Expression and purification of COMATOSE, a plant peroxisomal ABCD transporter, for functional and structural studies

Jack Wright

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

The University of Leeds

Faculty of Biological Sciences, School of Molecular and Cellular Biology

August 2022

Intellectual Property and Publication Statements

Acknowledgements

Firstly, I would like to thank my supervisors, Professor Alison Baker, Assistant Professor Stephen Muench, and Professor Adrian Goldman, for their expert guidance throughout this project and for their continued support during the tougher times. I am grateful to have had opportunity to work on this challenging yet rewarding project.

A special thank you to Dr David Carrier whose advice and expertise proved invaluable, especially at the beginning of the project, to David Klebl and Joshua White for their fantastic assistance in EM data collection, processing, and troubleshooting, and to all facility managers at the Astbury Biostructure Laboratory (Dan Maskell, Emma Hesketh, Martin Fuller, and Rebecca Thompson).

I am grateful to all members of Astbury level 6, both past and present, from the Muench, Goldman, Henderson, and Bayliss groups, for creating a happy working environment. Thank you to Dr Jamie Henderson for introducing me to the world of climbing, which undoubtedly kept me sane over the last four years; to Jessica Boakes, for always being a source of reason and sensibility; to Andreas Kiessling for our hilarious yet frustrating political discussions, and our insightful scientific ones; to Jannik Strauss and Claudia Stohrer for being fantastic climbing partners and always listening to my ranting; to Dr Steve Harborne for introducing me to sour beer; and a special thank you to Dr Maren Thomsen, who provided the spark for my interest in research science.

Last but certainly not least, I want to thank all my friends and family for their support throughout this journey, without whom this project would not have been possible. Thank you to my friends from home who I sorely miss, to my university friends who make me laugh every day, and to my partner, Aya, for her unwavering belief in me and limitless support during the good times and the bad; I am eternally grateful. And most importantly of all; this thesis is dedicated to the memory of my mother, Victoria Wright, who passed away shortly before its submission.

Abstract

ABC transporters represent one of the largest and most diverse protein superfamilies and are found in all known organisms. They couple the hydrolysis of ATP to the transport of different substrates across various biological membranes reflecting their critical role in many different cellular functions. The peroxisomal family D ABC transporter COMATOSE (CTS), which is expressed in *Arabidopsis thaliana,* is responsible for transporting acyl-CoAs into the peroxisome where they are used in beta oxidation. Importantly, the transporter possesses intrinsic thioesterase activity whereby the acyl-CoA substrates are cleaved during transport across the peroxisomal membrane. While this activity has been demonstrated biochemically, its nature remains unclear due to a lack of structural data on CTS. Reported here is the development and optimisation of purification protocols suitable for structure determination of CTS using cryo-EM, as well as the successful reconstitution of the purified and active transporter into proteoliposomes. Optimisation of the solubilisation and purification of an ATP-hydrolysis deficient mutant ($CTS^{D606N/E607Q}$) enabled the first 3D reconstruction of $CTS^{D606N/E607Q}$, revealing the transporters overall fold from a low-resolution map, and confirmed the steps required for generating a high-resolution structure in future studies. Using the same purification method, the wild-type and mutant were reconstituted into proteoliposomes where CTS showed substrate stimulated ATPase activity, opening the possibility of studying the transporter in a totally isolated system. These results pave the way for future studies on CTS to place the biochemistry of thioesterase activity into a structural framework and understand the transport cycle as well as the transporters broad substrate specificity. This work will synergise with the growing body of work on other family D members, such as human ABCD1 and yeast Pxa1p/Pxa2p, and help to explain early CTS observations in its native *A. thaliana*, thereby contributing to a more comprehensive understanding of this intriguing family of ABC transporters.

Contents Page

Abbreviations

2,4-D	2,4-dichlorophenoxyacetic acid
2,4-DB	4-(2,4-dichlorophenoxy)butyric acid
ABC	ATP-binding cassette
ACOT activity	Acyl-CoA thioesterase activity
ADP	Adenosine-diphosphate
AMN	Adrenomyleoneuropathy
AMP	Adenosine-monophosphate
AMP-PNP	Adenosine 5'-(β,γ-imido)triphosphate
ATP	Adenosine-triphosphate
ATP-γ-S	Adenosine 5'-[γ-thio]triphosphate
CHS	Cholesterol hemisuccinate
CMC	Critical micelle concentration
Cryo-EM	Cryo electron micropscopy
crYOLO	Cr-You Only Look Once
CTF	Contrast transfer function
CYMAL-4	4-Cyclohexyl-1-Butyl-β-D-Maltoside
DDM	N-dodecyl-β-D-Maltoside
DGNG	Decyl glucose neopentyl glycol
DMSO	Dimethlysulfoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
eGFP	Enhanced green fluorescent protein
EGTA	Ethyleneglycol-*bis*-(β-aminoethyl)-N,N,N',N'-tetraacetic acid
EtOH	Ethanol

FC-14	Fos-choline-14
FFA	Free fatty acid
GFP	Green fluorescent protein
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IAA	Indole-3-acetic acid
IBA	Indole-3-Butyric acid
IMAC	Immobilised metal affinity chromatography
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
JA	Jasmonic acid
kDa	Kilo Dalton
LB	Lysogeny broth
LDS	Lithium dodecyl sulphate
LMNG	Lauryl maltose neopentyl glycol
MDR	Multidrug resistance
$MgCl_2$	Magnesium chloride
M_w	Molecular weight
MWCO	Molecular weight cut off
NaCl	Sodium Chloride
NaOH	Sodium hydroxide
NBD	Nucleotide-binding domain
NBS	Nucleotide binding site
OGNG	Octyl glucose neopentyl glycol
PBST	Phosphate-buffered saline with tween
PDB	Protein data bank
pH	Potential of hydrogen
PMP	Peroxisomal membrane protein
PVDF	Polyvinylidene difluoride

ROS	Reactive oxygen species
RT	Room temperature
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide electrophoresis
SEC	Size-exclusion chromatography
SEM	Standard error of the mean
SMA	Styrene maleic acid
TB	Terrific Broth
TBST	Tris-buffered saline with Ttween
T_m	Melting temperature
TMD	Transmembrane domains
Tris-HCl	Tris(hydroxymethyl)aminoethane hydrochloride
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

List of Figures

List of Tables

1 Introduction

1.1 A History of ABC Transporters

The movement of molecules and ions across biological membrane is an essential process in all organisms that in many cases is mediated by membrane transport proteins. ABC transporters represent one of the oldest and largest superfamilies of primary active transporters being found in all kingdoms of life where they utilise the free energy released from the hydrolysis of adenosine triphosphate (ATP) to shuttle a wide-range of substrates across various biological membranes, behaving as importers or exporters. The field of ABC transporter is now vast, with many members of the family being well characterised biochemically and structurally, and with many different models describing their mechanisms. The account below provides a snapshot of the history of this important family of proteins, highlighting key studies and seminal findings.

Studies into ABC transporters began in the 1970's after the discovery of the substrate-binding proteins (SBPs), HisP from *Salmonella typhimurium* (Ames & Lever, 1970) and MalK from *Escherichia coli* (Hazelbauer, 1975) Around the same time, (Berger & Heppel, 1974) demonstrated biochemically that SBP-dependent transport systems derive their energy from the hydrolysis of ATP. In 1976 Juliano & Ling (1976) reported over-expression of a membrane-bound glycoprotein with concomitant altered drug permeability in Chinese hamster ovary (CHO) cells selected for resistance to colchicine. This was the first report of the now much-studied and archetypal ABC exporter known as P-glycoprotein (P-gp) which established the role of ABC transporters in multi-drug resistance (MDR). After cloning and sequencing *mdr1* (the gene encoding P-gp; (Gros *et al*, 1986; Riordan *et al*, 1985) it was clear that P-gp contained six transmembrane domains followed by an ATP binding site, intriguingly, in duplicate. Homology between the amino acid sequences of P-gp and a variety of bacterial transporters (hisP, malK, oppD, and PstB) – with particular respect to the

highly conserved nucleotide binding motifs, Walker A and Walker B (Walker *et al*, 1982) – provided a basis for the deducing the evolutionary origin of ATP binding proteins (Higgins *et al*, 1986). Over the next few years, it became clear that these particular ATP-binding proteins defined a larger superfamily (Ames, 1986; Higgins *et al*, 1990; Hyde *et al*, 1990) which was designated the ABC transporter superfamily. These ABC transporters represent a distinct subset of ABC proteins because not all ABC proteins possess a membrane spanning domain (Hopfner, 2016; Srikant, 2020). In subsequent years, a number of key findings were published that highlighted the importance of this superfamily, not only in terms of MDR, but for their role in various biochemical processes and pathologies.

Identification of the *cftr* gene (Kerem *et al*, 1989; Riordan *et al*, 1989) and its sequencing soon after (Zielenski *et al*, 1991) confirmed that the gene product was indeed an ABC transporter, and proved invaluable in the study of cystic fibrosis (CF), by pinpointing the F508Δ mutation present in ~90% of CF patients. Similarly, Mosser *et al* (1993) suggested that a partially deleted gene found in patients suffering from X-linked adrenoleukodystrophy (X-ALD) resulted in a dysfunctional peroxisomal ABC transporter. Further genetic analysis supported the role of this gene in the pathology of X-ALD (Braun *et al*, 1995). The first model describing the mechanism of transport by P-gp was published in 1992 (Higgins & Gottesman, 1992) and became known as the hydrophobic vacuum cleaner model, whereby the structurally diverse but generally hydrophobic substrates of P-gp accumulate in lipid-bilayers and are flipped by the transporter from the inner-leaflet of the membrane to the outer-leaflet or extracellular space after ATP binding and hydrolysis. Further developments were proposed in a mini-review in 1995 (Senior *et al*, 1995).

The first structural information pertaining to any ABC transporter came in 1997 from low-resolution electron microscopy data of P-gp solubilised from CHO cell membranes (Rosenberg *et al*, 1997). Despite this low resolution, the model generated was consistent with the available genetic and biochemical data (e.g., the flippase

model) and provided a basis for further experiments to delineate ABC transporter mechanisms. The following year the crystal structure of the nucleotide-binding domain from the bacterial histidine transporter HisP was solved at 1.8 Å (Hung *et al*, 1998) which provided the first high-resolution data describing the binding mode of ATP in ABC transporters. Subsequently, the crystal structure of the ABC protein MJ0796 confirmed that the nucleotide binding domains (NBDs) of ABC transporters form a 'sandwich dimer' when binding nucleotide (Smith *et al*, 2002) which is a critical and now well-understood step in the catalytic cycle of all ABC transporters. This landmark paper was crucial in the development of many subsequent transporter models because it outlined the allosteric effects that sandwich dimer formation and ATP hydrolysis might have on the TMDs of ABC transporters.

The first structural insight into a full-length ABC exporter came from the crystal structure of MsbA, a bacterial lipid flippase that was reported at 4.8 Å (Chang & Roth, 2001). Whilst valuable in the early structural studies, the authors later retracted their publication due to technical errors in obtaining the structure solution (see the commentary Matthews (2007)). In their landmark paper, Locher *et al* (2002) reported the first full length crystal structure of an ABC importer, BtuCD, which demonstrated that ABC importers adopt an architecture distinctly different from ABC exporters (comparison with the retracted MsbA structure). By bringing together the genetic, biochemical and structural data that was now available, Higgins & Linton (2004) put forward their pivotal ATP switch model that even today still influences the development of new mechanistic models. Later, the crystal structure of Sav1866, a bacterial homolog of the mammalian P-gp, was solved (Dawson & Locher, 2006) representing the second full-length high-resolution structure of an ABC exporter (but the first full-length and *correctly solved* one).

After completion of the human genome project and publication of the first sequences (Lander *et al*, 2001; Venter *et al*, 2001), humans were found to express 48 different ABC transporters that could be sub-divided in seven families, A – G. With the

availability of sequences and more advanced cloning technologies, ABC transporter research progressed rapidly, resulting in the enormous catalogue of ABC transporter research seen today. Some of the key discoveries in the history of ABC transporter research are highlighted in figure 1-1.

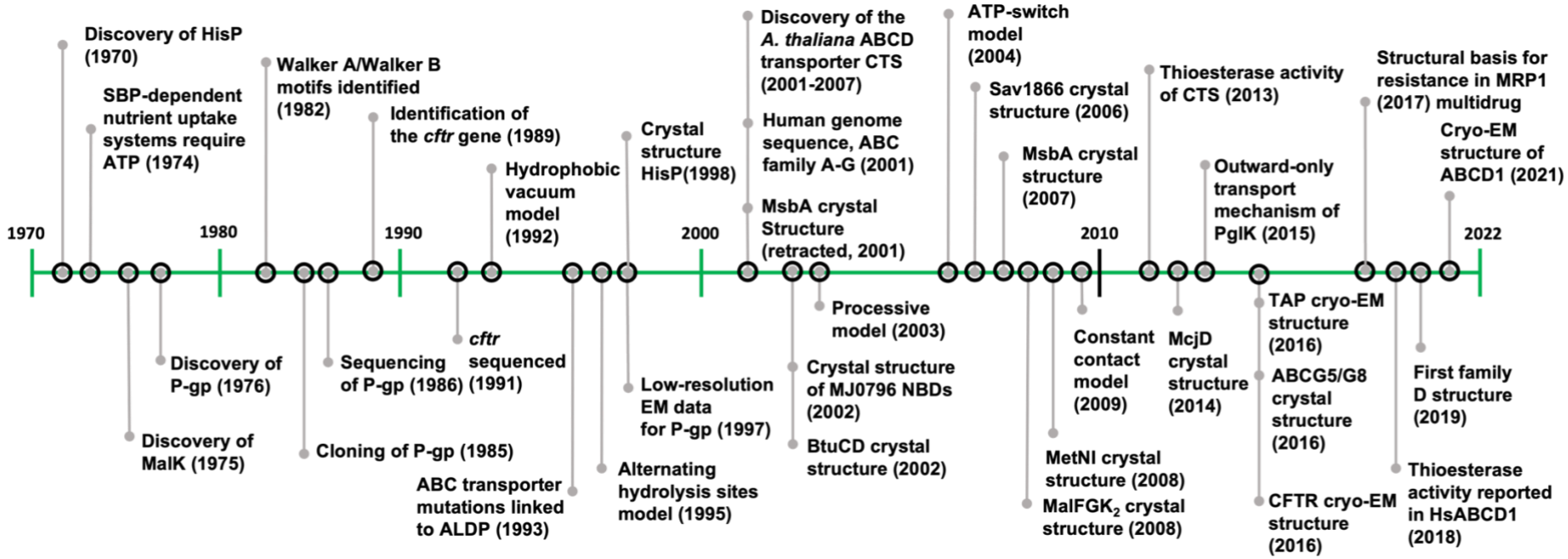


Figure 1-1. A timeline of some of the key discoveries in the field of ABC transporter between 1980 and the present day. All citations can be found in text (Ames & Lever, 1970; Berger & Heppel, 1974; Chang & Roth, 2001; Chen et al, 2021; Choudhury et al, 2014; Dawson & Locher, 2006; De Marcos Lousa et al, 2013; Engelen et al, 2012; Footitt et al, 2002; Hazelbauer, 1975; Higgins & Gottesman, 1992; Higgins & Linton, 2004; Hung et al., 1998; Janas et al, 2003; Johnson & Chen, 2017; Jones & George, 2009; Juliano & Ling, 1976; Kadaba et al, 2008; Lee et al, 2016a; Locher et al., 2002; Mosser et al., 1993; Okamoto et al, 2018; Oldham et al, 2016b; Perez et al, 2015; Riordan et al., 1989; Rosenberg et al., 1997; Senior et al., 1995; Venter et al., 2001; Walker et al., 1982; Wang et al, 2021; Ward et al, 2007; Xu et al, 2019; Zhang & Chen, 2016).

1.2 ABC Transporter Sub-families and Domain Arrangement

ABC transporters may be separated into two discrete categories: ABC importers and ABC exporters. The former move substrates from the extracellular space to the intracellular space, and the latter move substrates from the intracellular space to the extracellular space or from the cytosol into the lumen of organelles (though HsABCD4 may be an exception (Xu *et al.*, 2019)). ABC importers are found almost exclusively in prokaryotes where they primarily mediate the uptake of environmental nutrients such as sugars, amino acids, vitamins, and metal ions (Cui & Davidson, 2011; Fulyani *et al*, 2016; Higgins *et al.*, 1990; Naoe *et al*, 2016; Santos *et al*, 2018). ABC exporters on the other hand are involved in a wide variety of functionalities. For example in prokaryotes they are involved in efflux-mediated antibiotic resistance (Bolhuis *et al*, 1996), secretion of virulence factors (Lewis *et al*, 2012; McDevitt *et al*, 2006; Morgan *et al*, 2017; Omori & Idei, 2003), and bacterial cell wall synthesis (Cuthbertson *et al*, 2010), whereas in eukaryotes they are involved in MDR (Chen & Tiwari, 2011; Johnson & Chen, 2017; Juliano & Ling, 1976; Yin & Zhang, 2011), lipid transport (Footitt *et al.*, 2002; Nyathi *et al*, 2010; Perez *et al.*, 2015; Tarling *et al*, 2013), bile salt and cholesterol export (Berge *et al*, 2000; Štefková *et al*, 2004; Wang *et al*, 2020b), antigen presentation (Nöll *et al*, 2017; Oldham *et al.*, 2016b) and ion transport (Liu *et al*, 2017; Ostedgaard *et al*, 2001).

ABC transporters all adopt a common core architecture consisting of two TMDs spanning the membrane lipid-bilayer and two NBDs located in the cytoplasm (Fig. 1-2). Thus, the family members can be described based upon their domain arrangement (Biemans-Oldehinkel *et al*, 2006). ABC importers typically comprise 4 individually encoded polypeptide chains, that together form a ‘full-size’ transporter, with an additional SBD that is often necessary to deliver substrate to the protein (Bao *et al*, 2013; Hollenstein *et al*, 2007). In ABC exporters the encoded polypeptides can take the form TMD-NBD or NBD-TMD, and are designated as ‘half-transporters’ requiring

dimerisation for function e.g., Sav1886 (Dawson & Locher, 2006). Conversely, they may also take the form TMD-NBD-TMD-NBD or NBD-TMD-NBD-TMD, and are designated as 'full transporters' requiring no dimerisation for function, e.g. Multidrug resistance transporter 1 (MRP1) (Johnson & Chen, 2017) and the cystic fibrosis conductance regulator (CFTR) (Gadsby *et al*, 2006). Moreover, the resulting functional transporters assembled from dimerised 'half transporters' are described as 'homodimers' or 'heterodimers' depending on the sequence identity between the two halves of the protein, i.e., homodimers comprise of two identical 'half transporters' and heterodimers consist of two different 'half transporters'. Another arrangement is the 'fused-heterodimer' that folds from a single polypeptide chain with asymmetry between the two halves of the transporter, e.g., CTS (Baker *et al*, 2015; Dietrich *et al*, 2009; Footitt *et al.*, 2002). In homodimers, both nucleotide binding sites (NBSs) are identical and thus both can hydrolyse nucleotide. Conversely, in heterodimers there can be a deviation between the amino acid sequences of the two halves of the protein, resulting in one consensus NBS (hydrolysable), and one degenerate NBS (nonhydrolysable) (Bienengraeber *et al*, 2000; Hohl *et al*, 2012; Hohl *et al*, 2014; Timachi *et al*, 2017) the presence of one consensus and one degenerate NBS results in a asymmetric transporter (section 1.5.3).

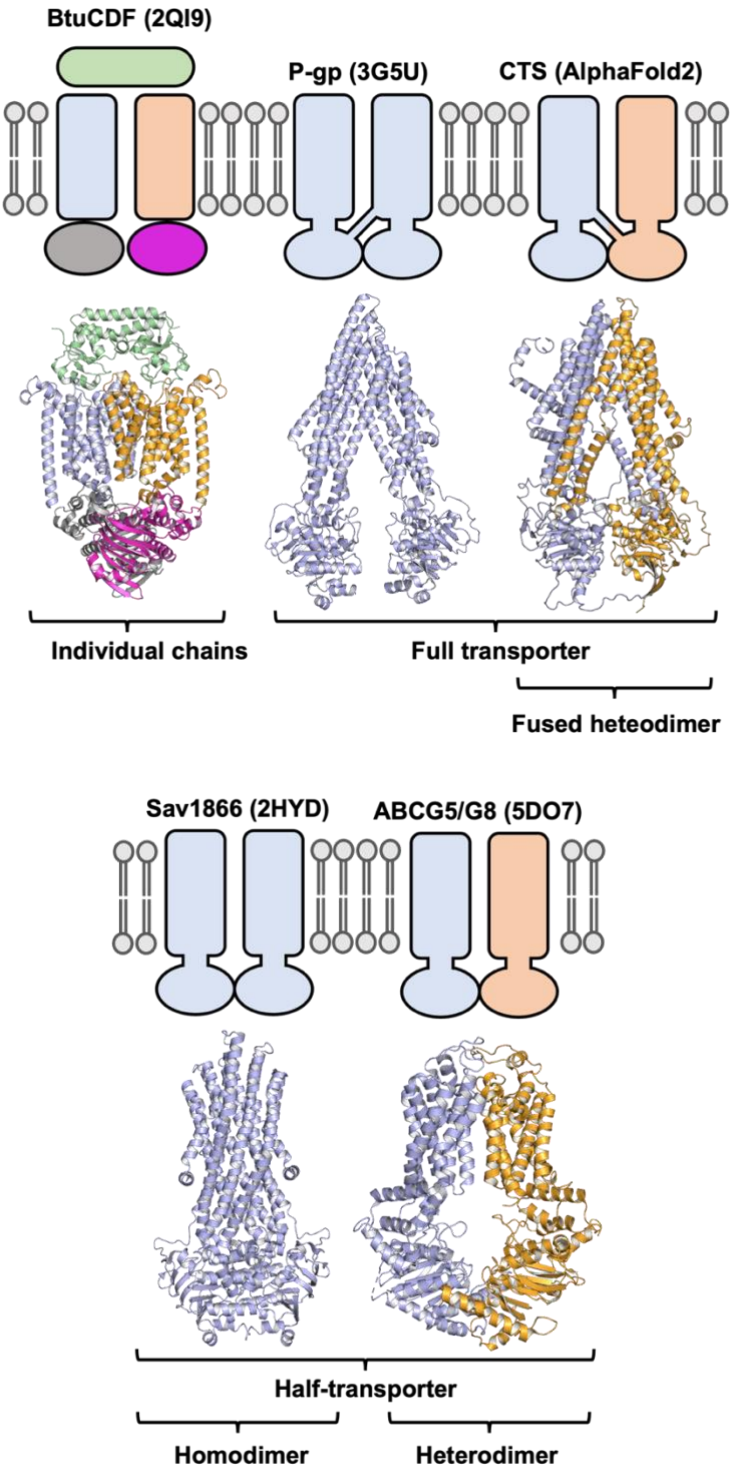


Figure 1-2. Domain arrangement in ABC transporters. The difference between full transporters, half transporters, homo & heterodimers, and importers formed from multiple polypeptide chains is illustrated. PDB IDs are in brackets, except for COMATOSE (CTS) which is the predicted model from AlphaFold2 (Jumper et al, 2021).

1.3 ABC Transporter Classification and Diversity

Given its large size, the ABC transporter superfamily has been categorised into several subcategories. Several different categorisations have been used for this, including the family A-G nomenclature – which was based on NBD and TMD sequence homology, gene structure, and domain order (Dean *et al*, 2001) – the division of ABC exporters vs. ABC importers, e.g. Sav1866 vs. BtuCD, (Dawson & Locher, 2006; Locher *et al.*, 2002), and further still, the different types of exporters, e.g. P-gp vs. ABCG2 (Aller *et al*, 2009; Taylor *et al*, 2017), and the different types of importers (Rice *et al*, 2014).

Whilst this division has been useful for categorising ABC transporters, there can be significant difference between transporters placed into the same category and therefore, a new classification has been proposed. Given the structural data now available for so many different ABC transporters, Thomas *et al* (2020) have proposed a new classification method that is based upon the structural homology of the ABC transporters TMDs. By combining phylogenetic and structural data, this new classification system divides ABC transporters into 7 different groups, each based on their TMD fold. This method – which maintains the previous A-G classification within type IV and V subcategories – can be applied universally across the three domains of life and is therefore able to incorporate newly discovered ABC transporters seamlessly. Whilst classification in this way requires structural data, this nomenclature can be used in addition to previously defined classification methods. The current members of type I-III are exclusively importers, and current members of Types VI and VII possess distinct features are led to their separation into these groups. Therefore, going forward this work will be framed from the perspective of ABC exporters, or more broadly, the newly classified Type IV transporters (though some insights will come from type V members).

As this new classification is based upon TMD fold, Type members span from a range of different organisms with diverse functionality. Within Type IV and V are transporters from humans (still following the A-G nomenclature), various Gram-positive and Gram-negative bacteria, yeast, and plants. Human Type IV transporters include, but are not limited to, the archetypal multidrug resistance protein P-gp (Esser *et al*, 2017), the bile-salt exporter ABCB11 (Wang *et al.*, 2020b), the ATP-sensitive potassium-channel regulators SUR1/2 (Sikimic *et al*, 2018), the lysosomal cobalamin importer ABCD4 (Xu *et al.*, 2019), the chloride ion channel CFTR (Liu *et al.*, 2017; Zhang & Chen, 2016), and the antigen presentation transporter TAP1/2 (Hohl *et al.*, 2012; Lehnert & Tampé, 2017; Oldham *et al*, 2016a). Several bacterial transporters reside within Type IV including another multidrug resistance transporter MsbA (Chang & Roth, 2001; Josts *et al*, 2018) and the peptide transporter McjD (Bountra *et al*, 2017) both from *Escherichia coli*, the lipid flippase Pglk from *Campylobacter jejuni* (Perez *et al*, 2017; Perez *et al*, 2019), the type-1 secretion system component PrtCD (Morgan *et al.*, 2017), and the siderophore transporter IrtAB (Arnold *et al*, 2020) from *Mycobacterium tuberculosis*. To date, the only yeast transporter classified in this way is the mitochondrial iron-sulphur cluster transporter Atm1 (Srinivasan *et al*, 2014). Type V transporters differ from Type IV in that they do not present with domain-swapped architecture. This class includes the promiscuous exporter ABCG2 (Kerr *et al*, 2011; Manolaridis *et al*, 2018; Taylor *et al.*, 2017), its sterol-effluxing G family member ABCG5/G8 (Lee *et al.*, 2016a), the floppases Wzm-WztN (Caffalette & Zimmer, 2021) and TarGH (Chen *et al*, 2020), and the exporter ABCA1 (Qian *et al*, 2017).

To date, the new classification system does not include any plant ABC transporters due to the lack of structural data. Nonetheless, plant ABC transporters possess a wide variety of functionalities; indeed in *A. thaliana* there are over 100 ABC transporters (Sánchez-Fernández *et al*, 2001; Verrier *et al*, 2008) and new plant transporters are being discovered all the time (Do *et al*, 2021). For example, *A.*

thaliana AtABCG28 is a half transporter that contributes to the growth of pollen tubes by transporting polyamines (a source of reactive oxygen species (ROS)) from the cytosol into vesicles that then move and fuse with the growing tip; in this way ABCG28 protects the growing tip from oxidative damage. (Do *et al*, 2019). Plant ABC transporters also play a role in detoxification; AtABCC1 protects from arsenic toxicity by sequestering it as phytochelatin-arsenic conjugates in the vacuole (Song *et al*, 2010). Pathogen resistance is also mediated by ABC transporters, where fungal infection of *Nicotiana benthamiana* is combated by ABCG1- and ABCG2-mediated secretion of capsidiol, a compound that is toxic to fungus (Shibata *et al*, 2016). These examples and tables 1.1 and 1.2 highlight the diversity in ABC transporters and their subfamilies. For an in-depth overview of plant ABC transporters see Do *et al.* (2021). Currently, the newly proposed classification system may be used in conjunction with previously established systems and therefore does not detract from or supersede them now and is not expected to in future. In this way, ABC transporters without any structural data – such as those from plants - can still be classified appropriately depending on the context in which they are discussed.

Subfamily	Transporters	Example Members	Topology	General function	Associated Diseases
A (type V)	ABCA1-12 (full transporters)	ABCA1	ECD-TMD-NBD- ECD-TMD-NBD	Lipid transport	Tangier disease
		ABCA4		Retinal transport	Stargardt disease
B (type IV)	ABCB1-11 (half and full transporters)	ABCB1 (P-gp)	TMD-NBD-TMD-NBD	Multidrug resistance	Various cancers
		ABCB4	TMD-NBD-TMD-NBD	Bile salt transport	Familial intrahepatic cholestasis
C (Type IV)	ABCC1-13 (full transporters)	ABCC7 (CFTR)	TMD-NBD-TMD-NBD	Chloride ion transport	Cystic Fibrosis
		ABCC2	TMD0-TMD-NBD-TMD-NBD	Multidrug resistance	Dubin-Johnson syndrome
D (Type IV)	ABCD1-4 (half transporters)	ABCD1 (ALDP)	TMD-NBD (homodimer)	Acyl-CoA transport	X-ALD
		ABCD4		Cobalamin transport	B12 deficiency
G (Type V)	ABCG1-5 (half transporters)	ABCG2	NBD-TMD (homodimer)	Xenobiotic transport	Variou
		ABCG5/G8	NBD-TMD (heterodimer)	Cholesterol and sterol transport	Sitosterolemia

Table 1.1. Human ABC transporters and their associated diseases reflecting the diversity of ABC transporter functionality. Families E and F have been omitted since the fold of family E is distinctly different from typical ABC transporters, and because family F do not possess TMDs thereby eliminating their role as transporters; thus, family E and F are beyond the scope of this work. Adapted from Vasiliou et al (2009) and Tusnady et al (2006). ECD, extra-cellular domain; TMD, transmembrane domains; NBD, nucleotide-binding domain; TMD0, TM5 extension motif.

Subfamily	Example Members	General Function
A	ABCA9	Lipid transfer into the Endoplasmic Reticulum
B	ABCB1	Transport of auxin and precursors
	ABCB25	Transport of Iron-sulphur cluster precursors
	ABCB14	Regulation of stomatal aperture
C	ABCC5	Phytic acid transport
	ABCC2	Accumulation of defence molecules
	ABCC3	Sequestration of heavy metals and toxic organics
D	ABCD1	Acyl-CoA transport
G	ABCG34	Secretion of defence molecules
	ABCG28	Sequestering polyamines
	ABCG16	Secretion of pollen wall and coat materials
	ABCG29	Secretion of ligin precursors

Table 1.2. Examples of A. thaliana ABC transporter. This exemplifies the huge variety of different cellular process that ABC transporters from different families are involved in. Adapted from Do et al. (2021).

1.4 The Structural Biology of ABC Transporters

1.4.1 The Transmembrane Domains (TMDs)

Unlike ABC transporter NBDs, there is no significant sequence conservation found between the TMDs of transporters with unrelated function, and their remarkable diversity has led to the recognition of seven distinct TMD folds (Thomas *et al.*, 2020). In this classification, types I-III currently represent importers exclusively, whereas Types IV include exporters (including floppases), importers, an ion channel, and regulators, and Type V include both exporters and importers. While members of the Type VI and VII bear resemblance to Type V folds, discrete differences place them into their own Types that are beyond the scope of this work. Examples of Types I-V are highlighted in figure 1-3.

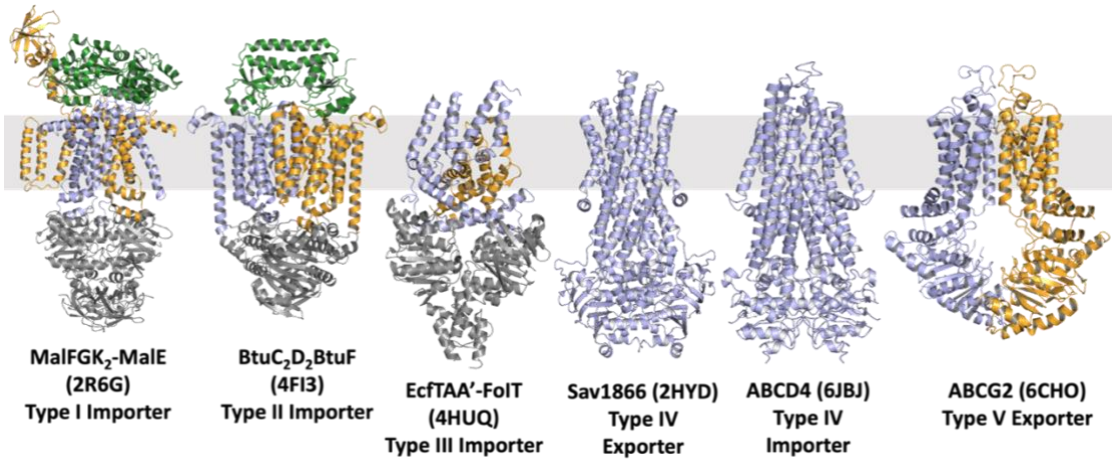


Figure 1-3. Examples of ABC transporter structures classified by their TMD folds. PDB IDs are in brackets. $MalFGK_2$-MalE and $BtuC_2D_2$BtuF are made from individual polypeptides; green, substrate binding proteins; blue and orange, TMDs; grey, NBDs. In EcfTaa'-FoIT; blue and orange, TMDs; grey, NBDs. Sav18866 and ABCD1 fold from a single polypeptide chain, blue. ABCG2 is a half transporter; blue and orange, transporter monomers.

There is considerable variation in the number of membrane-spanning helices found in different transporters, with some possessing as few as ten (Kadaba *et al.*, 2008) and others as many as twenty (Locher *et al.*, 2002). The packing of the membrane-spanning helices within the two TMDs defines the translocation pore – which may contain multiple substrate binding pocket – through which substrates are transported. In general, depending on the stage of the transport cycle, the translocation pore adopts an inward (facing the cytosol) or outward facing conformation (facing the extracellular space). However, it is important to note that other conformations also exist representing intermediate states of the proteins transport cycle that occur between the fully inward and fully outward states. This is exemplified by the outward occluded state of McjD (Choudhury *et al.*, 2014). The different TMD folds, as well as the varying number of membrane-spanning domains in different transporters, certainly reflects the diversity of ABC transporter substrates. So-called coupling helices, short alpha helical regions that sit in the intracellular loops between certain TM helices, connect the TMDs to the NBDs and therefore transmit conformational changes from the NBDs through the TMDs and *vice versa* (Ter Beek *et al*, 2014). Consequently, it is the coupling helices that allow the transporter to transition between inward and outward facing conformations. In most cases, the fold of ABC exporters, which was initially defined by Sav1866 after its crystal structure was solved by Dawson & Locher (2006), often present with a 'domain-swapped' architecture meaning that two helices from one monomer (usually TM4 and 5) of the transporter interact through their coupling helix with the NBD of the opposite monomer (Fig. 1-4).

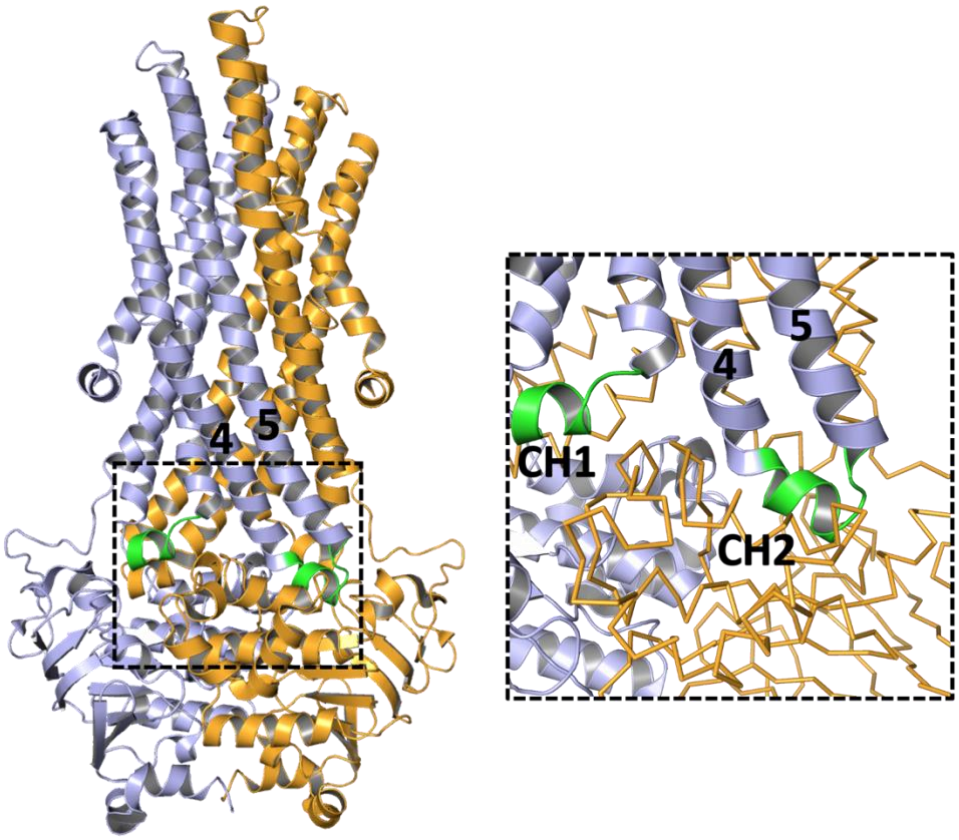


Figure 1-4. The structure of Sav1866 highlighting type IV ABC exporter 'domain-swapped' architecture. The dashed box on the complete structure (left) is expanded on the right, showing TM4 and TM5 of one monomer (purple) interacting via their coupling helix (CH2, green) with the NBD of the opposing monomer (orange, ribbon representation). Adapted from Dawson & Locher (2007). CH, coupling helix.

Unlike Type I-III transporters, which generally have their NBDs close to the membrane, the NBDs of Type IV and V transporters typically extend ~25 Å into the cytosol (though there are notable exceptions e.g., ABCG2) and have been shown in very close proximity (i.e. dimerised), or widely spaced apart (up to ~44 Å, though the physiological relevance of this observation remains controversial (Perez *et al.*, 2015; Wright *et al*, 2018).

1.4.2 The Nucleotide Binding Domains (NBDs)

In all ABC transporters, dimerisation of the NBDs is essential for ATP binding and hydrolysis and is mediated by distinct interactions between several critical motifs. These motifs are distributed across the monomers of the NBDs, which themselves adopt a conserved L-shaped fold comprising two distinct sub-domains: the RecA-like domain (RecA is a recombinase involved in recombination and DNA repair) and a unique alpha helical domain (Ford & Hellmich, 2020). The motifs include the Walker A (or P-loop), Walker B, and signature motifs (or C-loop), as well as the Q-loop, A-loop, D-loop, and the switch region. The crystal structure of the inactive but stable dimer formed by the E171Q mutant of the ABC protein MJ0796 revealed how these motifs interact with each other and with nucleotide (Smith *et al.*, 2002) in a true ABC transporter for the first time (Fig. 1-5).

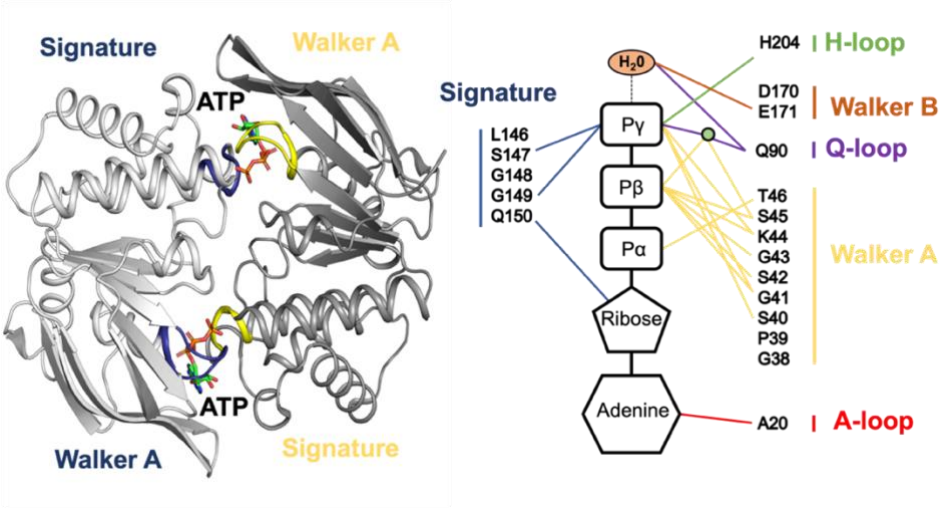


Figure 1-5. The 3D structure of the dimerised NBDs of the E171Q mutant of MJ0796. This diagram illustrates how ATP is sandwiched between the two NBDs. Left, the 3D model of MJ0796. ATP is shown by ball-and-stick representation between the two NBDs interacting with the Walker A motif of one NBD and the signature motif of the other. Right, contact diagram showing the interaction of MJ0796 E171Q residues with ATP. Na^+, green circle. Adapted from Smith et al. (2002).

It revealed the formation of two composite NBSs created by the co-ordination of two ATP molecules sandwiched between the Walker A and Walker B motifs of one NBD monomer and the signature motif and D-loop of the other. This 'head to tail' orientation established the basis for cross talk between the two halves of ABC transporters during the transport cycle. Within the RecA-like domain are the Walker A (G/AXXXXGKT/S, where X is any amino acid) and Walker B motifs (ΦΦΦΦDE, where Φ is any hydrophobic amino acid) (Walker *et al.*, 1982), the former being responsible for binding the γ- and β-phosphates of ATP and the latter coordinating a Mg^{2+} ion and providing the catalytic glutamate necessary for hydrolysis. Within the alpha helical domain is the well-conserved signature motif (typically LSGGQ) responsible for orienting nucleotide for hydrolysis. The highly conserved A-loop ~25-amino acids upstream of the Walker A motif provides an essential aromatic ring capable of stabilising nucleotide by forming π-π stacking interactions with the adenine ring of ATP (Ambudkar *et al*, 2006).

The Q-loop (containing a conserved glutamine), found in the alpha helical domain, is primarily involved in interaction with the TMDs, and is conformationally variable implicating this motif in the coupling of hydrolysis to conformational change in the TMDs (Dalmas *et al*, 2005; Zolnerciks *et al*, 2014). Whilst research into the D-loop – also known as the dimerisation loop – has yielded conflicting results that have not yet been reconciled, it is clear that this motif has a role in coordinating and activating a nucleophilic water molecule, and is thereby critical for inducing a hydrolysis competent state (Jones & George, 2012). Moreover, the D-loops (sequence SALD) of each monomer interact at the dimer interface, as well as with the Walker A of their respective monomers, perhaps mediating allosteric cross talk between NBSs (Zaitseva *et al*, 2005; Zaitseva *et al*, 2006). Whilst some models suggest a glutamate-driven general base mechanism for ATP hydrolysis, others have suggested a substrate assisted mechanism by which the switch histidine motif (or H-loop) behaves as a 'linchpin' that stabilises a transition state between the ATP, water molecules,

Mg^{2+}, and amino acid during the hydrolysis reaction (Locher, 2016; Zaitseva *et al.*, 2005; Zhou *et al*, 2013). A summary of key motifs is provided in table 1.3.

Motif	Consensus Sequence (N-C)	Function
A-loop	F/KxY (Ambudkar *et al.*, 2006)	ATP binding (base stacking)
Walker A	(G/A)XXXXGK(T/S) (Walker *et al.*, 1982)	ATP binding (phosphate binding)
Walker B	ΦΦΦΦDE (Geourjon *et al*, 2001)	ATP hydrolysis
Q-loop	ΦV(S/P)Q (Szollosi *et al*, 2018)	TMD-NBD interaction
Signature	LSGGQ (Schmitt & Tampe, 2002)	ATP binding (phosphate binding) NBD-NBD interaction
D-loop	SALD (Grossmann *et al*, 2014)	NBD-NBD interaction
H-loop	ΦAHRL (Zaitseva *et al.*, 2005)	ATP hydrolysis

Table 1.1. The sequence and function of the conserved motifs in the NBDs of ABC transporters. Φ denoted hydrophobic residues, and X denotes any amino acid. Adapted from Szollosi et al. (2018).

1.5 ABC Transporter Mechanics and Models

1.5.1 Alternate Access

In its most basic form, the catalytic cycle of ABC exporters (or the newly defined Type IV and V transporters) is generally described by the alternate access model first proposed by Jardetzky (1966) for P-type ATPases and the switch model (Higgins & Linton, 2004). These models propose that ABC transporters alternate between inward and outward facing conformational states to move substrates across the membrane. The cycle begins with transporter in the apo- inward-facing state with NBDs separated and substrate capable of binding at the TMDs. Upon substrate binding, the affinity of the NBDs for ATP increases and ATP binding results in the formation of a closed sandwich dimer and a transition state intermediate. Through the coupling helices, dimer formation transmits conformational change to the TMDs causing a transition from the inward- to the outward-facing state and expulsion of the substrate to the opposite side of the membrane. Hydrolysis of ATP and the release of P_i leads to a disengagement of the NBD sandwich dimer and restores the transporter to the basal state. However, the exact order of steps in this generalised model (Fig. 1-6) is not unclear and it does not consider individual transporter characteristics.

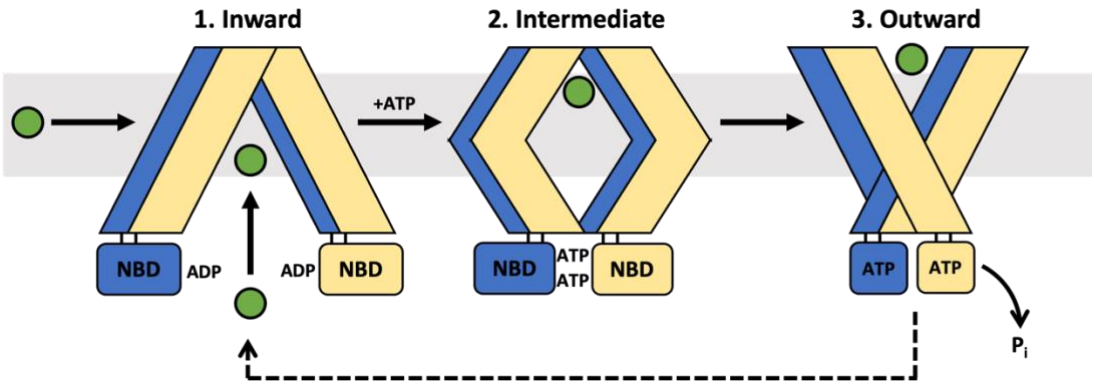


Figure 1-6. The generalised alternate access mechanism of ABC exporters. The transporter beings the inward-facing state ready to accept substrates between its TMDs. Substrate may enter from the intracellular space or from the membrane. Binding of substrate causes closure of the TMDs, and concomitant sandwiching of ATP between the NBDs. Dimer formation causes rearrangement of the TMDs by driving conformational changes through the coupling helices and expulsion of the substrate on the opposite side of the membrane. ATP is hydrolysed and the release of inorganic phosphate drives the dimerised NBDs apart restoring the transporter to the basal state. Blue and yellow, individual transporter monomers; green, generic substrate; grey, membrane. Adapted from Locher (2016).

Consequently, this has led to the development of many different mechanisms, each attempting to describe the specifics of the transporter being studied. Given the variety of models now published describing ABC transporter catalytic cycles and substrate/inhibitor binding modes (Januliene & Moeller, 2020; Johnson & Chen, 2017; Kawaguchi *et al*, 2021; Manolaridis *et al.*, 2018; Nosol *et al*, 2020a; Olsen *et al*, 2020; Perez *et al.*, 2017; Sun *et al*, 2021; Wang *et al.*, 2021), it is evident that transporter nuances eliminate the possibility of a one-size-fits-all model, and therefore, it is beyond the scope of this work to discuss each in detail (some important mechanisms are highlighted in table 1.4). That said, there are several general aspects of ABC

exporters presented below (section 1.5.2 and 1.5.3) that are critical to understanding how these transporters function and are particularly relevant to the work presented in this thesis on the family D ABC transporter, COMATOSE (section 1.7).

Model	Transporter	Features
ATP switch (Higgins & Linton, 2004)	Several in citation	The transporter alternates between inward and outward facing states where conformational change and substrate release is mediated by nucleotide binding. ATP hydrolysis resets the cycle.
Constant contact (Jones & George, 2009)	MJ0796	NBDs are semi-open with one NBS binding ATP while the other is occluded. The NBDs remain in contact while the NBS alternate between binding and hydrolysis.
Outward-only (Perez *et al.*, 2019)	PglK	Flipping of lipid-linked oligosaccharides occurs by substrate loading directly from an outward facing transporter and translocation via a 'substrate hunting' mechanism.
Ion transport (Zhang *et al*, 2018)	CFTR	CFTR functions as an ion channel that is regulated by the phosphorylation of a R domain that sits between the TMDs.
Processive Clamp (Janas *et al.*, 2003)	MdI1p	The association and dissociation of NBDs depends on the status of bound nucleotide at each NBS.
Alternating catalytic sites (Senior *et al.*, 1995)	P-gp	Each binding and hydrolysis of ATP occurs at one NBS during each cycle of the transport process.

Table 1.2. Different ABC transporter mechanisms highlighting key models. Despite the binding and hydrolysis of nucleotide clearly being a universal feature of ABC transporters, the variety of nuances observed for different transporters almost certainly eliminates the possibility of a one-size-fits-all model.

1.5.2 Substrate Promiscuity

Whilst some ABC exporters are very specific in what they transport (Bountra *et al.*, 2017), it is intriguing that many others are able to transport a wide variety of different substrates (McCormick *et al*, 2015). Whilst this action has been well characterised biochemically, the exact nature of this substrate promiscuity was, until recently, poorly understood due to lack of substrate-bound structures. The first real insight explaining this behaviour was gained from the substrate-bound cryo-EM structure of ABCC1, or MRP1 (Johnson & Chen, 2017). MRP1 has been shown to play a role in chemotherapeutic drug resistance (Chen & Tiwari, 2011; Cole, 2014) capable of transporting vincristine, doxorubicin, etoposide, and methotrexate, though this list is not exhaustive. Additionally, the transporter has been shown to transport anti-HIV drugs (Meaden, 2002), antibiotics, anti-depressants (Lee *et al*, 2010), statins (Knauer *et al*, 2010) pro-inflammatory molecules such as leukotriene C4 (LTC_4) and antioxidants such as glutathione (Cole & Deeley, 2006; He *et al*, 2012).

The structure of MRP1 solved by Johnson & Chen (2017) – which is bound to the transporter's physiological substrate LTC_4 – showed for the first time how one transporter is capable of accepting such a wide range of structurally dissimilar amphipathic substrates. Two distinct conformational states were reported, one in the apo-form and one bound to LTC_4 wherein substrate is accommodated in a bipartite cavity i.e., divided into two distinct regions. The positively charged P-pocket coordinates the glutathione moiety of LTC_4 and the hydrophobic H-pocket accommodates the arachidonic acid-derived hydrophobic lipid tail. By comparing the apo and LTC_4-bound structures, it is evident that the binding of LTC_4 causes global conformational changes to the transporter; the NBDs are closer together in the LTC_4-bound structure than in the apo-form, with several residues repositioning their side chains to contact the substrate. LTC_4 therefore acts

as a bridge between the two halves of the transporter, stabilising a conformation in which the NBDs are closer together and aligning them for binding and then hydrolysis of nucleotide i.e., in agreement with the generalised model in figure 1-6.

Moreover, residues of the P-pocket are found in both TMDs of the transporter meaning that substrates lacking a significantly hydrophobic region can still bridge the two halves of the transporter. That residues of the bipartite cavity reposition their side chains upon substrate binding implies MRP1 plasticity that may be extrapolated to other promiscuous transporters, where substrate binding causes twisting and/or unwinding of the TMDs and a tuning of the surface topology of the substrate binding cavity to accommodate a range of amphipathic substrates. The bipartite nature of the MRP1 binding pocket has also been observed in the bacterial exporter MsbA (Lee *et al*, 2013; Mi *et al*, 2017) where the LPS substrate occupies a hydrophobic pocket and a hydrophilic cavity bifurcated by a ring of residues interacting with the LPS glucosamine moiety. Further discussion of this interaction can be found in a review by Wright *et al.* (2018).

The plasticity of the binding cavity in MRP1 has also been invoked to explain the substrate promiscuity of P-gp. Two cryo-EM structures of P-gp in lipid nanodiscs have been recently solved; human P-gp bound to one molecule of Taxol and human-mouse chimeric P-gp bound to two copies of the inhibitor zosuquidar (Alam *et al*, 2019). Interestingly, both molecules occupy the same binding pocket, begging the question of how P-gp distinguishes between stimulatory Taxol and inhibitory zosuquidar. The binding of Taxol reduces the gap between the NBDs thereby priming them for nucleotide binding and hydrolysis. However, comparison between the two structures revealed that there are small but significant structural changes confined mainly to the second half of the transporter that result in an outward shift of coupling helix 1 in the zosuquidar-bound structure. That two molecules of zosuquidar are present contributes to an increase in the distance between the NBDs thereby reducing the ATPase activity of the

transporter explaining the inhibitory action of this compound compared to Taxol. This description of substrate-binding plasticity is line with early studies of P-gp drug-binding (Loo *et al*, 2003; Martin *et al*, 2000b; Martin *et al*, 2001) that describe how substrate binding to the TMDs can induce long range conformational changes to the NBDs thereby affecting transporter activity.

The nature of the substrate binding site(s) of promiscuous transporters can vary significantly e.g. ABCG2 has a flat substrate binding pocket (Taylor *et al.*, 2017) whereas the binding pocket of P-gp exhibits plasticity in its size and shape. Thus, different promiscuous transporters may adopt different modes of substrate specificity. It is important to note that like MRP1, residues from both halves of the P-gp interact with Taxol (Alam *et al.*, 2019). It has been suggested that 3D complementarity of the substrate to the binding site and the strength of the associated contacts is a more important determinant of substrate specificity than distinct substrate binding pockets. This may also be true of MRP1 (discussed above) in that substrates of different shape, charge, and hydrophobicity can adopt a variety of binding modes within the substrate-binding cavity due to local residue side-chain repositioning to bring the NBDs together for nucleotide binding. This model was initially dubbed the 'induced fit mechanism' (Loo *et al.*, 2003) and was also supported by nine structures of P-gp in which the opening and closing of the transporter results in continuous changes to the surface topology of the drug-binding pocket between the TMDs (Esser *et al.*, 2017) thereby creating a space capable of accepting a huge variety of substrates. This is an extension of early work that suggested P-gp contains multiple distinct binding pocket (Martin *et al*, 2000a; Martin *et al.*, 2000b; McCormick *et al.*, 2015; Scala *et al*, 1997).

Multiple nanodisc-reconstituted structures of ABCG2 from the Locher Group have revealed how this transporter accommodates its substrates and inhibitors (Jackson *et al*, 2018; Manolaridis *et al.*, 2018; Orlando & Liao, 2020; Taylor *et al.*, 2017). The substrate binding pocket of ABCG2 forms a narrow slit-like

cavity lined by TM2 and TM5 from the two TMDs that is suitable for accommodating flat, polycyclic, hydrophobic compounds. The cavity is divided into two regions by a 'leucine plug' created from L554 and L554' of opposing monomers, with the two regions bearing differences in relative hydrophobicity. This hydrophobicity difference drives substrate expulsion in a peristatic-like mechanism with NBD closure and ATP binding & hydrolysis resetting the cycle.

The polyspecificity of ABCG2 has been assessed by comparing structure and inhibitor bound structures of the transporter. In the substrate-bound structure, one copy of estone-3-sulphate occupies the binding cavity in one binding mode. Conversely, in the inhibitor-bound structure, two copies of the inhibitor MZ29 are found almost completely filling the space. By comparing the two structures, it is clear that the inhibitor makes many more contacts with the TMD residues than does the substrate. In another inhibitor-bound structure, a single copy of a different and much larger inhibitor, MB136, also fills the cavity entirely making similar contacts to MZ29. Similar to that reported above for P-gp, together these results indicate that substrate size and binding affinity are the deciding factors in which compounds are substrates and which are inhibitors i.e., inhibitors are typically large than substrates and have a greater affinity for the cavity than do substrates. This allows them to behave as a wedge, preventing the opening of the 'leucine plug' thereby eliminating any movement from cavity 1 to 2, locking the transporter in an inward-facing conformation, and ultimately arresting the transporter. This may help to explain how, despite its substrate binding cavity appearing to accept flat polycyclic compounds, ABCG2 can transport many more structurally dissimilar substrates, including tyrosine kinase inhibitors and steroid conjugates (Mo & Zhang, 2012; Wu *et al*, 2019).

1.5.3 NBD Asymmetry

In many heterodimeric ABC exporters, there is asymmetry between the two halves of the transporter (Dietrich *et al.*, 2009; Hohl *et al.*, 2014; Procko *et al*, 2006; Procko *et al*, 2009). This is most apparent at the NBDs where the sequence of the Walker B glutamate, the switch histidine, and signature motifs of one NBD deviates from the consensus sequence resulting in a degenerate NBS incapable of binding or hydrolysing nucleotide. Thus, consensus and degenerate NBSs are not functionally equivalent. Although ATP is almost exclusively hydrolysed at the consensus site (Procko *et al.*, 2009), more recent findings have clearly demonstrated a functional role for NBS degeneracy (Stockner *et al*, 2020). For example, degeneracy in the signature motif of ABC transporters may weaken nucleotide binding thereby reducing ATPase activity; indeed, this has been speculated on with respect to the LSPGE signature motif of ABCD4 (Xu *et al.*, 2019). Despite the nature of NBD degeneracy being unclear, its conservation among different families of ABC transporters clearly highlights its physiological importance.

The apo and AMP-PNP-bound crystal structures of TM287/288 revealed that even in the absence of nucleotide, the NBDs remain in physical contact (Hohl *et al.*, 2014) a finding that was confirmed by DEER spectroscopy measurements (Timachi *et al.*, 2017). Importantly, it was suggested that the D-loop of NBD2 allosterically couple the degenerate and consensus NBSs thereby allowing crosstalk between the NBDs, though the exact role for this coupling is has not been established. These studies also suggest that ATP is likely to remain unhydrolysed at the degenerate site which agrees with the constant-contact model (Jones & George, 2009) and the finding that in CFTR, ATP remains bound to the degenerate site for ~50x longer than the gating cycle of the transporter (Aleksandrov *et al*, 2002; Tsai *et al*, 2010) and that ATP binding to the degenerate site is essential for chloride conductance.

NBD non-equivalence has also been reported for the TAP transporter (Procko *et al.*, 2006) as well as its homolog from *Thermus thermophilus,* TmrAB (Zutz *et al*, 2011) but has also been shown for the homodimeric transporter, MsbA, contributing to the conflicting body of evidence for the role of asymmetry in this transporter (Mittal *et al*, 2012). Whilst a wide body of research exists detailing the observation of NBD asymmetries, the structural basis of its functionality remains to be established before it can be worked convincingly into new mechanistic models.

1.5.4 The Outward-only Model

In a major deviation from the alternate access model, the outward-only model was proposed to describe mechanism of the homodimeric bacterial transporter, PglK, in flipping lipid-linked oligosaccharides (LLOs) across the membrane (Perez *et al.*, 2015; Perez *et al.*, 2017; Perez *et al.*, 2019). In these studies, five structures of PglK were solved: two apo-inward states, one ADP-bound outward-occluded state, one ADP-bound inward state with an inhibitory nanobody, and a semi-occluded ATP-γ-S-bound state. Initially, the outward-only model was proposed based on the first three of these structures.

Given that the NBDs in the apo-inward facing states are extremely widely separated and the millimolar concentrations of ATP in the cell, it was proposed that these two states are irrelevant in the transport cycle. Rather, the LLO polyprenyl tail interacts with the surface of PglK as well as the external helix domain. ATP replaces ADP and the pyrophosphate moiety interacts with the arginine residues on the outward-facing cavity of the ATP-bound state. The remainder of the oligosaccharide then follows the pyrophosphate moiety. ATP hydrolysis drives closure of the TMDs, pushing the substrate into the extracellular space, with the transporter then adopting an the outward-occluded conformation bound to ADP. The energy gained from ATP hydrolysis is used to break the interaction between the pyrophosphate moiety and the arginine residues (Perez et al., 2015). Further developments in understanding Pglk

were made with the ADP-bound inward state bound by an inhibitory nanobody (Perez *et al.*, 2017). Here, nanobody binding to one NBS inhibited hydrolysis at the second NBS, showing how NBD coupling is important in the transport cycle of PglK. In later studies, the outward-only model was shown to be energetically feasible (Perez *et al.*, 2019), and the semi-occluded ATP-γ-S-bound state and MD simulations revealed how the transporter may employ a 'substrate hunting' mechanism to load LLO substrates into the outward-facing cavity. It has been proposed that given the low concentration of LLO in the membrane compared to phospholipids, that 'substrate hunting' may be used to increase the concentration of LLO around the transporters outward/membrane-facing cavity, and in turn, facilitate the movement of the sugar group into the transporter's translocation pathway. Therefore, the processive 'outward-only' mechanism of PglK may be the most energetically feasible method to transport such sparse, large, and complex substrates across the bilayer. This mechanism may be applicable to other flippases.

1.6 An Overview of Family D ABC Transporters

1.6.1 Peroxisomes

Peroxisomes are ubiquitous subcellular single membrane bound organelles present in almost all eukaryotic cells. Whilst originally named for their role in the production of hydrogen peroxide, peroxisomes also carry out several other biochemical processes some of which are organism specific and some of which are universal (Islinger *et al*, 2018; Kao *et al*, 2018; Wanders, 2014). The latter is best exemplified by the highly conserved process of β-oxidation in which fatty acyl-CoA esters are broken down two carbons at a time to yield acetyl-CoA.

In mammals, beta oxidation occurs in both the mitochondria and the peroxisome, whereas in plants and yeast, it occurs solely in the peroxisome. In mammalian beta oxidation, MCFAs and LCFAs enter the mitochondria via interaction with CAT1 and CAT2 on the mitochondrial membranes and following transport are broken down to acetyl-CoA that - among various other biological roles - primarily enters the citric acid cycle. Alternatively, VLCFAs (>22 carbon chain length) are broken down in the peroxisome. These are transported across the peroxisomal membranes by ABC transporters ABCD1-3 (Okamoto *et al.*, 2018) in the form of acyl-CoA esters, which are cleaved during the transport process to into FFA and CoA. It is this thioesterase activity that inspired the work presented in this thesis and is discussed in more detail in section 1.7.2. The acetyl-CoA generate in mammalian peroxisomal beta oxidation is moved back into the mitochondria via the carnitine shuttle for further beta oxidation. Unlike mammals, beta oxidation in plants and yeast occurs solely in the peroxisomes (Fig. 1-7). For example, in *A. thaliana*, acyl-CoA substrates are transported into the peroxisome via the family D ABC transporter COMATOSE (CTS). CTS cleaves the acyl-CoA during the transport cycle as the molecule is moved across the peroxisomal membrane. The acyl-CoA synthetases LACS6 and LACS7 then re-esterify the FFA and CoA back to acyl-CoA where it is then subject to beta-oxidation. The acetyl-CoA

generated here is used in the glyoxylate cycle by conversion to malate via malate synthase, and further conversion into gluconeogenic precursors. In this way, the glyoxylate cycle allows plants and yeast to create sugars from fatty acid thereby allowing them to grow on two-carbon compounds such as acetate (Kunze *et al.*, 2006).

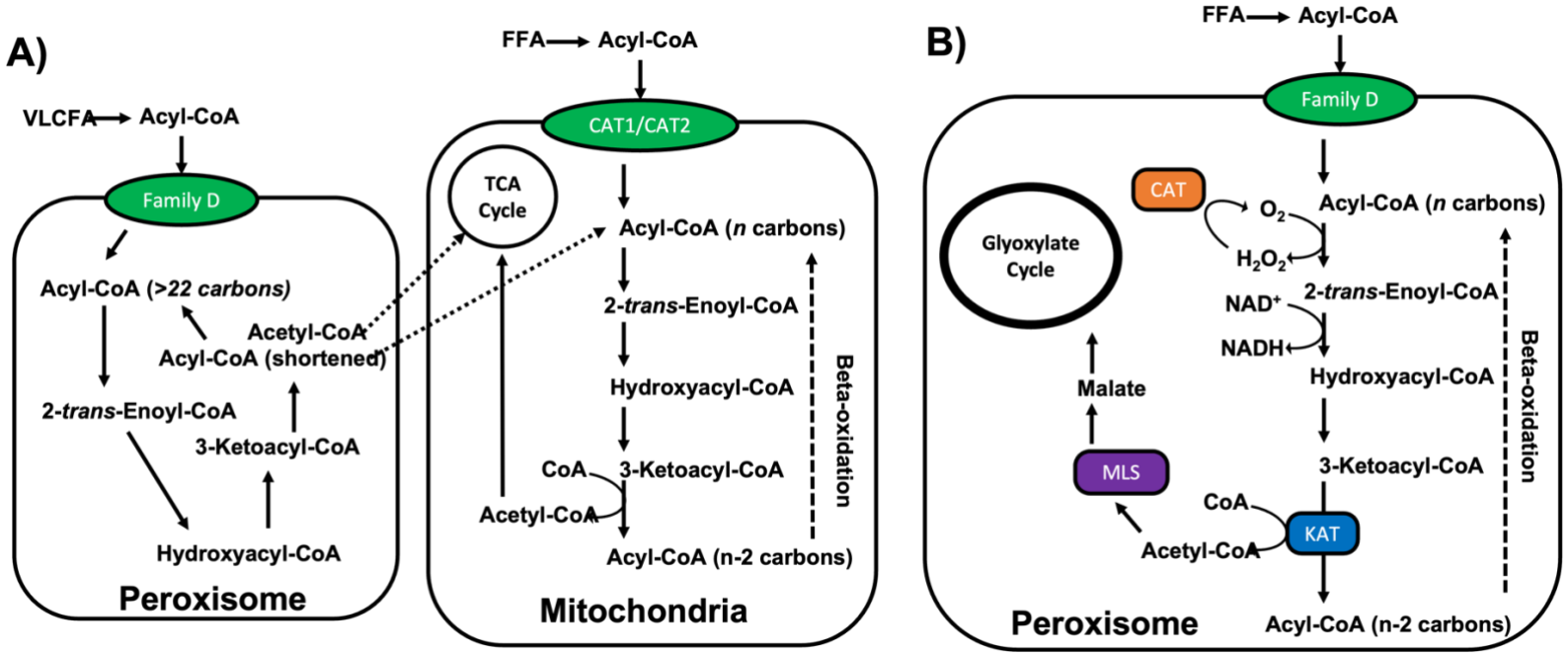


*Figure 1-7. A schematic of beta-oxidation in mammals, plants, and yeast. In mammals (**A**) this process occurs in the mitochondria and peroxisomes. In plants and yeast (**B**) this process occurs exclusively in the peroxisome where acyl-CoAs are broken down two carbons at a time yielding hydrogen peroxide, NADH, and Acetyl-CoA which then enters the glyoxylate cycle. The transport of substrates into the peroxisome is mediated by Family D ABC transporters. Hydrogen peroxide is removed by the action of catalase (CAT, orange), acetyl-CoA is generated from 3-ketoacyl-CoA by the action of 3-ketoayl-CoA thiolase (KAT, blue), and acetyl-CoA is converted to malate by malate synthase (MLS, purple) for entry into the glyoxylate cycle (black circle). Adapted from Baker et al (2006).*

In addition to straight-chained saturated and unsaturated fatty acids, many different substrates can be processed through peroxisomal β-oxidation (Baker *et al.*, 2006; Poirier *et al*, 2006) including jasmonic acid (JA) in plants (Theodoulou *et al*, 2005) and certain prostaglandins and leukotrienes in mammals (Wanders & Waterham, 2006). For these substrates to get into the peroxisome, specific transporters are required. Insertion of these so-called peroxisomal membrane proteins (PMPs) into the peroxisomal membrane is mediated by the PEX proteins, PEX3, PEX16 (SEE1 in *A. thaliana*), and PEX19 (Cross *et al*, 2016). All three PEX proteins are essential in mammals and plants whereas PEX16 is absent in yeast (Mayerhofer, 2016). PEX19 resides in the cytosol where it binds to a broad range of PMPs via their peroxisomal targeting signal (PTS) which can be N- or C-terminally located (Girzalsky *et al*, 2010). This PEX19-PMP complex then docks with PEX3 on the peroxisomal membrane (Fang *et al*, 2004), an interaction which requires farnesylation of PEX19 at its C-terminal CKQQ motif (Emmanouilidis *et al*, 2017; Gotte *et al*, 1998). This trimeric complex of PEX19-PMP-PEX3 then allows for direct insertion of the PMP into the peroxisomal membrane in a thermodynamically favourable, ATP-independent manner (Cross *et al.*, 2016). In plants and mammals PEX16 appears to function as a receptor that accepts the PEX19-PMP-PEX3 complex (Hua *et al*, 2015; Kim & Mullen, 2013).

This is the process by which the substrates involved in beta oxidation cross the peroxisomal membrane. As briefly mentioned, above a significant body of work has determined that it is the family D ABC transporters that are responsible for moving substrates across the peroxisomal membrane for beta oxidation (Baker *et al.*, 2015; De Marcos Lousa *et al.*, 2013; Le *et al*, 2021; Nyathi *et al.*, 2010; Nyathi *et al*, 2012; van Roermund *et al*, 2012; van Roermund *et al*, 2021). Why such specialised transport systems exist in peroxisomal membranes but not plasma membranes is discussed in section 1.7.2. It is these family D transporters, with a specific focus on

the *A. thaliana* transporter COMATOSE (section 1.7), that will be the focus of the remainder of this thesis.

1.6.2 ABCDs in Yeast

In yeast there are two peroxisomal 'half' ABCD transporters (with TMD-NBD topology), Pxa1p and Pxa2p both of which function to transport fatty acyl-CoAs into the peroxisome. This was established by Hettema *et al* (1996) who noted that deletion of Pxa1p or Pxa2p in yeast did not affect growth on media containing glucose, acetate, or glycerol, or the fatty acids laurate or myristate, whereas growth on oleate (C18:1) or palmitate (C16:0) was reduced (a similar finding was noted for Pxa1 by (Shani *et al*, 1995)). Intriguingly, in yeast cells expressing Pxa1p/Pxa2p where the acyl-CoA synthetase, Faa2p – which is responsible for activating MCFA in the peroxisomal lumen – is knocked out, β-oxidation continues at a normal rate, implying that the fatty acyl-CoA esters necessary for β-oxidation can be imported though a different route. When both Pxa1p or Pxa2p *and* Faa2p are knocked out, growth on MCFA and LCFA was abolished, and β-oxidation was severely reduced, implying that there are two different import routes and that MCFA require peroxisomal activation.

It has been shown that the active transporter is a heterodimer composed from one Pxa1p subunit and one Pxa2p subunit (Shani & Valle, 1996; van Roermund *et al.*, 2012). This interaction was later shown to be mediated through the C-terminal region of Pxa2p but not Pxa1p and was impaired through mutations in this region that correspond to equivalent mutations in human ABCD1 (Chuang *et al*, 2014). This finding is in line with earlier work highlighting that ABCD1 homodimerisation – and its heterodimerisation with ABCD2 and ABCD3 – is mediated by the last 87 C-terminal residue of the former (Hillebrand *et al*, 2007; Liu *et al*, 1999).

The active transporter prefers transporting long chain fatty acyl-CoA esters, e.g., C18:1 (Hettema *et al.*, 1996; Shani & Valle, 1996; Verleur *et al*, 1997), with short and medium chain fatty acids entering the yeast peroxisome through a different route

before their esterification to acyl-CoA by Faa2p (above). Moreover, ^{18}O labelling of beta oxidation intermediates indicated that the Pxa1p/Pxa2p complex hydrolyses its acyl-CoA substrates prior to transport into peroxisomes, with subsequence re-esterification through at least functional interaction with acyl-CoA synthetases Faa2p and/or Fat1p (van Roermund *et al.*, 2012). This is in line with a mechanism that first was proposed (Fulda *et al*, 2004) and demonstrated experimental in plants (De Marcos Lousa *et al.*, 2013) and subsequently demonstrated in the human transporters ABCD1-4 (Kawaguchi *et al.*, 2021; Okamoto *et al.*, 2018). Indeed, the *A. thaliana* peroxisomal ABC transporter COMATOSE has been shown to physically interact with the acyl-CoA synthetases LACS6/7. The consequence of this intriguing cleavage mechanism is discussed further in sections 1.7.1 and 1.7.2.

1.6.3 ABCDs in Humans

Humans encode four family D ABC transporters: HsABCD1 (or adrenoleukodystrophy protein, ALDP), HsABCD2 (or adrenoleukodystrophy-related protein, ALDPR), and HsABCD3 (or PMP70) all of which are localised to peroxisomes (Kawaguchi & Morita, 2016) and ABCD4, a cobalamin transporter which is localised to lysosomes (Coelho *et al*, 2012). All three peroxisomal transporters are 'half transporters' with TMD-NBD topology, that after dimerisation function to transport a distinct but overlapping set of different acyl-CoA esters into the peroxisome (Kawaguchi & Morita, 2016). Studies in yeast show a partial rescue of β-oxidation in the yeast knock out pxa1p/pxa2pΔ when transformed with HsABCD1, HsABCD2, or HsABCD3, where D1 preferentially transports C24:0 and C26:0 CoA esters, D2 preferentially transports C22:0, C22:6, and C24:6 CoA esters (all VLCFAs), and D3 preferentially transports long-chain unsaturated-, long branched chain-, and long-chain dicarboxylic CoA esters (Van Roermund *et al*, 2014; Van Roermund *et al*, 2011).

Interestingly, unlike yeast ABCD transporters that appear only to function as heterodimers, results are conflicted regarding the nature of HsABCD1/D2/D3 oligomerisation. Yeast two-hybrid experiments showed the formation of homo- and heterodimers between ABCD1, ABCD2, and ABCD3, (Liu *et al.*, 1999) specifically through their C-terminal halves. Fluorescence resonance energy transfer (FRET) experiments partially corroborated these findings, indicating that ABCD1 and ABCD3 can indeed form heterodimers, but that ABCD1 homodimers predominate (Hillebrand *et al.*, 2007). Moreover, active homo- and heterodimers of ABCD1 and ABCD2 have also been reported (Geillon *et al*, 2014). On the other hand, whilst Tanaka *et al* (2002) report that ABCD1 and ABCD3 stably interact on rat liver peroxisomal membranes, Guimarães *et al* (2004) report that in mouse livers, the same proteins adopt a predominantly homodimeric form. Moreover, the formation of ABCD1/D2 tetramers has also been observed by Geillon *et al* (2017) begging the question of which form – dimer or tetramer – predominates *in vivo*. In this study, anionic calixarene-based detergent C4C8 was used to solubilise mixtures of ABCD1 and ABCD2-eGFP into their constituent monomers and dimers, followed by Co-IP to assess interaction of the solubilised half-transporters. They found that a homo-tetrameric interaction was favoured over hetero-tetrameric interaction but could not provide any explanation for the functional significance of this oligomerisation. Whether tetramerisation has a role in substrate specificity is still unclear.

Intriguingly, in addition to the tetramers, native-PAGE results also reveal the presence of ABCD1/D2 assembles of ~720 kDa. Similar results were previously reported for the *Arabidopsis thaliana* peroxisomal ABCD transporter, CTS (De Marcos Lousa *et al.*, 2013). After solubilisation of *A. thaliana* membranes with β-DDM followed by blue native-PAGE in the first dimension, then SDS-PAGE in the second dimension, CTS was found in a complex of ~700 kDa. Thus, a deeper understanding of oligomerisation of family D transporters may contribute to a better understanding of the transport cycle and/or substrate specific. If oligomerisation is a necessity, then

one may speculate that mutations that abrogate oligomerisation may contribute to the pathophysiology of X-ALD (discussed in section 1.8), paving the way for development of novel treatment strategies.

To date, structural studies on peroxisomal family D ABC transporters are limited to that of ABCD1, with several structures reported in conventional publications (Chen *et al*, 2022; Le *et al.*, 2021; Wang *et al.*, 2021) and two in the BioRxiv (Jia *et al*, 2021; Xiong *et al*, 2021). Whilst these structure are certainly informative, with some being ATP-bound and some bound to substrate, no insights to transporter oligomerisation were reported, though several mechanistic models were presented. It is important to note that despite several reports of ABC transporter oligomerisation, may such studies investigate oligomer formation using protein that, because of solubilisation in detergent, is not in its native membrane environment. Thus, native interactions that lead to oligomerisation may be disrupted, and as a result, the nuances associated with oligomer formation may be overlooked. Given the importance of the lipid bilayer composition in the function of membrane proteins, ABC transporter oligomer formation would be best studied *in vivo o*r by extracting the transporters from their native membranes and retaining native lipids, perhaps using reagents such as SMA co-polymer or Ultrasolute™ amphipols.

ABCD4, unlike other HsABCD transporters, does not possess an N-terminal H0 motif. This motif has been shown to interact with PEX19 (section 1.6.1), thereby facilitating ABC1-3 insertion into the peroxisomal membrane (Lee *et al*, 2014). The absence of the H0 motif in ABC4 means that this transporter is instead localised to the membrane of lysosomes. Recently, the structure of HsABCD4 was solved using cryo-EM (Xu *et al.*, 2019). The 3.6 Å structure was obtained in the ATP-bound outward-facing state displaying a domain-swapped architecture typical of ABC exporters. However, unlike ABC exporters that behave in a *cis*-acting manner, *i.e.,* that allow both substrate and ATP access to the transporter from the

cytosolic side, HsABCD4 appears to work a *trans*-acting manner, effluxing substrate from the lysosome lumen into the cytosol (Coelho *et al.*, 2012).

This is supported by work from Kitai *et al* (2021) who showed that ABCD4 is able to transport cobalamin from the inside to the outside of liposomes in an ATP-dependent manner; thus, ABCD4 is an unusual example of a eukaryotic transporter behaving as an importer. Further still, HsABCD4 possesses a signature motif distinct from the canonical LSGGQ, instead having LSPGE, though the necessity of this deviation is unclear. Like Pxa1/Pxa2 and first reported in the *A. thaliana* transporter CTS (discussed later in section 1.7), ABCD1-4 have been shown to cleave their substrates during the transport cycle, drawing more parallels between the family D transporters. However, Xu *et al.* (2019) do not make reference to this in their paper and provide no explanation for the origin of this activity in a structural framework. Further discussion of this activity is made in section 1.7.2.

1.6.4 ABCDs in Trypanosomes

ABCDs are also found in many Protists, such as in the hydrogenosome of *Trichomonas Vaginalis* (Kay *et al*, 2012), in Leishmania (Sauvage *et al*, 2009), and in Trypanosomes. As the latter's ABCDs are best characterised from Protists, the rest of this section will focus on Trypanosomes.

Trypanosomes – unicellular protozoan parasites transmitted to humans by the Tsetse fly – contain unique organelles known as glycosomes, that have been shown to belong to the same organelle family as peroxisomes (Moyersoen *et al*, 2004), containing the enzymes necessary for β-oxidation, and other peroxisome-associated metabolic process. However, they also characteristically contain the enzymes of the glycolytic and gluconeogenic pathways (Acosta *et al*, 2019). In *Trypanosoma brucei* three ABCD family transporters have been identified, GAT1, GAT2, and GAT3 (Yernaux *et al*, 2006) all of which are 'half-transporters'. Studies on these transporters are rare, however work by Igoillo-Esteve *et al* (2011) has been paramount in their

initial characterisation. Knockdown of GAT1 expression but not GAT3 in procyclic *T. brucei* – i.e., the form that resides in the host fly's midgut containing high concentrations of lipids – is lethal to parasites growing on glucose-free medium thereby indicating a role for GAT1 in lipid metabolism (Igoillo-Esteve *et al.*, 2011).

The same study showed that glycosomes purified from procyclic *T. brucei* cells incorporated radiolabelled oleoyl-CoA in a concentration- and ATP-dependent manner, compared to GAT1 knockdowns in which oleoyl-CoA incorporation was significantly reduced. Together these findings indicated that oleoyl-CoA, and perhaps other fatty acyl-CoAs are substrates of GAT1. To date, the function of GAT2 and GAT3 is unknown. Given the essential nature of least GAT1 to trypanosomes, a deeper understanding of how these transporters works is of high importance due to their apparent similarity to other family D transporters.

1.6.5 ABCDs in Plants

Arabidopsis thaliana encodes just one peroxisomal ABCD transporter, known as AtABCD1 or COMATOSE (CTS), which will be the focus of the rest of this thesis. This transporter possesses many of the ABC transporter nuances discussed earlier, including a unique fused-heterodimer topology, strong substrate promiscuity, non-equivalent NBDs/asymmetry, and the ability to cleave its acyl-CoA substrate during the transport cycle (De Marcos Lousa *et al.*, 2013; Dietrich *et al.*, 2009; Footitt *et al*, 2007).

1.7 COMATOSE

1.7.1 The Biochemical Background of COMATOSE

CTS (AtABCD1) is the only plant peroxisomal ABC transporter and shares 34% sequence identity with its human homolog, HsABCD1 (ALDP). It was initially discovered independently by four different groups (Footitt *et al.*, 2002; Hayashi *et al*, 2002; Hooks *et al*, 2007; Zolman *et al*, 2001) hence its multiple names (CTS, PED3, ACN2, and PXA1). CTS is encoded as a 'full transporter' that folds from a single polypeptide chain – with the arrangement TMD-NBD-TMD-NBD – into a fused heterodimer, where genetic evidence suggests it functions to transport a range of different substrates into the peroxisome for β-oxidation (Footitt *et al.*, 2002). In plants, CTS is involved in several physiological processes including germination, seedling establishment, root growth, and fertility (Footitt *et al.*, 2007; Footitt *et al.*, 2002; Linka *et al*, 2008; Pinfield-Wells *et al*, 2005).

Initially, the *cts* locus in *A. thaliana* was identified as a positive regulator of germination; mutations of this locus prevented the transition from embryo development to germination, and epistatic genetic interactions were observed between *cts* and loci known to regulate embryo maturation (Russell *et al*, 2000). The inability of *cts* mutant seedlings to establish unless provided with an exogenous energy source indicated their inability to metabolise stored lipids (Zolman *et al.*, 2001). This was supported by further work in which *cts* mutants were found to accumulate acyl-CoAs, specifically those of chain length C20:1 and C22:1 (Footitt *et al.*, 2002). This, and confirmation that β-oxidation in *cts* mutants is not impaired indicated that failure to metabolise lipids was due to defects in the transport of substrates into the peroxisome (Footitt *et al.*, 2002). Further to this, resistance of these mutants – as well as *pxa1 and ped3 mutants* – to the root-stunting auxins (2,4-dichlorophenoxyacetic acid (2,4-D) and Indole-3-butyric acid (IBA), both of which are converted to their bioactive forms (4-(2,4-dichlorophenoxy)butyric acid

(2,4-DB) and Indole-3-acetic acid (IAA) respectively) via β-oxidation inside the peroxisome, supports the notion that CTS transports substrates into the peroxisome for beta-oxidiation (Footitt *et al.*, 2002; Hayashi *et al.*, 2002; Zolman *et al.*, 2001). Later work demonstrated that in *cts* mutants the levels of jasmonic acid (JA) are significantly reduced (Theodoulou *et al.*, 2005). Jasmonic acid is part of a class of compound known as jasmonates play a key role in plant defence, metabolism, and development. The synthesis of JA begins in the chloroplast where linolenic acid is converted into an intermediate OPDA. Exported ODPA is then converted into JA inside the peroxisome.

Evidence suggests that CTS contributes to the biosynthesis of JA by non-exclusively mediating the uptake of *cis*-(+)-12-oxo-phytodienoic acid (OPDA), or OPDA acyl-CoA, into the peroxisome (Theodoulou *et al.*, 2005). Interestingly, while previous reports suggested that the 'forever dormant' phenotype of *cts* mutants was the result of defective lipid mobilisation (Russell *et al.*, 2000), more recent work supported the idea that this phenotype results from increased levels of OPDA – i.e., accumulated OPDA that is not transported by CTS and synergises with abscisic acid which is a negative regulator of germination (Dave *et al*, 2011). Together these data strongly suggest a role for CTS in transporting a wide-range of structurally dissimilar substrates into the peroxisome for β-oxidation (Fig. 1-8).

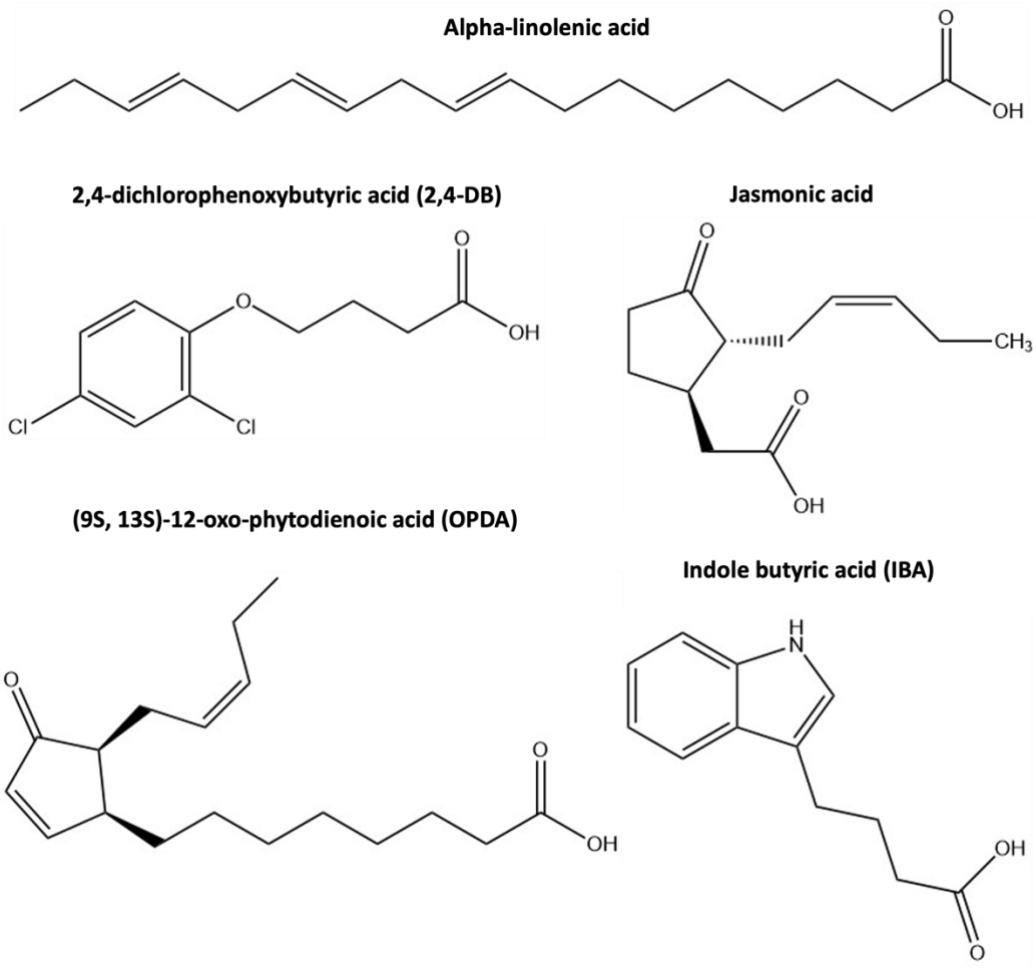


Figure 1-8. The chemical structure of CTS substrates shown as free acids. The true form of these substrates is likely as CoA esters.

From a series of CTS mutants multiple transporter insights been delineated. For example, the NBDs of CTS are functionally distinct. Mutations in the Walker A (K487A) and Walker B (D606N) motifs of NBD1 prevent *A. thaliana* seeds from germinating, whereas the equivalent mutations in NBD2 (Walker A K1136A, and Walker B D1276N) do not (Dietrich *et al.*, 2009). Moreover, mutations in the EAA motif (also known as the coupling helix linking the TMDs to the NBDs) of TMD1 (E300D; contacting NBD2) resulted in a phenotype similar to wild type, whereas mutations in the EAA motif of TMD2 (E954D; contacting NBD1) result in an auxin precursor resistant phenotype (Dietrich *et al.*, 2009). This is in line with reports that ABC transporters sometimes possess two non-equivalent NBDs, the dimerisation of which

results in one consensus NBS and one degenerate NBS (section 1.5.3) where the latter is capable of binding but not hydrolysing nucleotide. For example, the NBS in NBD2 of CTS has been shown not to be crucial in planta, and thus does not appear to have any positive function (Dietrich *et al.*, 2009). However, it may be speculated that the reason CTS possesses this non-functional NBS is due to its interaction with acyl-CoA synthetases; by having only one functional NBS, the rate of nucleotide hydrolysis may be reduced, thereby slowing down the transporter so that the correct acyl-CoA synthease can be recruited to accept the newly cleaved acyl-CoA. Alternatively, given the complexity of some molecules accepted by CTS, it may be that multiple transport cycles are required to move a single molecule of substrate. Perhaps, the non-functional NBS still binds nucleotide but does not hydrolyse it and effectively prevents two NBDs from ever fully separating. When substrate binds, multiple hydrolysis reactions occur at the functional NBS, driving conformational change through the TMDs to processively or peristaltically push substrate through the translocation pathway. By preventing the transporter from opening, the non-functional NBS ensures that the transporter does not adopt an inward facing conformation before substrate has been completely transported.

Whilst capable of germination, *A. thaliana* seedlings bearing the CTS mutation S810N – which is located in helix 2 of TMD2 – cannot establish (Hayashi *et al.*, 2002). The same mutation in CTS expressed in yeast (CTS^{S810N}) results in a marked reduction of oleate breakdown and cannot rescue β-oxidation of oleate in a *pxa1p/pxa2p*Δ mutant (De Marcos Lousa *et al.*, 2013). Interestingly, when expressed in Sf9 insect cell membrane, CTS^{S810N} retains ATPase activity but shows a significant reduction in intrinsic thioesterase activity (Carrier *et al*, 2019; De Marcos Lousa *et al.*, 2013). This activity, which is introduced below and discussed in section 1.7.2, has been shown to be necessary for β-oxidation (Carrier *et al.*, 2019), and thus clearly plays and important yet poorly understood role in the function of CTS during seedling establishment.

Whilst the *A. thaliana* pxa1/ped3/*cts* mutant series has inferred that the transporter can accept structurally dissimilar substrates, strong biochemical evidence has now established that the CTS primarily accepts its substrates in the form of acyl-CoAs (Carrier *et al.*, 2019; Nyathi *et al.*, 2010). Despite this now appearing conclusive, it was not always the case. Conflicting data regarding the nature of the lipids transported by CTS was published, with some reports suggesting that the true substrates were FFAs (Fulda *et al.*, 2004; Linka *et al.*, 2008) and other reporting acyl-CoAs (Footitt *et al.*, 2002; Nyathi *et al.*, 2010; Verleur *et al.*, 1997). Initial studies showed *cts* mutants being defective in lipid mobilisation (Footitt *et al.*, 2002) as demonstrated by their accumulation of acyl-CoAs, specifically those of chain length C20:1, C20:0, and C22:1.

Moreover, it was shown that CTS can complement a *pxa1/pxa2* mutant in yeast, restoring the ability of the mutant to grow on oleate and to metabolise a range of fatty acids (Nyathi *et al.*, 2010). Crucially, the ATPase activity of CTS could only be stimulated by acyl-CoA substrate, and not by FFAs (Nyathi *et al.*, 2010) indicating that the former are the true substrates. This together with the fact that yeast pxa1/pxa2 has been shown to transport acyl-CoAs and not FFAs (Verleur *et al.*, 1997) and the trypanosome transporter GAT1 has been shown to transport acyl-CoAs (Igoillo-Esteve *et al.*, 2011) implies a similar function for CTS, supporting acyl-CoAs as the true substrates.

However, for β-oxidation to occur in *A. thaliana*, long-chain the acyl-CoA synthetases LACS6 and LACS7 must be present. These enzymes are responsible for esterifying transported FFAs to CoA inside the peroxisome, resulting in the prerequisite acyl-CoAs that enter β-oxidation. Fulda *et al.* (2004) showed that in *A. thaliana lacs6/7* mutants present with a similar phenotype to *cts* mutants in that they are defective in lipid mobilisation, seedling establishment, and root growth. These mutants cannot break down fatty acids because the enzyme required for FFA activation to acyl-CoA is not present. Additionally, *A. thaliana* expresses two

peroxisomal adenine nucleotide carrier, PNC1 and PNC2, which are solely responsible for supplying peroxisomes with ATP (in exchange for ADP or AMP). When knocking down the expression PNC1/2 in *A. thaliana*, there is a concomitant increase in the accumulation of fatty acids, likely due to the inability of ATP-dependent LACS6/7 acyl-CoA synthetases to esterify the FFA to acyl-CoA so that they can enter β-oxidation (Linka *et al.*, 2008). Together these finding support FFAs being the true substrates of CTS. To reconcile these apparently contradictory findings, a hypothesis was put forward (De Marcos Lousa *et al.*, 2013) which has more recently been corroborated by two studies on the human homologs of CTS (Kawaguchi *et al.*, 2021; Okamoto *et al.*, 2018).

1.7.2 The Thioesterase Activity of COMATOSE

To reconcile the apparently contradictory evidence regarding transporter of FFA vs. acyl-CoAs a hypothesis was put forward that states the following: *acyl-CoAs are the substrates of CTS, and these substrates are cleaved to FFA and CoA during the transport cycle*. In this model, FFAs are esterified to acyl-CoA outside of peroxisome and are subsequently transported into the peroxisome by CTS (Fig. 1-9). Substrate enters the transporter between the TMDs, followed by nucleotide binding and dimerization of the NBDs. ATP hydrolysis likely provides the power stroke for transport of substrate across the membrane with concomitant substrate cleavage. FFA is released into the peroxisomal lumen where is re-esterified back to acyl-CoA by acyl-CoA synthetases LAC6/7 for activation and entry into β-oxidation.

While this thioesterase hypothesis was originally formulated by Fulda *et al.* (2004), indirect evidence supporting the hypothesis was reported in yeast where VLCFA-CoAs were apparently hydrolysed by the Pxa1/Pxa2 complex prior to transporter of the FFA moiety into the peroxisome, and that Pxa1/Pxa2 functionally interacts with the acyl-CoA synthetases Fat1p and Faa2p on the inner surface of the peroxisome, presumably to accept the FFA product of cleavage for re-esterification (van

Roermund *et al.*, 2012). Direct evidence was provided by De Marcos Lousa *et al.* (2013) who showed that CTS makes a physical interaction with LACS6/7 in *A. thaliana*, but when expressed in Sf9 insect cell membranes in the absence of LACS6/7, possesses ATP stimulated thioesterase activity. More recently, thioesterase activity has also been reported for HsABCD1 as well as HsABCD2, D3, and D4 showing that this activity is present in both symmetrical and asymmetrical ABCD transporters (Kawaguchi *et al.*, 2021; Okamoto *et al.*, 2018). This acyl-CoA cleavage process has been designated 'intrinsic thioesterase activity' (acyl-coA thioesterase activity (ACOT activity)). A schematic showing the proposed mode of acyl-CoA transport by CTS in *A. thaliana* is shown in figure 1-9.

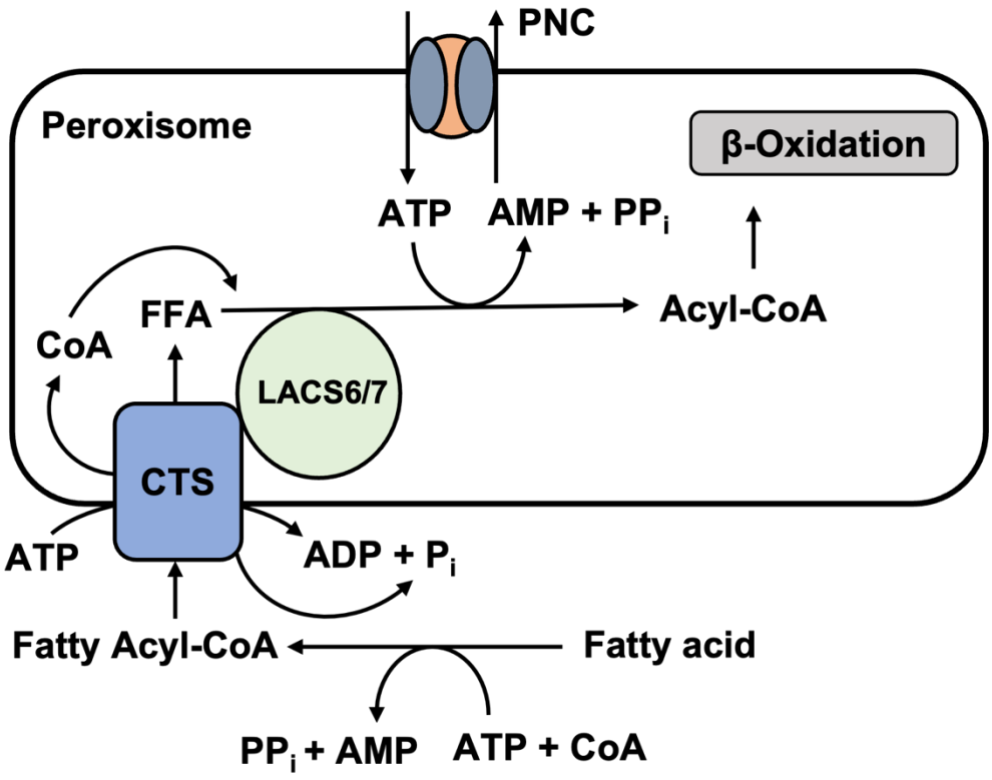


Figure 1-9. A diagram of the proposed model for the transport of acyl-CoAs into peroxisomes via CTS. CTS accepts fatty acyl-CoAs on the cytosolic side of the peroxisomal membrane and cleaves the substrate during the transporter cycle via an unknown mechanism and free fatty acids are released into the peroxisome. Free fatty acids are re-esterified back to acyl-CoA by LACS6 and LACS7 before undergoing beta-oxidation. After substrate cleavage, studies by van Roermund et al. (2021) in yeast have shown that the CoA moiety enters the peroxisomal lumen along with FFA. Adapted from De Marcos Lousa et al. (2013).

Whilst intriguing, the nature of this activity in CTS remains unclear because CTS has no homology to any known thioesterase and does not possess any accessory domains. There exists two types of acyl-CoA thioesterase, type I and type II. Type I belong to the α/β-hydrolase family, whereas the latter belong to the 'hot-dog' fold family. Type I thioesterase are not found in plants, yeast, or insect cells, whereas type II are ubiquitous, begging the question of whether CTS could have evolved a

similar thioesterase catalytic triad as in type II acyl-CoA thioesterases albeit within a different fold (Carrier *et al.*, 2019). Using a CTS homology model (Dietrich *et al.*, 2009) and searching for similar residues to those making up the catalytic triad of type II thioesterases, Carrier *et al.* (2019) identified three residues that are potentially involved in CTS substrate cleavage; D863, Q864, and T867 all of which are located on helix 9 (Fig. 1-10). By mutating these residues to alanine and assessing ATPase and thioesterase activity, it was demonstrated that thioesterase activity was reduced by all mutations and that thioesterase activity is dependent on ATPase activity but not *vice versa*. Crucially, when these CTS mutants were expressed in Pxa1/Pxa2Δ yeast cells, the loss of thioesterase activity also resulted in a loss of β-oxidation, indicating that substrate cleavage is essential for transport of substrates. However, it should be noted that while these residues could play a catalytic role, they may instead be involved in the conformational dynamics of the transporter.

Other residues involved in the transport cycle are those that bind but do not hydroylse substrate. Given the complexity and breadth of substrates transported by CTS and given that these acyl-CoA substrates are cleaved during transport, it would appear likely that there are many different substrate binding sites within the TMDs of CTS. Alternatively, TMD rearrangement during the transport cycle may provide plasticity that allows the transporter to accommodate a range of substrates (like that discussed for MRP1 and P-gp in section 1.5.2), by tuning the topology of the internal cavity. Indeed, the internal cavity would have to accommodate not only the fatty acid moiety of the substrate, but also the CoA moiety; the amphipathic nature of such compounds and therefore their potentially wide range of binding modes, will render the substrate recognition process incredibly complex. Finally, the ratio between CTS ATPase and thioesterase activity has been reported at around 6:1 (Carrier *et al.*, 2019; De Marcos Lousa *et al.*, 2013), similar to that reported for human ABCDs. This may reflect the complexity of transporting complex amphipathic substrates, if the transporter behaves in a processive manner as speculated on earlier (section 1.7.1).

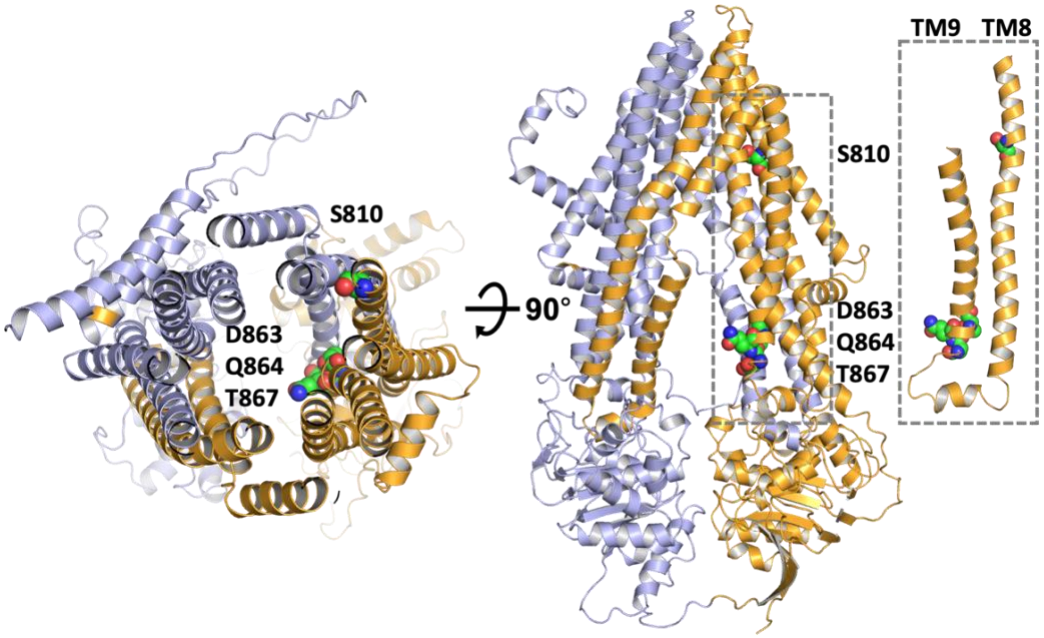


Figure 1-10. The predicted model of CTS generated by AlphaFold2. Residue S810 (TM8; section 1.7.1), and the putative triad residues D863, Q864, and T867 (TM9) highlighted as spheres. The dashed box is an expanded view of TM8-9. Adapted from Carrier et al. (2019).

1.8 X-Linked Adrenoleukodystrophy (X-ALD)

Defects in peroxisomal function are perhaps best exemplified by X-linked adrenoleukodystrophy (X-ALD). This is the most common inherited peroxisomal disorder (~1:21,000 males) caused by mutations in HsABCD1 resulting in defective β-oxidation and accumulation of VLCFAs (particularly C26:0) throughout the body (Wiesinger *et al*, 2013) most notably in the central nervous system, adrenal cortex, and testicular Leydig cells. Interestingly, expression of the disease varies widely, with multiple phenotypes being described and no reliable way to predict which form of the disease an individual will develop. Two phenotypes predominate: the milder adrenomyeloneuropathy (AMN; observed in both males and females) and the more severe cerebral ALD (CALD; typically observed in males). Crucially, there appears to be no correlation between mutations in the ABCD1 gene and disease severity. This

is best exemplified by the identification of 5 different phenotypes in 6 different males of the same family having the same destabilising missense mutation, P484R (Berger *et al*, 1994) and that monozygotic twins suffering from X-ALD have been described bearing entirely different clinical symptoms (Korenke *et al*, 1996). In addition, even complete loss of ABCD1 due to large deletions or frameshift mutations can result in the milder, AMN, form of the disease.

Despite at least 643 ABCD1 mutations having been identified (Kemp *et al*, 2001), the aetiology of X-ALD, specifically the demyelination of nerve cells, is still unclear. The accumulation of VLCFA results in a range of pathologies, the results of which have been outlined extensively in other work (Berger *et al*, 2014; Engelen *et al.*, 2012; Kemp *et al*, 2012). Whilst this list is non-exhaustive, with novel biochemical defects still being discovered (Petroni *et al*, 2019), the obvious categories into which these defects can be placed demonstrates how catastrophic the accumulation of VLCFAs can be, and helps in understanding the lack of a genotype-phenotype relationship.

Whilst many X-ALD mutations lead to a loss or reduction in the level of ABCD1 that is expressed, the pathophysiology of some phenotypes may be explained by the thioesterase activity of family D ABC transporters. Whilst HsABCD1 is a homodimer and the plant homolog CTS is a fused heterodimer, sequence identity between the two proteins is ~36% and both transporters have been shown to possess ATPase-dependent thioesterase activity (De Marcos Lousa *et al.*, 2013; Okamoto *et al.*, 2018). That a loss of thioestase activity from CTS blocks β-oxidation implies that it is an essential step in the transport cycle, and therefore, mutations in HsABCD1 that abrogate or reduce its thioesterase activity may contribute to the array of biochemical defect observed in X-ALD. If true, one could argue that, as hypothesised for CTS (De Marcos Lousa *et al.*, 2013), a defect in the thioesterase activity may lead to a failure of the peroxisome to recruit the correct acyl-CoA synthetases. This limits the peroxisome's ability to organise its metabolic needs and in turn, this spatial and

temporal disorganisation of metabolism may lead to a devastation of overall homeostatic cell signalling resulting in the biochemical defects expressed in table 1.5. Interestingly, residues in CTS that appear to contribute to the proteins ATPase-dependent thioesterase activity (D863, Q864, and T867) have been identified at the equivalent positions in HsABCD1 in patients suffering from X-ADL (Carrier *et al.*, 2019) potentially reinforcing the importance of this activity in normal physiological function.

Pathology	Observation	Reference
Oxidative Stress	Excess VLCFAs and/or siRNA knockdown of ABCD1 in murine oligodendrocytes induces oxidative stress.	(Baarine *et al*, 2012)
	ABCD1-knckout mice possess proteins modified by oxidative stress because of damage by accumulated malonaldehyde in the spinal cord.	(Fourcade *et al*, 2008)
Membrane defects	Turnover of C26:0 from model phospholipid vesicles is orders of magnitude slower than for shorter chain fatty acids and calorimetric studies showed that inclusion of C26:0 in a model membrane disrupts the overall membrane stability.	(Ho *et al*, 1995)
	Increased micro-viscosity observed in the membrane of adrenocortical cells exposed to VLCFA (C24:0 and C26:0) and in erythrocytes of X-ALD patients.	(Knazek *et al*, 1983; Whitcomb *et al*, 1988)
Energy-generation defects	Exposure of murine astrocytes to C22:0, C24:0, and C26:0 causes mitochondrial depolarisation and impairs the electron transport chain.	(Kruska *et al*, 2015)
	Oxidative damage to glycolysis and TCA cycle enzymes in the spinal cords of ABCD1 knock out mice and pyruvate kinase in X-ALD fibroblasts, with concomitant reduction in ATP and NAHD concentrations.	(Galino *et al*, 2011)
Cell death and stress	Excess VLCFAs causes death of astrocytes and myelin-producing oligodendrocytes.	(Hein *et al*, 2008)
	C24:0-lysoPC but not C16:0-lysoPC causes activation of microglial cell and apoptosis in the cortex of wild-type mice indicating cytotoxicity is chain-length dependent.	(Eichler *et al*, 2008)

Table 1.3. Biochemical defects associated with X-ALD divided by pathology type.

1.9 Aims and Objectives

Whilst the ACOT activity of CTS has been confirmed through several biochemical studies and corroborated by studies on its human and yeast homologs, there is no definitive explanation of exactly how the transporter is capable of substrate cleavage. This is primarily because of a lack of structural information pertaining to family D ABC transporters, and because previous studies have relied upon heterologous expression of CTS for biochemical investigation. Therefore, to address this research gap, the overarching aim of this study was to determine the 3D structure of substrate-bound CTS and/or CTS mutants using cryo-EM, to elucidate the nature of CTS ACOT at a molecular level, and to reconstitute active CTS into proteoliposomes to be used as a tool for future biochemical studies. In determining the structure of CTS, it would be possible to identify the residues involved in substrate binding and/or those that have a role in thioesterase activity and gain novel insights into the CTS transport cycle. This information could then be used to place previous biochemical findings into a structural framework. Similarly, the structural data would be useful in directing new biochemical studies that become possible through CTS reconstitution, thereby providing a total isolated system to investigate specific questions. The following objectives were considered in order to achieve the overall aim and gaining a mechanistic understanding of CTS ACOT activity:

1. To optimise the expression, solubilisation, and purification of CTS and/or CTS mutants in pursuit of a sample of sufficient quality and quantity for cryo-EM studies. This includes optimisation of large-scale CTS expression, determination of optimal solubilisation conditions using detergent and non-detergent alternatives, and optimisation of buffer composition for affinity and SEC purification.
2. To collect a sufficiently large cryo-EM data set on purified CTS and/or CTS mutants that could be used to generate a 3D reconstruction of the transporter.

3. To reconstitute CTS and/or CTS mutants into proteoliposomes and measure the proteins transport activity. This work will form the basis of future biochemical studies on CTS such as transport assays.

These objectives outline the hypothesis of this study, which is that upon obtaining a high-resolution model of substrate-bound CTS it will be possible to precisely identify residues involved in substrate binding and cleavage, and that such data will complement future biochemical studies such as transport assays to build up a picture of the catalytic cycle of the transporter.

2 Materials and Method

2.1 Materials

2.1.1 Bacterial strains, plasmids, Sf9 cells, growth media

Three bacterial strains were used in this study: OmniMAX™, BL21 Gold (DE3), and DH10Bac™ competent cells. OmniMAX™ and BL21 Gold (DE3) cells were routinely cultured in LB media or on LB agar plates supplemented with appropriate antibiotics, and DH10Bac™ cells were cultured in LB media containing kanamycin (50 µg/ml), tetracycline (10 µg/ml), and gentamycin (7 µg/ml; from here known as DH10 media) or LB agar plates supplemented with kanamycin (50 µg/ml), tetracycline (10 µg/ml), and gentamycin (7 µg/ml), IPTG (0.16 mM), and X-Gal (100 µg/ml; from here known as DH10 agar plates). LB media was prepared by dissolving 10 g tryptone, 10 g NaCl, and 5 g yeast extract in a total volume of 1L of MilliQ water and autoclaving. For LB agar, 15 g of agar was added per litre of LB media. For large scale expression in BL21 Gold (DE3), Terrific Broth (TB) was used. TB was prepared by dissolving 20 g tryptone, 24 g yeast extract, 4 ml glycerol, and 100 ml filtered sterilised phosphate buffer (170 mM KH_2PO4, 720 mM K_2HPO_4 in a total volume of 1L of MilliQ water and autoclaving. The plasmids used in this study are listed in table 2.1. Sf9 insect cells were purchased from ThermoFisher Scientific and routinely cultured to no greater than 50 passages in either Gibco™ Sf-900 serum-free media (Thermofisher Scientific) or Insect-XPRESS™ protein-free insect cell media with L-glutamine (Lonza). There was no noticeable reduction in protein quality or yield when using high passage cells.

Plasmid	Resistance	Information	Source
pFastBac-CTS-eGFP-StrepII	Carbenicillin	Wild-type CTS with a C-terminal eGFP and double streptag.	Dr David Carrier
pFastBac-CTS_K487A-eGFP-StrepII	Carbenicillin	Lysine to alanine mutation in the Walker A motif. This mutation reduces the affinity of the NBDs for ATP.	Dr David Carrier
pFastBac-CTS_DNEQ-eGFP-StrepII	Carbenicillin	Aspartic acid to asparagine and glutamic acid to glutamine in the Walker B motif at positions 606 and 607, respectively.	Dr David Carrier
pMSP1D1	Kanamycin	Membrane scaffold protein 1D1.	Sligar Lab (USA)
pMSP1E3D1	Kanamycin	Membrane scaffold protein 1E3D1.	Sligar Lab (USA)

Table 2.1. Plasmids used in this study.

2.2 Chemicals, reagents, and antibiotics

Most chemicals used in this study were purchased either from Sigma-Aldrich or Melford, with the exception of detergents, most of which were purchased from Anatrace and lipids which were purchased from Avanti® Polar Lipids.

2.2.1 Buffers

Buffers used routinely throughout this study are listed in table 2.2. All buffers were filtered sterilised prior to use (0.45 µm or 0.22 µm filters, Merck Millipore).

Buffer	Components
Homogenisation buffer	20 mM Tris, pH 7.4 or 20 mM HEPES, pH 7.5, 0.5 mM $MgCl_2$, 5 mM EGTA, 1X Proteoloc™ protease inhibitors (Expedeon), 1 mM AESBF, 1 mM DTT
Buffer A	20 mM Tris-HCl, pH 7.4, 200 mM NaCl, 20% Glycerol (v/v), 1.5 mM $MgCl_2$
Solubilisation Buffer A	20 mM Tris-HCl, pH 7.4, 200 mM NaCl, 20% Glycerol (v/v), 1.5 mM NaCl, 1% DDM/0.2% CHS/0.2% asolectin (w/v), 1 mM PMSF, 1X Proteoloc™ protease inhibitors (Expedeon)
Wash Buffer A	20 mM Tris-HCl, pH 7.4, 200 mM NaCl, 5% Glycerol (v/v), 1.5 mM NaCl, 0.1% DDM/0.2% CHS/0.2% asolectin - 0.05% DDM/0.002% CHS/0.002% asolectin (w/v)
Buffer B	20 mM HEPES, pH 7.5, 200 mM NaCl, 20% Glycerol (v/v), 2 mM $MgCl_2$
Solubilisation Buffer B	20 mM HEPES pH 7.5, 200 mM NaCl, 20% Glycerol (v/v), 2 mM $MgCl_2$, 1 mM DTT, 4 mM ATP, 1% LMNG/0.1% CHS (w/v), 1X Proteoloc™ protease inhibitors (Expedeon), 1 mM AEBSF
Wash Buffer B	20 mM HEPES pH 7.5, 150 mM NaCl, 5% glycerol (v/v), 2 mM $MgCl_2$, 0.1% - 0.06% digitonin (w/v)
Proteoliposome Buffer	50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$
PBS	100 mM NaCl, 2.7 mM KCl,10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4
PBST	PBS + 0.1% Tween 20
TBST	50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% Tween 20 (v/v)

Table 2.2. Buffer solutions used routinely in this study. Tris buffer was used if the protein was to be purified using DDM. HEPES buffer was used if the protein was to be purified using LMNG/digitonin.

2.3 Molecular Biology Techniques

2.3.1 Sequencing of Plasmids

15 µl of each plasmid (50 ng/µl) was mixed with 2 µl of the appropriate primer (10 µM) in Mix2Seq tubes (Eurofins Scientific) and sent to Eurofins Scientific for Sanger sequencing.

2.3.2 Transformation of Plasmid DNA into OmniMAX™ Competent Cells

50 µl of OmniMAX™ competent cells were thawed on ice and combined with 1 µl of vector (100-1000 ng/µl) in a pre-chilled 1.5 ml tube. The mixture was incubated on ice for 20-30 minutes, then heat-shocked at 42°C for 45-60 seconds. The mixture was returned to ice for 2-3 minutes, and the cells recovered by addition of 500-1000 µl LB media and incubation at 37°C (1h, 220 revolutions per minute (RPM)). Cells were spread on LB agar plates supplemented with the appropriate antibiotic.

2.3.3 Purification of Plasmid DNA

Overnight cultures of OmniMAX™ cells were prepared by inoculation of 10 ml of LB media (supplemented with appropriate antibiotics) with a single colony from a LB agar plate onto which transformed cells has previously been spread. The culture was then incubated at 37°C overnight (~16 h, 220 RPM). From here, plasmid purification was carried out using a NucleoSpin™ plasmid purification kit (Macherey-Nagel™) according to the manufacturer's instructions.

2.3.4 DNA Concentration Determination

Plasmid DNA concentration was measured using DS-11 spectrophotometer (Denovix®). The instrument was blanked by applying 1 µl of DNA sample buffer and measuring the absorbance at 260 nm. After the blank measurement, the sample concentration was measured in the same way and the concentration reported in ng/µl.

2.3.5 Generation of Bacmids

2.3.5.1 DH10Bac Transformation and Blue/White Screening

50 µl of DH10Bac™ cells or DH10Bac™ carrying a YFP reporter sequence (provided by the Goldman Group from the University of Helsinki; here on referred to as 'DH10Bac-Y') were thawed on ice and combined with 2 µl of pFastBac vector in a pre-chilled 1.5 ml tube. The mixture was incubated on ice for 20 minutes, then heat-shocked at 42°C for 45 seconds. The mixture was returned to ice for 2-3 minutes, and the cells recovered by addition of 400 µl LB media and incubation at 37°C (overnight, 220 RPM). The following day, cells were diluted 1:10, 1:100, and 1:1000 and spread onto DH10 agar plates. The plates were incubated at 37°C for 24h. Blue and white colonies were picked and streaked onto DH10 agar plates using a P200 pipette tip, which was then used to pipette up and down into 2 ml of DH10 media. These cultures and plates were incubated at 37°C (overnight, 220 RPM) – any cultures whose colonies on the corresponding area of the DH10 agar plates were blue were discarded. Any cultures whose colonies on the corresponding area of the DH10 agar plates were white were used for bacmid purification (known as positive cultures).

2.3.6 Bacmid Purification

Positive cultures from blue/white screening were spun down (2,900 g, 5 minutes, 10°C), and the pellet resuspended in 300 µl of resuspension buffer (Macherey Nagel NucleoSpin kit). The cells were lysed by addition of 300 µl lysis buffer (from the same kit), and following a 5-minute incubation, the precipitate was pelleted (13,000 RPM, 15 minutes, 4C). The supernatant was transferred to a new 1.5 ml tube and spun down (13,000 RPM, 5 minutes, ambient temperature). The supernatant was again transferred to a clean 1.5 ml tube. 700 µl of 100% isopropanol was added and gently mixed with the supernatant by creating waves (not full inversion). After another centrifugation step (13,000 RPM, 10 minutes, ambient temperature), the supernatant was carefully decanted being careful not to touch the transparent pellet. 200 µl of 70%

ethanol was added dropwise onto the tube wall opposite the pellet. The tube was spun again (13,000 RPM, 5 minutes, ambient temperature), the ethanol removed, and 50 µl of fresh 70% ethanol added.

2.3.7 Transfection of Bacmids into Sf9 Cells

From each round of blue/white screening, two positive clones were carried forward for transfection. In a total volume of 3 ml, $1x10^6$ Sf9 cells growing in the log phase were seeded into the wells of a six-well tissue culture plate in the format below.

Clone 1'	Clone 2'	Cell Control
Clone 1''	Clone 2''	Media Only

The cells were allowed to adhere for at least 1h prior to addition of bacmid DNA. 30 µl of sterile MilliQ water was added to each purified bacmid and the tube gentle tapped 20 times on the hood surface. For each transfection 10 µl of transfection reagent (X-TremeGene HP DNA, Roche) was added to 100 µl of Sf9 media (the transfection solution). 20 µl of resuspended bacmid was diluted into 200 µl of Sf9 media (the bacmid mix). Then 100 µl of the transfection solution was added to the bacmid mix and incubated for 30 minutes at RT (transfection solution). 150 µl of the transfection solution was then added to each of two wells per clone (i.e., wells Clone 1' and Clone''). The plate was then incubated at 27°C for 3 days at which point the media was collected (containing the low titre virus, V_0) and stored in black Falcon tubes. Fresh media was added to the cells and the plate incubated for a further 3 days.

2.3.8 Maintenance of Sf9 Cells

Sf9 cells were cultured in glass flasks and maintained at a density between $0.5\text{-}8x10^6$ cells/ml. Cells were counted every, or every other day and diluted in Sf9 media according to their density confirmed after counting.

2.3.9 Amplification of Baculovirus

To amplify low titre V_0 viruses to higher titre V_1 viruses, 25 ml of Sf9 cells at $1x10^6$ cells/ml were infected with 3 ml of V_0 and incubated for 24 h (27°C, 120 RPM). If the cell density $>1.6x10^6$ cells/ml, cells were diluted back to $1x10^6$ cells/ml. If the cell density $<1.6x10^6$ cells/ml, no dilution was made. In both cases, cells were incubated for a further 24 h. Cells were counted again; at this point the cells should have stopped dividing due to infection with virus. If so, the cells were incubated for a further 24 h, then spun down and the virus containing supernatant collected and filtered into black falcon tubes using a 0.22 µm syringe filter (V_1). If the cells had continued to double, the culture was discarded since infection had not been established and thus, the culture would not yield sufficient virus for downstream applications.

2.4 Expression of CTS and CTS Mutants

2.4.1 Expression Optimisation in Sf9 Cells

2 ml of Sf9 cells at $1x10^6$ cells/ml were added to each of the wells of a 24-deep-well block. Increasing amounts of V_1 virus were add to each of the wells and the block was incubated for 4 days (27°C, 120 RPM). Each day, 200 µl of culture was collected from each well and the cells pelleted (2,900 g, 5 minutes, ambient temperature). Cells were resuspended in 100 µl of homogenisation buffer, sonicated (2 seconds, 50% intensity), and 30 µl was collected and made up to 60 µl with 15 µl 4X Bolt™ LDS sample buffer (Invitrogen™), 6 µl 10X Bolt™ reducing agent (Invitrogen™), and 9 µl MilliQ. Samples were incubated at ambient temperature for ~30 minutes prior to SDS-PAGE.

2.4.2 SDS-PAGE

SDS-PAGE was carried out using 4-20% Mini-PROTEAN® or Criterion™ TGX™ pre-cast gels (Bio-Rad) with running buffer containing Tris (25 mM), glycine (192 mM) and SDS (0.1% (w/v)). All gel samples were prepared using 4X Bolt™ LDS sample buffer (Invitrogen™) and 10X Bolt™ reducing agent (Invitrogen™) and were

incubated for ~30 minutes at ambient temperature prior to gel running. Gels were pre-run at 100V for 10 minutes prior to loading (this was necessary because these gels do not contain SDS prior to running). After loading, gels were run at 100V for ~90 minutes until the dye front reached the bottom of the gel. Proteins were visualised either by imaging in-gel eGFP fluorescence using a gel doc fitted with appropriate excitation and emission filters (Ex. 488 nm, Em. 509 nm), by staining with Quick Coomassie Stain (Generon), or by western blotting (section 2.4.3).

2.4.3 Western Blotting

Proteins were blotted to PVDF membranes (Bio-Rad) using a Trans-blot Turbo Transfer system (Bio-Rad), then blocked in 3% (w/v) BSA in 1X TBST for either 1h at RT, or overnight at 4°C. Primary antibody (Strep-tactin HRP conjugate (IBA; 2-1502-001), 1:5000 in 3% (w/v) BSA in 1X PBST, or anti-GFP (Abcam; Ab290, 1:2000 in 3% (w/v) milk in 1X TBST) was added and incubated for 1h at ambient temperature, then the membrane was washed with 3 x 15 ml of 1X PBST or TBST (depending on the antibody). For anti-GFP, the membrane was probed with secondary antibody (Jackson ImmunoResearch, goat α-rabbit-HRP; 111-035-003) 1:10,000 in 3% milk (w/v) in 1X TBST) for 1h at RT or overnight at 4°C, then washed with 3 x 15 ml of 1X TBST. For both antibodies, proteins were visualised using Clarity Western ECL substrate according to the manufacturer's instructions (Bio-Rad).

2.4.4 Large-scale expression of CTS and CTS mutants

V_1 virus stocks were used to infect Sf9 cells for expression of CTS and CTS mutants. The amount of V_1 required for large scale expression was calculated by extrapolation from the results of small-scale expression tests (section 2.4.1). 300-400 ml volumes of Sf9 cells in 1L glass flasks were infected with 80-100 µl of V_1 virus, supplemented with antibiotic-antimycotic (Gibco®), and incubated for 3 days (27°C, 120 RPM).

2.5 Harvesting of Sf9 Cells Expressing CTS or CTS Mutants

Sf9 cells expressing CTS and CTS mutants were pooled into 600-800 ml volumes and harvested by centrifugation (600 g, 20 minutes, 4°C). Pellets were resuspended in homogenisation buffer (containing TRIS or HEPES if purifying in DDM or LMNG/digitonin, respectively) supplemented with Proteoloc™ EDTA-free protease inhibitors (Expedeon) and AEBSF (30 ml homogenisation buffer/800 ml cells), frozen in LN2 and stored at -80°C.

2.6 Membrane Preparation for CTS and CTS mutants

Harvested cells (section 2.5) were thawed at RT then supplemented with DTT (3 mM). 30 ml of cells were lysed by 40 passes in a tight-fitting glass homogeniser. Unbroken cells and debris were removed by centrifugation (2000 g, 20 minutes, 4°C), and the supernatant collected and stored on ice. The pellet was resuspended in 30 ml of homogenisation buffer, homogenised again, and spun down (2000 g, 20 minutes, 4°C). All supernatants were combined and subject to a high-speed centrifugation step (100,000 g, 1h, 4°C), and the pelleted membranes resuspended in a minimal volume of Buffer A containing AEBSF (1 mM; typically, 2 ml buffer/membrane pellet) using a positive-displacement pipette. Protein concentration in the membranes was determined using a Pierce™ BCA assay kit (Thermofisher Scientific) according to the manufacturer's instructions. Resuspended membranes were snap-frozen by dropping into LN2. Frozen membrane pellets were stored at -80°C.

2.7 Detergent Screening

500 µl of CTS membranes (~15 mg) were incubated either in either n-dodecyl-β-D-maltopyranoside (DDM), octyl glucose neopentyl glycol (OGNG)l, lauryl maltose neopentyl glycol (LMNG), decyl glucose neopentyl glycol (DGNG), or 4-cyclohexyl-1-butyl-β-D-maltoside (CYMAL-4) (2%, w/v) supplemented with CHS (0.4%, w/v), and asolectin (0.4%, w/v) for 2 hrs at 4°C. These detergent and lipid concentrations were used because they have been shown to efficiently extract membrane proteins from

the membrane (Stetsenko & Guskov, 2017). Asolectin, a soy-bean phospholipid mixture, was included to generate detergent-lipid micelles that may improve the stability of CTS and/or allow for a milder solubilisation than with detergent alone. 50 µl was then collected (total protein sample) and the remainder spun down (100,000 g, 30 minutes, 4°C). 50 µl of the soluble fraction was collected, and the insoluble pellet was resuspended in 50 µl of Buffer A. All samples were prepared for SDS-PAGE by mixing with 2X loading dye (prepared from 4% (w/v) SDS, 0.2% (w/v) bromophenol blue, and 20% (v/v) glycerol) supplemented with 10 mM DTT then heated at 60°C for 10 minutes. Samples were run on an SDS-PAGE gel then subject to western blot using an anti-strep antibody (section 2.4.3).

2.8 T_m determination

CTS membranes were diluted to 2 mg/ml in Buffer A and solubilised in 2% DDM for 2h at 4°C in a total volume of 3.5 ml. Insoluble material was removed by centrifugation (100,000 g, 1h, 4°C) and the soluble material was carried forward either undiluted (in 20% glycerol) or diluted in Buffer A such that the glycerol concentration was reduced to 10% or 5%. Samples were divided into 30 x 100 µl aliquots and heated at 4, 20, 30, 40, 50, 60, 70, 80, 90, and 100°C for 10 minutes in triplicate. Samples were placed back on ice for 30 minutes then spun down (20,000 g, 30 minutes, 4°C. 40 µl of each supernatant was prepared for SDS-PAGE (section 2.4.2), and then run on two 26-well Criterion™ gels. Surviving GFP signal was detected by in-gel GFP fluorescence. The bands intensities were determined using densitometry, then averaged per set of triplicates. The average signal from the 4°C triplicate was designated as 100% GFP survival, and the average triplicate intensities for the other temperature conditions compared against this. By plotting the percentage survival vs. temperature using a non-linear regression and a four-parameter dose-response curve for inhibitors, the T_m was determined. T_m plots were created in GraphPad Prism.

2.9 Optimisation of Solubilisation Concentration

To determine the optimal concentration of membrane at which to solubilise CTS, 200 µl of CTS membranes were solubilised at either 2, 4, 8, or 10 mg/ml using 1% or 2% (w/v) DDM for 2h (4°C, end-over-end rotation). FC-14 was used as a positive control (solubilising at 2 mg/ml). 50 µl of total protein from each condition was spun for soluble and insoluble fraction (100,000 g, 30 minutes, 4°C). Both the soluble and insoluble fractions from each condition were prepared for SDS-PAGE in a total volume of 70 µl. Equal amounts of membrane were loaded onto an SDS-PAGE gel and protein visualised by in-gel GFP and western blot (section 2.4.2 and 2.4.3).

2.10 Purification of CTS and CTS Mutants in Detergent

2.10.1 Solubilisation in DDM/CHS or LMNG/CHS

Frozen membranes were thawed at RT and diluted into either Solubilisation Buffer A (20 mM Tris-HCl pH 7.4, 200 mM NaCl, 20% glycerol, 1.5 mM $MgCl_2$, 1% DDM/0.2% CHS/0.2% asolectin (w/v), 1 mM PMSF, 1X Protloc™ protease inhibitors (Expedeon) or Solubilisation Buffer B (20 mM HEPES pH 7.5, 200 mM NaCl, 20% glycerol, 2 mM $MgCl_2$, 1 mM DDT, 1% LMNG/0.1% CHS (w/v), 4 mM ATP, 1 mM AEBSF, 1X Protloc™ protease inhibitors (Expedeon), and incubated for 2h (4°C, rolling agitation). Insoluble material was removed by centrifugation (100,000 g, 1h, 4°C), and the supernatant stored on ice for immediate purification.

2.10.2 Streptactin Affinity Purification – DDM/CHS/asolectin

Solubilised material (section 2.10.1) was diluted 4X in 20 mM Tris-HCl pH 7.4, 200 mM NaCl, 1.5 mM $MgCl_2$ to dilute glycerol and detergent concentrations to 5% and 0.25%, respectively. Streptactin resin (IBA, Fast Flow High Capacity; 1 ml) was washed with MilliQ water (10 column volumes (CV)), then equilibrated with 10CV of 20 mM Tris-HCl pH 7.4, 200 mM NaCl, 1.5 mM $MgCl_2$, 5% glycerol, 0.25% DDM/0.05% CHS/0.05% asolectin (w/v). Soluble material was incubated with equilibrated Streptactin resin overnight (4°C, stirring). Material was applied to an

affinity column and unbound material allowed to flow-through. Resin was washed with 2CV Wash Buffer A + 0.5% DDM/0.1% CHS/0.1% asolectin (w/v), 2CV Wash Buffer A + 0.1% DDM/0.02% CHS/0.02% asolectin (w/v), and 2CV Wash Buffer A + 0.05% DDM/0.01% CHS/0.01% asolectin (w/v). CTS was eluted in 8CV of Wash Buffer A + 0.05% DDM/0.01% CHS/0.01% asolectin (w/v), 5 mM desthiobiotin. All samples prepared as described were used immediately for any and all further experiments without freezing or long-term storage.

2.10.3 Streptactin Affinity Purification – LMNG/CHS

Solubilised material (section 2.10.1) was diluted 2X in 20 mM HEPES, 200 mM NaCl, 2 mM $MgCl_2$ to dilute glycerol and detergent concentrations to 10% and 0.5%, respectively (w/v). Streptactin resin (IBA, Fast Flow High Capacity; 1 ml) was washed with MilliQ water (10CV), then equilibrated with 10CV of 20 mM HEPES pH 7.5, 200 mM NaCl, 2 mM $MgCl_2$, 5% glycerol, 0.5% LMNG/0.05% CHS (w/v). Soluble material was incubated with equilibrated Streptactin resin overnight (4°C, stirring). Material was applied to an affinity column and unbound material allowed to flow-through. The resin was washed either with 10CV Wash Buffer B containing 0.1% digitonin (w/v) then 10CV Wash Buffer B containing 0.06% digitonin (w/v) or 5 x 5 ml Wash Buffer B containing 0.06% digitonin (w/v). CTS was eluted in 8 CV Wash Buffer B + 0.06% digitonin (w/v) + 5 mM desthiobiotin. All samples prepared as described were used immediately for any and all further experiments without freezing or long term storage.

2.10.4 Size-exclusion Chromatography (SEC)

All SEC was carried out using an ÄKTA Pure system (GE Healthcare) with either Superdex 200 5/150 or 10/300 Increase columns, or Superose 6 5/150 or 10/300 Increase columns at 4°C. 10/300 columns were ran at a flow rate of 0.75 ml/min and 5/150 columns were ran at a flow rate of 0.2 ml/min. These chromatography columns are used to separate biomolecules, most often proteins, by their size to improve the purity of the sample prior to downstream application. In structural biology for example,

the purity of a protein sample for crystallisation may need to be >95%. This is often lower for electron microscopy studies; however, it is still desirable. Prior to SEC, columns were washed with 1.2 CV water, and equilibrated with 1.2CV of an appropriate buffer.

2.11 Proteoliposome Reconstitution

Liposomes were prepared from asolectin and CHS. 31.5 mg of asolectin and 8.5 mg of CHS were weighed into a glass tube and dissolved in a minimal volume of 100% chloroform by bath-sonication and vortexing. Lipid films were created around the tube walls by drying under a gentle stream of nitrogen then dried under vacuum overnight. Lipid films were resuspended to a concentration of 20 mg/ml in 2 ml of 20 mM HEPES pH 7.5, 200 mM NaCl, 5% glycerol, 2 mM $MgCl_2$, 1 mM DTT by vortexing followed by bath-sonication for 1h at 4°C. Liposomes were extruded through a 0.22 µm filter pre-equilibrated in 20 mM HEPES pH 7.5, 200 mM NaCl, 5% glycerol, 2 mM $MgCl_2$ for 11 passes, and destabilised by the addition of Triton X-100 at a concentration of 0.8% (v/v, i.e., 16 µl per 2 ml lipid resuspension). Liposome destabilisation is the weakening of the interactions between lipid monomers of the liposome by detergent. This step is critical as without it, the membrane protein cannot be inserted into the liposome. This concentration of Triton X-100 was chosen due to its reported success in the formation of proteoliposomes (Simeonov *et al*, 2013; Stefan *et al*, 2020). Detergent purified CTS (i.e., in digitonin) was combined with destablised liposome at a 1:100 ratio (w/w) at a final CTS concentration of 100 µg/ml and a final liposome concentration of 10 mg/ml, then incubated for 1h at 4°C with rolling. A negative control was also prepared by omitting CTS. 200 mg of Biobeads SM-2 (Biorad), pre-equilibrated in 20 mM HEPES pH 7.5, 200 mM NaCl, 5% glycerol, 2 mM $MgCl_2$, were then added and the mixture incubated for 3 hrs at 4°C with rolling. A further 200 mg of biobeads SM-2 were added and samples incubated overnight at 4°C. Samples were transferred into fresh tubes and 400 mg of biobeads SM-2 were added and incubated for 3 hrs at 4°C.

Proteoliposomes were collected by ultracentrifugation (262,000 g/20 minutes/4°C) and resuspended in 50 mM Tris-HCL pH 7.5, 10 mM $MgCl_2$ to a concentration of 1 µg/µl, and used immediately for ATPase assays.

2.12 ATPase Assay on CTS Proteoliposomes

The ATPase activity of CTS was determined in proteoliposomes using a spectrophotormetric assay based on that reported by Chifflet *et al* (1988). 2.5 µg of CTS-proteoliposomes were diluted in buffer containing 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$ and supplemented by the addition of 20 µM or 40 µM C18:0-CoA (dissolved in 5% hydroxypropylated β-cyclodextrin) or 20 mM AlF_3 and incubated for at 37°C for 5 minutes. The reaction was initiated by the addition of 4 mM ATP (Na_2ATP dissolved in 7 mM $MgSO_4$ pH 7.4) in a total volume of 40 µl and allowed to continue for at 37°C for 30 minutes. The reaction was stopped by the addition of 40 µl 12% SDS. Colour development was initiated by the addition of 80 µl of a 1:1 mix of 6% (w/v) ascorbic acid (dissolved in 1M HCL) and 1% (w/v) ammonium molybdate and incubation for at RT for 5 minutes. Then 120 µl of a solution of 2% (w/v) sodium citrate, 2% (w/v) sodium metaarsenite, 2% (w/v) acetic acid was added and incubated at 37°C for 15 minutes. Colour development was measured at 800 nm and the release of inorganic phosphate calculated by comparison to a standard curve of inorganic phosphate and reported in nmol P_i/mg/min.

2.13 Purification of CTS in Styrene Maleic Acid Copolymer (SMA)

CTS membranes were solubilised 2 mg/ml in 50 mM Tris-HCl pH 7.4, 500 mM NaCl, 5% glycerol, and 2.5% (w/v) SMA for 2 hrs at RT in a total volume of 2 ml. Soluble material was collected by centrifugation (100,000 g, 30 minutes, 4°C). 600 µl of soluble material was bound to 50 µl of streptactin resin or 20 µl of GFP-trap resin (Chromotek) pre-equilibrated with 50 mM Tris-HCl pH 7.4, 500 mM NaCl, 5% glycerol (v/v). Binding continued overnight (4°C with end-over-end agitation). Both the streptactin resin and GFP-trap were washed with 4CV 50 mM Tris-HCl pH 7.4, 500

mM NaCl, 5% glycerol (v/v). Elution from the streptactin resin was carried out using 4CV 50 mM Tris-HCl pH 7.4, 500 mM NaCl, 5% glycerol (v/v) + 10 mM desthiobiotin. Elution from the GFP-trap resin was attempted using 1 µg HRV3C protease/20 µl GFP-Trap resin. Samples were prepared for SDS-PAGE (section 2.4.2) then proteins visualised by in-gel GFP and anti-streptag western blot (2.4.3).

2.14 Nanodisc Experiments

2.14.1 Expression of MSP1D1 & MSP1E3D1

pMSP1D1 and pMSP1E3D1 were transformed into OmniMAX™ competent cells as in section 2.1.2 and plasmids purified according to section 2.1.3. For expression, newly purified plasmids were transformed into BL21 Gold (DE3) competent cells (Agilent) and single colonies used to inoculate 1L of TB media. Cells were incubated until OD_{600} reached 0.6 (30°C, 220 RPM) then induced with IPTG (1 mM) and incubated for a further 3 hrs. Cells were collected by centrifugation (5000 g, 15 minutes, 4°C), the media discarded, and pellets frozen at –20°C.

2.14.2 Purification of MSP1D1 & MSP1E3D1

Cell pellets (equivalent to 500 ml of cells) were resuspended to a final volume of 16 ml in PBS containing PMSF (1 mM), DNase (2.5 mg), and Triton X-100 (1%, v/v). The resuspension was incubated on ice for 30 minutes, sonicated (30 seconds on, 30 seconds off, 12X, 50% intensity) then centrifuged (30,000 g, 30 minutes, 4°C). The supernatant was then incubated for 1h with 2 ml of Ni-NTA resin pre- equilibrated with 5CV of PBS. The resin was applied to a flow-column and the flow-through collected. The resin was washed with 5CV of 40 mM Tris pH 8.0, 300 mM NaCl, 1% Triton X-100, 5CV of 40 mM Tris pH 8.0, 300 mM NaCl, 50 mM Sodium Cholate, 20 mM imidazole, 5CV of 40 mM Tris pH 8.0, 300 mM NaCl, 50 mM imidazole, and then protein eluted with 12 x 1ml fractions of 40 mM Tris pH 8.0, 300 mM NaCl, 400 mM imidazole. The purified proteins were then dialysed overnight against 1L of 20 mM Tris pH 7.4, 100 mM NaCl, 0.5 mM EDTA using a SnakeSkin dialysis membrane (10

kDa molecular weight cut-off (MWCO); Thermofisher Scientific) with two exchanges. Both proteins were concentrated to ~9.5 mg/ml using a 5 kDa MWCO centrifugal concentrator (Santorius), snap-frozen as 50 µl aliquots, then stored at -80°C.

2.14.3 Preparation of POPC Lipid Scks

5 mg aliquots of POPC lipids (Avanti Polar Lipids; synthetic source) dissolved in chloroform were dried under a nitrogen stream in glass bijou bottles then stored under nitrogen in sealed glass bulbs at -20°C. Immediately prior to use, POPC lipids were dissolved to a concentration of 657 µM in buffer containing 20 mM Tris-HCl pH 7.5, 200 mM NaCl, 0.5% (w/v) DDM.

2.14.4 Reconstitution of CTS into POPC Nanodiscs

CTS was purified as in section 2.8.2 with the following alterations: Streptactin XT Superflow® high capacity resin instead of Strep-tactin® Superflow® high capacity resin, and elution using biotin instead of desthiobiotin. Purified CTS was combined with MSP1D1 or MSP1E3D1 and POPC lipids at varying ratios in 20 mM Tris-HCl pH 7.5, 200 mM NaCl, 2% (v/v) glycerol in a total volume of 2.5 ml then incubated on ice for 1h. ~500 mg biobeads (Biorad) were added to each sample and incubated for 1h (4°C, end-over-end mixing). The biobeads were then replaced with fresh biobeads, and the samples were incubated overnight (4°C, end-over-end mixing). The sample was separated from the biobeads using a thin-tipped pipette tip and centrifuged to obtain soluble and insoluble (37,000 g, 30 minutes, 4°C) for assessment by SDS-PAGE and negative stain EM.

2.15 Electron Microscopy

2.15.1 Negative-stain Electron Microscopy Grid Preparation

Negative stain electron microscopy was used to assess whether the purified protein samples were suitable for cryo-EM by looking for expected particles (of correct size and shape), the state of aggregation, and the degree of homogeneity/heterogeneity. Carbon-coated copper EM grids (400 mesh; Agar Scientific Ltd) were glow-

discharged using an easiGlow™ instrument (Pelco©) for 60 seconds. Sample was applied (3 µl), allowed to sit for 30 seconds, and then blotted away by touching the side of the grid to a disc of Whatman paper. Uranyl acetate (3 µl, 1% (w/v)) was applied and immediately blotted away by touching the side of the grid to Whatman paper. Uranyl acetate was applied again and allowed to sit for 10 seconds then blotted away as before. The grid was air dried under a lamp for ~30 seconds. Grids were imaged on a FEI F20 microscope operating at 120kV with data collected on a CMOS camera.

2.15.2 Cryo-electron Microscopy Grid Preparation and Data Collection

In this study, the following grids were used: carbon-coated copper Quantifoil R1.2/1.3 200 or 300 mesh grids (Agar Scientific Ltd), carbon-coated gold Quantifoil R1.2/1.3, 300 mesh grids (Agar Scientific Ltd), or Ultrafoil® grids. All grids were glow-discharged (40 mA, 60 seconds) using a GlowQ Plus system (Quorum Technologies Ltd). Samples–(0.5 - 0.9 mg/ml) were prepared using a Vitrobot plunge freezing device according to the manufacturer's instructions. 3 µl of sample was applied to the grid, blotted for 3 seconds at 100% humidity/4°C, before plunging into liquid ethane. Grids were loaded into a Titan Krios transmission electron microscopy (FEI) operating at 300 kV for screening and data collection (University of Leeds, Astbury Biostructure Laboratory) using either a K2® (GATAN) or Falcon 4 camera (Thermo Scientific) and data was collected use the EPU automated acquisition software.

2.15.3 Cryo-EM Data Processing

To process cryo-EM data, it is first necessary to correct the micrographs for motion drift because of stage movement and beam induced motion, and to restore missing information by estimating the contrast in the micrographs by carrying out CFT estimation. To do this, the micrographs were subject to motion-correction and CTF estimation using RELION's implementation of MotionCor2 (Zheng *et al*, 2017) and

either CTFFind 4.1 (Rohou & Grigorieff, 2015) or gCTF (Zhang, 2016). These pre-processed micrographs were then used for particle picking in crYOLO1.6.1 (Wagner *et al*, 2019). Particle picking is necessary to separate expected particle from aggregates and debris so that downstream processing steps can be carried out on a clean set of particles (i.e., a group or groups of particles representing the protein of interest, possible in different conformations.

To do this, crYOLO's general model, gmodel_phosnet_202005_N63_c17.h5 was used. Particle picking can be made more stringent or less stringent by adjusting the picking threshold. For all data sets, the picking threshold was set to 0.1. For the largest data set (purified as in section 2.8.3, table 5.2), a picking threshold of 0.3 was also used (crYOLO default). Once particles have been picked, they must be extracted from the micrographs *in* silico such that they can be classified independently from the entire micrograph. Extraction of each particle is carried out in a 'box' of defined size that surrounds the particle entirely without including too much of the background grid surface. Picked particles were extracted in a box size of either 180 or 256 pixels and subjected to reference-free 2D classification using a mask diameter of 220 Å. Here the mask diameter defines the size of the space used for the 2D classification, and the 2D classification defines an iterative process of particle alignment to generate more detailed classes from stacks of extracted particles. For the largest data set, a subset of the best 2D class averages were used for 3D classification using the model of ABCD4 as a template. This template was generated in ChimeraX using the molmap command to create a 15 Å density map from the ABCD4 model. 3D classification was carried out using C1 and C2 symmetry with a mask diameter of 220 Å.

3 Expression and Solubilisation Optimisation of COMATOSE for Structural Studies by Cryo-EM

3.1 Introduction

The initial observation that COMATOSE possesses intrinsic acyl-CoA thioesterase (ACOT) activity (De Marcos Lousa *et al.*, 2013) is now supported by several further studies including those on CTS expressed in yeast (Carrier *et al.*, 2019), the yeast transporter pxa1/pxa2 (van Roermund *et al.*, 2021), and the ABCDs in humans (Kawaguchi *et al.*, 2021; Okamoto *et al.*, 2018). Moreover, homology modelling and further biochemical studies have identified residues in CTS that may contribute to this activity (Carrier *et al.*, 2019) and it has also been shown that HsABCD1 ACOT activity is an essential for acyl-CoA transport (Kawaguchi *et al.*, 2021). Moreover, in yeast, peroxisomal acyl-CoA synthetase activity is required for ABCD-dependent transport of acyl-CoAs (van Roermund *et al.*, 2012). Despite these reports, there is still no definitive explanation of exactly *how* the transporter is capable of substrate cleavage on a molecular level, primarily because at the beginning of the project, there was no structural data available for CTS, or for any other family D ABC transporter. Therefore, the overarching aim of this project was to determine 3D structure of substrate-bound CTS (and/or relevant mutants) using cryo-EM and to evaluate this structure in the context of the biochemical studies to reveal the nature of CTS ACOT activity.

The structural work is paramount, primarily because it will allow the identification residues and/or regions of the transporter involved in substrate binding and cleavage. Once such residues are identified, other structural experiments will build up a picture of the catalytic cycle, for example by trapping the transporter in specific conformations and observing changes in substrate binding, by observing structural changes between pre- and post-substrate cleavage states, or by capturing intermediates that potentially allude to novel transport mechanisms that are not currently under consideration when exploring the transport of large amphipathic molecules. Of

course, such studies will also allow for the development of new biochemical experiments to validate the observations made and guide further studies.

Prior to this project, an expression and purification protocol for CTS^{WT} had been devised by previous post-doctoral researchers (Dr David Carrier and Dr Carine De Marcos Lousa), for use in biochemical studies. However, this protocol was not suitable for cryo-EM studies primarily because the yield of protein required for cryo-EM was much greater than previously required and because the buffer conditions were not suitable; therefore it needed to be optimised for such work. To carry out structural studies of membrane proteins by cryo-EM, two basic, but major criteria, must be met. Firstly, the protein must be expressed in a suitable expression system and be produced at a high enough yield (i.e., ≥ 0.5 mg/L culture). The identification of a suitable expression system is the first critical step toward structural studies because several factors must be considered, including the type of post-translational modifications (PTMs) and if disulphide bonds are required (CTS contains 10 cysteine residues), the ease of handling the expression system, and the mass of protein required for downstream experiments (Pandey *et al*, 2016). Regarding the latter, membrane proteins are generally expressed, even in the ideal system, at orders of magnitude lower levels than soluble proteins.

Secondly, the protein must be solubilised efficiently, and in such a way that does not overly complicate downstream cryo-EM grid preparation and screening/data collection experiments (Kampjut *et al*, 2021). Optimisation of this step is critical because significant losses can be incurred if a non-optimal detergent/detergent-free alternative is used, which compounds with the poor expression of membrane proteins stated above (Lenoir *et al*, 2018). Typically, an efficiency of >50% (i.e., a 50:50 split between soluble and insoluble) is desirable given that structural studies often require large amounts of material. Moreover, the physical and chemical properties of different detergents used for solubilisation must be considered because the detergent that best solubilises the transporter may not be the best detergent with which to carry out

downstream studies. For example, detergents can be classified as ionic, zwitterionic, or non-ionic. Ionic detergents such as SDS, whilst extremely efficient at solubilisation, are often denaturing and consequently are not useful in purifying the protein of interest in an active and folded form. Conversely, non-ionic detergents are more 'mild' than ionic detergents, and whilst they do not provide such a strong solubilisation efficiency, they typically solubilise the protein of interest in its folded form, preserving its structure. Zwitterionic detergents are an intermediate form of detergent, being more denaturing than non-ionic detergents but less than ionic ones. It is necessary to decide on the solubilisation method most applicable for downstream experiments. Moreover, the unique critical micelle concentration (CMC) of the detergent should be carefully considered, because, below this concentration the detergent will be incapable of solubilising the transporter, and similarly, detergents with high CMC values are generally not suitable for structural studies due to the excess of empty micelles in solution (Ratkeviciute *et al*, 2021). One should also consider the extent to which the chosen detergent may delipidate the protein of interest; it is well known that endogenous lipids are often necessary for membrane protein activity (Kalipatnapu & Chattopadhyay, 2005; Neumann *et al*, 2017) and thus, despite as chosen detergent solubilising the protein well, it may not be a suitable detergent for the proteins activity due to it stripping away requisite lipids.

Additionally, it is often useful to study mutants of the protein to investigate its structure, where some mutants may be utilised because they adopt a specific conformation (e.g., being locked open or closed, or in some transition state) (Manolaridis *et al.*, 2018; Song *et al*, 2021), or because they possess residues that alter substrate binding/transport (Bin Kanner *et al*, 2021; Khunweeraphong *et al*, 2017; Meier *et al*, 2021; Srikant *et al*, 2020). These approaches allow for a more controlled investigation of the protein by letting researchers analyse certain aspect of its catalytic cycle. In doing so, it is possible to build up a detailed model of the protein's

mechanism of action and inform not only on the physiological processes involved, but also those states or mutations can lead to pathology.

Considering these factors, this chapter details the experiments used to optimise the expression and solubilisation of CTS^{WT} and two mutants – CTS^{K487A} and $CTS^{D606N/E607Q}$ – with the aim of using such constructs for structural studies by cryo-EM. This work represents the beginning of structural studies on CTS.

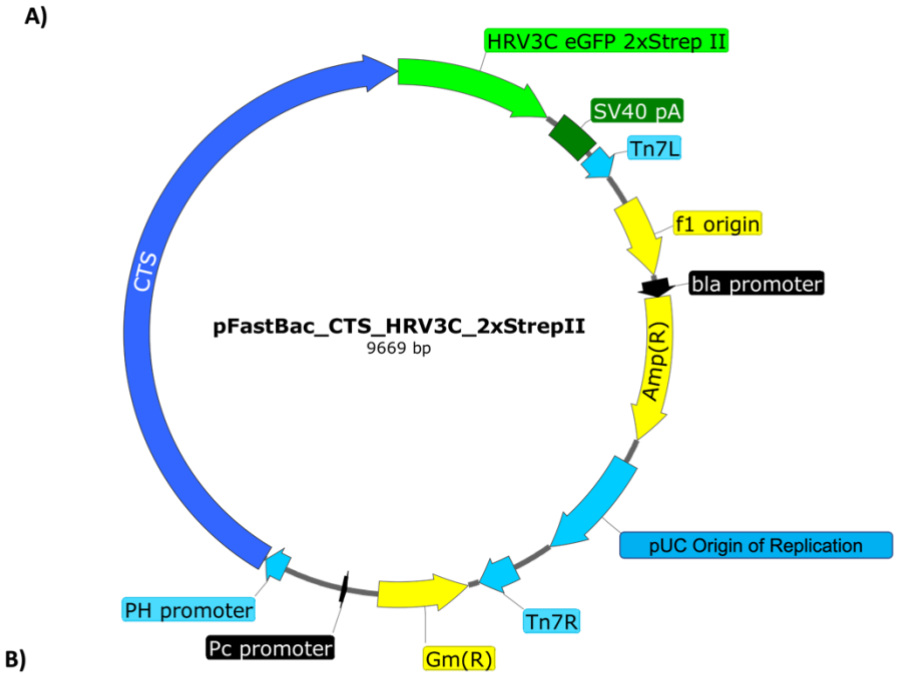
3.2 Results

3.2.1 Expression of CTS^{WT}, CTS^{K487A}, and $CTS^{D606N/E607Q}$

In this study, three different constructs were studied: CTS^{WT}, $CTS^{D606N/E607Q}$ and CTS^{K487A}. CTS^{WT} represents the wild-type version of CTS. $CTS^{D606N/E607Q}$ carries mutations in the Walker-B motif of NBD, where the replacement of the Mg^{2+} coordinating aspartate residue and catalytic glutamate residue renders the mutant hydrolysis-deficient. This mutant was chosen for structural characterisation because whilst it should be capable of binding nucleotide, the nature of its mutations mean that bound nucleotide will not be hydrolysed and therefore the protein would most likely be locked in an outward-facing conformational state. Mutants like this have been used frequently in the study of ABC transporters (Manolaridis *et al.*, 2018; Song *et al.*, 2021) to force a conformational state and/or to stabilise the transporter. On the other hand, CTS^{K487A} carries a lysine-to-alanine substitution mutation in the Walker-A motif of NBD1. This mutant was chosen for structural studies because intriguingly, it has reduced, but not completely abolished ATPase activity compared to CTS^{WT} (De Marcos Lousa *et al.*, 2013; Nyathi *et al.*, 2010).

How this occurs is unclear, but an explanation may be provided by a 3D model. All three constructs used in this study were kindly provided by Dr David Carrier at the beginning of the project and all were expressed using the Baculovirus expression system. Despite being time-consuming and relatively expensive, this system can express high yields of eukaryotic proteins was therefore suitable for expression of

CTS^{WT}. The full-length constructs consisted of CTS^{WT} (or either mutant) linked by a HRV3C cleavable linker to an eGFP tag followed by a 2xStrepII tag. By taking advantage of the construct's C-terminal eGFP tag (Fig. 3-1) protein expression was monitored at both the expression and purification stages, and through the 2xStrepII tag, the protein was purified using Streptactin affinity resins. When expressed, the full-length constructs yielded a ~182 kDa product (including the eGFP and StrepII tag).



CTS^{WT}	TGPNGSGKSSLFRVLGGLWPLVSGHIVKPGVGSDLNKEIFYVPQRPYMAVGTLRDQLIYP	480–539
CTS^{K487A}	TGPNGSGASSLFRVLGGLWPLVSGHIVKPGVGSDLNKEIFYVPQRPYMAVGTLRDQLIYP	480–539
$CTS^{D606N/E607Q}$	TGPNGSGKSSLFRVLGGLWPLVSGHIVKPGVGSDLNKEIFYVPQRPYMAVGTLRDQLIYP	480–539
CTS^{WT}	LTSGQESELLTEIGMVELLKNVDLEYLLDRYQPEKEVNWGDELSLGEQQRLGMARLFYHK	540–599
CTS^{K487A}	LTSGQESELLTEIGMVELLKNVDLEYLLDRYQPEKEVNWGDELSLGEQQRLGMARLFYHK	540–599
$CTS^{D606N/E607Q}$	LTSGQESELLTEIGMVELLKNVDLEYLLDRYQPEKEVNWGDELSLGEQQRLGMARLFYHK	540–599
CTS^{WT}	PKFAILDECTSAVTTDMEERF	600–620
CTS^{K487A}	PKFAILDECTSAVTTDMEERF	600–620
$CTS^{D606N/E607Q}$	PKFAILNQCTSAVTTDMEERF	600–620

Figure 3-1. CTS construct information. ***A)*** *Plasmid map for pFastBac_CTS_HRV3C_2xStrepII showing the HRV3C-cleavable C-terminal eGFP tag and 2xStrepII tags* ***B)*** *Sequence alignment of* CTS^{WT} CTS^{K487A}, *and* $CTS^{D606N/E607Q}$ *showing the residues where they differ (red,* CTS^{K487A}*; blue,* $CTS^{D606N/E607Q}$*).*

After transfection of the freshly prepared bacmids into adherent Sf9 cells and incubation for 60 hrs, expression of the corresponding proteins was assessed using fluorescence microscopy and western blotting. Observation of the Sf9 cells on their culture plate by eGFP-specific fluorescence microscopy indicated successful expression of all three constructs (Fig. 3-2a) evident by green fluorescence. The greater signal observed from CTS^{WT} and CTS^{K487A} compared to $^{CTSD606N/E607Q}$ was due to CTS^{WT} and CTS^{K487A} bacmids being generated from DH10 competent cells carrying a YFP reporter sequence (DH10Bac-Y), whereas $CTS^{D606N/E607Q}$ was expressed in the absence of this reporter. Because the band pass filters used in conventional fluorescence microscopy cannot differentiate between eGFP and YFP due to the protein's overlapping emission spectra, signal is generated from both proteins and observed as increased fluorescence. To the confirm specific expression of each construct, anti-GFP western blotting of cell lysates was used (Fig. 3-2b). This corroborated the fluorescent microscopy results and confirmed expression by the presence of CTS/mutant specific bands at ~190 kDa.

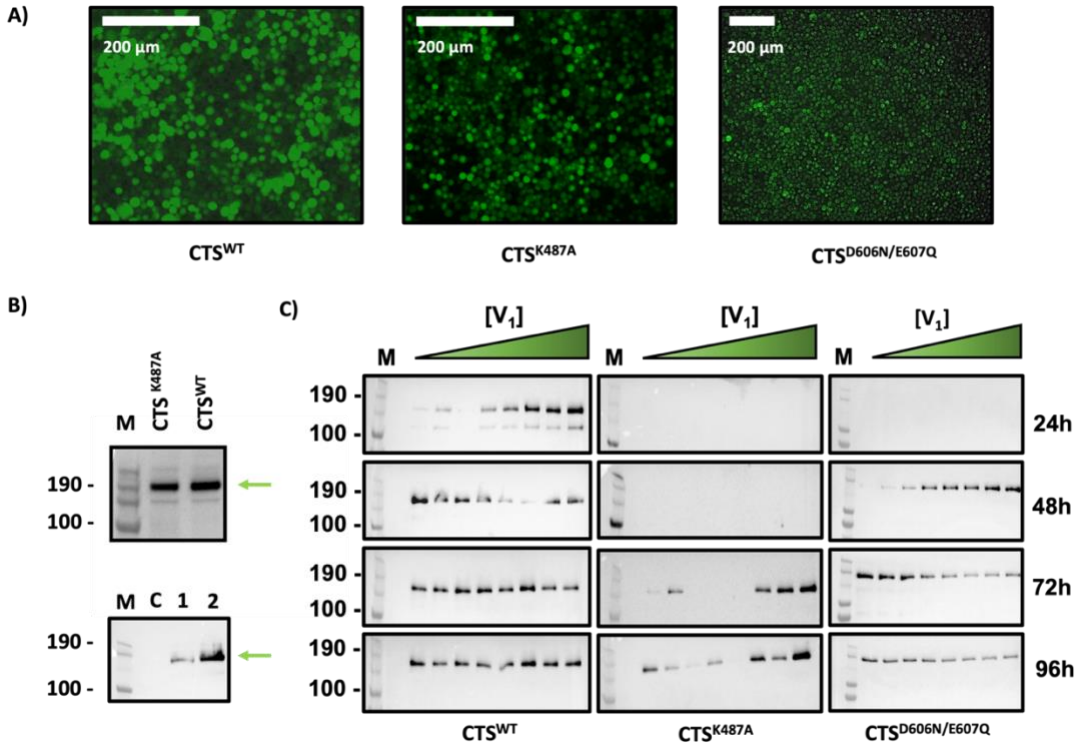


Figure 3-2. Transfection and small-scale expression testing for CTS^{WT}, CTS^{K487A}, and $CTS^{D606N/E607Q}$ representative of multiple experiments. ***A)*** *Fluorescence microscopy on adherent Sf9 cells after transfection with bacmids encoding either CTS^{WT}, CTS^{K487A}, or $CTS^{D606N/E607Q}$ showing expression of all three constructs.* ***B)*** *Anti-GFP western blot of cell lysates from Sf9 cells shown in panel A, confirming expression of all constructs.* ***Top panel****,CTS^{WT} & CTS^{K487A};* ***bottom panel****, $CTS^{D606N/E607Q}$.* ***C)*** *Anti-GFP western blots showing viral titre testing of all three constructs. Cells were grown for up to 96 hrs in the presence of increasing amounts of V_1 virus and an aliquot harvested every 24 hrs. Aliquots were then lysed and subject to anti-GFP western blotting to identify the optimal volume of V_1 virus for large scale expression.* ***M****, molecular weight marker;* ***C****, empty cell control;* ***1 & 2****, $CTS^{D606N/E607Q}$ clones;* ***[V_1],*** *amount of V_1 virus;* ***green arrow****, CTS-specific bands.*

After protein expression was confirmed, the isolated low-titre viral stock (V_0) was amplified to a high-titre viral stock (V_1) for large scale expression. However, it was first necessary to determine the optimal volume of V_1 that would be used to express each construct in large-scale suspension cultures. In small-scale (i.e., using 24-well deep well plates), Sf9 cells were incubated with increasing volumes of construct-specific V_1 virus and incubated for 96h. Aliquots were collected after 24, 48, 72, and 96 hrs, and the level of expression at each timepoint assessed by anti-GFP western blotting (Fig. 3-2c). The volume of V_1 virus chosen for subsequent large-scale expression was balanced between the volume of virus used and the shortest incubation time (e.g., if 1 ml of V_1 virus was required after 48 hrs to obtain good expression, but 100 µl of V_1 virus was required after 72 hrs, then 72 hrs was chosen). Therefore, whilst the optimal volume of V_1 virus for each construct varied, the optimal expression timepoint for all constructs was 72h. Once small-scale expression testing had identified conditions suitable for large-scale expression, the yield of each CTS construct from large-scale cultures was quantified to assess whether it would be enough for structural studies. Each construct was expressed in 1L of Sf9 cells and after 72 hrs, the cells were harvested, and a 1 ml fraction collected for quantification by anti-GFP western blotting. After lysis, a fraction was run on an SDS-PAGE gel and by using recombinant eGFP standards of known concentration and densitometry, the yield for each construct was estimated to be ~0.5 mg/1L of Sf9 cell culture (Fig. 3-3) meaning it was suitable for structural studies by cryo-EM (Parker & Newstead, 2014; Tate, 2001; Unger & Peleg, 2012). During each large-scale expression, this 1 ml fraction was always collected and used to assess the expression quantity and quality.

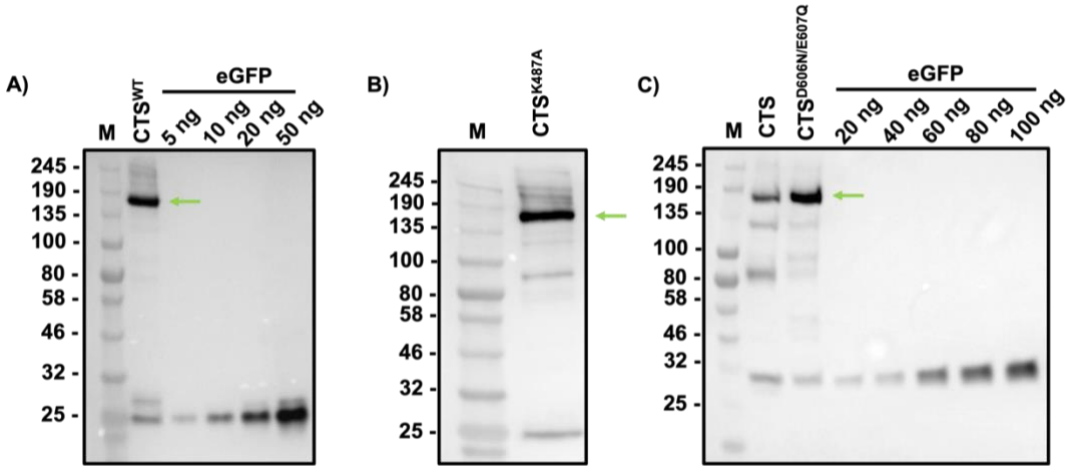


Figure 3-3. Western blot of large-scale expression testing for CTS and CTS mutants.CTS^{WT}, ***B)*** *CTS^{K487A} and* ***C)*** *CTS$^{D606N/E607Q}$. By using eGFP standards and densitometry the yield of all constructs expressed from 1L of Sf9 cell culture was calculated at around 0.5 mg/1L of Sf9 cell culture.* ***M****, molecular weight marker;* ***eGFP*** *(~25 kDa band), enhanced green fluorescent protein standards of known concentration, reported in nanograms (ng);* ***green arrows****, CTS/mutant specific bands. Bands at 25-32 kDa eGFP. Lower anti-GFP reactive bands likely represent degradation products and higher anti-GFP reactive bands.*

3.2.2 Detergent and Solubilisation Optimisation on CTS^{WT}

Given that solubilisation is a critical step in the purification of membrane proteins and one in which a significant amount of protein can be lost, a detergent optimisation screen was carried out to compare some commonly used detergents (Fig. 3-4). Quantification was carried out using densitometry, showing that N-dodecyl-β-D-maltopyranoside (DDM) was the most effective at solubilising CTS^{WT} from Sf9 cell membranes with an efficiency of ~64%, closely followed by the widely used detergent Lauryl maltose neopenyl glycol (CTSG), with an efficiency of ~43%. CTS was solubilised very poorly by (4-cyclohexyl-1-butyl-β-D-maltoside (Cymal-4) and octyl glucose neopentyl glycol (OGNG), and not at all by dodecyl glucose neopentyl glycol (DGNG). However, it should be noted that the band intensity of the pellet for OGNG

was weaker than for DGNG. This may have simply been caused by improper resuspension of the pellet for the OGNG condition, therefore giving the appearance of reduced signal by western blot, and therefore it is difficult to estimate the relative solubilisation efficiency of OGNG. Nonetheless, given that DDM and LMNG were clearly the best at solubilising CTS^{WT}, these detergents were chosen for purification and further work. It should also be noted that no differences in solubilisation were observed during purification experiments when the protein was solubilised overnight, for 2 hrs, or 1hr.

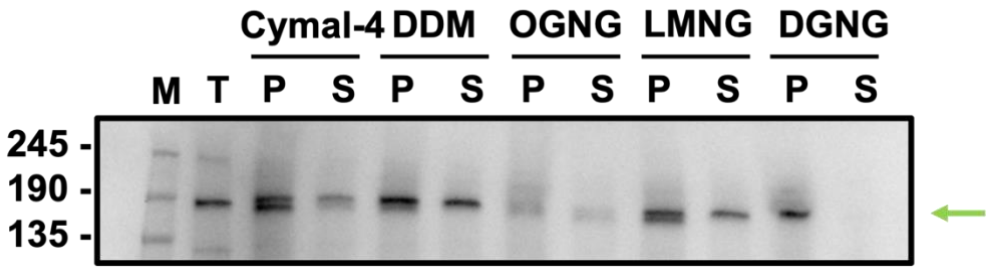


Figure 3-4. Solubilisation optimisation for CTS^{WT}. Western blot comparison between the solubilisation of Cymal-4, DDM, and three NG-class detergents in extracting CTS^{WT} from Sf9 cell membranes. Sf9 cell membranes were incubated with each detergent at a concentration of 2% (w/v) and the pellet and supernatant fractions compared by densitometry, showing that CTS^{WT} is most efficiently solubilised by DDM (~64% efficiency). ***M****, molecular weight marker;* ***T****, total protein;* ***P****, pellet;* ***S****, supernatant.*

A study by Geillon *et al.* (2017) showed that extraction of HsABCD1 from peroxisomal membranes was increased if peroxisomes were incubated with ATP before incubation with detergent. Given that structural studies often require significant amounts of protein, this simple modification to a solubilisation protocol was investigated as a method to improve CTS^{WT} yield for structural studies going forward.

Therefore, a similar experiment was carried out to see if incubation of CTS^{WT} Sf9 cell membranes with ATP would improve CTS^{WT} extraction using DDM. The experiment was carried out in triplicate, and as well as monitoring the solubilisation efficiency by measuring in-gel GFP fluorescence (Fig. 3-5b), the GFP signal of the resuspended pellet and supernatant was monitored by GFP fluorimetry (Fig. 3-5a). From this experiment, it was found incubation with 4 mM ATP did not improve CTS^{WT} extraction and instead DDM extracted CTS^{WT} from Sf9 cell membranes equally well whether the membranes had been incubated with ATP or not.

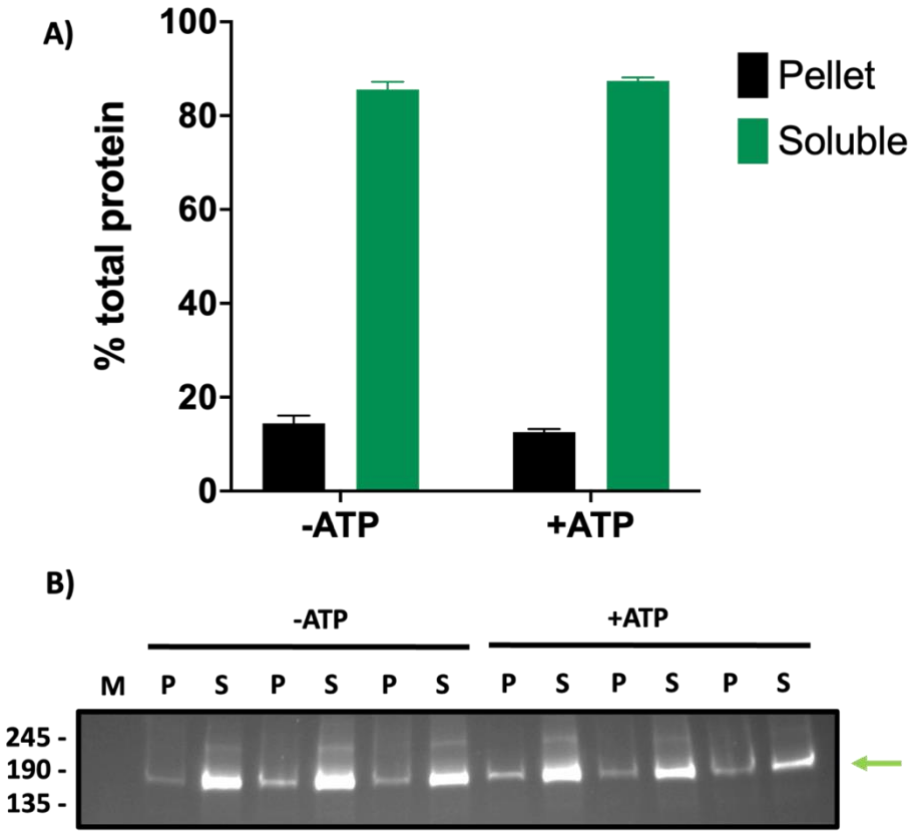


Figure 3-5. Incubation of Sf9 cell membranes containing CTS^{WT} with buffer containing ATP prior to solubilisation does not increase extraction yield. ***A)*** *Quantification of the solubilisation efficiency of DDM to extract CTS^{WT} from Sf9 cell membranes after incubation with or without ATP (4 mM) and $MgCl_2$ (2 mM), as measured by GFP fluorimetry, showing no difference between -ATP and +ATP conditions.* ***B)*** *In-gel GFP fluorescence of triplicate samples used for GFP fluorimetry quantification in panel A.* ***M****, molecular weight marker;* ***T****, total protein;* ***P****, pellet;* ***S****, supernatant. Error bars represent Standard Error of the Mean (SEM).*

In the previously developed purification protocol (Carrier *et al.*, 2019; De Marcos Lousa *et al.*, 2013) CTS^{WT} was solubilised using 2% (w/v) DDM at a membrane concentration of 2 mg/ml (i.e., total protein in the membranes). Whilst this can be used to extract sufficient quantified of CTS^{WT} from the membrane, it would not be suitable for large-scale purification due to the large dilutions required to dilute down the membranes and the associated expenditure on large amounts of expensive detergents. Therefore, based on the solubilisation trials, it was decided that further solubilisation would be carried out using DDM, but that the optimal membrane concentration for solubilisation, as well as the optimal concentration of DDM would be investigated. Sf9 cell membranes bearing CTS^{WT} were divided into aliquots and solubilised at varying concentrations (2 mg/ml, 4 mg/ml, 8 mg/ml, or 10 mg/ml), and with either 1% (/v) or 2% (w/v) DDM. By comparing the amount of CTS^{WT} in the soluble fraction of each condition and comparing this to the soluble fraction obtained using the zwitterionic detergent FC-14 (which was not chosen for further studies due to being a more harsh detergent than the others , it was determined that CTS^{WT} was solubilised equally well at 10 mg/ml of membranes as it was at 2 mg/ml of membranes (Fig. 3-6). Moreover, there appeared to be no distinguishable difference between using 1% or 2% DDM. Consequently, the purification protocol was modified taking these factors into account, where CTS^{WT} was subsequently solubilised from membranes at a total protein concentration of 10 mg/ml using 1% (w/v) DDM.

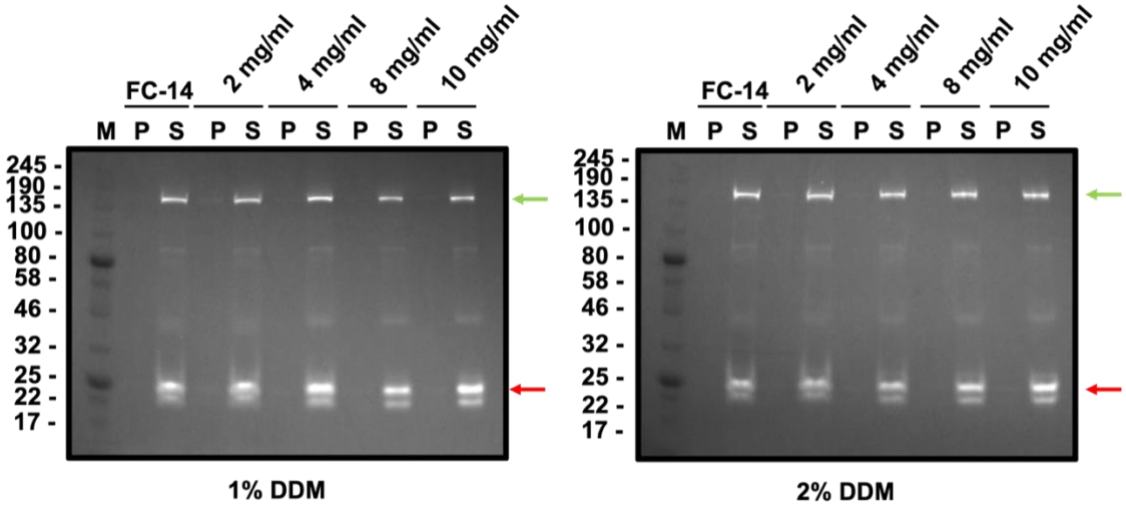


Figure 3-6. Solubilisation optimisation of CTS^{WT}. Sf9 cell membranes containing CTS^{WT} were diluted to varying total protein concentrations and solubilised using 1% or 2% (w/v) DDM. Comparison of the pellet and soluble fractions of each condition after solubilisation using in-gel GFP fluorescence demonstrated negligible differences in solubilisation efficiency between all conditions. ***Left panel****, solubilisation with 1% (v/v) DDM;* ***right panel****, solubilisation with 2% (w/v) DDM.* ***FC-14****, fos-choline 14;* ***P****, pellet;* ***S****, supernatant;* ***M****, molecular weight marker;* ***green arrows****, CTS^{WT} bands;* ***red arrows****, free GFP.*

Previous studies carried out in the Baker lab have shown that effective solubilisation of CTS^{WT} requires glycerol at a concentration of 20% (personal communication with previous lab members, Dr David Carrier, and Dr Carine De Marcos Lousa). Whilst this does not pose a problem at the early stages of purification, it is not suitable for EM studies because such concentrations of glycerol drastically reduce the contrast in EM micrographs (Sgro & Costa, 2018) Therefore, a significant change to the previous protocol was required to balance both stability and the suitability for downstream structural analysis. Therefore, to assess how well CTS^{WT} tolerates low glycerol concentrations, a GFP-based thermal stability assay was carried out. This method – which is based on the IMPROvER approach (Harborne *et al*, 2020) and similar to that described by Ashok *et al* (2015) – was used because the protein's melting temperature can be determined even in crude protein extracts after solubilisation by taking advantage of the eGFP-tag (Fig. 3-7).

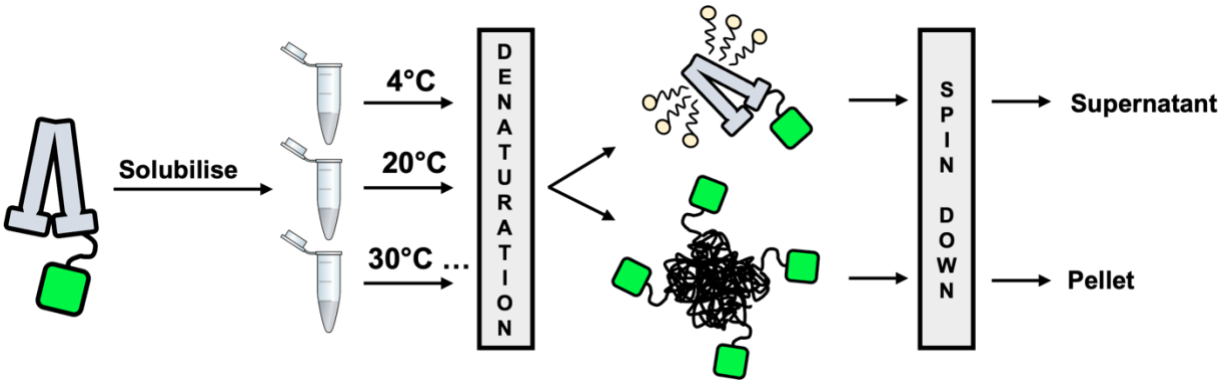


Figure 3-7. GFP-based thermal stability assay methodology. CTS is solubilised from its surrounding membrane using a detergent of choice and divided into aliquots that are heated at increasing temperatures. Above its melting temperature, CTS will begin to denature, leaving the more stable GFP-tag in a folded state. Heat-precipitated/denatured protein is removed from each aliquot by centrifugation, and the protein that remains in the supernatant in each condition is assessed by in-gel GFP fluorescence. ***Grey****, CTS^{WT};* ***green****, GFP-tag;* ***yellow****, detergent molecules.*

Using this method, three different glycerol concentrations were investigated: 5%, 10%, and 20% (v/v). CTS^{WT} (10 mg/ml) was solubilised from Sf9 cell membranes in 20 mM Tris pH 7.4, 200 mM NaCl, 1.5 mM $MgCl_2$, 5 mM EGTA 2% (w/v) DDM, and either 5%, 10%, or 20% glycerol, then divided into aliquots for incubation in triplicate at 4-100°C. After incubation, triplicate samples for each condition were subject to SDS-PAGE and bands visualised using in-gel GFP fluorescence. Using densitometry, the triplicate GFP signal intensity of each condition was compared to the 4°C control condition, and from this, the relative abundance of GFP remaining in solution after each incubation was determined. By plotting this against temperature and fitting the data using a four-parameter dose-response curve by non-linear least-squares fitting, it was possible to determine the T_m of the protein because the inflection point of the curve represents the T_M of the protein at which half of the protein is denatured (model parameters are presented in table 3.1).

Model Parameters		**Experimental Conditions**		
		5% glycerol	**10% glycerol**	**20% glycerol**
Best-fit Values	Bottom	-0.03	-0.025	-0.030
	Top	1.0	1.0	1.0
	LogIC50	55.7	59.7	55.0
	Hill Slope	-0.05	-0.06	-0.05
	Span	1.05	1.05	1.07
Goodness of Fit	Degrees of Freedom	26	26	26
	R squared	0.9879	0.9913	0.9913
	Sum of Squares	0.06899	0.05090	0.05090

Table 3.1. Model parameters for the IMPROvER-based GFP thermal stability assay to determine the T_M of CTS^{WT} in solution containing different concentrations of glycerol.

The T_m of CTS^{WT} in the 20%, 10%, and 5% glycerol conditions was calculated to be 55.7°C, 59.7°C, and 55.0°C, respectively (Fig. 3-8). Given the small difference between the T_m of CTS^{WT} at 20% and 5% glycerol, this data suggests that CTS^{WT} can tolerate glycerol concentrations as low as 5% and maintain stability comparable to when in solution with 20% glycerol. However, in the 10% glycerol condition there is a ~5°C increase in T_M compared to the 5% and 20% conditions. Typically, this would suggest a strong stabilising effect, but given that this condition sits between the 20% and 5% conditions and that their melting temperatures were similar, the T_M for the 10% condition does not appear to fit the trend (i.e., no difference in T_M between 20% and 5% glycerol). The significance of this observation is discussed in section 3.3.

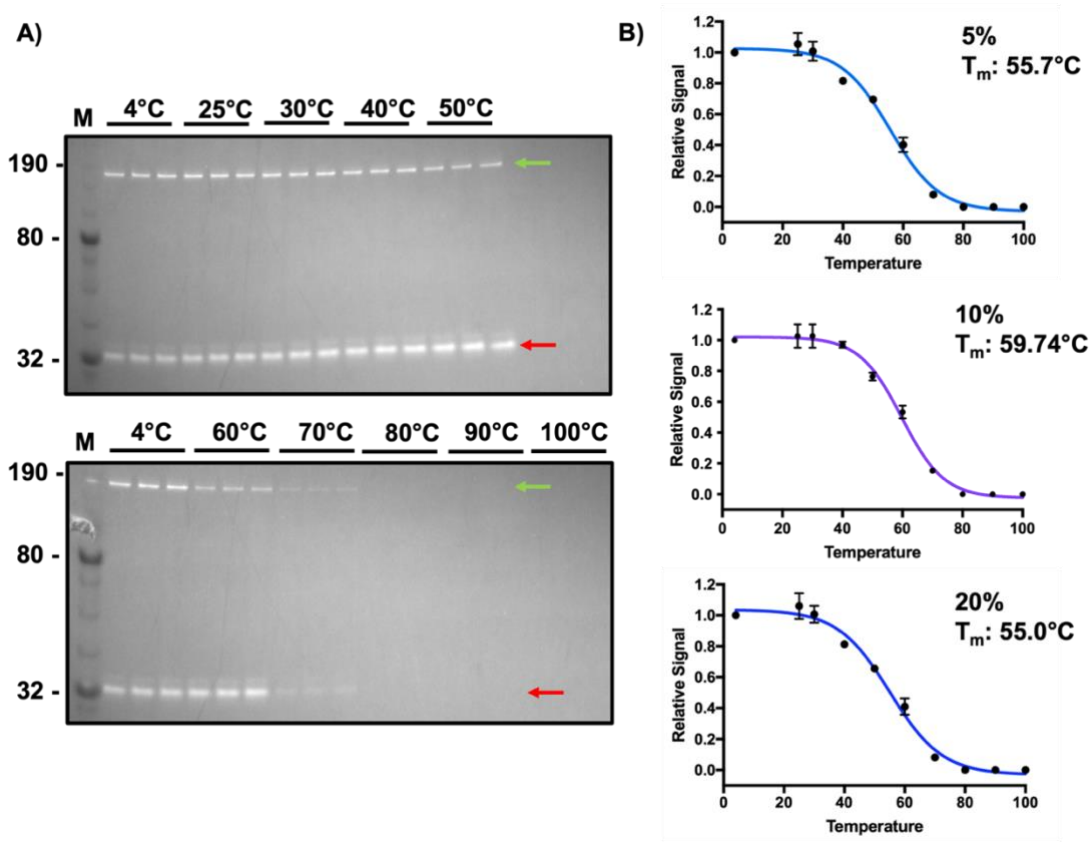


Figure 3-8. GFP-based thermal stability assay investigating the effect of glycerol on the T_m of CTS^{WT}. After DDM-solubilisation of CTS^{WT} from Sf9 cell membranes, the soluble fraction was equally divided into aliquots and each aliquot was incubated at a specific temperature. After incubation, the pellet and soluble fractions were separated, and the soluble fraction was run on SDS-PAGE. The amount of CTS^{WT} remaining in solution (i.e., in the soluble fraction) in each condition then observed by in-gel GFP fluorescence. ***A)*** *An example in-gel GFP fluorescence SDS-PAGE of CTS^{WT} membranes incubated at different temperatures, in buffer containing 5% glycerol. From the gel, the denaturation of CTS^{WT} can be observed between 50-60°C.* ***B)*** *Melting curves for CTS in buffer containing 5% (top), 10% (middle), and 20% glycerol (bottom).* ***Green arrows****, CTS^{WT} bands;* ***red arrows****, free eGFP bands;* ***M****, molecular weight marker. N=1, and therefore the glycerol effect is currently an inferred effect until proven in further studies.*

3.3 Discussion

To date, studies on CTS have focused primarily on the biochemical characterisation of the transporter (De Marcos Lousa *et al.*, 2013; Dietrich *et al.*, 2009; Footitt *et al.*, 2007; Footitt *et al.*, 2002; Nyathi *et al.*, 2010; Nyathi *et al.*, 2012; Theodoulou *et al.*, 2005), which ultimately led to the observation of CTS ACOT activity and identification of residues that may be involved (Carrier *et al.*, 2019). Despite this work, it is not clear where in this protein this activity is derived from or how the candidate residues thought to be involved in substrate hydrolysis contribute to the transport cycle (Carrier *et al.*, 2019). To address these outstanding questions, a new direction was pursued through structural biology. Accordingly, the work in this chapter details the very beginning of structural studies on CTS^{WT}, outlining the modifications made to existing protocols to ensure that the expression and solubilisation of CTS^{WT} was suitable for large scale purification of the transporter, and for downstream cryo-EM studies.

Early on, it was necessary to ensure that CTS^{WT} could be expressed at levels suitable for EM. Despite cryo-EM being much less reliant on high protein concentrations than x-ray crystallography, the fact that CTS^{WT} degrades so easily meant that the initial yield after expression would need to be high enough to offset these losses such that after purification, a concentration of ≥ 0.5 mg/ml could be achieved and, in a buffer that is suitable for EM studies (Drulyte *et al*, 2018; Kampjut *et al.*, 2021; Sgro & Costa, 2018). Because of the effects of vitrification during cryo-EM grid preparation, the concentration of protein on the grid is often not representative of that measured in solution (Klebl *et al*, 2020), and therefore, this concentration represents the lower value required for cryo-EM. Typically, cryo-EM grids are prepared at concentrations of ~ 1 mg/ml, though there is considerable variation (e.g., the GPCR Smoothened between 5 – 10 mg/ml (Qi *et al*, 2019) and the glucagon-like peptide-1 receptor between 15 – 20 mg/ml (Cong *et al*, 2021)). That considered, after expression testing and consideration of protein loss during the solubilisation and concentration steps, a

yield of ~ 0.5mg/L of culture was considered to be suitable for cryo-EM studies (indeed, after purification CTS^{WT} could be concentrated to 0.5 – 1 mg/ml; pages 114 and 122).

However, it is known that membrane protein overexpression can lead to the generation of mis-folded, aggregated, and inactive protein (Ferrer-Miralles *et al*, 2015; Tate, 2001; Wagner *et al*, 2007), with incredibly low yields. Moreover, the construct design is critical for the successful expression of the protein of interest to alleviate the above (Edavettal *et al*, 2012). Thus, whilst the amount of CTS generated from Sf9 cells in this study was deemed suitable for cryo-EM, it would be beneficial to understand how much of the CTS that is expressed is correctly folded, and indeed whether different constructs would behave better or worse in that respect that the constructs used in this study. Future work may wish to focus on these tasks to improve CTS expression. For example, a series of different constructs could be produced, incorporating different affinity tags, including stabilising mutations, deleting predicted disordered regions, truncations, or fusions. Once generated, such constructs could be expressed in Sf9 cells and their expression levels compared, and the quality of the expressed protein assessed using a simple technique such as that described by Thomas & Tate (2014). They found that membrane proteins generated in insect cells were expressed as a mixture of folded and unfolded proteins, whereas mammalian cells were more efficient at expressing the same proteins in a folded state. To date, CTS expression has not been attempted in mammalian cells as there is precedent for expressing the plant protein in Sf9 cells as this expression system has an optimal growth temperature (27°C) similar to that of CTS's host organism (*A. thaliana*, ~23°C). However, the study discussed above suggests that an investigation into mammalian expression may be useful.

The choice of detergent and efficiency of extraction is a critical part of membrane protein purification. Indeed, the extraction of ABC transporter from their native membrane using detergents is well documented (Ambudkar, 1995; Cai & Gros, 2003;

Meyer *et al*, 1994; Poolman *et al*, 2005; Sharom, 1995; Wright *et al.*, 2018) with many detergents showing varying extraction efficiencies on different ABC transporters. Thus, whilst previous work had used DDM to solubilise CTS^{WT}, solubilisation optimisation was necessary to ensure that the maximum extraction efficiency was achieved. The detergents used for screening and their concentrations were chosen based upon their success in solubilising various channels and transporters for structural studies. Moreover, these detergents were chosen because – apart from the similarity between OGNG and DGNG (differing only in their hydrophobic tail length) – they have different chemical structures. Therefore, it would be possible to eliminate such detergents that are likely not compatible with CTS should different detergents be required in future studies.

Of the five detergents tested, CTS^{WT} was most efficiently solubilised by those with the longest tail groups (DDM having a single 12-carbon chain and LMNG having two 12-carbon chains), followed by CYMAL-4 baring the shortest tail group (albeit it being an aryl chain). Interestingly, although neither OGNG nor DGNG solubilised CTS^{WT} well, there was a distinct different in their solubilisation efficiencies despite their tail length differing by only 2 carbons. However, both detergents solubilised CTS^{WT} less efficiently than the shortest chain detergent, CYMAL-4, which like DDM and LMNG, has a maltoside headgroup. Therefore, it appears detergents with glucose heads-groups are not efficient for solubilising CTS^{WT}. If in future studies new CTS constructs are pursued then further detergent screening may be carried but instead of utilising western blotting as the sole measure of solubilisation efficiently, it may be used in conjunction with fluorescence-size-exclusion chromatography (FSEC). This technique, which takes advantage of the constructs GFP tag, is capable not only of informing on expression level, but also on sample monodispersity, and approximate molar mass of unpurified samples (Kawate & Gouaux, 2006). This method has been used to aid in the study of many different integral membrane proteins, including ABC transporters (Ellinger *et al*, 2013).

Lastly, the total protein concentration at which the CTS^{WT}-containing membranes would be solubilised was 2 mg/ml according to the original protocol. Experiments showed that solubilisation of CTS^{WT} from Sf9 cell membranes was equally efficient at 2 mg/ml and 10 mg/ml total protein concentration, and with either 1% or 2% (w/v) DDM. Solubilisation at 10 mg/ml drastically reduced the amount of membrane dilution required as well as detergent used during solubilisation, and overall streamlined the downstream purification process.

The report by Geillon *et al.* (2017) showing an increased yield of soluble human ABCD1 after incubation of peroxisomes with solution containing ATP prior to DDM solubilisation, represented a potential simple method to improve the yield of CTS^{WT} during solubilisation despite CTS being extracted well by DDM (~64%-80% efficiency). In their study Geillon *et al.* (2017) it was hypothesised that this observation resulted from the binding of ATP to the transporter causing the transporter to adopt the more compact outward-facing that is more easily extracted from the membrane than the inward-facing state (here on referred to as the 'conformation hypothesis') or that ATP induces movement of peroxisomal ABCD transporters outside of detergent-resistant lipid rafts/microdomains (here on referred to as the 'microdomain hypothesis'). However, CTS^{WT} is known to be active when in Sf9 cell membranes (Carrier *et al.*, 2019; De Marcos Lousa *et al.*, 2013) and therefore able to bind and hydrolyse ATP, meaning it is cycling through inward and outward facing states, respectively. Despite the DDM extraction efficiency of CTS^{WT} from Sf9 cell membranes being higher than that for the extraction of human ABCD1 from peroxisomal membrane, the observation that the extraction efficiency of CTS^{WT} does not improve when membranes are incubated with buffer containing ATP contrasts with the conformation hypothesis. However, it should be noted that in future studies, more pronounced effect may be observed if non-hydrolysable ATP analogues are used due to their ability to lock the protein into the outward-facing state.

Moreover, like peroxisomes, Sf9 insect cells are also known to possess lipid rafts (Avisar *et al*, 2005), again contrasting the microdomain hypothesis. It should be noted that ABCD1 exists in peroxisomal membranes primarily as a homo-tetramers (Geillon *et al.*, 2017) and although CTS has been shown to form oligomers in *A. thaliana* membranes of at least three CTS molecules (De Marcos Lousa *et al.*, 2013), its oligomeric state in Sf9 cell membranes in unknown. Therefore, the different DDM extraction efficiencies between CTS^{WT} in this study and ABCD1 by Geillon *et al.* (2017) may have a basis in the transporter's oligomeric state. If the conformation hypothesis is true, repeating this experiment with the inactive $CTS^{D606N/E607Q}$ mutant or a non-hydrolysable ATP analogue such as AMP-PNP, that is incapable of changing conformations, the solubilisation efficiency should increase. Unfortunately, due to time constrains this experiment was not carried out.

Finally, using an IMPROvER-based thermal stability assay (Harborne *et al.*, 2020), the thermal stability of CTS^{WT} was assessed in buffers containing different amount of glycerol. Whilst high glycerol concentrations are necessary for solubilisation of CTS, this is not generally suitable for cryo-EM studies, especially those proteins or complexes below 500 kDa (Drulyte *et al.*, 2018) due to it reducing the contrast observed in EM micrographs. The T_M of CTS^{WT} in the 5% and 20% glycerol conditions was reported at ~55°C whereas the T_M of the 10% condition was reported at ~59°C. This increase in the apparent T_M of CTS^{WT} was unexpected due to this condition being in-between the 5% and 20% conditions that reported similar melting temperatures. This is intriguing because all GFP-thermal stability assays were carried out on the same batch of CTS^{WT} protein and in each experiment (5%, 10%, and 20% glycerol conditions), all temperature challenges were carried out in triplicate and the data processed in the same way. Whilst care should be taken when considering this 10% condition, it should be noted that the primary goal of this experiment was to ensure that a lower glycerol concentration, could be accommodated. Indeed, this was suggested by the 5% condition. Moreover, despite the increased T_M observed in the

10% condition, it would not have been suitable to use 10% glycerol in EM studies due to the large reduction in contrast, and especially not for relatively small proteins (in terms of EM) such as ABC transporters. However, this finding is interesting given that it may have application to non-structural studies for further characterisation of CTS^{WT} e.g., in transport assays or biophysical studies. Indeed, 4°C is a large increase when considering protein melting temperatures (Alexandrov *et al*, 2008; Ashok *et al.*, 2015), Ultimately, it was decided that 5% glycerol condition would be used for subsequent purification of CTS^{WT} due to the T_M between this and the 20% condition being negligible, and because a balance between protein stability and contrast in EM was required. Care should be taken however, as this experiment was carried out once per condition, and thus, the effect of glycerol is inferred until proven by further experimentation.

In summary, prior to this project, CTS^{WT} had not been studied using structural biology techniques, and therefore, existing protocols required modification to ensure that CTS^{WT} could be produced at a sufficient quantity and quantity for downstream purification processes and structural studies. Therefore, the work presented in this chapter demonstrates the establishment of a robust expression and solubilisation protocol for CTS^{WT} (and CTS mutants), confirming that these constructs are amenable for study by cryo-EM, and lays the foundation for purification and electron microscopy studies presented in subsequent chapters.

4 Studies Towards the Structure of Wild-Type COMATOSE (CTS^{WT})

4.1 Introduction

Before ABC transporter were studied using structural biology techniques, elegant biochemical experiments were carried out that laid the foundations for deciphering the catalytic transport cycle. Complementary experiments by Loo & Clarke (1995) and (al-Shawi *et al*, 1994) utilising cysteine-reactive N-ethylmaleimide - which inactivates the ATPase activity of the protein – and cysteine-free mutants showed that Mg-ATP binds to both NBSs of P-gp, and that the NBDs physically interact. Careful vanadate trapping experiments showed that orthovanadate completely inhibits P-gp ATPase activity by trapping the nucleotide in the NBS (Urbatsch *et al*, 1995b), and that trapping of nucleotide at one NBS prevents hydrolysis at the other, suggesting an alternating hydrolysis mechanism (Urbatsch *et al*, 1995a; van Veen *et al*, 2000). Further trapping experiments allowed Urbatsch *et al.* (1995a) to outline a scheme for Mg-ATP hydrolysis. Building on this, several groups –utilising a variety of different methods including spectroscopy, protease accessibility, and biophysical techniques – reported conformational changes in the TMDs of various ABC transporters upon ATP binding (Manciu *et al*, 2003; Sonveaux *et al*, 1999; Vigano *et al*, 2000). Whereas other biochemical experiments have shown that substrate binding to the TMDs increases the affinity of the NBDs for nucleotide (Gao *et al*, 2000; Qu *et al*, 2003). This then questions which comes first in the transport cycle, nucleotide binding or substrate binding. These data, as well as many other elegant experiment and observations are summarised by Senior *et al.* (1995) and Higgins & Linton (2004), and lead to the development of the ATP switch model. Thus, these importance of biochemical experiments on deciphering the catalytic cycle of ABC transporters cannot be understated. It was this work that allowed the results of structural studies to be inferred and placed into context.

Before the so-called ‘resolution revolution’ of cryo-EM, the majority of protein structures – both soluble and membrane – were solved using X-ray crystallography. Indeed, the first structure of an ABC exporter was solving using this method (Dawson & Locher, 2006), which was followed by crystal structures from several important ABC transporters including MsbA (Ward *et al.*, 2007) P-gp (Jin *et al*, 2012), TM287-TM288 (Hohl *et al.*, 2012), and MalFGK (Kadaba *et al.*, 2008). The crystal structures of these and various other ABC transporters were vitally important in understanding how these transporters function, and the knowledge gained from these studies laid the foundation for various mechanistic models describing the catalytic cycle of different transporters and their substrate specificities (Hohl *et al.*, 2014; Locher, 2016; Perez *et al.*, 2015; Ward *et al.*, 2007). However, two major lines of questioning remained; the first was that of substrate polyspecificity (i.e., how can a single transporter bind and transporter a wide range of chemical variable substrates, e.g., P-gp), and the second questioned the physiological relevance of specific transporter conformations, especially those obtained from crystal structures (Wright *et al.*, 2018). Around this time, technological advancement in the field of EM was taking off, which ultimately led to the resolution revolution (Kuhlbrandt, 2014). The resolution-revolution began with the development of direct-electron-detectors (DEDs) (Bammes *et al*, 2012; Jin *et al*, 2008; Milazzo *et al*, 2011) that, unlike photographic film or CCDs, can record single electron events in rapid succession. With this improved frame rate, ‘micrograph movies’ could be recorded during data collection and used to compensate for sample movement during each exposure. At the same time, many other developments arose, including software for automated data collection, motion- and CTF correction algorithms, particle picking algorithms, single particle analysis software packages, and crucially, improved accessibility of EM to structural biologists in general. Thus, structural biologist studying ABC transporters looked to cryo-EM as a complementary method to x-ray crystallography a means of answer their outstanding questions. Indeed, the number of protein structures solved by cryo-EM has risen dramatically

over recent years, whilst the number of crystal structures solved has remained relatively constant (Fig. 4-1).

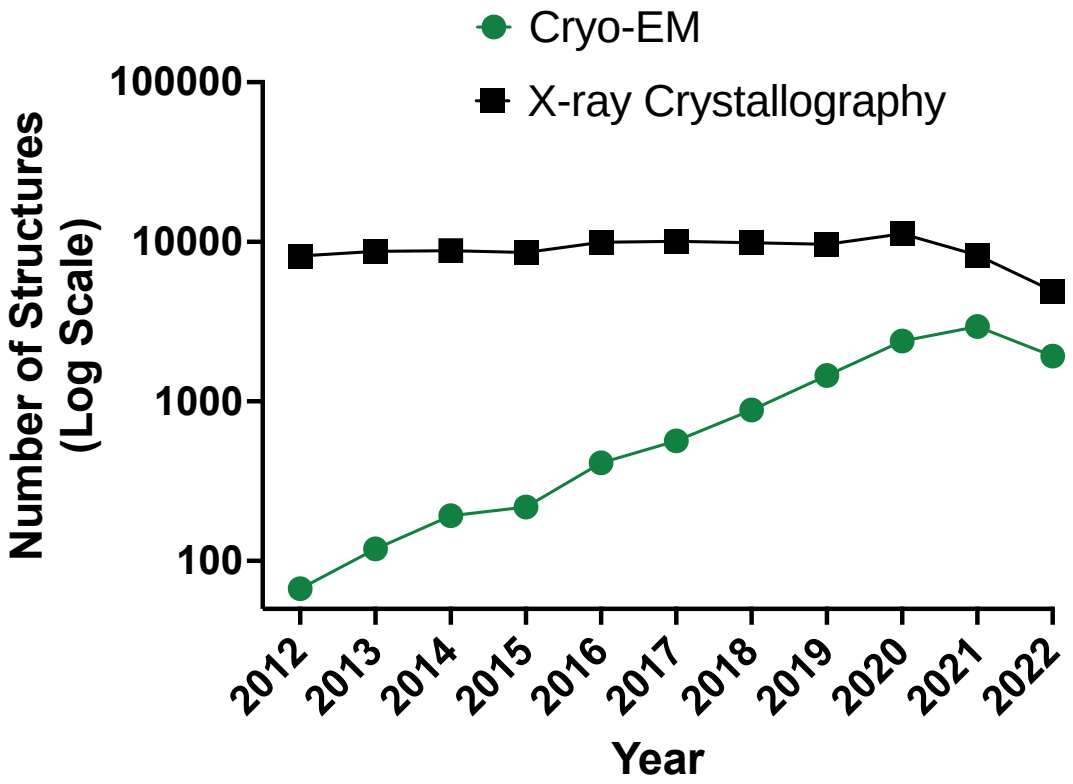


Figure 4-1.The number of structures deposited in the PDB per year by method. Whilst the number of crystal structures solved has remained constant over the last ~10 years, the number of structures determined by cryo-EM has steadily increased. Data (from the PDB) is accurate as of July 2022.

However, x-ray crystallography is still an incredibly powerful structural biology tool. For example, it is better suited than cryo-EM to provided high-resolution information as a function of time, where the use of x-ray free-electron laser light sources and allowed the development of serial femtosecond crystallographic analyses of various proteins (Keable *et al*, 2021; Liu & Lee, 2019). Moreover, x-ray crystallography is currently the best method to obtain true-atomic-resolution structures of proteins and is therefore more capable than cryo-EM in investigating the role of hydrogen-bonding

in macromolecules; this has recently been demonstrated by ultra-high resolution structures of the light-driven pump bacteriorhodopsin (Borshchevskiy *et al*, 2022). Finally, x-ray diffraction experiments can be carried out on protein within the lipidic cubic phase (LCP), showing that the crystal structure of membrane proteins (the structure of >700 solved in LCP) can be solved in a faithful membrane mimetic environment (Li & Caffrey, 2020).

However, cryo-EM has several key advantages over x-ray crystallography when studying ABC transporters. Firstly, while x-ray crystallography is capable of structure determination at incredibly high resolution, its major limitation is the necessity of forming crystals. Whilst some proteins crystallise readily, other proteins are much more recalcitrant, and this is especially true of membrane proteins usually because the presence of detergent can limit crystal packing and interrupt crystal contacts; in the best case this leads to poorly diffracting crystals, and in the worst case prohibits crystallisation entirely. Poor crystal packing typically results in a high crystal solvent content, which limits the resolution obtained during a diffraction experiment. All these problems are alleviated by cryo-EM since crystallisation is not necessary.

Secondly, despite their highly dynamic nature, ABC transporters packed into a crystal lattice can only adopt a single conformation. Conversely, during cryo-EM sample preparation, particles are suspended in vitreous ice and can adopt any conformation (Drulyte *et al.*, 2018; Naydenova & Russo, 2017). Because multiple conformational states may be identified in a single data set (Hofmann *et al*, 2019), this allows for the identification of rare conformation states, and the build-up of a more discrete view of the transporters catalytic cycle. Secondly, because of the ability cryo-EM to identify different conformational states, by adding substrate to transporters prior to making grids, this technique is well suited to obtaining substrate-bound structures of highly dynamic ABC transporters (Johnson & Chen, 2017; Manolaridis *et al.*, 2018; Scortecci *et al*, 2021; Wang *et al.*, 2021). Thirdly, most membrane proteins are generally expressed at low yield, with those that are eukaryotic and/or mammalian often

expressing even more poorly (Kesidis *et al*, 2020; Tate, 2001). Crystallisation screening often requires high concentrations of sample in relatively large volumes (ideally ~100 µl at >10 mg/ml). These sample requirements are often not obtainable for membrane proteins given their poor expression levels and propensity to degrade and aggregate during the purification process. Alternatively, preparation of a single cryo-EM grid typically requires just 3 µl of sample at ~1 mg/ml (Kampjut *et al.*, 2021; Schmidli *et al*, 2019), therefore, some proteins that will not concentrate well or crystallise at all can be studied using cryo-EM.

Importantly, it has been suggested that the use of detergents in structural studies of ABC transporters can cause proteins to adopt physiological irrelevant conformations, especially for crystal structures (Guo, 2020; Locher, 2016; Wright *et al.*, 2018). Detergent-free solubilisation methods such as nanodiscs (Ritchie *et al*, 2009) and styrene maleic acid (SMA) co-polymer (Lee *et al*, 2016b) may be used to overcome this problem due to their more native conditions. For example, SMA co polymer is capable of extracting proteins from the membrane whilst retaining their native surrounding lipids that may be critical for the proteins function. Nanodiscs require the purified protein to be reconstituted, but the addition of exogenous lipids during nanodisc formation provides a more native environment than detergent alone. Moreover, it has been shown that proteins solubilised in this way can be crystallised (Broecker *et al*, 2017; Nikolaev *et al*, 2017), this is non-trivial and requires its own optimisation. Alternatively, both methods – but predominantly nanodiscs – have been used successful in cryo-EM. Finally, whilst bigger is generally better for cryo-EM studies, the structure of many small proteins has been solved using this technique (Fan *et al*, 2019; Gopalasingam *et al*, 2019) – such as the 52 kDa streptavidin – meaning that CTS at 182 kDa (with the eGFP tag) is certainly amenable to cryo-EM. For these reasons, it was decided that structural studies on CTS^{WT} would be approached using cryo-EM. CTS^{WT} is a fragile ABC transporter that is subject to extensive proteolysis in its linker region that joins its two pseudo halves (Nyathi *et al.*,

2010). Despite expressing fairly well for a eukaryotic membrane protein (0.5 mg/L culture, section 3.2), it does not concentrate well and instead degrades significantly when the concentration is ≥ 1 mg/ml (observed by Dr David Carrier and in this work), therefore eliminating x-ray crystallography as an option. Moreover, to develop an understanding of the thioesterase activity exhibited by CTS^{WT} it would be ideal to solve the structure with substrate bound or capture the transporter in different conformation states. It is likely that significant conformational changes occur during substrate binding and whilst crystallisation may be capable of capturing one such state, the dynamic nature of ABC transporters means this challenge is better suited to cryo-EM.

Given that no structural data exists for COMATOSE, this chapter outlines the purification strategy used to purify CTS^{WT} – in detergent and detergent-free systems – for structural studies by cryo-EM, with the aim of solving a high-resolution structure of the transporter in an active form.

4.2 Results

4.2.1 Solubilisation and Purification of COMATOSE using SMA Co-polymer and Nanodiscs

Having established a detergent solubilisation method, it was decided that non-detergent methods would also be investigated for comparison and in an attempt to surmount the limitations of the former. Thus, attempts were made to purify CTS^{WT} using two detergent-free methods: SMA, which is capable of extracting membrane proteins directly from their native membrane and retaining native lipids (Lee *et al.*, 2016b) which may be critical for the proteins function (Qian *et al*, 2021; Stieger *et al*, 2021), and nanodiscs (Ritchie *et al.*, 2009), which require protein reconstitution after purification.

4.2.1.1 SMA Co-polymer

The first was the use of SMA co-polymer to solubilise CTS directly from Sf9 cell membranes for subsequently purification using the either the StrepII tag, or the eGFP tag. Unfortunately, whilst the SMA was capable of solubilising CTS with an efficiency of ~40%, SMA-solubilised CTS^{WT} was not able to sufficiently able to bind either Streptactin resin, or GFP-trap resin, with large amounts of CTS being found in the flow-through (Fig. 4-2). Whilst small amounts of protein were eluted from the streptactin resin, the lack of resin binding meant that it could not be purified in large amounts by affinity chromatography. Attempts to remove free-SMA after solubilisation by either dialysis or dilution did not improve resin-binding. This resin-binding problem after solubilisation is well-known to those in the field, but the cause of this is currently unclear. Anecdotal evidence suggests instead of modifying buffer composition away from conditions known to be successful (i.e., 500 mM NaCl, pH 8), greater consideration should be given to construct design, primarily the location of the protein's affinity tag and the nature of the tag itself (the Strep II tag appears to be particularly problematic) as both appear to be important factors. Finally, solubilisation

of CTS^{WT} using batches of newly prepared SMA abolished solubilisation entirely (data not shown). Therefore, no further experiments were carried out using SMA and instead, attempts were made to reconstitute detergent purified-CTS into lipid nanodiscs.

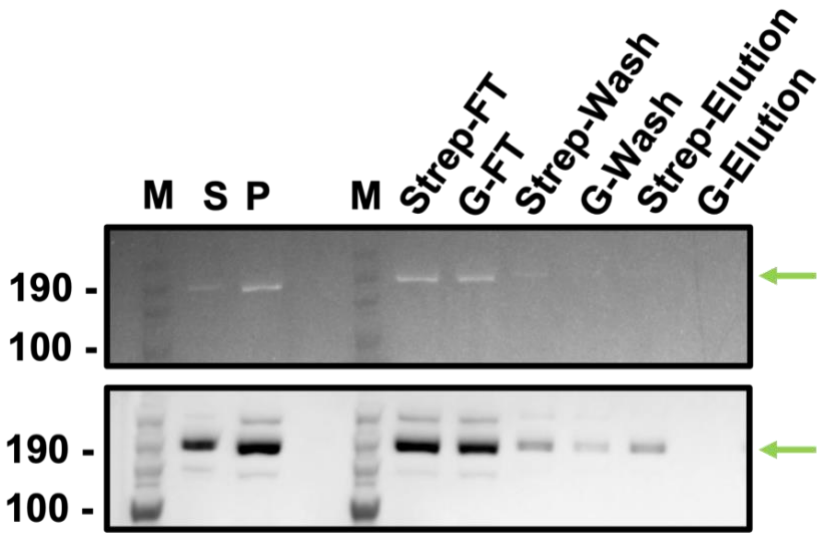


Figure 4-2. Small-scale purification of CTS^{WT} solubilised in SMA co-polymer. Whilst CTS^{WT} can be solubilised by SMA (~40% soluble), it does not efficiently bind to affinity resins for purification as evident by the intense flow-through band after incubation with either Streptactin or GFP-trap resin. ***Top panel****, in-gel GFP fluorescence;* ***bottom panel****, anti-GFP western blot;* ***M****, molecular weight marker;* ***S****, supernatant;* ***P****, pellet;* ***Strep****, streptactin resin;* ***G****, GFP-trap resin;* ***FT****, flow-through.*

4.2.1.2 Nanodiscs

Two different membrane scaffold proteins were used to reconstitute purified CTS^{WT} into lipid nanodiscs: MSP1D1 (diameter 9.5 nm) and MSP1E3D1 (diameter 12.1 nm) (Ritchie *et al.*, 2009). Optimisation of the protein:scaffold:lipid ratio is critical for the formation of lipid nanodiscs, and their formation is typically assessed by SEC (Ritchie *et al.*, 2009). However, this process can be time-consuming and more importantly, due to the low expression level of CTS^{WT}, dividing the purified sample in order to assess many different protein:scaffold:lipid ratios would mean that each SEC run

would not be capable of easily assessing the efficiency of reconstitution due to the low amount of protein injected onto the column. Instead, a gel-based approach was used by preparing different protein:lipid:scaffold mixtures, and after incubation and bio-bead detergent extraction, observing the amount of CTS^{WT} left in solution (i.e., after removal of detergent, whatever remains in solution must be being solubilised by a nanodisc). Several different protein:scaffold:lipid ratios were investigated for their ability to reconstitute CTS^{WT}; these were chosen to cover a wide enough range of permutations that had been used successfully for other ABC exporters (e.g. (Josts *et al.*, 2018). However, only one ratio – the ratio 1:3:40 using MSP1E3D1 – yielded a sufficient amount of CTS^{WT} in the soluble fraction (table 4.1).

MSP1D1		**MSP1E3D1**	
Ratio	**CTS in pellet or soluble?**	**Ratio**	**CTS in pellet or soluble?**
1:0:40	pellet	**1:0:40**	pellet
1:3:40	pellet	**1:3:40**	majority soluble
1:3:60	pellet	**1:3:60**	majority pellet, some soluble
1:3:120	pellet	**1:3:80**	majority pellet, some soluble
1:5:40	pellet	**1:3:120**	pellet
		1:1.5:40	majority pellet, some soluble
		1:1.5:60	majority pellet, some soluble
		1:1.5:80	majority pellet, some soluble
		1:1.5:120	majority pellet, some soluble

Table 4.1. Different protein:scaffold:lipid ratios tested for reconstitution of CTS^{WT} into lipid nanodiscs.

After running the samples on SDS-PAGE (Fig. 4-3), in-gel GFP fluorescence showed distinct degradation of CTS^{WT}, presumably occurring during the reconstitution incubation step, with free-eGFP visible at the bottom of the gel most prominently in each soluble fraction. Bands between 80 kDa – 46 kDa likely represent impurities from the MSP1E3D1 purification. The MSP1E3D1 1:3:40 ratio appeared to successful

reconstitute CTS^{WT} into a nanodisc as evidenced by the majority of protein remaining in the soluble fraction after reconstitution (Fig 4-3). However, it should be noted that due to time constraints, this experiment was not repeated.

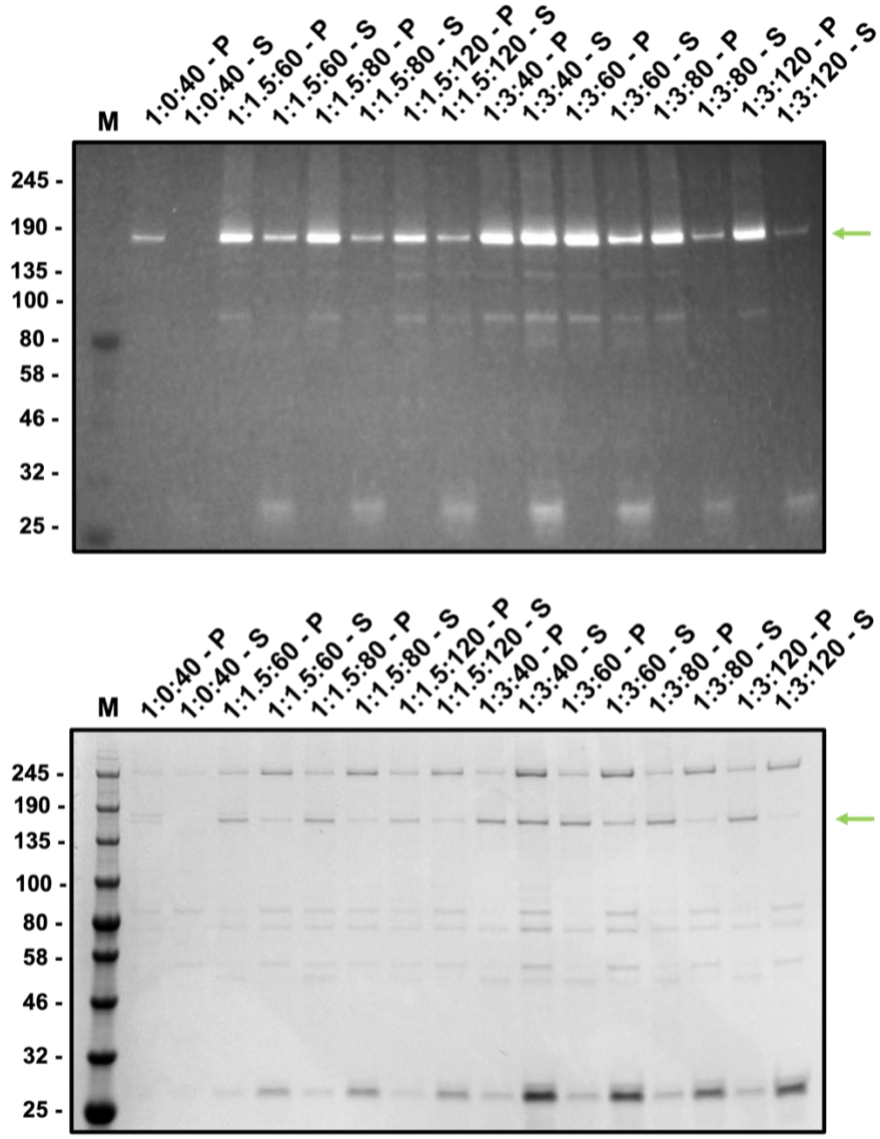


Figure 4-3. MSP1E3D1 nanodisc optimisation by SDS-PAGE. By comparing different ratios of CTSWT:scaffold:lipid after incubation and subsequent detergent removal using bio-beads, the ratio 1:3:40 appeared to result in the formation of a CTS^{WT}-nanodisc, with substantial amounts of CTS^{WT} remaining in the soluble fraction (i.e. the fraction in which CTS exists after detergent removal and solubilisation by a nanodisc). However, further experiments could not repeat this in large-scale (N=1). ***Top panel****, in-gel GFP fluorescence;* ***bottom panel****, Coomassie-staining of the gel in the top panel.* ***M****, molecular weight marker;* ***P****, pellet;* ***S****, soluble;* ***green arrow****, full length CTS bands;* ***top bands****, represent a co-purifying contaminant expressed in Sf9 insect cells.*

Therefore, this sample was assessed by negative-stain EM (Fig. 4-4). Despite the sample containing a few extra bands, given the size of CTS^{WT} relatively to these (except the 245kDa band), it was expected that CTS^{WT} would be easily observed if nanodisc formation had occurred properly and therefore, it would be possible to confirm whether or not nanodisc purification using this method was suitable. However, the expected particles were not observed (see figure 4-4 for an example of negatively stained nanodisc-reconstituted $MalFGK_2$ (Fabre *et al*, 2017)) and instead, a polydisperse sample was observed with significant heterogeneity and large, often aggregated particles, amongst a background of smaller particles. The large particles may represent the ~245 kDa protein observed by Coomassie-stained SDS-PAGE that was later confirmed as an unknown contaminant protein intrinsic to Sf9 cells that has a strong affinity for Streptactin XP Superflow High Capacity Resin but not Strepactin Superflow High Capacity Resin (see gel image in appendix; the XT resin was used during nanodisc experiments in an attempt to improve resin binding of CTS^{WT} as it was reported to have a higher binding affinity than the standard High Capacity Resin). Since a suitable reconstitution condition was not identified despite the devotion of a significant amount of time, nanodisc reconstitution was not pursued further and instead alternative methods were sought (section 4.2.2 and 4.2.3).

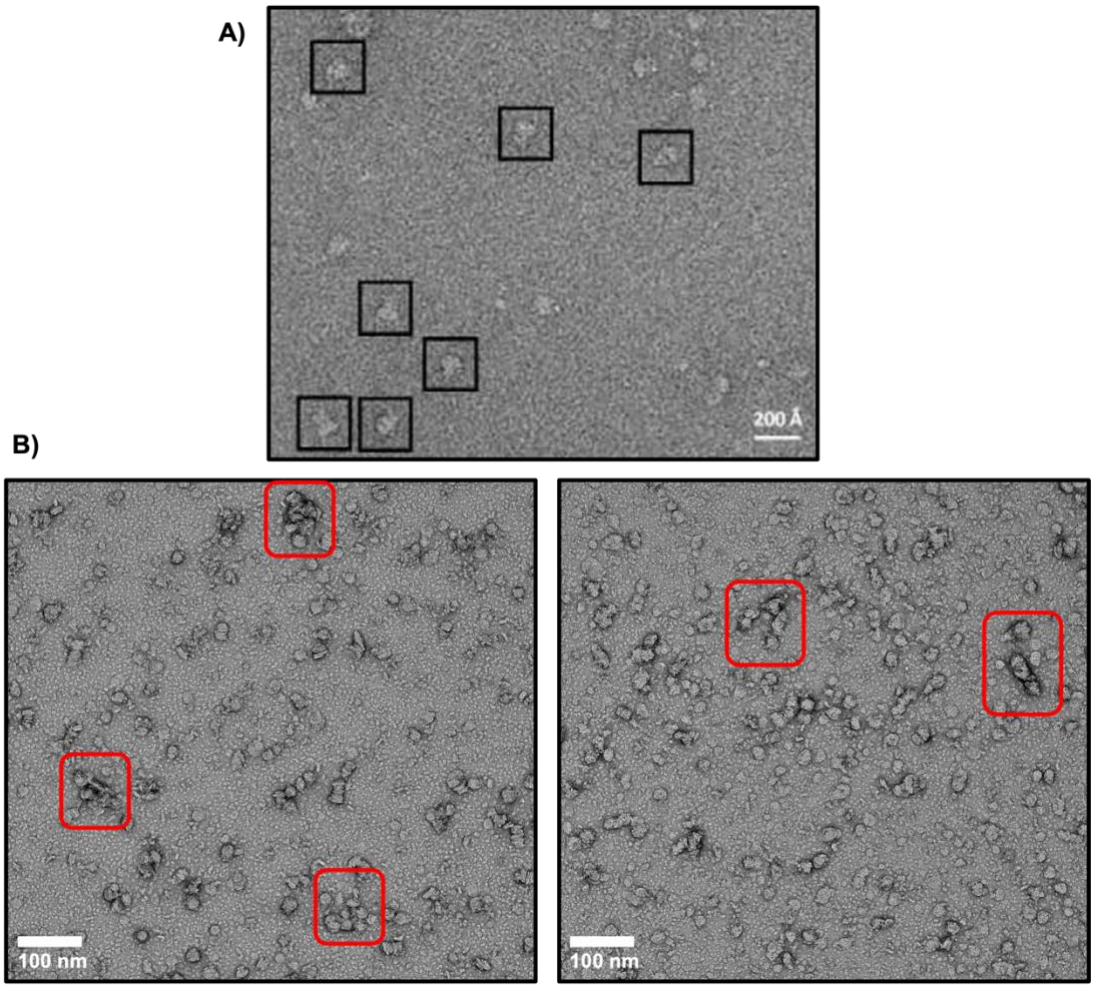


Figure 4-4. Negative stain EM assessment of CTS^{WT} after nanodisc reconstitution using MSP1E3D1 at a ratio of 1:3:40 (protein:scaffold:lipid). After purification, CTS^{WT} was combined with MSP1E3D1 and POPC lipids at a ratio of 1:3:40 for reconstitution. CTS^{WT} was separated from empty nanodisc and free scaffold protein by affinity chromatography then subject to negative stain EM. ***A****) A representative negative-stain electron micrograph of nanodisc-reconstituted $MalFGK_2$ showing expected particles (black boxes) for comparison to CTS^{WT}.* ***B****) Despite sample clean-up, the expected particles were not observed with grids showing distinct heterogeneity (aggregated particles are highlighted in red boxes) and polydispersity. Each panel represents a different grid square of the same grid. Scale bar = 100 nm.*

4.2.2 ATPase Activity on CTS Proteoliposomes

Prior to the work reported here, the ATPase activity of CTS^{WT} (and various mutants) had been reported from ATPase activity assays carried out on whole yeast peroxisomes (Carrier *et al.*, 2019; Nyathi *et al.*, 2010; Nyathi *et al.*, 2012) or in insect cell membranes (Carrier *et al.*, 2019; De Marcos Lousa *et al.*, 2013) . However, given that structural studies usually require purified protein, it was necessary to determine whether CTS^{WT} retains activity after purification. To do this, efforts were focused on reconstitution of the detergent purified protein (purified as in section 4.2.3) into proteoliposomes to carry out ATPase activity assays. Two proteins were used for the ATPase assays: CTS^{WT} and $CTS^{D606N/E607Q}$. $CTS^{D606N/E607Q}$ is a hydrolysis-deficient Walker-B mutant (section 3.2.1).

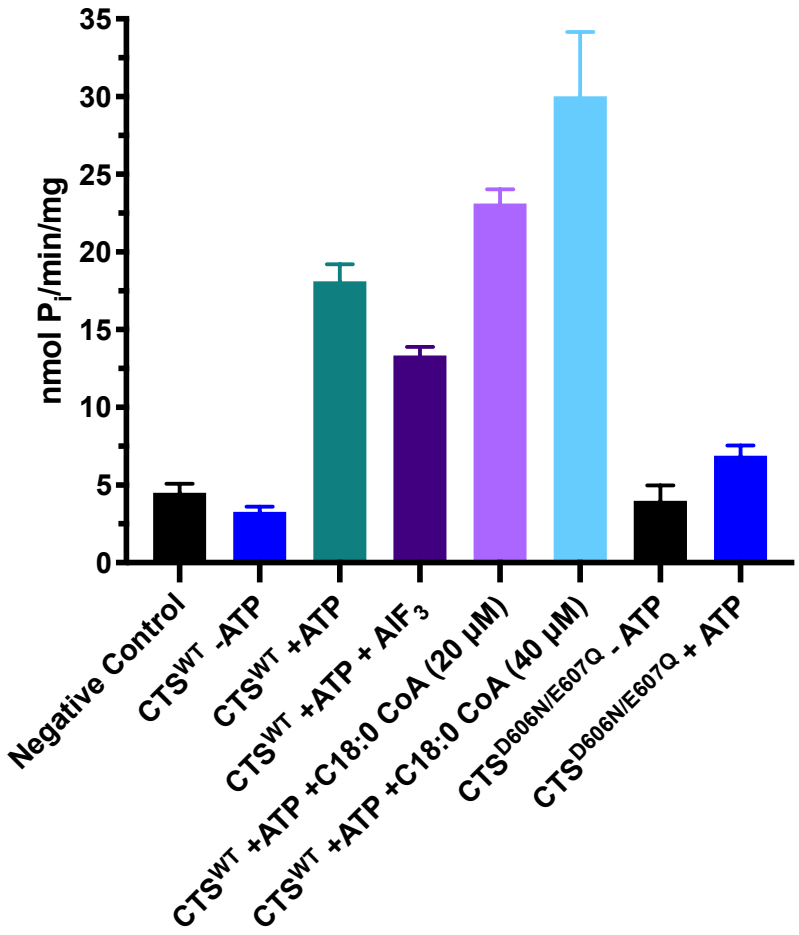


Figure 4-5. Reconstitution and ATPase activity assay of CTS. ***A)*** *An example of proteoliposome reconstitution of* CTS^{WT} *confirmed by sucrose density gradient centrifugation.* ***Left panel****,* CTS^{WT} *proteoliposomes after centrifugation in a sucrose density gradient, showing a discrete band of* CTS^{WT}*-containing proteoliposomes in the 10% fraction (white band; gradient 2.5% - 60% top to bottom).* ***Right panel****, Coomassie-stained SDS-PAGE showing a sample from each layer in the sucrose gradient, confirming that* CTS^{WT} *floats with the liposomes in the 10% layer.* ***B)*** *ATPase activity assay of* CTS^{WT} *and* $CTS^{D606N/E607Q}$*.* CTS^{WT} *shows substrate-stimulated ATPase activity with sensitivity to aluminium fluoride, whereas the catalytic mutant* $CTS^{D606N/E607Q}$ *does not. Each condition was measured in triplicate; errors bars are reported as SEM.* ***M****, molecular weight marker; percentages represent concentrations of sucrose in each layer of the gradient;* $\boldsymbol{AlF_3}$*, aluminium fluoride. N=1.*

As expected, when comparing the negative control (i.e., empty liposomes not containing CTS) to CTS^{WT} in the absence of ATP (CTS^{WT} -ATP), there was no distinguishable difference in the amount of inorganic phosphate released (Fig 4-5b). This provided a baseline for the assay at ~ 4 nmol P_i/mg/min. The activity of CTS^{WT} was confirmed by the observation of inorganic phosphate release at a rate of ~18 nmol P_i/mg/min. when in the presence of ATP (CTS^{WT} +ATP), a ~4.5-fold increase compared to the CTS^{WT} -ATP condition. The ATPase activity of CTS^{WT} was reduced when incubated with the transition-state inhibitor (CTS^{WT} +ATP +AlF_3), aluminium fluoride, immediately prior to the assay. This agrees with previous findings that the ATPase activity of peroxisomes purified from yeast expressing CTS^{WT} is inhibited by aluminium fluoride (Nyathi *et al.*, 2010). Interestingly, whilst the aforementioned study reported total ATPase inhibition (restoring the detectable activity to that of the negative control baseline), in the assay presented here, some ATPase activity remains. Concentrations of C18:0-CoA between 20 μM and 40 μM increased the rate of inorganic phosphate release to between ~22 – 37 nmol P_i/mg/min., a ~2-fold increase on the CTS^{WT} +ATP condition in the absence of substrate. However, this is still very low compared to other ABC transporters and compared to CTS in peroxisomes (Nyathi *et al.*, 2010) (i.e., not purified and reconstituted). It may be that the maximum ATPase activity of CTS can only be observed when the protein is reconstituted in the presence of acyl-CoA synthetases (e.g. LACS6/7) which CTS is known to make physical interaction with (De Marcos Lousa *et al.*, 2013). Thus, given that this is an N=1 experiment, future studies will be required to validate this result conclusively. Due to time constraints, it was not possible to include a condition measuring the activity of mutant $CTS^{D606N/E607Q}$ in the presence of substrate though this condition will be important in future activity experiments on reconstituted CTS. Compared to CTS^{WT}, the catalytic mutant $CTS^{D606N/E607Q}$ did not show any substantial ATPase activity in the presence of ATP, with an inorganic phosphate release rate of ~ 7.5 nmol P_i/mg/min., a ~2.4-fold decrease compared to wild-type CTS^{WT}. By

carrying out the assay using CTS -ATP, CTS +ATP, CTS +ATP +C18:0-CoA, and negative control conditions at 10- and 20-minute time points (originally 30 minutes), the activity measured was confirmed to be in the assay's linear range (see appendix). Taken together, these data suggest that after detergent purification and reconstitution, wild-type CTS still possesses substrate-dependent ATPase activity.

4.2.3 Purification of CTS^{WT} using DDM

Initial attempts to purify CTS^{WT} using detergents were made using DDM (Fig. 3-4 showing DDM as best suited for this). After solubilisation of CTS^{WT} from Sf9 cell membranes using a mixture of DDM, CHS, and asolectin and protease inhibitors (Proteoloc protease inhibitor cocktail, and AEBSF), the protein was bound to Strepactin Superflow High Capacity Resin (IBA) for purification. CTS^{WT} was eluted off the streptactin resin into buffer containing 20 mM Tris pH 7.4, 200 mM NaCl, 5% glycerol (v/v), 0.05% DDM/0.005% CHS/0.005% asolectin (w/v), and 5 mM desthiobiotin. Although the concentration of CTS^{WT} eluting from the streptactin resin was low, it was still visible by Coomassie staining (Fig. 4-6a). Whilst highly sensitive silver-staining after SDS-PAGE confirmed that the protein was highly pure after a one-step purification (Fig. 4-6b), after pooling of these fractions and concentration, CTS^{WT} displayed significant degradation and precipitation (Fig. 4-6a, right panel). This was evidenced not only by the increased intensity of bands present prior to concentration – which alone would simply indicate concentration of these minor contaminants – but also new bands between ~135 – 80 kDa, streaking, and a different intensity ratio between the two lowest bands observed in the silver-stain gel (i.e., concentration resulted in more degradation into the ~80 kDa band). Therefore, this limited the final achievable concentration of full-length CTS^{WT} in solution to ~ 0.6 - 1 mg/ml.

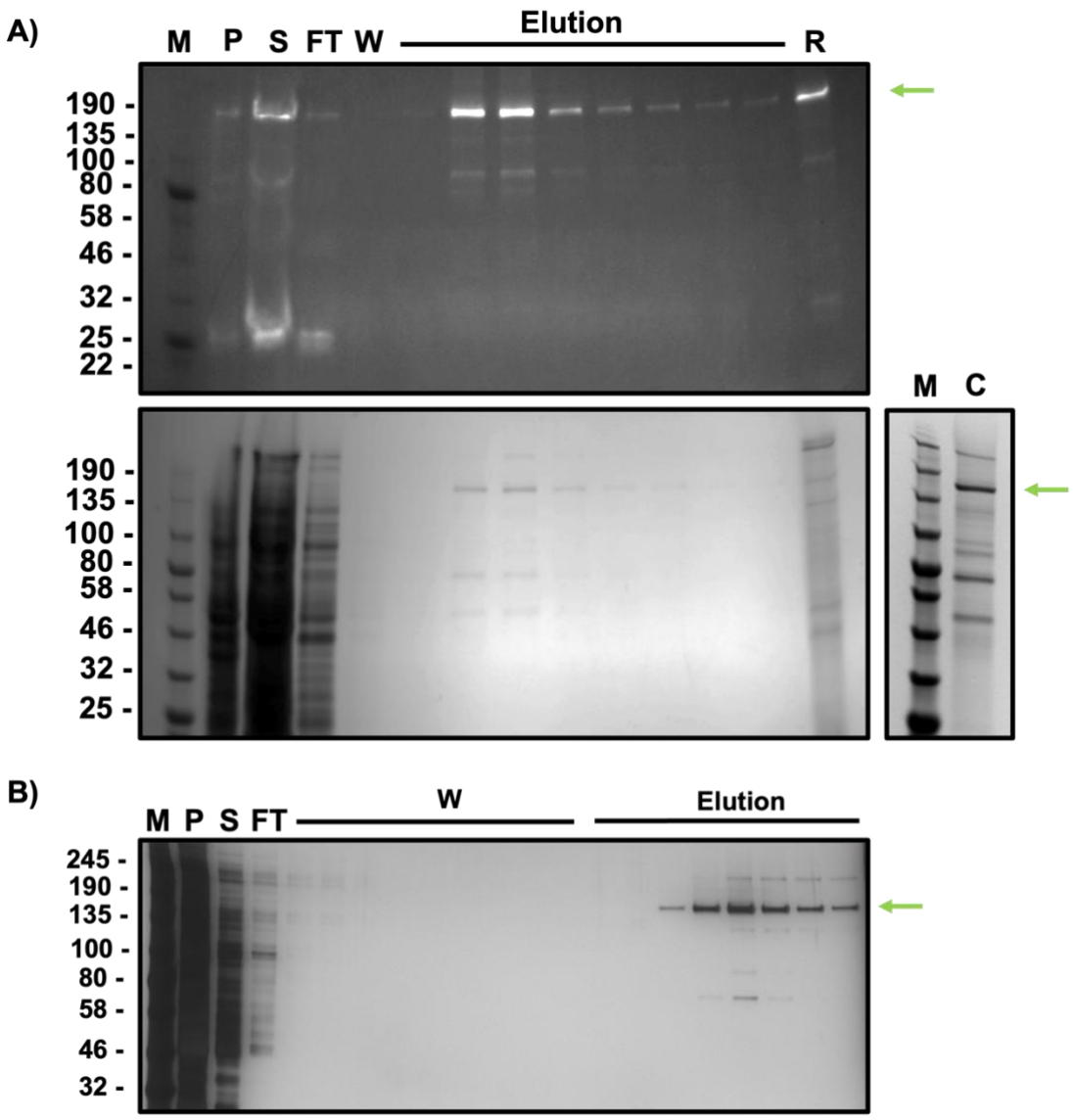


Figure 4-6. Purification of CTS^{WT} in DDM. ***A)*** *SDS-PAGE of Streptactin affinity chromatography purification of CTS^{WT} in DDM.* ***Top panel****, streptactin affinity purification showed by SDS-PAGE in-gel GFP;* ***bottom left panel****, affinity purification showed by SDS-PAGE stained with Coomassie;* ***bottom right panel****, SDS-PAGE stained with Coomassie showing affinity purified CTS^{WT} after concentration showing substantial degradation* ***B)*** *Silver-staining showing high purity of CTS after one-step streptactin affinity purification.* ***M****, molecular weight marker;* ***P****, pellet;* ***S****, supernatant;* ***FT****, Flow-through;* ***W****, washes,* ***green arrows****, CTS^{WT} band;* ***R****, resin sample. A and B are different gels from different purifications.*

Although ~ 1 mg/ml is suitable for cryo-EM studies, a higher concentration is desirable, mostly because of the number of particles required for high-resolution reconstructions. Therefore, to see if this concentration could be improved, three different centrifugal concentrators were investigated to assess how they affected the degradation of CTS^{WT}. After purification, eluted CTS^{WT} was pooled and equally divided between three 100 kDa molecular weight cut-off (MWCO) concentrators i.e., Amicon® (regenerated cellulose), Proteus X-spinner (polyethersulphone), and Vivaspin® (polyethersulphone), and concentrated until each sample reached ~ 100 µl. The concentration of each sample was then measured by absorbance at 280 nm and sample degradation was compared by running each sample on SDS-PAGE. From this, the Amicon® concentrator appeared to yield more protein but given the more pronounced degradation, this is likely the result of an erroneous increased amount of starting material compared to the other concentrators. Therefore, no meaningful difference between the ability of the three different concentrators to concentrate CTS^{WT} was noted, but substantial degradation was observed in all samples. Due to the limitations of A280 (nanodrop) measurements for particulate samples, in future studies it would be useful to quantitively assess the extent of CTS degradation using techniques such as dynamic light scattering. Nonetheless, the degradation problem appears to be intrinsic to CTS^{WT} (Fig. 4-7).

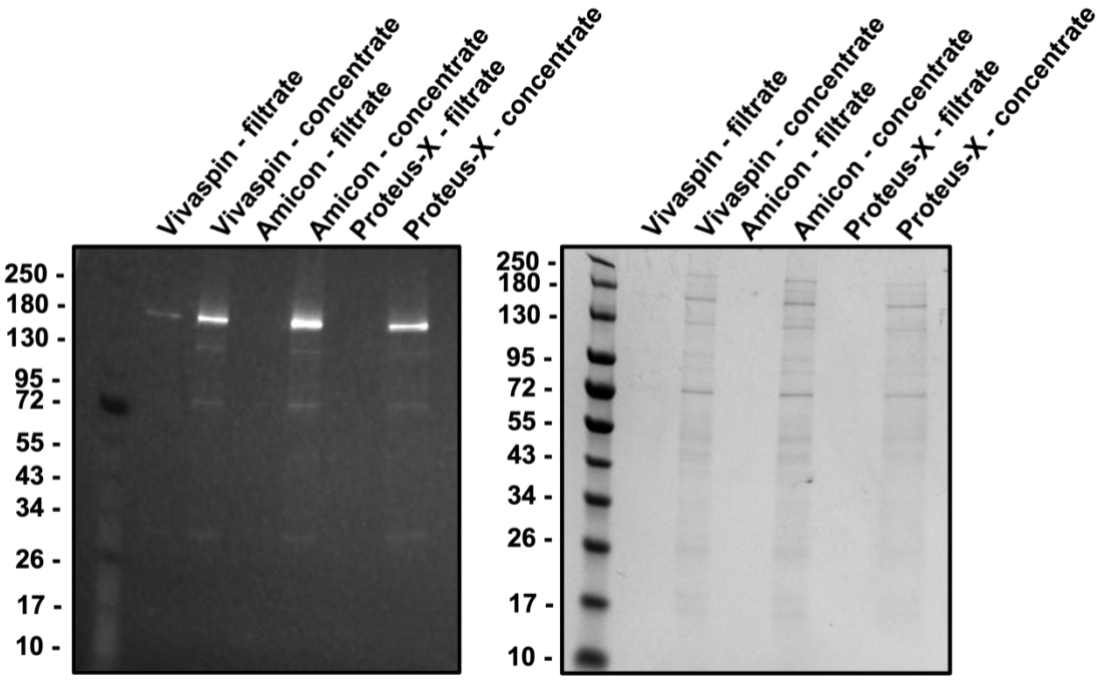


Figure 4-7. Investigating the effects of different concentrators on the degradation of purified CTS^{WT}. After concentration of CTS^{WT}, the three different concentrators tested all performed approximately equally, with all showing significant degradation of CTS^{WT} (the degree of degradation after concentration is highlighted by comparison to Fig. 4-6a).

To produce a sample of sufficient quality for cryo-EM, the sample must be highly pure, monodisperse and homogeneous (although some heterogeneity is *always* present). So, a size-exclusion chromatography (SEC) step was employed to clean up the sample after affinity purification and concentration to remove degradation products, aggregates, and excess empty micelles (Fig. 4-8). In this case, micelles would be present from DDM and from the phospholipid mixture, asolectin. Whilst this lipid is useful for the solubilisation of CTS^{WT} (personal communication with Dr David Carrier), its inclusion in buffers downstream would hinder EM studies due to it compromising a mixture of phospholipids that can form micelles of various sizes in solution. After concentration, purified CTS^{WT} was injected onto a Suprose 6 1/150 increase SEC column equilibrated with 20 mM Tris pH 7.4, 200 mM NaCl, 2.5% glycerol (v/v), 0.03%

DDM (w/v), to reduce the DDM concentration to ~3X CMC, and entirely remove asolectin and desthiobiotin.

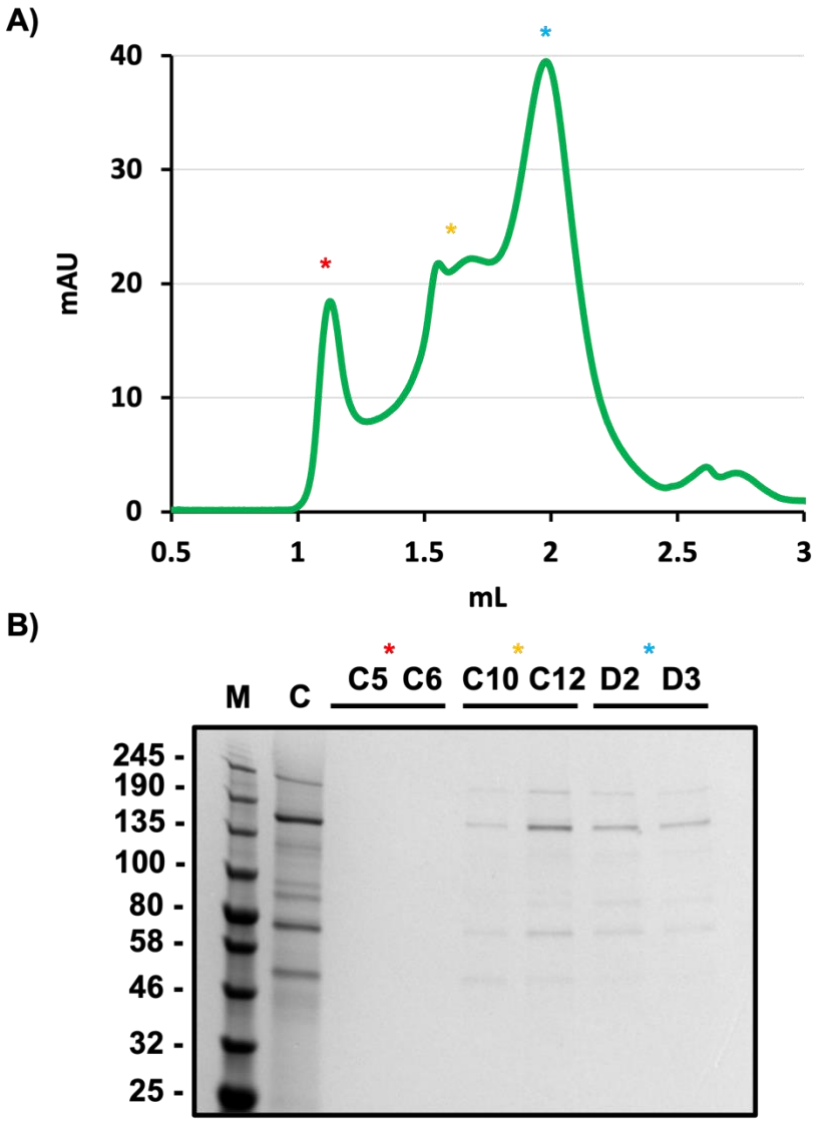


Figure 4-8. Size-exclusion chromatography on CTS^{WT} in DDM. ***A)*** *SEC chromatogram of DDM-solubilised CTS^{WT} showing three peaks (highlighted with red, yellow, and blue asterisks).* ***B)*** *SDS-PAGE showing concentrated CTS^{WT} prior to SEC, and two fractions from each SEC peak (C5 & 6, C10 & 12, and D2 & D3), with CTS^{WT} in fractions C10, C12, D2, and D3.* ***M****, molecular weight marker;* ***C****, concentrated CTS prior to SEC.*

After a significant void peak came a doublet-peak followed by a larger peak; two fractions collected from under each peak confirmed the presence of CTS^{WT} in the doublet and the largest peak but not the void, thereby confirming that CTS^{WT} had been successful purified away from any aggregates and at the same time, the amount of empty micelle in solution reduced. According to SDS-PAGE of the SEC fractions, CTS^{WT} was most prominent in the right-most peak of the doublet in fraction C12. Also present in this lane was a band running between 58 - 80 kDa. The nature of this band, which is almost always present when running purified CTS^{WT} on SDS-PAGE, can be speculated on by considering the structure of CTS^{WT}. Unlike its homodimeric human homologs (i.e., ABCD1-4), CTS^{WT} is a fused heterodimer where the two asymmetric halves of the transporter are joined by a disordered linker. Thus, it seems that this band appears to be a bifurcation of CTS^{WT} somewhere in this linker region, and its presence before SEC *and* afterwards highlights the fragility of CTS^{WT}.

Whilst C12 looked like a promising sample – best exemplified by comparing the concentrated CTS^{WT} sample before and after SEC – negative stain EM comparing fractions C11, C12, D2, and D3 confirmed that the sample was still too heterogeneous, with the difference between CTS^{WT} and empty micelles being too small to discern (Fig. 4-9). To reduce the presence of empty micelles, a further SEC purification was carried out, lowering the final DDM concentration, from 0.03% to 0.025%. However, this experiment was not capable of improving negative stain grid quality and instead increased aggregation. Therefore, different methods were investigated in an attempt to begin cryo-EM studies.

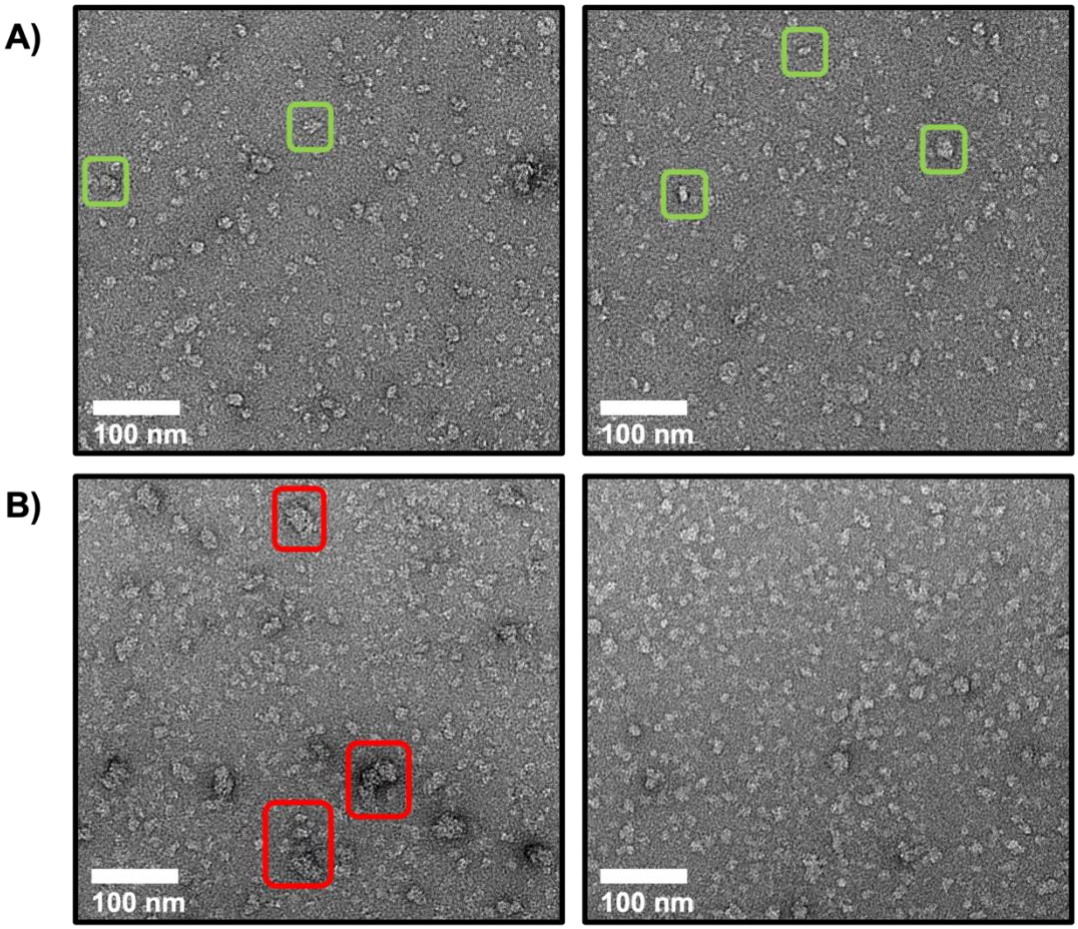


Figure 4-9. Negative stain on SEC-purified CTS^{WT} in DDM. ***A)*** *Two micrographs from fraction C12 from figure 4-8. Although this fraction contained the highest concentration of CTS^{WT}, high background and heterogeneity (green boxes showing particles of different shape and size) made distinguishing CTS^{WT} particles difficult.* ***B)*** *Peak fraction from a further SEC purification to reduce DDM concentration from 0.03% to 0.025%. Reducing the concentration of DDM to reduce the number of empty micelles present did not improve the grid quality, and instead, more sample aggregation is evident (red boxes). Scale bars = 100 nm. Micrographs in A) and B) collected at 50K and 55k magnification, respectively.*

4.2.4 Purification of CTS^{WT} using LMNG/Digitonin

Due to unsuccessful attempts to purify DDM-solubilised CTS^{WT} to a quality suitable for cryo-EM, new methods were sought after. It was around this time that the structure of the cobalamin transporter, ABCD4, was solved, revealing the first structure for a family D ABC transporter (Xu *et al.*, 2019). When comparing the purification protocol of ABCD4 to that used for CTS^{WT}, it was clear that the protocols were very similar, with the only major difference being the detergents used (table 4.2). However, detergent screening had previously confirmed that like ABCD4, CTS^{WT} could be appropriately solubilised using LMNG, meaning the only significant difference between the two protocols was the exchange and elution into digitonin. Additionally, given that CTS^{WT} and ABCD4 belong to the same family, and that purification using SMA and nanodiscs had already been attempted, it was decided that CTS^{WT} would be purified using a protocol similar to that for ABCD4.

	CTS^{WT}	**ABCD4**
Expression system	Sf9 insect cells	HEK 293F cells
Solubilisation	20 mM Tris-HCl, pH 7.4, 200 mM NaCl, 20% glycerol (v/v), 2 mM $MgCl_2$, 1% DDM/C0.1% CHS (w/v), 2 h, 4°C 1 mM DTT 1X Proteoloc™ Protease Inhibitor Cocktail 1X AEBSF	20 mM Tris-HCl pH 7.5, 150 mM NaCl, 20% glycerol (v/v), 2 mM $MgCl_2$, 2% LMNG/0.2% CHS (w/v), 5 mM ATP, 1.5 h, 8°C
Affinity resin	Streptactin resin	Flag resin
Elution buffer	20 mM Tris-HCl, pH 7.4, 200 mM NaCl, 5% glycerol (v/v), 2 mM $MgCl_2$, 0.05% DDM/0.005% CHS (w/v)	25 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5% glycerol (v/v), 2 mM MgCl2, 0.06% digitonin (w/v)
SEC	Superose® 6 Increase 5/150 20 mM Tris-HCl, pH 7.4, 150 mM NaCl, 5% glycerol (v/v), 2 mM $MgCl_2$, 0.03%-0.025% DDM/0.003%-0.0025% CHS (w/v)	Superdex® 200 Increase 10/300 25 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5% glycerol (v/v), 2 mM MgCl2, 0.06% digitonin (w/v) 1 mM TCEP

Table 4.2. Comparison between purification protocols for CTS^{WT} and ABCD4 (Xu et al., 2019).

Sf9 cell membranes containing CTS^{WT} were solubilised in buffer containing 1% LMNG/0.1% CHS (w/v) supplemented with 4 mM ATP. After solubilisation, CTS^{WT} was bound to strepactin affinity resin, washed, and eluted in 20 mM HEPES pH 7.5,

200 mM NaCl, 2.5% glycerol, 2 mM MgCl2, 0.06% digitonin + 5 mM desthiobiotin (Fig. 4-10a). The protein was concentrated (~ 0.9 mg/ml) and then assessed by SDS-PAGE; this showed that whilst degradation occurred, it was less pronounced than when CTS^{WT} was purified in DDM, with more protein being retained after concentration (Fig. 4-10b). However, it should be noted that unlike the purification in DDM, during this purification, ATP was present during the solubilisation step which may have improved the protein's stability.

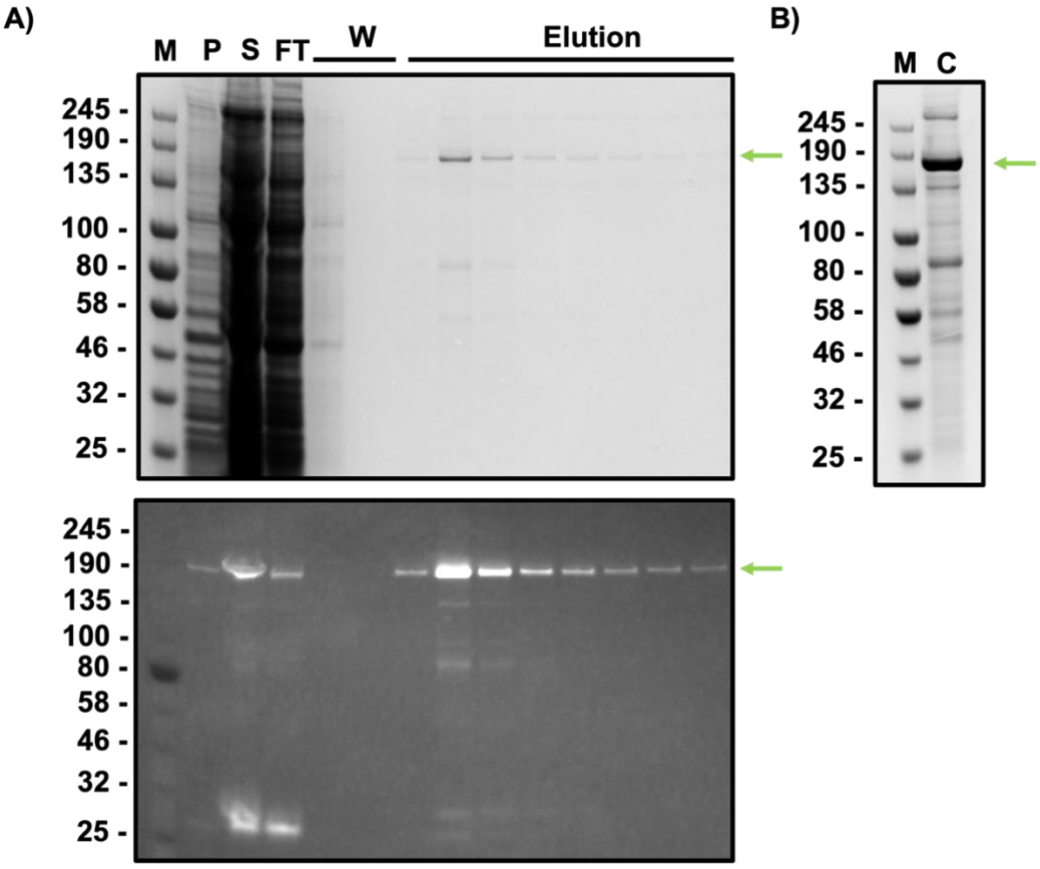


Figure 4-10. Purification of CTS^{WT} in LMNG/Digitonin. ***A) Top panel****, coomassie-stained SDS-PAGE of affinity purified CTS^{WT}, showing highly pure elution fractions.* ***Bottom panel****, corresponding in-gel GFP image.* ***B)*** *Concentrated elution fractions. Whilst CTS^{WT} still degrades using this protocol, its relative abundance compared to degradation products is higher than when purified in DDM.* ***M****, molecular weight marker;* ***P****, pellet;* ***S****, supernatant;* ***W****, washes;* ***C****, concentrated protein.*

Initial efforts to subject the concentrated protein to SEC were hindered by precipitation of the digitonin in the SEC buffer. Due to this, and due to the sample being relatively clean with CTS^{WT} being the dominant species in solution, the first cryo-EM grids were prepared using standard Quantifoil 1.2/1.3 300 mesh grids. During screening of these grids (Fig. 4-11) it was noted that streaks were visible across the grid; this was mostly likely caused by improper formation of vitreous ice. Additionally, on some areas of the grid the ice was too thick to image through, and other areas suffered from surface contamination or cracking of the carbon support. Nonetheless, it was evident that there was a reasonable distribution of particles on these grids, and therefore this represented a promising first attempt at cryo-EM studies on CTS^{WT}. Consequently, efforts were focused on cryo-EM grid optimisation.

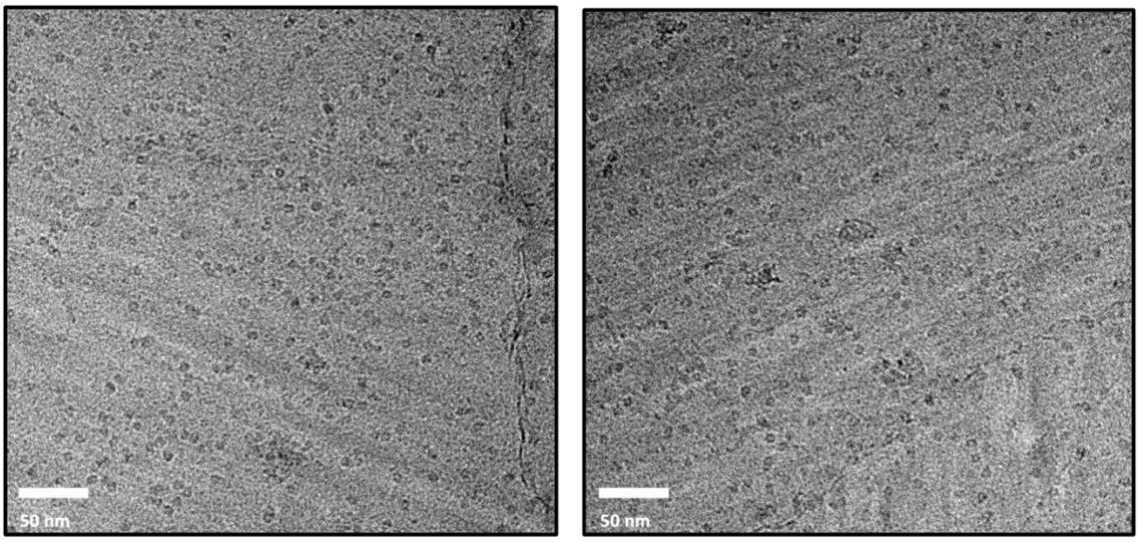


Figure 4-11. The first cryo-EM screening of CTS^{WT} purified in digitonin. Two micrographs taken from the same grid are shown, with a reasonable distribution of particles being observed in both micrographs amongst a background of improperly formed vitreous ice. Images collected at 130,000x magnification. Scale bars = 50 nm.

After testing different blotting times to optimise ice-thickness, CTS^{WT} appeared to be most well distributed on grids prepared using a blot time of 3 seconds. Unfortunately, after repeat purification of CTS^{WT} and preparation of the samples on the same grids, further screening revealed that it was not possible to reproduce the particle distribution observed in figure 4-11. Whilst the streaking problem was eliminated, further screening on Quantifoil grids (Fig. 4-12) revealed a significant sticking of particles to the carbon support of the grid.

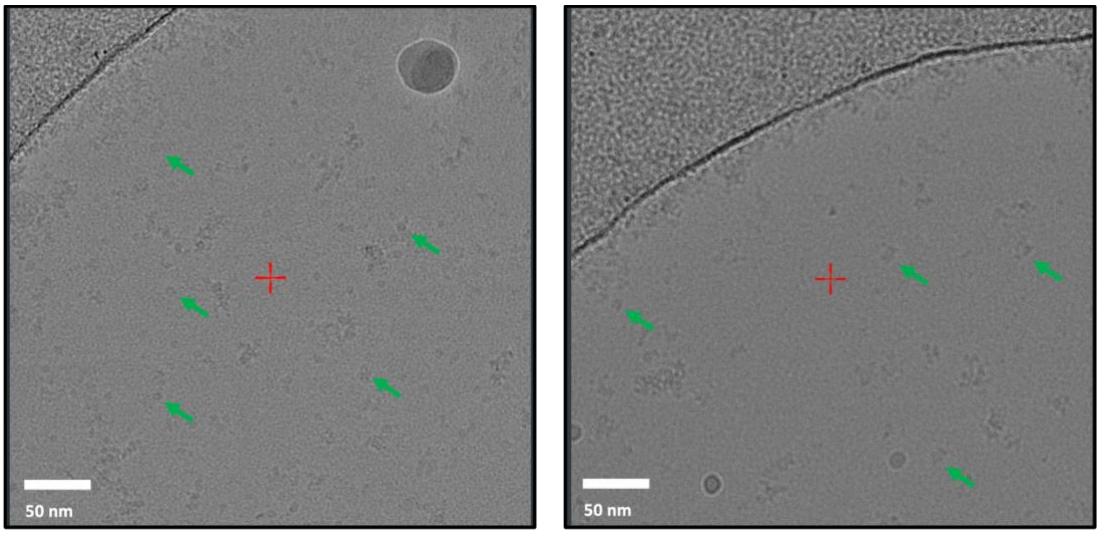


Figure 4-12. Cryo-EM screening of CTS^{WT} purified in digitonin on Quantifoil grids. Two micrographs taken from different grid squares show particles sticking to the carbon support. Despite this, particles can be observed, albeit with some aggregation. Images collected at 130,000x magnification. Green arrows point towards representative particles that would be picked for analysis. Scale bars = 50 nm.

Despite this severely limiting the number of particles distributed across the grid hole, collection of a small data set (table 4.3) resulted in the first 2D classes for CTS^{WT} (Fig. 4-13). From ~ 43k particles, these classes appear to represent multiple views of the protein, including top-down and sideways views. In the top-down views (Fig. 4-13b) an annulus of density can be seen surrounding a slightly less dense object in the centre; this may represent CTS^{WT} surrounded by the digitonin micelle as viewed from above.

CTS^{WT} (Quantifoil) Data collection and processing	
Magnification	130,000x
Microscope/Detector	Titan Krios/K2 (counting)
Voltage (kV)	300
Defocus range (μm)	-1.5 to -3.6
Total dose ($e^-/Å^2$)	82.1
Pixel size (Å)	1.14
Micrographs (No.)	1,120
Particles (No.)	43,571

Table 4.3. Cryo-EM data collection parameters for CTS^{WT} on Quantifoil grids.

In the sideways classes (Fig. 4-13a) there appears to be a larger bulbous region of density that is distal to a smaller region of density. In the sideways classes, this smaller region is observed as a single dense region (Fig. 4-13a, bottom panels), or is divided into two smaller densities (Fig. 4-13a, top panels); this small region most likely represents the NBDs of CTS^{WT}. Conversely, the bulb probably represents the digitonin micelle as this feature is observed in all structures of ABC transporters solved in detergent and is not indicative of the expected structure such proteins. A region of density that appears to represent the eGFP tag can be observed in two classes (Fig 4-13a, top panels) but is best observed in the class highlighted in green. The reason for the absence of this density in the other sideways views is unclear, though it may be explained by the flexibility between CTS^{WT} and the eGFP tag. When comparing the sideways 2D classes of CTS to the 2D classes of ABCD4 (Xu *et al.*,

2019), it was clear that there were strong similarities, despite the difference in particle number per each 2D classification (43k particles for CTS^{WT} vs. 1.5 million for ABCD4). Unfortunately, it was not possible to definitively observe any secondary structure in these classes, most likely due to the limited number of particles comprising each group of classes and because of sample heterogeneity. Consequently, attempts to generate an *ab initio* reconstruction from this data set were unsuccessful.

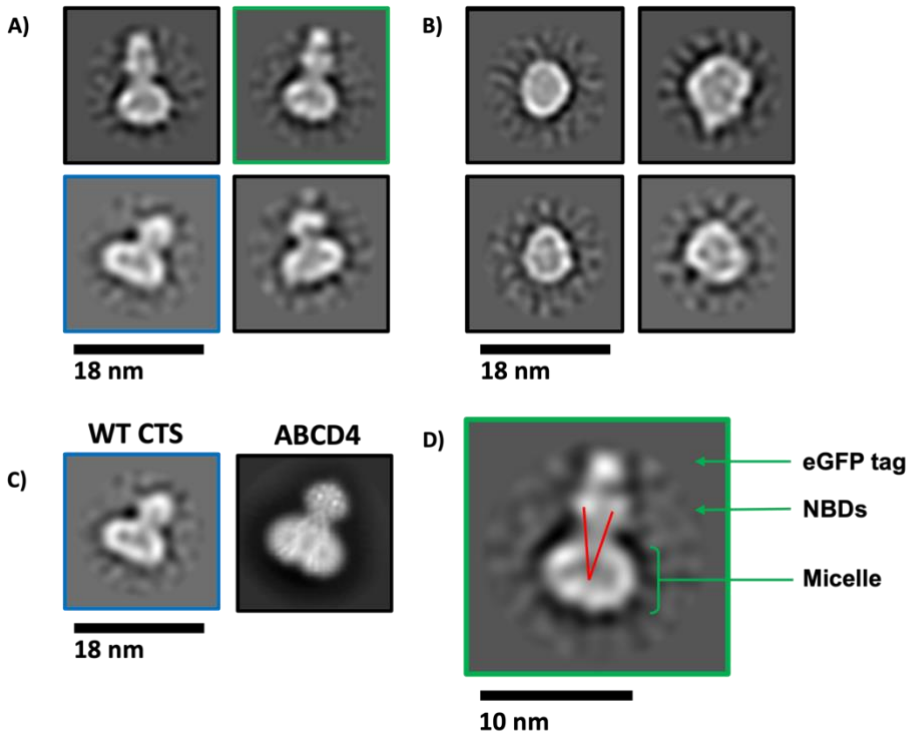


Figure 4-13. 2D Classification of CTS^{WT} from Quantifoil grids. ***A)*** *Four different 2D classes representing sideways views of CTS.* ***B)*** *Four different 2D classes possibly representing top-down views of CTS^{WT}.* ***C)*** *A comparison of a single CTS^{WT} 2D class from panel A (blue box, left) with 2D classes from the solved structure of ABCD4 (right), showing strong similarity.* ***D)*** *Expanded 2D class from panel A (green box) with annotations highlighting the expected features.* ***eGFP****, enhanced green fluorescence protein;* ***NBDs****, nucleotide binding domains; Red 'V' shape, putative location of TMDs.*

To prevent CTS sticking to the carbon support on Quantifoil grids (Fig. 4-12), purified CTS^{WT} was applied to Ultrafoil® grids (Fig. 4-14). These grids were developed to reduce sample movement during cryo-EM data collection (Passmore & Russo, 2016; Russo & Passmore, 2014), and have the added advantage of limiting particle sticking to carbon supports. Screening of these grids confirmed that particle distribution was modestly improved, and particle aggregation reduced compared to Quantifoils.

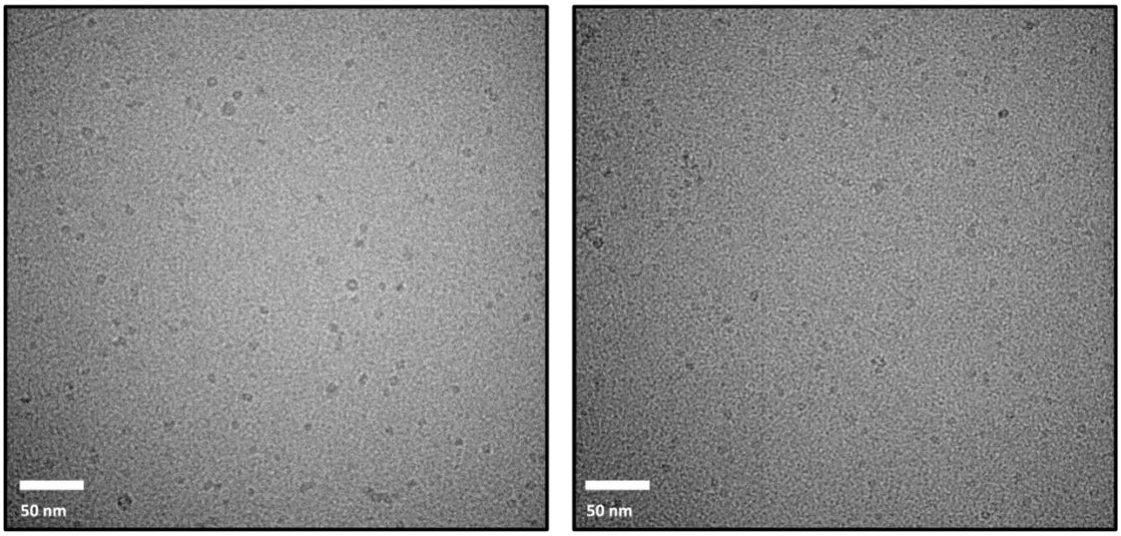


Figure 4-14. Cryo-EM screening of CTS^{WT} purified in digitonin on Ultrafoil® grids. Two micrographs obtained from the same Ultrafoil® grid showed improved particle distribution and reduced aggregation compared to Quantifoil grids. Images collected at 130,000x magnification. Scale bars = 50 nm.

Consequently, a small data set was collected (table 4.4) and the picked particles subject to 2D classification. 2D classification refers to the grouping particles into a series of classes, each of which represents the average of many hundreds, or thousands of similar particles picked from many micrographs. RELION, the most widely used programme for cryo-EM processing, does this using a maximum likelihood method (Scheres, 2012a, b). Beginning with a set of random reference images, each particle is compared to the reference by a series of translations and

rotations, after which probabilities are calculated to assess the degree of matching. By combining similar groups of particles and calculating their 2D Fourier Transform, an average 2D image of the particles is formed. These 2D images are then taken to be references, and the process is iterated over until all particles have been classified (Sigworth, 2016). The resulting series of classes is often used as a quality check to assess, for example, sample heterogeneity and particle orientation (side-ways, top-down, etc.) Moreover, it can be used to discard bad particles/contaminants or as a pre-processing step prior to 3D reconstruction.

Unfortunately, unlike the 2D classes obtained from particles picked off Quantifoil grids – where two different orientations were obtained (top down and sideways) – the 2D classes from particles picked off Ultrafoil® grids appeared, in the vast majority of cases, to show particles in one orientation (Fig. 4-15). Of all classes, only one appeared to represent a side-ways view of CTS^{WT}. In this class (Fig. 4-15, green box), a large density that may be the digitonin micelle is visible, with two faintly visible densities extending away from it (perhaps representing the NBDs). However, this is unconvincing and given the similarity of other class, it is most likely that they represent empty digitonin micelles. Furthermore, even if these classes contained CTS^{WT} particles, a 3D reconstruction would be impossible due to missing information from multiple angles of the protein (it would not satisfy the projection-slice theorem (Nogales & Scheres, 2015)). Therefore, *ab initio* reconstruction was not attempted on this data.

CTS^{WT} (Ultrafoil grids) Data collection and processing	
Magnification	130,000x
Microscope/Detector	Titan Krios/K2 (counting)
Voltage (kV)	300
Defocus range (μm)	-1.5 to -3.6
Total dose (e^{-}/Å^2)	72
Pixel size (Å)	1.14
Micrographs (No.)	2,646
Particles (No.)	89,342

Table 4.4. Cryo-EM data collection parameters for CTS^{WT} on Ultrafoil grids.

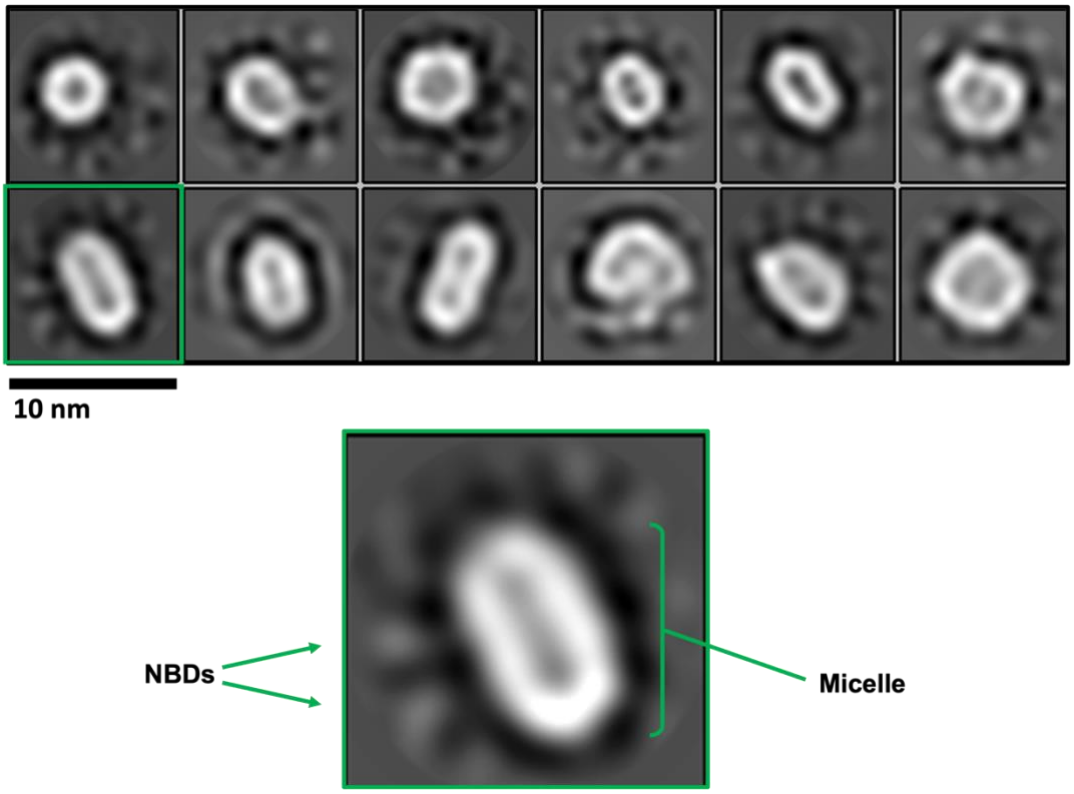


Figure 4-15. 2D Classification of CTS^{WT} from Quantifoil grids. Top panels, representative 2D classes of CTS^{WT} classes from Ultrafoil® grids appearing to show a predominance of empty digitonin micelles. Bottom panel, expanded view of the 2D class in the green box above, potentially showing a CTS in a digitonin micelle with NBDs pointing to the left.

To reduce sample heterogeneity and clean up the sample prior to grid preparation, a SEC step was utilised. Due to problems with digitonin precipitating the SEC buffer at both low temperatures and at RT, glyco-diosgenin (GDN) was used as an alternative. GDN is known as a synthetic substitute for digitonin that has a few key advantages over the latter. It is, cheaper, non-toxic, has a smaller CMC, and does not suffer from the batch-to-batch variability that digitonin does (digitonin is extracted from *Digitalis purpurea)*. Crucially, GDN is highly water-soluble (>10%) meaning it would be more easily applicable to SEC than digitonin (~5% soluble upon heating to 95°C). Unfortunately, and despite these advantages, efforts to perform SEC on CTS^{WT} in GDN were unsuccessful due to a total loss of protein through on-column precipitation.

Attempts were made to eliminate this problem by altering the concentration of GDN during different SEC experiments, however, the protein was lost each time and could not be purified using these experimental conditions.

4.3 Discussion

The work present in this chapter details the efforts made in purifying CTS^{WT} in an active form to determine its 3D structure using cryo-EM. Prior to these studies, CTS^{WT} had not been subject to any structural studies and therefore its purification needed to be optimised to generate protein of a suitable quality for grid preparation. Additionally, grid preparation itself also required significant optimisation such that a high-resolution model of CTS^{WT} could be determined. This non-trivial work, which spanned over several years, represented a serious challenge and exemplified all the difficulties of working with membrane proteins in a single project. Despite these challenges, this work resulted in an optimised solubilisation and purification protocol for CTS^{WT} that is suitable for structural studies by cryo-EM and showed that the transporter can be purified and reconstituted in an active form. Consequently, the work presented here was vitally important in paving the way for further studies on CTS^{WT} and provided a basis for studying the hydrolysis-deficient mutant, $CTS^{D606N/E607Q}$, which as can be seen in section 5, proved more fruitful.

4.3.1 ATPase Activity Assay

The ATPase activity of ABC transporters solubilised and purified in detergents varies widely, with some transporters showing high activity, and others showing low or total inactivity (Callaghan *et al*, 1997; Chavan *et al*, 2013; Infed *et al*, 2011; Timachi *et al.*, 2017). The reconstitution of CTS^{WT} into proteoliposomes for activity assays represented a significant milestone in the study of CTS because, for the first time, it was shown that the purified transporter possesses a similar level of substrate-stimulated ATPase activity to the transporter in Sf9 cell membranes (Carrier *et al.*, 2019). Moreover, the activity of reconstituted CTS^{WT} was comparable to the activity of its human homolog, ABCD1, when this transporter was reconstituted into liposomes (Kawaguchi *et al.*, 2021; Okamoto *et al.*, 2018).

The hydrolysis of ATP at the NBDs of ABC transporters is dependent on NBD dimerisation and the formation of composite nucleotide binding sites. The increased activity of CTS^{WT} observed in the presence of ATP therefore indicates that the NBDs are competent in binding and hydrolysing nucleotide even after the transporter has been solubilised, purified, and reconstituted. Conversely, the NBD1 Walker-B mutant $CTS^{D606N/E607Q}$, that lacks the catalytic glutamate necessary for ATP hydrolysis, expectedly showed no substantial increase in activity when in the presence of ATP. Indeed, this mutant is inactive *in vivo* where it prevents the germination of *A. thaliana* seedlings (Dietrich *et al.*, 2009). This is typical of such mutant used to study ABC transporters; the activities of CTS^{WT} and $CTS^{D606N/E607Q}$ reported here are similar to that report for proteoliposome-reconstituted ABCD4 and its hydrolysis deficient mutant, $ABCD4^{E549Q}$ (Xu *et al.*, 2019), respectively. Moreover, since NBD dimerisation/nucleotide hydrolysis in ABC transporters is linked to substrate transport, the increased activity of CTS^{WT} observed in the presence of increasing concentrations of C18:0 CoA indicated that after reconstitution the transporter was still capable of binding substrate (i.e., the bound substrate converts the transporter from an inward/open conformation with NBDs separated, to an outward/closed conformation, with NBDs in close proximity with a high affinity for ATP). A comparison between the ATPase activity of CTS, ABCD1, and ABCD4 is provided in table 4.5.

Transporter	Approximate ATPase Activity		
	Liposomes +ATP/+ATP +40 µM Substrate (nmol P_i/min/mg protein)	**Sf9 membranes +ATP/ +ATP +Substrate (nmol P_i/min/mg membranes)**	**Yeast Peroxisomes (strain BJ1991) +ATP/+ATP +10 µM Substrate (nmol P_i/min/mg protein)**
CTS^{WT} (this study)	18/30	-/-	-/-
$CTS^{D606N/E607Q}$ (this study)	7.5/-	-/-	-/-
CTS^{WT} (Carrier *et al.*, 2019)	-/-	24/28	-/-
CTS^{WT} (Nyathi *et al.*, 2010)	-/-	-/-	50/40
ABCD1 (Okamoto *et al.*, 2018)	7/-	-/-	-/-
ABCD1(Kawaguchi *et al.*, 2021)	11/-	-/-	-/-
$ABCD4^{WT}$ (Xu *et al.*, 2019)	13/-	-/-	-/-
$ABCD4^{E549Q}$ (Xu *et al.*, 2019)	6/-	-/-	-/-

Table 4.5. CTS ATPase Activity in Proteoliposomes compared to other studies. The ATPase activity reported for proteoliposome-reconstituted ABCD1 and ABCD4 compared to the ATPase activity reported for proteoliposome-reconstituted CTS^{WT} and $CTS^{D606N/E607Q}$, CTS^{WT} in Sf9 cell membranes, or yeast peroxisomes.

Interestingly, the activity of reconstituted CTS^{WT} was higher than expected after incubation with the transition-state inhibitor AlF_3, which arrests the transporter by forming a stable Mg-ADP-inhibitor complex. This compound has been shown to reduce the ATPase activity of CTS^{WT} in peroxisomes (Nyathi *et al.*, 2010) ~5 fold and

in Sf9 cell membranes ~7 fold (De Marcos Lousa *et al.*, 2013). In contrast, the activity of reconstituted CTS^{WT} was reduced only ~1.6 fold by AlF_3. This may be due to the yeast peroxisomes and Sf9 cell membranes containing a much lower concentration of CTS^{WT} than the reconstituted proteoliposomes, and, because they are not a source of purified protein, will likely contain other endogenous ATPases that may also be resistant to AlF_3 that are not present after reconstitution. However, the activity of ABCD1 in proteoliposome was reduced ~4 fold by AlF_3 (Kawaguchi *et al.*, 2021; Okamoto *et al.*, 2018). Nonetheless, there is still a clear difference between the activity of reconstituted CTS^{WT} in the presence and absence of AlF_3 and strong evidence for substrate stimulated ATPase activity.

This is important because it means that this reconstitution method can now be used to carry out transport assays on CTS^{WT} for further biochemical studies. For example, this system can be used to investigate the importance of specific CTS residues such as D863/Q864/T867, the triplet originally identified through homolog modelling as a potential catalytic triad (Carrier *et al.*, 2019), the intriguing S810N mutant that germinates *in vivo* but does not establish or break down fatty acids (Dietrich *et al.*, 2009) and retains ATPase activity but has reduced ACOT activity (Carrier *et al.*, 2019), and K487A which prevents germination and has severely reduced ATPase and ACOT activity (Carrier *et al.*, 2019; Dietrich *et al.*, 2009). Through reconstitution and transport assays, these mutants can be studied in an isolated system to pin-point their exact function with respect to substrate binding, cleavage, and transport.

Crucially, this reconstitution method may be used to carry out transport assays that are better placed to investigate the transporters thioesterase activity than the original ACOT activity assay (De Marcos Lousa *et al.*, 2013). This assay was carried out on CTS-containing Sf9 cell membranes where thioesterase activity was detected using Ellman's agent (DTNB). The reaction of DTNB with the free sulfhydryl groups of cleaved acyl-CoAs yields TNB^- that is subsequently ionised to the yellow compound TNB^{2-} which is detected spectrophotometrically. This assay however is not sensitive

enough to detect a single turn-over event. Alternatively, Kawaguchi *et al.* (2021) have investigated the ACOT activity of ABCD1 by using a highly sensitive fluorescence-based transport assay that works by tracking location of the fluorescent NDB-C16 moiety of NBD-C16-CoA after incubation with proteoliposome-reconstituted ABCD1 (i.e., is NBD-C16-CoA cleaved by ABCD1, and if so, does the NBD-C16 reside outside or inside the liposome). Using this assay, the group were able to show that the fatty acid moiety of NBD-C16-CoA does end up inside the liposomes and gets there by ABCD1 thioesterase activity. Now that CTS^{WT} can be reconstituted into a liposome, this assay may be utilised to further characterise the ACOT activity of CTS^{WT}. However, such transport assays for acyl-CoA substrates may be troubled by a number of issues. For example, the hydrophobicity of such compounds can lead to their passive permeability into the liposome or absorption/sequestration in the liposome membrane. In both cases, this leads to difficulties in ascertaining the true rate of transport since the final location of substrate cannot be easily identified. Overcoming these issues may require modification of substrates and/or alteration of the liposome lipid composition. Another issue is that of low CTS activity; to ensure accurate and precise quantification of transport kinetics, it would be necessary to improve the activity of CTS, which may be done by pre-loading the liposomes with CTS's acyl-CoA synthase partners, or by identifying acyl-CoA from plants that stimulate the maximum activity possible.

However, it may be possible to carry out a fluorescence-anisotropy-based transport assay, such as that described by Stefan *et al.* (2020). Modification of these methods can therefore inform events such as the order of the catalytic cycle, how substrate binding affects conformational changes, and/or the nature of NBD asymmetry in CTS and futile ATPase activity. It should be noted that the orientation of CTS^{WT} in the proteoliposome membrane after reconstitution was not determined, and therefore the true rate of ATP hydrolysis is likely higher than reported (i.e., if the NBDs face the inside the proteoliposome, they cannot bind nucleotide and substrate cannot enter

the transporter). Therefore, it would be beneficially to carry out partial proteolysis to determine the topology of CTS^{WT} after reconstitution and prior to transport assays to measure transport kinetics more accurately.

4.3.2 Purification and EM on CTS^{WT}

Overall, the purification of CTS^{WT} was approached using three different strategies: SMA co-polymer, nanodiscs, and detergents. The decision to use SMA was made after considering two of the polymer's key characteristics. Firstly, SMA is an effective chelator of divalent cation such as magnesium and calcium, where concentrations of these cations above 5 mM precipitate the SMALP (Lee *et al.*, 2016b). This poses a problem for ABC transporters which require Mg^{2+} for ATP hydrolysis and therefore may be inactive when solubilised in SMA. However, SMA is capable of extracting membrane proteins from the directly from the membrane along with their native lipids (Qiu *et al*, 2018) and secondly, and to maintain membrane protein solubility in the absence of detergent. This would be helpful for studies by cryo-EM because there would be no detergent in solution to complicate downstream purification steps, and no micelles to complicate particle picking during cryo-EM studies (Rubinstein, 2007). For this reason, SMA was investigated for CTS^{WT}.

Unfortunately, whilst CTS^{WT} could be solubilised using SMA 2000, the solubilised transporter could not bind to any resins to be purified. This intriguing observation is counter to many reports demonstrating the affinity purification of other SMA-solubilised membrane proteins (Sun *et al*, 2018; Unger *et al*, 2021), including various ABC transporters including P-gp, ABCG2, MRP1, and CFTR (Gulati *et al*, 2014). It has been shown that SMA will precipitate out of solution in the presence of divalent cations or at a pH < 7, and that the highly negative charge of the SMALP can lead to non-specific binding to affinity resins, unless ≥ 500 mM NaCl is included in the binding buffer (Hardy *et al*, 2019; Lee *et al.*, 2016b). Despite taking all these factors into account during attempted purification, the binding of SMA-solubilised CTS^{WT} to resin

was still poor. Therefore, given that all the constructs studied by Gulati *et al.* (2014) used his-tags (and one his/GFP-tag construct) it was considered whether the poor resin-binding of CTS^{WT} was due to the 2xStrepII/eGFP-tag. It was hypothesised that the 2xStrepII/GFP-tag was wrapped up in the SMALP, thereby being excluded from solution and leaving it inaccessible to the resins or that the tag may have ended up between the NBDs of CTS^{WT} thereby being blocked from interacting with the resin. Interestingly, Bersch *et al* (2017) reported a similar streptactin resin-binding problem with their SMA-solubilised strep-tagged construct and that this was improved by removing free SMA after solubilisation. However, attempts to carry this out for CTS^{WT} were hampered by a newly prepared batch of SMA that abrogated solubilisation of CTS^{WT} entirely. This implied that the quality SMA co-polymer prepared from the styrene maleic anhydride precursor can vary significantly, and this batch-to-batch variability that can hinder downstream applications. However, the difference between different SMA co-polymer preparations was not investigated in this study.

As with SMA, the overarching rationale for reconstitution of CTS^{WT} into lipid nanodiscs was the absence of detergent. However, nanodiscs represented a more attractive approach since unlike SMA, they are not sensitive to divalent cations, and because many different ABC transporter had already been studied after nanodisc reconstitution (Hofmann *et al.*, 2019; Jackson *et al.*, 2018; Josts *et al.*, 2018; Mi *et al.*, 2017). A detailed comparison between structures solved in detergent vs. nanodiscs is provided by Wright *et al.* (2018). Moreover, the ATPase activity of ABC transporters reconstituted into nanodiscs is often greater than that of ABC transporters solubilised in detergent (Nandigama *et al*, 2019), indicting a role for the surrounding lipids in modulating transporter activity (Kawai *et al*, 2011), which may be illuminated through structural studies. Optimisation of protein:scaffold:lipid ratios for reconstitution of CTS^{WT} only yielded one condition where a substantial amount of nanodiscs formed, though it is unclear why the expected particles could not be observed by negative stain EM despite the sample being relatively clean and subject to ultracentrifugation.

One parameter was that not varied was the lipid used for reconstitution. POPC was chosen because it has been used extensively for nanodisc reconstitution of several types of membrane proteins (Denisov *et al*, 2019; Her *et al*, 2016). However, in a recently published paper, the structure of the CTS homolog, HsABCD1, was solved in a nanodisc containing a mixture of brain-polar-lipids and cholesterol at a 1:10:350 ratio of HsABCD1:scaffold:BLP/cholesterol (Le *et al.*, 2021). This mixture has also been used for ABCG2 (Manolaridis *et al.*, 2018) and Pg-p (Nosol *et al.*, 2020a). This ratio used for HsABCD1 is significantly higher than those tested for CTS^{WT} and therefore it would be interesting to investigate these conditions (ratio and lipid composition) for CTS^{WT}.

More success was had from work on CTS^{WT} using different detergents. Using both DDM and LMNG/digitonin, it was possible to appropriately solubilise and purify CTS^{WT} to a level suitable for EM studies although the final concentration of full-length CTS^{WT} after purification was always a hindrance. Through this study it was noted that CTS^{WT} is a particularly fragile and labile protein that degrades easily. If not supplemented with many protease inhibitors, CTS^{WT} will degrade during the membrane preparation stage, and after this, will degrade over time even when on ice and substantially during concentration. One notable degradation product of CTS^{WT} is one that on SDS-PAGE runs ~half the distance of full-length CTS^{WT}. This product was first identified in a peroxisomal fraction that was isolated from yeast cells that had been transformed with a plasmid bearing the *cts* gene (Nyathi *et al.*, 2010). Originally this product was suggested to result from cleavage somewhere in the predicted flexible linker between the two halves of CTS^{WT} by an endogenous yeast protease. However, in the work presented here, the product has been observed during expression of CTS^{WT} in Sf9 insect cells, as well as its purification in both DDM and LMNG/digitonin, even after separation from endogenous proteases. This is exemplified by the observation of the product even after sample clean up by size-exclusion chromatography when assessed by coomassie staining, and by specific in-gel GFP fluorescence. Thus, it

can be supposed that the proteolysis of the linker region contributes to the intrinsic fragility of CTS^{WT}. However, whilst this cleavage is visible by SDS-PAGE, studies by Nyathi *et al.* (2012) showed that co-expression of the two pseudo-halves of CTS results in a function transporter possessing ATPase activity, implying that even if the linker is cleavage, the two halves of CTS would still be associated. This type of observation has been reported widely for many membrane proteins e.g., P-gp (Nuti *et al*, 2000).

When purified in DDM, CTS^{WT} displayed significant heterogeneity on negative stain grids. Whilst DDM has been used to solve the structure of several membrane proteins using cryo-EM, including Complex I (Parey *et al*, 2018), the human epithelial sodium channel (Noreng *et al*, 2018), the LH1-RC complex (Tani *et al*, 2020), and the V-ATPase (Tsunoda *et al*, 2018), the small size of CTS^{WT} compared to many of these proteins made distinguishing CTS^{WT} from empty micelles difficult, due to DDM micelles being ~70-92 kDa in size and CTS^{WT} being ~182 kDa. (Chaptal *et al*, 2017). Moreover, DDM drastically reduces contrast in EM micrographs; this problem is highlighted by Tsunoda *et al.* (2018), who resorted to the use of a phase plate to determine the structure of *Eh*V-ATPase. The difficulty in distinguishing like-sized particles in negative stain micrographs is highlighted nicely by Gewering *et al* (2018). Purification in LMNG/digitonin resulted in the most success, with CTS^{WT} 2D classes being generated from cryo-EM data sets. When purified this way, much less degradation was observed, perhaps because the presence of ATP in the solubilisation buffer kept the transporter in a more compact outward-facing conformation and less liable to proteolysis (Geillon *et al.*, 2017), thereby reducing sample heterogeneity.

Optimisation of CTS^{WT} on different grids led to the collection of two data sets. The use of Quantifoil grids resulted in 2D classes for CTS^{WT} that when compared to the published 2D classes for hydrolysis-deficient ABCD4 (Xu *et al.*, 2019), clearly showed distinct side-ways views of CTS^{WT}. Given that this is wild-type protein one may expect the transporter to adopt various conformation states. However, from the 2D classes, it appears that CTS^{WT} was in the outward-facing state since no classes were observed with distinctly separate NBDs (this difference between NBDs close and NBDs apart can indeed be observed after 2D classification). The reason for this may be because CTS^{WT} was solubilised in the presence of ATP; since many ABC transporters are inactive in detergent, it may be that once ATP bound to CTS^{WT} it was not hydrolysed and was therefore carried forward through the purification. Unpublished data from the Baker group has shown that CTS^{WT} possess very low ATPase activity when purified in DDM; this may also be true for CTS^{WT} in digitonin though has not been confirmed experimentally. Despite there being several 2D class averages representing side-ways views of CTS^{WT} after subset selection, ultimately a 3D model could not be reconstructed. This was most likely due to the limited number of a particles extracted from the micrographs, i.e., 43k particles in total translating into an average of ~38 particles per micrographs, with subsequent 2D classification and subset selection reducing the total number of useable particles even further.

Comparatively, the data set that was used to solve the structure of ABCD4 contained ~550 particles per micrograph and a total of ~1.5 million particles. The sheer lack of particles in the CTS^{WT} data set probably resulted from two major factors: the first was the low concentration of CTS^{WT} applied to the grid, and the second was the sticking of CTS^{WT} particles to the carbon support of the grid. Since CTS^{WT} rarely concentrates above ~0.8 mg/ml, attempts were made to limit the sticking of CTS^{WT} to the carbon support using Ultrafoil grid. Whilst this was slightly improved, the absence of suitable sideways 2D classes in this data set, perhaps because of preferred orientation on this grid type or interaction of CTS^{WT} at the air-water-interface (and therefore damaged

particles), meant that *ab initio* reconstruction would be unsuccessful despite this data set containing twice the number of particles compared to the Quantifoil data set. At the same time, at this point in the project it was not possible to SEC purify CTS^{WT} in digitonin due to precipitation problems, and therefore the lack of a SEC purification step certainly contributed to sample heterogeneity, and abrogated 3D reconstruction. Multiple attempts to circumvent this problem by running SEC experiments using various concentrations of the digitonin substitute, GDN, were all unsuccessful due to loss of protein on the column. Initially, the reason for this loss was thought to be caused by using a GDN concentration too close to CMC, however, increasing this to 5x CMC did not help, despite this concentration being used successfully for other transporters (Diederichs *et al*, 2020). Recently, the cryo-EM structure of SLC26A9 was solved using a GDN concentration 10x CMC (Walter *et al*, 2019). Therefore, it would be worth investigating the impact of higher GDN concentrations on $CTS^{WT.}$

In summary, the work presented here laid all the groundwork for structural studies, representing several years of challenging work to establish a reproducible and reliable solubilisation, purification, and reconstitution method for CTS^{WT} (Fig. 4-16). It has been demonstrated that CTS^{WT} is amenable to cryo-EM studies and whilst it was not possible to generate a 3D model of the wild-type protein, the 2D classes generated – that are similar to those published for human homolog ABCD4 – provided a promising outlook for further structural studies. Crucially, in this work a protocol for the reconstitution of CTS into proteoliposomes was established showing that the transporter possesses substrate stimulated ATPase activity. This was important as it opens the possibility of carrying out novel biochemical studies on CTS, such as transport assays, in a totally isolated system. Finally, by considering and evaluating the results presented here, further studies were pursued focusing on $CTS^{D606N/E607Q}$, the catalytic mutant of CTS^{WT}. By building upon the work presented here, and carrying out further cryo-EM studies, the first 3D model of $CTS^{D606N/E607Q}$ was generated and is presented in section 5.

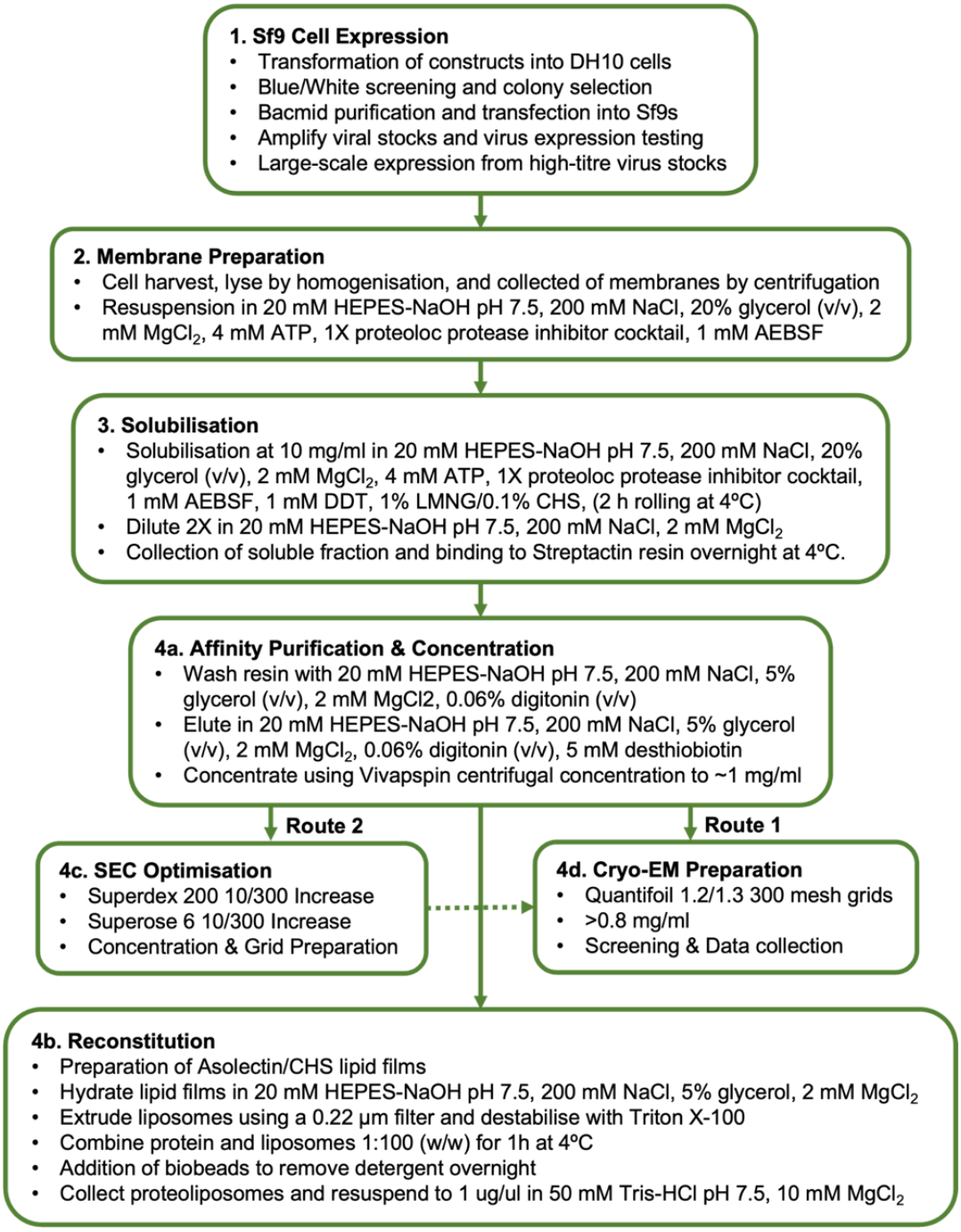


Figure 4-16. A workflow of the expression, solubilisation, purification methods used to study CTS using cryo-EM, and the reconstituted method used to measure the transporters ATPase activity. Route 1 represents grid preparation after affinity purification. Route 2 represents grid preparation after affinity purification and sample clean-up via SEC. Route 2 is to be optimised in future studies.

5 Studies Towards the Structure of CTS$^{D606N/E607Q}$

5.1 Introduction

To move substrates across the membrane ABC transporters cycle through several different conformational states as they bind substrate and hydrolyse ATP. It is necessary to obtain structural information for each of these states build up a picture of the entire catalytic pathway, and although this is not-trivial, multiple conformational states of a single transporter can be obtained from a single cryo-EM data set (Hofmann *et al.*, 2019; Mi *et al.*, 2017). However, the physical flexibility of ABC transporters can make structural studies difficult for several reasons: flexibility often reduces the stability of the protein leading to significant degradation, it can limit the final resolution of any structure obtained due to heterogeneity in the data set, and because one may wish to study a specific conformation which is rarely observed in the wild-type data set. Therefore, many structural studies on ABC transporters are carried out on inactive mutants, created by mutating the catalytic glutamate residue of the Walker B motif to glutamine (Caffalette & Zimmer, 2021; Jackson *et al.*, 2018; Liu *et al*, 2021; Manolaridis *et al.*, 2018). Though these mutant transporters are still capable of binding nucleotide, this mutation abolishes ATP hydrolysis, often locking the protein in a closed (outward facing) conformation (Lusvarghi *et al*, 2021; Manolaridis *et al.*, 2018).

For these reasons and given that studies on CTS^{WT} had not yet yielded a 3D model, structural studies by cryo-EM were pursued using the Walker B double mutant of CTS, CTS$^{D606N/E607Q}$. By drawing on all the findings presented in previous result chapters (section 3 and 4) and applying them to CTS$^{D606N/E607Q}$ the first 3D maps of CTS were generated using cryo-EM. Despite these maps being low resolution, the results presented in this chapter demonstrate that these can be obtained even from limited data sets, and by outlining the experiments carried out and those required to move from low to high resolution therefore paves the way for further experimental

optimisation to generate the first high-resolution structure of the plant peroxisomal ABC transporter.

5.2 Results

5.2.1 Purification of CTS$^{D606N/E607Q}$ in LMNG/digitonin

As with CTS^{WT}, CTS$^{D606N/E607Q}$ was solubilised in LMNG and exchanged into digitonin on the affinity column. After elution from the affinity resin, CTS$^{D606N/E607Q}$ appeared to behave in a similar way to CTS^{WT}, although it was clear that CTS$^{D606N/E607Q}$ degraded less severely during the concentration step (Fig. 5-1). Densitometry determined that after concentration, the ratio between CTS^{WT} and the bifurcated cleavage product (discussed in chapter 4.3.2) was 2.45. In comparison, the ratio between CTS$^{D606N/E607Q}$ and the bifurcated cleavage product was 5.44. This shows that the mutant protein is retained at a higher level than CTS^{WT} during concentration since there is less degradation into the cleavage product, implying that CTS$^{D606N/E607Q}$ is more stable than CTS^{WT}.

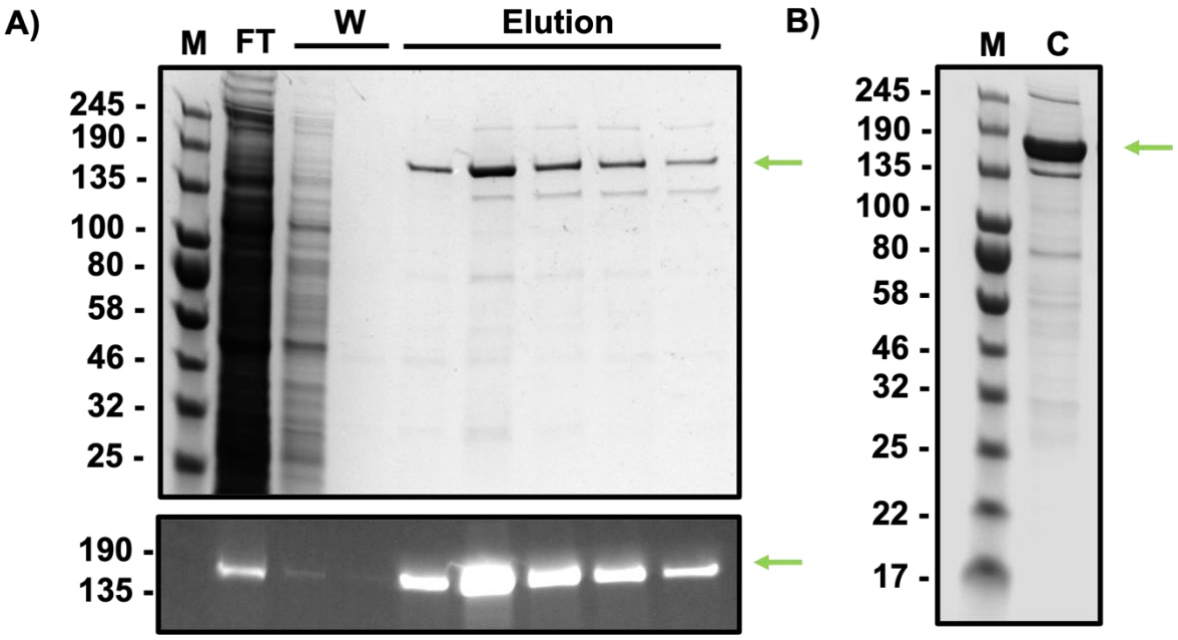


Figure 5-1. A typical purification of CTS$^{D606N/E607Q}$ in LMNG/Digitonin. ***A)*** *Top panel, SDS-PAGE of Streptactin affinity purification of CTS$^{D606N/E607Q}$ eluted in digitonin. Bottom panel, corresponding in-gel GFP image.* ***B)*** *Concentrated elution fractions, showing CTS$^{D606N/E607Q}$ as the dominant species with a small amount of presumed degradation. This figure is representative of multiple purifications of CTS$^{D606N/E607Q}$ (N=>5).*

When using this purification protocol, it was desirable to use a SEC clean-up step to separate CTS$^{D606N/E607Q}$ away from degradation products and aggregates to improve sample heterogeneity when preparing cryo-EM grids. However, it was noted that upon preparing large volumes of SEC buffer, the digitonin would precipitate in solution (oddly, this did not occur when preparing small volumes <1 ml). This meant that initially, it was not possible to clean up the CTS$^{D606N/E607Q}$ sample using SEC. Because CTS precipitates if frozen and stored at -80°C, and because the sample was relatively clean after concentration, it was decided that despite not being SEC purified cryo-EM grids would be prepared, and efforts would be made to further purify '*in silico*', i.e., to discard poor particles/aggregates and separate empty micelles from CTS$^{D606N/E607Q}$ particles.

5.2.2 Initial Attempts at Cryo-EM on CTS$^{D606N/E607Q}$

Initial screening of CTS$^{D606N/E607Q}$ on Quantifoil grids revealed a reasonable particle distribution (data not shown) that prompted the collection of a small data set (table 5.1). From this data set particles were picked using crYOLO and then subject to 2D classification in RELION 3.1.1, which showed several classes that appeared to represent CTS$^{D606N/E607Q}$ (Fig. 5-2).

CTS$^{D606N/E607Q}$ Data collection and processing - 1	
Magnification	130,000x
Microscope/Detector	Titan Krios/K2 (counting)
Voltage (kV)	300
Defocus range (μm)	-1.5 to -3.6
Total dose (e^-/Å^2)	82.1
Pixel size (Å)	1.07
Micrographs (No.)	1,910
Particles (No.)	75,177

Table 5.1. Cryo-EM data collection parameters for CTS$^{D606N/E607Q}$ on Quantifoil grids.

Like CTS^{WT}, the classes from CTS$^{D606N/E607Q}$ showed a larger bulbous region, probably representing the digitonin micelle, distal to a smaller density probably representing the NBDs (Fig. 5-2a). Each region – the NBDs and the micelle surrounding the TMDs – provides a similar level of contrast due to their being comprised of a similar atomic composition compared to the supporting vitreous ice. From this 2D classification, a subset selection was carried out to choose distinctly sideways classes, and to discard bad particles, aggregates, or suspected empty micelles. This was done manually after developing an intuition for good and bad particles over time Although this is relatively subjective, it provides a quick means of assessing multiple different 2D class subsets.

Subsequent 2D classification of these sub-selected particle stacks (Fig. 5-2b) revealed secondary structure of the TMDs in at least two 2D class averages, with either one strong region of density or two separate – but equally strong – densities representing the NBDs. However, like CTS^{WT}, it was not possible to generate an *ab initio* 3D model of $CTS^{D606N/E607Q}$ from these particle stacks as the number of particles in these most detailed 2D classes after subset selection was only ~6500. Despite this, it was evident that the 2D classes from $CTS^{D606N/E607Q}$ were clearer and more detailed than those obtained for CTS^{WT}. Thus, it was reasoned that two major factors were preventing the formation of at least a low-resolution 3D model of $CTS^{D606N/E607Q}$: sample heterogeneity and the size of the data set. Regarding the former, efforts were focused on sample clean-up via SEC.

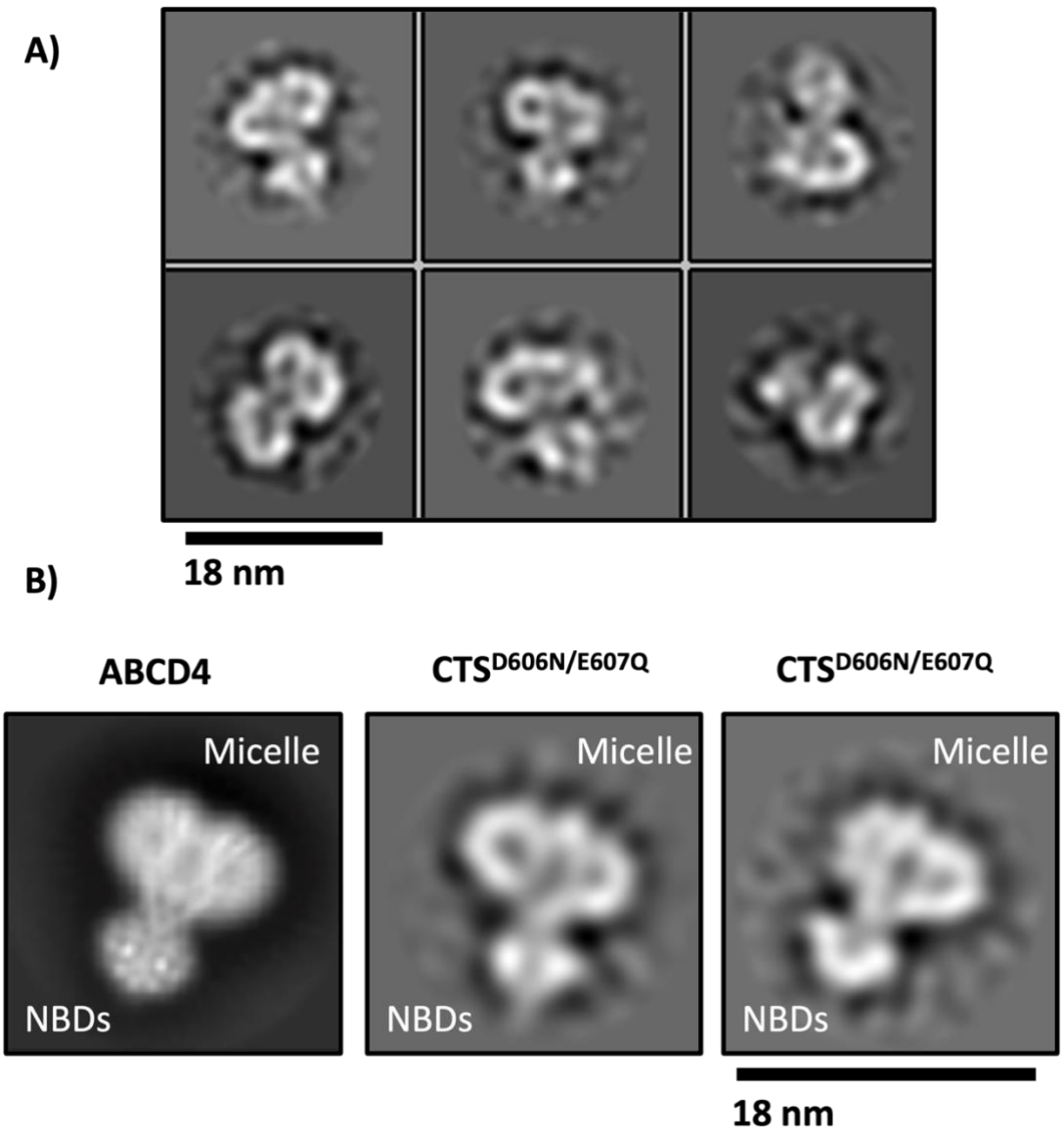


Figure 5-2. 2D Classification of CTS$^{D606N/E607Q}$ purified in digitonin. ***A)*** *Representative 2D class averages of CTS$^{D606N/E607Q}$ obtained from 2D classification of ~75k particles, with some potential sideways views evident.* ***B)*** *2D classification of particle stacks obtained after subset selection of classes averages from the 2D classification in panel A, showing more well-defined views of CTS$^{D606N/E607Q}$* (middle and right) and a comparison to the 2D class from the published structure of ABCD4 (left).

5.2.3 $CTS^{D606N/E607Q}$ Size-exclusion Chromatography

Attempts were made to ensure the continued solubility of digitonin in the desired SEC buffer by heating to ~95°C, cooling of the buffer to 4°C, and filtering off any precipitate prior to the SEC run as recommended in the product description of digitonin provided by Calbiochem (SKU 300410-250). Whilst this worked for a short period of time – i.e., < 30 minutes – the digitonin precipitated solution after this and did not remain in solution for long enough to carry out even the shortest of SEC experiments e.g., using a Superose 6 5/150 Increase column with a 3 ml bed volume, and even precipitated a second time after the first filtration. Carrying out the SEC run at room-temperature was also considered, but the digitonin precipitated once again even at ~18-20°C. These problems were puzzling because many cryo-EM structures of membrane proteins had been solved in digitonin including the ABC transporters, ABCD4 (Xu *et al.*, 2019), MRP1 (Johnson & Chen, 2017), and ABCA1 (Qian *et al.*, 2017). Finally, I was kindly provided with a suitable protocol for SEC using digitonin by Dr Valeria Kalienkova from the Paulino group at the University of Groningen which had previously been used to determine the structure of TMEM16A (Paulino *et al*, 2017) and TMEM16F (Alvadia *et al*, 2019). This protocol called for the incubation of any SEC buffer containing digitonin at 4°C for at least five days before filtration. After doing so, it was possible to keep the digitonin in solution for ≥ 24 hrs and therefore the SEC clean up step was possible. After affinity purification and concentration to ~0.5 mg/ml, 500 µl of $CTS^{D606N/E607Q}$ was injected onto a Superdex 200 10/300 Increase column equilibrated in 20 mM HEPES, 150 mM NaCl, 4% glycerol, 1.5 mM, 0.05% (w/v) digitonin (Fig. 53).

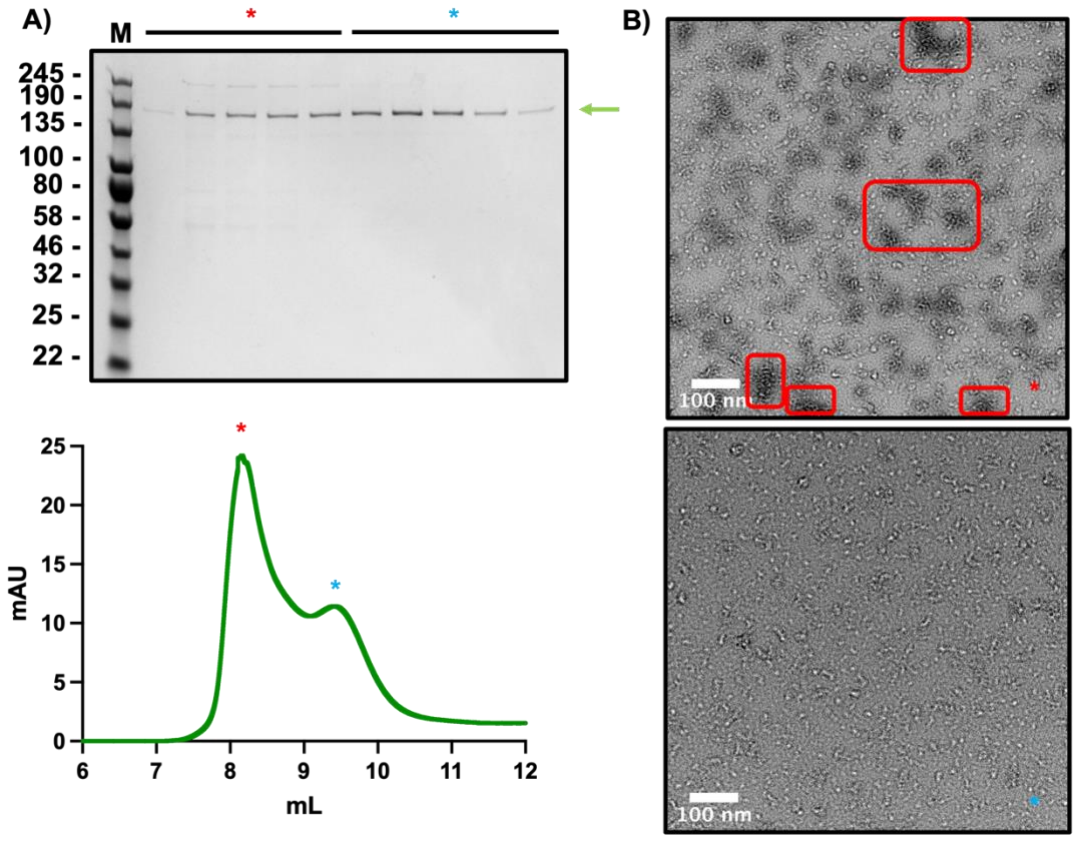


Figure 5-3. SEC purification of CTS$^{D606N/E607Q}$ in digitonin using a Superdex 200 10/300 Increase column. ***A)*** *Top panel, Coomassie-stained SEC fractions of CTS$^{D606N/E607Q}$showing high-purity. Bottom panel, SEC chromatogram of CTS$^{D606N/E607Q}$ showing two peaks but with poor separation (red and blue asterisks).* ***B)*** *Negative stained micrographs from fractions C2 (top panel, red asterisk in A), and C5 (bottom panel, blue asterisk in A). Red boxes highlight regions of aggregation from the void peak that are noticeably absent from the shoulder peak.*

Two peaks were observed, one at ~ 8 ml (the void volume) and another – a shoulder peak – at around 9.5 ml (Fig. 5-3a). Samples were collected from fractions under both peaks and assessed by SDS-PAGE, with CTS$^{D606N/E607Q}$ being observed in all fractions. Fractions from the two peaks were then assessed by negative stain EM, which revealed that the particles under each peak looked distinctly different (Fig. 5-3b, top panel). In the void peak, the majority of particles were expectedly aggregated, with clumps of particles surrounded by deep staining of uranyl acetate being observed

across the micrographs. Conversely, particles from the shoulder peak showed no aggregation and instead, particles were distributed evenly across the micrographs, with strong homogeneity (Fig 5-3b, bottom panel). This result was important as it demonstrated that $CTS^{D606N/E607Q}$ could now be SEC purified in digitonin, which would significantly improve sample heterogeneity when preparing cryo-EM grids, and potentially allow for a 3D model to be generated. However, despite the purity of $CTS^{D606N/E607Q}$ after SEC (e.g., in the shoulder peak as observed by SDS-PAGE; Fig 5-1, top panel), the poor separation between the two peaks during the SEC experiment meant that further optimisation would be necessary to ensure the sample would be optimal for cryo-EM grid preparation. To better separate these two peaks, the experiment was run again, but using a different SEC column – a Superose 6 10/300 Increase – with a wider separation range (Fig. 5-4).

Interestingly, instead of separating these two peaks, a completely different chromatogram trace was observed, characterised by a broad hump across the majority of the useable column volume (i.e., ignoring the void). Fractions from this hump were assessed by SDS-PAGE, showing that $CTS^{D606N/E607Q}$ had eluted in fractions under this entire volume (Fig. 5-4a, top panel). As before, peak fractions were assessed by negative stain EM, where consecutive elution fractions looked similar in terms of particle number, particle distribution, and homogeneity, though with markedly different staining depths (Fig. 5-4b). When compared to the negative stain micrographs obtained from the previous SEC experiment using the Superdex 200 10/300 Increase column (Fig. 5-3b, bottom panel), there was clearly a substantial difference in particle size and shape. From these two experiments, it was clear that despite the broad peak observed after purification over the Superose 6 column, that these particles looked better than those from the Superdex 200 purification. Potential reasons for this difference are discussed in section 5.2.

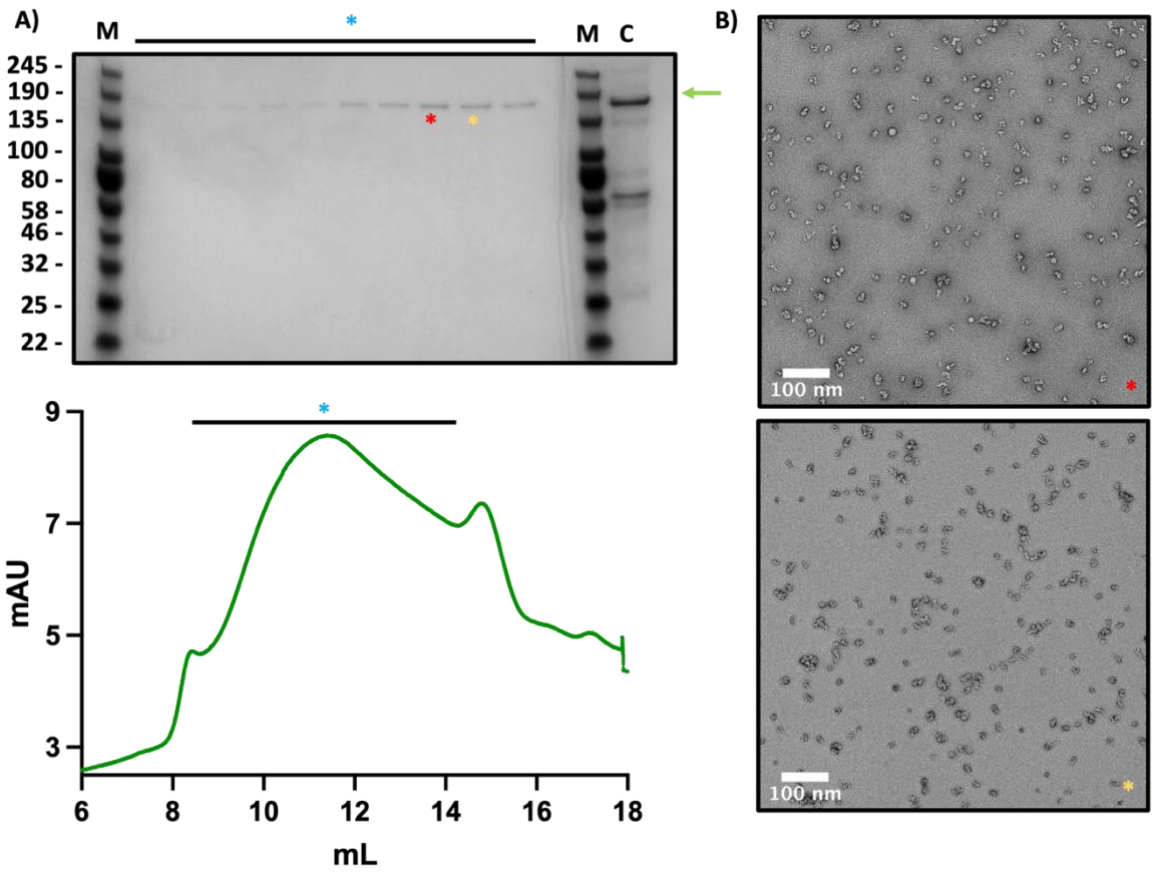


Figure 5-4. SEC purification of CTS$^{D606N/E607Q}$ in digitonin using a Superose 6 10/300 Increase column. ***A)*** *Top panel, Coomassie-stained SEC fractions of CTS$^{D606N/E607Q}$. Bottom panel, SEC chromatogram of CTS$^{D606N/E607Q}$ characterised by a broad hump-peak (blue asterisk).* ***B)*** *Negative stain micrographs on peak fractions from the hump-peak (top panel, red asterisk in A; bottom panel, yellow asterisk in A).*

The second factor limiting the generation of a 3D model of $CTS^{D606N/E607Q}$ was the limited size of the data set obtained during a cryo-EM experiment. Whilst sample clean-up is undoubtedly the most important step in the generation of a high-resolution cryo-EM map, the amount of information available for 3D reconstruction is also important (i.e., the absolute number of useable particles that the software can access when attempting to build a map, either *ab initio* or from a 3D classification). Therefore, given that reasonable 2D class could be obtained from non-SEC purified $CTS^{D606N/E607Q}$ and due to time-constraints that limited the optimisation of the SEC step, it was decided that a much larger data set would be collected to increase the number of particles available for 3D reconstruction to generate a model of $CTS^{D606N/E607Q}$.

5.2.4 $CTS^{D606N/E607Q}$ Large Scale Data Collection

Before a large-scale data set was collected, several grids were screened to identify the most promising grid type. $CTS^{D606N/E607Q}$ was affinity purified as before and applied to both Quantifoil 1.2/1.3 300 mesh grids and Ultrafoil grids for screening prior to data collection (Fig. 5-5). Ultrafoil grids were used in an attempt to improve particle distribution during the large-scale data collection to enhance the number of useable particles. From this screening, particles could only be observed on Quantifoil grids, with the vitreous ice on the Ultrafoil grids being too thick to image through (images not collected for this reason). On Quantifoil grids, the expected particles could be observed based upon their size and shape compared to other ABC transporters in digitonin (Johnson & Chen, 2017; Xu *et al.*, 2019), with roughly spherical particles most likely representing the relatively small CTS transporter in a digitonin micelle. Although these particles were at a lower concentration than desired – and despite some aggregates and contamination – the grid was deemed suitable for an initial small data collection. A very small data set was collected – amounting to ~1 hr of automated micrograph collection – and processed in crYOLO and RELION 3.1.1 to

see if the expected 2D classes could be obtained. After this was confirmed (data not shown), a much larger data set was collected on this same grid for ~60 hrs (table 5.2).

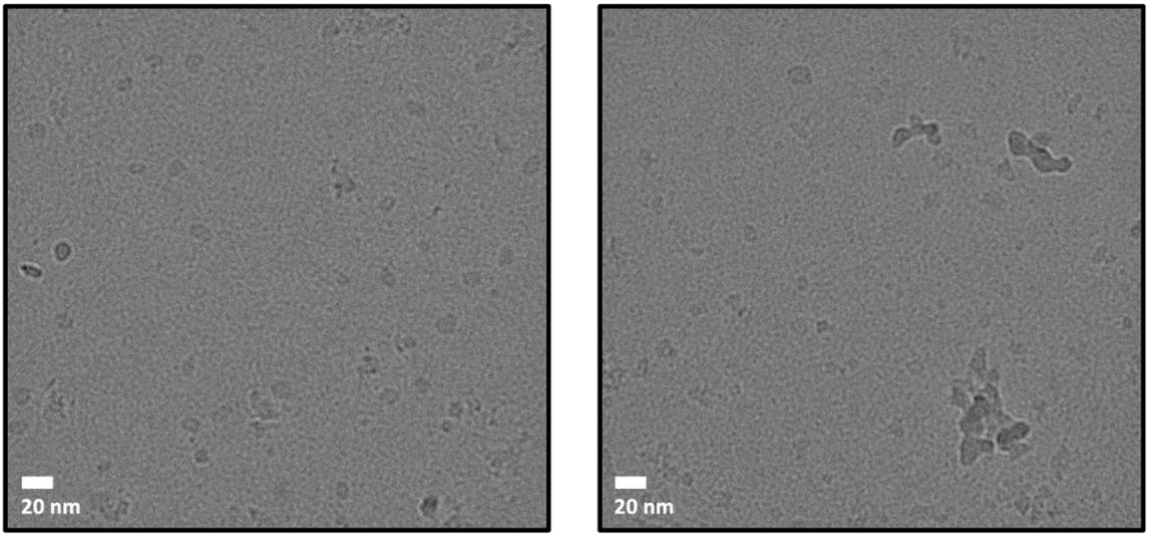


Figure 5-5. Representative micrographs from the second cryo-EM grid screening session on affinity-purified CTS$^{D606N/E607Q}$. CTS$^{D606N/E607Q}$ was again screened on Quantifoil 1.2/1.3 grids, showing a similar particle distribution to the first screening session. This grid was ultimately subject to the largest data collection for CTS$^{D606N/E607Q}$. Scale bar = 20 nm.

$CTS^{D606N/E607Q}$ Data collection and processing - 2	
Magnification	96,000x
Microscope/Detector	Titan Krios/Falcon 4 (counting)
Voltage (kV)	300
Defocus range (μm)	-1.5 to -3.6
Total dose ($e^-/Å^2$)	42.7
Pixel size (Å)	0.85
Micrographs (No.)	12,084
Particles (No. picked from crYOLO, threshold 0.3)	106,606
Particles (No. picked from crYOLO, threshold 0.1)	271,234

Table 5.2. Cryo-EM data collection parameters for the largest data set collected for $CTS^{D606N/E607Q}$. From this data, two sets of particles were picked using the crYOLO software, 106,606 particles from a picking threshold of 0.3 (default), and 271,234 particles from a less stringent picking threshold of 0.1.

To process this data, two different sets of particles were picked and processed independently. The micrographs were subject to particle picking using crYOLO's general model, either with a default picking threshold of 0.3 or the less stringent threshold of 0.1. Until this point, all other data sets (CTS^{WT} and $CTS^{D606N/E607}$) had been subject to picking using the 0.1 threshold. However, it was questioned whether a more stringent picking process would yield a cleaner particle stack, and therefore better, more defined 2D classes, and perhaps a 3D model. Subsequently, both set of particles (here on referred to as Pick0.3 or Pick0.1) were extracted and subject to 2D classification in RELION 3.1.1 (Fig. 5-6).

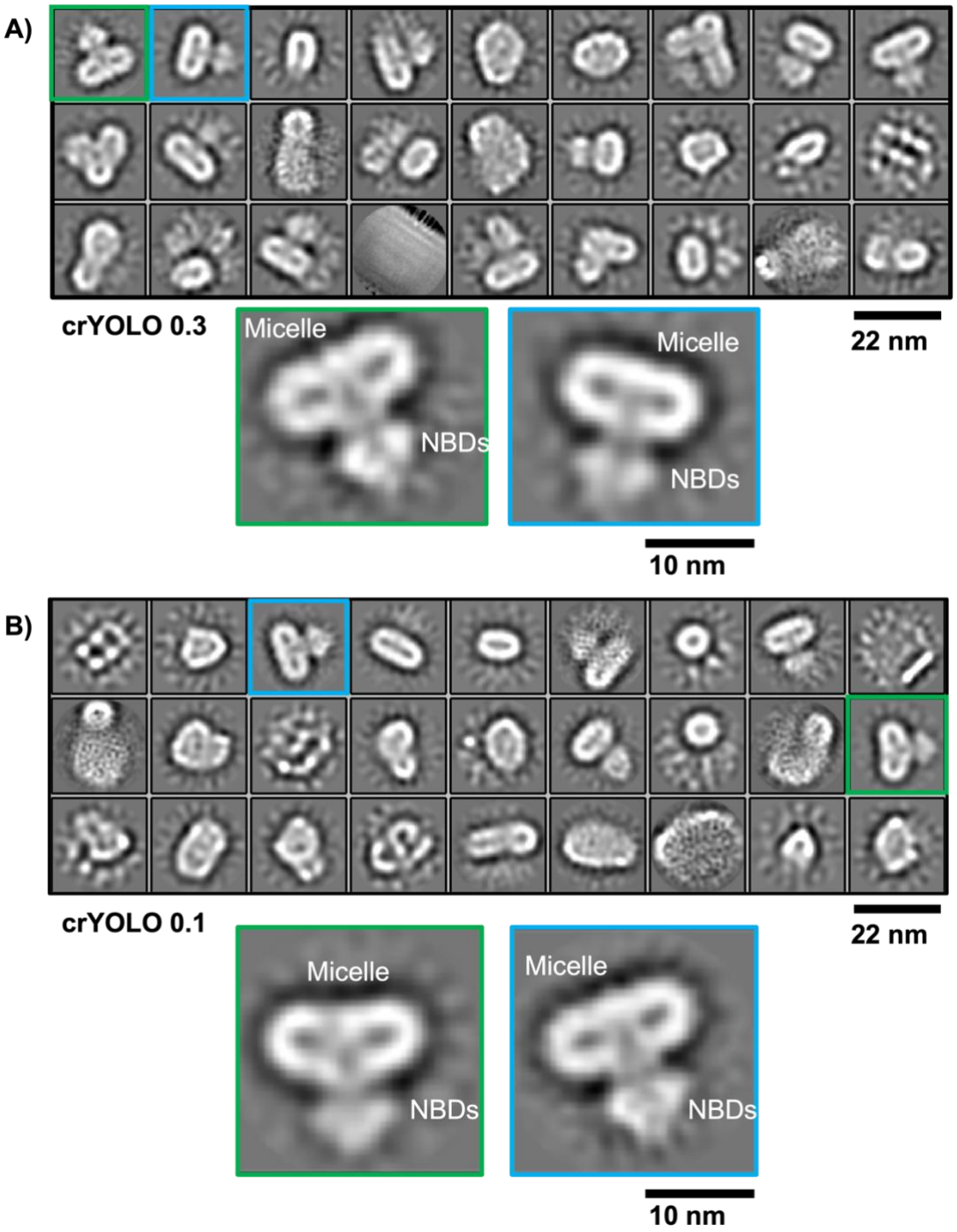


Figure 5-6. 2D classification of particles picked from the largest data set collected for CTS$^{D606N/E607Q}$ using two picking thresholds. 0.3 (default, ***panel A****) or 0.1 (less stringent,* ***panel B****). Expanded views of the best 2D class averages from each set of particles are highlighted by green and blue boxes. Strong density can be observed for the digitonin micelle and NBDs. Density representing the TMDs is also visible.*

For both Pick0.3 and Pick0.1, the multiple expected class averages were observed representing apparent sideways views of $CTS^{D606N/E607Q}$. These classes were the most detailed 2D classes obtained for CTS yet, with multiple classes showing strong density for the digitonin micelle, the NBDs, and secondary structure in the TMD region. As expected, – because $CTS^{D606N/E607Q}$ is a hydrolysis deficient mutant of CTS^{WT} – there were no classes that showed the transporter in a distinctly inward-facing state with NBDs separated. At least four of the most detailed classes from Pick0.3 and Pick0.1 showed $CTS^{D606N/E607Q}$ with an elongated, but less pronounced density, around the NBDs compared to the other classes. This is most likely due to the flexibility of the eGFP tag, because, depending on the orientation of the particles in the vitreous ice, it may or may not be visible. When comparing the 2D classes obtained from Pick0.3 to those obtained from Pick0.1, the classes are very similar, with no substantial difference despite the large difference in particle number (106k vs 276k, respectively).

Given the detailed 2D classes obtained from both Pick0.3 and Pick0.1, a subset of the best classes from each classification were selected for further processing and 3D reconstruction. In both cases, ~90% of picked particles were discarded during subset selection, indicating significant sample heterogeneity. This certainly contributed to unsuccessful *ab initio* reconstruction, but by using the cryo-EM structure of ABCD4 (the human cobalamin transporter) as a template, a 3D classification was successfully carried out on both subset particle stacks from Pick0.3 (14,463 particles total, Fig 5-8) and Pick0.1 (32,174 particles, Fig 5-9). The workflow for cryo-EM data processing of this data set is presented in figure 5-7.

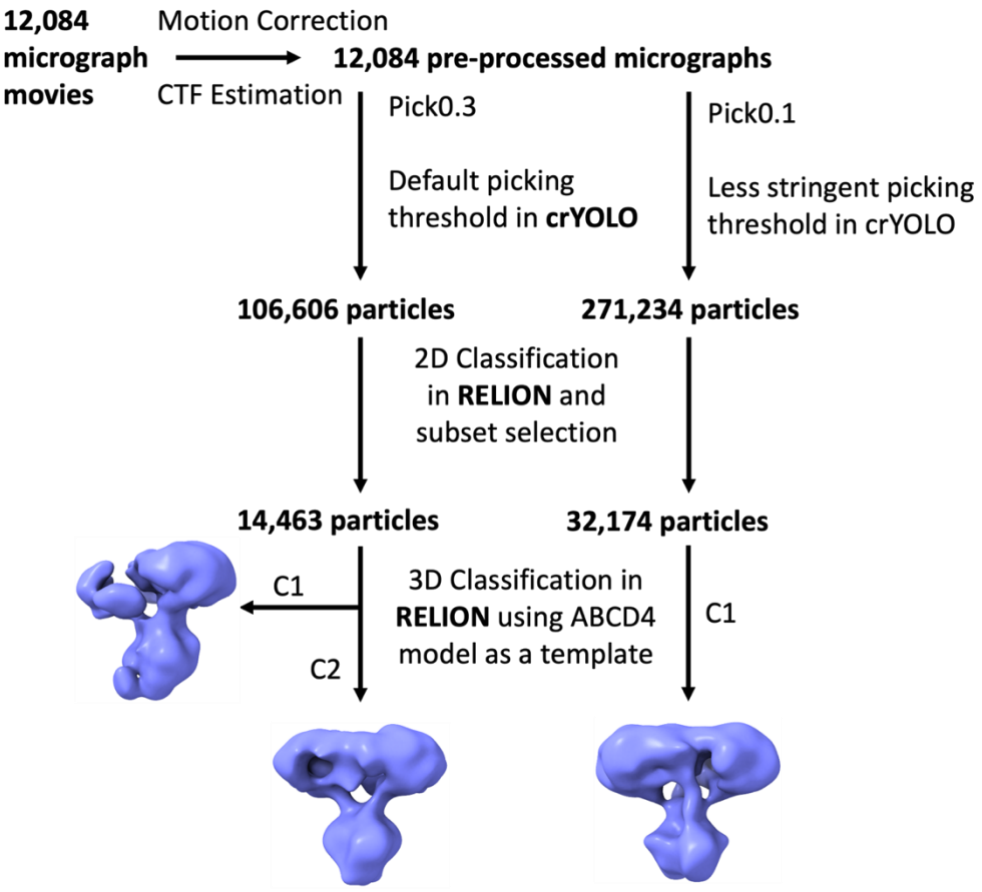


Figure 5-7. A flowchart for cryo-EM data processing of CTS$^{D606N/E607Q}$.

Given the pseudo-symmetrical nature of CTS as a fused heterodimer, the 3D classification on Pick0.3 was carried out by applying C1 and C2 symmetry to compare the two processing parameters. For Pick0.3, 3D classification using C1 symmetry (i.e., no symmetry) yielded two classes of obvious noise (as compared to the 2D classes used for 3D classification and known structures of ABC transporters), and a third class that more closely resembled the expected shape of a low-resolution ABC transporter, albeit with a large chunk of density missing (Fig. 5-8b, top right). By applying C2 symmetry (2-fold rotational symmetry)), three further 3D classes were obtained with one yielding a more convincing map of CTS$^{D606N/E607Q}$ (Fig. 5-8b, top left). This therefore represented the first 3D model of the plant peroxisomal ABC transporter, COMATOSE.

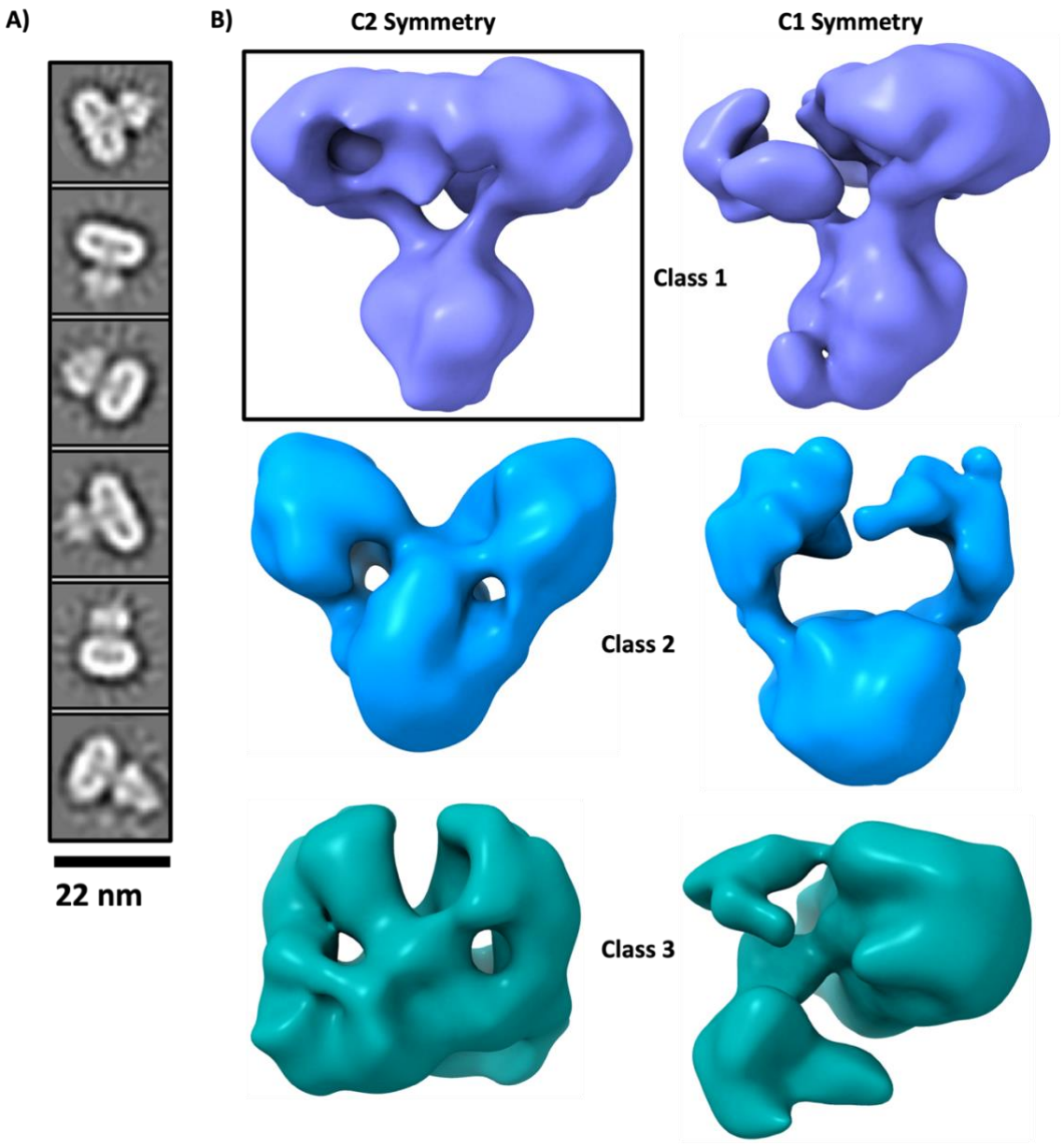


Figure 5-8. CTS$^{D606N/E607Q}$ 3D classification of particles picked using a crYOLO picking threshold of 0.3 (default). ***A)*** *2D class averages from Fig 5-6a that were selected for 3D classification.* ***B)*** *3D maps generated from 3D classification of particles picking using a picking threshold of 0.3, either using C1 or C2 symmetry. Left, C2 symmetry. Right, C1 symmetry. By comparison, it was clear that for each 3D classification, one out of three maps represented a low-resolution ABC transporter, with the map generated after application of C2 symmetry being the most convincing (boxed).*

Despite the low resolution of this maps at ~18 Å, it was clear that the transporter occupied an outward-facing state with NBDs in proximity as should be expected for this mutant. Given that C2 symmetry yielded the best map for Pick0.3, the particle stack obtained from the subset selection of 2D class averages from Pick0.1 was subject to 3D classification using only C2 symmetry (Fig. 5-9).

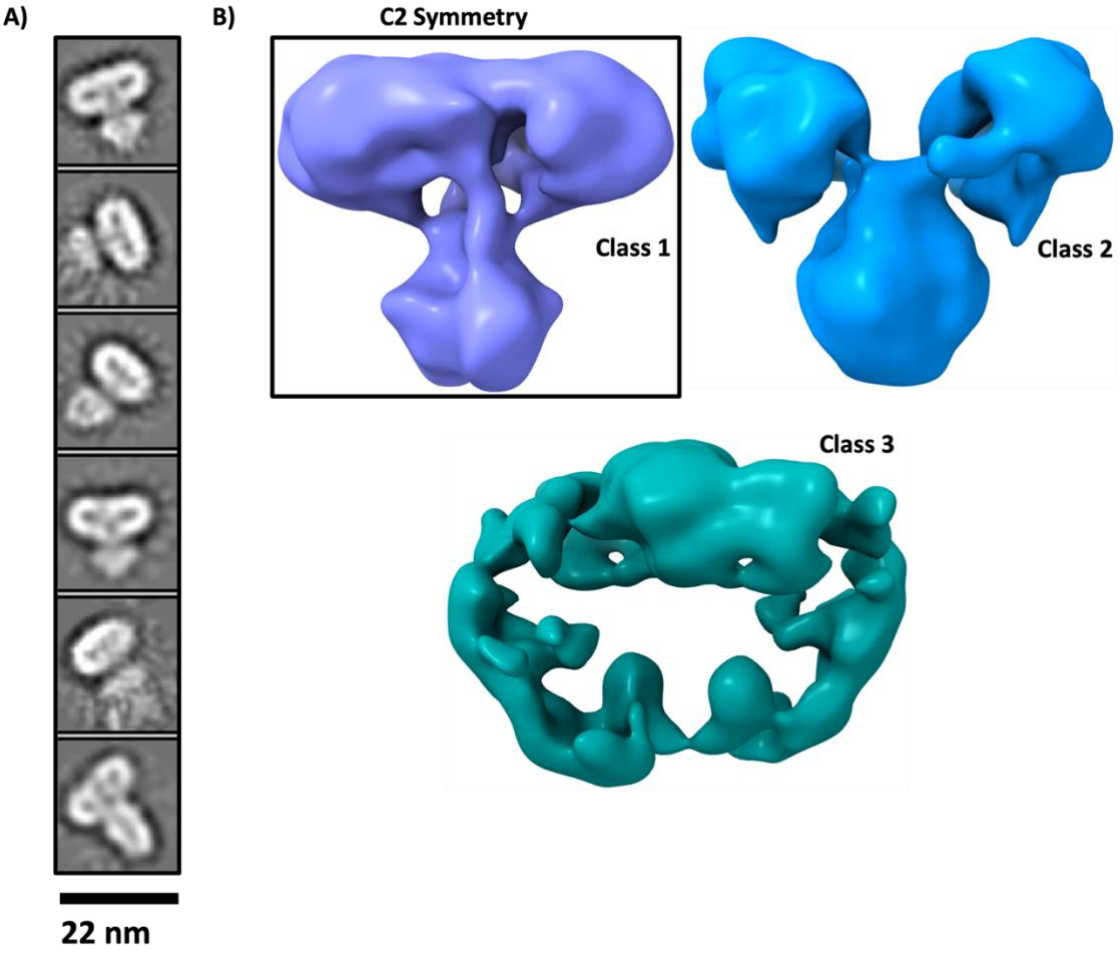


Figure 5-9. CTS$^{D606N/E607Q}$ 3D classification of particles picked using a crYOLO picking threshold of 0.1. ***A)*** *2D class averages from Fig 5-6b that were selected for 3D classification.* ***B)*** *3D maps generated from 3D classification of particles picking using a picking threshold of 0.1, yielding one class of the expected shape and size (boxed).*

As with Pick0.3, the 3D classification for Pick0.1 yielded one of three classes with the expected shape for an ABC transporter. By comparing the 3D models of Pick0.3 and Pick0.1 to the 3D model of the published structure of ABCD4, despite the obvious difference in resolution, the major features of $CTS^{D606N/E607Q}$ were evident (Fig. 5-10), including the digitonin micelle and the dimerised NBDs. The digitonin micelle in both $CTS^{D606N/E607Q}$ maps appears more oblate than the micelle surrounding ABCD4, though this is likely due to the difference in resolution. Some density for the TMDs is present (evidenced by comparison to solved structures of ABC transporters), though in both models, there is missing density where the TMDs enter the digitonin micelle.

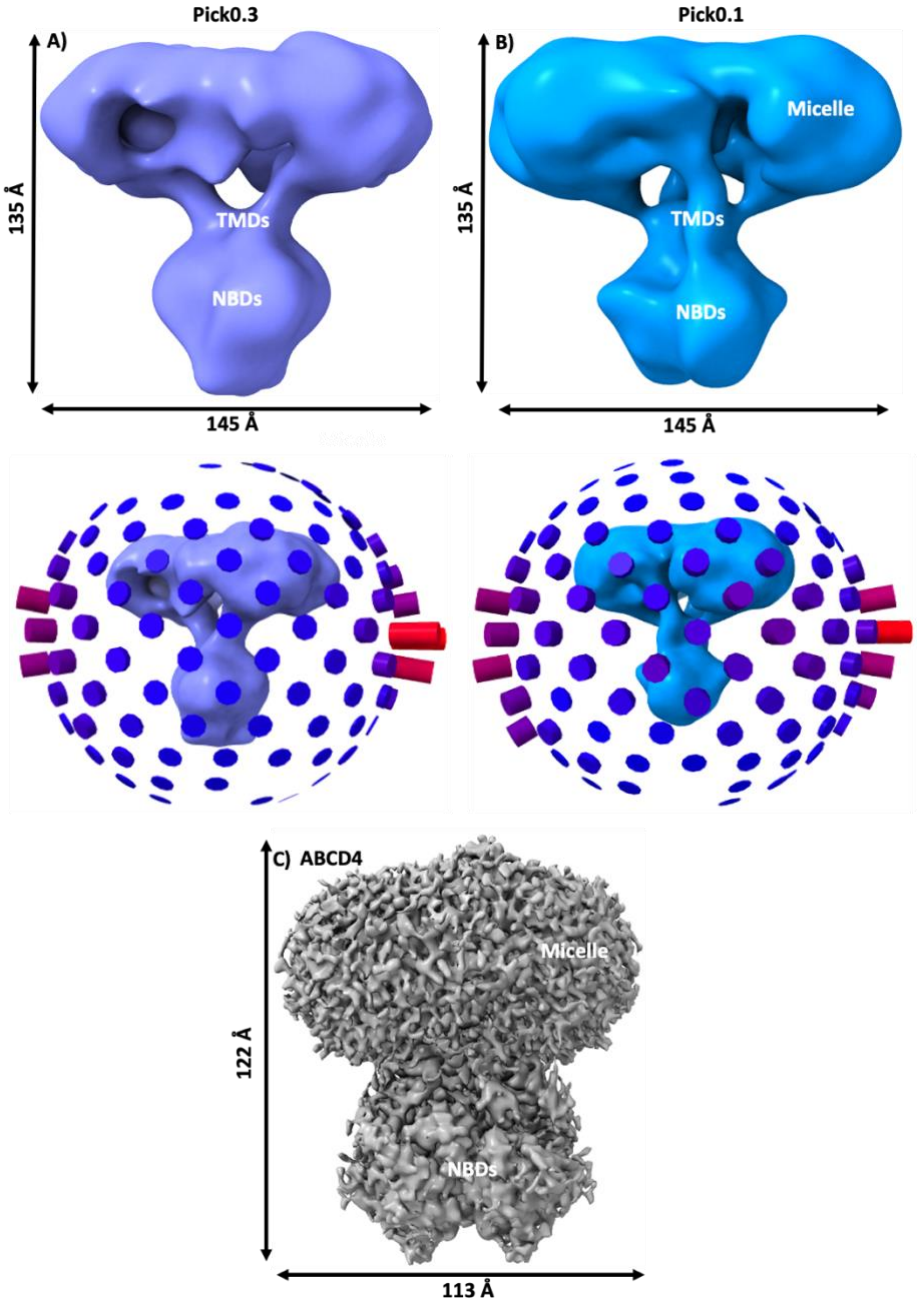


Figure 5-10. 3D models of CTS$^{D606N/E607Q}$ as determined by cryo-EM. ***A)*** *the 3D model generated using a picking threshold of 0.3.* ***B)*** *the 3D model generated using a picking threshold of 0.1. Top panels, equally contoured maps of CTS$^{D606N/E607Q}$ each at a resolution of ~18 Å. Bottom, the angular distribution of particle views that contributed to each 3D model.* ***Red****, more particles;* ***blue****, fewer particles.* ***C)*** *The published map of ABCD4 for comparison.*

Nonetheless, the NBDs, TMDs, and detergent micelle are further highlighted when fitting the structure of ABCD4 into the maps generated for $CTS^{D606N/E607Q}$ (Fig. 5-11) and by comparing the maps to the predicted model of CTS created by DeepMind's 3D protein structure prediction tool, AlphaFold2 (Jumper *et al.*, 2021) (Fig. 5-12). Initially, 3D classification was carried out only on particles obtained from obviously sideways 2D class averages. However, there were several 2D class averages that potentially represented top-down views of $CTS^{D606N/E607Q}$, and if they did, it was contemplated whether their inclusion in further 3D classification would result in a map with improved resolution. Unfortunately, 3D classification including these putative top-down views yielded map of significantly worse resolution (not shown), perhaps due to inclusion of empty micelles, or because these putative views could not be precisely aligned. The angular distribution of the 3D classified particle for both Pick0.3 and Pick0.1 (Fig. 5-10) showed that – despite being chosen for their side-on orientation – these classes containing almost no information for top or bottom views of the particles. This certainly limited the amount of information available for a higher resolution reconstruction. For Pick0.3, of the initial 14,463 particles 6,780 went into the best class (46.9%), whereas for Pick0.1, of the initial 32,174 particles, that number was 11,138 (34.6%). Whilst these percentages are common for 3D classification, it is typically expected that the other classes at least some-what represent the protein of interest. However, this is obviously not true for the other classes generated during 3D classification for $CTS^{D606N/E607Q}$.

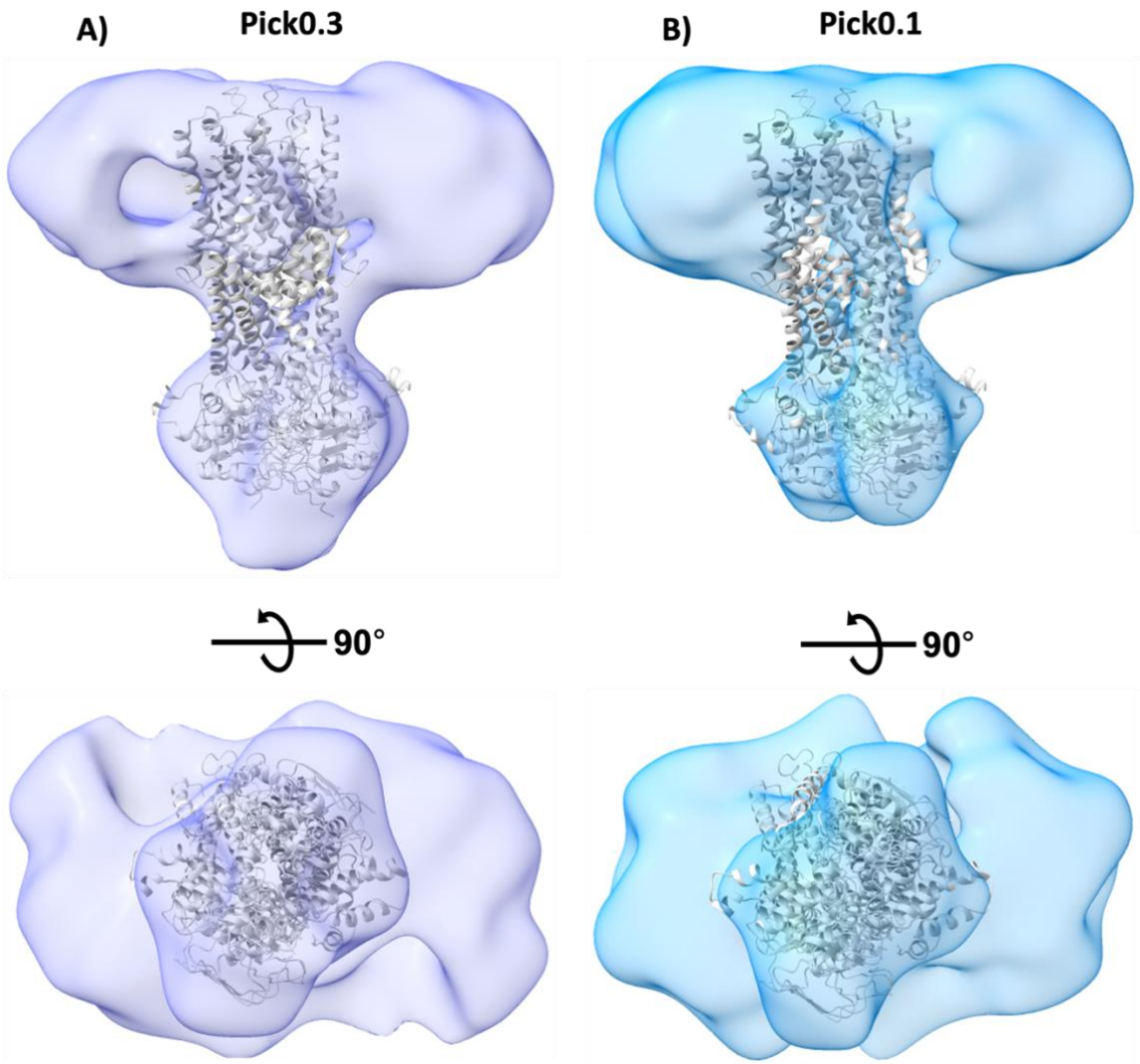


Figure 5-11. The maps generated for Pick0.3 and Pick0.1 fitted with the published model of ABCD4. ***A)*** *Pick0.3.* ***B)*** *Pick0.1. By fitting the structure of ABCD4 into the* $CTS^{D606N/E607Q}$ *maps, it is possible to clearly distinguish the orientation of the transporter's key features within the digitonin micelle.*

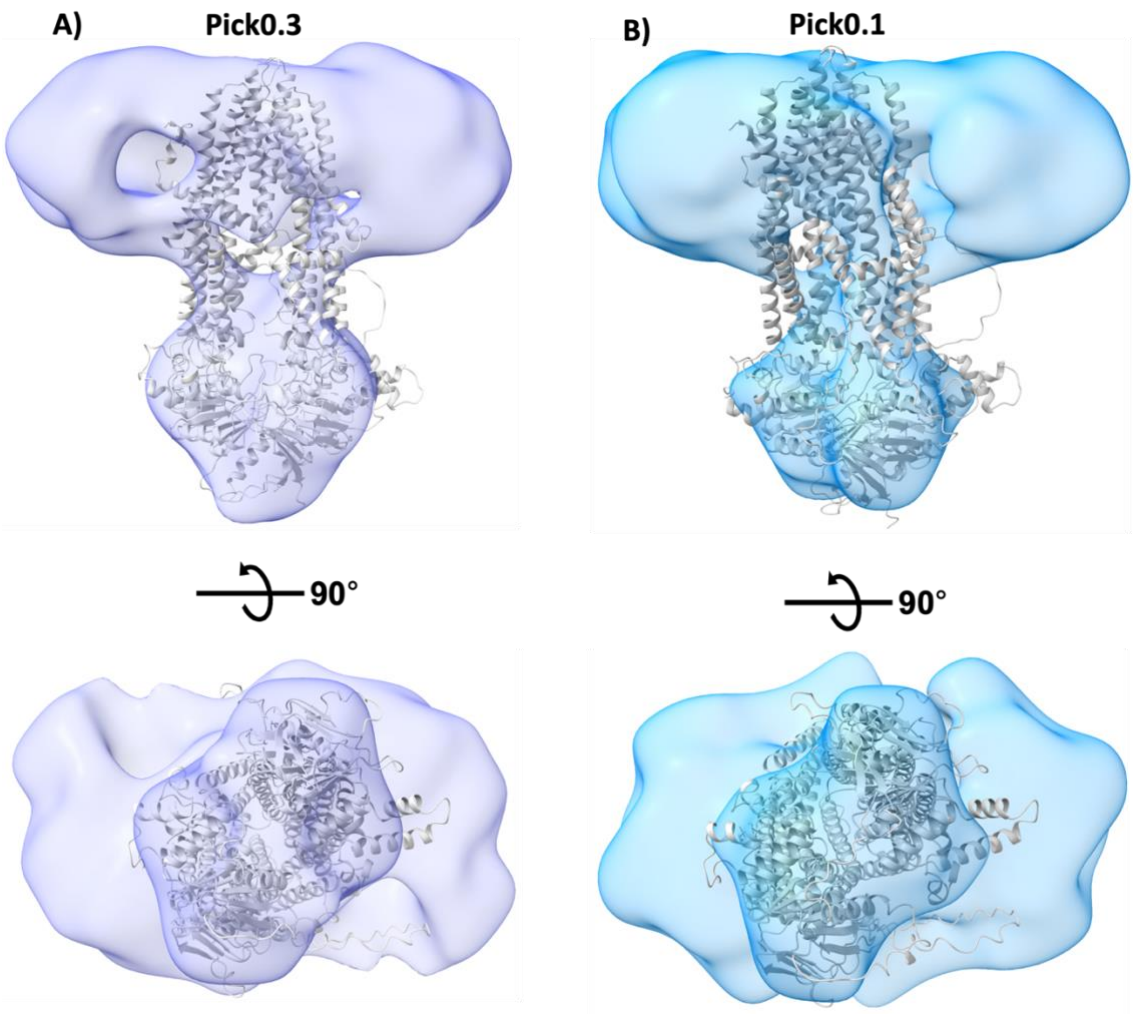


Figure 5-12. The maps generated for Pick0.3 and Pick0.1 fitted with the predicted model of CTS generated by AlphaFold2. The AF2 model of CTS is good, with the majority of residues assigned with confidence or very high confidence with the exception of the linker region and a predicted disordered region at the N-terminus. The model aligns with the C22:0-bound cryo-EM structure of ABCD1 with a root-mean-square deviation (all-atom) of 2.9 Å but does not predict substrate binding sites. ***A)*** *Pick0.3.* ***B)*** *Pick0.1. The fitting of the AlphaFold2 predicted CTS model into the maps for* $CTS^{D606N/E607Q}$ *bears strong similarities to the fitting of the experimentally determined model of ABCD4 into the* $CTS^{D606N/E607Q}$ *maps (Fig. 5-11).*

5.3 Discussion

Before this project began, the function of COMTOSE had been investigated primarily using biochemical analysis, which as discussed in section 1.7 and 1.7.2, led to the observation of COMATOSE thioesterase activity (De Marcos Lousa *et al.*, 2013). However, it was difficult to elucidate the nature of this activity in the absence of high-resolution structural information for the transporter. Indeed, prior to 2019, no structural data existed for any family D ABC transporter, reflecting the challenges associated with working on this family of membrane proteins. The first structure of a family D ABC transporter was that of ABCD4 (Xu *et al.*, 2019). Therefore, the work carried out over the course of this project represented the first attempts to characterise COMATOSE in a structural framework thereby contextualising the earlier biochemical findings with respect to ACOT activity. Consequently, the results presented in this final chapter represents the application of the robust solubilisation and purification protocol established in sections 3 and 4, to the study of a CTS mutant, $CTS^{D606N/E607Q}$, resulting in the first cryo-EM 3D model of the plant peroxisomal transporter.

Unlike CTS^{WT}, which was purified using either DDM or later LMNG/digitonin, $CTS^{D606N/E607Q}$ was purified using only LMNG/digitonin. When using this latter purification method, $CTS^{D606N/E607Q}$ degraded much less severely during the concentration step than CTS^{WT}, an observation that was reproducible during multiple purification attempts. This was a critical step forward in reproducibly producing this challenging protein, which is often the most difficult aspect of membrane protein studies, especially those studies focused on structure determination. This reduced degradation may be explained by considering the difference between the two constructs. During their transporter cycle, ABC transporters exchange between inward- and outward-facing conformations via sandwich dimerisation of ATP between the two NBDs (Higgins & Linton, 2004; Locher, 2016). The presence of ATP during solubilisation means that both CTS^{WT} and $CTS^{D606N/E607Q}$ likely bind ATP, however, in

$CTS^{D606N/E607Q}$ the substitution of the catalytic glutamate for asparagine means that – like other ABC transporters (Kim & Chen, 2018; Liu *et al.*, 2021) - this transporter is incapable of hydrolysing ATP (section 4.2.1), and therefore remains in a closed or outward-facing conformation. In this state, the protein is likely less susceptible to proteolysis in its linker region (Nyathi *et al.*, 2012), and perhaps due to its more compact structure compared to CTS^{WT} (which despite low activity in detergent (at least for DDM, communication with Dr David Carrier), is likely still hydrolysing *some* ATP and therefore moving between open and closed conformational states), is more stable during concentration. Despite the improved concentration compared to CTS^{WT}, a clean-up by SEC was desired to improve sample homogeneity when preparing cryo-EM grids. It is clear from the available literature that the vast majority of high-resolution ABC transporter structures determined using cryo-EM are purified in this way (Hofmann *et al.*, 2019; Johnson & Chen, 2017; Mi *et al.*, 2017; Wang *et al.*, 2020b; Xu *et al.*, 2019; Zhang *et al.*, 2018).

Initial SEC clean-up of $CTS^{D606N/E607Q}$ in digitonin demonstrated that whilst the transporter could be subject to SEC, further optimisation would be necessary to separate the two observed peaks. Interestingly, repeating SEC purification using a higher resolution column yielded a different trace entirely, with a broad hump across the entire column volume. This unexpected observation may have resulted from an incomplete exchange of LMNG on the affinity column during purification. LMNG is known to have an incredibly low 'off-rate', meaning that the exchange between the detergent monomers in solution vs. the detergent monomers in the micelle occurs very slowly compared to other detergents, e.g., four-fold lower than for DDM (Chung *et al*, 2012). This, coupled to the fact that LMNG has an incredibly low CMC at just ~0.001% means that it can be difficult to get complete exchange between LMNG and other detergents, or to remove LMNG entirely, hence the development of the LMNG removal method, GraDeR (Hauer *et al*, 2015). This method works by centrifuging LMNG-purified sample through a gradient comprising an increasing concentration of

glycerol and a decreasing concentration of LMNG. Therefore, the broad hump observed when purifying $CTS^{D606N/E607Q}$ may have been caused by an incomplete exchange between LMNG into digitonin, where the resulting fractions are populated by $CTS^{D606N/E607Q}$ solubilised in either detergent. Moreover, LMNG is known to form micelles ranging from 235 kDa – 622 kDa, with unconventional shapes and mobility properties (Chaptal *et al.*, 2017), which may contribute to the broad hump observed by SEC. However, this result should be treated cautiously as the experiment was carried out just once and so further experimentation is required to evaluate this explanation.

Despite time-constraints that limited optimisation of the SEC clean-up step, affinity-purified $CTS^{D606N/E607Q}$ was subject to cryo-EM sample preparation, screening, and data collection in small scale to determine if this sample would be amenable to structural studies. Three different grid types were used for optimisation: Quantifoil 1.2/1.3 200 mesh, Quantifoil 1.2/1.3 300 mesh, and UltraFoil grids. The most success was had from the Quantifoil grids, though it should be noted that particle concentration was always limited to ~0.8 mg/ml and that particle distribution was non-optimal, often with particles sticking to the carbon film or being poorly distributed. Ultrafoil grids were used to alleviate these problems (Russo & Passmore, 2016), but unfortunately, very few particles could be found while screening, apparently the result of poor vitreous ice formation or the ice being too thick for the electron beam to pass through. Nonetheless, screening of Quantifoil grids confirmed that this sample was amenable to structural studies, where after a small data collection and subsequent data processing, the first 2D classes for this transporter were obtained. This data set was pre-processed, first by 'CTF estimation' via the CTFFind 4 program (Rohou & Grigorieff, 2015) to restore missing information that results from the necessary application of defocus to micrographs during data collection, and secondly, by 'motion-correction' via RELION 3.1.1's implementation of the MotionCor2 algorithm (Zivanov *et al*, 2019) to reduce beam-induced & stage-induced blurring of

micrographs (Rawson *et al*, 2016). Particles were then picked from these pre-processed micrographs using crYOLO which was chosen because of its ability to automatically pick suitable particles even from heterogeneous data sets – as this was expected to be – and to exclude contaminants such as hexagonal ice and aggregates. Moreover, in crYOLO the picking threshold can be altered to compare different sets of particles. Even from the initial poorly detailed 2D classes, when compared to the 2D classes from solved structures of other ABC transporters such as ABCD4 (Xu *et al.*, 2019), ABCB1 (Nosol *et al*, 2020b), and ABCB6 (Wang *et al*, 2020a), it was evident that some classes represented $CTS^{D606N/E607Q}$. Given that these 2D classes could be obtained from an initial data set of ~43k particles, it was reasoned that a larger data set, even without SEC purification would yield a 3D model.

Further screening and processing from 1 h of data yielded ~12,000 particles using crYOLO's default picking threshold and the expected 2D classes, and therefore, a larger collection over 60 hrs was estimated by simple scaling to yield ~700,000 particles. Whilst this number is significantly lower than the often >2 million particles picked when pursuing a high-resolution reconstruction of an ABC exporter, it is still ~10x greater than the original $CTS^{D606N/E607Q}$ data set (table 5.1) and there have been reports of high-resolution structures solved from a little as 250k particles (Johnson & Chen, 2017). However, this did not scale proportionately, and instead of ~700,000 particles being picked, only ~106,000 particles ended up in the default picking (Pick0.3) particle stack. Even the less stringent picking threshold (Pick0.1) yielded ~2.5x fewer particles than expected. This may have been caused by grid storage and handling. After the grid was first screened and 1 h of data collected, it was then returned to storage in LN2. The same grid was then retrieved from storage and put back into the microscope for the second and larger data collection. During this handling, the grid may have warmed or partially thawed and re-frozen, causing the original vitreous ice to be replaced with hexagonal crystalline ice, and/or causing particles to re-distribute in the remaining ice (Cabra & Samso, 2015). Additionally,

when collecting data for 1 h just one grid square was used, whereas when data was collected for 60 hrs, many gird squares were collected on. Moreover, there is often a substantial variation in ice-thickness across a single grid which is dependent on several factors including the blot time, blot force, and the blotting method used (Drulyte *et al.*, 2018) Even the irregular pattern of fibres in the blotting paper used can affect ice thickness and be detrimental to samples (Armstrong *et al*, 2020) further increasing grid variability and limiting reproducibility (Sgro & Costa, 2018). Therefore, some grid squares simply may have had thinner, more suitable vitreous ice than others. Another factor to consider is the effect that the air-water interface has during grid preparation. Between blotting and plunge freezing, the proteins suspended in the thin aqueous film formed by the blotting process are subject to Brownian motion, causing the molecules to interact with the air-water interface thousands of times per second (Taylor & Glaeser, 2008), where they are at risk of unfolding and denaturation (D'Imprima *et al*, 2019; Glaeser & Han, 2017). This effect is more or less pronounced depending on the thickness of ice that is formed during vitrification and occurs as quickly as six milliseconds (Klebl *et al.*, 2020). The article by Klebl *et al.* (2020) highlights the variety of factors that affect grid preparation, and why no 'silver bullet' has been found. In the case of $CTS^{D606N/E607Q}$ here, there may have been a substantial variation in ice thickness across the grid where in some cases interactions with the air-water interface led to the denaturation of useable particles therefore limiting the number of particles to be picked and processed downstream.

Despite this unfortunate particle scaling issue, the data set resulted in the largest number of particles picked for CTS so far. Particles were picked using two different picking thresholds, and each particles stack contributing to the most detailed 2D classes ever obtained for CTS, which resulted not only from an increased number of particles picked, but also from optimisation of the particle extraction parameter. Initially particles were extracted as in the original data set (table 5.1) in a box size of 180 pixels, however, these classes were noisy and lacked the detailed expected from

particle stacks of this size despite the data being expectedly heterogeneous. Re-extraction using a box size of 256 pixels substantially improved this, yielding 2D classes that appeared promising for the generation of a 3D model.

At the time of this data processing, the only structure available for a family D ABC transporter was that of ABCD4. Given that *ab initio* reconstruction was not possible, most likely because of the low number of particles and lack of necessary orientations, and because 3D classification requires an initial reference map, the model of ABCD4 was the obvious starting point. To reduce bias in the reconstruction, the ABCD4 model was first converted to a map file using the MolMap tool in ChimeraX then low-pass filtered to 30Å during the 3D classification. Moreover, this model fortunately represented a hydrolysis-deficient mutant of ABCD4 bearing an E549Q mutation similar to the E607Q mutation in $CTS^{D606N/E607Q}$, which further reduced reconstruction bias given that $CTS^{D606N/E607Q}$ should adopt a similar conformation to this ABCD4 mutant. This 3D classification yielded the first 3D models for $CTS^{D606N/E607Q}$ (one from each particle stack). Whilst the resolution of these maps was low, they provide a basis for further cryo-EM optimisation in pursuit of a high-resolution structure.

Interestingly, in both Pick0.3 and Pick0.1, ~90% of the initially picked particles were discarded during the subset selection process after initial 2D classification, and the remaining particles stacks used for 3D classification of Pick0.3 and Pick0.1 yield similar 3D maps. Given that the maps looked similar despite Pick0.1 containing almost twice as many particles as Pick0.3, there is an argument to use the less stringent particle picking parameter when processing future CTS data sets as the higher threshold (Pick0.3), was missing particles that certainly contributed to the Pick0.1 map. This would be magnified if processing an even larger data set and might cause information-rich particles to be excluded. At around the time of writing, the structure of ABCD1 was solved by cryo-EM (Wang *et al.*, 2021). Despite the low sequence identity between HsABCD1 and HsABCD4 (~28%), the overall fold of the two nucleotide bound transporters is very similar (Wang *et al.*, 2021), with a root-

mean-square-deviation of just 1.87 Å. Moreover, from fitting the predicted AlphaFold2 structure of CTS into the $CTS^{D606N/E607Q}$ maps, it is clear that the predicted model fits in a similar fashion to ABCD4. Therefore, 3D classification of the $CTS^{D606N/E607Q}$ particle stacks using the nucleotide bound structure of HsABCD1 (Wang *et al.*, 2021) or the AlphaFold2 predicted model of CTS as a template would not be expected to substantially improve the resolution of the ~18 Å maps obtained for $CTS^{D606N/E607Q}$ using ABCD4 as a template.

Nonetheless, the maps presented here for $CTS^{D606N/E607Q}$ confirm for the first time that the COMATOSE transporter adopts a similar fold to other family D ABC transporters, probably with the domain-swapped architecture (Szollosi *et al.*, 2018) of other ABC exporters that was predicted for CTS by homology modelling (Dietrich *et al.*, 2009). However, the resolution does not allow any inference to be made regarding the intriguing functional asymmetry of CTS first reported by Dietrich *et al.* (2009). Unlike its other family members, CTS is a fused heterodimer, that possesses non-equivalent NBDs where the Walker A motif of NBD1 is dispensable for function; further insights into this feature of CTS will be gleaned only from high-resolution studies (~3-4 Å).

At the time of writing, multiple structures of ABCD1 were solved (Chen *et al.*, 2022; Jia *et al.*, 2021; Le *et al*, 2022; Wang *et al.*, 2021; Xiong *et al.*, 2021). By comparing the methods used to purify and determine the structure of ABCD1 (excluding work by Jia *et al.* (2021) due to a lack of information detailing methods) to the methods used for CTS (table 5.3), it is evident that the major limiting factor for $CTS^{D606N/E607Q}$ was indeed the SEC clean-up step, which unfortunately due to time constraints, could not be optimised. It is known from the SEC purification presented in figure 5-3 that a substantial amount of aggregation is present in the sample after concentration, which certainly contributed to sample heterogeneity. As with all solved structures of ABCD1, the SEC step would have eliminated these aggregates thereby creating a more homogeneous sample.

Another consideration is the nature of the protein's tag. Whilst $CTS^{D606N/E607Q}$ has a HRV3C cleavable eGFP tag, it was not cleaved off using the purification and instead, the rationale was that because CTS is a psuedosymmetric fused heterodimer, inclusion of the tag would break this pseudosymmetry and simplify image processing. Moreover, the tag would simply increase the size of the CTS particles and therefore simplify particle picking. However, it is clear from the 3D classification of both Pick0.3 and Pick0.1 that C2 symmetry is more suitable than C1 anyway, and therefore cleavage of the tag may have been the more suitable approach. Indeed, none of the constructs used to determine the structures ABCD1 had an eGFP tag, and though the construct used by Le *et al.* (2022) did include a YFP tag, it was ultimately cleaved off during the affinity purification. Therefore, the eGFP tag that is fused to $CTS^{D606N/E607Q}$ by a flexible linker may have contributed to the low-resolution map by reducing the accuracy of particle alignment against the template during 3D classification.

Additionally, it should also be noted that of all the reported structures of family D ABC transporters (ABCD4 and various structures of ABCD1), all are from humans and therefore, – with the exception of the study by (Jia *et al.*, 2021) – have been expressed in HEK cells. Interestingly however, it has been reported that the expression of unstable ABCD1 mutants in fibroblasts can be improved by low-temperature culture (Zhang *et al*, 2011). Thus, when comparing insect cells vs. HEK cell expression systems (27°C and 37°C, respectively), the lower temperature used during insect cells expression may indeed be advantageous. Whilst this certainly appears suitable for a plant protein that has evolved to function at 23°C (the optimal growth temperature of *A. thaliana*) and despite the strong precedent for expression in insect cells, it may be interesting to empirically assess the expression and stability of CTS produced in HEK cells given that its closest homologs have been produced in this system.

	CTS$^{D606N/E607Q}$	ABCD1 (Wang *et al.*, 2021)	ABCD1 (Le *et al.*, 2021)	ABCD1 (Chen *et al.*, 2021)
Expression system	Sf9 insect cells	HEK-293S GnTI−	HEK 293	HEK 293F
Solubilisation	1%LMNG/0.1%CHS	1% LMNG/0.1%CHS	1% DDM/0.2% CHS	1% LMNG/0.1%CHS
Affinity Purification	Streptactin resin	Flag resin	Rho-1D4 antibody conjugated Sepharose 4B resin	Flag resin
Elution	5 mM desthiobiotin 20 mM HEPES pH 7.5, 200 mM NaCl, 2 mM $MgCl_2$, 5% glycerol (v/v) 0.06% digitonin (w/v)	Flag peptide 25 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5% glycerol (w/v), 2 mM MgCl2, 0.06% digitonin	Tag cleavage 25 mM HEPES pH 7.5, 150 mM NaCl, 20% glycerol (v/v), 0.02% DDM/0.004% CHS (w/v) Followed by nanodisc reconstitution	Flag peptide 25 mM Tris-HCl pH 7.5, 150 mM NaCl, 10% glycerol (v/v),1 mM DTT, 0.06% digitonin (w/v)
SEC?	n/a	Superose 6 column 25 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5% glycerol (v/v), 2 mM MgCl2, 0.06% digitonin (w/v)	G4000swxl SEC column 25 mM HEPES pH 7.5, 150 mM NaCl	Superdex 200 Increase 10/300 column 25 mM Tris-HCl pH 7.5, 150 mM NaCl, 1mM DTT, 0.06% digitonin (w/v).

Table 5.3. Methods used to study different ABCD Transporters. Comparison of the methods used to purify CTS$^{D606N/E607Q}$ in this study vs. methods used to purify ABCD1 for structural studies by cryo-EM. Yellow, similar methods by row. Green, significantly different methods by row.

In summary, this chapter represents a significant step forward in the study of COMATOSE as prior to this project, no structural information for the transporter existed. The maps presented here represent first ever structural information for this intriguing transporter, revealing its overall architecture and confirming that it adopts a similar fold to its other family members. Moreover, these insights are further validated when comparing these $CTS^{D606N/E607Q}$ maps to the model predicted for CTS by AlphaFold2. Finally, this work has established the baseline for all future cryo-EM studies on CTS and has highlighted the steps necessary to push these initial low-resolution maps into high-resolution models within the near future to finally elucidate the nature of the transpor’er's thioesterase activity.

6 Conclusion and Outlook

The study of the peroxisomal ABC transporter COMATOSE has spanned ~20 years, from its discovery until the present day. In that time, it's essential role in *A. thaliana* beta oxidation has been well established (section 1.7) (Carrier *et al.*, 2019; Dave *et al.*, 2011; De Marcos Lousa *et al.*, 2013; Dietrich *et al.*, 2009; Footitt *et al.*, 2007; Footitt *et al.*, 2002; Hooks *et al.*, 2007; Nyathi *et al.*, 2010; Nyathi *et al.*, 2012; Russell *et al.*, 2000; Theodoulou *et al.*, 2005) although the nature of its most intriguing characteristic – its intrinsic thioesterase activity (De Marcos Lousa *et al.*, 2013) – has not yet been elucidated due to a lack of structural data for CTS and until recently, no structural data existed for any family D transporter (first reported 2019, ABCD4 (Xu *et al.*, 2019)). Around the time that CTS ACOT activity was reported, the field of cryo-EM was beginning its resolution revolution (Kuhlbrandt, 2014), whereby developments in EM hardware and software was pushing the boundaries of what EM could do, resulting in the first near atomic-resolution protein structures using this method (Bartesaghi *et al*, 2014). As this period continued (and still does), smaller proteins were able to be studied using cryo-EM (Khoshouei *et al*, 2017; Khoshouei *et al*, 2016), as well as those that has previously resisted crystallisation (for study by x-ray crystallography) and/or were highly dynamic (Zhang *et al*, 2017). Given that ABC transporters fell into these categories, determination of their structures by cryo-EM proliferated rapidly (highlighted by figure 4-1, section 4.1). Therefore, at this crossroad it appeared that structural studies of CTS could now be pursued, and as such, this thesis represents the first work carried out on CTS from a structural perspective, using cryo-EM to elucidate the nature of CTS ACOT activity.

Therefore, the first major aim of the project was to determine if CTS was amenable to study by cryo-EM. At the beginning of the project, a series of protocols that had previously been used for expressing and purifying CTS were evaluated for their amenability to cryo-EM studies. Over the course of the project, significant

modifications and improvements were made to these protocols through empirical testing, which included optimised expression testing for large-scale expression of CTS^{WT} and mutants in Sf9 insect cells, optimised detergent solubilisation and assessment of non-detergent alternatives, and modification of buffer composition to improve protein stability and ease down-stream purification and cryo-EM grid preparation. This work culminated in a robust and reproducible expression, solubilisation, and purification protocol for CTS^{WT} and $CTS^{D606N/E607Q}$, and was shown through cryo-EM grid screening to be clearly suitable for further cryo-EM studies.

Building upon these developments, a second major aim of this project was to determine if CTS, after purification, was still in an active form. This would be vital for future studies on the transporter, because until now biochemical studies on CTS (e.g., the study on its ACOT activity), had only been reported in isolated Sf9 insect cell membranes or isolated yeast peroxisomes. To do this, a proteoliposome reconstitution protocol was optimised, whereby the detergent purified transporter (purified using the previously established method, section 2.10.3) was inserted into asolectin/CHS liposomes and the ATPase activity measured in the presence of substrate (oleoyl-CoA) or inhibitor ($AlF3_3$). This confirmed that after reconstitution the transporter possessed substrate stimulated ATPase activity and sensitivity to the known ATPase inhibitor, AlF_3. The CTS mutant $CTS^{D606N/E607Q}$, which had been presumed as hydrolysis-deficient, was shown as such. Together, these data showed that CTS could be reconstituted in an active form, opening up the possibility of studying the catalytic and transport cycle of CTS in a totally isolated system, and thereby provided a new method for investigating the nature of CTS ACOT activity that was previously inaccessible, and one that would complement future structural studies.

The final aim of this study was to determine the structure of CTS using cryo-EM, to place the biochemical observations of the transporter in a structural framework. To do this, the established purification protocol was employed to purify $CTS^{D606N/E607Q}$ in

digitonin and cryo-EM data collected in the Astbury Biostructure Laboratory, resulting in the determination of two maps of CTS$^{D606N/E607Q}$. This represented the first structural data ever obtained for CTS and for a plant ABC transporter.

These two maps were obtained from a single CTS$^{D606N/E607Q}$ data set using two different particle picking parameters and by using the recently solved structure of ABCD4 as a template (Xu *et al.*, 2019) for 3D classification. Although the maps obtained for CTS$^{D606N/E607Q}$ were of low resolution, from the experiments carried out during this study it is now clear what experiments need to be done to obtain a high-resolution model of CTS. Despite several grid types being utilised during this study, it was not always possible to observe CTS particles, or image through the ice due to poor vitreous ice formation. Therefore, further work is needed to determine the optimal grid for CTS data collection. This will improve particle distribution on the grid and hopefully increase the number of particle orientations observed in the ice, both of which are key parameters for high-resolution single particle cryo-EM (Armstrong *et al.*, 2020; Drulyte *et al.*, 2018; Kampjut *et al.*, 2021; Sgro & Costa, 2018).

Moreover, due to time constraints, it was not possible to optimise a SEC step to clean up the sample prior to cryo-EM grid preparation, and therefore, experiments that surmount this issue are of paramount importance going forward as the extent of sample heterogeneity (i.e., heterogeneity should be reduced) will determine if a high-resolution 3D reconstruction is obtainable. Overcoming these limitations will be necessary for the *ab initio* 3D reconstruction of CTS that was not possible in this study. Despite these limitations, the maps obtained for CTS$^{D606N/E607Q}$ were indeed convincing as compared to other ABC transporters (e.g., nucleotide-bound ABCD4 (Wang *et al.*, 2021) and the newly solved structure of nucleotide-bound ABCD1 (Wang *et al.*, 2021)) clearly showing the typically features of these proteins purified in detergent, such as the TMDs surrounded by a detergent micelle with the NBDs protruding out below. Interestingly, it is clear from these maps and from 2D classifications that the transporter adopted the outward facing state with NBDs in

close proximity, as was expected from the hydrolysis deficient mutant (i.e., the transporter was nucleotide bound). This map confirmed for the first time that CTS does indeed share a similar fold to other ABC exporters and therefore, according to the new ABC transporter classification, belongs to the Type IV family (Thomas *et al.*, 2020). However, it should be noted that since the maps of CTS were generated using ABCD4 as a template, there is likely some bias in the 3D classification and tends towards a map of a similar shape to ABCD4. Nonetheless, this was partially overcome by the template being low-pass filtered to 30 Å before 3D classification, thereby removing a significant amount of high-resolution bias and lending validity to the maps generated. Ultimately, a more homogeneous data set that allows for an *ab initio* reconstruction of CTS will likely corroborate the maps presented here.

Now that the fold of CTS has been confirmed and it is indeed similar to other family D transporters, it is possible to consider how CTS may carry out its ACOT activity. This consideration can be divided into several questions: what is the exact mechanism of acyl-CoA cleavage? How do substrates bind CTS? What is the fate of the CoA moiety after cleavage? Why cleave at all? Firstly, whilst the residues identified by Carrier *et al.* (2019), are clearly functionally important (as is the previously described S810 residue (Dietrich *et al.*, 2009)) their molecular interaction with substrate and their function with respect to substrate cleavage is unknown. A detailed understanding of their contribution to the transport cycle, either as part of a catalytic triad or as substrate binders, or residues important in protein stability or movement will require the high-resolution substrate-bound structure of CTS. Only one possible mechanism of substrate cleavage has currently been suggested and that is by Kawaguchi *et al.* (2021) who showed that after incubation of human ABCD1 with NBD-C16-CoA, the NBD-C16 moiety is covalently bound to the transporter via acylation, and that this covalent intermediate is formed by thioester linkage of the NBD-C16 to cysteine residues of ABCD1.

However, the exact cysteine residues which are involved in NBD-C16 binding were not identified. By using an ABCD1 homology model to assesses the environment in which the three ABCD1 cysteines reside, the binding was hypothesised to be mediated via Cys39 or Cys88 instead of Cys121 because of the latter's predicted location on a loop facing the peroxisomal lumen. However, the recently solved structure of ABCD1 by Wang *et al.* (2021) (discussed below), has revealed the true position of each cysteine residue, showing that Cys121 sits close to the substrate binding pocket, thereby implicating it as the NBD-C16 binder. However, it is unclear from current studies whether the cysteine residues of ABCD1, and by extension those residues in other family D transporters, play a catalytic or conformational role in the transport cycle. This is also true of the S810 residue in CTS, which sits in a similar position to Cys121 in ABCD1 (i.e., near a suspected substrate binding pocket (discussed below)), and has been shown to be important for CTS ACOT activity (Carrier *et al.*, 2019; Dietrich *et al.*, 2009). Like the cysteines in ABCD1, the role of S810 in the transport cycle of CTS is unclear and although its position alludes to a role in conformational change, further insights will require a high-resolution 3D model to shed light upon its this. It is worth noting that a structural superposition of the cryo-EM structure of ABCD1 bound to C22:0 (PDB 7VZB) and the AF2 model of CTS shows a root-mean-square-deviation (all-atom) of 2.9 Å (alignment generated in PyMOL using the 'align' command). Thus, it appears that insights gained from structures of ABCD1 can be used to make plausible inferences about CTS.

Additionally, several cryo-EM structures of human ABCD1 have been reported (Chen *et al.*, 2022; Jia *et al.*, 2021; Le *et al.*, 2022; Wang *et al.*, 2021; Xiong *et al.*, 2021). Of these, in the study by Wang *et al.* (2021) two structures were solved, one outward-facing ATP-bound state, and an inward-facing state bound to oleoyl-CoA. In the latter, two molecules of oleoyl-CoA are found, one bound to each TM3-6 bundle in the two halves of the transporter. Interestingly, in the ATP-bound state these molecules are replaced by ambiguous density that may be the free fatty acid product of the

transporters thioesterase activity, though it does not appear that any cysteines are involved in the binding, as discussed by Kawaguchi *et al.* (2021). The acyl-CoAs in this structure lie parallel to the TM helices, interacting with a single TMD bundle. However, another structure by Chen *et al.* (2022) shows two molecules of C22:0-CoA bound to the transporter, each with its CoA moiety interacting in planar fashion with a hydrophilic cavity in one TMD domain, and the long acyl chain extending across the to the opposite TMD interacting with a hydrophobic cleft. This binding mode, which is distinctly different from the oleoyl-CoA bound structure by Wang *et al.* (2021) is different again from the binding of C26:0 and C26:0-CoA in an ABCD1 structure reported by Xiong *et al.* (2021) which indicated that while the CoA moiety is bound in a cytosolic vestibule, the acyl chain protrudes out into the lipid bilayer through a small opening in one TM bundle, a mechanism which was hypothesised by Le *et al.* (2022) . This variety of binding modes – which may be related to the flexibility of acyl-chains – and the fact that family D ABC transporters collectively transport a variety of different molecules indicates that family D transporters may possess multiple substrate binding sites, which also may be true of CTS.

By identifying the residues involved in substrate binding in these ABCD1 structures and comparing these residues to the equivalent residues in CTS (Fig. 6-1), it is possible to identify conversed residues in CTS and place them in a CTS homology model to identify potential substrate binding pockets (Fig. 6-2). From this, it is clear that many of the resides involved in acyl-CoA binding in ABCD1 are shared by CTS or that CTS possesses strongly chemically similar equivalent residues. Given that the CoA moiety of different acyl-CoA is shared, it may be expected that residues contacting the CoA are the most similar, with more variation being observed in residues contacting the variable acyl-chain. However, it appears that there is a general variability in the residues involved in binding acyl-CoA with no substantial conservation of CoA binding residues compared to acyl-chain binders. Nonetheless, the identification of conserved substrate binding residues in CTS indicates that CTS

may share some of the substrate binding modes that have been revealed by structures of ABCD1.

It is interesting that none of the residues involved in substrate binding in ABCD1 and that are conserved in CTS are in close proximity to the triad of residues identified by Carrier *et al.* (2019) as being critical for ATPase and ACOT activity, all of which reside closer to the TMD-NBD interface than deep within the TMDs. A detailed understanding of substrate binding and cleavage by CTS will require a high-resolution substrate bound structure; this will be necessary to determine if CTS substrates occupy specific binding pockets, or whether the physiochemical characteristics of the substrate are a more important factor.

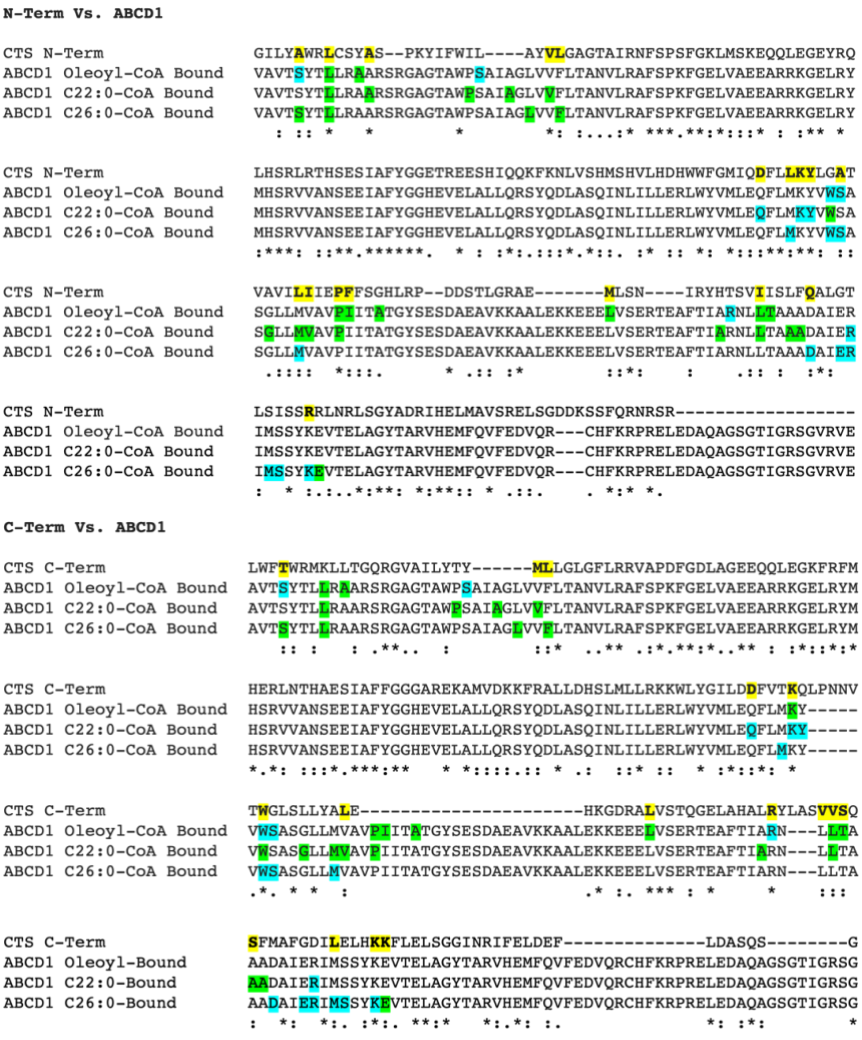


Figure 6-1. Sequence alignment of CTS with ABCD1 to identify residues involved in substrate binding. Substrate-binding residues in three different structures of ABCD1 were identified in the sequence alignment (green, residues contacting the acyl-chain; blue, residues contacting the CoA moiety). ***Top****, sequence alignment between the N-terminus of CTS (residues 236-439) and ABCD1 (residues 222-458).* ***Bottom****, sequence alignment between the C-terminus of CTS (residues 892-1081) with ABCD1 (residues 223-454). CTS residues that are identically conserved (asterisk) or strong conserved (colon) are highlighted in yellow. Weakly similar residues (.) are not highlighted. Oleoyl-CoA bound structure, Wang et al. (2021); C22:0-CoA bound structure, Chen et al. (2022); 26:0-CoA structure, Xiong et al. (2021). Sequences obtained from UniProt.*

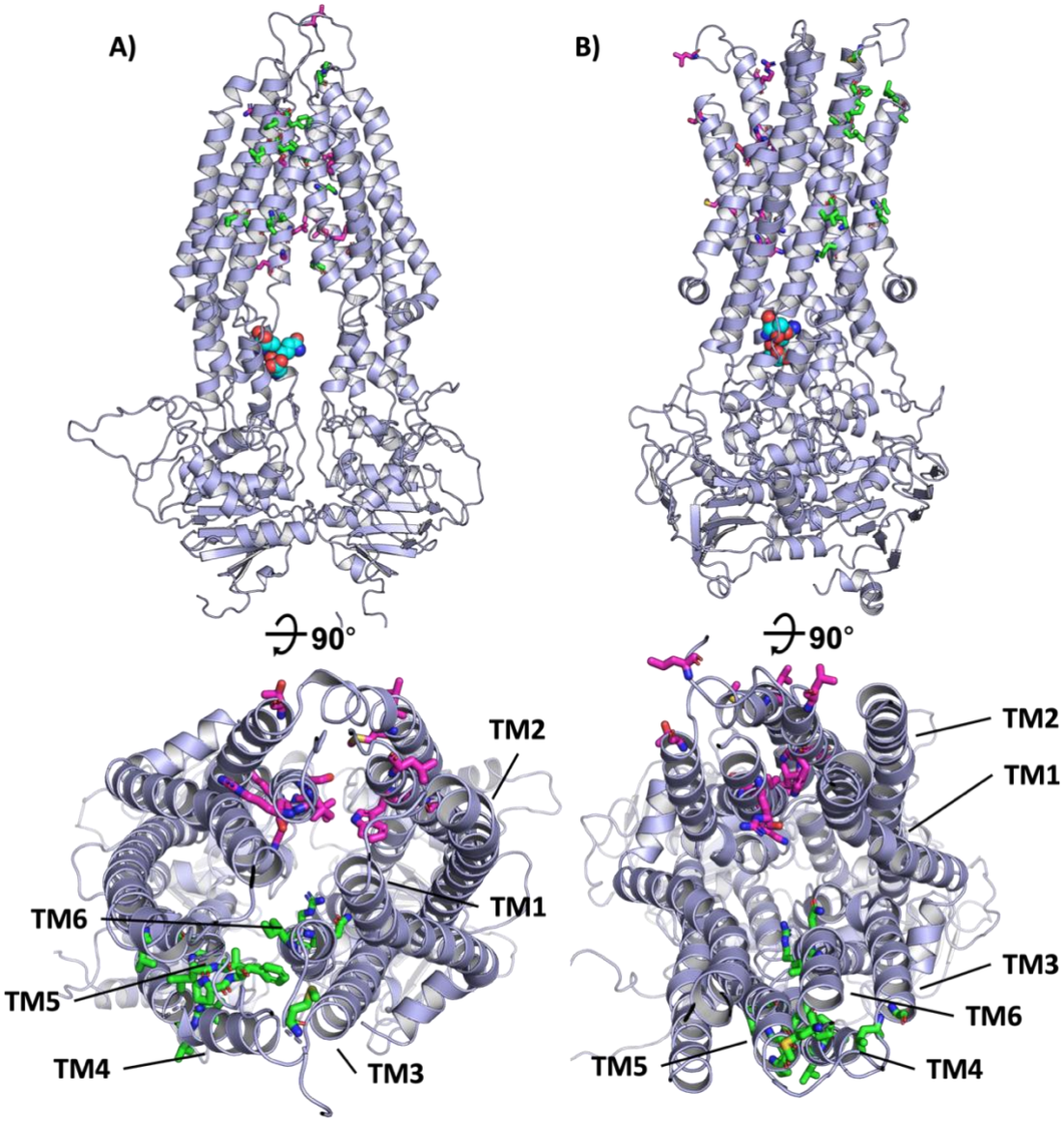


*Figure 6-2. Inward- and outward-facing CTS homology models showing potential substrate binding residues identified by comparison to substrate-bound structures of ABCD1. Substrate binding residues in ABCD1 have identically conserved or strongly conversed equivalent residues in CTS that may represent CTS substrate binding pockets. Potential substrate-binding residues are highlighted per half of the transporter (**green**, N-term; **pink**, C-Term). **A)** Inward facing state. **B)** Outward facing state. Helices from one bundle of TMD helices are annotated (TM1-6). Bottom panels in A and B represents a top-down view of the TMDs. Putative triad residues D863, Q864, and T867 are shown as spheres in the top panels.*

Of all currently solved structures of ABCD1, the majority have been solved in detergent (Chen *et al.*, 2022; Wang *et al.*, 2021; Xiong *et al.*, 2021) with just one study reporting structures in lipid nanodiscs (Le *et al.*, 2022). In the report by Wang *et al.* (2021), detergent-purified oleoyl-CoA-bound ABCD1 was observed in an inward conformation with NBDs extremely widely separated. Chen *et al.* (2022) reported a similar structure but instead bound to C22:0-CoA, with the transporter still in the inward-facing state but with NBDs in much closer proximity. It is important to note than this partial closure of the NBDs may be caused by C22:0-CoA bridging the two halves of the transporter (whereas oleoyl-CoA is exclusive to each TMD bundle) similar to that reported for MRP1 (Johnson & Chen, 2017). Xiong *et al.* (2021) reported four inward-facing states, only one of which has NBDs in close proximity despite one being a hydrolysis deficient mutant (that is expected to have NBDs locked together), and two others being bound to substrate. In contrast to these detergent-purified structures, there is no wide-open inward-facing state observed for apo-ABCD1 in lipid nanodiscs (Le *et al.*, 2022), which is in line with previous reports regarding the conformational relevance (Wright *et al.*, 2018). Therefore, caution should be exercised when making mechanistic interpretations from widely open inward-facing states. Indeed, such caution is why in this study CTS was investigated for reconstitution into lipid nanodiscs, and later the hydrolysis deficient CTS$^{D606N/E607Q}$ mutant was investigated, that according to the 3D maps obtained, did indeed adopt a more physiologically relevant conformation.

Another outstanding question regarding the ACOT activity of CTS is concerned with the fate of the CoA moiety after substrate cleavage, i.e., does it also enter the peroxisome along with the FFA moiety? Initially, it was unknown whether after substrate cleavage, CoA was transported into the peroxisome along with the FFA or if it was imported into the peroxisome via a different route. However, work by van Roermund *et al.* (2021) has confirmed that in whole yeast peroxisomes, the CoA moiety is transported into the peroxisome after substrate transport and cleavage by

Pxa1p/Pxa2. Given results from cross-kingdom complementation experiments (Carrier *et al.*, 2019; De Marcos Lousa *et al.*, 2013; van Roermund *et al*, 2008; Van Roermund *et al.*, 2014; Van Roermund *et al.*, 2011) and the similarities between other family D transporters, it is likely that the same is true of CTS. Until now however, studies on CTS in isolated *A.thaliana* peroxisomes have been limited due to their fragility and the difficulty in purifying them away from contaminants (Nyathi *et al.*, 2010). Despite the former study overcoming this by heterologous expression of CTS, the authors also highlighted the desire for a reconstitution system applicable for studying CTS. Eleven years later, this thesis has shown that a suitable purification and reconstitution protocol has now been established and that the substrate stimulated activity of CTS has been confirmed after reconstitution (as well the lack of activity for $CTS^{D606N/E607Q}$). Therefore, it is now possible to unequivocally investigate the fate of CoA during the CTS transport cycle using this proteoliposome system, and beyond that, answer many more questions about the biochemistry of CTS to shed new light on previous observations.

For example, the reconstitution system may be used to investigate the following question: why cleave the substrate at all? On the face of it, the cleavage process seems energetically wasteful. However, the amount of ATP generated from the beta-oxidation of just one molecule of palmitate massively outweighs the amount used during the transport cycle. Instead, energy expenditure during the cleavage process may be used as a means of peroxisomal regulation or control. For example, CoA transport into the peroxisome by Pxa1/Pxa2 in yeast (van Roermund *et al.*, 2021) after acyl-CoA cleavage allows for the hydrophobic and hydrophilic moieties of the substrate to be translocated separately. Moreover, in addition to acyl-CoA synthetases LACS6 and LACS7, *A. thaliana* is known to express a large repertoire of other acyl-CoA synthetases (Shockey *et al*, 2002) and acyl-activating enzymes (Shockey & Browse, 2011). Given that CTS makes physical interactions with LACS6 and LACS7 (De Marcos Lousa *et al.*, 2013), it is conceivable that the transporter can

interact with other acyl-CoA synthetases, such that the substrate cleavage products can be prioritised, processed, and metabolised according to the requirements of the plant. Thus, the cleavage process may be viewed as an energy expenditure mechanism as a means of peroxisomal self-regulation. For example, the transport and cleavage of myristoyl-[^{3}H]-CoA into yeast peroxisomes is reduced – but not completely abolished – when the yeast acyl-CoA synthase Faa2p is knocked out (van Roermund *et al.*, 2021), perhaps indicating that without acyl-CoA synthetases, the transporter continues to cleave substrate, causing FFA to accumulate in the membrane, arresting the transporter and upsetting peroxisomal regulation.

Alternatively, Kawaguchi *et al.* (2021) speculates that the cleavage mechanism may simply be a solution to the difficult task of transporting large amphipathic molecules across membranes. Ultimately, this is still an unanswered question, but one that can be investigated for CTS using reconstitution system; for example, it may be possible to investigate the transport of substrates across the proteoliposome membrane where the lumen has been pre-loaded with a specific acyl-CoA synthetase (or synthetases) of interest to investigate specific interactions or kinetic experiments. The same system can be used to investigate the nature of specific mutations, ones that lead to germination and establishment defects in plants (e.g., K487A, & S810N) (Dietrich *et al.*, 2009) or equivalent mutations in ABCD1 that causes the devastating pathology X-ALD (section 1.8).

In summary, over the course of this PhD a series of major developments have been made to the study of the COMATOSE transporter that are distinct from the primarily biochemical work that came before, and which place an emphasis on the study of CTS from a single molecule perspective. At the beginning of this project, structural studies on CTS had never been pursued. Now, and after significant purification optimisation, reported here are the first cryo-EM maps for the plant peroxisomal transporter purified in detergent, showing the transporters overall fold and key features. Whilst the primary aim was a high-resolution structure, this work has shown

that CTS is amenable to cryo-EM studies and has outlined the future experiments needed to obtain such a result. The aforementioned biochemical studies on CTS previously relied upon heterogenous expression of the transporter due to the absence of a suitable reconstitution method. The secondary aim of this project was realised in full, as the proteoliposome reconstitution method reported here – showing that purified and reconstituted CTS possesses substrate stimulated ATPase activity – establishes this system as a tool for future biochemical work on CTS. Together, these developments will be indispensable for future studies on this important and intriguing peroxisomal ABC transporter.

7 References

- Acosta H, Burchmore R, Naula C, Gualdrón-López M, Quintero-Troconis E, Cáceres AJ, Michels PAM, Concepción JL, Quiñones W (2019) Proteomic analysis of glycosomes from Trypanosoma cruzi epimastigotes. *Molecular and Biochemical Parasitology* 229: 62-74
- al-Shawi MK, Urbatsch IL, Senior AE (1994) Covalent inhibitors of P-glycoprotein ATPase activity. *J Biol Chem* 269: 8986-8992
- Alam A, Kowal J, Broude E, Roninson I, Locher KP (2019) Structural insight into substrate and inhibitor discrimination by human P-glycoprotein. *Science* 363: 753-756
- Aleksandrov L, Aleksandrov AA, Chang XB, Riordan JR (2002) The first nucleotide binding domain of cystic fibrosis transmembrane conductance regulator is a site of stable nucleotide interaction, whereas the second is a site of rapid turnover. *Journal of Biological Chemistry* 277: 15419-15425
- Alexandrov AI, Mileni M, Chien EYT, Hanson MA, Stevens RC (2008) Microscale Fluorescent Thermal Stability Assay for Membrane Proteins. *Structure* 16: 351-359
- Aller SG, Yu J, Ward A, Weng Y, Chittaboina S, Zhuo R, Harrell PM, Trinh YT, Zhang Q, Urbatsch IL *et al* (2009) Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science* 323: 1718-1722
- Alvadia C, Lim NK, Clerico Mosina V, Oostergetel GT, Dutzler R, Paulino C (2019) Cryo-EM structures and functional characterization of the murine lipid scramblase TMEM16F. *Elife* 8
- Ambudkar SV (1995) Purification and reconstitution of functional human P-glycoprotein. *J Bioenerg Biomembr* 27: 23-29
- Ambudkar SV, Kim IW, Xia D, Sauna ZE (2006) *FEBS Letters* **580**: 1049-1055
- Ames GF, Lever J (1970) Components of histidine transport: histidine-binding proteins and hisP protein. *Proceedings of the National Academy of Sciences of the United States of America* 66: 1096-1103
- Ames GFL (1986) Bacterial Periplasmic Transport Systems: Structure, Mechanism, and Evolution. *Annual Review of Biochemistry* 55: 397-425
- Armstrong M, Han BG, Gomez S, Turner J, Fletcher DA, Glaeser RM (2020) Microscale Fluid Behavior during Cryo-EM Sample Blotting. *Biophys J* 118: 708-719
- Arnold FM, Weber MS, Gonda I, Gallenito MJ, Adenau S, Egloff P, Zimmermann I, Hutter CAJ, Hürlimann LM, Peters EE *et al* (2020) The ABC exporter IrtAB imports and reduces mycobacterial siderophores. *Nature* 580: 413-417
- Ashok Y, Nanekar R, Jaakola VP (2015) Defining thermostability of membrane proteins by western blotting. *Protein Eng Des Sel* 28: 539-542
- Avisar D, Segal M, Sneh B, Zilberstein A (2005) Cell-cycle-dependent resistance to Bacillus thuringiensis Cry1C toxin in Sf9 cells. *J Cell Sci* 118: 3163-3171
- Baarine M, Andréoletti P, Athias A, Nury T, Zarrouk A, Ragot K, Vejux A, Riedinger JM, Kattan Z, Bessede G *et al* (2012) Evidence of oxidative stress in very long chain fatty acid - Treated oligodendrocytes and potentialization of ROS production using RNA interference-directed knockdown of ABCD1 and ACOX1 peroxisomal proteins. *Neuroscience* 213: 1-18
- Baker A, Carrier DJ, Schaedler T, Waterham HR, Van Roermund CW, Theodoulou FL (2015) Peroxisomal ABC transporters: Functions and mechanism. *Biochemical Society Transactions* 43: 959-965
- Baker A, Graham IA, Holdsworth M, Smith SM, Theodoulou FL (2006) Chewing the fat: β-oxidation in signalling and development. *Trends in Plant Science* **11**: 124-132
- Bammes BE, Rochat RH, Jakana J, Chen DH, Chiu W (2012) Direct electron detection yields cryo-EM reconstructions at resolutions beyond 3/4 Nyquist frequency. *J Struct Biol* 177: 589-601
- Bao H, Dalal K, Wang V, Rouiller I, Duong F (2013) The maltose ABC transporter: Action of membrane lipids on the transporter stability, coupling and ATPase activity. *Biochimica et Biophysica Acta - Biomembranes* 1828: 1723-1730
- Bartesaghi A, Matthies D, Banerjee S, Merk A, Subramaniam S (2014) Structure of beta-galactosidase at 3.2-A resolution obtained by cryo-electron microscopy. *Proc Natl Acad Sci U S A* 111: 11709-11714

- Berge KE, Tian H, Graf GA, Yu L, Grishin NV, Schultz J, Kwiterovich P, Shan B, Barnes R, Hobbs HH (2000) Accumulation of dietary cholesterol in sitosterolemia caused by mutations in adjacent ABC transporters. *Science* 290: 1771-1775
- Berger EA, Heppel LA (1974) Different mechanisms of energy coupling for the shock sensitive and shock resistant amino acid permeases of Escherichia coli. *Journal of Biological Chemistry* 249: 7747-7755
- Berger J, Forss-Petter S, Eichler FS (2014) Pathophysiology of X-linked adrenoleukodystrophy. *Biochimie* 98: 135-142
- Berger J, Molzer B, Faé I, Bernheimer H (1994) X-linked adrenoleukodystrophy (ALD): A novel mutation of the ALD gene in 6 members of a family presenting with 5 different phenotypes. *Biochemical and Biophysical Research Communications* 205: 1638-1643
- Bersch B, Dorr JM, Hessel A, Killian JA, Schanda P (2017) Proton-Detected Solid-State NMR Spectroscopy of a Zinc Diffusion Facilitator Protein in Native Nanodiscs. *Angew Chem Int Ed Engl* 56: 2508-2512
- Biemans-Oldehinkel E, Doeven MK, Poolman B (2006) ABC transporter architecture and regulatory roles of accessory domains. *FEBS Letters* 580: 1023-1035
- Bienengraeber M, Alekseev AE, Abraham MR, Carrasco AJ, Moreau C, Vivaudou M, Dzeja PP, Terzic A (2000) ATPase activity of the sulfonylurea receptor: a catalytic function for the K ATP channel complex. *The FASEB Journal* 14: 1943-1952
- Bin Kanner Y, Ganoth A, Tsfadia Y (2021) Extracellular mutation induces an allosteric effect across the membrane and hampers the activity of MRP1 (ABCC1). *Sci Rep* 11: 12024
- Bolhuis H, van Veen HW, Molenaar D, Poolman B, Driessen AJ, Konings WN (1996) Multidrug resistance in Lactococcus lactis: evidence for ATP-dependent drug extrusion from the inner leaflet of the cytoplasmic membrane. *The EMBO Journal* 15: 4239-4245
- Borshchevskiy V, Kovalev K, Round E, Efremov R, Astashkin R, Bourenkov G, Bratanov D, Balandin T, Chizhov I, Baeken C *et al* (2022) True-atomic-resolution insights into the structure and functional role of linear chains and low-barrier hydrogen bonds in proteins. *Nat Struct Mol Biol*: 440-450
- Bountra K, Hagelueken G, Choudhury HG, Corradi V, El Omari K, Wagner A, Mathavan I, Zirah S, Yuan Wahlgren W, Tieleman DP *et al* (2017) Structural basis for antibacterial peptide self-immunity by the bacterial ABC transporter McjD. *The EMBO Journal* 36: 3062-3079
- Braun A, Ambach H, Kammerer S, Rolinski B, Stockler S, Rabl W, Gartner J, Zierz S, Roscher AA (1995) Mutations in the gene for X-linked adrenoleukodystrophy in patients with different clinical phenotypes. *American Journal of Human Genetics* 56: 854-861
- Broecker J, Eger BT, Ernst OP (2017) Crystallogenesis of Membrane Proteins Mediated by Polymer-Bounded Lipid Nanodiscs. *Structure* 25: 384-392
- Cabra V, Samso M (2015) Do's and don'ts of cryo-electron microscopy: a primer on sample preparation and high quality data collection for macromolecular 3D reconstruction. *J Vis Exp* 95: 52311
- Caffalette CA, Zimmer J (2021) Cryo-EM structure of the full-length WzmWzt ABC transporter required for lipid-linked O antigen transport. *Proceedings of the National Academy of Sciences* 118: e2016144118
- Cai J, Gros P (2003) Overexpression, purification, and functional characterization of ATP-binding cassette transporters in the yeast, Pichia pastoris. *Biochim Biophys Acta* 1610: 63-76
- Callaghan R, Berridge G, Ferry DR, Higgins CF (1997) The functional purification of P-glycoprotein is dependent on maintenance of a lipid-protein interface. *Biochimica et Biophysica Acta - Biomembranes* 1328: 109-124
- Carrier DJ, van Roermund CWT, Schaedler TA, Rong HL, Ijlst L, Wanders RJA, Baldwin SA, Waterham HR, Theodoulou FL, Baker A (2019) Mutagenesis separates ATPase and thioesterase activities of the peroxisomal ABC transporter, Comatose. *Scientific Reports* 9: 10502
- Chang G, Roth CB (2001) Structure of MsbA from E. coli: A homolog of the multidrug resistance ATP binding cassette (ABC) transporters. *Science* 293: 1793-1800

- Chaptal V, Delolme F, Kilburg A, Magnard S, Montigny C, Picard M, Prier C, Monticelli L, Bornert O, Agez M *et al* (2017) Quantification of Detergents Complexed with Membrane Proteins. *Sci Rep* 7: 41751
- Chavan H, Khan MM, Tegos G, Krishnamurthy P (2013) Efficient purification and reconstitution of ATP binding cassette transporter B6 (ABCB6) for functional and structural studies. *J Biol Chem* 288: 22658-22669
- Chen L, Hou WT, Fan T, Liu B, Pan T, Li YH, Jiang YL, Wen W, Chen ZP, Sun L *et al* (2020) Cryo-electron Microscopy Structure and Transport Mechanism of a Wall Teichoic Acid ABC Transporter. *mBio* 11: e02749-02719
- Chen Z-P, Xu D, Wang L, Zhou C-Z, Hou W-T, Chen Y (2021) Structural basis of substrate recognition and translocation by human ABCD1. *bioRxiv*: 2021.2009.2024.461565
- Chen ZP, Xu D, Wang L, Mao YX, Li Y, Cheng MT, Zhou CZ, Hou WT, Chen Y (2022) Structural basis of substrate recognition and translocation by human very long-chain fatty acid transporter ABCD1. *Nat Commun* 13: 3299
- Chen ZS, Tiwari AK (2011) Multidrug resistance proteins (MRPs/ABCCs) in cancer chemotherapy and genetic diseases. *FEBS Journal* 278: 3226-3245
- Chifflet S, Torriglia A, Chiesa R, Tolosa S (1988) A method for the determination of inorganic phosphate in the presence of labile organic phosphate and high concentrations of protein: application to lens ATPases. *Anal Biochem* 168: 1-4
- Choudhury HG, Tong Z, Mathavan I, Li Y, Iwata S, Zirah S, Rebuffat S, Van Veen HW, Beis K (2014) Structure of an antibacterial peptide ATP-binding cassette transporter in a novel outward occluded state. *Proceedings of the National Academy of Sciences of the United States of America* 111: 9145-9150
- Chuang CY, Chen LY, Fu RH, Chen SM, Ho MH, Huang JM, Hsu CC, Wang CC, Chen MS, Tsai RT (2014) Involvement of the carboxyl-terminal region of the yeast peroxisomal half ABC transporter Pxa2p in its interaction with Pxa1p and in transporter function. *PLoS ONE* 9: e104892
- Chung KY, Kim TH, Manglik A, Alvares R, Kobilka BK, Prosser RS (2012) Role of detergents in conformational exchange of a G protein-coupled receptor. *J Biol Chem* 287: 36305-36311
- Coelho D, Kim JC, Miousse IR, Fung S, Du Moulin M, Buers I, Suormala T, Burda P, Frapolli M, Stucki M *et al* (2012) Mutations in ABCD4 cause a new inborn error of vitamin B12 metabolism. *Nature Genetics* 44: 1152-1155
- Cole SPC (2014) Multidrug resistance protein 1 (mrp1, abcc1), a "multitasking" atp-binding cassette (abc,) transporter. *Journal of Biological Chemistry* 289: 30880-30888
- Cole SPC, Deeley RG (2006) Transport of glutathione and glutathione conjugates by MRP1. *Trends in Pharmacological Sciences* 27: 438-446
- Cong Z, Chen LN, Ma H, Zhou Q, Zou X, Ye C, Dai A, Liu Q, Huang W, Sun X *et al* (2021) Molecular insights into ago-allosteric modulation of the human glucagon-like peptide-1 receptor. *Nat Commun* 12: 3763
- Cross LL, Ebeed HT, Baker A (2016) Peroxisome biogenesis, protein targeting mechanisms and PEX gene functions in plants. *Biochim Biophys Acta* 1863: 850-862
- Cui J, Davidson AL (2011) ABC solute importers in bacteria. *Essays in Biochemistry* 50: 85-99
- Cuthbertson L, Kos V, Whitfield C (2010) ABC Transporters Involved in Export of Cell Surface Glycoconjugates. *Microbiology and Molecular Biology Reviews* 74: 341-362
- D'Imprima E, Floris D, Joppe M, Sanchez R, Grininger M, Kuhlbrandt W (2019) Protein denaturation at the air-water interface and how to prevent it. *Elife* 8
- Dalmas O, Orelle C, Foucher AE, Geourjon C, Crouzy S, Di Pietro A, Jault JM (2005) The Q-loop disengages from the first intracellular loop during the catalytic cycle of the multidrug ABC transporter BmrA. *Journal of Biological Chemistry* 280: 36857-36864
- Dave A, Hernández ML, He Z, Andriotis VME, Vaistij FE, Larson TR, Graham IA (2011) 12-Oxo-phytodienoic acid accumulation during seed development represses seed germination in Arabidopsis. *Plant Cell*: 583-599
- Dawson RJP, Locher KP (2006) Structure of a bacterial multidrug ABC transporter. *Nature* 443: 180-185
- Dawson RJP, Locher KP (2007) Structure of the multidrug ABC transporter Sav1866 from Staphylococcus aureus in complex with AMP-PNP. *FEBS Letters* 581: 935-938

- De Marcos Lousa C, Van Roermund CWT, Postis VLG, Dietrich D, Kerr ID, Wanders RJA, Baldwin SA, Baker A, Theodoulou FL (2013) Intrinsic acyl-CoA thioesterase activity of a peroxisomal ATP binding cassette transporter is required for transport and metabolism of fatty acids. *Proceedings of the National Academy of Sciences of the United States of America* 110: 1279-1284
- Dean M, Rzhetsky A, Allikmets R, 2001. The human ATP-binding cassette (ABC) transporter superfamily, Genome Research. pp. 1156-1166.
- Denisov IG, Schuler MA, Sligar SG (2019) Nanodiscs as a New Tool to Examine Lipid-Protein Interactions. *Methods Mol Biol* 2003: 645-671
- Diederichs KA, Ni X, Rollauer SE, Botos I, Tan X, King MS, Kunji ERS, Jiang J, Buchanan SK (2020) Structural insight into mitochondrial beta-barrel outer membrane protein biogenesis. *Nat Commun* 11: 3290
- Dietrich D, Schmuths H, De Lousa CM, Baldwin JM, Baldwin SA, Baker A, Theodoulou FL, Holdsworth MJ (2009) Mutations in the arabidopsis peroxisomal ABC transporter COMATOSE allow differentiation between multiple functions in planta: Insights from an allelic series. *Molecular Biology of the Cell* 20: 530-543
- Do T, Martinoia E, Lee Y, Hwang J (2021) 2021 update on ATP-binding cassette (ABC) transporters: how they meet the needs of plants. *Plant Physiology* 187: 1876-1892
- Do THT, Choi H, Palmgren M, Martinoia E, Hwang JU, Lee Y (2019) Arabidopsis ABCG28 is required for the apical accumulation of reactive oxygen species in growing pollen tubes. *Proceedings of the National Academy of Sciences of the United States of America* 116: 12540-12549
- Drulyte I, Johnson RM, Hesketh EL, Hurdiss DL, Scarff CA, Porav SA, Ranson NA, Muench SP, Thompson RF (2018) Approaches to altering particle distributions in cryo-electron microscopy sample preparation. *Acta Crystallogr D Struct Biol* 74: 560-571
- Edavettal SC, Hunter MJ, Swanson RV (2012) Genetic construct design and recombinant protein expression for structural biology. *Methods Mol Biol* 841: 29-47
- Eichler FS, Ren JQ, Cossoy M, Rietsch AM, Nagpal S, Moser AB, Frosch MP, Ransohoff RM (2008) Is microglial apoptosis an early pathogenic change in cerebral X-linked adrenoleukodystrophy? *Annals of Neurology* 63: 729-742
- Ellinger P, Kluth M, Stindt J, Smits SH, Schmitt L (2013) Detergent screening and purification of the human liver ABC transporters BSEP (ABCB11) and MDR3 (ABCB4) expressed in the yeast Pichia pastoris. *PLoS One* 8: e60620
- Emmanouilidis L, Schütz U, Tripsianes K, Madl T, Radke J, Rucktäschel R, Wilmanns M, Schliebs W, Erdmann R, Sattler M (2017) Allosteric modulation of peroxisomal membrane protein recognition by farnesylation of the peroxisomal import receptor PEX19. *Nature Communications* 8: 14635
- Engelen M, Kemp S, De Visser M, Van Geel BM, Wanders RJA, Aubourg P, Poll-The BT (2012) X-linked adrenoleukodystrophy (X-ALD): Clinical presentation and guidelines for diagnosis, follow-up and management. *Orphanet Journal of Rare Diseases* 7: 51
- Esser L, Zhou F, Pluchino KM, Shiloach J, Ma J, Tang WK, Gutierrez C, Zhang A, Shukla S, Madigan JP *et al* (2017) Structures of the Multidrug Transporter P-glycoprotein Reveal Asymmetric ATP Binding and the Mechanism of Polyspecificity. *Journal of Biological Chemistry* 292: 446-461
- Fabre L, Bao H, Innes J, Duong F, Rouiller I (2017) Negative Stain Single-particle EM of the Maltose Transporter in Nanodiscs Reveals Asymmetric Closure of MalK2 and Catalytic Roles of ATP, MalE, and Maltose. *J Biol Chem* 292: 5457-5464
- Fan X, Wang J, Zhang X, Yang Z, Zhang JC, Zhao L, Peng HL, Lei J, Wang HW (2019) Single particle cryo-EM reconstruction of 52 kDa streptavidin at 3.2 Angstrom resolution. *Nat Commun* 10: 2386
- Fang Y, Morrell JC, Jones JM, Gould SJ (2004) PEX3 functions as a PEX19 docking factor in the import of class I peroxisomal membrane proteins. *J Cell Biol* 164: 863-875
- Ferrer-Miralles N, Saccardo P, Corchero JL, Xu Z, Garcia-Fruitos E (2015) General introduction: recombinant protein production and purification of insoluble proteins. *Methods Mol Biol* 1258: 1-24

- Footitt S, Dietrich D, Fait A, Fernie AR, Holdsworth MJ, Baker A, Theodoulou FL (2007) The COMATOSE ATP-binding cassette transporter is required for full fertility in Arabidopsis. *Plant Physiology* 144: 1467-1480
- Footitt S, Slocombe SP, Larner V, Kurup S, Wu Y, Larson T, Graham I, Baker A, Holdsworth M (2002) Control of germination and lipid mobilization by COMATOSE, the Arabidopsis homologue of human ALDP. *EMBO Journal* 21: 2912-2922
- Ford RC, Hellmich UA (2020) What monomeric nucleotide binding domains can teach us about dimeric ABC proteins. *FEBS Lett* 594: 3857-3875
- Fourcade S, López-Erauskin J, Galino J, Duval C, Naudi A, Jove M, Kemp S, Villarroya F, Ferrer I, Pamplona R *et al* (2008) Early oxidative damage underlying neurodegeneration in X-adrenoleukodystrophy. *Human Molecular Genetics* 17: 1762-1773
- Fulda M, Schnurr J, Abbadi A, Heinz E, Browse J (2004) Peroxisomal Acyl-CoA Synthetase Activity Is Essential for Seedling Development in Arabidopsis thaliana. *Plant Cell* 16: 393-405
- Fulyani F, Schuurman-Wolters GK, Slotboom DJ, Poolman B (2016) Relative rates of amino acid import via the ABC transporter GlnPQ determine the growth performance of Lactococcus lactis. *Journal of Bacteriology* 198: 477-485
- Gadsby DC, Vergani P, Csanády L (2006) The ABC protein turned chloride channel whose failure causes cystic fibrosis. *Nature* 440: 477-483
- Galino J, Ruiz M, Fourcade S, Schlüter A, López-Erauskin J, Guilera C, Jove M, Naudi A, García-Arumí E, Andreu AL *et al* (2011) Oxidative damage compromises energy metabolism in the axonal degeneration mouse model of X-adrenoleukodystrophy. *Antioxidants and Redox Signaling* 15: 2095-2107
- Gao M, Cui HR, Loe DW, Grant CE, Almquist KC, Cole SP, Deeley RG (2000) Comparison of the functional characteristics of the nucleotide binding domains of multidrug resistance protein 1. *J Biol Chem* 275: 13098-13108
- Geillon F, Gondcaille C, Charbonnier S, Van Roermund CW, Lopez TE, Dias AMM, De Barros JPP, Arnould C, Wanders RJ, Trompier D *et al* (2014) Structure-function analysis of peroxisomal ATP-binding cassette transporters using chimeric dimers. *Journal of Biological Chemistry* 289: 24511-24520
- Geillon F, Gondcaille C, Raas Q, Dias AMM, Pecqueur D, Truntzer C, Lucchi G, Ducoroy P, Falson P, Savary S *et al* (2017) Peroxisomal ATP-binding cassette transporters form mainly tetramers. *J Biol Chem* 292: 6965-6977
- Geourjon C, Orelle C, Steinfels E, Blanchet C, Deleage G, Di Pietro A, Jault JM (2001) A common mechanism for ATP hydrolysis in ABC transporter and helicase superfamilies. *Trends Biochem Sci* 26: 539-544
- Gewering T, Januliene D, Ries AB, Moeller A (2018) Know your detergents: A case study on detergent background in negative stain electron microscopy. *J Struct Biol* 203: 242-246
- Girzalsky W, Saffian D, Erdmann R (2010) Peroxisomal protein translocation. *Biochim Biophys Acta* 1803: 724-731
- Glaeser RM, Han BG (2017) Opinion: hazards faced by macromolecules when confined to thin aqueous films. *Biophys Rep* 3: 1-7
- Gopalasingam CC, Johnson RM, Chiduza GN, Tosha T, Yamamoto M, Shiro Y, Antonyuk SV, Muench SP, Hasnain SS (2019) Dimeric structures of quinol-dependent nitric oxide reductases (qNORs) revealed by cryo-electron microscopy. *Sci Adv* 5: eaax1803
- Gotte K, Girzalsky W, Linkert M, Baumgart E, Kammerer S, Kunau WH, Erdmann R (1998) Pex19p, a farnesylated protein essential for peroxisome biogenesis. *Mol Cell Biol* 18: 616-628
- Gros P, Croop J, Housman D (1986) Mammalian multidrug resistance gene: Complete cDNA sequence indicates strong homology to bacterial transport proteins. *Cell* 47: 371-380
- Grossmann N, Vakkasoglu AS, Hulpke S, Abele R, Gaudet R, Tampe R (2014) Mechanistic determinants of the directionality and energetics of active export by a heterodimeric ABC transporter. *Nat Commun* 5: 5419
- Guimarães CP, Domingues P, Aubourg P, Fouquet F, Pujol A, Jimenez-Sanchez G, Sá-Miranda C, Azevedo JE (2004) Mouse liver PMP70 and ALDP: Homomeric interactions prevail in vivo. *Biochimica et Biophysica Acta - Molecular Basis of Disease* 1689: 235-243

- Gulati S, Jamshad M, Knowles TJ, Morrison KA, Downing R, Cant N, Collins R, Koenderink JB, Ford RC, Overduin M *et al* (2014) Detergent-free purification of ABC (ATP-binding-cassette) transporters. *Biochem J* 461: 269-278
- Guo Y (2020) Be Cautious with Crystal Structures of Membrane Proteins or Complexes Prepared in Detergents. *Crystals (Basel)* 10: 86
- Harborne SPD, Strauss J, Boakes JC, Wright DL, Henderson JG, Boivineau J, Jaakola VP, Goldman A (2020) IMPROvER: the Integral Membrane Protein Stability Selector. *Sci Rep* 10: 15165
- Hardy D, Bill RM, Rothnie AJ, Jawhari A (2019) Stabilization of Human Multidrug Resistance Protein 4 (MRP4/ABCC4) Using Novel Solubilization Agents. *SLAS Discov* 24: 1009-1017
- Hauer F, Gerle C, Fischer N, Oshima A, Shinzawa-Itoh K, Shimada S, Yokoyama K, Fujiyoshi Y, Stark H (2015) GraDeR: Membrane Protein Complex Preparation for Single-Particle Cryo-EM. *Structure* 23: 1769-1775
- Hayashi M, Nito K, Takei-Hoshi R, Yagi M, Kondo M, Suenaga A, Yamaya T, Nishimura M (2002) Ped3p is a peroxisomal ATP-binding cassette transporter that might supply substrates for fatty acid β-oxidation. *Plant and Cell Physiology* 43: 1-11
- Hazelbauer GL (1975) Maltose chemoreceptor of Escherichia coli. *Journal of Bacteriology* 122: 206-214
- He SM, Li R, R. Kanwar J, Zhou SF (2012) Structural and Functional Properties of Human Multidrug Resistance Protein 1 (MRP1/ABCC1). *Current Medicinal Chemistry* 18: 439-481
- Hein S, Schönfeld P, Kahlert S, Reiser G (2008) Toxic effects of X-linked adrenoleukodystrophy-associated, very long chain fatty acids on glial cells and neurons from rat hippocampus in culture. *Human Molecular Genetics* 17: 1750-1761
- Her C, Filoti DI, McLean MA, Sligar SG, Alexander Ross JB, Steele H, Laue TM (2016) The Charge Properties of Phospholipid Nanodiscs. *Biophys J* 111: 989-998
- Hettema EH, van Roermund CW, Distel B, van den Berg M, Vilela C, Rodrigues-Pousada C, Wanders RJ, Tabak HF (1996) The ABC transporter proteins Pat1 and Pat2 are required for import of long-chain fatty acids into peroxisomes of Saccharomyces cerevisiae. *The EMBO Journal* 15: 3813-3822
- Higgins CF, Gallagher MP, Hyde SC, Mimmack ML, Pearce SR, 1990. Periplasmic binding protein-dependent transport systems: the membrane-associated components, Philosophical transactions of the Royal Society of London Series B: Biological sciences. pp. 353-364; discussion 364.
- Higgins CF, Gottesman MM (1992) Is the multidrug transporter a flippase? *Trends in Biochemical Sciences* 17: 18-21
- Higgins CF, Hiles ID, Salmond GP, Gill DR, Downie JA, Evans IJ, Holland IB, Gray L, Buckel SD, Bell AW *et al* (1986) A family of related ATP-binding subunits coupled to many distinct biological processes in bacteria. *Nature* 323: 448-450
- Higgins CF, Linton KJ (2004) The ATP switch model for ABC transporters. *Nature Structural and Molecular Biology* 11: 918-926
- Hillebrand M, Verrier SE, Ohlenbusch A, Schäfer A, Söling HD, Wouters FS, Gärtner J (2007) Live cell FRET microscopy: Homo- and heterodimerization of two human peroxisomal ABC transporters, the adrenoleukodystrophy protein (ALDP, ABCD1) and PMP70 (ABCD3). *Journal of Biological Chemistry* 282: 26997-27005
- Ho JK, Moser H, Kishimoto Y, Hamilton JA (1995) Interactions of a very long chain fatty acid with model membranes and serum albumin: Implications for the pathogenesis of adrenoleukodystrophy. *Journal of Clinical Investigation* 96: 1455-1463
- Hofmann S, Januliene D, Mehdipour AR, Thomas C, Stefan E, Bruchert S, Kuhn BT, Geertsma ER, Hummer G, Tampe R *et al* (2019) Conformation space of a heterodimeric ABC exporter under turnover conditions. *Nature* 571: 580-583
- Hohl M, Briand C, Grütter MG, Seeger MA (2012) Crystal structure of a heterodimeric ABC transporter in its inward-facing conformation. *Nature Structural and Molecular Biology* 19: 395-402
- Hohl M, Hürlimann LM, Bo̎hm S, Scho̎ppe J, Grütter MG, Bordignon E, Seeger MA (2014) Structural basis for allosteric cross-talk between the asymmetric nucleotide binding sites of

a heterodimeric ABC exporter. *Proceedings of the National Academy of Sciences of the United States of America* 111: 11025-11030
- Hollenstein K, Frei DC, Locher KP (2007) Structure of an ABC transporter in complex with its binding protein. *Nature* 446: 213-216
- Hooks MA, Turner JE, Murphy EC, Johnston KA, Burr S, Jarosławski S (2007) The Arabidopsis ALDP protein homologue COMATOSE is instrumental in peroxisomal acetate metabolism. *Biochemical Journal* 406: 399-406
- Hopfner KP (2016) Invited review: Architectures and mechanisms of ATP binding cassette proteins. *Biopolymers* 105: 492-504
- Hua R, Gidda SK, Aranovich A, Mullen RT, Kim PK (2015) Multiple Domains in PEX16 Mediate Its Trafficking and Recruitment of Peroxisomal Proteins to the ER. *Traffic* 16: 832-852
- Hung LW, Wang IX, Nikaido K, Liu PQ, Ames GFL, Kim SH (1998) Crystal structure of the ATP-binding subunit of an ABC transporter. *Nature* 396: 703–707
- Hyde SC, Emsley P, Hartshorn MJ, Mimmack MM, Gileadi U, Pearce SR, Gallagher MP, Gill DR, Hubbard RE, Higgins CF (1990) Structural model of ATP-binding proteing associated with cystic fibrosis, multidrug resistance and bacterial transport. *Nature* 346: 362-365
- Igoillo-Esteve M, Mazet M, Deumer G, Wallemacq P, Michels PAM (2011) Glycosomal ABC transporters of Trypanosoma brucei: Characterisation of their expression, topology and substrate specificity. *International Journal for Parasitology* 41: 429-438
- Infed N, Hanekop N, Driessen AJ, Smits SH, Schmitt L (2011) Influence of detergents on the activity of the ABC transporter LmrA. *Biochim Biophys Acta* 1808: 2313-2321
- Islinger M, Voelkl A, Fahimi HD, Schrader M (2018) The peroxisome: an update on mysteries 2.0. *Histochemistry and Cell Biology* 150: 443-471
- Jackson SM, Manolaridis I, Kowal J, Zechner M, Taylor NMI, Bause M, Bauer S, Bartholomaeus R, Bernhardt G, Koenig B *et al* (2018) Structural basis of small-molecule inhibition of human multidrug transporter ABCG2. *Nature Structural and Molecular Biology* 25: 333-340
- Janas E, Hofacker M, Chen M, Gompf S, van der Does C, Tampe R (2003) The ATP hydrolysis cycle of the nucleotide-binding domain of the mitochondrial ATP-binding cassette transporter Mdl1p. *J Biol Chem* 278: 26862-26869
- Januliene D, Moeller A (2020) Cryo-EM of ABC transporters: an ice-cold solution to everything? *FEBS Lett* 594: 3776-3789
- Jardetzky O (1966) Simple allosteric model for membrane pumps *Nature* 211: 969-970
- Jia Y, Zhang Y, Lei J, Yang G (2021) Structure insights of the human peroxisomal ABC transporter ALDP. *bioRxiv*: 2021.2009.2024.461756
- Jin L, Milazzo AC, Kleinfelder S, Li S, Leblanc P, Duttweiler F, Bouwer JC, Peltier ST, Ellisman MH, Xuong NH (2008) Applications of direct detection device in transmission electron microscopy. *J Struct Biol* 161: 352-358
- Jin MS, Oldham ML, Zhang Q, Chen J (2012) Crystal structure of the multidrug transporter P-glycoprotein from Caenorhabditis elegans. *Nature* 490: 566-569
- Johnson ZL, Chen J (2017) Structural Basis of Substrate Recognition by the Multidrug Resistance Protein MRP1. *Cell* 168: 1075-1085.e1079
- Jones PM, George AM (2009) Opening of the ADP-bound active site in the ABC transporter ATPase dimer: Evidence for a constant contact, alternating sites model for the catalytic cycle. *Proteins: Structure, Function and Bioinformatics* 75: 387-396
- Jones PM, George AM, 2012. Role of the D-loops in allosteric control of ATP hydrolysis in an ABC transporter, Journal of Physical Chemistry A, 11 ed., pp. 3004-3013.
- Josts I, Nitsche J, Maric S, Mertens HD, Moulin M, Haertlein M, Prevost S, Svergun DI, Busch S, Forsyth VT *et al* (2018) Conformational States of ABC Transporter MsbA in a Lipid Environment Investigated by Small-Angle Scattering Using Stealth Carrier Nanodiscs. *Structure* 26: 1072-1079.e1074
- Juliano RL, Ling V (1976) A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants. *BBA - Biomembranes* 455: 152-162

- Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Zidek A, Potapenko A *et al* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* 596: 583-589
- Kadaba NS, Kaiser JT, Johnson E, Lee A, Rees DC (2008) The high-affinity E. coli methionine ABC transporter: Structure and allosteric regulation. *Science* 321: 250-253
- Kalipatnapu S, Chattopadhyay A (2005) Membrane protein solubilization: recent advances and challenges in solubilization of serotonin1A receptors. *IUBMB Life* 57: 505-512
- Kampjut D, Steiner J, Sazanov LA (2021) Cryo-EM grid optimization for membrane proteins. *iScience* 24: 102139
- Kao YT, Gonzalez KL, Bartel B (2018) Peroxisome function, biogenesis, and dynamics in plants. *Plant Physiology* 176: 162–177
- Kawaguchi K, Morita M (2016) ABC Transporter Subfamily D: Distinct Differences in Behavior between ABCD1-3 and ABCD4 in Subcellular Localization, Function, and Human Disease. *Biomed Res Int* 2016: 6786245
- Kawaguchi K, Mukai E, Watanabe S, Yamashita A, Morita M, So T, Imanaka T (2021) Acyl-CoA thioesterase activity of peroxisomal ABC protein ABCD1 is required for the transport of very long-chain acyl-CoA into peroxisomes. *Sci Rep* 11: 2192
- Kawai T, Caaveiro JMM, Abe R, Katagiri T, Tsumoto K (2011) Catalytic activity of MsbA reconstituted in nanodisc particles is modulated by remote interactions with the bilayer. *FEBS Letters* 585: 3533-3537
- Kawate T, Gouaux E (2006) Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins. *Structure* 14: 673-681
- Kay C, Woodward KD, Lawler K, Self TJ, Dyall SD, Kerr ID (2012) The ATP-binding cassette proteins of the deep-branching protozoan parasite Trichomonas vaginalis. *PLoS Negl Trop Dis* 6: e1693
- Keable SM, Kolsch A, Simon PS, Dasgupta M, Chatterjee R, Subramanian SK, Hussein R, Ibrahim M, Kim IS, Bogacz I *et al* (2021) Room temperature XFEL crystallography reveals asymmetry in the vicinity of the two phylloquinones in photosystem I. *Sci Rep* 11: 21787
- Kemp S, Berger J, Aubourg P (2012) X-linked adrenoleukodystrophy: Clinical, metabolic, genetic and pathophysiological aspects. *Biochimica et Biophysica Acta - Molecular Basis of Disease* 1822: 1465-1474
- Kemp S, Pujol A, Waterham HR, Van Geel BM, Boehm CD, Raymond GV, Cutting GR, Wanders RJA, Moser HW, 2001. ABCD1 mutations and the X-linked adrenoleukodystrophy mutation database: Role in diagnosis and clinical correlations, Human Mutation. pp. 499-515.
- Kerem BS, Rommens JM, Buchanan JA, Markiewicz D, Cox TK, Chakravarti A, Buchwald M, Tsui LC (1989) Identification of the cystic fibrosis gene: Genetic analysis. *Science* 245: 1073-1080
- Kerr ID, Haider AJ, Gelissen IC (2011) The ABCG family of membrane-associated transporters: You don't have to be big to be mighty. *British Journal of Pharmacology* 164: 1767-1779
- Kesidis A, Depping P, Lode A, Vaitsopoulou A, Bill RM, Goddard AD, Rothnie AJ (2020) Expression of eukaryotic membrane proteins in eukaryotic and prokaryotic hosts. *Methods* 180: 3-18
- Khoshouei M, Radjainia M, Baumeister W, Danev R (2017) Cryo-EM structure of haemoglobin at 3.2 A determined with the Volta phase plate. *Nat Commun* 8: 16099
- Khoshouei M, Radjainia M, Phillips AJ, Gerrard JA, Mitra AK, Plitzko JM, Baumeister W, Danev R (2016) Volta phase plate cryo-EM of the small protein complex Prx3. *Nat Commun* 7: 10534
- Khunweeraphong N, Stockner T, Kuchler K (2017) The structure of the human ABC transporter ABCG2 reveals a novel mechanism for drug extrusion. *Sci Rep* 7: 13767
- Kim PK, Mullen RT (2013) PEX16: a multifaceted regulator of peroxisome biogenesis. *Front Physiol* 4: 241
- Kim Y, Chen J (2018) Molecular structure of human P-glycoprotein in the ATP-bound, outward-facing conformation. *Science* 359: 915-919

- Kitai K, Kawaguchi K, Tomohiro T, Morita M, So T, Imanaka T (2021) The lysosomal protein ABCD4 can transport vitamin B12 across liposomal membranes in vitro. *J Biol Chem* 296: 100654
- Klebl DP, Gravett MSC, Kontziampasis D, Wright DJ, Bon RS, Monteiro DCF, Trebbin M, Sobott F, White HD, Darrow MC *et al* (2020) Need for Speed: Examining Protein Behavior during CryoEM Grid Preparation at Different Timescales. *Structure* 28: 1238-1248 e1234
- Knauer MJ, Urquhart BL, Meyer Zu Schwabedissen HE, Schwarz UI, Lemke CJ, Leake BF, Kim RB, Tirona RG (2010) Human skeletal muscle drug transporters determine local exposure and toxicity of statins. *Circulation Research* 106: 297-306
- Knazek RA, Rizzo WB, Schulman JD, Dave JR (1983) Membrane microviscosity is increased in the erythrocytes of patients with adrenoleukodystrophy and adrenomyeloneuropathy. *Journal of Clinical Investigation* 72: 245-248
- Korenke GC, Wilichowski E, Hunneman DH, Hanefeld F, Fuchs S, Krasemann E, Doerr HG (1996) Cerebral adrenoleukodystrophy (ALD) in only one of monozygotic twins with an identical ALD genotype. *Annals of Neurology* 40: 254-257
- Kruska N, Schönfeld P, Pujol A, Reiser G (2015) Astrocytes and mitochondria from adrenoleukodystrophy protein (ABCD1)-deficient mice reveal that the adrenoleukodystrophy-associated very long-chain fatty acids target several cellular energy-dependent functions. *Biochimica et Biophysica Acta - Molecular Basis of Disease* 1852: 925-936
- Kuhlbrandt W (2014) Cryo-EM enters a new era. *Elife* 3: e03678
- Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, Fitzhugh W *et al* (2001) Initial sequencing and analysis of the human genome. *Nature* 409: 860-921
- Le LTM, Thompson JR, Dang PX, Bhandari J, Alam A (2021) Structures of the human peroxisomal fatty acid transporter ABCD1 in a lipid environment. *bioRxiv*: 2021.2009.2004.458904
- Le LTM, Thompson JR, Dang PX, Bhandari J, Alam A (2022) Structures of the human peroxisomal fatty acid transporter ABCD1 in a lipid environment. *Commun Biol* 5: 7
- Lee A, Asahina K, Okamoto T, Kawaguchi K, Kostsin DG, Kashiwayama Y, Takanashi K, Yazaki K, Imanaka T, Morita M (2014) Role of NH2-terminal hydrophobic motif in the subcellular localization of ATP-binding cassette protein subfamily D: Common features in eukaryotic organisms. *Biochemical and Biophysical Research Communications* 453: 612-618
- Lee JY, Kinch LN, Borek DM, Wang J, Wang J, Urbatsch IL, Xie XS, Grishin NV, Cohen JC, Otwinowski Z *et al* (2016a) Crystal structure of the human sterol transporter ABCG5/ABCG8. *Nature* 533: 561-564
- Lee SC, Bennett BC, Hong WX, Fu Y, Baker KA, Marcoux J, Robinson CV, Ward AB, Halpert JR, Stevens RC *et al* (2013) Steroid-based facial amphiphiles for stabilization and crystallization of membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America* 110: E1203-E1211
- Lee SC, Knowles TJ, Postis VLG, Jamshad M, Parslow RA, Lin Y-p, Goldman A, Sridhar P, Overduin M, Muench SP *et al* (2016b) A method for detergent-free isolation of membrane proteins in their local lipid environment. *Nature Protocols* 11: 1149-1162
- Lee SH, Lee MS, Lee JH, Kim SW, Kang RH, Choi MJ, Park SJ, Kim SJ, Lee JM, Cole SPC *et al* (2010) MRP1 polymorphisms associated with citalopram response in patients with major depression. *Journal of Clinical Psychopharmacology* 30: 116-125
- Lehnert E, Tampé R (2017) Structure and dynamics of antigenic peptides in complex with TAP. *Frontiers in Immunology* 8: 10
- Lenoir G, Dieudonne T, Lamy A, Lejeune M, Vazquez-Ibar JL, Montigny C (2018) Screening of Detergents for Stabilization of Functional Membrane Proteins. *Curr Protoc Protein Sci* 93: e59
- Lewis VG, Ween MP, McDevitt CA (2012) The role of ATP-binding cassette transporters in bacterial pathogenicity. *Protoplasma* 249: 919-942
- Li D, Caffrey M (2020) Structure and Functional Characterization of Membrane Integral Proteins in the Lipid Cubic Phase. *J Mol Biol* 432: 5104-5123

- Linka N, Theodoulou FL, Haslam RP, Linka M, Napier JA, Neuhaus HE, Webera APM (2008) Peroxisomal ATP import is essential for seedling development in Arabidopsis thaliana. *Plant Cell* 20: 3241-3257
- Liu F, Lee J, Chen J (2021) Molecular structures of the eukaryotic retinal importer ABCA4. *Elife* 10
- Liu F, Zhang Z, Csanády L, Gadsby DC, Chen J (2017) Molecular Structure of the Human CFTR Ion Channel. *Cell* 169: 85-95.e88
- Liu H, Lee W (2019) The XFEL Protein Crystallography: Developments and Perspectives. *Int J Mol Sci* 20
- Liu LX, Janvier K, Berteaux-Lecellier V, Cartier N, Benarous R, Aubourg P (1999) Homo- and heterodimerization of peroxisomal ATP-binding cassette half- transporters. *Journal of Biological Chemistry* 274: 32738-32743
- Locher KP (2016) Mechanistic diversity in ATP-binding cassette (ABC) transporters. *Nature Structural and Molecular Biology* 23: 487-493
- Locher KP, Lee AT, Rees DC (2002) The E. coli BtuCD structure: A framework for ABC transporter architecture and mechanism. *Science* 296: 1091-1098
- Loo TW, Bartlett MC, Clarke DM (2003) Drug binding in human P-glycoprotein causes conformational changes in both nucleotide-binding domains. *Journal of Biological Chemistry* 278: 1575-1578
- Loo TW, Clarke DM (1995) Covalent modification of human P-glycoprotein mutants containing a single cysteine in either nucleotide-binding fold abolishes drug-stimulated ATPase activity. *J Biol Chem* 270: 22957-22961
- Lusvarghi S, Durell SR, Ambudkar SV (2021) Does the ATP-bound EQ mutant reflect the pre- or post-ATP hydrolysis state in the catalytic cycle of human P-glycoprotein (ABCB1)? *FEBS Lett* 595: 750-762
- Manciu L, Chang XB, Buyse F, Hou YX, Gustot A, Riordan JR, Ruysschaert JM (2003) Intermediate structural states involved in MRP1-mediated drug transport. Role of glutathione. *J Biol Chem* 278: 3347-3356
- Manolaridis I, Jackson SM, Taylor NMI, Kowal J, Stahlberg H, Locher KP (2018) Cryo-EM structures of a human ABCG2 mutant trapped in ATP-bound and substrate-bound states. *Nature* 563: 426-430
- Martin C, Berridge G, Higgins CF, Mistry P, Charlton P, Callaghan R (2000a) Communication between multiple drug binding sites on P-glycoprotein. *Mol Pharmacol* 58: 624-632
- Martin C, Berridge G, Mistry P, Higgins C, Charlton P, Callaghan R (2000b) Drug binding sites on P-glycoprotein are altered by ATP binding prior to nucleotide hydrolysis. *Biochemistry* 39: 11901-11906
- Martin C, Higgins CF, Callaghan R (2001) The vinblastine binding site adopts high- and low-affinity conformations during a transport cycle of P-glycoprotein. *Biochemistry* 40: 15733-15742
- Matthews BW (2007) Five retracted structure reports: Inverted or incorrect? *Protein Science* 16: 1013–1016
- Mayerhofer PU (2016) Targeting and insertion of peroxisomal membrane proteins: ER trafficking versus direct delivery to peroxisomes. *Biochim Biophys Acta* 1863: 870-880
- McCormick JW, Vogel PD, Wise JG (2015) Multiple Drug Transport Pathways through Human P-Glycoprotein. *Biochemistry* 54: 4374-4390
- McDevitt CA, Collins RF, Conway M, Modok S, Storm J, Kerr ID, Ford RC, Callaghan R (2006) Purification and 3D Structural Analysis of Oligomeric Human Multidrug Transporter ABCG2. *Structure* 14: 1623-1632
- Meaden ER (2002) P-glycoprotein and MRP1 expression and reduced ritonavir and saquinavir accumulation in HIV-infected individuals. *Journal of Antimicrobial Chemotherapy* 50: 583-588
- Meier G, Thavarasah S, Ehrenbolger K, Hutter CAJ, Hürlimann LM, Barandun J, Seeger MA (2021) Deep mutational scan of a drug efflux pump reveals its structure-function landscape. *bioRxiv*: 2021.2010.2001.462730

- Meyer TH, van Endert PM, Uebel S, Ehring B, Tampe R (1994) Functional expression and purification of the ABC transporter complex associated with antigen processing (TAP) in insect cells. *FEBS Lett* 351: 443-447
- Mi W, Li Y, Yoon SH, Ernst RK, Walz T, Liao M (2017) Structural basis of MsbA-mediated lipopolysaccharide transport. *Nature* 549: 233-237
- Milazzo AC, Cheng A, Moeller A, Lyumkis D, Jacovetty E, Polukas J, Ellisman MH, Xuong NH, Carragher B, Potter CS (2011) Initial evaluation of a direct detection device detector for single particle cryo-electron microscopy. *J Struct Biol* 176: 404-408
- Mittal A, Böhm S, Grütter MG, Bordignon E, Seeger MA (2012) Asymmetry in the homodimeric ABC transporter MsbA recognized by a DARPin. *Journal of Biological Chemistry* 287: 20395-20406
- Mo W, Zhang JT (2012) Human ABCG2: Structure, function, and its role in multidrug resistance. *International Journal of Biochemistry and Molecular Biology* 3: 1-27
- Morgan JLW, Acheson JF, Zimmer J (2017) Structure of a Type-1 Secretion System ABC Transporter. *Structure* 25: 522-529
- Mosser J, Douar AM, Sarde CO, Kioschis P, Feil R, Moser H, Poustka AM, Mandel JL, Aubourg P (1993) Putative X-linked adrenoleukodystrophy gene shares unexpected homology with ABC transporters. *Nature* 361: 726-730
- Moyersoen J, Choe J, Fan E, Hol WGJ, Michels PAM (2004) Biogenesis of peroxisomes and glycosomes: Trypanosomatid glycosome assembly is a promising new drug target. *FEMS Microbiology Reviews* 28: 603-643
- Nandigama K, Lusvarghi S, Shukla S, Ambudkar SV (2019) Large-scale purification of functional human P-glycoprotein (ABCB1). *Protein Expr Purif* 159: 60-68
- Naoe Y, Nakamura N, Doi A, Sawabe M, Nakamura H, Shiro Y, Sugimoto H (2016) Crystal structure of bacterial haem importer complex in the inward-facing conformation. *Nature Communications* 7: 13411-13411
- Naydenova K, Russo CJ (2017) Measuring the effects of particle orientation to improve the efficiency of electron cryomicroscopy. *Nat Commun* 8: 629
- Neumann J, Rose-Sperling D, Hellmich UA (2017) Diverse relations between ABC transporters and lipids: An overview. *Biochim Biophys Acta Biomembr* 1859: 605-618
- Nikolaev M, Round E, Gushchin I, Polovinkin V, Balandin T, Kuzmichev P, Shevchenko V, Borshchevskiy V, Kuklin A, Round A *et al* (2017) Integral Membrane Proteins Can Be Crystallized Directly from Nanodiscs. *Crystal Growth & Design* 17: 945-948
- Nogales E, Scheres SH (2015) Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity. *Mol Cell* 58: 677-689
- Nöll A, Thomas C, Herbring V, Zollmann T, Barth K, Mehdipour AR, Tomasiak TM, Brüchert S, Joseph B, Abele R *et al* (2017) Crystal structure and mechanistic basis of a functional homolog of the antigen transporter TAP. *Proceedings of the National Academy of Sciences of the United States of America* 114: E438-E447
- Noreng S, Bharadwaj A, Posert R, Yoshioka C, Baconguis I (2018) Structure of the human epithelial sodium channel by cryo-electron microscopy. *Elife* 7: 39340
- Nosol K, Romane K, Irobalieva RN, Alam A, Kowal J, Fujita N, Locher KP (2020a) Cryo-EM structures reveal distinct mechanisms of inhibition of the human multidrug transporter ABCB1. *Proceedings of the National Academy of Sciences of the United States of America* 117: 26245-26253
- Nosol K, Romane K, Irobalieva RN, Alam A, Kowal J, Fujita N, Locher KP (2020b) Cryo-EM structures reveal distinct mechanisms of inhibition of the human multidrug transporter ABCB1. *Proc Natl Acad Sci U S A* 117: 26245-26253
- Nuti SL, Mehdi A, Rao US (2000) Activation of the human P-glycoprotein ATPase by trypsin. *Biochemistry* 39: 3424-3432
- Nyathi Y, Lousa CDM, Van Roermund CW, Wanders RJA, Johnson B, Baldwin SA, Theodoulou FL, Baker A (2010) The Arabidopsis peroxisomal ABC transporter, comatose, complements the Saccharomyces cerevisiae pxa1 pxa2Δ mutant for metabolism of long-chain fatty acids and exhibits fatty Acyl-CoA-stimulated ATpase activity. *Journal of Biological Chemistry* 285: 29892-29902

- Nyathi Y, Zhang X, Baldwin JM, Bernhardt K, Johnson B, Baldwin SA, Theodoulou FL, Baker A (2012) Pseudo half-molecules of the ABC transporter, COMATOSE, bind Pex19 and target to peroxisomes independently but are both required for activity. *FEBS Letters* 586: 2280-2286
- Okamoto T, Kawaguchi K, Watanabe S, Agustina R, Ikejima T, Ikeda K, Nakano M, Morita M, Imanaka T (2018) Characterization of human ATP-binding cassette protein subfamily D reconstituted into proteoliposomes. *Biochemical and Biophysical Research Communications* 496: 1122-1127
- Oldham ML, Grigorieff N, Chen J (2016a) Structure of the transporter associated with antigen processing trapped by herpes simplex virus. *eLife* 5: e21829
- Oldham ML, Hite RK, Steffen AM, Damko E, Li Z, Walz T, Chen J (2016b) A mechanism of viral immune evasion revealed by cryo-EM analysis of the TAP transporter. *Nature* 529: 537-540
- Olsen JA, Alam A, Kowal J, Stieger B, Locher KP (2020) Structure of the human lipid exporter ABCB4 in a lipid environment. *Nat Struct Mol Biol* 27: 62-70
- Omori K, Idei A (2003) Gram-negative bacterial atp-binding cassette protein exporter family and diverse secretory proteins. *Journal of Bioscience and Bioengineering* 95: 1-12
- Orlando BJ, Liao M (2020) ABCG2 transports anticancer drugs via a closed-to-open switch. *Nat Commun* 11: 2264
- Ostedgaard LS, Baldursson O, Welsh MJ (2001) Regulation of the Cystic Fibrosis Transmembrane Conductance Regulator Cl- Channel by Its R Domain. *Journal of Biological Chemistry* 276: 7689-7692
- Pandey A, Shin K, Patterson RE, Liu XQ, Rainey JK (2016) Current strategies for protein production and purification enabling membrane protein structural biology. *Biochem Cell Biol* 94: 507-527
- Parey K, Brandt U, Xie H, Mills DJ, Siegmund K, Vonck J, Kuhlbrandt W, Zickermann V (2018) Cryo-EM structure of respiratory complex I at work. *Elife* 7: e39213
- Parker JL, Newstead S (2014) Method to increase the yield of eukaryotic membrane protein expression in Saccharomyces cerevisiae for structural and functional studies. *Protein Sci* 23: 1309-1314
- Passmore LA, Russo CJ (2016) Specimen Preparation for High-Resolution Cryo-EM. *Methods Enzymol* 579: 51-86
- Paulino C, Neldner Y, Lam AK, Kalienkova V, Brunner JD, Schenck S, Dutzler R (2017) Structural basis for anion conduction in the calcium-activated chloride channel TMEM16A. *Elife* 6: e26232
- Perez C, Gerber S, Boilevin J, Bucher M, Darbre T, Aebi M, Reymond JL, Locher KP (2015) Structure and mechanism of an active lipid-linked oligosaccharide flippase. *Nature* 524: 433-438
- Perez C, Köhler M, Janser D, Pardon E, Steyaert J, Zenobi R, Locher KP (2017) Structural basis of inhibition of lipid-linked oligosaccharide flippase PglK by a conformational nanobody. *Scientific Reports* 7: 46641-46641
- Perez C, Mehdipour AR, Hummer G, Locher KP (2019) Structure of Outward-Facing PglK and Molecular Dynamics of Lipid-Linked Oligosaccharide Recognition and Translocation. *Structure* 27: 669-678.e665
- Petroni A, Paroni R, Aloisi AM, Blasevich M, Haman N, Fessas D (2019) Thermogenic flux induced by lignoceric acid in peroxisomes isolated from HepG2 cells and from X-adrenoleukodystrophy and control fibroblasts. *Journal of Cellular Physiology* 234: 18344-18348
- Pinfield-Wells H, Rylott EL, Gilday AD, Graham S, Job K, Larson TR, Graham IA (2005) Sucrose rescues seedling establishment but not germination of Arabidopsis mutants disrupted in peroxisomal fatty acid catabolism. *Plant Journal* 43: 861-872
- Poirier Y, Antonenkov VD, Glumoff T, Hiltunen JK (2006) Peroxisomal β-oxidation-A metabolic pathway with multiple functions. *Biochimica et Biophysica Acta - Molecular Cell Research* 1763: 1413-1426
- Poolman B, Doeven MK, Geertsma ER, Biemans-Oldehinkel E, Konings WN, Rees DC (2005) Functional analysis of detergent-solubilized and membrane-reconstituted ATP-binding cassette transporters. *Methods Enzymol* 400: 429-459

- Procko E, Ferrin-O'Connell I, Ng SL, Gaudet R (2006) Distinct Structural and Functional Properties of the ATPase Sites in an Asymmetric ABC Transporter. *Molecular Cell* 24: 51-62
- Procko E, O'Mara ML, Bennett WFD, Tieleman DP, Gaudet R (2009) The mechanism of ABC transporters: general lessons from structural and functional studies of an antigenic peptide transporter. *The FASEB Journal* 23: 1287-1302
- Qi X, Liu H, Thompson B, McDonald J, Zhang C, Li X (2019) Cryo-EM structure of oxysterol-bound human Smoothened coupled to a heterotrimeric Gi. *Nature* 571: 279-283
- Qian H, Zhao X, Cao P, Lei J, Yan N, Gong X (2017) Structure of the Human Lipid Exporter ABCA1. *Cell* 169: 1228-1239 e1210
- Qian P, Croll TI, Hitchcock A, Jackson PJ, Salisbury JH, Castro-Hartmann P, Sader K, Swainsbury DJ, Hunter CN (2021) Cryo-EM structure of the dimeric Rhodobacter sphaeroides RC-LH1 core complex at 2.9 A: the structural basis for dimerisation. *Biochem J* 478: 3923–3937
- Qiu W, Fu Z, Xu GG, Grassucci RA, Zhang Y, Frank J, Hendrickson WA, Guo Y (2018) Structure and activity of lipid bilayer within a membrane-protein transporter. *Proc Natl Acad Sci U S A* 115: 12985-12990
- Qu Q, Russell PL, Sharom FJ (2003) Stoichiometry and affinity of nucleotide binding to P-glycoprotein during the catalytic cycle. *Biochemistry* 42: 1170-1177
- Ratkeviciute G, Cooper BF, Knowles TJ (2021) Methods for the solubilisation of membrane proteins: the micelle-aneous world of membrane protein solubilisation. *Biochem Soc Trans* 49: 1763-1777
- Rawson S, Iadanza MG, Ranson NA, Muench SP (2016) Methods to account for movement and flexibility in cryo-EM data processing. *Methods* 100: 35-41
- Rice AJ, Park A, Pinkett HW (2014) Diversity in ABC transporters: Type I, II and III importers. *Critical Reviews in Biochemistry and Molecular Biology* 49: 426-437
- Riordan JR, Deuchars K, Kartner N, Alon N, Trent J, Ling V (1985) Amplification of P-glycoprotein genes in multidrug-resistant mammalian cell lines. *Nature* 316: 817-819
- Riordan JR, Rommens JM, Kerem BS, Alon NOA, Rozmahel R, Grzelczak Z, Zielenski J, Lok SI, Plavsic N, Chou JL *et al* (1989) Identification of the cystic fibrosis gene: Cloning and characterization of complementary DNA. *Science* 245: 1066-1073
- Ritchie TK, Grinkova YV, Bayburt TH, Denisov IG, Zolnerciks JK, Atkins WM, Sligar SG (2009) Chapter 11 - Reconstitution of membrane proteins in phospholipid bilayer nanodiscs. *Methods Enzymol* 464: 211-231
- Rohou A, Grigorieff N (2015) CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* 192: 216-221
- Rosenberg MF, Callaghan R, Ford RC, Higgins CF (1997) Structure of the multidrug resistance p-glycoprotein to 2.5 nm resolution determined by electron microscopy and image analysis. *Journal of Biological Chemistry* 272: 10685-10694
- Rubinstein JL (2007) Structural analysis of membrane protein complexes by single particle electron microscopy. *Methods* 41: 409-416
- Russell L, Larner V, Kurup S, Bougourd S, Holdsworth M (2000) The Arabidopsis COMATOSE locus regulates germination potential. *Development* 127: 3759-3767
- Russo CJ, Passmore LA (2014) Electron microscopy: Ultrastable gold substrates for electron cryomicroscopy. *Science* 346: 1377-1380
- Russo CJ, Passmore LA (2016) Ultrastable gold substrates: Properties of a support for high-resolution electron cryomicroscopy of biological specimens. *J Struct Biol* 193: 33-44
- Sánchez-Fernández R, Davies TGE, Coleman JOD, Rea PA (2001) The Arabidopsis thaliana ABC Protein Superfamily, a Complete Inventory. *Journal of Biological Chemistry* 276: 30231-30244
- Santos JA, Rempel S, Mous STM, Pereira CT, Beek JT, de Gier JW, Guskov A, Slotboom DJ (2018) Functional and structural characterization of an ECF-type ABC transporter for vitamin B12. *eLife* 7: e35828
- Sauvage V, Aubert D, Escotte-Binet S, Villena I (2009) The role of ATP-binding cassette (ABC) proteins in protozoan parasites. *Mol Biochem Parasitol* 167: 81-94

- Scala S, Akhmed N, Rao US, Paull K, Lan LB, Dickstein B, Lee JS, Elgemeie GH, Stein WD, Bates SE (1997) P-glycoprotein substrates and antagonists cluster into two distinct groups. *Mol Pharmacol* 51: 1024-1033
- Scheres SH (2012a) A Bayesian view on cryo-EM structure determination. *J Mol Biol* 415: 406-418
- Scheres SH (2012b) RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* 180: 519-530
- Schmidli C, Albiez S, Rima L, Righetto R, Mohammed I, Oliva P, Kovacik L, Stahlberg H, Braun T (2019) Microfluidic protein isolation and sample preparation for high-resolution cryo-EM. *Proc Natl Acad Sci U S A* 116: 15007-15012
- Schmitt L, Tampe R (2002) Structure and mechanism of ABC transporters. *Curr Opin Struct Biol* 12: 754-760
- Scortecci JF, Molday LL, Curtis SB, Garces FA, Panwar P, Van Petegem F, Molday RS (2021) Cryo-EM structures of the ABCA4 importer reveal mechanisms underlying substrate binding and Stargardt disease. *Nat Commun* 12: 5902
- Senior AE, Al-Shawi MK, Urbatsch IL (1995) The catalytic cycle of P-glycoprotein. *FEBS Letters* 377: 285-289
- Sgro GG, Costa TRD (2018) Cryo-EM Grid Preparation of Membrane Protein Samples for Single Particle Analysis. *Front Mol Biosci* 5: 74
- Shani N, Valle D (1996) A Saccharomyces cerevisiae homolog of the human adrenoleukodystrophy transporter is a heterodimer of two half ATP-binding cassette transporters. *Proc Natl Acad Sci U S A* 93: 11901-11906
- Shani N, Watkins PA, Valle D (1995) PXA1, a possible Saccharomyces cerevisiae ortholog of the human adrenoleukodystrophy gene. *Proc Natl Acad Sci U S A* 92: 6012-6016
- Sharom FJ (1995) Characterization and functional reconstitution of the multidrug transporter. *J Bioenerg Biomembr* 27: 15-22
- Shibata Y, Ojika M, Sugiyama A, Yazaki K, Jones DA, Kawakita K, Takemoto D (2016) The Full-Size ABCG Transporters Nb-ABCG1 and Nb-ABCG2 Function in Pre- and Postinvasion Defense against Phytophthora infestans in Nicotiana benthamiana. *Plant Cell* 28: 1163-1181
- Shockey J, Browse J (2011) Genome-level and biochemical diversity of the acyl-activating enzyme superfamily in plants. *Plant J* 66: 143-160
- Shockey JM, Fulda MS, Browse JA (2002) Arabidopsis contains nine long-chain acyl-coenzyme a synthetase genes that participate in fatty acid and glycerolipid metabolism. *Plant Physiol* 129: 1710-1722
- Sigworth FJ (2016) Principles of cryo-EM single-particle image processing. *Microscopy (Oxf)* 65: 57-67
- Sikimic J, McMillen TS, Bleile C, Dastvan F, Quast U, Krippeit-Drews P, Drews G, Bryan J (2018) ATP binding without hydrolysis switches sulfonylurea receptor 1 (SUR1) to outward-facing conformations that activate K ATP channels. *Journal of Biological Chemistry* 294: 3707-3719
- Simeonov P, Werner S, Haupt C, Tanabe M, Bacia K (2013) Membrane protein reconstitution into liposomes guided by dual-color fluorescence cross-correlation spectroscopy. *Biophys Chem* 184: 37-43
- Smith PC, Karpowich N, Millen L, Moody JE, Rosen J, Thomas PJ, Hunt JF (2002) ATP binding to the motor domain from an ABC transporter drives formation of a nucleotide sandwich dimer. *Molecular Cell* 10: 139-149
- Song G, Zhang S, Tian M, Zhang L, Guo R, Zhuo W, Yang M (2021) Molecular insights into the human ABCB6 transporter. *Cell Discov* 7: 55
- Song WY, Park J, Mendoza-Cozatl DG, Suter-Grotemeyer M, Shim D, Hortensteiner S, Geisler M, Weder B, Rea PA, Rentsch D *et al* (2010) Arsenic tolerance in Arabidopsis is mediated by two ABCC-type phytochelatin transporters. *Proc Natl Acad Sci U S A* 107: 21187-21192
- Sonveaux N, Vigano C, Shapiro AB, Ling V, Ruysschaert JM (1999) Ligand-mediated tertiary structure changes of reconstituted P-glycoprotein. A tryptophan fluorescence quenching analysis. *J Biol Chem* 274: 17649-17654

- Srikant S (2020) Evolutionary history of ATP-binding cassette proteins. *FEBS Lett* 594: 3882-3897
- Srikant S, Gaudet R, Murray AW (2020) Selecting for Altered Substrate Specificity Reveals the Evolutionary Flexibility of ATP-Binding Cassette Transporters. *Curr Biol* 30: 1689-1702 e1686
- Srinivasan V, Pierik AJ, Lill R (2014) Crystal structures of nucleotide-free and glutathione-bound mitochondrial ABC transporter Atm1. *Science* 343: 1137-1140
- Stefan E, Hofmann S, Tampe R (2020) A single power stroke by ATP binding drives substrate translocation in a heterodimeric ABC transporter. *Elife* 9: e55943
- Štefková J, Poledne R, Hubáček JA (2004) ATP-binding cassette (ABC) transporters in human metabolism and diseases. *Physiological Research* 53: 235-243
- Stetsenko A, Guskov A (2017) An Overview of the Top Ten Detergents Used for Membrane Protein Crystallization. *Crystals* 7: 197
- Stieger B, Steiger J, Locher KP (2021) Membrane lipids and transporter function. *Biochim Biophys Acta Mol Basis Dis* 1867: 166079
- Stockner T, Gradisch R, Schmitt L (2020) The role of the degenerate nucleotide binding site in type I ABC exporters. *FEBS Lett* 594: 3815-3838
- Sun C, Benlekbir S, Venkatakrishnan P, Wang Y, Hong S, Hosler J, Tajkhorshid E, Rubinstein JL, Gennis RB (2018) Structure of the alternative complex III in a supercomplex with cytochrome oxidase. *Nature* 557: 123-126
- Sun Y, Wang J, Long T, Qi X, Donnelly L, Elghobashi-Meinhardt N, Esparza L, Cohen JC, Xie XS, Hobbs HH *et al* (2021) Molecular basis of cholesterol efflux via ABCG subfamily transporters. *Proc Natl Acad Sci U S A* 118
- Szollosi D, Rose-Sperling D, Hellmich UA, Stockner T (2018) Comparison of mechanistic transport cycle models of ABC exporters. *Biochim Biophys Acta Biomembr* 1860: 818-832
- Tanaka AR, Tanabe K, Morita M, Kurisu M, Kasiwayama Y, Matsuo M, Kioka N, Amachi T, Imanaka T, Ueda K (2002) ATP binding/hydrolysis by and phosphorylation of peroxisomal ATP-binding cassette proteins PMP70 (ABCD3) and adrenoleukodystrophy protein (ABCD1). *J Biol Chem* 277: 40142-40147
- Tani K, Kanno R, Makino Y, Hall M, Takenouchi M, Imanishi M, Yu LJ, Overmann J, Madigan MT, Kimura Y *et al* (2020) Cryo-EM structure of a Ca(2+)-bound photosynthetic LH1-RC complex containing multiple alphabeta-polypeptides. *Nat Commun* 11: 4955
- Tarling EJ, Vallim TQdA, Edwards PA (2013) Role of ABC transporters in lipid transport and human disease. *Trends in Endocrinology and Metabolism* 24: 342-350
- Tate CG (2001) Overexpression of mammalian integral membrane proteins for structural studies. *FEBS Lett* 504: 94-98
- Taylor KA, Glaeser RM (2008) Retrospective on the early development of cryoelectron microscopy of macromolecules and a prospective on opportunities for the future. *J Struct Biol* 163: 214-223
- Taylor NMI, Manolaridis I, Jackson SM, Kowal J, Stahlberg H, Locher KP (2017) Structure of the human multidrug transporter ABCG2. *Nature* 546: 504-509
- Ter Beek J, Guskov A, Slotboom DJ (2014) Structural diversity of ABC transporters. *Journal of General Physiology* 143: 419-435
- Theodoulou FL, Job K, Slocombe SP, Footitt S, Holdsworth M, Baker A, Larson TR, Graham IA (2005) Jasmonic acid levels are reduced in COMATOSE ATP-binding cassette transporter mutants. Implications for transport of jasmonate precursors into peroxisomes. *Plant Physiology* 137: 835-840
- Thomas C, Aller SG, Beis K, Carpenter EP, Chang G, Chen L, Dassa E, Dean M, Duong Van Hoa F, Ekiert D *et al* (2020) Structural and functional diversity calls for a new classification of ABC transporters. *FEBS Lett* 594: 3767-3775
- Thomas JA, Tate CG (2014) Quality control in eukaryotic membrane protein overproduction. *J Mol Biol* 426: 4139-4154
- Timachi MH, Hutter CA, Hohl M, Assafa T, Bohm S, Mittal A, Seeger MA, Bordignon E (2017) Exploring conformational equilibria of a heterodimeric ABC transporter. *Elife* 6: e20236
- Tsai MF, Li M, Hwang TC (2010) Stable ATP binding mediated by a partial NBD dimer of the CFTR chloride channel. *Journal of General Physiology* 135: 399-414

- Tsunoda J, Song C, Imai FL, Takagi J, Ueno H, Murata T, Iino R, Murata K (2018) Off-axis rotor in Enterococcus hirae V-ATPase visualized by Zernike phase plate single-particle cryo-electron microscopy. *Sci Rep* 8: 15632
- Tusnady GE, Sarkadi B, Simon I, Varadi A (2006) Membrane topology of human ABC proteins. *FEBS Lett* 580: 1017-1022
- Unger L, Ronco-Campana A, Kitchen P, Bill RM, Rothnie AJ (2021) Biological insights from SMA-extracted proteins. *Biochem Soc Trans* 49: 1349-1359
- Unger T, Peleg Y (2012) Recombinant protein expression in the baculovirus-infected insect cell system. *Methods Mol Biol* 800: 187-199
- Urbatsch IL, Sankaran B, Bhagat S, Senior AE (1995a) Both P-glycoprotein nucleotide-binding sites are catalytically active. *J Biol Chem* 270: 26956-26961
- Urbatsch IL, Sankaran B, Weber J, Senior AE (1995b) P-glycoprotein is stably inhibited by vanadate-induced trapping of nucleotide at a single catalytic site. *J Biol Chem* 270: 19383-19390
- van Roermund CW, Visser WF, Ijlst L, van Cruchten A, Boek M, Kulik W, Waterham HR, Wanders RJ (2008) The human peroxisomal ABC half transporter ALDP functions as a homodimer and accepts acyl-CoA esters. *FASEB J* 22: 4201-4208
- van Roermund CWT, Ijlst L, Majczak W, Waterham HR, Folkerts H, Wanders RJA, Hellingwerf KJ (2012) Peroxisomal fatty acid uptake mechanism in Saccharomyces cerevisiae. *The Journal of biological chemistry* 287: 20144-20153
- Van Roermund CWT, Ijlst L, Wagemans T, Wanders RJA, Waterham HR (2014) A role for the human peroxisomal half-transporter ABCD3 in the oxidation of dicarboxylic acids. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids* 1841: 563-568
- van Roermund CWT, L IJ, Baker A, Wanders RJA, Theodoulou FL, Waterham HR (2021) The Saccharomyces cerevisiae ABC subfamily D transporter Pxa1/Pxa2p co-imports CoASH into the peroxisome. *FEBS Lett* 595: 763-772
- Van Roermund CWT, Visser WF, Ijlst L, Waterham HR, Wanders RJA (2011) Differential substrate specificities of human ABCD1 and ABCD2 in peroxisomal fatty acid β-oxidation. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids* 1811: 148-152
- van Veen HW, Margolles A, Muller M, Higgins CF, Konings WN (2000) The homodimeric ATP-binding cassette transporter LmrA mediates multidrug transport by an alternating two-site (two-cylinder engine) mechanism. *EMBO J* 19: 2503-2514
- Vasiliou V, Vasiliou K, Nebert DW (2009) Human ATP-binding cassette (ABC) transporter family. *Hum Genomics* 3: 281-290
- Venter CJ, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA *et al* (2001) The sequence of the human genome. *Science* 291: 1304-1351
- Verleur N, Hettema EH, Van Roermund CWT, Tabak HF, Wanders RJA (1997) Transport of activated fatty acids by the peroxisomal ATP-binding-cassette transporter Pxa2 in a semi-intact yeast cell system. *European Journal of Biochemistry* 249: 657-661
- Verrier PJ, Bird D, Burla B, Dassa E, Forestier C, Geisler M, Klein M, Kolukisaoglu U, Lee Y, Martinoia E *et al* (2008) Plant ABC proteins--a unified nomenclature and updated inventory. *Trends Plant Sci* 13: 151-159
- Vigano C, Margolles A, van Veen HW, Konings WN, Ruysschaert JM (2000) Secondary and tertiary structure changes of reconstituted LmrA induced by nucleotide binding or hydrolysis. A fourier transform attenuated total reflection infrared spectroscopy and tryptophan fluorescence quenching analysis. *J Biol Chem* 275: 10962-10967
- Wagner S, Baars L, Ytterberg AJ, Klussmeier A, Wagner CS, Nord O, Nygren PA, van Wijk KJ, de Gier JW (2007) Consequences of membrane protein overexpression in Escherichia coli. *Mol Cell Proteomics* 6: 1527-1550
- Wagner T, Merino F, Stabrin M, Moriya T, Antoni C, Apelbaum A, Hagel P, Sitsel O, Raisch T, Prumbaum D *et al* (2019) SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Commun Biol* 2: 218
- Walker JE, Saraste M, Runswick MJ, Gay NJ (1982) Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *The EMBO Journal* 1: 945-951

- Walter JD, Sawicka M, Dutzler R (2019) Cryo-EM structures and functional characterization of murine Slc26a9 reveal mechanism of uncoupled chloride transport. *Elife* 8: e46986
- Wanders RJ, Waterham HR (2006) Biochemistry of mammalian peroxisomes revisited. *Annu Rev Biochem* 75: 295-332
- Wanders RJA (2014) Metabolic functions of peroxisomes in health and disease. *Biochimie* 98: 36-44
- Wang C, Cao C, Wang N, Wang X, Wang X, Zhang XC (2020a) Cryo-electron microscopy structure of human ABCB6 transporter. *Protein Sci* 29: 2363-2374
- Wang L, Hou WT, Chen L, Jiang YL, Xu D, Sun L, Zhou CZ, Chen Y (2020b) Cryo-EM structure of human bile salts exporter ABCB11. *Cell Research* 30: 623–625
- Wang R, Qin Y, Li X (2021) Structural basis of acyl-CoA transport across the peroxisomal membrane by human ABCD1. *Cell Res* 0: 1–4
- Ward A, Reyes CL, Yu J, Roth CB, Chang G (2007) Flexibility in the ABC transporter MsbA: Alternating access with a twist. *Proceedings of the National Academy of Sciences of the United States of America* 104: 19005-19010
- Whitcomb RW, Linehan WM, Knazek RA (1988) Effects of long-chain, saturated fatty acids on membrane microviscosity and adrenocorticotropin responsiveness of human adrenocortical cells in vitro. *Journal of Clinical Investigation* 81: 185-188
- Wiesinger C, Kunze M, Regelsberger G, Forss-Petter S, Berger J (2013) Impaired very long-chain acyl-CoA β-oxidation in human X-linked adrenoleukodystrophy fibroblasts is a direct consequence of ABCD1 transporter dysfunction. *Journal of Biological Chemistry* 288: 19269-19279
- Wright J, Muench SP, Goldman A, Baker A (2018) Substrate polyspecificity and conformational relevance in ABC transporters: New insights from structural studies. *Biochemical Society Transactions* 46: 1475-1484
- Wu CP, Lusvarghi S, Wang JC, Hsiao SH, Huang YH, Hung TH, Ambudkar SV (2019) Avapritinib: A Selective Inhibitor of KIT and PDGFRα that Reverses ABCB1 and ABCG2-Mediated Multidrug Resistance in Cancer Cell Lines. *Molecular Pharmaceutics* 16: 3040-3052
- Xiong C, Jia L, Shen M, Xiong W, Xiong L, Wang T, Zhou D, Liu Z, Tang L (2021) ATP and Substrate Binding Regulates Conformational Changes of Human Peroxisomal ABC Transporter ALDP. *BioRxiv*
- Xu D, Feng Z, Hou WT, Jiang YL, Wang L, Sun L, Zhou CZ, Chen Y (2019) Cryo-EM structure of human lysosomal cobalamin exporter ABCD4. *Cell Research* 29: 1039-1041
- Yernaux C, Fransen M, Brees C, Lorenzen S, Michels PAM (2006) Trypanosoma brucei glycosomal ABC transporters: Identification and membrane targeting. *Molecular Membrane Biology* 23: 157-172
- Yin J, Zhang J (2011) Multidrug resistance-associated protein 1 (MRP1/ABCC1) polymorphism: From discovery to clinical application. *Journal of Central South University (Medical Sciences)* 36: 927-938
- Zaitseva J, Jenewein S, Jumpertz T, Holland IB, Schmitt L (2005) H662 is the linchpin of ATP hydrolysis in the nucleotide-binding domain of the ABC transporter HlyB. *EMBO Journal* 24: 1901-1910
- Zaitseva J, Oswald C, Jumpertz T, Jenewein S, Wiedenmann A, Holland IB, Schmitt L (2006) A structural analysis of asymmetry required for catalytic activity of an ABC-ATPase domain dimer. *EMBO Journal* 25: 3432-3443
- Zhang K (2016) Gctf: Real-time CTF determination and correction. *J Struct Biol* 193: 1-12
- Zhang X, De Marcos Lousa C, Schutte-Lensink N, Ofman R, Wanders RJ, Baldwin SA, Baker A, Kemp S, Theodoulou FL (2011) Conservation of targeting but divergence in function and quality control of peroxisomal ABC transporters: an analysis using cross-kingdom expression. *Biochem J* 436: 547-557
- Zhang Y, Sun B, Feng D, Hu H, Chu M, Qu Q, Tarrasch JT, Li S, Sun Kobilka T, Kobilka BK *et al* (2017) Cryo-EM structure of the activated GLP-1 receptor in complex with a G protein. *Nature* 546: 248-253

- Zhang Z, Chen J (2016) Atomic Structure of the Cystic Fibrosis Transmembrane Conductance Regulator. *Cell* 167: 1586-1597.e1589
- Zhang Z, Liu F, Chen J (2018) Molecular structure of the ATP-bound, phosphorylated human CFTR. *Proc Natl Acad Sci U S A* 115: 12757-12762
- Zheng SQ, Palovcak E, Armache JP, Verba KA, Cheng Y, Agard DA (2017) MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* 14: 331-332
- Zhou Y, Ojeda-May P, Pu J (2013) H-loop histidine catalyzes ATP hydrolysis in the E. coli ABC-transporter HlyB. *Physical Chemistry Chemical Physics* 15: 15811-15815
- Zielenski J, Rozmahel R, Bozon D, Kerem Bs, Grzelczak Z, Riordan JR, Rommens J, Tsui LC (1991) Genomic DNA sequence of the cystic fibrosis transmembrane conductance regulator (CFTR) gene. *Genomics* 10: 214-228
- Zivanov J, Nakane T, Scheres SHW (2019) A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *IUCrJ* 6: 5-17
- Zolman BK, Silva ID, Bartel B (2001) The Arabidopsis pxa1 mutant is defective in an ATP-binding cassette transporter-like protein required for peroxisomal fatty acid β-oxidation. *Plant Physiology* 127: 1266-1278
- Zolnerciks JK, Akkaya BG, Snippe M, Chiba P, Seelig A, Linton KJ (2014) The Q loops of the human multidrug resistance transporter ABCB1 are necessary to couple drug binding to the ATP catalytic cycle. *FASEB Journal* 28: 4335-4346
- Zutz A, Hoffmann J, Hellmich UA, Glaubitz C, Ludwig B, Brutschy B, Tampé R (2011) Asymmetric ATP hydrolysis cycle of the heterodimeric multidrug ABC transport complex TmrAB from Thermus thermophilus. *Journal of Biological Chemistry* 286: 7104-7115

8 Appendix

8.1 SF9 Cell Lysate ‘Purification’

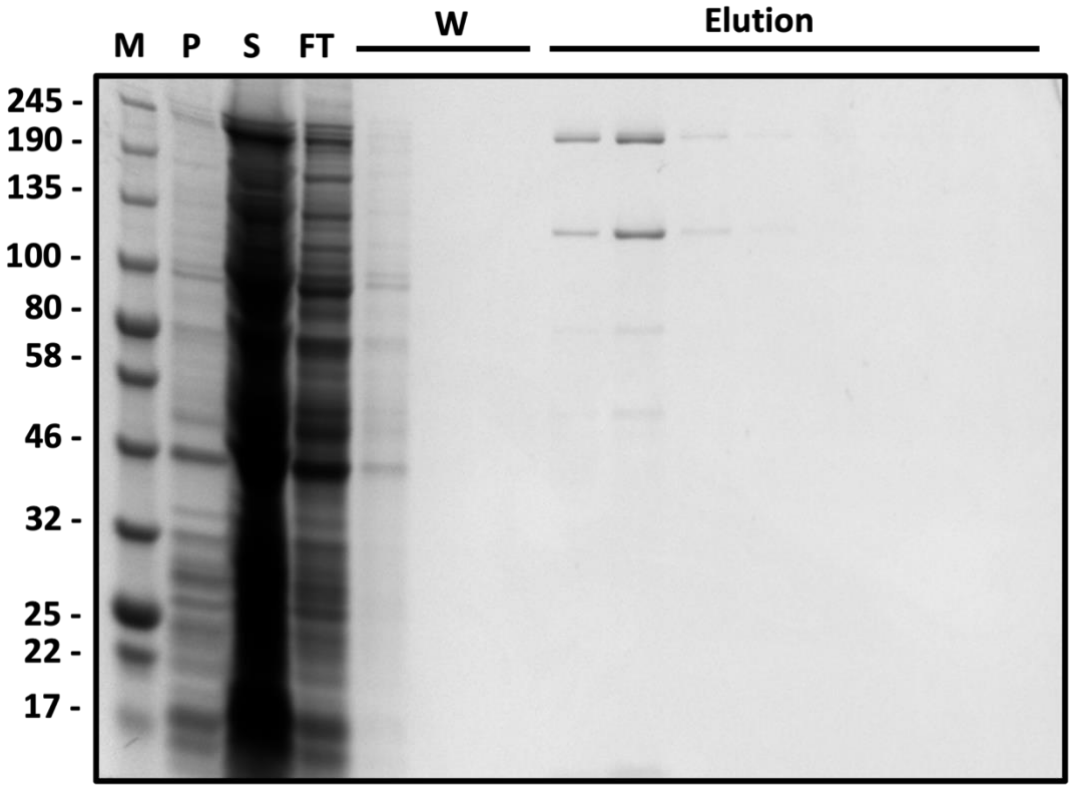


Figure 8-1. Mock 'purification' running Sf9 cell lysate over Streptactin XT Superflow resin. Sf9 cells contain a protein that has a strong affinity this resin. CTS would normally be expected to run just below 190 kDa (in between the two major bands observed here. The contaminants identity is unknown.

8.2 ATPase Assay Time-Points

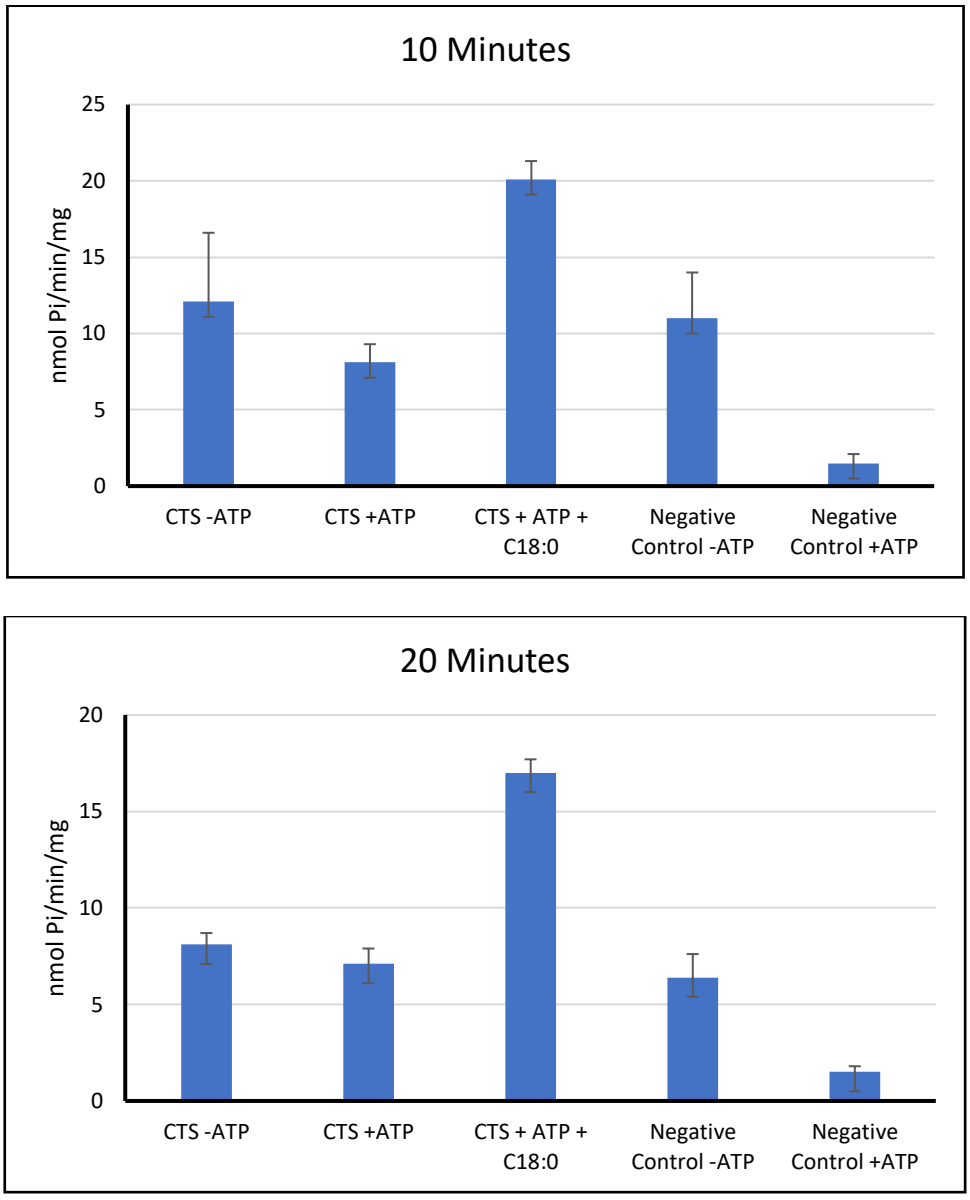


Figure 8-2. Time-point ATPase assays. ATPase assays carried out at different time points confirming that incubation with substrate stimulates the activity of CTS and that the recorded phosphate release is within the linear range of the assay as compared to figure 4.5 . Top, 10-minute incubation; bottom, 20-minute incubation. Error bars represent SD from triplicate measurements. N=1.