2, 6, and 7 and an opposing upwards motion of TM1 overall showing a small rotation of the panel domain. When the panel domain was kept fixed, the core domain shows a movement towards the panel domain making the intracellular cavity smaller. Additionally, the core domain shows a downwards movement similar to the panel domain. To sum up, all superimpositions showed a closure of the intracellular cavity in monomer A which needs to occur to switch between the inward and outward facing conformation and vice versa. Interestingly, the two latter superimpositions show a downward movement of either core or panel in monomer A which is slightly contradictory as you would assume an upwards movement toward the extracellular site to occur allowing the opening of the cavity at the extracellular site. It needs to be mentioned here that the superimpositions are subjective and probably do not reflect the actual movement (see discussion for more detail).

Sodium binding in the E260A low sodium structure

The crystallization condition did not contain any sodium ions, so the final sodium concentration was estimated to be around 3.5 mM. When the electron density was investigated, there was no density in Na1 indicating that Na1 is not occupied by sodium (see figure 4.17). In Na2, there was some weak positive density visible in Na2. It was difficult to say whether sodium is in there or not, but when sodium was modeled in, only the sodium in monomer B came up as a B-factor outlier. As the density was much less than for monomer A, it would seem at best to be only partially occupied. There were no clear sodium-related differences visible between the two monomers and monomer A showed the main differences with one sodium bound, which is unfortunately contradictory to the fact that monomer A showed the main difference.

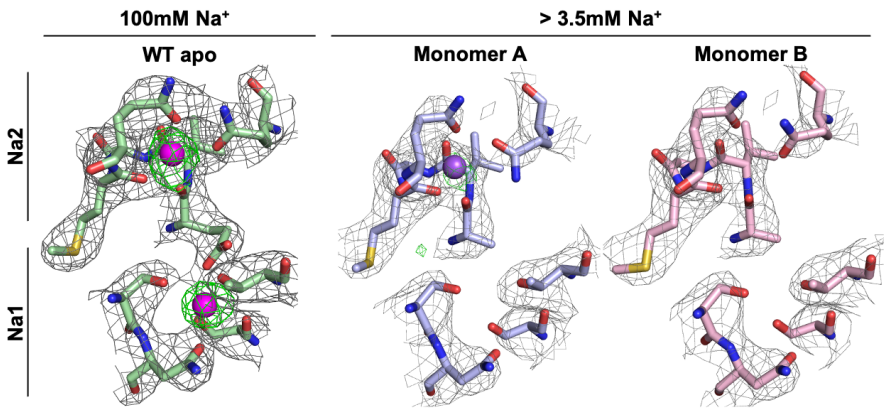


Figure 4.17: Comparison of sodium binding in the WT apo sodium structure and the E260A structure: Sodium binding sites are shown as stick representation (WT apo sodium- green, Monomer A- light blue, Monomer B- pink). 2mFo-DFc map is shown at 1 sigma (grey mesh). The mFo-DFc map was calculated with no sodium ions present during the refinement and the resulting positive density is shown at +3 sigma (green mesh). Final sodium concentrations are indicated at the top.

4.2 B-factor analysis of a high resolution WT inhibitor structure reveals flexible areas within the protein

Almost at the end of the PhD, a high-resolution dataset of the wild type protein was obtained when trying to crystallize the protein with an inhibitor. Unfortunately, the inhibitor was not visible in the electron density, but overall, the resolution was better than before and it was decided to use this dataset for a B-factor analysis to further investigate the flexible areas in the protein and to support the findings from the previously described structures regarding flexible areas within $ASBT_{NM}$. A single crystal grew in an initial MemMeso screen with 0.1 M Sodium chloride 0.1 M HEPES 7, 30 % v/v PEG 300, and 0.1 M Calcium chloride dihydrate in the condition and two data sets were collected from the same crystal. The resolution was cutoff at 2.1 Å and the structure was refined to a final R-factor of 0.19 and a R-free of 0.23. Overall, the inhibitor structure overlays well with the WT apo sodium structure with an RMSD 0.345 Å and the only major difference was found in TM6 and EL3 similar to the E260A structure

where a part of the extracellular part of TM6 has unfolded and has moved towards the core domain (see figure 4.18). Additionally, the extracellular part has slightly tilted away from the protein towards the membrane. On the intracellular part, the helix has moved closer to the core domain. Again, the crystal packing is quite tight with TM1 of an adjacent monomer being in close proximity with EL3 which could be the reason for the difference.

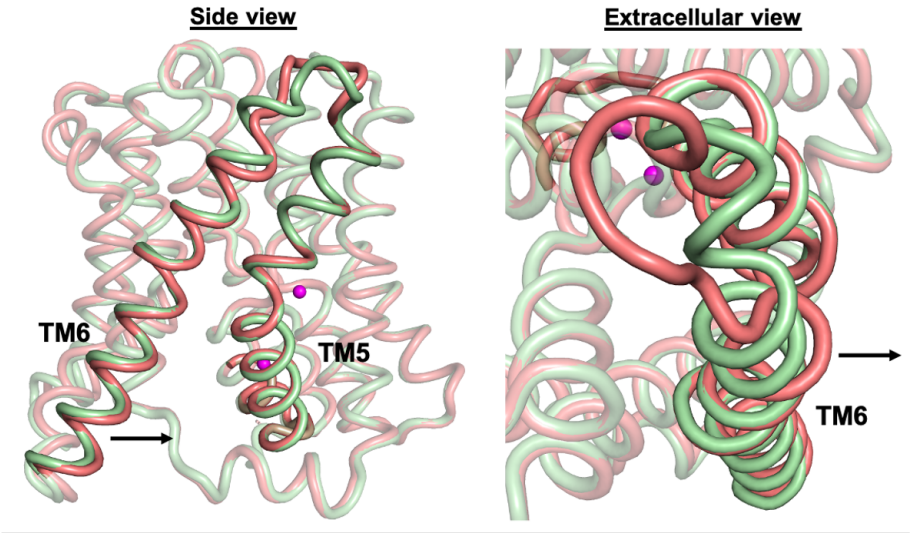


Figure 4.18: Differences in TM6 in the WT inhibitor structure and the WT apo sodium structure: Left image shows a side view of TM6 highlighting the differences between the inhibitor structure (red) and the WT apo structure (green). The extracellular part of TM6 has unwound and became unstructured. The intracellular part of TM6 shows a movement towards the core domain with the arrow indicating the direction of movement. Right image shows an extracellular view of TM6. Arrow indicates the outward movement of TM6 in the inhibitor structure. Sodium ions are shown as purple spheres.

B-factors are indicators for the movement of atoms with high B-factors indicating more flexible atoms and vice versa. Therefore, the analysis of the B-factors can give useful information about flexible areas and can support the findings from before. In agreement with the findings before, the panel domain shows higher B-factors in general as seen by the green to yellow-colored areas in comparison to the core domain (see figure 4.19). As discussed in the WT apo structure, the intracellular halves of TM 6 and 7 including IL3 show the highest B-factors in the panel domain indicating their flexible nature. The core domain shows a quite rigid center which is made out of TM 4/5 and 9/10 and includes the two sodium binding sites which are both occupied in the structure. TM3 and 8 in the core

domain which are connected to the panel domain show a slightly higher flexibility if compared to the other helices in the core domain. As they are connecting the panel domain it is likely that they move during the transition between the outward and inward conformation. Overall, the B-factor analysis results fit with the results from before showing more flexibility in the panel domain and the helices connecting the panel and the core domain (see core and panel domain analysis of the wild type apo structure with sodium (subsection 4.1.1) and the E260A structure (subsection 4.1.3) for more information).

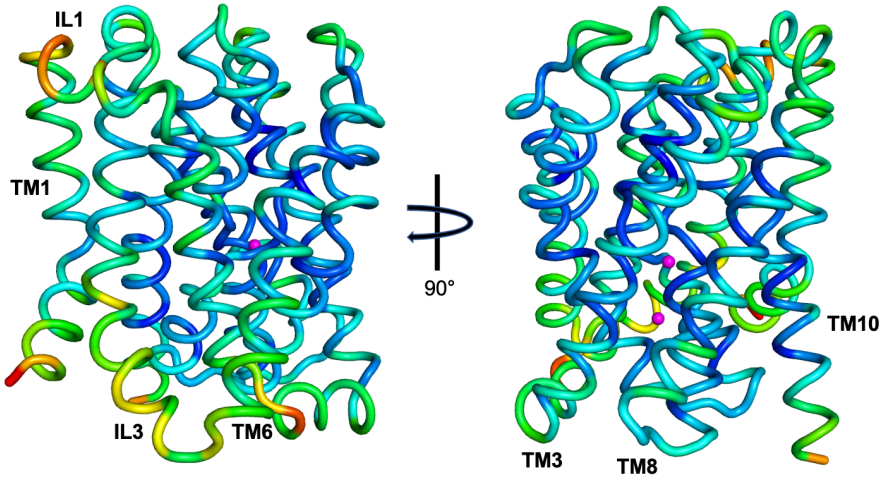


Figure 4.19: B-factor analysis of the WT inhibitor structure: Ribbon representation of the WT inhibitor structure colored according to the B-factors (blue – low B-factors to red – high B-factors). Areas with higher B-factors are labelled.

Table 4.4: Data collection and refinement statistics for the $ASBT_{NM}$ wild type inhibitor structure: Crystals were grown in 0.1 M Sodium chloride, 0.1 M HEPES 7, 30 % v/v PEG 300 and 0.1 M Calcium chloride dihydrate. Data was processed using DIALS and AIMLESS in CCP4i2 (Project et al. 1994). Structure was refined using Phenix (Adams et al. 2010).

Crystal	WT with inhibitor
Wavelength	0.9999Å
Beamline	I24/ DLS
Oscillation range	0.10°
Exposure time	0.01 s
Transmission	14.30 %
Beam size	7x7 μm
Space group	$P2_12_12_1$
Resolution	42.94 - 2.1 Å(2.175 - 2.1)
CC_{50}	0.91 (0.605)
Cell dimensions	49.4791 80.6094 86.4526 90 90 90
Number of measured reflections	208154 (20889)
Number of unique reflections	20810 (2033)
Completeness	99.56 % (99.22)
Multiplicity	10.0 (10.2)
$I/\sigma I$	17.91 (1.53)
R_{merge}	0.4187 (1.062)
Refinement	
Number of reflections	20732 (2027)
R-factor	0.1938
R-free	0.2303
R.m.s.d from ideal values	
Number of non-hydrogen atoms	2431
Bond lengths	0.011
Bond angles	1.04
Ramachandran favored	99.02 %
Ramachandran allowed	0.98 %
Ramachandran outliers	0.00 %
Rotamer outliers	0.42 %

4.3 Two low resolution wild type structures indicate a different conformation of TM1

As mentioned before, a low resolution wild type dataset was collected in the MemMeso D2 condition which does not contain any sodium ions (0.1 M Lithium sulfate, 0.1 M MES pH 6.5, 30% v/v PEG 400). Unfortunately, several attempts of optimizing the crystal condition by changing the pH, salt or precipitant concentration did not yield in better diffracting crystals. Therefore, the better resolution datasets obtained from addition of either sodium or taurocholate were discussed before. Even though the resolution is low, it was decided to include this data set in here, as it shows a different feature similar to another dataset which will be discussed a bit later.

The dataset was cut off at a resolution of 3.9 Å. After the Molecular replacement, the electron density for the two monomers was investigated and interestingly, it showed a difference in the electron density for the intracellular halve of TM1 for one monomer. As TM1 is flexible, Molecular replacement was always performed with a model missing the first 15 residues on TM1. Figure 4.20a shows the density for the missing first half of TM1 after one round of refinement performed in Phenix using grouped B-factors and XYZ (reciprocal) refinement. Additional restraints were applied such as NCS, secondary structure restraints and the WT apo structure as a reference model.

Table 4.5: Data collection and refinement statistics for the $ASBT_{NM}$ wild type apo no sodium low resolution structure: Crystals were grown in 0.1 M Lithium sulfate, 0.1 M MES pH 6.5 and 30% v/v PEG 400. Data was processed using DIALS and AIMLESS in CCP4i2 (Project et al. 1994). Structure was refined using Phenix (Adams et al. 2010).

Crystal	WT apo no sodium low resolution
Wavelength	0.9686Å
Beamline	I24/ DLS
Oscillation range	0.15°
Exposure time	0.050 s
Transmission	14.50 %
Beam size	20x20 μm
Space group	$P12_11$
Resolution	69.73 - 3.9Å(4.04 - 3.9)
CC_{50}	0.952 (0.699)
Cell dimensions	72.8047 86.7058 80.4459 90 106.712 90
Number of measured reflections	28485 (8192))
Number of unique reflections	8866 (2508)
Completeness	98.19 % (99.22)
Multiplicity	3.2 (3.3)
$I/\sigma I$	2.0 (0.8)
R_{merge}	0.324 (1.051)
Refinement	
Number of reflections	8718 (872))
R-factor	0.3165
R-free	0.3745
R.m.s.d from ideal values	
Number of non-hydrogen atoms	4489
Bond lengths	0.004
Bond angles	0.857
Ramachandran favored	97.36 %
Ramachandran allowed	2.31 %
Ramachandran outliers	0.33 %
Rotamer outliers	1.72 %

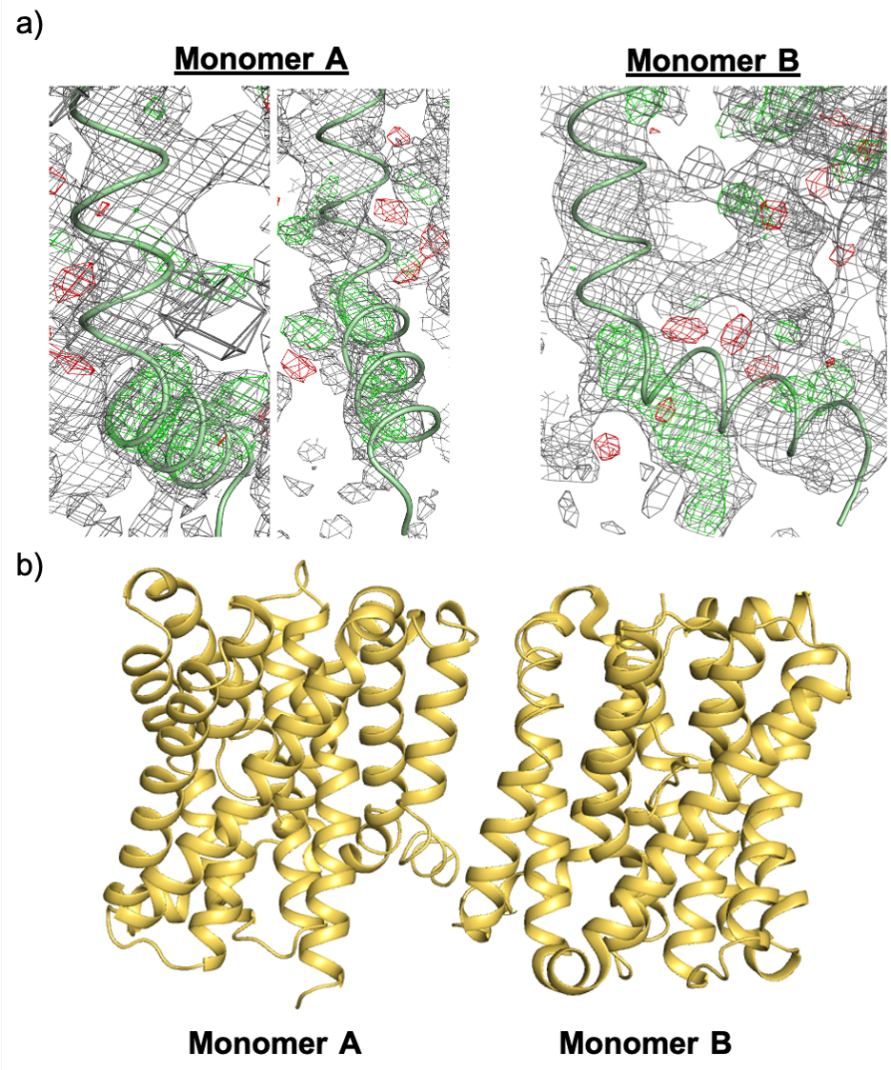


Figure 4.20: Electron density comparison for TM1 in the low resolution WT apo structure: a) Electron density maps after the molecular replacement lacking the first 15 residues of TM1. 2mFo-DFc map is shown at 1 sigma (grey mesh). The mFo-DFc is shown at +3 sigma (green mesh). TM1 of the WT apo with sodium is shown in green. TM1 fits nicely into the electron density of monomer A (left) but not into the electron density of monomer B. b) Asymmetric unit of the low resolution WT apo structure showing the labelling of the two monomers.

For Monomer A, TM1 of the WT apo structure fits nicely into the electron density clearly showing the ''normal" bending of TM1. However, for monomer B the electron density and TM1 do not overlay well indicating a change in the intracellular part of TM1. It looks like that TM1 is straighter in monomer B. Interestingly, in $ASBT_{YF}$ TM1 shows also a straighter conformation and there-

fore it is possible that TM1 can adapt different conformations which might be important for transport activity. To further investigate the differences in TM1, the structure was further refined and residues for TM1 were built into the density. Overall, the two monomers in the asymmetric unit are quite similar to the WT apo structure with a RMSD of 0.233 Å (monomer A) and 0.244 Å for monomer B (see figure 4.20 for crystal packing and monomer labelling). As mentioned before, the WT apo structure was used as a reference model during refinement as the resolution is quite low. Therefore, it is not surprising that the structures resemble each other more in comparison to other structures as the resolution does not allow to differentiate between small conformational changes and the low resolution structure is forced by restraints from the WT apo structure. Because of that, only the difference in TM1 will be shortly discussed for this structure. Figure 4.21 shows the movement of TM1 of monomer B in comparison to the WT apo structure. TM1 has moved towards TM10 (core domain) making the binding pocket smaller on the intracellular site. This movement is performed by a "straightening" act, overall decreasing the degree of tilt between the upper and lower part of TM1. The asymmetric unit shows that both monomers are inverted and TM1 of the other monomer is in close contact with TM1 of monomer B (see figure 4.21b). The superimposition with the WT apo structure shows that in this packing the bent, intracellular half of TM1 would clash with the extracellular part of TM1 of monomer A. Therefore, the crystal packing might be responsible for the different conformation of TM1. Nevertheless, this supports the idea that TM1 is quite flexible in $ASBT_{NM}$.

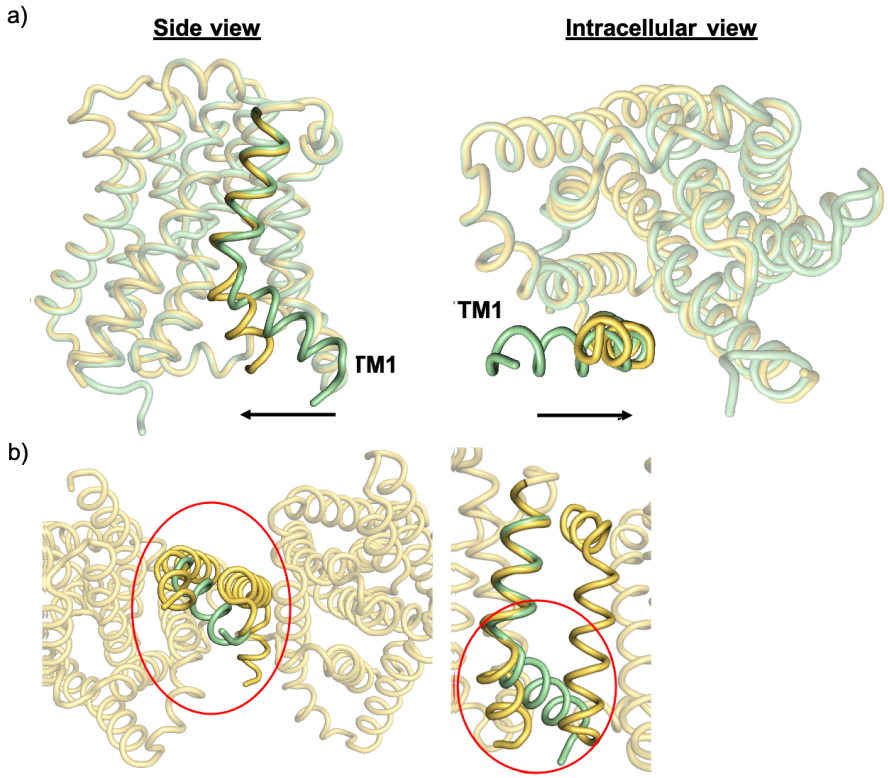


Figure 4.21: Differences in the intracellular part of TM1 in the low resolution WT apo structure: a) Superimposition of low resolution WT apo structure (yellow) and WT apo with sodium (green). TM1 moves towards the core domain as seen from the side and from the intracellular site. Arrow indicates direction of movement. b) Crystal packing shows that TM1s of monomer B and the adjacent molecule are in close contact. When overlaid with TM1 of the WT apo with sodium structure (green), TM1 would clash with the adjacent monomer.

Later on, a data set was collected from a crystal grown in a different condition containing 0.1 M Sodium chloride, 0.1 M MES pH 6, 30% v/v PEG 400, and 0.1 M Calcium chloride dihydrate. Interestingly, even the crystallization condition is different the space group is similar with the cell parameters only showing small changes. The resolution and statistics are slightly better for this data set maybe indicating that some component in the condition supporting this crystal packing. The resolution was cut off at 3.5 Å. After refinement, it was clear that the structure shows the same conformation of TM1 for one monomer in the asymmetric unit while the other monomer shows a similar bending as normal. As both

structures were obtained from different crystallization conditions, it is unlikely that the different conformation of TM1 is caused by any chemical compound in the condition and therefore, both structures are mentioned in this thesis.

Overall, monomer A is more similar to the WT apo structure with a RMSD of 0.363 Å showing only minor differences which are in similar areas as described before for other structures and therefore will not be discussed here. For convenience, the nomenclature for the monomers are adopted from the structure before (see figure 4.21 for asymmetric unit and monomer labelling). Monomer B differs more from the WT apo structure with a RMSD of 0.489 with the most striking difference being TM1. If the whole monomers are superimposed (see figure 4.22a), the upper half of TM1 in the calcium structure moves slightly downwards, while the intracellular half is less bent and points more towards TM10. On the intracellular side, TM1 moves toward the substrate binding pocket. When investigated closer, the kink of TM1 seems to be caused by a hydrogen bond between Trp18 and Thr14 in the WT apo structure. It was hypothesized that the breakage of the hydrogen bond might be responsible for the straightening of TM1 but interestingly the distance between the two residues does not change in both structures contradicting the straightening movement which should result in different distances between the residues in both structures. Next, only TM1 of each structure was overlaid and it showed that they are more similar than thought before. Instead of the proposed straightening act which was assumed because of the superimposition before, the difference is caused by a movement of TM1 around the kink area which acts like a hinge region. The upper part of TM1 moves towards the membrane plane (away from the protein) while the intracellular half moves towards the core domain (see figure 4.22d).

Table 4.6: Data collection and refinement statistics for the $ASBT_{NM}$ wild type apo calcium structure: Crystals were grown in 0.1 M Sodium chloride, 0.1 M, MES pH 6, 30% v/v PEG 400 and 0.1 M Calcium chloride dihydrate. Data was processed using DIALS and AIMLESS in CCP4i2 (Project et al. 1994). Structure was refined using Phenix (Adams et al. 2010).

Crystal	WT apo calcium
Wavelength	0.9795Å
Beamline	I24/ DLS
Oscillation range	0.15°
Exposure time	0.050 s
Transmission	15.04 %
Beam size	20x20 μm
Space group	$P12_11$
Resolution	70.42 - 3.5 Å(3.625 - 3.5)
CC_{50}	0.914 (0.892)
Cell dimensions	73.4556 87.2386 81.2962 90 106.523 90
Number of measured reflections	23387 (5493))
Number of unique reflections	12515 (2944)
Completeness	99.26 % (98.64)
Multiplicity	1.9 (1.9)
$I/\sigma I$	2.3 (0.9)
R_{merge}	0.158 (0.422)
Refinement	
Number of reflections	12469 (1217)
R-factor	0.2744
R-free	0.3113
R.m.s.d from ideal values	
Number of non-hydrogen atoms	4537
Bond lengths	0.004
Bond angles	0.83
Ramachandran favored	98.37 %
Ramachandran allowed	1.47 %
Ramachandran outliers	0.16 %
Rotamer outliers	0 %

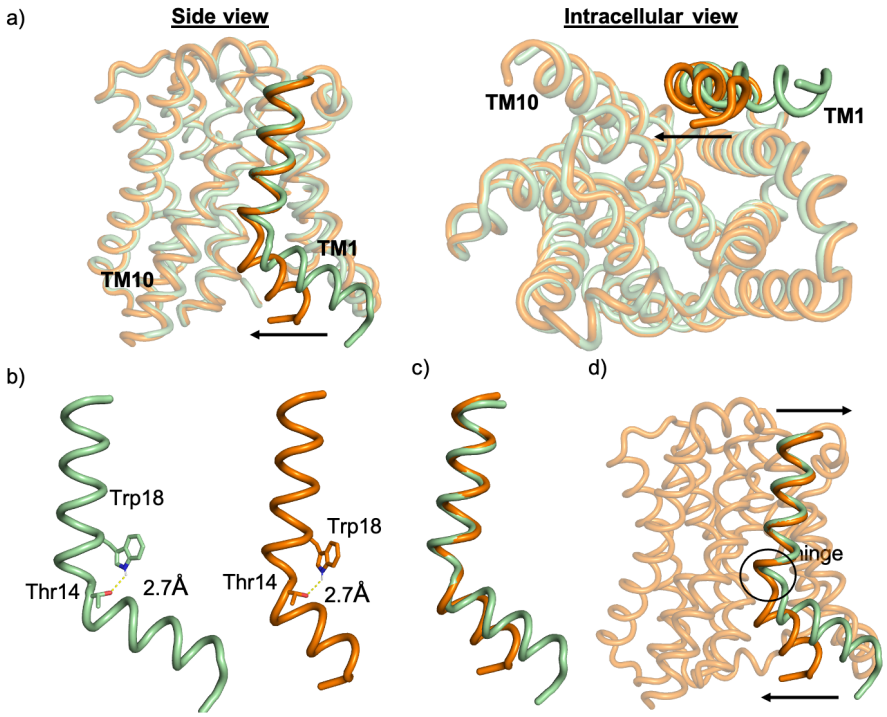


Figure 4.22: TM1 analysis in the WT calcium structure: a) Superimposition of monomer B of the WT calcium structure (orange) with the WT apo sodium structure (green). TM1 is highlighted and movement between structures is indicated by an arrow. b) Distance of Trp18 and Thr14 on TM1 in the WT calcium and the WT apo sodium structure c) Separate superimposition of TM1 from both structures. Superimposition was performed solely on TM1. d) Image illustrating the asymmetric movement of TM1. The movement direction of TM1 is indicated by arrows. The hinge region is shown as a circle.

To check if this movement is only present in TM1 or can be seen in the whole panel domain, the differences in the panel domain were next investigated. Contradictory to TM1, the panel domain shows a movement towards the membrane away from the protein on the intracellular side by TM6/7 and IL3 (see figure 4.23). As both helices are not directly connected to TM1, TM2 was separately investigated, but did not show any differences. EL1 seems to compensate for the movement of TM1 and does not transfer it to TM2 (see figure 4.23).

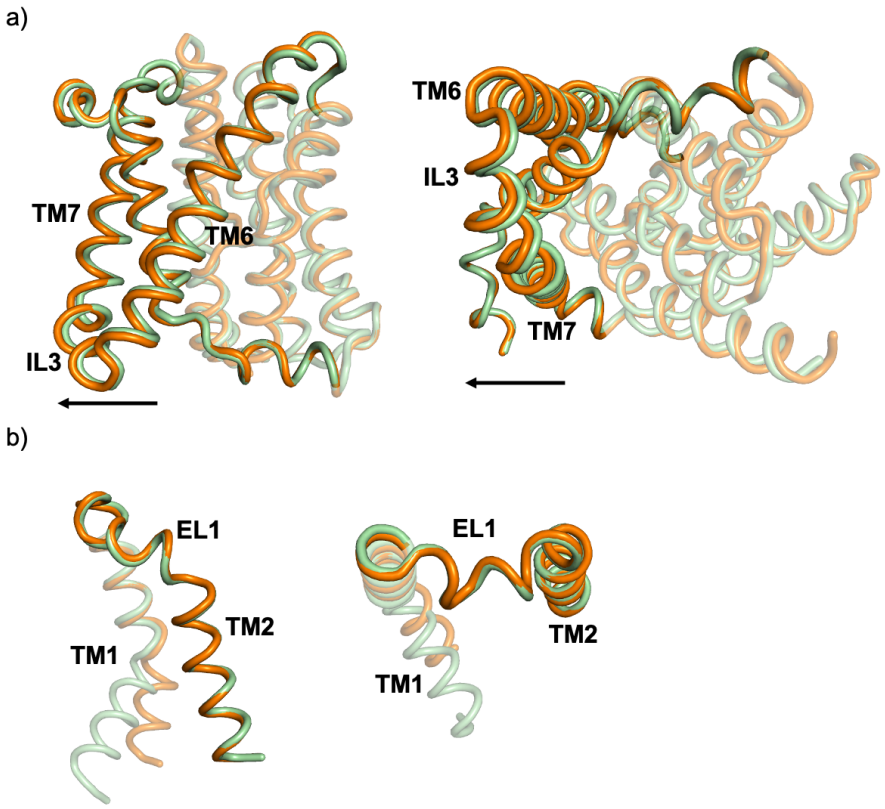


Figure 4.23: Comparison of the panel domain of the WT calcium structure with the WT apo sodium structure: a) Side view of the panel domain highlighting the differences between the WT calcium structure (orange) and the WT apo sodium structure (green). Helices showing differences are labelled and arrows indicate the movement direction of the WT calcium structure. b) Differences in EL1 and TM2 of both structures.

4.4 Discussion

In this chapter, different new structures of $ASBT_{NM}$ wild type and an E260A mutant were introduced and discussed. Unfortunately, the structures only showed the protein in the inward-facing conformation but nevertheless some differences were seen and can be used for the better understanding of the transport mechanism by $ASBT_{NM}$.

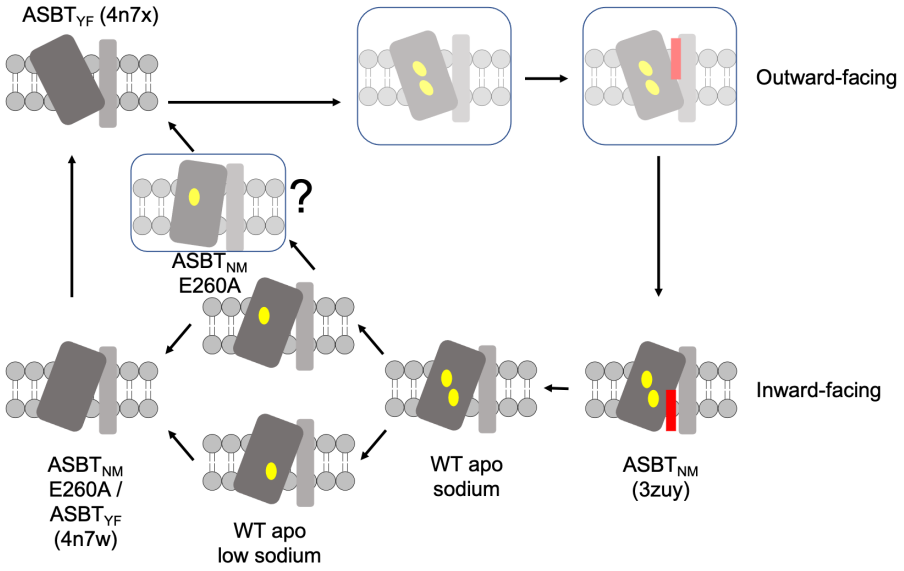


Figure 4.24: Schematic transport cycle of $ASBT_{NM}$ with newly elucidated conformations: Different conformations of $ASBT_{NM}$ are shown which occur during one complete transport cycle. The transporter is shown in grey with the core domain in dark grey and the panel domain in light grey. The two sodium ions are depicted as yellow spheres and the taurocholate is shown as a red bar. Conformations for which structural information are known are labelled with the name and the PDB code in brackets. Conformations which are unknown are transparent. Structures which were solved in this thesis are labeled with their names as they were introduced in this chapter. The E260A conformation is in transparent and marked with a question mark as it is not sure if the conformation occurs naturally.

Coming back to the figure of the proposed transport cycle of $ASBT_{NM}$ some new states can be added to the transport cycle by structures elucidated in this thesis (see figure 4.24). An apo sodium state can be added which represents the transporter after the substrate was released from the substrate-bound state (3zuy). The apo state showed a slightly smaller substrate cavity which probably

results from the taurocholate release which is quite a big substrate. Interestingly, the apo state still showed the kink in TM1 for $ASBT_{NM}$ indicating that the kink is not caused by taurocholate binding. $ASBT_{YF}$ does not show the kink in TM1 in the inward-facing apo conformation and therefore the substrate or sodium release could have been responsible for the straightening of TM1. TM1 is not well conserved between $ASBT_{NM}$ and $ASBT_{YF}$ and Thr14 is not conserved in $ASBT_{YF}$ which interacts with Trp18 (conserved in $ASBT_{YF}$) in $ASBT_{NM}$ and seems to induce/ support the kink. An easy experiment to verify the hypothesis that Thr14 and Trp18 are responsible for the kink in $ASBT_{NM}$ would be to design a Thr14Ala mutation and obtain structural data from it. Overall, from the structural differences in TM1 between $ASBT_{NM}$ and $ASBT_{YF}$ and from the low conservation, it can be hypothesized that TM1 is probably not functionally important for transport.

Next, a different apo structure was obtained which showed some differences in sodium binding and the structures can be included in the transport model. The two structures showed some lower occupancy in either Na1 or Na2 and therefore it is likely that in these two monomers the sodium ions started to unbind from their binding sites. Both structures did not show major changes in the area surrounding the sodium binding sites indicating that sodium release does not induce major conformational changes in the protein. Unfortunately, at this resolution it is difficult to make any statement about movements of the side chains in the sodium binding sites and some higher resolution data would be needed to comment on differences in side chain locations. The E260A structure (one monomer) supported this idea and showed no structural changes even though Na1 was empty and Na2 only showed really weak density for Na2. Overall, this data suggests that the transporter can have both sodium sites singly occupied but no conclusion can be made about the order in which the sodium ions are released. Nevertheless, the data indicates that sodium release from Na1 does not trigger an unwinding of TM4b as seen in $ASBT_{YF}$ (see figure 4.25). Two residues of Na1 are located on TM4b and therefore it was assumed that sodium release in Na1 is responsible for the unwinding. Na2 is not directly connected to TM4b and therefore it is unlikely that release of sodium in Na2 causes the unwinding.

Overall, the data suggests that the change in TM4b does not occur in $ASBT_{NM}$ and indicates another structural difference between $ASBT_{NM}$ and $ASBT_{YF}$. It needs to be mentioned here that a citrate molecule was found to bind close to TM4b in the inward-facing conformation of $ASBT_{YF}$ which came from the crystallization buffer and could be responsible for the unwinding of TM4b. To verify that $ASBT_{NM}$ does not show the unwinding, a complete sodium free $ASBT_{NM}$ structure would be useful. One possibility would be to try to purify the protein in a sodium-free condition. For that purpose, sodium chloride could be substituted by potassium chloride in the purification buffer. Nevertheless, it is known from experiments (data not shown) that the protein is really unstable in buffer without sodium and therefore some additives might need to be added to the buffer to stabilize the protein. The CPM assay could be used to identify buffer conditions in which the protein is more stable. Glycerol would be a good first try as it is known to stabilize proteins (Vagenende et al. 2009).

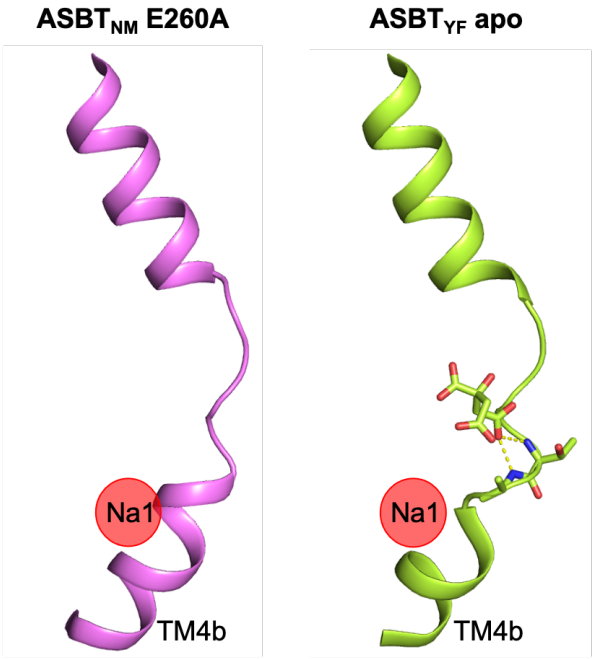


Figure 4.25: Differences in TM4 between the E260A structure and the $ASBT_{YF}$ inward-facing structure: TM4 of the E260A structure is shown in pink (left) and TM4 of $ASBT_{YF}$ in green (right). The location of the Na1 site is shown as a red sphere as known from the $ASBT_{NM}$ wild type structure (3zuy). Na1 is not occupied in the E260A structure and in the $ASBT_{YF}$ inward-facing structure. The bound citrate molecule is shown as sticks.

One monomer in the asymmetric unit of the E260A structure showed some

more conformational changes with the substrate binding cavity being smaller and the location of the two domains to each other being slightly altered. As Zhou et al proposed for $ASBT_{YF}$, the transition between inward- and outward-facing conformation occurs because of a rotation movement of the core domain together with an upwards movement of the core domain (X. Zhou et al. 2014). Figure 4.26 shows a superimposition of the two E260A monomers with the outward-facing conformation of $ASBT_{YF}$. The superimposition was performed on the panel domain which allows the movement of the core domain. The superimposition reveals that the two E260A monomers show some differences on the intracellular site while the extracellular site is quite similar. The overlay of the three core domains shows the beginning of a rotation movement on the intracellular site with TM4b moving towards TM2 in the panel domain and TM10 moving slightly into the substrate binding site. Therefore, one monomer of the E260A structure might show the beginning of a transition state from the inward-facing to the outward-facing conformation. Surprisingly, the monomer showed density for sodium in the Na2 site and therefore it is not a complete apo (empty) state like the $ASBT_{YF}$ structure. It needs to be further investigated if the transporter can switch conformations with a sodium ion still bound or if the structure might be a crystallography artifact.

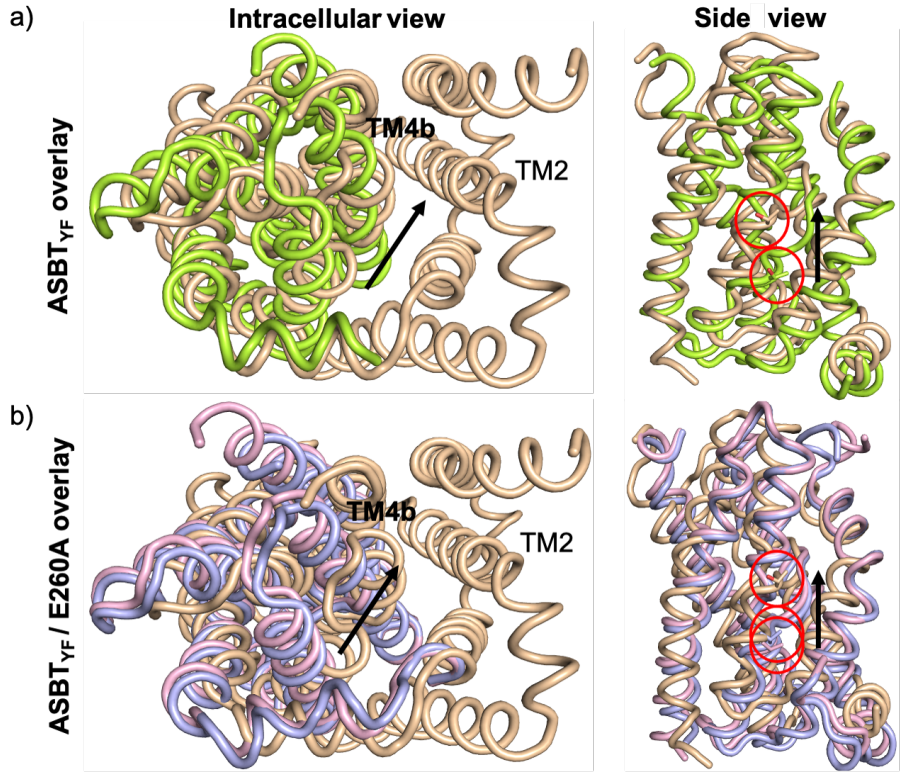


Figure 4.26: Comparison of the core domain location relative to the panel domain of the outward-facing conformation of $ASBT_{YF}$ with the inward-facing conformation of $ASBT_{YF}$ and with the E260A monomers: Superimpositions were performed on the panel domain. a) Overlay of the two $ASBT_{YF}$ structures. The outward-facing conformation of $ASBT_{YF}$ is colored in wheat. The inward-facing conformation of $ASBT_{YF}$ is colored in green. TM4b shows a movement towards the panel domain from the intracellular site. The side view shows an upwards movement of Thr112 (red circles) from the inward towards the outward-facing conformation. b) Overlay of the outward-facing conformation of the $ASBT_{YF}$ structure and the two monomers of the E260A structure. The outward-facing conformation of $ASBT_{YF}$ is colored in wheat. Monomer A of the E260A structure is colored in violet and monomer B in pink. TM4b in monomer A shows a small movement towards the panel domain in comparison to monomer B. TM4b moves towards TM2 during the transition from the inward- to the outward-facing conformation. Thr112 of the three individual structures are highlighted in red circles. Thr112 of monomer A shows a small upwards movement in comparison to monomer B. Overall, the movements in monomer A suggest the beginning of the transition from the inward- to outward-facing conformation.

Chapter 5: Functional characterization of taurocholate transport and binding

Hu et al used an in-cell transport assay to functionally characterize $ASBT_{NM}$ wild type and four different mutants (Hu et al. 2011). An in-cell assay has the advantage of the protein being inserted in the membrane in the correct orientation and a membrane consisting of natural lipids which were shown to play important roles in transport of certain transporters. Nevertheless, cells contain other transport systems which can influence the outcome of the transport assay. Especially for bile acids, there are certain transporters known, e.g AcrB (Drew et al. 2008) which pumps bile acids out of the cell as they are toxic for the cells. Additionally, in a cell-based assay the buffer components cannot be changed easily in the inside of the cell like pH or salt concentration. To further characterize $ASBT_{NM}$, a former PhD student and Postdoc has started to develop a liposome-based assay which was continued.

5.1 Continuation of a soybean lipid-based assay transport assay

5.1.1 Getting the assay working

Liposome were prepared according to the established protocol in our group (see material and methods for more information). Briefly, the soybean lipids were dissolved in 10 mM MOPS (pH 7), 5 mM $MgCl_2$, 200 mM KCl and applied to several rounds of sonication. After the sonication, the solution was spun down to remove undissolved lipids and the liposomes were applied to eight freeze-thaw cycles. Liposomes can be stored at this point until further usage. Reconstitution is performed by destabilizing the liposomes with sodium cholate and the detergent is removed by using a PD-10 desalting column which acts as a small size exclusion column separating the large proteoliposomes from free detergent. A first transport assay trial was performed with self-made liposomes and wild type protein. Empty liposomes were used as a negative control to which only buffer was added at the reconstitution step instead of protein. For the transport assay, proteoliposomes or empty liposomes were incubated with 137 mM NaCl and 666 nM valinomycin for 5 min at 37 °C. Uptake was initiated by the addition of ^{3}H-taurocholate and

the transport was performed for 5 min. Figure 5.1 shows the first transport assay which was performed. The counts were really low with only a small window of 60 CPM between the wild type containing liposomes and the empty liposomes. Trying to increase the wild type signal, a second assay was performed in which only ^{3}H-taurocholate was added to the liposomes without cold taurocholate. The wild type signal increased of about 120 CPM while the empty liposomes only about 30 CPM resulting in an overall window of about 150 CPM. These values were still really low if compared to former results and the assay was considered to be not working.

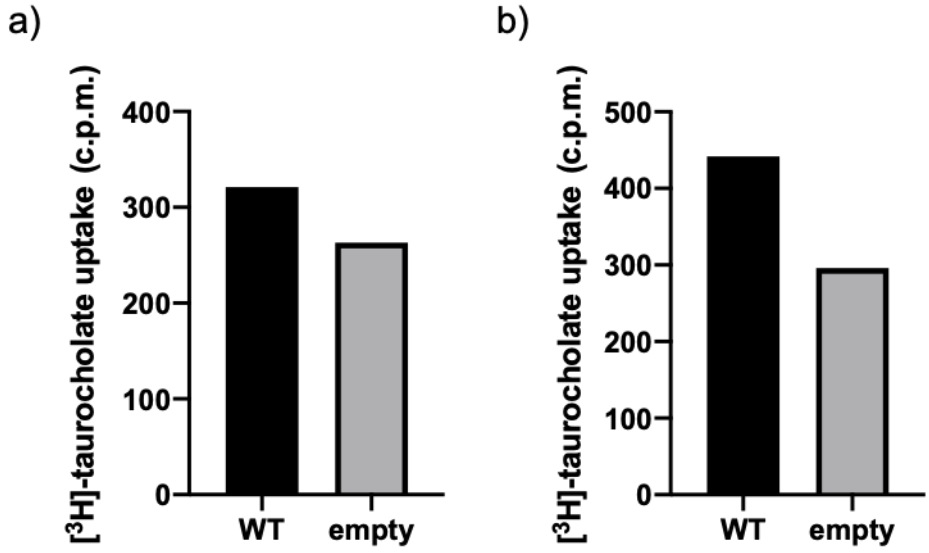


Figure 5.1: First trial of the soybean-based liposome transport assay: Internalized ^{3}H-taurocholate signal after 5 min of transport at 37 °C of wild type containing and empty liposomes. Liposomes were prepared for the first time according to the established protocol. a) A mixture of ^{3}H-taurocholate (0.209 μM) and cold taurocholate (0.8 μM) was added to initiate transport. b) Only 0.209 μM of ^{3}H-taurocholate was added to start the assay.

To find out if the purified protein or the liposomes were the problem, old liposomes which were prepared by a Postdoc in our group were thawed and the same batch of wild type protein was reconstituted. The results are shown in figure 5.2. The same wild type protein gave a signal of 7075 CPM and the empty liposomes of 2964 CPM. The wild type protein seemed to be active in the old liposomes, so it could be concluded that the problem were the liposomes. To note here, some experiments were only performed as a single experiment as the

radioactive compound is quite expensive and it was tried to reduce the excessive usage of it. The next batch of liposomes was prepared together with the help of the Postdoc who prepared the old batch of liposomes. Figure 5.2 b shows a comparison of the newly prepared liposomes against old liposomes. Again, the old liposomes showed transport activity if wild type protein was reconstituted into them with a mean signal of 5667 CPM compared to 1575 CPM for the empty control. Unfortunately, the freshly prepared liposomes showed again a lower signal with only 3015 CPM for the wild type-containing and 1730 CPM for empty liposomes. Nevertheless, the difference was higher this time than for the first trial which might indicate that the problem was a combination of the preparation and another problem.

Old liposomes were used to check if dialysis of the proteoliposomes increases the overall signal as it was found out by former experiments in the group. Indeed, the overnight dialysis increased the overall wild type to about 12 146 CPM while the empty signal only marginally increased as seen in figure 5.2 c. These results were consistent with the former results indicating that the assay gave reproducible results. After a while, the lipids were identified causing the problems and after ordering a new batch, the assay gave similar results when performed with newly prepared liposomes as seen in figure 5.3. Without dialysis, the wild type proteoliposomes gave 5221 CPM while 8776 CPM for dialyzed proteoliposomes. The empty liposomes showed signals of 1680 CPM and 2914 CPM.

In general, the dialysis resulted in a better window between the wild type and empty signal and it was continued with using only dialyzed liposomes for further experiments. A last comparison of old versus new proteoliposomes was performed and is shown in figure 5.3 b. The wild type signal was 8776 CPM and 10 655 CPM for new and old liposomes. The empty signal was almost the same for both, 2194 CPM and 2620 CPM respectively. Finally, the assay seemed to be giving reproducible results with respect to former obtained results by former members and the window between empty and wild type signal was considered as good enough for later investigation of $ASBT_{NM}$ mutants activity. At last, it was checked for how long the proteoliposomes could be stored at 4 °C after reconstitution. The wild type signal slightly decreased on the second day with

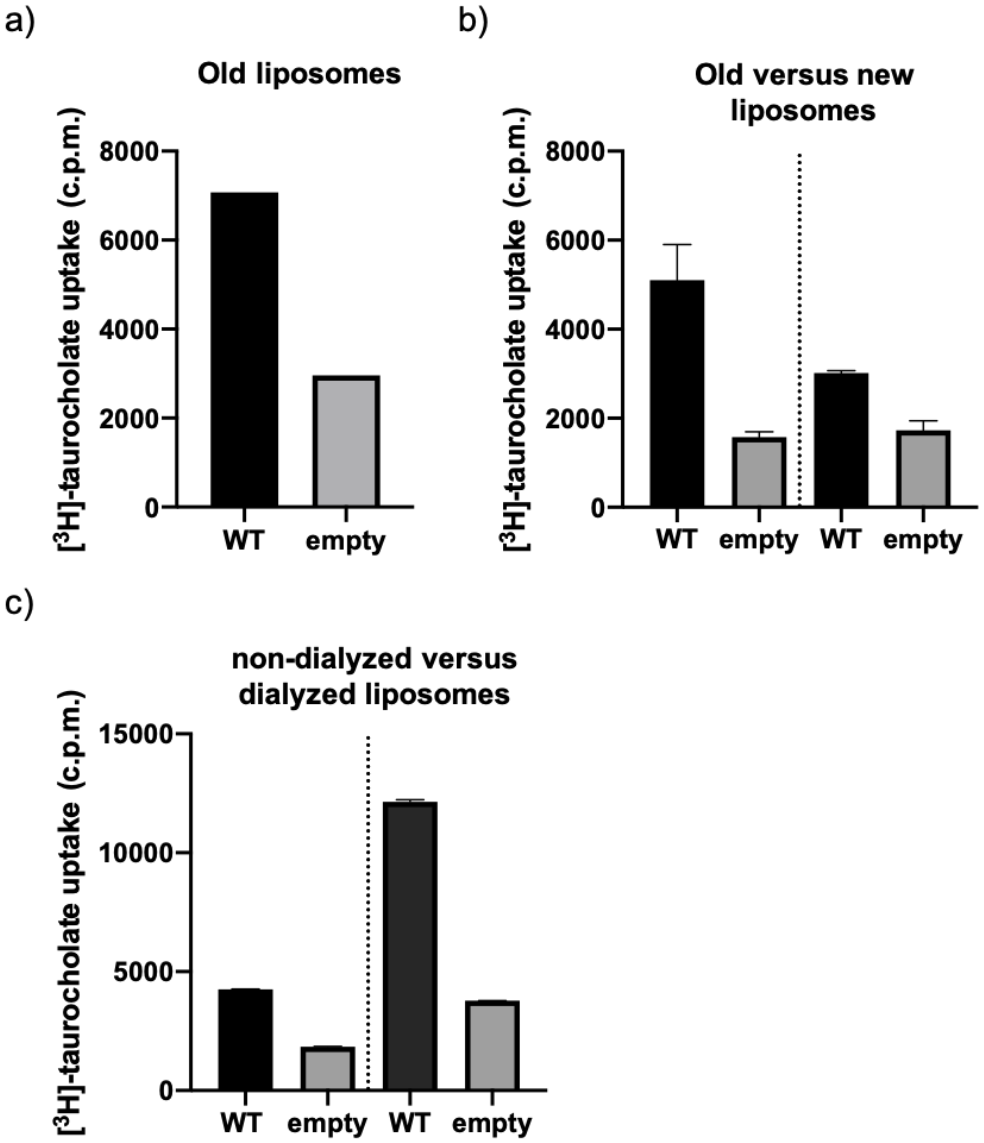


Figure 5.2: Transport assay comparing old liposomes versus newly prepared liposomes and dialyzed versus non-dialyzed liposomes: a) Internalized ^{3}H-taurocholate signal after 5 min of transport at 37 °C of wild type containing and empty liposomes prepared from an old batch of liposomes. b) Internalized ^{3}H-taurocholate signal of wild type containing and empty liposomes prepared with an old batch of liposomes (left) and newly prepared liposomes (right) c) Internalized ^{3}H-taurocholate signal of wild type containing and empty liposomes which were dialyzed (right) or not (left). Experiment was performed with old liposomes.

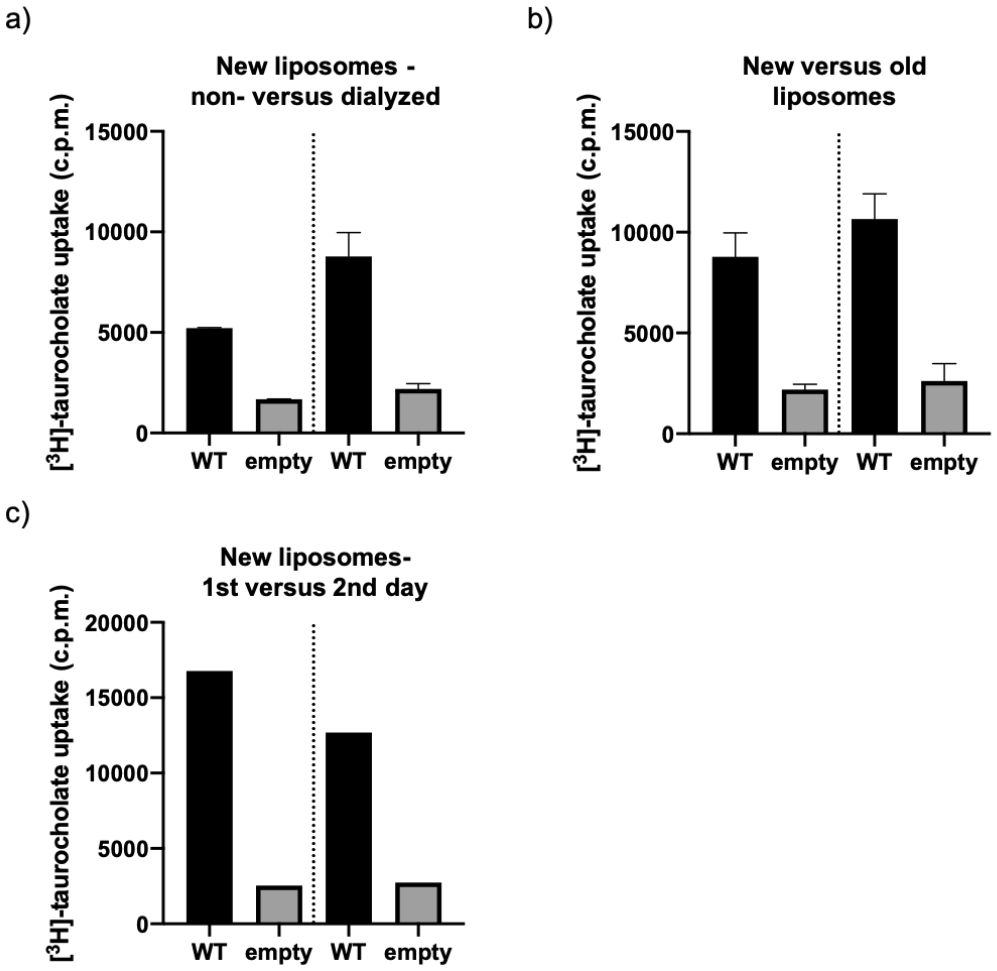


Figure 5.3: Comparison of the transport activity of newly prepared versus old liposomes with a new batch of lipids: a) Internalized ^{3}H-taurocholate signal after 5 min of transport at 37 °C of wild type containing and empty liposomes with newly bought lipids. Difference in signal between non-dialyzed (left) and dialyzed liposomes are shown. b) Comparison of the transport activity of wild type containing and empty liposomes reconstituted in old (right) and newly prepared liposomes (left). Proteoliposomes were dialyzed overnight. c) Comparison of the signal between dialyzed proteoliposomes stored at 4 °C for one or two days after reconstitution.

the empty signal increasing. Nevertheless, the difference was still large enough after two days and it was decided that the proteliposomes can be used for two days, in case needed.

5.1.2 Time-course and negative controls

Next, a time-course was performed to check the time-dependent uptake of taurocholate into the liposomes and to investigate the time frame in which the transport follows a linear dependency. This needs to be known as for later experiments, the transport activity will be compared between a sodium and a no-sodium condition and between mutants and the comparison needs to be done at time points where the transport is linear. Both, wild type and empty liposomes, showed the expected increase in counts over time (see figure 5.4). The specific uptake was calculated by subtracting the empty signal from the wild type signal. The graph shows a linear dependency within the first five minutes which starts to flatten at about 10 min and reaches saturation at 20 min. This result was consistent with former results and it was concluded that accumulation after 5 min would be used for the transport assay.

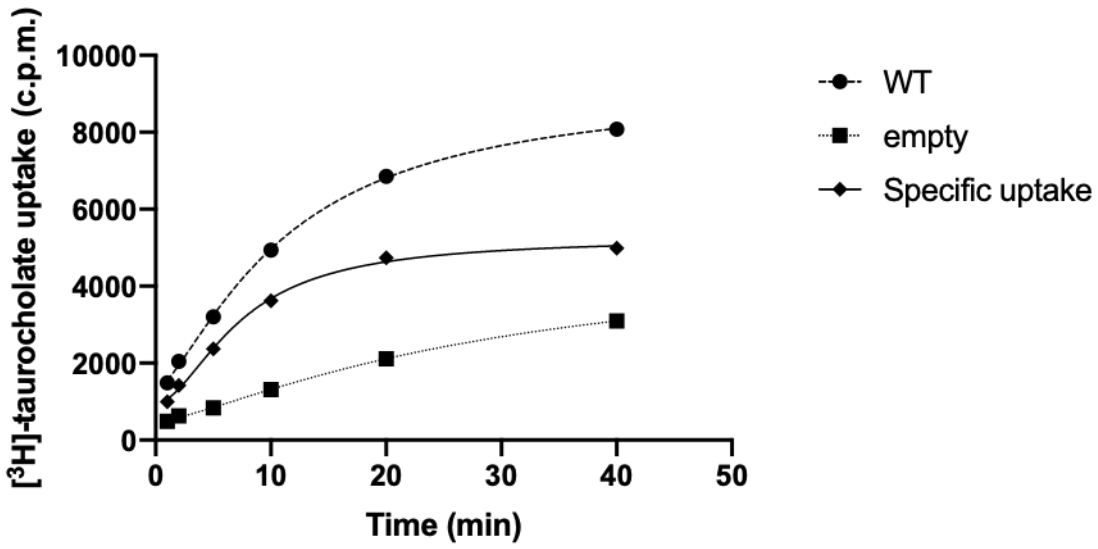


Figure 5.4: Time-dependent taurocholate uptake for wild type containing and empty liposomes: WT (black circles) and empty signal (black squares) are shown as dotted lines. The specific uptake (black diamond) was calculated by subtracting the empty from the WT signal and is shown as a black curve.

Next, it was continued with trying the necessary controls for the assay. First, a no sodium control was performed as $ASBT_{NM}$ is a sodium-coupled transporter which needs the energy of a sodium gradient to perform transport. As seen in figure 5.5 a, non-dialyzed wild type-containing liposomes showed only a small difference when 137 mM choline chloride was added instead of sodium chloride. Dialysis of the liposomes increased the differences between the normal wild type signal (16 779 CPM) to the no sodium control (7729 CPM). Figure 5.5b shows the reduction in transport activity as a percentage of the wild type activity. As estimated from two different repeats with the same proteoliposomes, the wild type protein showed a reduction of around 65% when choline chloride was added instead of sodium chloride. NapA was used as another negative control. NapA shares the same fold than $ASBT_{NM}$, but does not transport taurocholate. As shown in figure 5.5 c, NapA-containing liposomes showed only a marginally higher signal than empty liposomes, probably due to the fact that liposomes are a bit leakier if reconstituted with a protein. Even though the control experiments were only performed twice, they showed the expected results and were promising. Therefore, it was decided to put the the assay to the final test. It was planned to perform the needed amount of repeats of the control experiment at a later stage to make a significant statement.

5.1.3 The failure of the assay

As described at the beginning, the transport assay was hoped to be used for characterizing all $ASBT_{NM}$ mutants according to their transport activity. At the beginning, a few mutants were chosen to see if the assay allows detection of differences in their transport activity. N265A and E260L were chosen as they have shown reduction of transport activity in a range between 30 and 50% compared to the wild type. Additionally, Q77AQ264A was used as it was shown to have no transport activity at all (Sekiguchi 2014). Unfortunately, the assay did not show any differences in transport activity for the E260L and N265A mutant (see figure 5.6). Only for the Q77AQ264A mutant, an abolished transport activity could be detected in agreement with former results. The results are shown from two individual reconstitutions and the counts were normalized according to the

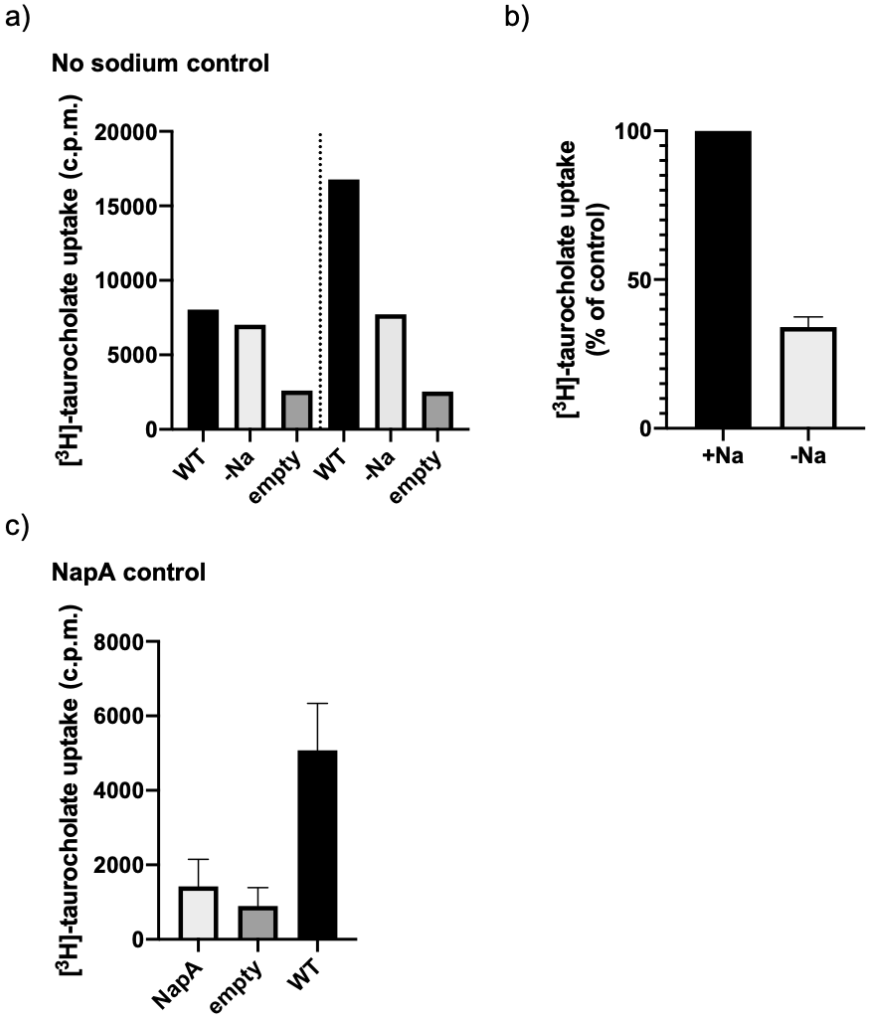


Figure 5.5: Negative controls for the transport assay (No sodium control and NapA control): a) Non-dialyzed proteoliposomes (left) and dialyzed liposomes (right) were checked for internalized ^{3}H-taurocholate signal after 5 min of transport at 37 °C of wild type containing and empty liposomes with either 137 mM NaCl (WT) or choline chloride (-Na) on the outside of the liposomes. b) No sodium (-Na) activity is shown as percentage of the activity with sodium (WT). Empty signal was subtracted from both signals before calculating the percentage. c) Comparison of the transport activity of wild type or NapA containing and empty liposomes.

protein content. In the end, it was decided to stop working on this transport assay as it did not show the hoped differences between the mutants and the assay had already been optimized (data not shown). Unfortunately, several attempts of optimizing the assay did not show any improvement. Therefore, it was decided to stop working on the soybean-transport assay and to try a different approach to see if better results could be obtained.

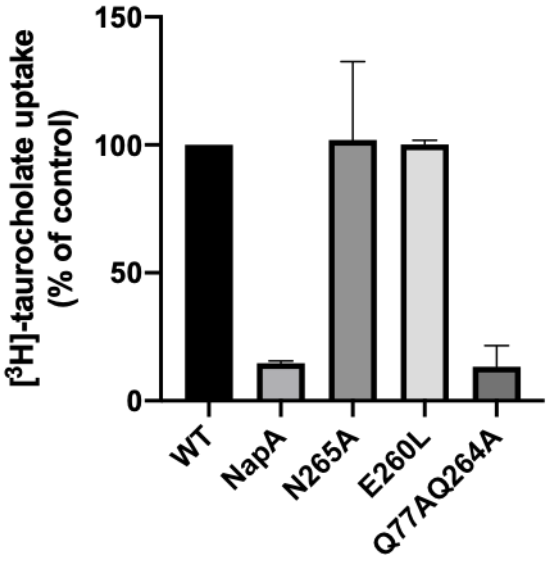


Figure 5.6: Transport activity of $ASBT_{NM}$ mutants using the soybean-based transport assay: Normalized ^{3}H-taurocholate uptake after 5 min of transport at 37 °C of wild type, NapA, N265A, E260L, and Q77AQ264A containing liposomes. The empty signal was subtracted from all values and the wild type signal was set to be 100%. The mean uptake is shown from two independent repeats of the experiment and the standard deviation.

5.2 Development of a new transport assay

After the soybean lipid-based transport assay did not show the hoped results, it was decided to look for other methods trying to get an assay set-up allowing to characterize the different $ASBT_{NM}$ mutants according their transport activity. Zhou et al. used a liposome-based transport assay for $ASBT_{YF}$ which is based on the usage of *E. coli* polar lipids and the usage of triton for the reconstitution step (X. Zhou et al. 2014). Given that both transporters share a high homology and have the same substrate, it was concluded that trying their protocol was the most

sensible way, also in regards to time as it was halfway through the PhD. Liposomes were prepared by evaporating the chloroform and resuspending the lipids in potassium phosphate buffer (pH 7.5) and 1.5% octyl glucoside. Liposomes were dialyzed overnight to remove the detergent and flash frozen. Reconstitution is performed by the addition of 0.12% Triton X-100. Destabilized liposomes are incubated with the protein in DM and the detergent is removed by the addition of Bio Beads. Proteoliposomes are harvested by ultracentrifugation on the next day and flash frozen until further usage. (see material and methods for more detailed information).

First, the reconstitution method was tried with freshly prepared liposomes and the reconstitution was checked by SDS-PAGE gel electrophoresis. Figure 7 shows the SEC chromatogram of wild type protein in 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, and 0.3% DM. The protein eluted as a single, monodisperse peak indicating that the protein is stable in DM and can be used for the reconstitution (see figure 5.7 a). The collected fractions were checked for their protein content using a NanoDrop and the fraction showing the highest concentration was used for reconstitution without any further concentration step. As shown in figure 5.7 b, the resuspended liposome pellet after the ultracentrifugation showed a band for $ASBT_{NM}$ wild type while the empty liposomes did not to which the same volume of buffer only was added. To note here, at this point it was not clear if the band showed reconstituted protein or aggregated protein as both would pellet during the ultracentrifugation. Nevertheless, the gel shows an example how a dilution series of known amounts of pure protein could be used to work out the protein amount in the liposomes. After showing the potential reconstitution, a transport assay was performed. The proteoliposomes were diluted to 5 mg/ml in 10 mM Hepes (pH 7.4), 100 mM NaCl, 1 mM$CaCl_2$, 1 mM $MgCl_2$ and 1 μM of taurocholate of which 0.2 μM was ^{3}H-taurocholate to start the transport. After 5 min, the transport was terminated by running the (proteo-)liposomes over a G-25 column as it was done before for the soybean-based liposomes. The results are shown in figure 5.7 c. The wild type proteoliposomes gave a signal of 11 218 CPM compared to only 1744 CPM for the empty liposomes.

It was continued with performing a time course experiment to identify the

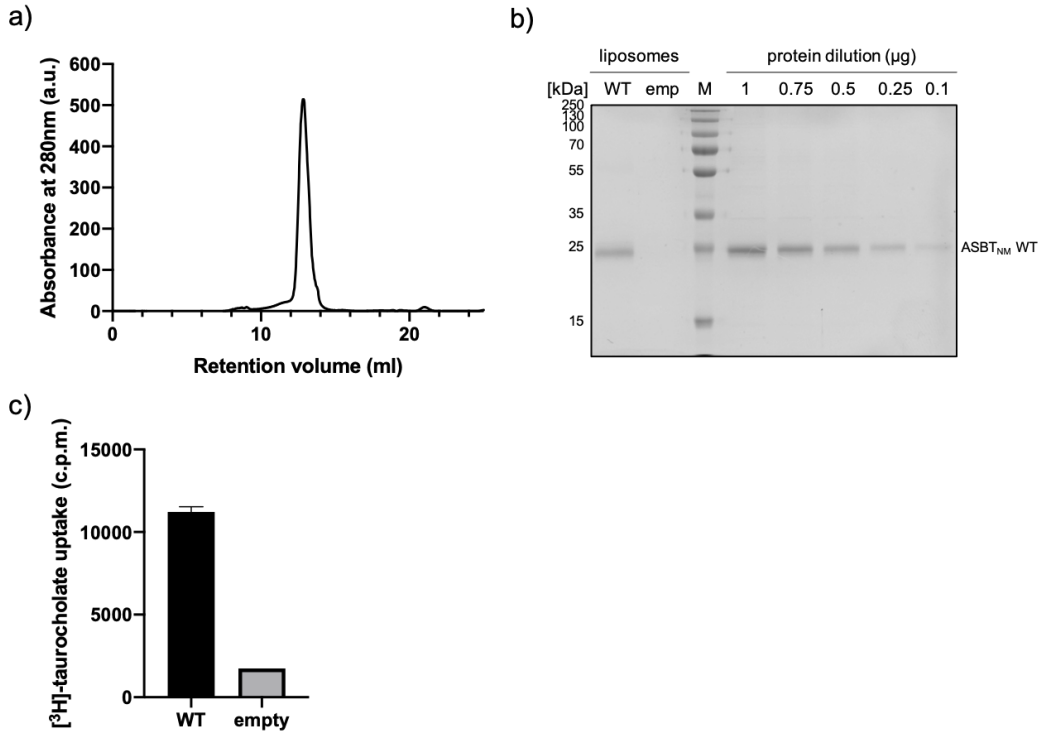


Figure 5.7: Size exclusion chromatogram of $ASBT_{NM}$ WT in DM buffer, SDS gel showing the reconstitution efficiency for the *E. coli* polar lipid-based assay and transport activity of WT and empty liposomes: a) SEC chromatogram of $ASBT_{NM}$ wild type in 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, and 0.3% DM. b) SDS-PAGE gel showing the samples of the WT-containing (WT) and empty (emp) liposomes (left side). Samples with known amounts of $ASBT_{NM}$ WT are shown on the right side (1 ug/0.75 ug/0.5 ug/0.25 ug/0.1 ug). Dilution series was used to calculate the protein amount. Molecular weight is shown on the left side of the gel. c) Internalized ^{3}H-taurocholate uptake after 5 min of transport at 37 °C of wild type containing and empty liposomes.

time range in which the transport follows a linear dependency which was mandatory for later experiments. The graph shows a steep increase of internalized ^{3}H-taurocholate within the first five minutes of transport which flattens after 10 min and reaching a maximum at 40 min (see figure 5.8 a). After 40 min, the curve starts to decline which could indicate that back transport of taurocholate from the lumen of the liposomes to the outside. In contrast to the soybean liposome assay, no valinomycin was added in this assay which causes an accumulation of positive charge in the lumen and might cause the reversal of transport after a certain time. Empty liposomes were used as a negative control and they showed a low signal compared to the wild type-containing liposomes indicating that the observed signal is because of the incorporated protein. The specific uptake was calculated by subtracting the empty signal from the wild type signal showing a similar trend than the wild type proteoliposomes. The transport followed a linear trend within the first two minutes as it was estimated based on the specific uptake curve. Next, a transport assay was performed with different amounts of taurocholate concentrations to determine the K_M of $ASBT_{NM}$ for taurocholate. The transport was performed for one minute (in the linear range as estimated before) and the transport rates were calculated for each concentration and plotted against the final concentrations of taurocholate. The transport follows a typical Michaelis-Menten kinetic as shown in figure 5.8 b. At lower concentrations, the graph shows a linear dependency meaning that by doubling concentrations the transport rate doubles as well. At higher concentrations, the graph starts to flatten and reaches a plateau indicating that at this concentration all transporter molecules are saturated and the transport rate cannot be increased anymore. The K_M was estimated to be 79 μM with a v_{max} of 267 pmol $mg^{-1}min^{-1}$. Overall, both values were in the expected range compared to a K_M of 52 and v_{max} of 453 as determined from in-cell assay. The results were promising and it was continued to investigate the sodium-dependency. Unfortunately, the no sodium control still showed 80% activity when choline chloride was added instead of sodium chloride as shown in figure 5.8 c.

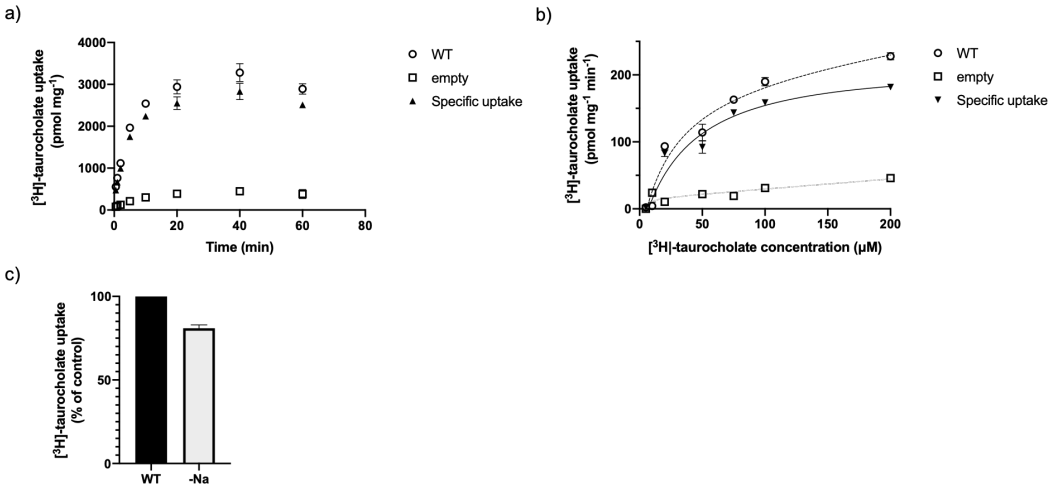


Figure 5.8: Time-dependent taurocholate uptake, Michaelis-Menten transport kinetic, and sodium dependency: a) Internalized ^{3}H-taurocholate signal of WT containing (white circles) and empty (white squares) liposomes. The specific uptake (black diamond) was calculated by subtracting the empty from the WT signal b) Michaelis-Menten transport kinetic of taurocholate uptake WT containing (white circles) and empty (white squares) liposomes. The specific uptake (black diamond) was calculated by subtracting the empty from the WT signal. c) Internalized ^{3}H-taurocholate signal of WT containing liposomes after 1 min of transport with either 100 mM NaCl (WT) or choline chloride (-Na). Results are normalized to the empty liposome signal.

5.2.1 Introduction of a vacuum manifold for washing the liposomes

One idea was that the no sodium control might show some unspecific binding of the taurocholate to the liposomes or to the protein. It was thought that the control might work when the samples are washed more excessively. When checking the literature, it was found out that vacuum manifolds are most commonly used for separating the liposomes from the free label. The (proteo-)liposomes are added onto a membrane and are washed by addition of wash buffer and by applying a vacuum. The liposomes will stick to the membrane because of their size while the free label and buffer are sucked through the membrane. The remaining radioactivity on the membranes are equal to the internalized substrate. The usage of a vacuum manifold has the advantage of offering the possibility to wash the liposomes with as much buffer as needed. A first try with the vacuum manifold was performed and was compared to the column usage as shown in figure 5.9. The sample was added to 1 ml of wash buffer (100 mM potassium phosphate buffer (pH 6), 100 mM LiCl) and after the vacuum was applied the membrane was washed with an additional 1 ml of wash buffer. The vacuum manifold did not show a large difference between the wild type, the no sodium, and the empty signal when compared to the column experiment performed in parallel. The empty liposomes showed a really high signal of around 5000 CPM for the vacuum manifold indicating that the label binds either unspecific to the membrane or to the liposomes and suggests that more wash steps are necessary to reduce the signal.

In a next experiment, the volume of wash buffer was further increased. The liposomes were added to 2 ml of wash buffer and washed with two times 2 ml of wash buffer. Additionally, the effect of 1 mM cold taurocholate in the wash buffer was investigated. Cold taurocholate could potentially compete with ^{3}H-taurocholate and therefore reduce the background further. As seen from figure 5.10 a, the increase of washing buffer reduced the difference between empty and wild type signal if compared to the experiment shown in figure 5.9. Overall, the WT signal only marginally decreased while the empty liposomes showed a larger difference indicating that the signal from the empty liposomes is caused by unspe-

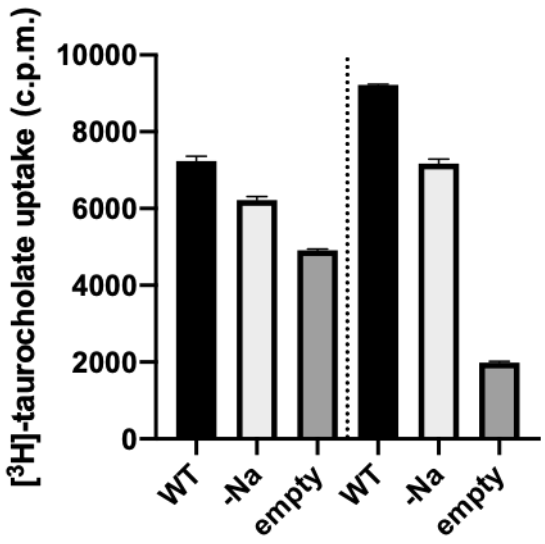


Figure 5.9: Comparison of the vacuum manifold and G-25 column to separate free label from the proteoliposomes: Internalized ^{3}H-taurocholate signal after 5 min of transport at 37 °C of wild type with (WT) and without sodium (-Na) and empty liposomes. Free ^{3}H-taurocholate was either remove by using a vacuum manifold (left) or by using a G-25 column (right). For the vacuum manifold, liposomes were added to 1 ml of wash buffer (100 mM potassium phosphate buffer (pH6), 100 mM LiCl) and washed with an additional 1 ml.

cific bound ^{3}H-taurocholate. Additionally, the addition of 1 mM cold taurocholate to the wash buffer decreased the empty signal further while the wild type signal stayed almost the same and supports the idea of the cold taurocholate competing with the ^{3}H-taurocholate. In the end, a publication was found where they are using basolateral membranes from liver to measure the taurocholate transport of NTCP (Novak et al. 1989) and their protocol was adopted. The membranes were prewashed with 3.5ml of wash buffer, then (proteo-)liposomes were added to 3.5 ml of wash buffer and three time washed with 3.5ml. The result is shown in figure 5.10 b. The wild type signal decreased slightly to about 5000 CPM while the empty signal was about 1600 CPM in comparison to the former experiment. To check how much of the radioactive label has bound unspecific to the membrane, one sample was added in the run in which only the same amount of label was added without liposomes to the membranes and the washing procedure was performed as described above. The membrane showed only a signal of 310 CPM which is negligible and it looked like this washing procedure was working.

To put the protocol with the vacuum manifold to a final test, a short time

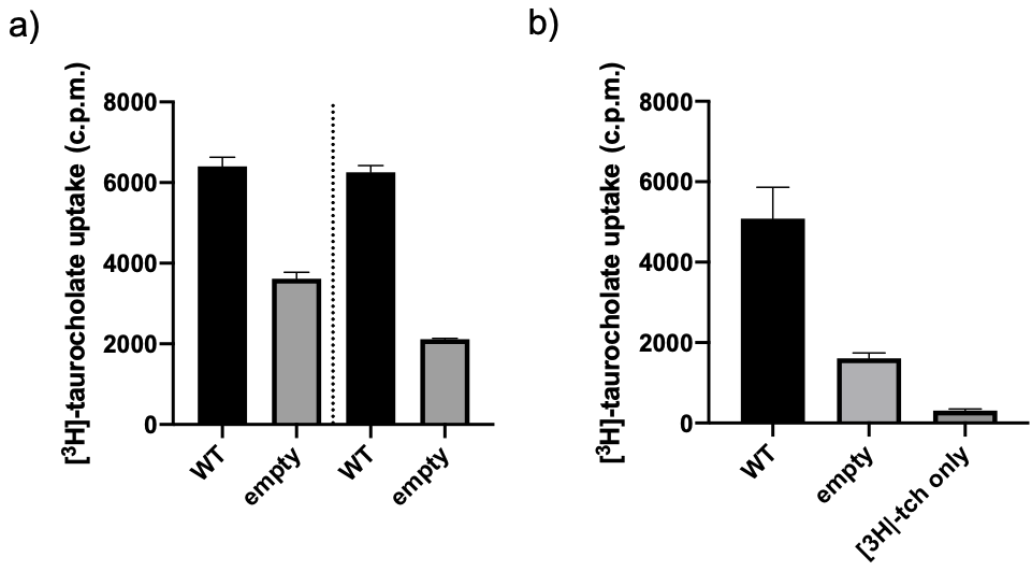


Figure 5.10: Comparison of increasing volumes of wash buffer used to separate free label from the proteoliposomes by the vacuum manifold: a) WT and empty liposomes signal when the liposomes were added to 2 ml of wash buffer (100 mM potassium phosphate buffer (pH 6), 100 mM LiCl) either with (left) or without 1 mM taurocholate (right) and washed with additional 2 ml. b) WT and empty liposomes signal. The membrane was prewashed with 3.5 ml of wash buffer (100 mM potassium phosphate buffer (pH 6), 100 mM LiCL, 1 mM taurocholate), liposome were added into 3.5 ml of wash buffer, and the membrane was washed with three times 3.5 ml of wash buffer.

course was performed with and without sodium to check again the linear range of transport for the vacuum manifold. As seen in figure 5.11, the wild type shows the normal trend showing a linear trend in the first 30 s of transport for the calculated specific uptake. The no sodium control unfortunately did not show a huge decrease in activity. At 30 s the transport activity was around 60% for the no sodium control. A last experiment was performed with two mutants to see if the mutants show any differences in comparison to the soybean-lipid based assay. The E260A and N265A mutant were tried, but again unfortunately the mutants did not show any decrease in transport activity. Figure 5.11 b shows the activity relativized to the protein amount and to the wild type activity. Surprisingly, both mutants showed a slightly higher activity which is un likely to be the case. As it was almost at the end of the PhD and some more promising results were obtained regarding a new substrate of $ASBT_{NM}$, it was decided to stop working on the transport assay and to concentrate on other experiments.

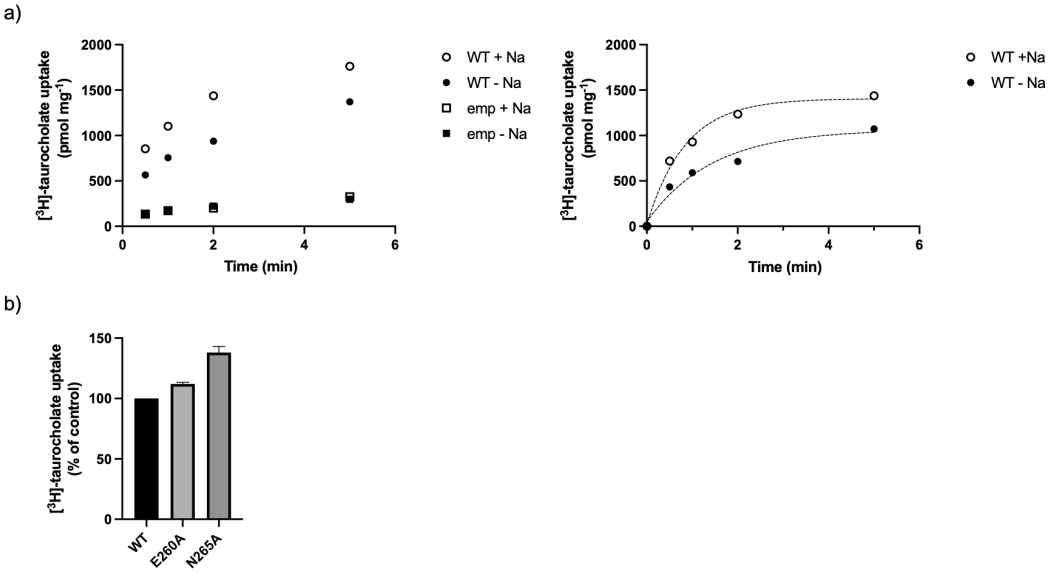


Figure 5.11: Time/sodium-dependent taurocholate uptake and differences in transport activity of $ASBT_{NM}$ E260A/ N265A a) Internalized ^{3}H-taurocholate signal of WT containing (white circles) and empty (white squares) liposomes with 100 mM sodium. The specific uptake (right) was calculated by subtracting the empty from the WT signal. No sodium signal of WT containing (black circles) and empty (black squares) liposomes with 100 mM choline chloride. The specific uptake (right) was calculated by subtracting the empty from the WT signal. b) Transport activity in percentage to the WT protein of $ASBT_{NM}$ E260A and N265A. Transport was performed for 30 s.

5.3 Binding trials of taurocholate to $ASBT_{NM}$ WT by Microscale thermophoresis

It was tried to use MST to measure the binding affinity of $ASBT_{NM}$ WT to taurocholate. Other techniques which are commonly used for determination of the K_D like Circular dichroism (CD) or Isothermal calorimetry (ITC) have been tried by a former PhD student (Sekiguchi 2014) but were shown to not work, probably because of the detergent nature of the taurocholate. MST is a method which measures the movement of a molecule in a temperature gradient. The movement is dependent on the hydration shell, the size of the molecule and the charge. Therefore, the signal changes upon the addition of a ligand to the protein as the binding usually changes one of the parameters. For the MST experiment, the protein is used as a GFP-fusion protein. The sample is filled into small capillaries and a laser is used to heat the sample in the capillary. The fluorescence is measured to follow the movement of the molecule out of the heated area. By measuring the differences in thermophoresis of the protein with different concentrations of the ligand, the affinity for the protein to the ligand can be determined. The general recommended procedure was followed using the software provided and after several prechecks, an experiment was performed with different amounts of taurocholate aiming to obtain a binding affinity. When checking the MST traces, it was clear that the difference between the apo protein and the taurocholate bound protein is really small indicating that substrate binding might not hugely affect the beforehand mentioned factors. Nevertheless, a K_D experiment was tried and figure 5.12 shows the obtained data points with the fitted curve by the program. The estimated KD was around 2.3 mM. As seen in the graph, there was no s-shaped curve obtained from the experiment even though 2.5 mM was used as the highest taurocholate concentration which should saturate the protein which was used at 50 nM. Several optimization trials were performed changing the protein concentration, the pH, and the buffer composition (data not shown) but unfortunately nothing helped to obtain a saturation at higher concentration and to obtain a more s-shape curve. Therefore, it was decided to not continue with the MST and focus on the transport assay which was discussed before.

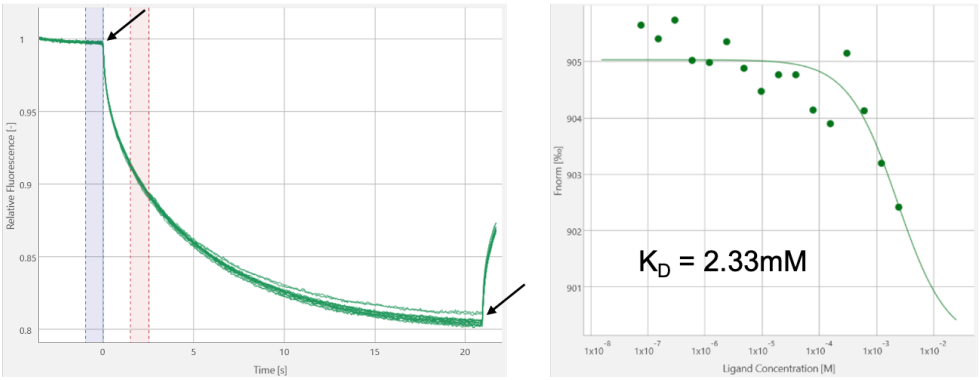


Figure 5.12: Microscale thermophoresis experiment of $ASBT_{NM}$ WT with different taurocholate concentrations: Left image shows the measured MST traces for the sixteen different taurocholate concentrations with the WT protein. Blue area indicates the fluorescence before heat induction and red area indicates the fluorescence after heat induction. Both values are used for calculation of the F_{norm} (right side). Arrow indicates the time point when the laser is turned on and switched off. ''Fitted" curve for estimation of the K_D of $ASBT_{NM}$ WT with taurocholate.

5.4 Discussion

In this chapter, the functional characterization of taurocholate transport and binding were described and unfortunately the experiments which were performed here have not been successful. It was planned to characterize $ASBT_{NM}$ wild type and mutants according to their transport activity by using a liposome-based transport assay. The results could have been useful to better understand hASBT which shares homology to $ASBT_{NM}$.

Two different transport assays were tried which are based on different lipids, a different reconstitution method and a different detergent removal method. Overall, transport activity of $ASBT_{NM}$ wild type was shown for both assays. The performed time-courses showed a time-dependent increase in internalized taurocholate while empty liposomes showed always a lower signal indicating that the reconstituted wild type protein is the cause for higher signal. Potentially, the reconstituted protein could make the liposomes leakier as the protein will move within the lipid-bilayer of the liposomes. This could lead to more unspecific bound taurocholate to the proteoliposomes in comparison to empty liposomes. To rule out this possibility, NapA was used as a second negative control in the soybean lipid-based transport assay. NapA has a similar fold than $ASBT_{NM}$ and will show similar movement in the bilayer. As the NapA control showed almost the same signal than empty liposomes, it can be concluded that the increase in signal is caused by $ASBT_{NM}$ wild type.

When comparing the two transport assays, the soybean lipid-based liposome assay showed overall a higher background in comparison to the *E.coli* polar-lipid based assay indicating that the soybean liposomes are not as tight. This could be caused by several factors. First, the used lipids were different for each assay. The soybean-lipid based assay uses L-α-Phosphatidylcholine from soybean which is a crude extract and can contain other lipids and impurities. One advantage of them is that they are cheaper but this might also explain the higher background for the empty liposomes. The impurities might loosen the lipid bilayer and make it easier for taurocholate to bind. Nevertheless, it needs to be mention here that L-α-Phosphatidylcholine are commonly used for liposome preparations

and transport assays with one example being the functional characterization of NapA in liposomes made of L-α-Phosphatidylcholine from soybean (Uzdavinys et al. 2017).

Another reason could be the usage of different detergents for the reconstitution of the proteins into the liposomes and different removal strategies of them. The soybean lipid-based assay uses sodium cholate with is removed by rapid dilution while the other assay uses Triton X-100 which is removed by Bio Beads. The removal of detergent by BioBeads is probably slower than the dilution method which might favors the formation of tighter bilayers but this is only a hypothesis and needs further investigation, for example by performing the reconstitution with both detergent and both removal strategies and then the empty signal could be compared to see if a pattern can be seen. Allen et al measured the remaining detergent in liposomes after extensive detergent removal by different methods (Allen et al. 1980). They showed that the two detergents show differences in their ability to remove them. Overall, the paper supports the used removal methods for the two different detergents. Cholate can be easily removed by gel filtration and dialysis, while Triton-X100 shows a slower removal by dialysis. BioBeads or gel filtration are described to be the best methods to remove Triton X-100 from liposomes.

Last, the two transport assays were performed in different buffers. The main factor which might explain the differences in the empty liposome signal is the usage of valinomycin in the soybean assay. Valinomycin is an ionophore which allows the selective transport of potassium ions. It was used to counteract the accumulation of the positive net charge in the inside of the liposomes which is build up during the transport of the two sodium ions per taurocholate molecule. As valinomycin diffuses through the membrane to transport potassium ions, it is likely that it makes the bilayer looser and makes it easier for the taurocholate to interact with the lipid bilayer.

Next, the sodium dependency of taurocholate transport was investigated for both transport assays, but unfortunately the sodium control was not working for both assays which is contradictory to the previously findings of Hu et al who showed that the transport is drastically reduced if choline chloride is added instead of

sodium chloride (Hu et al. 2011). They used a cell-based assay which differs from the liposome assay and that might cause the differences. In liposomes, the protein is reconstituted in the membrane in two different orientations. The transporter can either have the extracellular part facing the outside of the liposomes which would be the natural way or it can have the intracellular part facing the outside (non-natural). This is the main difference to the in-cell assay in which the protein will only have its natural orientation in the membrane. Even though it is unlikely, it could be that the transporter unspecifically transport taurocholate in the case of the non-natural orientation. An easy way for checking the orientation of the protein in the liposomes would be by reconstituting the GFP-tagged protein into the liposomes. Once reconstituted, the proteoliposomes can be treated with TEV protease which will only be able to cleave the GFP tag from the protein if it is orientated in the non-natural way (C-terminus with GFP facing the outside). The percentage of the two orientations can be determined by SDS-PAGE by quantifying how much of the total GFP-tagged protein has been cleaved. At one point, it was considered to perform this experiment, but it was decided against it as is known that large hydrophilic parts like the GFP tag can influence the orientation during reconstitution in the way that the large hydrophilic domains tend to point towards the outside of the liposomes because of the smaller lumen of the liposomes (Ritzmann et al. 2017). Therefore, reconstitution with the GFP-tagged protein would most likely not reflect the orientation distribution in the normal setup.

Last, the transport assay was performed with a few mutants to see if the assay at least allows to detect differences in the transport activity between the mutants. The N265A, E260A and E260L have shown about 50% reduction in the activity by the former PhD student in the soybean lipid-based assay. These results were not reproducible as discussed before. It is not really clear why the assay did not work for the mutants and therefore it was decided to stop working on it and focus on a more promising result. As it will be described in the next chapter, a new potential substrate for $ASBT_{NM}$ was identified which is more soluble and does not have any detergent character like the taurocholate. It would be useful to try the liposome-based transport assay with a more soluble substrate to check if it

works better for it. In general, liposome-based transport assay are commonly used for transporters and therefore it is unfortunate that the transport assay did not work for $ASBT_{NM}$.

Taurocholate binding was also tested by MST which was thought to be a useful method to determine binding affinities for the different substrate mutants of $ASBT_{NM}$. Unfortunately, the MST did not show any binding affinities in the range of the K_D determined by the in-cell transport assay (50 μM). The MST was performed with a GFP-tagged protein and that might be the reason why it did not work. The GFP is almost the same size than $ASBT_{NM}$. Therefore, it might be not possible to detect small changes in the protein because the signal is overshadowed by the GFP which will not change upon the addition of taurocholate. It could be tried to label the protein with a fluorescence dye but that would require the introduction of a cysteine at the intra- or extracellular site. Unfortunately, there was no time to do this experiment but it might be useful to try it in the future to see if the GFP-tag caused the problem. Another reason for the failure of the experiment could be the formation of mixed micelles upon the addition of taurocholate to the protein in DDM buffer. This would change the charge of the micelle as taurocholate is negatively charged and could affect the MST signal. It could be tried to use some alternatives for the detergent such as amphipols, SMALPs or nano disc which can keep membrane proteins in solution but might not interact with the taurocholate as much as a detergent micelle.

Chapter 6: Pantoate, a new potential substrate for $ASBT_{NM}$

As mentioned in the introduction, *Neisseria meningitidis* is a bacterium which colonize the nasopharynx and only enters the blood stream under certain conditions (Yazdankhah et al. 2004). Therefore, it is unlikely that the bacterium needs a transporter which transports bile acid into the cell. Moreover, bile acids are toxic for bacteria at high concentrations (Molinero et al. 2019). Because of that, it was hypothesized that the transporter might have a different substrate than bile acid. A new substrate could be beneficial for further characterization of the protein, in view of all the problems which have been described in this thesis regarding the usage of taurocholate in functional assays. In the $ASBT_{YF}$ inward-facing structure (4n7w), a citrate molecule from the crystallization buffer was found bound approximately in the middle of the membrane close to the cross over point of TM4 and 9 (see figure 6.1). This is the location where the substrate is supposed to bind as known from other transporter structures. The citrate molecule interacts intensively with the transporter (see figure 6.1). All residues interacting with the citrate molecules are conserved between $ASBT_{YF}$ and $ASBT_{NM}$ (see figure 6.2) and therefore citrate or citrate-like molecules were the first choice to test as possible binders/ substrates. Citrate interacts only with its two carboxyl groups which are four carbon atoms apart (see figure 6.1). Additionally, the hydroxyl group interacts with Asn259 and the backbone amide group of Gly261. The third carboxyl group of the citrate molecule does not seem to interact with the protein. Therefore, it was decided to try compounds from the Krebs cycle which citrate is a part of and a lot of the intermediates are C4-body substances which share similarities to the citrate molecule. To further identify potential substrates, the adjacent genes of $ASBT_{NM}$ and $ASBT_{YF}$ were investigated and it was looked for potential gene clusters which could hint towards other compounds (see figure 6.1). $ASBT_{YF}$ has two hemolysin-related genes downstream. These are proteins which causes the lysis of red blood cells and are probably not connected to the transporter. More interestingly, upstream adjacent to the transporter gene an aspartate kinase gene was found. Aspartate fits into the previously described pattern of C4-compounds with two carboxyl groups. Additionally, aspartate can be fed into the Krebs cycle via oxaloacetate. Because of this, aspartate was included into the substrate screen together with glutamate

and alpha-ketoglutarate. Unfortunately, no gene cluster could be identified for $ASBT_{NM}$ which has some lipid-related proteins and a ribosome pseudouridine synthase adjacent to it. The adjacent downstream gene is a hypothetical protein which has an unknown function.

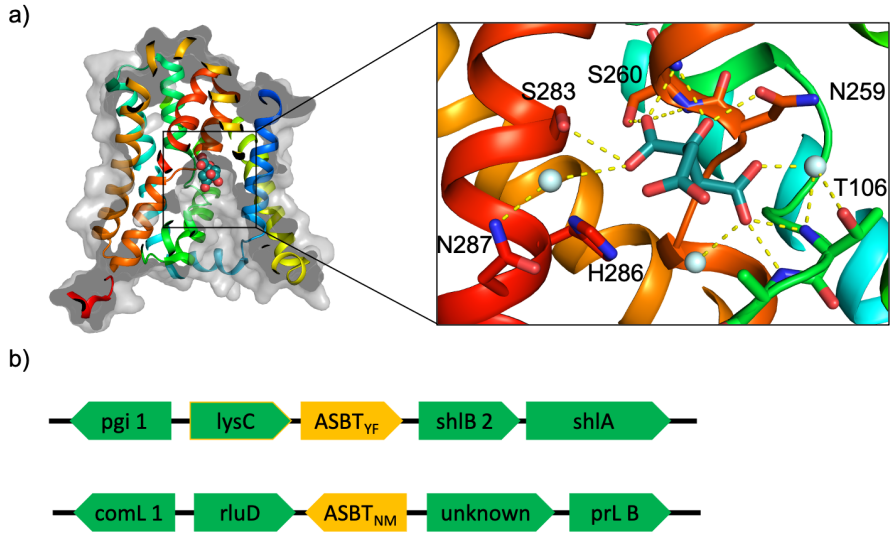


Figure 6.1: Side view of $ASBT_{YF}$ with citrate-bound and schematic representation of genes around $ASBT_{YF/NM}$: a) Surface representation of $ASBT_{YF}$ in the inward-facing conformation with helices differently colored starting with TM1 (dark blue) to TM10 (red). Enlargement of the bound-citrate molecule (blue) with residues interacting with it shown as sticks and water molecules as light blue spheres. b) Gene arrangement around the genes of $ASBT_{YF}$ (top) and $ASBT_{NM}$ (bottom). Two adjacent genes are shown in upstream and downstream direction. The genes are encoding the following proteins: shlB 2 - hemolysin activator protein, shl A – hemolysin, pgi - 1 glucose-6-phosphate isomerase, lysC - aspartate kinase III, rluD - ribosomal large subunit pseudouridine synthase D, comL1 – competence lipoprotein, unknown – hypothetical protein, prL B - putative rare lipoprotein B.

Finally, Ernst et al described a PanS transporter from *Salmonella enterica* which transports CoA-precursors into the cell (Ernst et al. 2015). More precisely, they identified pantoate, ketopantoate, and pantothenate as substrates. Figure 6.2 shows a sequence alignment of $ASBT_{NM}$, $ASBT_{YF}$, and PanS. PanS is almost identical to $ASBT_{YF}$ with a sequence identity of 83% and to a lesser extent to $ASBT_{NM}$ with around 44% identity. All residues which were shown to interact with the citrate are conserved in all three transporters and are marked with an asterisk in the alignment.

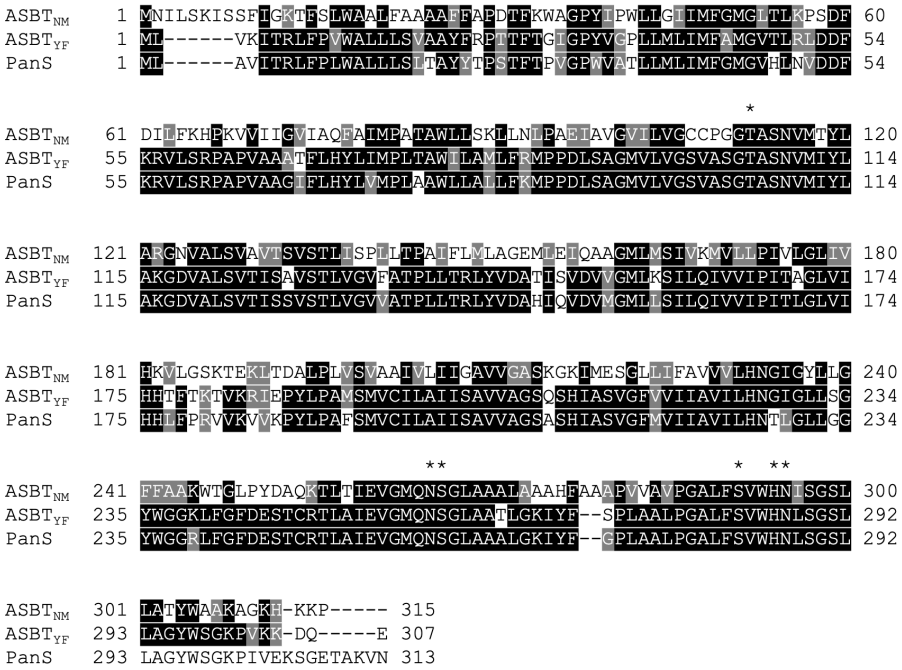


Figure 6.2: Sequence alignment of $ASBT_{NM}$, $ASBT_{YF}$, and PanS: Sequence alignment of $ASBT_{NM}$ (Uniprot: Q9K0A9) to $ASBT_{YF}$ (Uniprot: A0A380PV03), and PanS (Uniprot: Q8ZKL0) using T-coffee (*https* : *//www.ebi.ac.uk/Tools/msa/tcoffee/*) and Boxshade (*https* : *//embnet.vital−it.ch/software/BOX_form.html*). Identical residues are shown in black while functionally conserved residues are shown in grey. Residues which are interacting with the citrate molecule in $ASBT_{YF}$ are marked with an asterisk on the top.

6.1 Substrate screen for $ASBT_{NM}$

After accumulating a variety of potential substrates, the CPM assay was used to see if any of the compounds increase the melting temperature of the wild type protein as described by Majd et al (Majd et al. 2018). As a positive control, taurocholate was added which is a known substrate. The potential substrates were mixed with the wild type protein and were incubated for 30 min at RT to allow binding. Figure 6.3 shows the results for the substrate screen showing the difference in melting temperature to the wild type apo protein. As expected, taurocholate showed an increase in the melting temperature of about 3.5 °C and supports the idea of the assay to identify substrates. Interestingly, a second compound, pantoate, showed a similar stabilization like taurocholate. Other compounds showed less stabilization. Mainly, C4-compounds showed an increase in melting temperature while larger C5-compounds like glutamate and ketoglutarate

only showed a minimal effect. Overall, L-malate, L-aspartate, and oxaloacetate showed increase in the range of 1 to 1.5 °C, probably indicating that they have some functionalities which can interact with the protein, but to a lesser extent than pantoate. In contrast to the findings of Ernst et al, no stabilization was seen when ketopantoate or pantothenate were added indicating that they do not bind to the protein in this assay (Ernst et al. 2015).

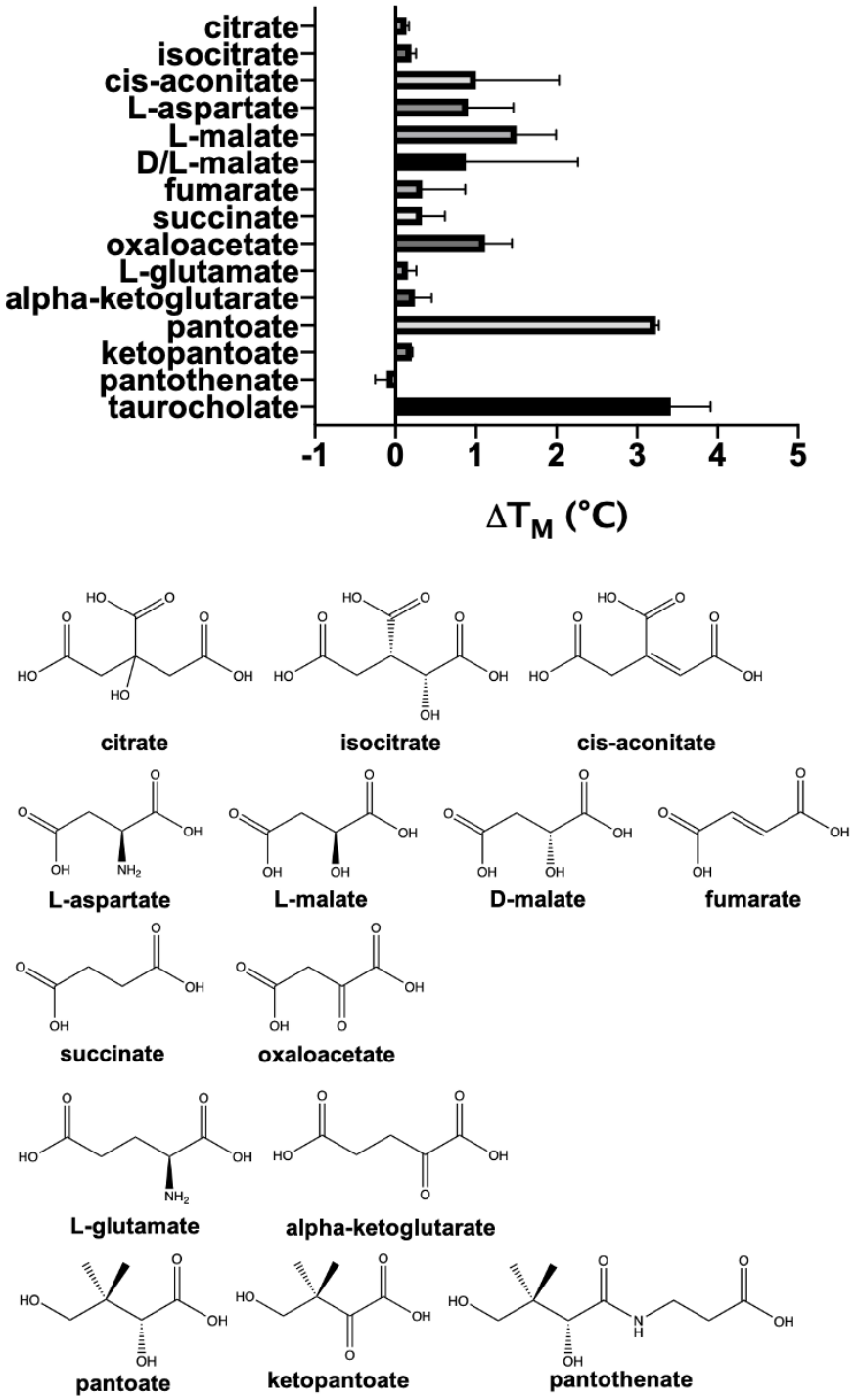


Figure 6.3: Thermostability of $ASBT_{NM}$ wild type with potential substrates: The difference in melting temperature (ΔT_M) is shown and was calculated by subtracting the melting temperature of the apo protein from the melting temperature for the wild type with each potential substrate at a final concentration of 1 mM. The mean is shown from three individual experiments with the standard deviation. Chemical structures of all compounds.

6.2 Confirmation of pantoate binding and affinity measurement by ITC

ITC is a biophysical method measuring the heat absorption or generation during the interaction between two molecules (Rajarathnam et al. 2014). During an ITC experiment, one binding partner is titrated into the sample cell containing the other binding partner. Upon the addition of the binding partner, the temperature either increases or decreases depending on the nature of the event (endothermic or exothermic). A second cell is kept at the same temperature as the sample cell and is used as reference. If the binding event is for example exothermic, the temperature increases in the sample cell and less power is needed to keep the sample cell at the same temperature than the reference cell. Vice versa applies for an endothermic reaction. These changes in power are measured over the period of the experiment and the graph usually consists of several spikes, one for each injection. By integrating the power over the time (area of each spike), the binding enthalpy can be calculated. At last, by plotting the enthalpy against the molar ratio of the titrated binder, the binding affinity (K_D) can be found out. Moreover, ITC further allows the determination of other constants like the entropy and the binding constant n (number of molecule bind).

It was decided to try to confirm the binding of pantoate to $ASBT_{NM}$ WT by ITC. The first initial trials were performed during a demonstration. The protein was concentrated to 15 mg/ml and dialyzed overnight. The dialysis buffer (20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.03% DDM) was used on the next day to dissolve the substrate to ensure that no differences in buffer composition interfere with the measurement. Firstly, a run was performed in which 5 mM pantoate was titrated into dialysis buffer to check if there is any major heat change detectable which would interfere with the later measurements. As seen in figure 6.4, only small heat changes are observed which are normal as the substrate dissolves in the buffer but are in the range of 0.05 μcal/s and therefore negligible.

For the first protein run, the protein was diluted to a final concentration of 30 μM in dialysis buffer and added to the sample cell while 5 mM pantoate solution was added to the syringe for the titration. The sample and reference cell were kept

at 25 °C during the experiment. After the first four injections, no major heat changes were observed and it was decided to increase the concentration of the protein. At 110 μM, the first two injections showed a larger decrease in power than the later injections indicating that some binding occurred in an exothermic manner but the differences were still really low.

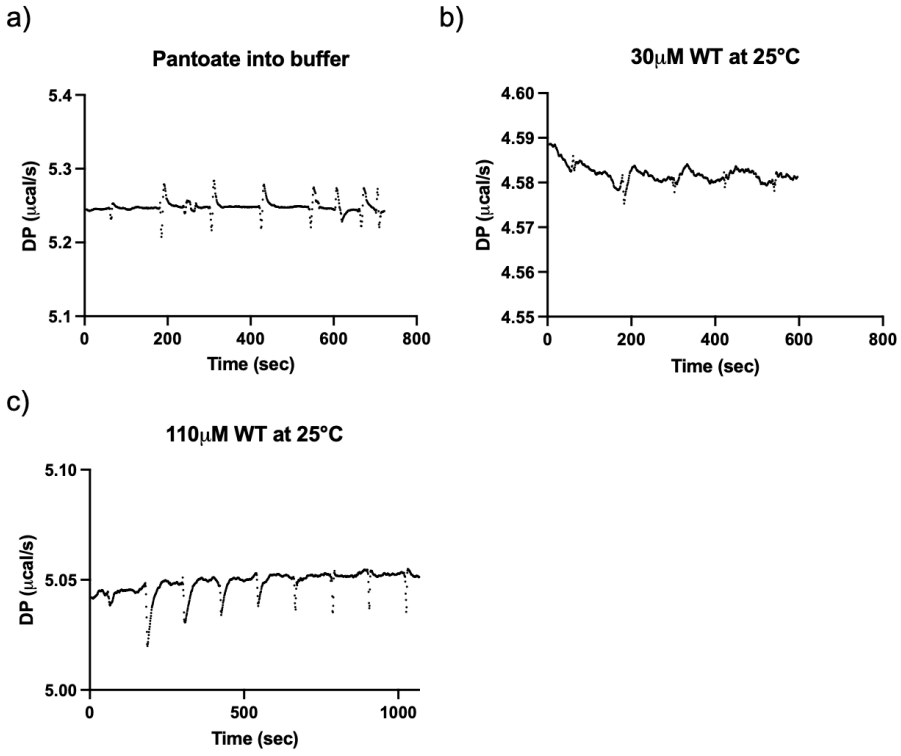


Figure 6.4: ITC heat flow profiles from the substrate-buffer control and the initial protein-substrate trials: a) ITC thermogram of 5 mM pantoate into protein buffer (20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.03% DDM). b) ITC thermogram of 5 mM pantoate into 30 μM $ASBT_{NM}$ WT at 25 °C (sample and reference cell) c) ITC thermogram of 20 mM pantoate into 110 μM $ASBT_{NM}$ WT at 25 °C (sample and reference cell).

In a last run, the protein concentration was again increased to 220 μM and the temperature was decreased to 10 °C which can improve the amplitude as the binding enthalpy is temperature dependent (Rajarathnam et al. 2014). Both factors nicely improved the experiment as seen in figure 6.5. The largest amplitude was about 0.7 μcal/s and changes in the peak amplitudes were observed until approximately the 13th injection allowing to plot a sigmoidal curve to estimate the K_D of the wild type protein for pantoate. The K_D for the wild type protein to pantoate was estimated to be around 124 μM. The binding is exothermic with a

binding enthalpy of −2.1 kcal/mol. To note here, the binding enthalpy was determined by extrapolating the curve during the analysis because experimental data is missing for lower substrate concentrations. The same applies for the molar ratio which cannot be directly determined from the plot and needs to be estimated from the extrapolated curve. The ITC experiment successfully confirmed the findings from the CPM assay and next it was tried to get a structure of the wild type with pantoate bound.

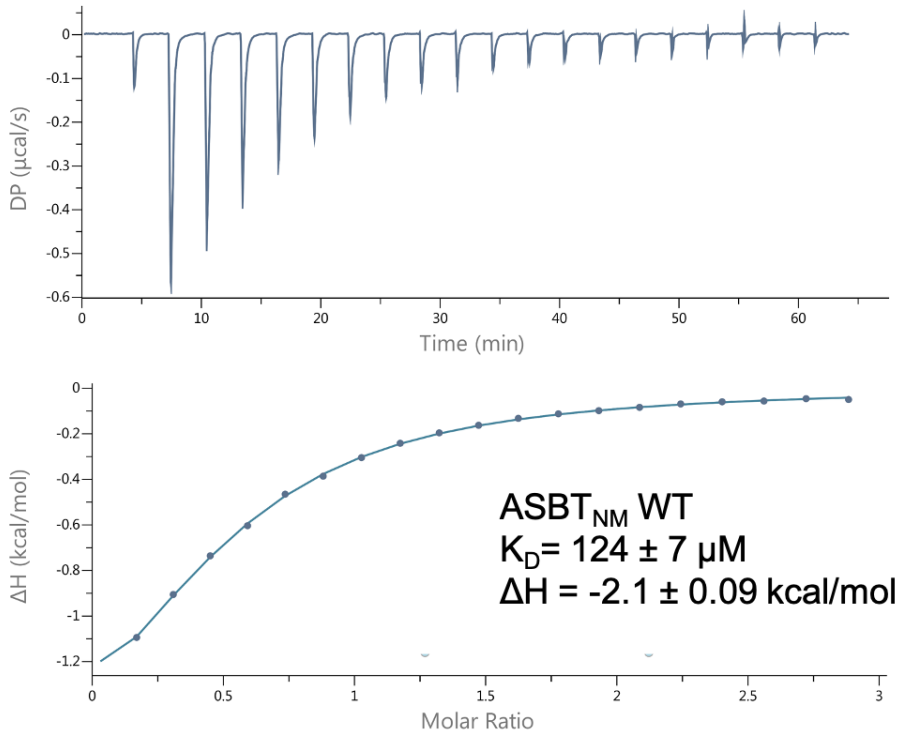


Figure 6.5: Binding affinity of $ASBT_{NM}$ wild type to pantoate estimated by ITC: ITC thermogram (top) of 5 mM pantoate into 220 μM ASBTNM WT at 10 °C (sample and reference cell). Graph showing the integrated heat changes (ΔH) versus the molar ratio. The estimated K_D and the binding (ΔH) enthalpy are shown in the graph.

6.3 A wild type structure with pantoate

Crystals were obtained by co-crystallization of 1 mM pantoate with the wild type protein at 25 mg/ml. The protein was incubated with the substrate for 30 min at RT and then used for LCP crystallization. After one week, crystals appeared in

the MemGold2 A1 condition which contains 0.2 M Magnesium chloride hexahydrate, 0.005 M Cadmium chloride hemi-(pentahydrate), 0.1 M Tris (pH 7.5), and 14% v/v PEG 500 MME. The crystals from the initial screen were screened for diffraction and a dataset was obtained from one of the crystals with the resolution being cutoff at 2.6Å.

After the molecular replacement, the electron density was investigated and there was clear positive density in the substrate binding site for the pantoate. Additionally, there was some more positive density which showed some alpha helical features (see figure 6.6). It was started to build in an alpha helix into the density and after a couple round of refinements, it was clear that the density belongs to TM1. At this resolution, density for some of the bulkier side chains on TM1 were visible (see figure 6.6 right) and helped to build the intracellular half of TM1. It was tried to connect the two parts of TM1, but they did not fit well together. When it was tried to linked both halves together, the r-factors increased and the bond angles/ length showed huge deviations from expected values. Therefore, it seems that TM1 has two different conformation in the crystal, but unfortunately the electron density was not good enough to build in the complete TM1 in the two different conformations. Therefore, the structure contains two separate parts of TM1 which are not connected.

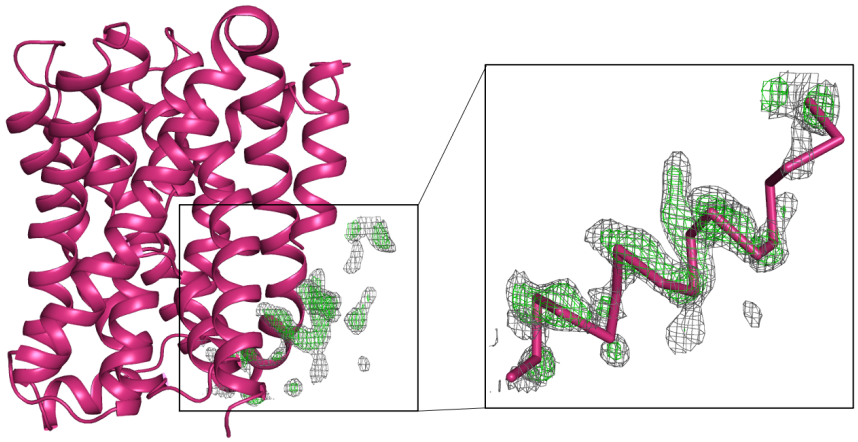


Figure 6.6: Electron density for TM1 in the substrate binding site: Structure after molecular replacement is shown in pink with the electron density found in the substrate binding cavity (left). 2mFo-DFc map is shown at 1 sigma (grey mesh). The mFo-DFc map is shown at +3 sigma (green mesh). Enlargement of the electron density in the cavity with a alpha helix shown in ribbon representation (right)

Figure 6.7 shows the finished structure with the two parts of TM1 showing that in one of the two conformations, TM1 moves into the substrate binding pocket and makes it inaccessible from the intracellular site (see figure 6.7). At the crossover point of TM4 and 9, clear positive density was observed close to the area where citrate was found in the $ASBT_{YF}$ structure. A pantoate molecule was modeled in and is shown in figure 6.7 (bottom image). The carboxyl group interacts with two backbone amino groups of Thr112 and Ala113 while the alpha hydroxyl group interacts with Asn265 and the backbone carbonyl group of Gly267. The rest of the pantoate molecule does not show any hydrogen bonds to the protein.

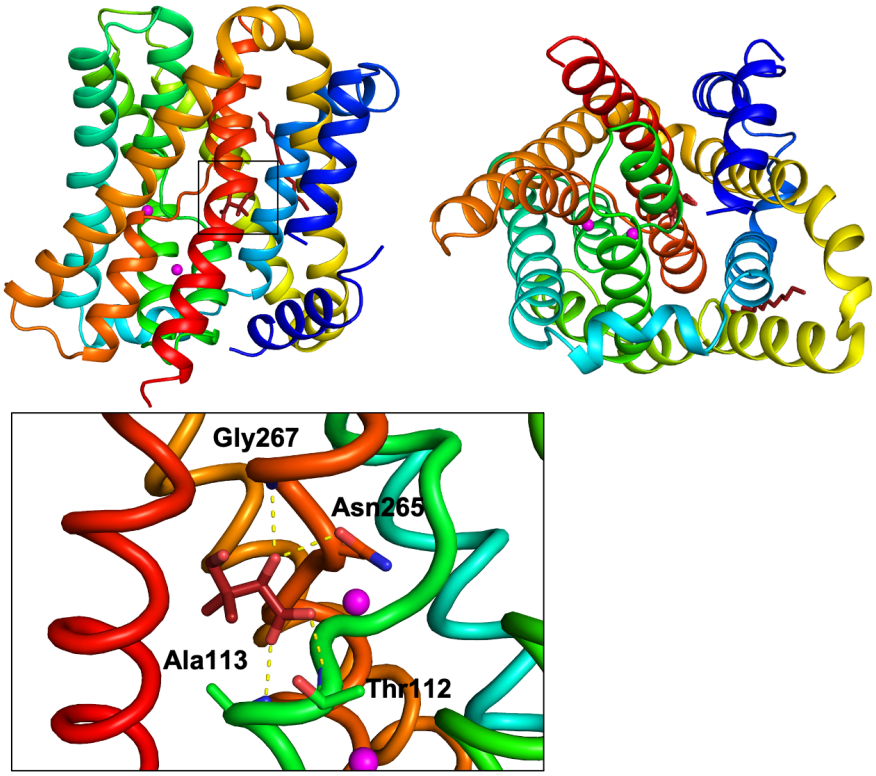


Figure 6.7: $ASBT_{NM}$ wild type structure with pantoate bound: Side view (top left) and intracellular view (top right) of the $ASBT_{NM}$ wild type structure in cartoon representation. Helices are colored differently starting with TM1 in dark blue to TM10 in red. Sodium ions are shown as pink spheres. Pantoate molecule and lipids are shown as sticks (firebrick). Enlargement of the pantoate in the binding site. Hydrogen bonds are shown as yellow dotted lines. Residues which are interacting with the pantoate molecule are shown as sticks and are labelled with their name and number.

The wild type pantoate structure overlays well with the wild type apo sodium structure with a RMSD of 0.448Å. Interestingly, for this structure the main dif-

ferences are found in the core domain and TM6. The extracellular half of TM6 shows again a slight outward movement and the unwinding of its first two extracellular turns similar to the inhibitor structure (see figure 6.8). The adjacent helices TM5 and TM4b have moved upwards compared to the wild type protein and the intracellular view shows a slight rotation movement of them towards IL1. Overall, these movements fit nicely with the idea of the transition between the outward-facing and inward-facing conformation as described before. Therefore, the structure might resemble a transition state, maybe even a potential inward-occluded state with regards of TM1 which blocks the binding site from the intracellular site. But it needs to be further investigated if the movement of TM1 into the substrate binding site occurs naturally or is just a crystallization artifact.

6.4 Affinity of substrate mutants to pantoate

After obtaining the structure and identifying the pantoate binding site, three mutants were checked for their binding affinity for pantoate by ITC which were shown to either interact or being in close contact to the pantoate molecule. Asn265 and Thr112 were mutated to an alanine to see the effect of each residue on the binding. Furthermore, a Thr112Val mutant was used as well which was designed by a former PhD student to investigate if the more hydrophobic substitution affects binding. The experiment was performed using the same condition than for the wild type protein. As seen in figure 6.9, no binding was observed for the Asn265Ala mutant. Asn265 interacts with the alpha-hydroxyl group of pantoate and the interaction seems to be vital. On the other hand, the Thr112Ala and Thr112Val mutants showed a slightly higher K_D with 78 μM and 86 μM. Thr112 is in close proximity to the carboxyl group of the pantoate but the side chain does not directly interact with it. Therefore, it is not surprisingly that the mutant did not show the same drastic effect than the Asn265Ala mutation. Nevertheless, it was expected to see a reduced affinity for at least the valine mutation as the introduction of a hydrophobic residue at the position was assumed to make it less favorable for the hydrophilic carboxyl group to come in close contact. Over-

Table 6.1: Data collection and refinement statistics for the $ASBT_{NM}$ wild type structure with pantoate: Crystals were grown in 0.2 M Magnesium chloride hexahydrate, 0.005 M Cadmium chloride hemi-(pentahydrate), 0.1 M Tris (pH 7.5) and 14% v/v PEG 500 MME. Data was processed using DIALS and AIMLESS in CCP4i2 (Project et al. 1994). Structure was refined using Phenix (Adams et al. 2010).

Data collection and refinement statistics for Wild type pantoate structure	
Crystal	Wild type apo with sodium
Wavelength	0.9795Å
Beamline	I24/ DLS
Oscillation range	0.10°
Exposure time	0.030 s
Transmission	15.62 %
Beam size	7x7 μm
Space group	C121
Resolution	45.95 - 2.6Å (2.693 - 2.6)
CC_{50}	0.91 (0.652)
Cell dimensions	85.2568 89.1829 53.7752 90 124.616 90
Number of measured reflections	63308 (4892)
Number of unique reflections	10241 (1012)
Completeness	99.70 % (98.25)
Multiplicity	6.2 (4.8)
$I/\sigma I$	10.62 (1.16)
R_{merge}	0.3325 (1.372)
Refinement	
Number of reflections	10221 (1011)
R-factor	0.2397
R-free	0.2662
R.m.s.d from ideal values	
Number of non-hydrogen atoms	2314
Bond lengths	0.004
Bond angles	0.72
Ramachandran favored	97.40 %
Ramachandran allowed	2.6 %
Ramachandran outliers	0 %
Rotamer outliers	1.29 %

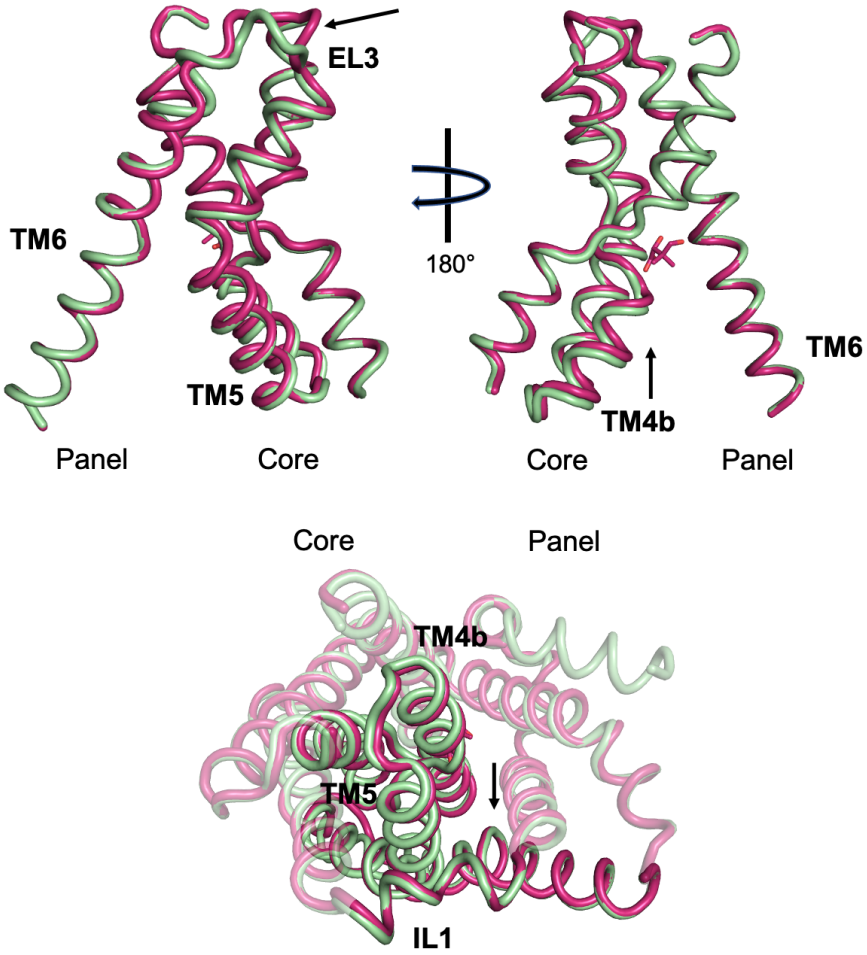


Figure 6.8: Differences in TM4, 5 and 6 between the WT pantoate and WT apo structure: Side view of TM4, 5, 6, and 9 showing the differences between the WT pantoate structure (pink) and the WT apo structure (green). Panel and core domain location are indicated below the structures. Arrows indicate the differences in EL3 (top left) and the upwards directed movement of TM4b on the pantoate structure (top right). Intracellular view of the superimposition with the movement of TM4b toward IL1 indicated by an arrow.

all, the results of the ITC for the mutants support the pantoate binding site as identified from the crystal structure.

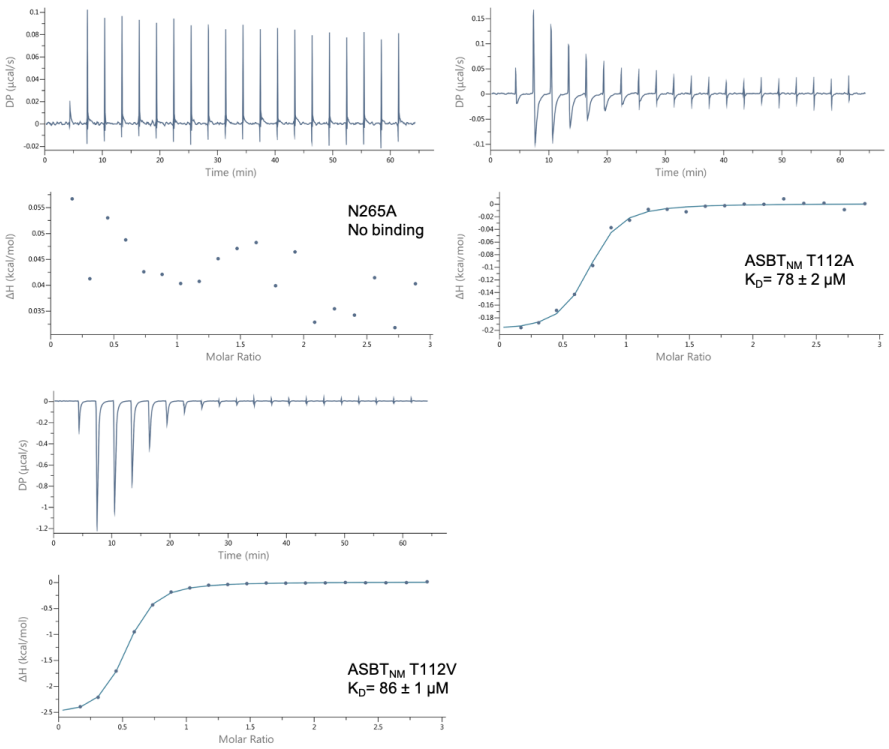


Figure 6.9: Binding affinity of $ASBT_{NM}$ T112A, T112V and N265A mutants to pantoate estimated by ITC: ITC thermogram (top) of 5 mM pantoate into different mutants with the bottom graph showing the integrated heat changes (ΔH) versus the molar ratio for the N265A mutant, the T112A mutant, and the T112V mutant . Estimated K_D values are shown in the graphs for T112A and T112V while no binding was detected for N265A.

6.5 Discussion

In this chapter, the identification of pantoate as a new potential substrate was described from the beginning of identifying it as a potential binder to $ASBT_{NM}$, the validation of binding by ITC, the structural proof by elucidating a pantoate-bound wild type structure and the validation of residues involved in binding by ITC.

The CPM assay which was introduced in the first chapter was used to screen different potential binders which have been selected according to criteria mentioned at the beginning of this chapter. Pantoate showed the highest stabilization among the different compounds which indicated that pantoate binds to the protein.

Interestingly, taurocholate and pantoate stabilized the protein to almost the same extent which is surprising in regards to their size difference. Taurocholate is quite big and it takes up almost all the space in the substrate cavity in the inward-facing conformation seemingly making the protein less flexible which presumably could cause an increase in thermostability as it could hinder unfolding. Additionally, the taurocholate molecule was shown to form hydrogen bonds with four residues in $ASBT_{NM}$ and also has some hydrophobic interactions with residues on the panel domain.

On the other hand, pantoate shows only three hydrogen bonds with the protein, but still shows the same effect than taurocholate. The simplest explanation for this observation could be that the melting temperature just cannot increase anymore just purely because the protein starts to unfold at this temperature, maybe at a different position which is not affected by substrate binding and allows the dye to enter the hydrophobic core and to react with the cysteines. Another reason could be that the two different substrates might lead to different populations of $ASBT_{NM}$ in solution, more precisely the substrates could favor different conformations in solution, for example one favors the outward-facing conformation while the other the inward-facing conformation. Therefore, it is possible that pantoate binding might favor the more stable conformation which shows a higher heat stability.

Nevertheless, the assay was working and pantoate was successfully identified as

a binder. The assay also showed differences in the melting temperature among the different compounds which differ in their size and their functional groups. After obtaining the crystal structure and elucidating the interactions, the other results from the CPM can now be evaluated and it can be tried to explain the observations in regards to the structure.

For example, L-malate is a naturally occurring intermediate in the Krebs cycle and it showed the second highest increase in melting temperature. In comparison to pantoate ((2R)-2,4-Dihydroxy-3,3-dimethylbutanoic acid), the alpha hydroxyl group shows a different configuration in L-malate ((S)-Hydroxybutanedioic acid). In the structure, the alpha hydroxyl group of pantoate interacts with Asn265. When L-malate is modeled in the binding site at the same position with the carboxyl group interacting with the protein backbone, the hydroxyl group of L-malate is not able to hydrogen bond with Asn265 which could explain the lower melting temperature for L-malate (see figure 6.10). Therefore, Asn265 might make the binding site specific for one stereoisomer and allow only binding of R-configurated alpha hydroxyl groups like in pantoate. Unfortunately, no other compounds contained a R-configurated alpha hydroxyl group to support this hypothesis. A racemate of D-/L-malate were checked but it gave really irreproducible results seen from the large error bars and therefore no conclusions could be made from it.

To confirm the hypothesis that the transporter shows a preference for R-configurated alpha hydroxyl groups, the CPM assay could be repeated including D-malate. D-malate is commercially available while no enantiomer for pantoate is available. By comparing the melting temperatures of D- to L-malate, the influence of the configuration of the alpha hydroxyl group could be easily tested. It needs to be mentioned here that the hypothesis is based on the structural information of the pantoate binding to the inward-facing conformation. It cannot be excluded that the binding of the pantoate to the outward-facing conformation is different or maybe that it does not bind to the outward-facing conformation. To investigate if the pantoate binding site is available from the extracellular site, a chimera model was prepared in which TM4 and TM9 together with the pantoate were superimposed onto the $ASBT_{YF}$ outward-facing structure. As seen in figure

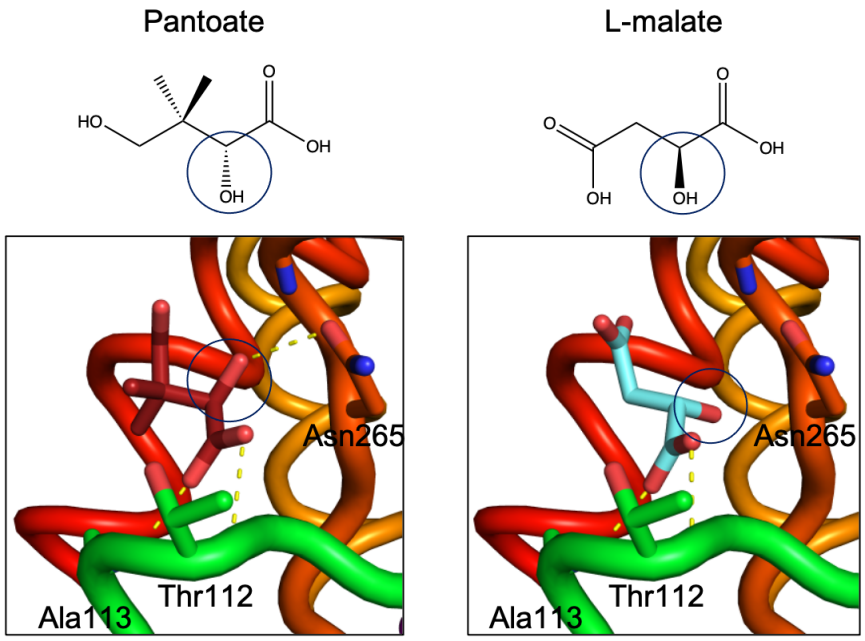


Figure 6.10: Comparison of pantoate and L-malate binding in the inward-facing conformation: Identified binding site of pantoate is shown with residues important for pantoate binding shown as sticks. Pantoate binding is shown in the left image and the right image shows the binding site with a L-malate molecule modeled into the binding site by keeping the position of the carboxyl group the same. Interactions are shown as yellow dotted lines. The blue circle indicates the hydroxyl group of pantoate and L-malate. Pantoate and L-malate structures are shown on the top of the images.

6.11, in the chimera structure the pantoate molecule is located directly in the extracellular cavity with all residues important for binding being accessible from the outside indicating that pantoate could bind to the outward-facing conformation using the same binding site than in the inward-facing conformation. This is in agreement with the alternating-access mechanism.

When overlaid with TM4b and TM9 from $ASBT_{YF}$, almost no difference is seen in the positioning of Asn265 and overall TM9 overlays nicely (see figure 6.12). On the other hand, TM4b shows some minor differences, mainly in the top part of TM4b and the discontinuous part. The top part of TM4b and the discontinuous loop from the outward-facing conformation are slightly located towards the extracellular site in the outward-facing conformation which would clash with the pantoate molecule.

One reason for it could be that the pantoate might bind slightly different to the

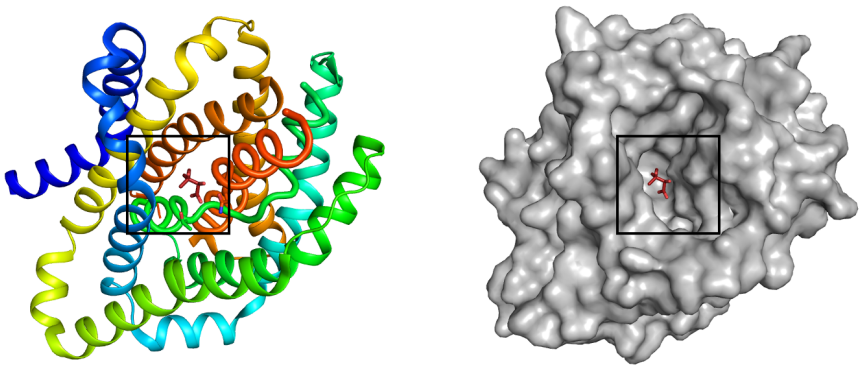


Figure 6.11: Chimera model of the $ASBT_{YF}$ outward-facing structure with TM4 and TM9 from $ASBT_{NM}$: Extracellular view of the chimera model in cartoon representation. TM4 and TM9 $ASBT_{NM}$ are shown in ribbon representation. Helices are colored differently starting with TM1 in dark blue to TM10 in red. Pantoate molecule is shown as sticks (firebrick). Surface representation of the chimera model (right) with pantoate.

outward-facing conformation to adapt for the differences in the binding site. Another reason could be that the pantoate induces some structural changes upon binding. Last, two residues of the Na1 site are located on TM4b and this could indicate some sodium dependency on the pantoate binding. The outward-facing conformation is sodium free and that might be the reason for the slightly altered conformation of TM4b and the loop region. Ser114 and Asn115 of Na1 are located next to Ala113 and Thr112 of the pantoate binding site. Therefore, it is possible that sodium binding in Na1 induces some minor changes in the substrate binding site regarding the positioning of the backbone amino groups of Thr112 and Ala113 to allow pantoate binding. This would be in agreement with the alternating-access mechanism in which ion binding primes the binding site. This is further supported by a B-factor analysis of TM4 which shows that TM4 is more rigid in the pantoate bound structure in comparison to TM4 of the outward-facing conformation where the loop and TM4 shows quite high B-factors indicating that the areas are quite flexible. It needs to be mentioned here that both helices are from different conformations and therefore the differences in B-factors could also result from differences in the overall tertiary structure instead of only sodium or pantoate binding and it needs to be further investigated for example by obtaining more structures of $ASBT_{NM}$.

To prove that sodium binding influences the affinity for the transporter to pantoate, an ITC experiment could be performed with the protein in buffer without sodium to see if the protein shows a lower affinity to pantoate without sodium like it was shown for Glt_{Ph} from *Pyrococcus horikoshii* (Hänelt et al. 2015).

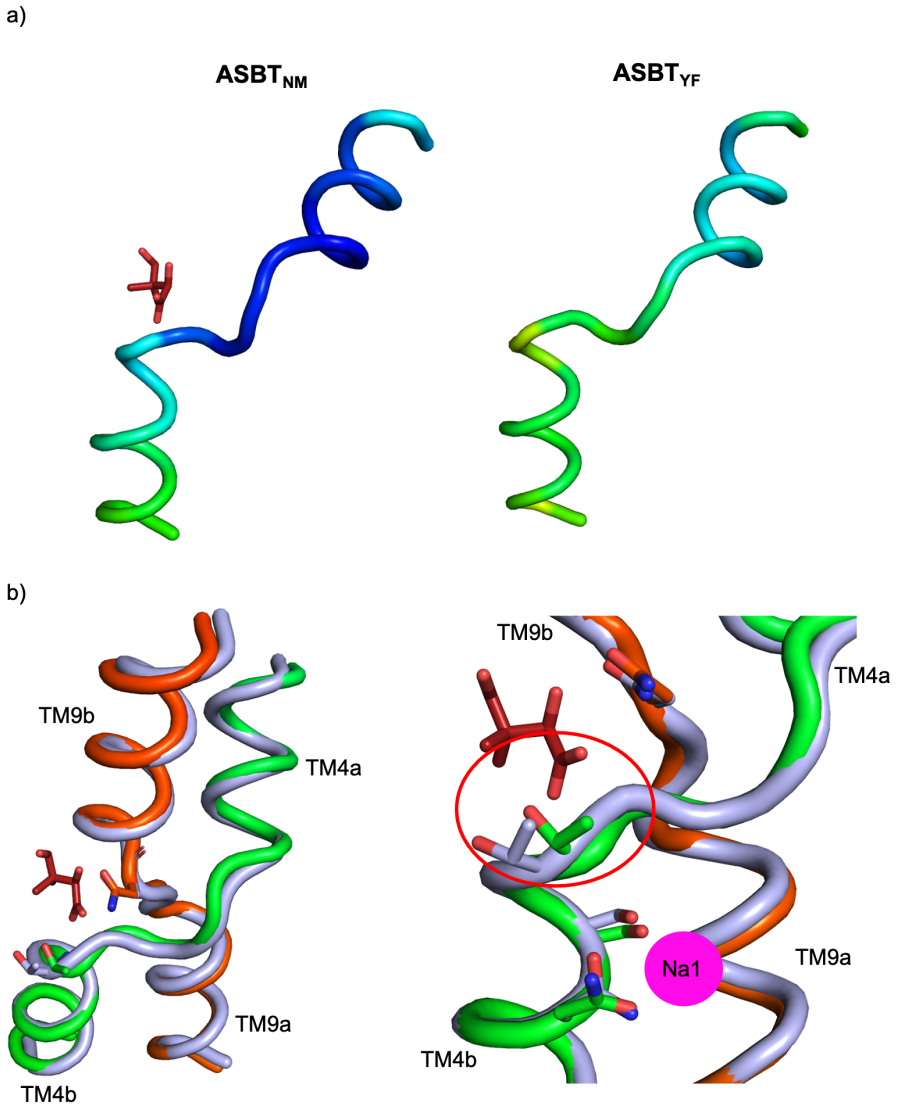


Figure 6.12: Differences in TM4 and TM9 between $ASBT_{YF}$ and $ASBT_{NM}$: a) B-factor analysis of TM4 of $ASBT_{NM}$ with pantoate bound and TM4 of the outward-facing conformation of $ASBT_{YF}$ b) Overlay of TM4 and TM9 of $ASBT_{NM}$ (green and orange) and $ASBT_{YF}$ (light blue). Difference is highlighted and the sodium 1 site is shown as a pink sphere.

6.5.1 $ASBT_{NM}$ plays a potential role in pantoate transport

Interestingly, ketopantoate and pantothenate did not show any increase in the melting temperature which indicates that they do not bind to the protein. Pantothenate is vitamin B5 and is a component of Coenzyme A which plays key roles in a lot of cellular processes. While higher organisms are not able to synthesize pantothenate, bacteria and plants are able to synthesized it de novo from beta alanine and a-ketoisovalerate which is an intermediate in the synthesis of branched amino acids (Ernst et al. 2015). A key enzyme is the keto pantoate reductase which catalyzes the formation of pantoate from keto pantoate. Ernst et al used an indirect approach to identify PanS as a pantoate transporter in which they used a bacterial strain which lacks the ketopantoate reductase and therefore the strain is dependent on the addition of exogenous pantothenate to survive which is transported into the cell by PanF, a pantothenate transporter (Ernst et al. 2015). The identified PanS transporter was shown to restore growth of the strain when solely pantoate, ketopantoate or pantothenate were exogenously added, while cells lacking the transporter did not show any growth. This indicates that the transporter allows the influx of all three compounds which can be used by the cell to synthesize pantothenate and CoA for survival. Contradictory to their findings, ketopantoate and pantothenate did not show any binding to $ASBT_{NM}$ in the CPM assay. This could be just simply because the proteins are not completely similar which could explain the different substrate specificity. Nevertheless, residues which were shown to bind pantoate are conserved between $ASBT_{NM}$ and PanS and the sequence around the binding site shows a high similarity as well which makes it unlikely that they differ so much in the substrate specificity. From a structural point of view, pantothenate binding seems to be unlikely as it is bigger than pantoate and the binding site is not big enough to accompany it. Additionally, the carboxyl group of pantoate was shown to facilitate the major interaction with the protein and in pantothenate, the carboxyl group is linked to a beta alanine via an amide bond which would not allow the same interaction.

Ketopantoate contains an alpha keto group instead of a hydroxyl group and the keto group reduces the rotational freedom at the alpha carbon atom. Ketopantoate was modeled into the binding site, so that its carboxyl interacts in a similar way with the protein than the pantoate (see figure 6.13). This results in the keto group moving away from the Asn265 and does not allow it to interact with it anymore which could be a reason why ketopantoate did not show any stabilization in the assay. Overall, the results from the CPM assay fit nicely with the structural data in explaining why ketopantoate and pantothenate are likely to not bind to the protein, but it needs to be further investigated. The binding of ketopantoate and pantothenate to the protein could be further investigated by ITC which was successfully used for the pantoate.

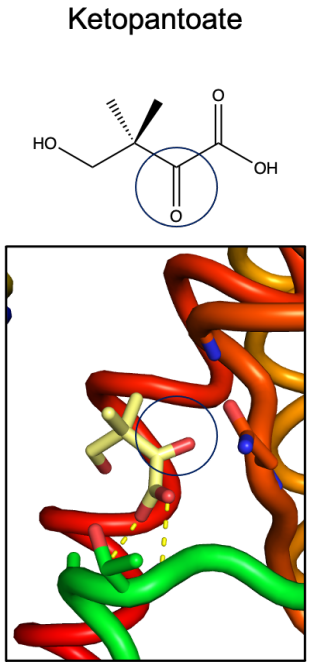


Figure 6.13: Model of ketopantoate binding to $ASBT_{NM}$: The ketopantoate (yellow) was modeled into the binding site by keeping the position of the carboxyl group the same. Interactions are shown as yellow dotted lines. The blue circle indicates the ketogroup of ketopantoate and L-malate. Residues important for pantoate binding are shown as sticks.

Last, $ASBT_{NM}$ shares similarity with the plant Bile Acid Sodium Symporter (BASS) which comprises six members. Four of their members have identified substrate. BASS1 is described as a plastidial pantoate transporter (Huang et al. 2018). It shares 39% identity and 91% similarity to $ASBT_{NM}$ with all pan-

toate binding related residues conserved. Interestingly, BASS2 and BASS6 have a smaller substrate with pyruvate and glycolate (Furumoto et al. 2011, South et al. 2017). Both contain a carboxyl group and either an alpha hydroxyl group or an alpha keto group which are the functional groups of pantoate interacting with $ASBT_{NM}$ and indicating a substrate pattern of different alpha keto or hydroxyl acids The sequence alignment reveals that Asn265 is not conserved among the BASS transporters (see figure 6.15). It is a serine in BASS2 to 6 with the exemption of BASS4, but overall BASS4 shows a really low sequence identity to the other transporters and therefore will be excluded in the discussion. Serine is smaller than asparagine but still contain a hydroxyl group which would be able to interact with an alpha hydroxyl or keto groups. Thr112 is also not conserved in BASS 2 to 6 and is a glutamine. But the side chain of Thr112 does not interact with the pantoate and is probably not as important for substrate binding. One interesting change in the sequence is alanine 113 to a leucine in BASS 3,5 and 6. The introduction of a bulky amino acid at this position makes the binding site smaller and might be responsible that BASS6 has glycolate as a substrate which only has a two carbon backbone. BASS2 does not have the Ala113Leu mutation and transports a slightly bigger substrate pyruvate which has a three carbon backbone (see figure 6.14). To prove the hypothesis, an Ala113Leu mutation could be prepared of $ASBT_{NM}$ and the binding affinities of it for pantoate, pyruvate and glycolate can be determined for the mutant by ITC or by the CPM assay. As mentioned before, Asn265 and Thr112 are also mutated in BASS2 and 6 and might be also necessary for binding.

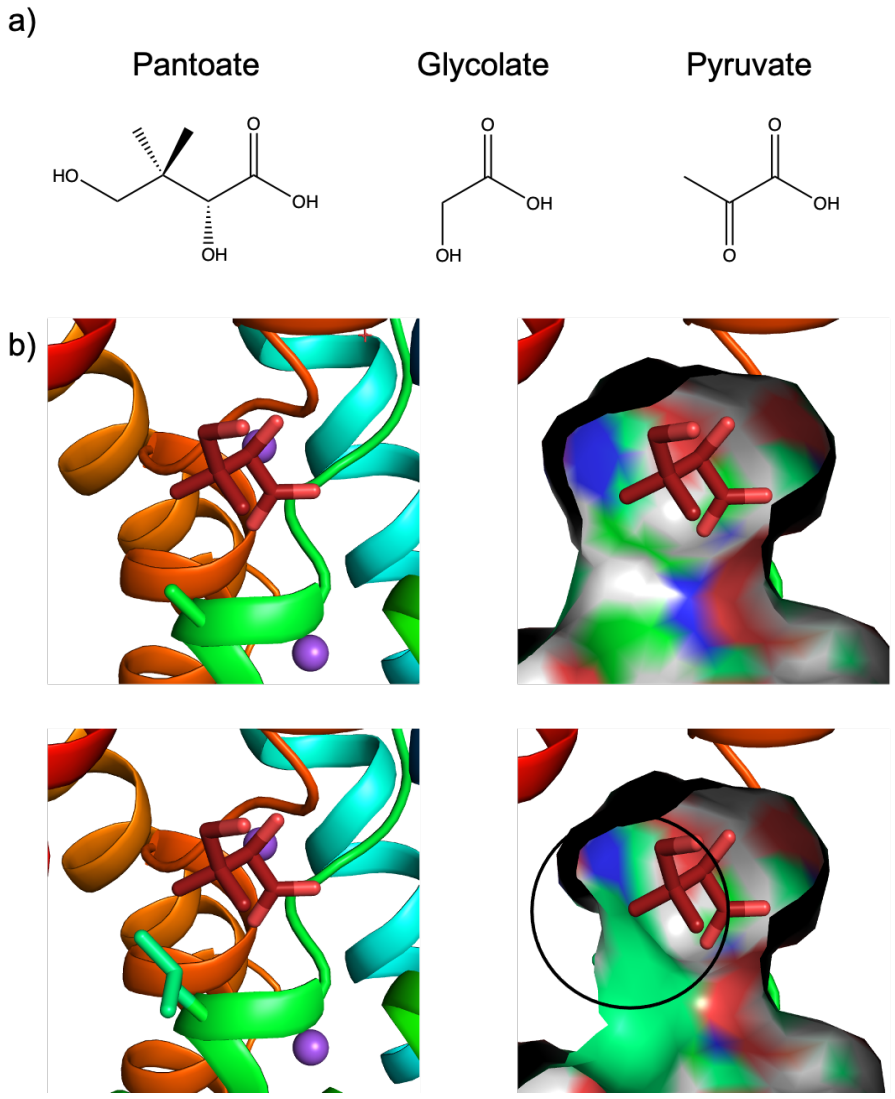


Figure 6.14: Structures of pantoate, glycolate and pyruvate/ Influence on the binding sites by an Ala113Leu mutation: a) Chemical structures of pantoate, glycolate and pyruvate. b) Substrate binding site with pantoate bound as cartoon representation (left) and surface representation (right) of the wild type and the Ala113Leu mutation.

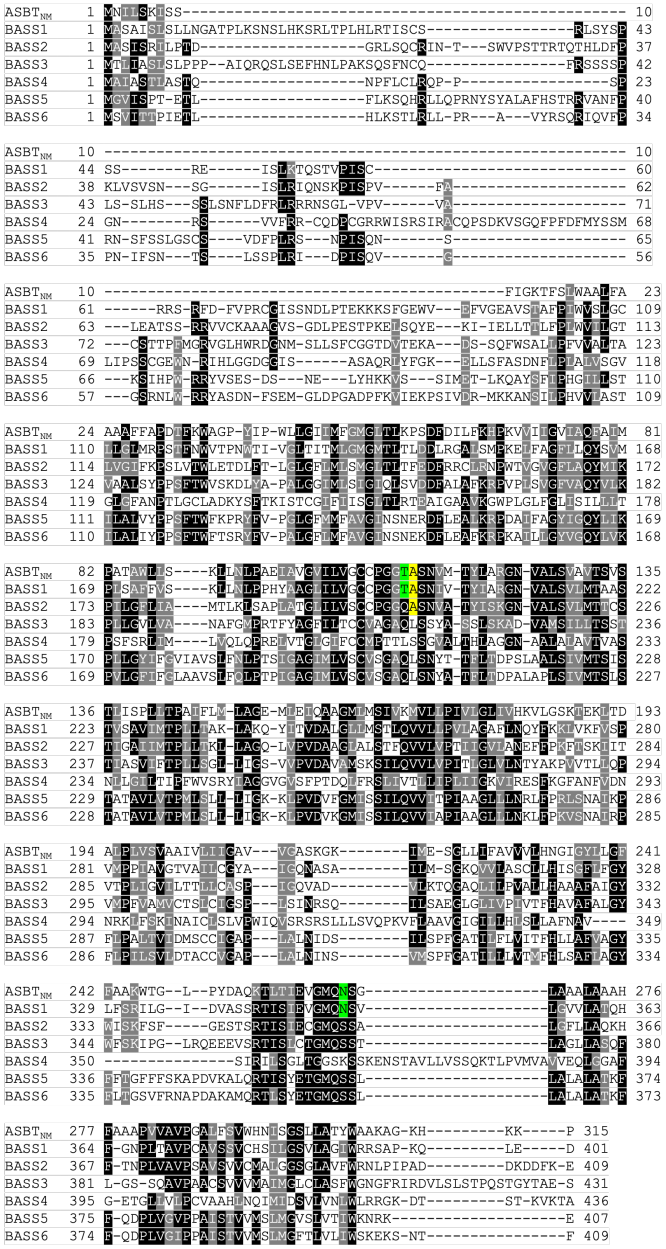


Figure 6.15: Sequence alignment of $ASBT_{NM}$ and the BASS transporter family from Arabidopsis thaliana : Sequence alignment of $ASBT_{NM}$ (Uniprot: Q9K0A9) to BASS1 (Uniprot: Q93YR2), BASS2 (Uniprot: Q1EBV7), BASS3 (Uniprot: Q8RXE8), BASS4 (Uniprot: F4IZC4), BASS5 (Uniprot: F4JPW1), and BASS6 (Uniprot: Q8VYY4) using T-coffee (*https* : *//www.ebi.ac.uk/Tools/msa/tcoffee/*) and Boxshade (*https* : *//embnet.vital−it.ch/software/BOX_form.html*). Identical residues are shown in black boxes while functionally conserved residues are shown in grey. Residues which are interacting with the pantoate molecule are colored in green. Ala113 is marked in yellow.

Chapter 7: Summary and outlook

In this thesis, different new crystal structures of $ASBT_{NM}$ in the inward-facing conformation were presented and discussed. A wild type apo structure was solved at 3.18Å with both sodium ions present. The apo sodium state will naturally occur after the release of taurocholate into the cytosol. The structure showed a slight movement of the panel domain towards the substrate cavity which is probably induced by taurocholate release as taurocholate is quite a large substrate. Next, a wild type structure in a low sodium condition was solved at 3.1Å. The structure is quite similar to the wild type apo structure but showed some difference in sodium occupancy in the different monomers of the asymmetric unit. This fits with the transport cycle as the two sodium ions will leave the protein after the substrate is released as the sodium concentration is lower on the cytosol in ASBT's natural environment. Interestingly, no pattern was observed for the sodium ion release and one monomer only showed sodium in Na1 while the other monomer showed it in Na2. The sodium release does not trigger major conformational changes in the protein. An E260A mutant structure was solved at 3.2Å and showed a potential transition state in which the core domain has started to rotate relative to the panel domain. This movement results in a slightly smaller substrate cavity and might represents the start of the transition between the inward- to outward-facing conformation. Contradictory, one sodium ion was bound in Na2 so the transporter is not completely sodium free. On the extracellular site, the structure showed an unwinding of TM6 in agreement with the $ASBT_{YF}$ structure. Additionally, a slight outward-directed movement of TM6 was observed which fits with the outward-facing structure of $ASBT_{YF}$. The movement of TM6 seems to be one requirement for the opening of the substrate cavity on the extracellular site. Overall, these structures allow more structural insights into the different states of the transporter in the inward-facing conformation and help to understand the transport mechanism. The functional characterization of taurocholate binding and transport did not work and it is likely that the failure is caused by the detergent character of the taurocholate. To overcome this problem, a new potential substrate was identified for $ASBT_{NM}$. Pantoate binds with a K_D of 124 μM to the wild type protein. The structure of pantoate bound to the wild type protein was solved at 2.6Å and the binding site was identified consisting of two

back bone amino groups and Asn265. Asn265 was identified as a key residue for pantoate binding and the mutation to an alanine completely abolished binding to the protein. The identification of pantoate as a potential binder opens up new possibilities to study $ASBT_{NM}$ and also allows to re-evaluate old experiments which have been performed with taurocholate as a substrate. As pantoate is a soluble substrate, experiments could be repeated with it to see if they work with a different substrate. Until now, only inward-facing conformations have been solved of $ASBT_{NM}$. Pantoate as a new substrate might allow to obtain a wild type structure in the outward-facing conformation. It needs to be noted here that the crystallization experiments with the pantoate have been performed almost at the end of the PhD and no other wild type crystal with pantoate were screened or optimized. Therefore, the crystallization with the wild type and pantoate should be continued. Additionally, mutants should be also tried to crystallize as they might favor the outward-facing conformation when pantoate is bound. To obtain an outward-facing conformation, it could also be tried to force the transporter in the outward-facing conformation by introducing disulfide bonds. This has been successfully done for NapA (Coincon et al. 2016). By introducing two cysteines in the protein, the authors were able to trap the transporter in an outward-facing conformation. As the outward-facing conformation for $ASBT_{YF}$ is known, it could be tried to introduce cysteines in $ASBT_{NM}$ to trap it in its outward-facing conformation. For example, two cysteines could be introduced on TM2 and TM4b which move towards each other in the outward-facing conformation. If this experiment is successful, the crystallization of the transporter with pantoate or taurocholate could be tried. From experiences in this thesis and from the soaking experiment of Zhou et al, it is questionable if a taurocholate-bound structure can be obtained by LCP (X. Zhou et al. 2014). Even though 1 mM taurocholate was added to the protein, no taurocholate molecule was observed in the binding site. It is likely that the taurocholate interacts with the LCP and might integrate into the LCP when the monoolein and the protein are mixed during the LCP formation. It might be necessary to go back to vapor diffusion crystallization for the taurocholate. As mentioned before, the functional characterization of $ASBT_{NM}$ was not successful for taurocholate but the transport assay might work

for a different substrate. A major experiment would be to show the transport of pantoate by $ASBT_{NM}$ as all the experiments in this thesis only showed binding. For that purpose, the transport assay could be tried to see if they show better results with the pantoate.

Chapter 8: Bibliography

Adams, Paul D, Pavel V Afonine, Gábor Bunkóczi, Vincent B Chen, Ian W Davis, Nathaniel Echols, Jeffrey J Headd, L-W Hung, Gary J Kapral, Ralf W Grosse-Kunstleve, et al. (2010). "PHENIX: a comprehensive Python-based system for macromolecular structure solution". In: *Acta Crystallographica Section D: Biological Crystallography* 66.2, pp. 213–221.

Agellon, Luis B and Enrique C Torchia (2000). "Intracellular transport of bile acids". In: *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids* 1486.1, pp. 198–209.

Agre, Peter (2006). "The aquaporin water channels". In: *Proceedings of the American Thoracic Society* 3.1, pp. 5–13.

Alberts, Bruce, Dennis Bray, Karen Hopkin, Alexander D Johnson, Julian Lewis, Martin Raff, Keith Roberts, and Peter Walter (2015). *Essential cell biology*. Garland Science.

Alcaiay, Myriam and Daniela Toniolo (1988). "CpG islands of the X chromosome are gene associated". In: *Nucleic acids research* 16.20, pp. 9527–9543.

Alexandrov, Alexander I, Mauro Mileni, Ellen YT Chien, Michael A Hanson, and Raymond C Stevens (2008). "Microscale fluorescent thermal stability assay for membrane proteins". In: *Structure* 16.3, pp. 351–359.

Allen, Theresa M, Alice Y Romans, Henri Kercret, and Jere P Segrest (1980). "Detergent removal during membrane reconstitution". In: *Biochimica et Biophysica Acta (BBA)-Biomembranes* 601, pp. 328–342.

Altschul, Stephen F, Warren Gish, Webb Miller, Eugene W Myers, and David J Lipman (1990). "Basic local alignment search tool". In: *Journal of molecular biology* 215.3, pp. 403–410.

Balakrishnan, Anand and James E Polli (2006). "Apical sodium dependent bile acid transporter (ASBT, SLC10A2): a potential prodrug target". In: *Molecular pharmaceutics* 3.3, pp. 223–230.

Baldwin, Stephen A (2000). *Membrane transport*. Oxford University Press.

Bhat, Eijaz Ahmed, Mohnad Abdalla, and Irfan A Rather (2018). "Key factors for successful protein purification and crystallization". In: *Glob. J. Biotechnol. Biomater. Sci* 4, pp. 1–7.

Carpenter, Elisabeth P, Konstantinos Beis, Alexander D Cameron, and So Iwata (2008). "Overcoming the challenges of membrane protein crystallography". In: *Current opinion in structural biology* 18.5, pp. 581–586.

Chiang, John YL (2013). "Bile acid metabolism and signaling". In: *Comprehensive Physiology* 3.3, pp. 1191–1212.

Chiang, JY (1998). "Regulation of bile acid synthesis". In: *Front biosci* 3.4, pp. d176–93.

Choudhury, Hassanul G, Zhen Tong, Indran Mathavan, Yanyan Li, So Iwata, Séverine Zirah, Sylvie Rebuffat, Hendrik W Van Veen, and Konstantinos Beis (2014). "Structure of an antibacterial peptide ATP-binding cassette transporter in a novel outward occluded state". In: *Proceedings of the National Academy of Sciences* 111.25, pp. 9145–9150.

Coincon, Mathieu, Povilas Uzdavinys, Emmanuel Nji, David L Dotson, Iven Winkelmann, Saba Abdul-Hussein, Alexander D Cameron, Oliver Beckstein, and David Drew (2016). "Crystal structures reveal the molecular basis of ion translocation in sodium/proton antiporters". In: *Nature structural & molecular biology* 23.3, pp. 248–255.

Craddock, Ann L, Martha W Love, Rebecca W Daniel, Lyndon C Kirby, Holly C Walters, Melissa H Wong, and Paul A Dawson (1998). "Expression and transport properties of the human ileal and renal sodium-dependent bile acid transporter". In: *American Journal of Physiology-Gastrointestinal and Liver Physiology* 274.1, G157–G169.

Da Silva, Tatiana Claro, James E Polli, and Peter W Swaan (2013). "The solute carrier family 10 (SLC10): beyond bile acid transport". In: *Molecular aspects of medicine* 34.2-3, pp. 252–269.

Dawson, Paul A (2018). "Bile formation and the enterohepatic circulation". In: *Physiology of the gastrointestinal tract*, pp. 931–956.

Dawson, Paul A, Melissa L Hubbert, and Anuradha Rao (2010). "Getting the mOST from OST: role of organic solute transporter, OSTα-OSTβ, in bile acid and steroid metabolism". In: *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids* 1801.9, pp. 994–1004.

De Duve, Christian and Robert Wattiaux (1966). "Functions of lysosomes". In: *Annual review of physiology* 28.1, pp. 435–492.

Doring, B, J Geyer, and E Petzinger (2006). "Cloning and functional characterisation of the human sodium-dependent organic anion transporter (SLC10A6): Transport of steroid sulfates". In: *Naunyn-Schmiedebergs Archives of Pharmacology*. Vol. 372. SPRINGER 233 SPRING STREET, NEW YORK, NY 10013 USA, pp. 15–15.

Drew, David, Mirjam M Klepsch, Simon Newstead, Ralf Flaig, Jan-Willem De Gier, So Iwata, and Konstantinos Beis (2008). "The structure of the efflux pump AcrB in complex with bile acid". In: *Molecular membrane biology* 25.8, pp. 677–682.

Elkins, Christopher A and Hiroshi Nikaido (2002). "Substrate specificity of the RND-type multidrug efflux pumps AcrB and AcrD of Escherichia coli is determined predominately by two large periplasmic loops". In: *Journal of bacteriology* 184.23, pp. 6490–6498.

Emsley, Paul and Kevin Cowtan (2004). "Coot: model-building tools for molecular graphics". In: *Acta crystallographica section D: biological crystallography* 60.12, pp. 2126–2132.

Ernst, Dustin C and Diana M Downs (2015). "The STM4195 gene product (PanS) transports coenzyme A precursors in Salmonella enterica". In: *Journal of bacteriology* 197.8, pp. 1368–1377.

Fernandes, Carla F, José R Godoy, Barbara Döring, Marcia CO Cavalcanti, Martin Bergmann, Ernst Petzinger, and Joachim Geyer (2007). "The novel putative bile acid transporter SLC10A5 is highly expressed in liver and kidney". In: *Biochemical and biophysical research communications* 361.1, pp. 26–32.

Forrest, Lucy R and Gary Rudnick (2009). "The rocking bundle: a mechanism for ion-coupled solute flux by symmetrical transporters". In: *Physiology* 24.6, pp. 377–386.

Fujisawa, T and M Mori (1996). "Influence of bile salts on β-glucuronidase activity of intestinal bacteria". In: *Letters in applied microbiology* 22.4, pp. 271–274.

Furumoto, Tsuyoshi, Teppei Yamaguchi, Yumiko Ohshima-Ichie, Masayoshi Nakamura, Yoshiko Tsuchida-Iwata, Masaki Shimamura, Junichi Ohnishi, Shingo Hata, Udo Gowik, Peter Westhoff, et al. (2011). "A plastidial sodium-dependent pyruvate transporter". In: *Nature* 476.7361, pp. 472–475.

Gasteiger, Elisabeth, Alexandre Gattiker, Christine Hoogland, Ivan Ivanyi, Ron D Appel, and Amos Bairoch (2003). "ExPASy: the proteomics server for in-depth protein knowledge and analysis". In: *Nucleic acids research* 31.13, pp. 3784–3788.

Gasteiger, Elisabeth, Christine Hoogland, Alexandre Gattiker, Marc R Wilkins, Ron D Appel, Amos Bairoch, et al. (2005). "Protein identification and analysis tools on the ExPASy server". In: *The proteomics protocols handbook*, pp. 571–607.

Geyer, J, CF Fernandes, B Döring, S Burger, JR Godoy, S Rafalzik, T Hübschle, R Gerstberger, and E Petzinger (2008). "Cloning and molecular characterization of the orphan carrier protein Slc10a4: expression in cholinergic neurons of the rat central nervous system". In: *Neuroscience* 152.4, pp. 990–1005.

Geyer, J, T Wilke, and E Petzinger (2006). "The solute carrier family SLC10: more than a family of bile acid transporters regarding function and phylogenetic relationships". In: *Naunyn-Schmiedeberg's archives of pharmacology* 372.6, pp. 413–431.

Godoy, José R, Carla Fernandes, Barbara Döring, Knut Beuerlein, Ernst Petzinger, and Joachim Geyer (2007). "Molecular and phylogenetic characterization of a novel putative membrane transporter (SLC10A7), conserved in vertebrates and bacteria". In: *European journal of cell biology* 86.8, pp. 445–460.

Hagenbuch, B, H Lübbert, B Stieger, and PJ Meier (1990). "Expression of the hepatocyte Na+/bile acid cotransporter in Xenopus laevis oocytes." In: *Journal of Biological Chemistry* 265.10, pp. 5357–5360.

Hagenbuch, Bruno and Paul Dawson (2004). "The sodium bile salt cotransport family SLC10". In: *Pflügers Archiv* 447.5, pp. 566–570.

Hagenbuch, Bruno and Peter J Meier (1996). “Sinusoidal (basolateral) bile salt uptake systems of hepatocytes”. In: *Seminars in liver disease*. Vol. 16. 02. © 1996 by Thieme Medical Publishers, Inc., pp. 129–136.

Hänelt, Inga, Sonja Jensen, Dorith Wunnicke, and Dirk Jan Slotboom (2015). “Low affinity and slow Na+ binding precedes high affinity aspartate binding in the secondary-active transporter GltPh”. In: *Journal of Biological Chemistry* 290.26, pp. 15962–15972.

Hofmann, AF and Donald M Small (1967). “Detergent properties of bile salts: correlation with physiological function”. In: *Annual review of medicine* 18.1, pp. 333–376.

Hofmann, Alan F (1999). “Bile acids: the good, the bad, and the ugly”. In: *Physiology* 14.1, pp. 24–29.

Hofmann, Alan F and Karol J Mysels (1992). “Bile acid solubility and precipitation in vitro and in vivo: the role of conjugation, pH, and Ca2+ ions.” In: *Journal of lipid research* 33.5, pp. 617–626.

Hu, Nien-Jen, So Iwata, Alexander D Cameron, and David Drew (2011). “Crystal structure of a bacterial homologue of the bile acid sodium symporter ASBT”. In: *Nature* 478.7369, pp. 408–411.

Huang, Lili, Michal Pyc, Saleh Alseekh, Donald R McCarty, Valérie de Crécy-Lagard, Jesse F Gregory III, Christopher S Henry, Alisdair R Fernie, Robert T Mullen, and Andrew D Hanson (2018). “A plastidial pantoate transporter with a potential role in pantothenate synthesis”. In: *Biochemical Journal* 475.4, pp. 813–825.

Hunefeld, FL (1840). “Die Chemismus in der thierischen Organization”. In: *Brockhaus: Leipzig*, pp. 158–163.

Hunte, Carola, Emanuela Screpanti, Miro Venturi, Abraham Rimon, Etana Padan, and Hartmut Michel (2005). “Structure of a Na+/H+ antiporter and insights into mechanism of action and regulation by pH”. In: *Nature* 435.7046, pp. 1197–1202.

Jardetzky, Oleg (1966). “Simple allosteric model for membrane pumps”. In: *Nature* 211.5052, pp. 969–970.

Jaskolski, Mariusz, Zbigniew Dauter, and Alexander Wlodawer (2014). "A brief history of macromolecular crystallography, illustrated by a family tree and its Nobel fruits". In: *The FEBS journal* 281.18, pp. 3985–4009.

Jeckelmann, Jean-Marc and Bernhard Erni (2020). "Transporters of glucose and other carbohydrates in bacteria". In: *Pflügers Archiv-European Journal of Physiology*, pp. 1–25.

Kaback, H Ronald, Irina Smirnova, Vladimir Kasho, Yiling Nie, and Yonggang Zhou (2011). "The alternating access transport mechanism in LacY". In: *The Journal of membrane biology* 239.1, pp. 85–93.

Kairam, Sanjay, Diana MacLean, Manolis Savva, and Jeffrey Heer (2012). "GraphPrism: Compact Visualization of Network Structure". In: *Advanced Visual Interfaces*. URL: `http://vis.stanford.edu/papers/graphprism`.

Karakus, Emre, Marie Wannowius, Simon Franz Müller, Silke Leiting, Regina Leidolf, Saskia Noppes, Stefan Oswald, Martin Diener, and Joachim Geyer (2020). "The orphan solute carrier SLC10A7 is a novel negative regulator of intracellular calcium signaling". In: *Scientific reports* 10.1, pp. 1–14.

Kendrew, JC, DICKERSON RE, STRANDBERG BE, HART RG, DAVIES DR, PHILLIPS DC, and SHORE VC (1960). "Structure of myoglobin: A three-dimensional Fourier synthesis at 2 A. resolution." In: *Nature* 185.4711, pp. 422–427.

Kendrew, John C, G Bodo, Howard M Dintzis, RG Parrish, Harold Wyckoff, and David C Phillips (1958). "A three-dimensional model of the myoglobin molecule obtained by x-ray analysis". In: *Nature* 181.4610, pp. 662–666.

Kouzuki, Hirokazu, Hiroshi Suzuki, Kousei Ito, Rui Ohashi, and Yuichi Sugiyama (1998). "Contribution of sodium taurocholate co-transporting polypeptide to the uptake of its possible substrates into rat hepatocytes". In: *Journal of Pharmacology and Experimental Therapeutics* 286.2, pp. 1043–1050.

Kramer, Werner, Frank Girbig, Ulrike Gutjahr, Simone Kowalewski, Karin Jouvenal, Günter Müller, Dominique Tripier, and Giinther Wess (1993). "Intestinal bile acid absorption. Na (+)-dependent bile acid transport activity in rabbit small intestine correlates with the coexpression of an integral 93-kDa and a

peripheral 14-kDa bile acid-binding membrane protein along the duodenum-ileum axis." In: *Journal of Biological Chemistry* 268.24, pp. 18035–18046.

Krogh, Anders, Björn Larsson, Gunnar Von Heijne, and Erik LL Sonnhammer (2001). "Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes". In: *Journal of molecular biology* 305.3, pp. 567–580.

Lee, Chiara, Hae Joo Kang, Anna Hjelm, Abdul Aziz Qureshi, Emmanuel Nji, Hassanul Choudhury, Konstantinos Beis, Jan-Willem de Gier, and David Drew (2014). "MemStar: a one-shot Escherichia coli-based approach for high-level bacterial membrane protein production". In: *FEBS letters* 588.20, pp. 3761–3769.

Lee, Chiara, Hae Joo Kang, Christoph Von Ballmoos, Simon Newstead, Povilas Uzdavinys, David L Dotson, So Iwata, Oliver Beckstein, Alexander D Cameron, and David Drew (2013). "A two-domain elevator mechanism for sodium/proton antiport". In: *Nature* 501.7468, pp. 573–577.

Lewis, Roger and Sherwood Gorbach (1972). "Modification of bile acids by intestinal bacteria". In: *Archives of internal medicine* 130.4, pp. 545–549.

Lippiat, Jonathan D (2008). *Potassium channels: methods and protocols*. Springer.

Loland, Claus J (2015). "The use of LeuT as a model in elucidating binding sites for substrates and inhibitors in neurotransmitter transporters". In: *Biochimica et Biophysica Acta (BBA)-General Subjects* 1850.3, pp. 500–510.

Majd, Homa, Martin S King, Shane M Palmer, Anthony C Smith, Liam DH Elbourne, Ian T Paulsen, David Sharples, Peter JF Henderson, and Edmund RS Kunji (2018). "Screening of candidate substrates and coupling ions of transporters by thermostability shift assays". In: *Elife* 7, e38821.

Mathieu, Khadija, Waqas Javed, Sylvain Vallet, Christian Lesterlin, Marie-Pierre Candusso, Feng Ding, Xiaohong Nancy Xu, Christine Ebel, Jean-Michel Jault, and Cédric Orelle (2019). "Functionality of membrane proteins overexpressed and purified from E. coli is highly dependent upon the strain". In: *Scientific reports* 9.1, pp. 1–15.

McPherson, Alexander (1991). "A brief history of protein crystal growth". In: *Journal of crystal growth* 110.1-2, pp. 1–10.

— (2004). “Introduction to protein crystallization”. In: *Methods* 34.3, pp. 254–265.

Michel, Hartmut and Dieter Oesterhelt (1980). “Three-dimensional crystals of membrane proteins: bacteriorhodopsin”. In: *Proceedings of the National Academy of Sciences* 77.3, pp. 1283–1285.

Molinero, Natalia, Lorena Ruiz, Borja Sánchez, Abelardo Margolles, and Susana Delgado (2019). “Intestinal Bacteria interplay with bile and cholesterol metabolism: implications on host physiology”. In: *Frontiers in physiology* 10, p. 185.

Monaco, Hugo L (2009). “The liver bile acid-binding proteins”. In: *Biopolymers: Original Research on Biomolecules* 91.12, pp. 1196–1202.

Mueckler, Mike and Bernard Thorens (2013). “The SLC2 (GLUT) family of membrane transporters”. In: *Molecular aspects of medicine* 34.2-3, pp. 121–138.

Nelson, David L, Albert L Lehninger, and Michael M Cox (2008). *Lehninger principles of biochemistry*. Macmillan.

Novak, Donald A, Frederick C Ryckman, and Frederick J Suchy (1989). “Taurocholate transport by basolateral plasma membrane vesicles isolated from human liver”. In: *Hepatology* 10.4, pp. 447–453.

Pettersen, Eric F, Thomas D Goddard, Conrad C Huang, Gregory S Couch, Daniel M Greenblatt, Elaine C Meng, and Thomas E Ferrin (2004). “UCSF Chimera—a visualization system for exploratory research and analysis”. In: *Journal of computational chemistry* 25.13, pp. 1605–1612.

Prieto, Ana I, Francisco Ramos-Morales, and Josep Casadesús (2004). “Bile-induced DNA damage in Salmonella enterica”. In: *Genetics* 168.4, pp. 1787–1794.

Project, Collaborative Computational et al. (1994). “The CCP4 suite: programs for protein crystallography”. In: *Acta crystallographica. Section D, Biological crystallography* 50.Pt 5, pp. 760–763.

Qin, Shanshan, Mengmeng Meng, Dehua Yang, Wenwen Bai, Yan Lu, Yao Peng, Gaojie Song, Yiran Wu, Qingtong Zhou, Suwen Zhao, et al. (2018). “High-throughput identification of G protein-coupled receptor modulators through

affinity mass spectrometry screening". In: *Chemical science* 9.12, pp. 3192–3199.

Quick, Matthias and Jonathan A Javitch (2007). "Monitoring the function of membrane transport proteins in detergent-solubilized form". In: *Proceedings of the National Academy of Sciences* 104.9, pp. 3603–3608.

Quick, Matthias, Anne-Marie Lund Winther, Lei Shi, Poul Nissen, Harel Weinstein, and Jonathan A Javitch (2009). "Binding of an octylglucoside detergent molecule in the second substrate (S2) site of LeuT establishes an inhibitor-bound conformation". In: *Proceedings of the National Academy of Sciences* 106.14, pp. 5563–5568.

Rajarathnam, Krishna and Jörg Rösgen (2014). "Isothermal titration calorimetry of membrane proteins—Progress and challenges". In: *Biochimica et Biophysica Acta (BBA)-Biomembranes* 1838.1, pp. 69–77.

Rath, Arianna, Mira Glibowicka, Vincent G Nadeau, Gong Chen, and Charles M Deber (2009). "Detergent binding explains anomalous SDS-PAGE migration of membrane proteins". In: *Proceedings of the National Academy of Sciences* 106.6, pp. 1760–1765.

Reyes, Nicolas, Christopher Ginter, and Olga Boudker (2009). "Transport mechanism of a bacterial homologue of glutamate transporters". In: *Nature* 462.7275, pp. 880–885.

Ritzmann, Noah, Johannes Thoma, Stephan Hirschi, David Kalbermatter, Dimitrios Fotiadis, and Daniel J Müller (2017). "Fusion domains guide the oriented insertion of light-driven proton pumps into liposomes". In: *Biophysical journal* 113.6, pp. 1181–1186.

Roberts, Michael S, Beatrice M Magnusson, Frank J Burczynski, and Michael Weiss (2002). "Enterohepatic circulation". In: *Clinical pharmacokinetics* 41.10, pp. 751–790.

Rouquette-Loughlin, Corinne E, Jacqueline T Balthazar, Stuart A Hill, and William M Shafer (2004). "Modulation of the mtrCDE-encoded efflux pump gene complex of Neisseria meningitidis due to a Correia element insertion sequence". In: *Molecular microbiology* 54.3, pp. 731–741.

Rouquette-Loughlin, Corinne, Steven A Dunham, Michael Kuhn, Jacqueline T Balthazar, and William M Shafer (2003). "The NorM efflux pump of Neisseria gonorrhoeae and Neisseria meningitidis recognizes antimicrobial cationic compounds". In: *Journal of bacteriology* 185.3, pp. 1101–1106.

Ryan, Renae M and Robert J Vandenberg (2016). "Elevating the alternating-access model". In: *Nature structural & molecular biology* 23.3, pp. 187–189.

Šarenac, Tanja M and Momir Mikov (2018). "Bile acid synthesis: from nature to the chemical modification and synthesis and their applications as drugs and nutrients". In: *Frontiers in pharmacology* 9, p. 939.

Schrödinger, LLC (Nov. 2015). "The PyMOL Molecular Graphics System, Version 1.8".

Sekiguchi, Yusuke (2014). "Structural and functional studies of the Apical Sodium Dependent Bile Acid Transporter". PhD thesis. Imperial College London.

Shimamura, Tatsuro, Mitsunori Shiroishi, Simone Weyand, Hirokazu Tsujimoto, Graeme Winter, Vsevolod Katritch, Ruben Abagyan, Vadim Cherezov, Wei Liu, Gye Won Han, et al. (2011). "Structure of the human histamine H 1 receptor complex with doxepin". In: *Nature* 475.7354, pp. 65–70.

Shimamura, Tatsuro, Shunsuke Yajima, Shun'ichi Suzuki, Nicholas G Rutherford, John O'Reilly, Peter JF Henderson, and So Iwata (2008). "Crystallization of the hydantoin transporter Mhp1 from Microbacterium liquefaciens". In: *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* 64.12, pp. 1172–1174.

Shneider, Benjamin L, Paul A Dawson, Donna-Marie Christie, Winita Hardikar, Melissa H Wong, Frederick J Suchy, et al. (1995). "Cloning and molecular characterization of the ontogeny of a rat ileal sodium-dependent bile acid transporter." In: *The Journal of clinical investigation* 95.2, pp. 745–754.

Skou, Jens Christian and Mikael Esmann (1992). "The na, k-atpase". In: *Journal of bioenergetics and biomembranes* 24.3, pp. 249–261.

Sonoda, Yo, Simon Newstead, Nien-Jen Hu, Yilmaz Alguel, Emmanuel Nji, Konstantinos Beis, Shoko Yashiro, Chiara Lee, James Leung, Alexander D Cameron, et al. (2011). "Benchmarking membrane protein detergent stability for im-

proving throughput of high-resolution X-ray structures". In: *Structure* 19.1, pp. 17–25.

South, Paul F, Berkley J Walker, Amanda P Cavanagh, Vivien Rolland, Murray Badger, and Donald R Ort (2017). "Bile acid sodium symporter BASS6 can transport glycolate and is involved in photorespiratory metabolism in Arabidopsis thaliana". In: *The Plant Cell* 29.4, pp. 808–823.

Stamp, Dennis and Gareth Jenkins (2008). "An overview of bile-acid synthesis, chemistry and function". In: *Bile acids: toxicology and bioactivity*, pp. 1–13.

Stephens, David S, Brian Greenwood, and Petter Brandtzaeg (2007). "Epidemic meningitis, meningococcaemia, and Neisseria meningitidis". In: *The Lancet* 369.9580, pp. 2196–2210.

Stieger, BBON, Brigitte O'Neill, and Peter J Meier (1992). "ATP-dependent bile-salt transport in canalicular rat liver plasma-membrane vesicles". In: *Biochemical Journal* 284.1, pp. 67–74.

Stieger, Bruno (2011). "The role of the sodium-taurocholate cotransporting polypeptide (NTCP) and of the bile salt export pump (BSEP) in physiology and pathophysiology of bile formation". In: *Drug transporters*, pp. 205–259.

Stillwell, William (2013). "Membrane transport". In: *An introduction to biological membranes*, p. 305.

Studier, F William (2005). "Protein production by auto-induction in high-density shaking cultures". In: *Protein expression and purification* 41.1, pp. 207–234.

Tanford, Charles (1982). "Simple model for the chemical potential change of a transported ion in active transport". In: *Proceedings of the National Academy of Sciences* 79.9, pp. 2882–2884.

Thanassi, David G, Luisa W Cheng, and Hiroshi Nikaido (1997). "Active efflux of bile salts by Escherichia coli." In: *Journal of bacteriology* 179.8, pp. 2512–2518.

Uzdavinys, Povilas, Mathieu Coinçon, Emmanuel Nji, Mama Ndi, Iven Winkelmann, Christoph von Ballmoos, and David Drew (2017). "Dissecting the proton transport pathway in electrogenic Na+/H+ antiporters". In: *Proceedings of the National Academy of Sciences* 114.7, E1101–E1110.

Vagenende, Vincent, Miranda GS Yap, and Bernhardt L Trout (2009). "Mechanisms of protein stabilization and prevention of protein aggregation by glycerol". In: *Biochemistry* 48.46, pp. 11084–11096.

Velkinburgh, Jennifer C van and John S Gunn (1999). "PhoP-PhoQ-regulated loci are required for enhanced bile resistance in Salmonella spp." In: *Infection and immunity* 67.4, pp. 1614–1622.

Weichselbaum, Anton (1887). *Ueber die Aetiologie der akuten Meningitis cerebrospinalis*. na.

Weinman, Steven A (1997). "Electrogenicity of Na (+)-coupled bile acid transporters." In: *The Yale journal of biology and medicine* 70.4, p. 331.

West, Kristy L, Tosca L Zern, Dustie N Butteiger, Bradley T Keller, and Maria Luz Fernandez (2003). "SC-435, an ileal apical sodium co-dependent bile acid transporter (ASBT) inhibitor lowers plasma cholesterol and reduces atherosclerosis in guinea pigs". In: *Atherosclerosis* 171.2, pp. 201–210.

Wong, Melissa H, Peter Oelkers, Ann L Craddock, and Paul A Dawson (1994). "Expression cloning and characterization of the hamster ileal sodium-dependent bile acid transporter." In: *Journal of Biological Chemistry* 269.2, pp. 1340–1347.

Yan, Nieng (2015). "Structural biology of the major facilitator superfamily transporters". In: *Annual review of biophysics* 44, pp. 257–283.

Yang, Zhengrong, Chi Wang, Qingxian Zhou, Jianli An, Ellen Hildebrandt, Luba A Aleksandrov, John C Kappes, Lawrence J DeLucas, John R Riordan, Ina L Urbatsch, et al. (2014). "Membrane protein stability can be compromised by detergent interactions with the extramembranous soluble domains". In: *Protein Science* 23.6, pp. 769–789.

Yazdankhah, Siamak P and Dominique A Caugant (2004). "Neisseria meningitidis: an overview of the carriage state". In: *Journal of medical microbiology* 53.9, pp. 821–832.

Zhou, Xiaoming, Elena J Levin, Yaping Pan, Jason G McCoy, Ruchika Sharma, Brian Kloss, Renato Bruni, Matthias Quick, and Ming Zhou (2014). "Structural basis of the alternating-access mechanism in a bile acid transporter". In: *Nature* 505.7484, pp. 569–573.