Structural and functional studies on a prokaryotic homologue of the mammalian SLC7 cationic amino acid transporters

Katharina Esther Julia Jungnickel

Department of Biochemistry

and

Brasenose College

A thesis submitted for the partial fulfilment of the degree of Doctor of Philosophy,

February 2018

Supervised by:

Prof. Elspeth Garman

Prof. Simon Newstead

Structural and functional studies on a prokaryotic homologue of the mammalian SLC7 cationic amino acid transporters

Katharina Esther Julia Jungnickel

Department of Biochemistry and Brasenose College, University of Oxford

Submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Abstract

The Amino acid/Polyamine/Cationic transporter superfamily is among the largest of the secondary active transporters. APC family members use either a proton or sodium gradient to drive substrate translocation across the membrane. Members of the APC superfamily are responsible for the uptake of amino acids and are found in all kingdoms of life. Amino acids, and their derivatives, are crucial for both prokaryotic and eukaryotic cell biology, being the building blocks of a variety of macromolecules and play important roles in metabolism and cell signalling. Therefore, the systems responsible for their transport are important targets of further study, in particular to gain greater insight into how these proteins recognise and discriminate between different types of amino acids. The transport mechanism of several APC members has been elucidated using a variety of structural, biophysical and functional approaches. However, these studies have focussed on sodium coupled transporters due to their function as neurotransmitter transporters. Here, the crystal structure of the first proton-coupled amino acid transporter of the APC superfamily, from *Geobacillus kaustophilus, Gk*ApcT, is presented. *Gk*ApcT closely related to the mammalian cationic amino acid transporters (CATs) of the SLC7 family, responsible for the transport arginine and lysine into the cell. The mechanism of transport was further investigated using liposome-based functional assays to characterise the substrate specificity and kinetics. The crystal structure was used to guide a detailed site-directed mutagenesis study, which identified several conserved residues involved in substrate binding and proton-coupling. The current structure serves as a template to further understand proton coupled amino acid transport in bacteria and to understand mammalian amino acid transport via the SLC7 and SLC36 families.

Acknowledgements

I would like to express my deep gratitude to Professor Elspeth Garman and Professor Simon Newstead, my research supervisors, for their guidance, support, time, encouragement and welcoming me into their research groups. It was truly a pleasure to be working with them. Without the two of them, I would not be the researcher that I am today.

I am happy I could and part of the Marie Curie ITN network NanoMem providing my research funding but through which I met so many wonderful people from all over Europe. It was a pleasure to have had Anthony Savill (Molecular Dimensions Ltd) as associate partner in the NanoMem, who gave me insight into product development and who supported me during my studies.

Additionally, I would like to thank Dr. Joanne Parker, for her help in the lab, taking the time to answer all the questions I ever had and for the fruitful discussions and advice.

My grateful thanks also extend to Dr. Ed Lowe for his help during the serial crystallography experiments, for being the encyclopaedia about every crystallography related topics, for taking his time to explain everything in detail and for his support.

Furthermore, many thanks to the members of the Newstead and Garman group, past and present, for having me as a group member and for making the time I spent there with them lovely.

I am especially grateful to Gurdeep, Oleg, Frank, Sebastian, and Richard for being there for me and making me laugh. It was a pleasure having you around in the lab and outside of it, and I will deeply miss you

I wish to thank my friends from home, Oxford, and other parts of the world for their support and encouragement throughout my study.

Finally, my deepest gratitude is to my family and Stephan for supporting me, for being there whenever necessary and for making me the person I am. I cannot explain how much it means to me to have you in my world.

TABLE OF CONTENTS

List of abbreviations

7.7MAG	1-(7Ztetradecenoyl)-rac-glycerol
7.8MAG	1-(7Z-pentadecenoyl)-rac-glycerol
ABC	ATP binding cassette
ADP	adenosine diphosphate
APC	amino acid/polyamine/organocation transporter
ATP	adenosine triphosphate
BCCT	betaine / carnitine / choline transporters
CAT	cationic amino acid transporter
CCCP	carbonyl cyanide m-chlorophenyl hydrazone
CD	circular dichroism
Chl	cholesterol
CHS	cholesterol hemisuccinate
CMC	critical micelle concentration
DDM	dodecyl β-maltopyranoside
DLS	Diamond Light Source
DM	decyl β-maltopyranoside
DNA	deoxyribonucleic acid
EAAT	excitatory amino acid transporter
EL	extracellular loop
eNOS	endothelial nitric oxide synthase
EPR	electro paramagnetic resonance
FSEC	fluorescent size exclusion chromatography
FWHM	full width at half maximum
GABA	γ-aminobutyric acid
GFP	green fluorescent protein
GPCR	G-protein coupled receptor
HAT	heteromeric amino acid transporter
ID	inner diameter
iNOS	inducible nitric oxide synthase

KAc	potassium acetate
KAN	kanamycin
KB	Kirkpatrick-Baez
KF	potassium fluoride
KNatart	potassium sodium tartrate
KPi	potassium phosphate
KO	knock-out
LAT	L-type amino acid transporter
LB	lysogenic broth
LCP	lipidic cubic phase
LDAO	n-lauryl dimethylamine n-oxide
LSP	lipidic sponge phase
MBP	maltose binding protein
MFS	major facilitator family
MLV	multilamellar vesicle
MMC	methyl mercury chloride
MO	monoolein, 1-(cis-9-octadecenoyl)-rac-glycerol
MR	molecular replacement
NaAc	sodium acetate
NaPi	sodium phosphate
NCS1	nucleobase:cation symporter-1
OGNG	2,2-dihexylpropane-1,3-bis-β-D-glucopyranoside
DMNG	2,2-dioctylpropane-1,3-bis-β-D-maltopyranoside
Ni-NTA	nickel nitrilotriacetic acid
nNOS	neuronal nitric oxide synthase
NO	nitric oxide
OD	outer diameter
OG	octylglucpyranoside
PC	phosphatidylcholine
PEG	polyethylene glycol
POPE	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine

POPG	1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol)
pyranine	trisodium 8-hydroxypyrene-1,3,6-trisulfonate
RMSD	root mean square deviation
SAD	single-wavelength anomalous dispersion
SBS	substrate binding site
SeMet	selenomethionine
SFX	serial femtosecond X-ray crystallography
SIR	single isomorphous replacement
SIRAS	single isomorphous replacement with anomalous scattering
SONICC	second order nonlinear imaging of chiral crystals
SX	serial X-ray crystallography at a synchrotron
TB	terrific broth
TEV	tobacco etch virus
TLS	translation-libration-screw
TM	transmembrane helix
ULV	unilamellar vesicle
UV	ultraviolet
WT	wildtype
XFEL	X-ray free electron laser
YAG	yttrium aluminium garnet

“It’s human nature to stretch, to go, to see, to understand.

Exploration is not a choice, really; it’s an imperative.”

Michael Collins

1 Membrane transport of Amino Acids

The Amino acid/Polyamine/Cationic (APC) superfamily is among the largest families of secondary active transporters with a diverse range of substrates such as amino acids, sugars and smaller peptides. In order to transport these organic compounds across the membrane, members of the APC superfamily have evolved to use different mechanisms.

For example, the most prominent member of the APC superfamily is LeuT, a transporter specific for leucine and alanine, and member of the neurotransmitter:sodium symporter (NSS) family (Yamashita *et al.*, 2005). LeuT couples substrate transport to two sodium ions that have been shown to bind in the centre of the protein (Yamashita *et al.*, 2005). MhsT, another transporter of the NSS family, works in a similar fashion, using the inwardly directed sodium gradient to couple uptake of bulky hydrophobic amino acids (Quick and Javitch, 2007). LeuT is one of the most studied APC members due to its relation to the human members of the NSS family, also referred to as SLC6 family (Hediger *et al.*, 2004). Human SLC6 transporters are found in the pre-synapse of neurons (Schloss *et al.*, 1992). There, they are responsible for Na^+/Cl^- coupled uptake of neurotransmitters released into the synaptic cleft (Schloss *et al.*, 1992). Due to this function, the human SLC6 transporters have been recognised as important targets for drug development.

The bacterial Mhp1 is another sodium-coupled transporter of the APC superfamily which belongs to the nucleobase:sodium symporter 1 (NCS1) family, transporting indolyl methyl- and benzylhydantoins (Suzuki and Henderson, 2006; Weyand *et al.*, 2008). The latter is the precursor for phenylalanine and tryptophan synthesis in the cell (Suzuki and Henderson, 2006). Mhp1 is thus an example of a member of the APC superfamily not transporting amino acids but their precursors, or as the sodium-coupled sugar transporter vSGLT non-amino acid substrates (Faham *et al.*, 2008), member of the solute sodium symporter (SSS) family. Members of this family use the inward electrochemical sodium gradient to harness uptake of sugars, amino acids and vitamins, vital for the cell (Wright, 2004).

BetP is a sodium-coupled transporter of the betaine / carnitine / choline transporters (BCCT) family involved in response to osmotic stress in bacteria through transport of the zwitterionic betaine with sodium ions (Farwick *et al.*, 1995; Perez *et al.*, 2011). Another transporter belonging to the BCCT family is the carnitine/γ-butyrobetaine exchanger CaiT which is not

involved in a stress response mechanism (Jung *et al.*, 2002). However, carnitine is important for fatty acid transport at the bacterial membrane (Rebouche and Seim, 1998).

AdiC and GadC, members of the APC family, are also antiporters, involved in the response mechanism to acidic stress. In AdiC the transport of arginine into the cell is coupled to the exchange for an agmatine molecule (Iyer *et al.*, 2003). Arginine is converted in the cell to agmatine, onto which a proton is transferred during this process (Iyer *et al.*, 2003). Therefore, excretion of agmatine out of the cell indirectly moves protons (virtual proton pump) away from the cell interior, regulating the internal pH (Iyer *et al.*, 2003). A similar mechanism is observed for GadC, the glutamate/γ-aminobutyric acid (GABA) exchanger, where glutamate is converted into GABA via decarboxylation and proton-transfer (Hersh *et al.*, 1996). GABA exits the cell via GadC in exchange for a glutamate. Both AdiC and GadC are pH-dependent and show the highest antiporter activity at low pHs (Fang *et al.*, 2007; Ma *et al.*, 2012a) which is not surprising since the cell needs to respond to acidic stress in these environments.

In 2009 another close homologue to AdiC was solved by X-ray crystallography, the archaeal amino acid transporter *Mj*ApcT from *Methanocaldococcus janaschii* (Shaffer *et al.*, 2009). This transporter is specific for alanine, serine, glutamate and glutamine, and although homologous to AdiC, it was proposed to be proton-coupled (Shaffer *et al.*, 2009).

Although many structures and functional data of APC superfamily members are available, only little is known about the proton-coupling mechanism within the APC superfamily facilitated, for example, by the human proton-assisted amino acid transporters (PATs) of the SLC36 family (Thwaites *et al.*, 1993; Boll *et al.*, 2003, 2004). Additionally, a proton-coupling mechansims has also been proposed for the cationic amino acid transporters (CATs) of the SLC7 family in *Arabidopsis thaliana* (Frommer *et al.*, 1995; Hammes *et al.*, 2006).

Most of the current research has been focussed on the sodium-coupled transporters due to their relationship with human neurotransmitter transporters as potential drug targets. Therefore, more research on the mechanism of proton-coupling and its differences to the sodium-coupled mechanism in terms of ion-binding sites and the potential interaction with the substrate is necessary. Insight into proton-coupled amino acid transporters might elucidate how they evolved to bind and release either protons, sodium ions, or do not require any ion interaction.

The examples given above have provided a small overview about the versatility of the APC superfamily regarding transported molecules and the mechanism they use for substrate translocation. The transporters of the APC superfamily fulfil a variety of functions such as providing nutrients for the cell, involvement in neurotransmitter regulation or stress response. Further details into the APC superfamily will be given in the following sections.

1.1 THE ELECTROCHEMICAL POTENTIAL

As described above, many APC superfamily members couple substrate transport to ion movement. This is necessary since the substrate is mostly transported against its electrochemical gradient. To harness the energy to perform this 'uphill' transport the APC transporters use the electrochemical gradient of ions such as sodium ions or protons across the membrane.

The electrochemical gradient $\Delta\tilde{\mu}$ of an ion is described by an electrical and a chemical component. Due to the different ion compositions on each side of the membrane, a membrane potential builds up across it, and this charge difference (the membrane potential $\Delta\Psi$) is the electrical component of the electrochemical gradient. The chemical component, is described by the difference in the ion concentration on both sides of the membrane (Δion, e.g. for proton gradient ΔpH). The electrochemical gradient of a cation (X^{m+}) can be expressed in the Nernst-equation:

$$\Delta\tilde{\mu}_{X^{m+}} = \Delta\Psi - 2.3\left(\frac{RT}{mF}\right)\log\left(\frac{[X^{m+}]_{int}}{[X^{m+}]_{ext}}\right)$$

where T is the absolute temperature, $\Delta\Psi$ is the electrical potential in mV, R is the universal gas constant equal to 8.314 J($K^{-1}mol^{-1}$), F the Faraday constant equal to 96,485 C(mol^{-1}), m the valence of the ion, and $[X^{m+}]$ the concentration of the ion in the internal (int) and external (ext) environment.

The electrochemical gradient of ions is the driving force for many cellular processes. For instance, one of the most important biosynthetic processes benefits from the proton electrochemical gradient $\Delta\tilde{\mu}H^{+}$, also referred to as the proton motive force, to drive the generation of adenosine tri-phosphate (ATP), the main energy source for the cell, facilitated by

F_1F_0-ATPases (Mitchell, 1985). Thus, this proton electrochemical gradient has to be maintained in order to ensure continuous ATP synthesis. To achieve this, protons have to be actively moved out of the cell, which additionally prevents the interior of the cell from acidifying, necessary to avoid endangering the viability of the cell (Baker-Austin and Dopson, 2007; Slonczewski *et al.*, 2009; Krulwich *et al.*, 2011).

To ensure a constant proton gradient, specific enzymes and enzyme complexes exist in all kingdoms of life to transport protons. The most well-known example is the electron transfer chain in the mitochondrion, where the transfer of electrons from electron donors to electron acceptors, with oxygen being the final electron acceptor (aerobic organisms), induced the transfer of protons across the membrane, which are in turn used to drive ATP production (Alberts, 2002) (Figure 1-1). Anearobic organisms use instead of oxygen other compounds such as nitrate as electron acceptors. Light energy is an additional source for producing proton gradients e.g. in the thylakoid (Alberts, 2002). Photosystem II, and the cytochrome b6f complex residing in the thylakoid membrane of the organelle, transfer protons into the chloroplast. The proton gradient then drives ATP synthesis (Alberts, 2002).

Ion gradients across the cellular membrane are facilitated by active transporters that transport ions against their chemical gradient. The sodium gradient is maintained by the sodium-potassium-exchanger, or Na^+/K^+-ATPase, transporting three sodium ions outside of the cell and importing two potassium ions under ATP consumption (Skou, 1957). This generates a sodium electrochemical gradient ($\Delta\tilde{\mu}Na^+$) across the membrane, and a negative membrane potential on the cytoplasmic side. Since the charged ions cannot freely pass the membrane, they have to diffuse passively through dedicated channels, or are transported actively using an energy source such as ATP, as described for the Na^+/K^+-ATPase (Skou, 1957).

Passive or 'downhill' transport of sodium, calcium or potassium via channels, for instance, are important mechanisms in neuron response (Voglis and Tavernarakis, 2006).

Besides the facilitative, or passive, transport of sodium and other ions, their electrochemical gradient also provides an energy source to take up important molecules for cell growth, and the ability to remove toxic molecules from within the cell. Maintenance of a constant pH and ion composition of the cytosol is thus necessary for the longevity of the cell (Baker-Austin and Dopson, 2007; Slonczewski *et al.*, 2009; Krulwich *et al.*, 2011).

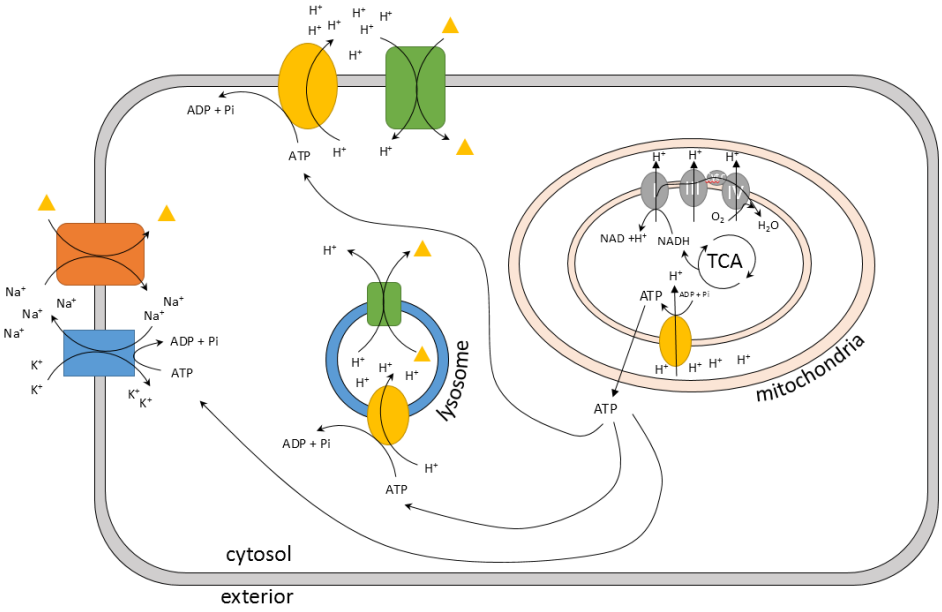


Figure 1-1: Pathway of nutrition uptake in the eukaryotic cell and its energy sources

Protons are transferred accros the membrane by the respiratory complex (grey). The generated proton motive force is in turn used to produce ATP via ATPases (yellow). The ATP produced is used in a whole range of cellular processes e.g. protein and DNA synthesis but also to generate a proton electrochemical gradient accross the lysosomal membrane via hydrolysis of the ATP to ADP. The energy of the hydrolysis is used for uphill transport of protons via ATPases. In the lysosome and the cell membrane, transporters (green) are localised that use the proton electrochemical gradient across the membrane to drive uptake of small molecules (yellow triangles) used for signalling, protein synthesis etc. In addition to ATP driven proton pumps, ATP hydrolysis also drives the sodium potassium exchanger (blue), generating a sodium electrochemical gradient between the exterior and the cytoplasm which is used by sodium-coupled transporters (orange) to drive uptake of nutrients and other small molecules.

A specific class of membrane proteins, the so-called secondary-active transporters, use the electrochemical gradients of protons or sodium ions, which vary depending on the cell type (*Table 1-1*), or that of the substrate to drive substrate transport across the membrane. It is this type of transporter to which the APC superfamily belongs. Secondary-active transporters show a variety of folds, substrate specificities and coupling mechanisms (reviewed in Boudker and Verdon, 2010).

One important substrate class are amino acids, fulfilling a variety of functions besides their more obvious role as protein building blocks, such as acting as carbon and nitrogen source, and their involvement in important cellular signalling and metabolic pathways (reviewed in Wu, 2009).

Table 1-1: Membrane potentials and pH gradients for different cellular membranes

Cell type	Membrane potential[1]	ΔpH[1]	reference
Cell membrane (eukaryotes)	-40 to -95 mV	< 0.1	(Hille, 2001)
Cell membrane (neutralophile bacteria)	-170 mV	+0.6	(Krulwich *et al.*, 2011)
Cell membrane (acidophile bacteria)	-256 mV	4.5	(Krulwich *et al.*, 2011)
Cell membrane (alkaliphile bacteria)	-50 mV	-2.2	(Krulwich *et al.*, 2011)
Neuron	-60 to -70 mV	+0.3 to 0.1	(Alberts, 2002), (Ruffin *et al.*, 2014)
Mitochondrion	175 mV	-0.8	(Alberts, 2002)
Lysosome	~20 mV	+2.5	(Casey *et al.*, 2010)
Chloroplast	50 mV	-0.8	(Alberts, 2002)

[1] potential- and pH-differences are given with respect to the cytoplasm

1.2 AMINO ACIDS AND THEIR REGULATORY ROLE IN THE CELL

Amino acids consist of a carboxyl and amino functional group and a side chain. The latter gives the amino acids their distinct properties by which they can be classified by polarity, charge, aromatic, branched, aliphatic, cationic, polar non-charged, anionic or the side chain contains sulphur groups (Figure 1-2). The wide range of different properties possessed by biological amino acids shows that they are a perfect tool kit for the generation of proteins with different functions (Doig, 2017). As a result, a variety of transporters have evolved that transport specific amino acids by recognizing their unique side chain properties. Some amino acid transporters have evolved to show a broad substrate spectrum while others are rather specific. Thus, the binding site of these transporters might be limited in its size to only transport small amino acids, or provide an electrochemical environment suitable for only positively or negatively charged compounds.

Besides being the structural unit of proteins, amino acids can act as an energy source for the cell upon breakdown of proteins, or act as metabolic regulators and signalling molecules.

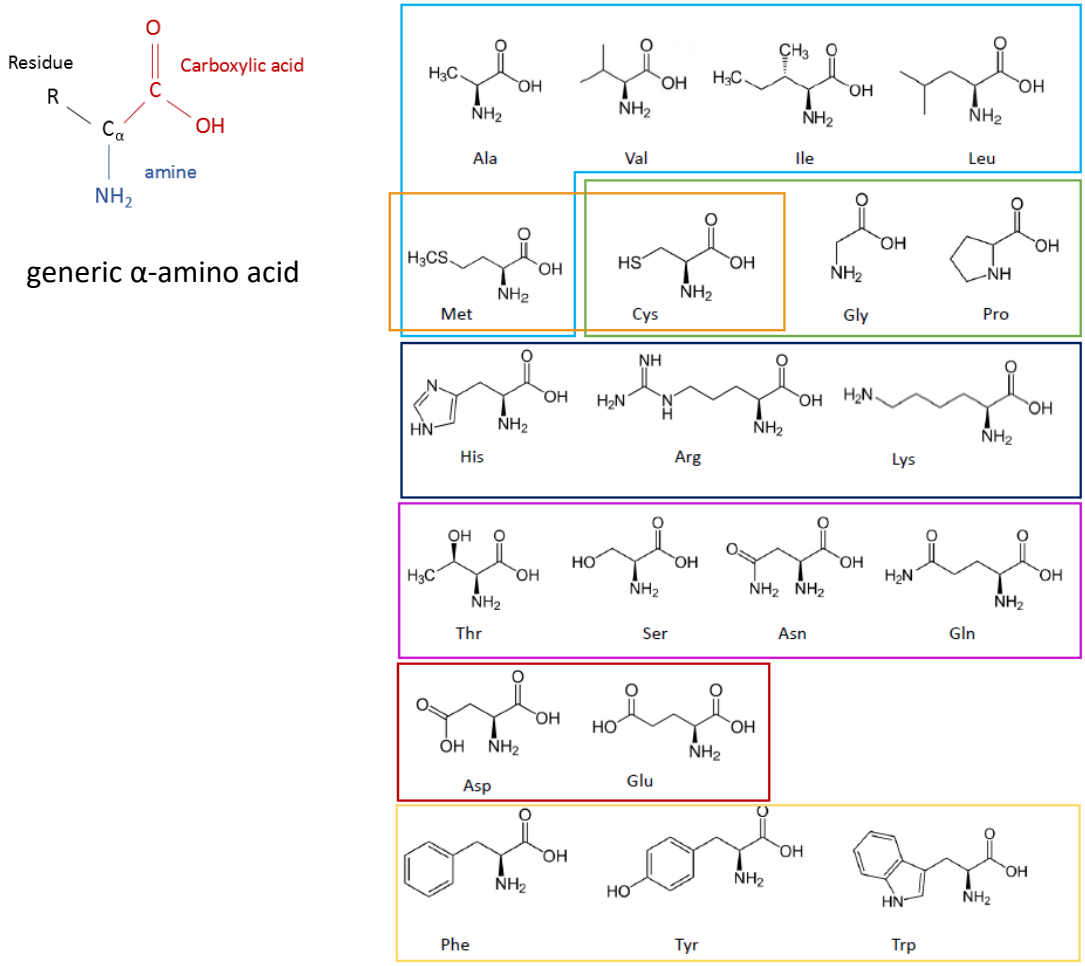


Figure 1-2: Properties of amino acids

Proteinogenic amino acids are α-amino acids with a carboxyl and amine group and a side chain (residue) attached to the Cα- atom (top left panel). The side chain gives each amino acid its specific property. They can be categorised by being negatively charged (red), positively charged (dark blue), aromatic (yellow), polar non-charged (pink), aliphatic (light blue), contain sulphur (orange) or have special side chains (green)

Amino acids can be derived through the degradation of proteins in special organelles, such as the lysosome (de Duve and Wattiaux, 1966), but they can also be produced biosynthetically. Microbes and plants can synthesise all proteinogenic amino acids, whereas humans and animals are unable to produce some of them. Therefore, a discrimination must be made between amino acids that the body can produce itself, referred to as non-essential amino acids, and the ones that have to be derived from diet, referred to as essential amino acids.

In humans, these nonessential amino acids are Ala, Arg, Asp, Asn, Cys, Gly, Glu, Gln, Pro, Ser, Tyr and the essential amino acids are His, Ile, Leu, Lys, Met, Phe, Thr, Trp, Val.

Although it is possible to biologically synthesise each amino acid (Figure 1-3), some metabolic pathways require many steps to reach the final product. For example, synthesis of methionine from aspartate involves 7 individual stages for which different enzymes are necessary (Figure 1-3). Thus, errors in translation and transcription of the genes encoding these enzymes result in a loss of function and the potential death of the cell. Some of these pathways have lost the ability to produce certain amino acids since the human diet includes a variety of amino acids. These are the essential amino acids that must be taken in via diet.

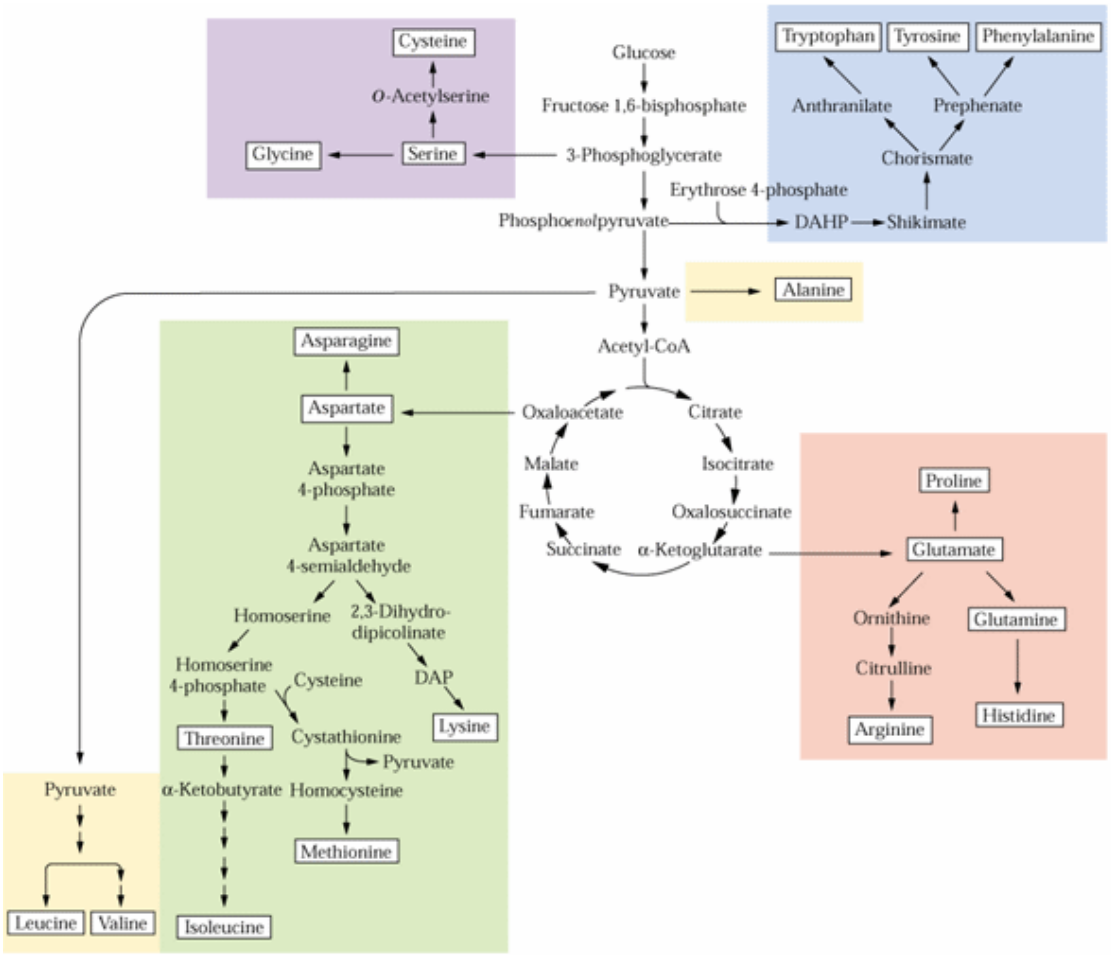


Figure 1-3: Amino acid synthetic pathway

Proteinogenic amino acids can be biosynthesised in specific pathway which are shown here (Hildebrand, 2016). Depending on their chemical origin, the pathways are coloured differently. There are leucine, valine and alanine that are derived from pyruvate (yellow box), others originate from oxaloacetate (green box), 3-phosphaglycerate (purple box), phosphoenolpyruvate (blue box) or α-ketoglutarate (red box). Synthesis of some amino acids includes several enzymatic steps, for example the production of isoleucine or methionine.

Examples of nonproteinogenic amino acids that play an important role in metabolic processes within an organism are ornithine, citrulline and 5-aminolevulenic acid, to name a few. Ornithine and citrulline play an essential for the urea cycle (Morris, 2002), a metabolic pathway to dispose of excess ammonia originating from amino acid catabolism which would otherwise be toxic to the organism in high doses. Citrulline is generated from ornithine and carbamoyl phosphate in the mitochondrion (Morris, 2002). An exchanger for citrulline and ornithine resides in the mitochondrial membrane to move ornithine into the mitochondrion, and citrulline out of it (Camacho *et al.*, 1999), so that the urea cycle can continue in the cytosol. Another example of a non-proteinogenic amino acid with an important metabolic function is 5-aminolevulenic acid. This organic molecule is the precursor for porphyrin synthesis which produces heme (mammals) or chlorophyll (plants) (Vavilin and Vermaas, 2002). Besides the metabolic importance of amino acids, they fulfil a range of other functions such as gene regulation, or signalling (Wu, 2009), which will be described in the next section in more detail, giving some examples of their influences.

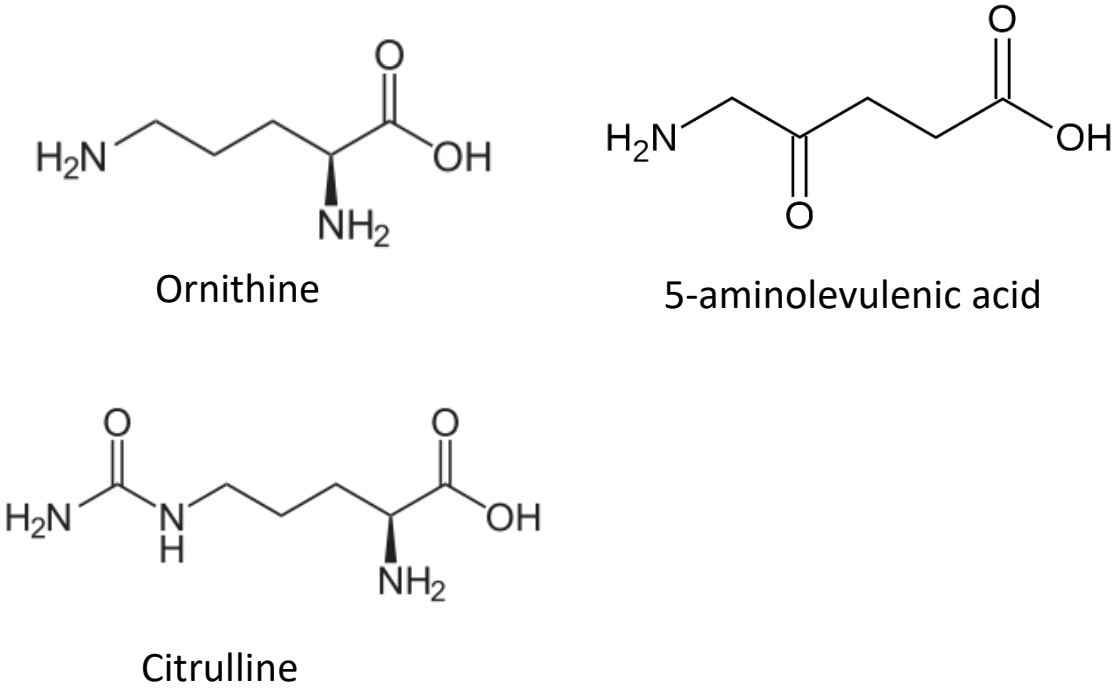


Figure 1-4: Chemical structures of ornithine, citrulline and 5-aminolevulenic acid

1.2.1 Amino acids as signalling molecules

Amino acids are not only the building blocks of proteins but also function as signalling molecules or their precursors. They are involved in a vast spectrum of regulatory and signalling pathways in the cell such as cell proliferation via the mTORC nutrient sensing complex, as neurotransmitters in the synaptic cleft or host defence mechanisms (reviewed in Henley, 2001 and Wu, 2009).

1.2.1.1 Amino acids and their derivatives as neuronal signalling molecules

The central nervous system is the key to understanding the world around us and our interaction with it. All important processes are regulated from this organ. The communication between individual neurons, which store and process information, is essential and is facilitated by neurotransmitters such as glutamate, serotonin, dopamine, glycine or γ-aminobutyric acid which activate ligand-gated ion channels and G-protein coupled receptors (GPCRs) (Alberts, 2002). A neurotransmitter can either excite or inhibit postsynaptic signalling cascades (Spencer and Baker, 1992). Neurotransmitters are stored in vesicles within the neuron and are released via calcium-induction into the synaptic cleft to be recognised by specific receptors, starting a signalling cascade in the receiving neuron (Figure 1-5) (Jessell and Kandel, 1993).

Glutamate is such a signalling molecule, and one of the most abundant amino acids (Newsholme *et al.*, 2003). It is released into the synaptic cleft and interacts with glutamate receptors on the receiving neuron. This generates an excitatory electrical response via activation of ion channels that reside in the membrane of the receiving neuron, changing the ion composition on the cytoplasmic side (Jessell and Kandel, 1993). The sudden change in ion concentration reverses the electrical polarisation of the membrane, creating an action potential that propagates along the axon of the neuron (Jessell and Kandel, 1993). The neurotransmitter is removed from the synaptic cleft via specific transporters, so that it does not continue to stimulate the receptors. Reuptake of glutamate is performed by members of the excitatory amino acid transporter (EAAT) family which belong to the SLC1 family of transporters in humans (Figure 1-5) (Fairman *et al.*, 1995; Slotboom *et al.*, 1999).

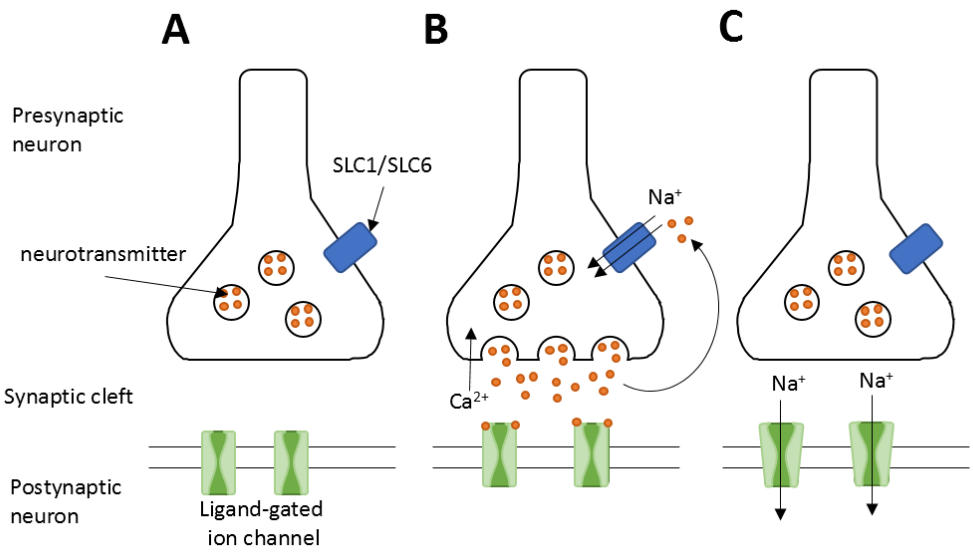


Figure 1-5: Electrical transmission at the synaptic cleft

Electrical signalling between the two neurons, schematically shown here, occurs in the gap junction or synaptic cleft. The neurotransmitter (small signalling molecule, orange circles) is released from the signalling neuron into the synaptic cleft, activating ion channels (green) and generating an action potential that traverses through the receiving neuron towards the next receiving neuron. Reuptake of the neurotransmitter is facilitated by the sodium coupled neurotransmitter transporters of the SLC1 and SLC6 family (blue).

Amino acids can also act as precursors for neurotransmitters. Dopamine, derived from phenylalanine, tyrosine and L-DOPA (Figure 1-6A), is associated with drug addiction (Wise, 1996). Dopaminergic pathways in the brain release dopamine and are mostly associated with reward-related behaviour, which is triggered by addictive drugs (Wise, 1996). Disorders such as attention deficit hyperactivity disorder, schizophrenia, or Parkinson's disease are related to dysfunctions within these dopaminergic pathways (Barbeau, 1970). Serotonin, also known as 5-hydroxytryptamine or 5-HT, is a tryptophan derivative (Figure 1-6B). This neurotransmitter is a regulator of sleep, appetite and mood via activating signalling cascades by interacting with serotonin receptors, G-protein-coupled receptors (GPCRs), and ligand-gated ion channels in the postsynaptic membrane (Jessell and Kandel, 1993). The effect on the neurotransmission is either excitatory or inhibitory. The serotonin transporter (SERT), which is a member of the SLC6 or neurotransmitter sodium symporter (NSS) family, is responsible for the reuptake of serotonin from the synaptic cleft.

In general, several neurological conditions are associated with neurotransmitter transporter malfunction such as depression, anxiety disorder, epilepsy and Parkinson's disease (Masson

et al., 1999). Drug development to treat these diseases targets the modulation of the neurotransmitter transport activity (Iversen, 2009). Structural information available for these transporters has elucidated their binding sites (Penmatsa *et al.*, 2013; Coleman *et al.*, 2016) which will further guide structure-based drug design.

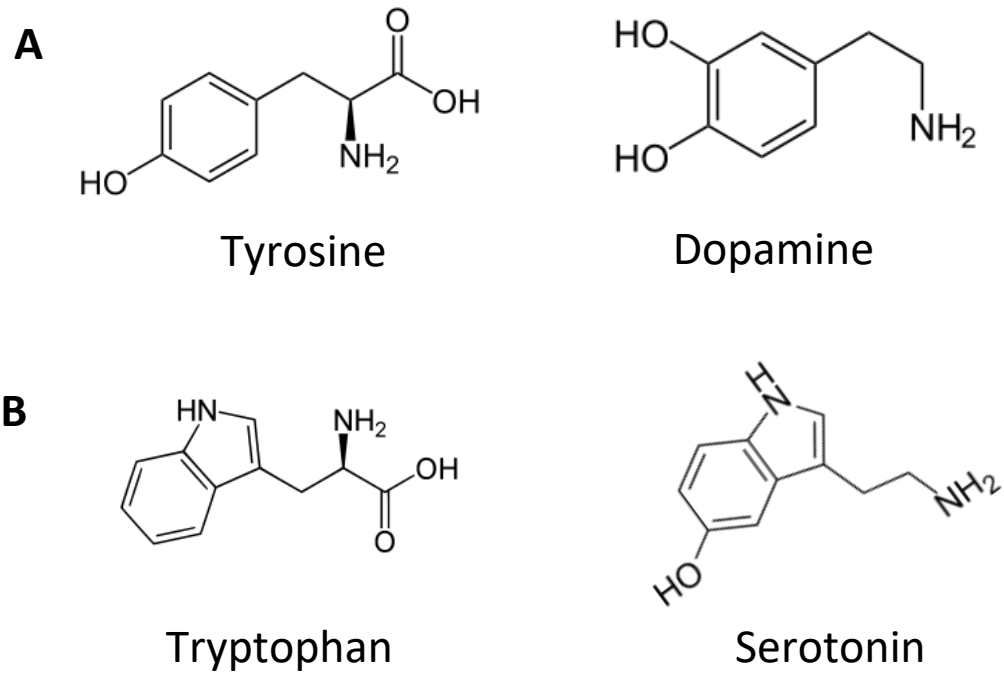


Figure 1-6: Chemical structures of amino acids, tyrosine and tryptophan, and their neurotransmitter derivatives, dopamine and serotonin

1.2.1.2 mTORC regulation by amino acids

The availability of nutrients is one key regulator of cell growth and cell division (Efeyan *et al.*, 2015). As this parameter is ever changing in our environment, signalling pathways have evolved that sense and respond to the varying pool of available nutrients (Efeyan *et al.*, 2015). Part of a nutrient sensing complex, fulfilling this central function, is the target of rapamycin (mTOR). mTOR is a serine-threonine kinase that forms distinct complexes with other regulatory proteins to either form mTORC1 or mTORC2 (Loewith *et al.*, 2002). mTORC1 stimulates cell growth while mTORC2 regulates cell proliferation (Loewith *et al.*, 2002), and inhibition of mTORC1 results in a decreased life span of mice (Harrison *et al.*, 2009). Furthermore, restricting the calorie consumption of Rhesus monkeys decreases the chance of them developing cancer, which shows the relation between diet and mTORC1

activity (Colman *et al.*, 2009). Amino acids stimulate mTORC1 via the amino acid/Rag pathway (Sancak *et al.*, 2008; Efeyan *et al.*, 2012), however, the exact mechanism of this activation is not yet known, and little has been elucidated as to which amino acids are responsible for the activation. It is known that leucine and arginine are important for mTORC1 activation but show no effect on mTORC1 activity when the cell is deprived of the other proteinogenic amino acids (Hara *et al.*, 1998). Additionally, amino acids do not only regulate the activity of the complex, but also its subcellular localisation. In the presence of amino acids, mTORC1 translocates from the cytoplasm to the lysosomal membrane (Efeyan *et al.*, 2012).

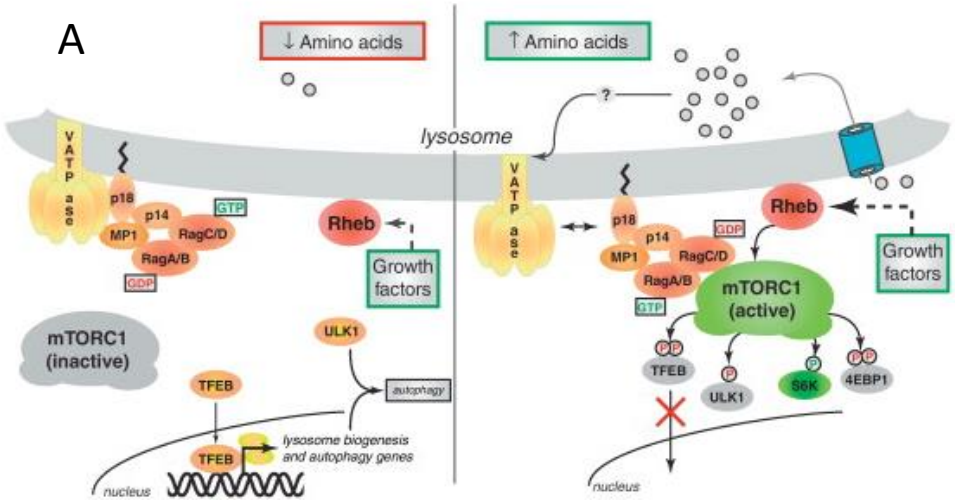


Figure 1-7: mTORC1 signalling via amino acid stimulation

In simple terms, a high availability of amino acids result in the activitateion of mTORC1 which induces protein and lipid synthsis amon other regulatory pathways. At low abundance of amino acids, the mTORC1 complex is inactivated. In detail, at low amino acid concentrations, the inactive conformation of the v-ATPase-Regulator-Rag GTPase complex (vATPase, p18, p14, MP1, RagA/B, RagC/D) cannot interact with mTORC1. Therefore, mTORC1 remains in the cytoplasm. Increasing concentrations of amino acid signal to the Rag GTPases via the v-ATPase. The Rag GTPases switch their loaded nucleotide (GDP to GTP and *vice versa*) resulting in their activation. The active Rag GTPases can now interact with mTORC1 and localise the complex to the lysosomal membrane. The small GTPase Rheb (Ras homologue enriched in the brain), regulated by growth factors, activates the kinase activity of mTORC1, phosphorylating several targets (TFEB, ULK1, S6K, 4EBP1) regulating protein synthesis, autophagy or lipid synthesis. Figure from Efeyan *et al.*, 2012.

It is proposed that a nutrient sensor activates mTORC1, although it is yet unknown which protein fulfils this function. Among the potential candidates are the proton-assisted amino acid transporters (PATs) of the SLC36 family. These transporters are also phylogenetically clustered within the APC superfamily and are predicted to have the same '5 + 5' inverted topology repeat fold as LeuT. The PATs are located in several places within the cell, not only at the plasma membrane but also the lysosomal and endosomal membrane (Sagne *et al.*, 2001). The human PAT1 and PAT4 have been identified to be required for activation of the

mTORC1 complex (Heublein *et al.*, 2010; Ögmundsdóttir *et al.*, 2012). Since the lysosome is the location of protein degradation, the proton electrochemical gradient across the lysosomal membrane is used by PAT1 or PAT4 to export amino acids coupled to proton movement into the cytosol thereby activating mTORC1. Additionally, the sodium-dependent amino acid transporter family, SLC38, has been identified as another regulator and key component of the amino acid sensing machine of mTORC1. Loss of SLC38A9 resulted in the impairment of amino acid activation of mTORC1 (Rebsamen *et al.*, 2015). This shows the importance not only of amino acids but also of their transporters in regulating the nutrient sensing pathways in the cell.

1.2.1.3 NO-signalling and the role of arginine transport

Nitric oxide is crucial for the pathways of the immune response. It is generated by immune cells such as macrophages, dendritic cells, mast cells, monocytes and many other cell types via three systems: the inducible nitric oxide synthase (iNOS), neuronal nitric oxide synthase (nNOS) and endothelial nitric oxide synthase (eNOS) (Forstermann and Sessa, 2012). Both nNOS and eNOS exist as preformed proteins in the cell whereas iNOS production is only regulated by cytokines (Bogdan, 2001; Forstermann and Sessa, 2012). Interestingly, at low levels the production of iNOS is upregulated via a positive feedback loop, but down regulated at high protein concentrations, probably to prevent overproduction of NO (Forstermann and Sessa, 2012). Additionally, other signalling pathways participate in the regulation of iNOS production, which makes the regulatory network of iNOS incredibly versatile. Another factor that influences the availability of NO in humans is the substrate of nitric oxide synthetase (NOS), the amino acid arginine, which can be derived in three ways:

1) dietary intake, which in humans is around 4-5 g/day (King *et al.*, 2008)
2) protein digest, performed in specialised compartmented organelles such as the lysosome (Liu *et al.*, 2012)
3) arginine biosynthesis from glutamine, ornithine or citrulline in compartmentalised tissues (Morris, 2007)

Since arginine cannot diffuse freely through the membrane into the cell, it has to be carried over the membrane by specific transporters. This is achieved by the SLC7 cationic amino

acid transporters (CATs) which exclusively transport cationic amino acids. Proteins in this family are widely expressed in all tissues, with specific expression patterns for each member. In macrophages, CAT2 is the transporter providing arginine, and it has been shown to be activated by treatment with lipopolysaccharides (Kakuda *et al.*, 1999) which simulate infection by bacteria. CAT2 mRNA levels are regulated by IFN-gamma (Kakuda *et al.*, 1999). There is another pathway of arginine metabolism which opposes that of NO production. The enzyme arginase can break down arginine to ornithine which is then used for protein synthesis. Depending on the preference of the macrophage for one pathway or the other, they can be defined as M1 (NOS) or M2 (arginase) macrophages (Mills *et al.*, 2000; Mantovani *et al.*, 2004).

Further downstream of the arginine metabolic pathway, spermine and spermidine are produced that downregulate the production of iNOS (Szabó *et al.*, 1994). NO signalling influences numerous processes which range from the production of cytokines (Eun *et al.*, 2000; Lecleire *et al.*, 2005; Macháček *et al.*, 2016) and vasodilation, to the stimulation of light flashing by fireflies (Trimmer, 2001).

NO's action in host defence mechanisms by macrophages result from its inhibitory effect on DNA repair processes and protein synthesis in the invading parasite. Inactivation of enzymes and the disruption of the parasites membrane via peroxidation, and reduction of exogenous arginine through induction of iNOS or arginase by the host immune cells, can lead to the death of the parasite. NO additionally influences tumour cells. It has been shown that NO producing macrophages mediate cytostasis and tumour cell death (Xu *et al.*, 2002). In contrast, indogenous NO production in tumour cells results in the upregulation of DNA damaging repair pathways. This mechanism protects the tumour cells against endogenous NO product, which would be toxic to them, and additionally against DNA damage, which might be induced by tumour therapy.

The underlying principles of NO signalling requires the availability of its precursor arginine in the cytoplasm of macrophages. As described above, the CAT transporters of the SLC7 family are one of the main regulators for arginine availability in these cells and any malfunction would result in an imbalance of the intracellular arginine concentration which would in turn possibly have a negative impact on the downstream signalling pathways such as host defence via NO.

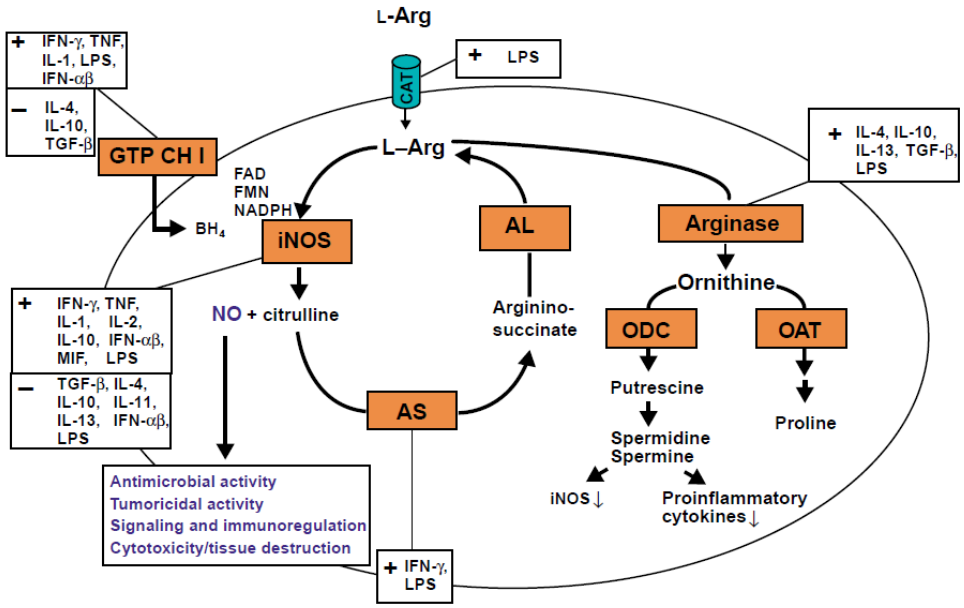


Figure 1-8: Regulation of NO signalling

Arginine is metabolised via two systems, iNOS or arginase. The latter uses arginine to produce ornithine, a precursor for proline and other cytokines that down-regulate iNOS (shown by arrows). iNOS uses arginine to produce nitric oxide and citrulline. Citrulline can be recycled to produce arginine via arginine synthase (AS) and AL. Inhibition and activation of the different enzymes involved in the pathways above are given with a negative or positive sign, respectively. ODC-ornithine decarboxylase, OAT - ornithine aminotransferase, AL - arginosuccinate lyase, GTP CH I – guanidine triphosphate cyclohydrolase I, IFN-γ – interferon gamma, FAD - flavin adenine dinucleotide, IL – interleukin, LPS – lipopolysaccharide, TGF-ß – transforming growth factor beta, MIF - macrophage migration inhibitory factor, TNF – tumour necrosis factor, FMN - flavin mononucleotide, and NADPH - nicotinamide adenine dinucleotide phosphate. Figure from Bogdan, 2001.

1.2.1.4 Secondary active amino acid transporters as a key component in amino acid signalling

The amino acid signalling pathways described above show how important these small molecules are for cell survival, growth, host immunity or neuronal activity. Nevertheless, to reach the specific location to fulfil their function or to remove them to prevent further signalling, the amino acids have to be transported across the membrane. In the previous sections, the transporters which are involved in these processes have been described. As mentioned in section 1.1, secondary-active transporters exist that use an electrochemical gradient to drive “uphill” substrate transport (reviewed in Forrest *et al.*, 2011). These transporters are involved in amino acid uptake in the cellular signalling processes described above, and are found in all kingdoms of life. The transport of the substrate and ion can be facilitated in the same direction, also known as symport, in opposite directions, named antiport, or without any ion-coupled movement, uniport (Figure 1-9).

Mostly, gradients of sodium ions or protons are used as the energy source for coupled transport, of which the latter is found predominantly in prokaryotes, whereas sodium-coupled transporters are mostly present in eukaryotes (Saier, 2000). Proton-driven transporters in eukaryotes are mainly located in tissues or organelles with high pH, such as the late endosomes, lysosomes, vacuoles and the thylakoid lumen (Table 1-2).

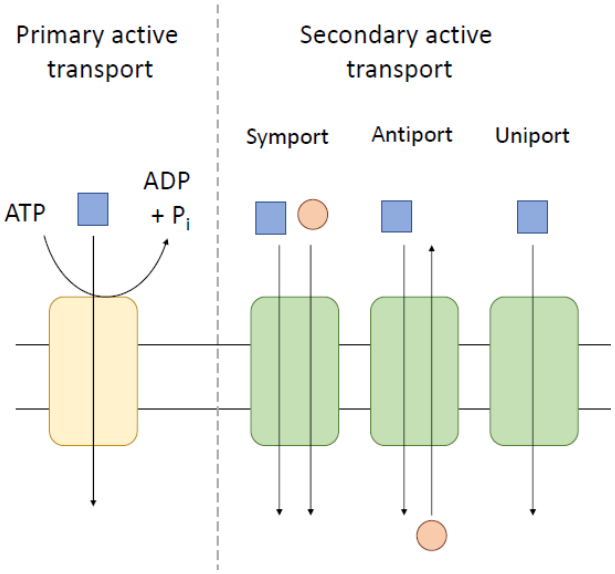


Figure 1-9: Types of active membrane transporters

Uphill transport of biomolecules needs an energy source. In all kingdoms of life, this energy source can either be derived by the hydrolysis of ATP to ADP (left) through a primary-active transport mechanism, or the translocation of the molecule is coupled to the downhill transport of an ion (right), called secondary active transport. The latter are divided into symport, antiport and uniport mechanism, depending on the direction of transport of substrate (blue square) and ion (red circle) and if any coupled transport occurs.

Table 1-2: pHs of different cellular compartments

Environment	pH
Cytosol	7.2*
Lysosome	4.7*
Vacuole	5.2**
Early endosome	6.3*
Late endosome	5.5*
Thylakoid lumen	5.0**

* in mammalian cells (Casey *et al.*, 2010)
** in *Arabidopsis thaliana* cells (Shen *et al.*, 2013)

Acidic environments are less prevalent in higher eukaryotic organisms, which might explain why sodium-coupled transporters are more dominant there. This is also the case for halophilic bacteria that have adapted to life in salty environments as their evolutionary niche. In

other prokaryotes, proton-coupled transport is the main source for the uptake of nutrition such as uptake of amino acids via the APC superfamily, transporters in the membrane, or small peptides via the major facilitator superfamily (MFS). These secondary active transporters share a common feature, the internal symmetry of their helical bundles (Figure 1-10), although different folds that have evolved to pursue a similar function, the transport of molecules across the membrane.

The common MFS transporters consist of 12 transmembrane helices (TM) with an internal '6+6' symmetry (Figure 1-10A). One of the first structurally studied examples is LacY (Abramson, 2003). Based on the different conformations in which LacY was captured, a previous hypothesis about the alternating access model of transporters (Jardetzky, 1966) was confirmed: the substrate binding site is accessible from one side of the membrane while inaccessible from the other during the transport cycle. Further investigation of other secondary active transporters revealed the other novel folds that are known to date (Figure 1-10). The crystal structure of the sodium-coupled amino acid transporter LeuT, showed a distinct fold from the MFS family and other proteins that share this three-dimensional helical arrangement, were also classified as part of the LeuT-fold family. Other folds that were observed include the NhaA fold, found for the Na^+/H^+-exchanger (Hunte *et al.*, 2005), the ATP/ADP carrier fold (Pebay-Peyroula *et al.*, 2003) and the Glt_{Ph} fold of the glutamate transporter of the EAAT family (Yernool *et al.*, 2004) (Figure 1-10).

Among the extensively studied groups of secondary active transporters, besides the MFS family, is the APC superfamily of which members share a similar fold to that of LeuT. They show a vast spectrum of transported substrates such as amino acids, sugars, neurotransmitters and other metabolic products.

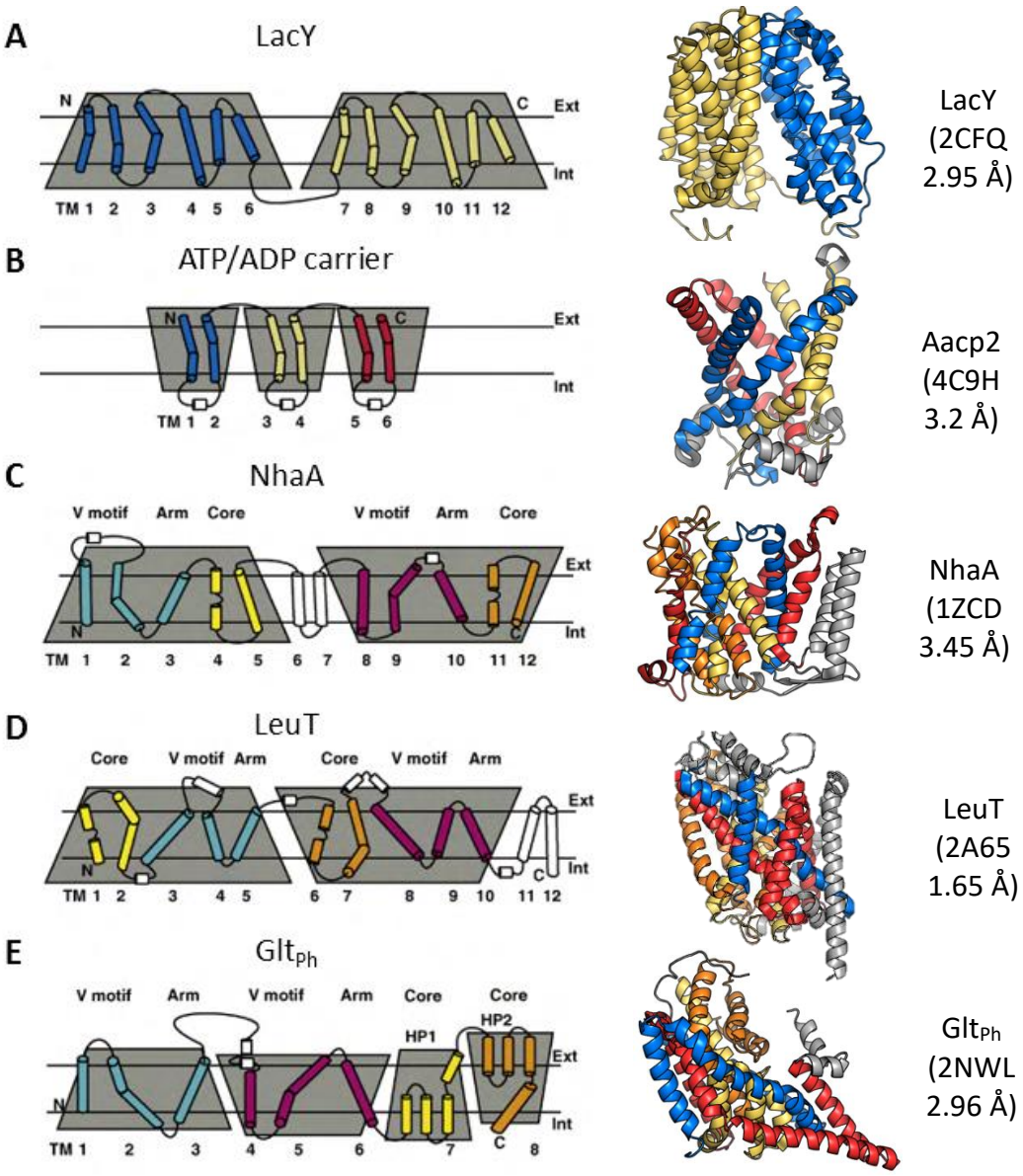


Figure 1-10: Topology and folds of different secondary active transporters

Secondary active transporters can be categorised by their different folds that have been determined from structural information obtained previously. What all the transporters have in common is an internal structural symmetry which is not reflected by high sequence similarity of the symmetrical counterparts. The different folds found are the LacY-fold (A) which the MFS transporters share, the ATP/ADP carrier fold, (B), the NhaA fold (C), the LeuT fold of the APC superfamily and the Glt$_{Ph}$ fold (E). Symmetrical parts are highlighted by the trapezoids. The N- and C-terminal core regions are coloured in blue and red and the other repeats are coloured in yellow and orange. Secondary structure that is not part of the repeats is coloured in white. For each folding family, a representative crystal structure is shown on the right. Figure adapted from Boudker and Verdon, 2010.

1.3 THE AMINO ACID-POLYAMINE CATION SUPERFAMILY

To fully understand a membrane transporter, obtaining information about the substrate binding site, ion binding and the conformational changes during the transport cycle is crucial.

Identification of the substrate binding site can be derived by structural data via a co-crystal structure of the transporter with a substrate. These data help to understand the specificities of the transporter and how they relate to other transporters within a protein family. The latter can be verified by mutational studies trying to modulate the substrate spectrum.

The structural data can also help to identify ion-binding sites for example for the sodium-coupled transporters, although this approach is difficult when proton-coupling is the mechanism used by the protein.

Capturing the transporter in different conformations via X-ray crystallography or performing molecular dynamics simulations can further shed light into the transport mechanism.

In general, the combination with structural, mutational and functional studies can however help to determine residues involved in ion-coupling and substrate binding. This combinatory approach has been successfully used on the APC superfamily to deepen the understanding on how they achieve ion-coupled translocation of their substrates, and how they can discriminate between the wide range of available organic molecules.

The APC superfamily is, after the MFS family, one of the largest protein families of membrane transporters (Reddy *et al.*, 2012; Wong *et al.*, 2012). Over the past few decades, many more members of this family have been identified (Jack *et al.*, 2000; Wong *et al.*, 2012; Vastermark *et al.*, 2014). Strikingly, members from different subfamilies show low sequence identity (< 20 %), explaining the difficulty to date in discovering potential members, and requiring the development of new bioinformatics tools (Vastermark *et al.*, 2014).

The APC superfamily is subdivided into several subfamilies (Figure 1-11) that share similar sequences and substrate specificities within one family. Members of this superfamily do not only transport amino acids but also other small molecules (Table 1-3). The NSS family transports neurotransmitters, such as dopamine and serotonin (Androutsellis-Theotokis *et al.*, 2003) in a sodium-coupled manner. One extensively studied member of this family is LeuT,

a sodium coupled alanine transporter which is also specific for leucine. Its functional similarity to the eukaryotic NSS (SLC6) transporters was the basis for understanding neurotransmitter transport at the synaptic cleft and the mechanism of antidepressants on the activity of these transporters. This was mainly elucidated via *in vitro* functional experiments together with mutational studies and co-crystallisation of this transporter with specific antidepressants (Zhou *et al.*, 2007). The APC family is a broad specific family of amino acid transporters and consists of exchangers, proton-coupled and sodium-coupled transporters (Table 1-3). Several structures, particularly of prokaryotic APC transporters, have been solved for different members of this family (Table A 1) with the transporters captured in different conformations and with different substrates in their binding sites. Extensive research on these proteins has enabled their transport mechanism to be understood, as well as establishing and verifying the translocation pathway, observing novelties between each member, and obtaining information on similarities and differences within the subfamilies. Knowledge of the prokaryotic APC transporters has been used to understand their human homologues due to their importance in the signalling pathways in the brain. This research first shed light on the human APC transporters before the first structural data could be collected on them, given the difficulty of obtaining stable and pure protein for these kinds of studies.

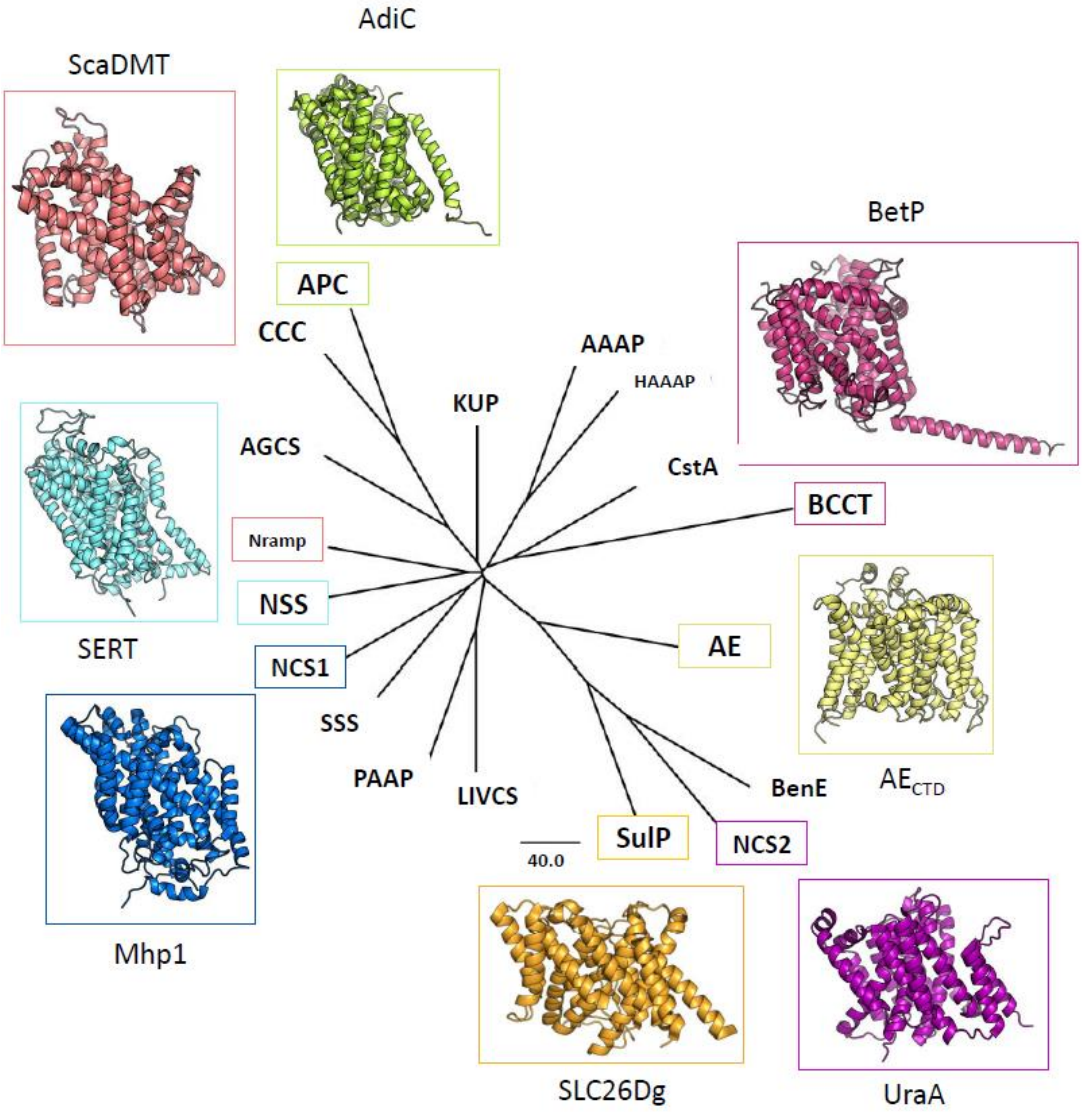


Figure 1-11: Phylogenic tree of the APC superfamily

The phylogenic tree of the APC superfamily based on their protein sequences and was generated based on a vigourous BLAST search with subsequent SuperFamilyTree generation and clustering, as described by Vastermark *et al.*, 2014. The APC superfamily consist of several sub families that show a different degree of evolutionary relationship with one another, as shown by the phylogenic tree. For a subset of them, structural information is available. Representative structures of these families are shown above, labelled with their protein name and highlighting to which family they belong. APC – Amino Acid-Polyamine-Organocation Family, CCC - Cation-Chloride Cotransporter Family, AGCS – Alanine or Glycine:Cation Symporter Family, Nramp - Metal Ion (Mn2+-iron) Transporter Family, NSS – Neurotransmitter:Sodium Symporter Family, NCS1 - Nucleobase:Cation Symporter-1 Family, NCS2- Nucleobase:Cation Symporter-2 Family, SSS - Solute:Sodium Symporter Family , PAAP - Putative Amino Acid Permease Family , LIVCS - Branched Chain Amino Acid:Cation Symporter Family, SulP - The Sulfate Permease Family, BenE – Benzoate:H+ Symporter Family, AE- Anion Exchanger Family, BCCT - Betaine/Carnitine/Choline Transporter Family, CstA - Peptide Transporter Carbon Starvation CstA Family, HAAAP - Hydroxy/Aromatic Amino Acid Permease Family, AAAP - Amino Acid/Auxin Permease Family and KUP - K+ Uptake Permease Family. Figure adapted from Vastermark *et al.*, 2014.

Table 1-3: Transport properties of the APC superfamily members from www.tcdb.org

TC family	SLC family	Substrates	Coupling mechanism
APC	SLC7	amino acids	H^+, exchangers, uniport
BCCT		betaine, carnitine, choline	H^+, Na^+, exchanger, uniport
AAAP	SLC36, SLC38	auxin, amino acids	H^+, Na^+
SSS	SLC5	sugars, amino acids, organocations, nucleosides, vitamins, inositols, urea, anions	Na^+
NSS	SLC6	neurotransmitters, amino acids, osmolytes	Na^+, Cl^-
AGCS		alanine, glycine	H^+, Na^+
CCC	SLC12	Na+, K+	Cl^-
NCS1		purines, pyrimidines and related metabolites	H^+
NCS2	SLC23	nucleobases, purines, pyrimidines	H^+, Na^+
HAAAP	SLC32	amino acids	H+
SulP	SLC26	sulphate, chloride,	H^+, HCO_3^-,
PAAP		alanine	Na^+
LIVCS		aliphatic, hydrophobic amino acids (leucine, isoleucine, valine)	H^+, Na^+
NRAMP		divalent metal cations (Mg2+, Fe2+)	H^+
CstA		peptides	-
KUP		K+	H^+
BenE		small aromatic compounds	H^+
AE	SLC4	HCO_3^-, Cl^-	Na^+/ H^+-dependent exchangers

1.3.1 Topology and fold

A common feature shared by the APC superfamily is the internal pseudo two-fold symmetry, since one portion of the core helical segment can be aligned with another portion of the protein by a 180° turn perpendicular to the membrane plane, also known as inverted-repeat topology (Figure 1-12). This property has also been observed for other transporter families and is thought to have its origin in a gene duplication (Keller *et al.*, 2014). The theory is that the transporter existed as a homodimer. Two identical monomers form this homodimer, related to each other by a two-fold symmetry (Keller *et al.*, 2014).

This in turn means that if both monomers originate from the same gene, both proteins were inserted into the membrane in the same orientation. To create a functional transporter, one monomer had to flip by 180° in the membrane. Another possibility is that both halves were translated sequentially and then inserted in opposite directions into the membrane. Gene duplication or gene translation of the transporter might thus have allowed a faster and more efficient insertion of a functional protein into the membrane (Khafizov *et al.*, 2010; Keller *et al.*, 2014).

Most members of the APC superfamily consist of 12 to 14 TMs of which the core regions show a symmetry relation also referred to as '5+5'-symmetry, but additionally a '7+7'-symmetry has been observed in other transporters (Västermark and Saier, 2014). When both symmetry-related segments are superimposed, the root mean square deviation (RMSD) between them is very small and ranges from 0.1 to 6 Å, although they share a very low sequence identity (< 20 %).

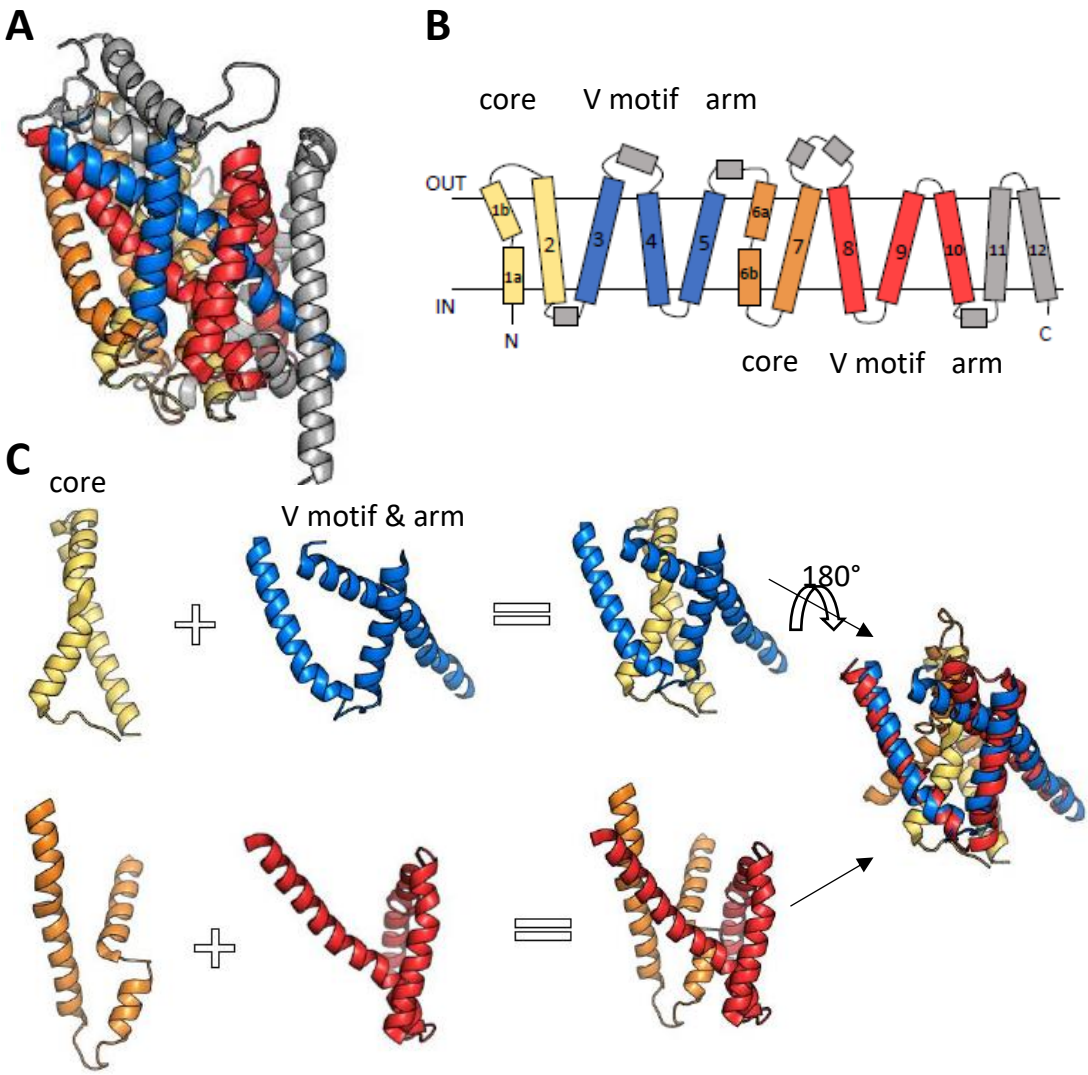


Figure 1-12: Symmetrical arrangement of helices in LeuT like transporter

Structural analysis of APC superfamily members revealed that the proteins consist of two to three helix bundles that are symmetric with respect to one another. Proteins with the LeuT-fold have a transmembrane core region of 10 helices, of which the first five helices and the second five helices are related by a pseudo two-fold symmetry. Additional N- or C-terminal helices are found. An example of an APC family member, LeuT, is shown here.
Side view of LeuT (2A65, 1.65 Å) embedded in membrane (A). The topology diagram shows the helices forming the core region (yellow and orange) and the V motif & arm region (blue and red) and the additional transmembrane helices 11 & 12 (grey) (B). One part of the core and V motif & arm domain form one symmetrical subunit of LeuT (C). Both symmetrical parts, consisting of 5 TMs, are structurally similar and can be superimposed by rotating one symmetrical half by 180° perpendicular to the membrane plane (C).

At the interface of two helices that cross in the centre of the protein, they are interrupted by an unwound section. The substrate binding site is located between these two sections which open a polar environment within the transmembrane domain of the transporter, due to the exposed carbonyl and amine groups (Figure 1-13). This polar environment enables

substrate and ion binding. The loop regions formed by the unwound sections contain important residues for substrate binding but also offer flexibility to orchestrate the conformational changes important for the transport cycle (Screpanti and Hunte, 2007). In each loop segment, highly conserved residues (particularly glycines) are located, indicating the importance of flexibility at this region of the protein.

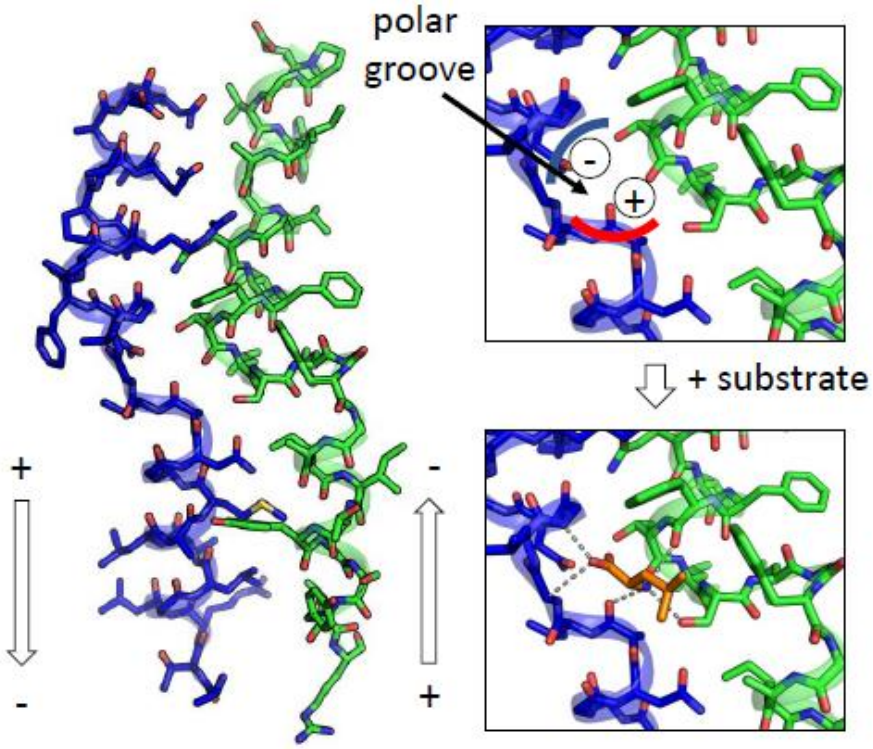


Figure 1-13: Unwound segments of the two discontinuous helices form one major part of the substrate binding site

The discontinuous helices TM1 (blue) and TM6 (green), here shown from a crystal structure of LeuT (2A65, 1.65 Å) have opposite dipole moments (left, arrows marked with negative and positive sign). The unwound segments of those helices provide a polar environment, as the backbone at this region does not form a hydrogen binding network with the other backbone atoms of the same helix (top right). The polar environment (indicated by the arrow) is necessary for the interaction with ions and the substrate, here leucine (orange), forming hydrogen bonds with side chain residues and the backbone of the discontinuous segments.

1.3.2 Transport mechanism

Before the first crystal structures of membrane transporters were published, theories were developed on how the protein could actually facilitate the translocation of the substrate across the membrane (Jardetzky, 1966). It was thought that the transporter was open to one side of the membrane bilayer in its substrate-free form, and that this state might be the energetically favoured conformation of the protein. Upon binding of a substrate and ion, conformational changes into an occluded state are induced. In this state, the substrate binding site would not be accessible from either side of the membrane. Further conformational rearrangements within the protein open the substrate binding site to the other side of the membrane so that the substrate could be released. Then, the transporter would return to its initial starting state. This cycle is known as the "alternating access" mechanism, and has now been supported by crystal structures that have captured different proteins in various states (Forrest *et al.*, 2008; Shimamura *et al.*, 2010; Kaback *et al.*, 2011; Weyand *et al.*, 2011; Perez *et al.*, 2012; Solcan *et al.*, 2012). Nevertheless, some details about how the actual movement of individual helices is facilitated remains unclear.

Since the transporter consists of two halves, there was the possibility that both halves move against each other (rocker-switch model) or that one half stays rigid while the other performs the major movements (rocking-bundle model). Interestingly, there seems to be no single unique transport mechanism which is common to all, or at least most, transporters (Table 1-4). The increasing number of conformations in which different transporters have been crystallised has revealed that even within the same protein family each transporter exhibits individual conformational changes.

Additionally, alternative transport models have been suggested: the "elevator mechanism" and the "toppling mechanism". Both mechanisms are based on information gained from intensive structural studies to support these unusual translocation mechanisms. For both, the substrate no longer stays fixed within the membrane but moves with the transporter itself to be released on the other side of the membrane. The "elevator mechanism" has been identified for Na^+/H^+- exchangers, with the substrate binding to the protein on the extracellular side.

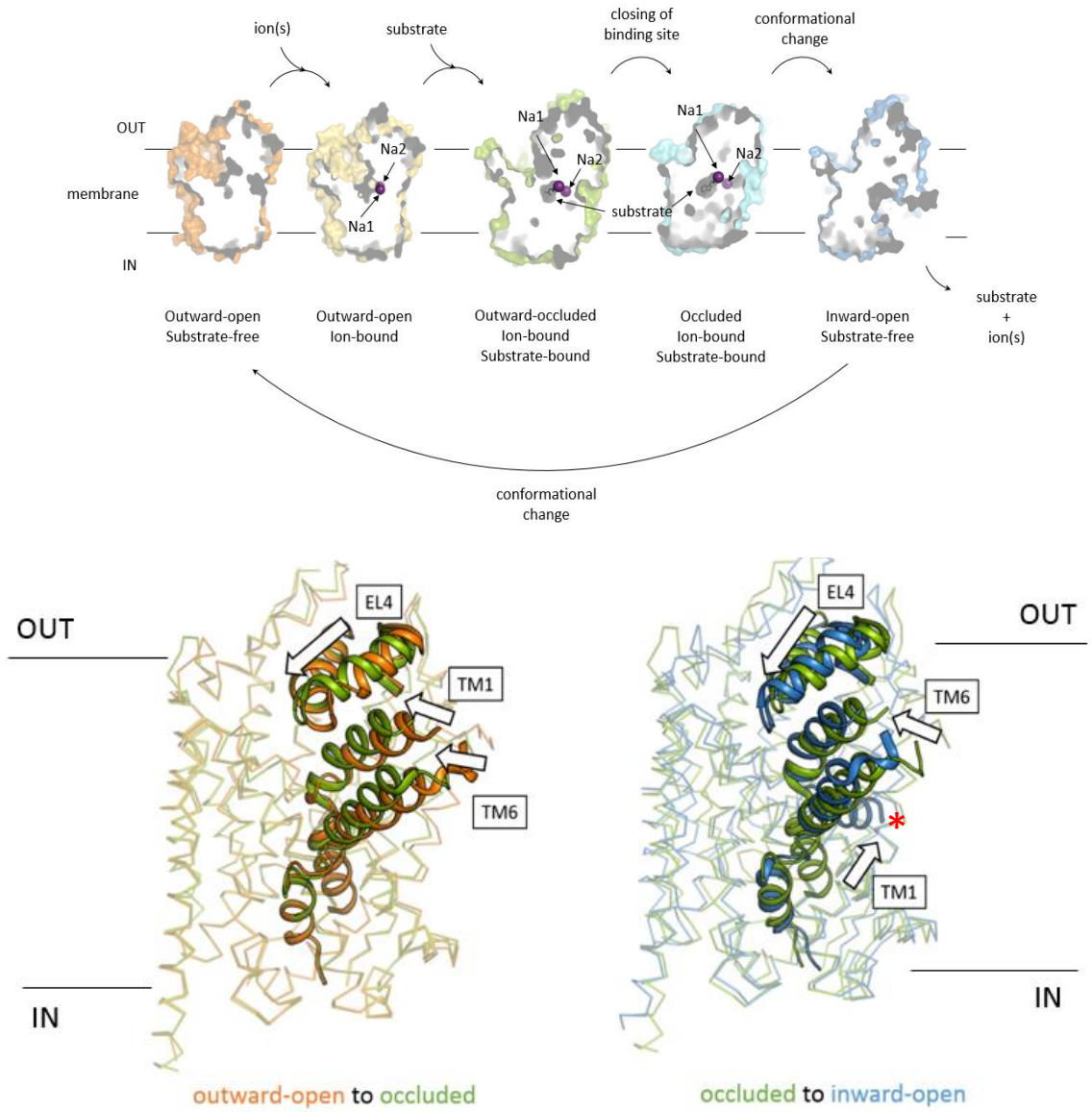


Figure 1-14: Alternating access mechanism based on published structures o LeuT

Crystal structures of LeuT, a sodium-independent neutral amino acid transporter, have supported the alternating access mechanism and revealed the different conformations that this transporter can adopt during the transport cycle (top). The protein is open to one side of the membrane (outward open), thus opening the binding site for the substrate and ion to bind in the middle of the transmembrane region of the transporter (outward open, ion bound). The binding event triggers conformational changes within the protein which closes the substrate binding site (occluded, ion and substrate bound). This is occluded from both sides of the membrane, and is followed by the opening of the protein to the other side of the membrane (inward occluded, substrate and ion bound) and release of the substrate (inward-open).
Comparison of the outward-open (PDB-ID: 5JAE, orange) and occluded state (PDB-ID: 2A65, green) revealed the movement of EL4, closing the substrate binding site, whereas TM1 and TM6 also bend towards the centre of the protein (bottom). To change from the occluded to the inward-open state (PDB-ID: 3TT3, blue) to release the substrate, TM1 bends aside, in a “dog leg”-movement (red star in bottom panel) and EL4, TM1 and TM6 move further towards the centre.

The binding site closes over the substrate and the protein slides within the membrane to the intracellular side and then releases the substrate. The 'sliding' or elevator motion can be facilitated by an interaction with another membrane protein.

This association with another protein complex has also been observed for the vitamin transporters, and revealed the "toppling" of the vitamin bound protein within the membrane. This is only possible because the two loops closing over the vitamin binding site present a hydrophobic surface which in turn allows the protein to insert into the hydrophobic part of the membrane. This movement is referred to as "toppling". The protein interacts with an ABC protein and binds to it, which pushes apart the two loops that close the binding site, releasing the substrate (Lee *et al.*, 2013; Slotboom, 2013; Swier *et al.*, 2016).

Table 1-4: Different types of transport mechanisms

Protein	Mechanism
LacY	Rocker-switch mechanism
LeuT	Rocking bundle mechanism
NhaA	Elevator mechanism
FolT	Toppling mechanism

1.3.3 Substrate binding sites

Members of the APC superfamily transport small molecules such as amino acids, sugars or nucleotides (Jack *et al.*, 2000; Yamashita *et al.*, 2005; Shaffer *et al.*, 2009; Shimamura *et al.*, 2010). Some transporters can recognize a wide range of substrates whereas others are specific for only a few. Thus, the substrate binding site is very divergent among the members of the APC superfamily, although the same overall protein fold is shared between them. Therefore, biochemical as well as structural data must be gathered to give insight into substrate recognition and substrate transport.

Based on structural data, discontinuous helices residing in the central cavity of the transporter have been found to be involved in coordinating the substrate at the unwound segments, and glycines located there are highly conserved across the family. The residues in

this region determine the electrochemical environment necessary for substrate binding. Targeted mutational studies on different APC transporters, such as the sodium-coupled transporters LeuT and Mhp1, have resulted in a complete loss of substrate or ion binding, or a change in substrate or ion specificity (Androutsellis-Theotokis *et al.*, 2003; Zomot *et al.*, 2007; Tavoulari *et al.*, 2011; Piscitelli and Gouaux, 2012; Simmons *et al.*, 2014). These studies have revealed important residues involved in both processes, and comparison of them with other family or superfamily members could give insight into differences and similarities of substrate binding.

Only a small number of APC superfamily members have been co-crystallised with a substrate or in different conformations, such as Mhp1 and LeuT. The latter has been crystallised in complex with different amino acids and neurotransmitters, and has shown a vast ability to accommodate different substrates into its substrate binding site. Also, mutations at this site can change LeuT's specificity from leucine towards tryptophan, based on comparisons with a tryptophan specific NSS member, TnaT (Piscitelli and Gouaux, 2012). Analysis of the sequence alignment between LeuT and the distantly related TnaT revealed that residues involved in substrate binding are mainly conserved between the two proteins, but found two exceptions, Phe259 and Ile359 of LeuT which are Val and Gln in TnaT, respectively. A homology model of TnaT based on LeuT showed that the substrate binding site was changed in its steric and electrostatic environment (Piscitelli and Gouaux, 2012). It was reasoned that substitution of the corresponding residues in LeuT would modulate the protein to transport tryptophan (Piscitelli and Gouaux, 2012). The single point mutation LeuT-Ile359Gln showed tryptophan transport activity although the double mutant did not (Figure 1-15) (Piscitelli and Gouaux, 2012).

Another example for which more than one conformation has been structurally obtained is AdiC, an arginine-agmatine exchanger, transporting arginine in one direction and agmatine in the opposite direction, which is also known as antiport. In fact, it acts as so-called "virtual" proton pump, since the intracellular decarboxylation of arginine to agmatine results in the release of a proton. This mechanism is used to respond to large changes in acidity outside the cell and thus acts as a rescue mechanism, with the highest AdiC activity being observed at a pH around 2 (Fang *et al.*, 2007).

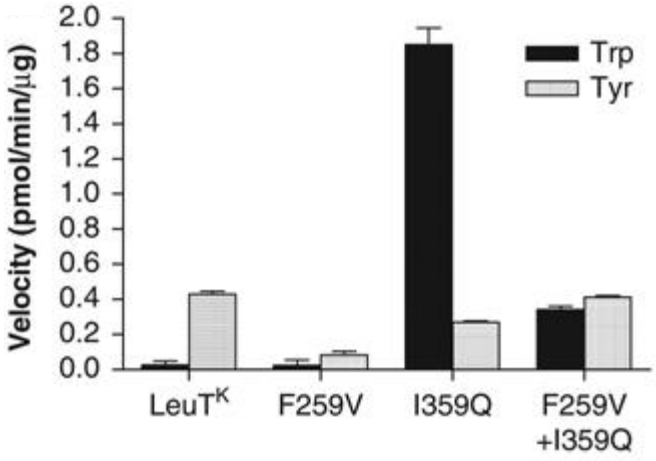


Figure 1-15: Modulation of LeuT's substrate specificity

Liposome-based assays monitoring the uptake of radiolabelled tyrosine or tryptophan by LeuT, and the mutants based on the comparison to the binding site of TnaT. Figure from Piscitelli and Gouaux, 2012.

The role of Tyr74 on the intracellular side of AdiC (Figure 1-16) has been studied and *in vitro* assays show that it acts as a pH sensor (Wang *et al.*, 2014). AdiC is a good example for illustrating that the same fold (LeuT fold in this case) is used for different translocation mechanisms, and moreover, based on sequence alignments to other members of the APC superfamily, the substrate specificity of a given transporter may be difficult to determine. Co-crystallised with arginine, the structure of AdiC revealed important residues involved in substrate binding. The bound arginine makes contacts to TM1 (Ile23, Ser26, Gly27), TM3 (Ala96, Cys97, Asn101, Met104), TM6 (Trp202, Ile205), TM8 (Trp293) and TM10 (Ser357) (Gao *et al.*, 2010). It has been noted that the substrate binding site is surrounded by several aromatic residues that close the substrate cavity above the substrate, acting as a 'gate' (Zhao *et al.*, 2013) but also form interactions with the substrate via Π-interactions, for example Phe253 in LeuT (Figure 1-17A), Trp117 in MhsT (Figure 1-17B), or Trp293 in AdiC (Figure 1-17D).

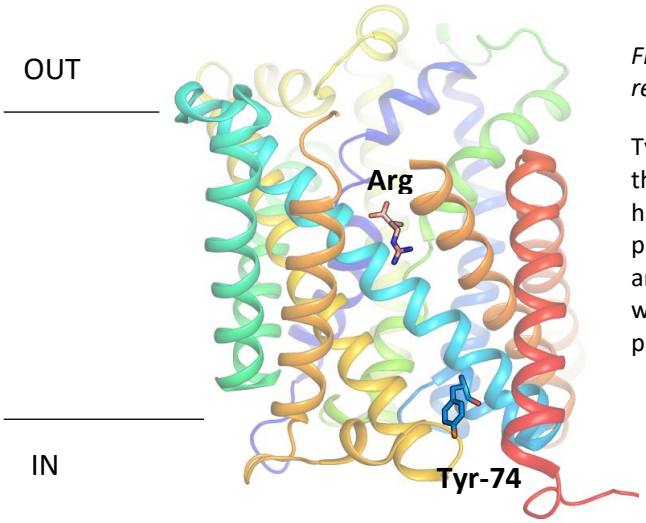


Figure 1-16: Location of the pH-sensing residue Tyr74 of AdiC

Tyr74 in AdiC (3OB6, 3.0 Å) is located on the intracellular side of the protein and has been proposed to be involved in a pH-sensing mechanism. Tyr74 is labelled and highlighted in the figure together with the bound substrate arginine (Arg, pink).

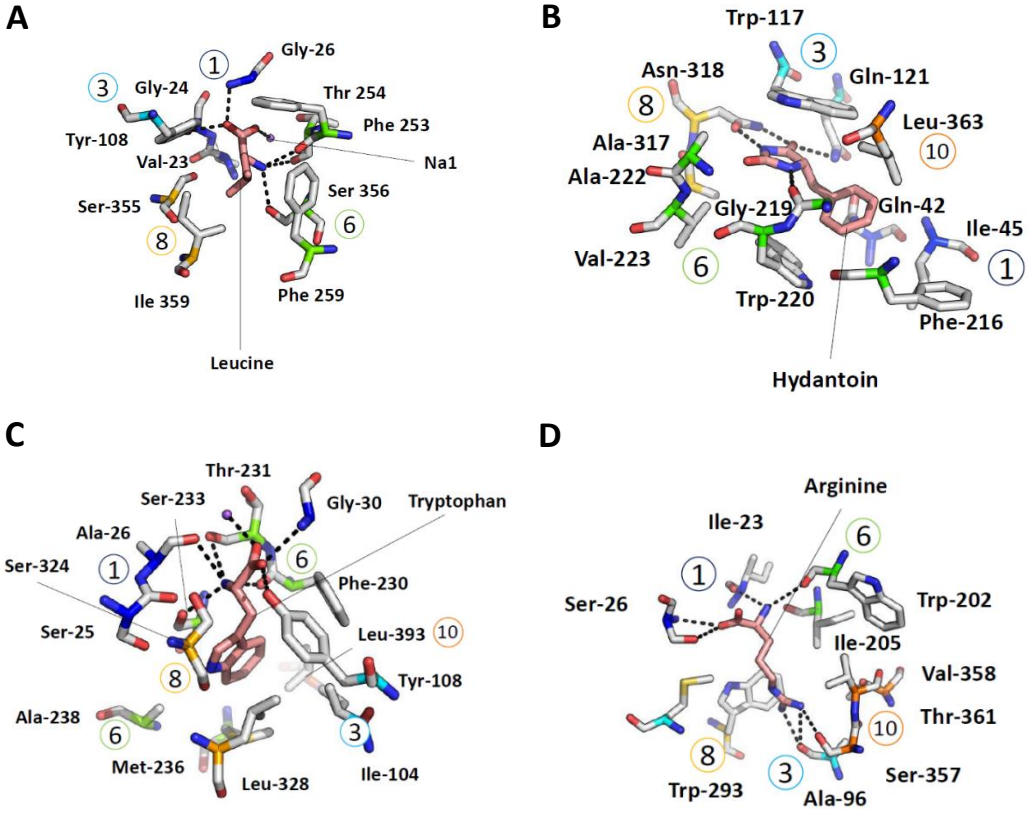


Figure 1-17: Examples of substrate binding in APC superfamily members

Residues interacting with the substrate (labelled, light pink) are shown as sticks together with the substrate (shown in light pink). Additionally, sodium ions are shown as small spheres (purple) in LeuT (2A65, 1.65 Å) bound to Leu (A), Mhp1 (4D1B, 3.8 Å) (B), MhsT (4US4, 2.6 Å) (C) and AdiC (3OB6, 3.0 Å) (D). The corresponding transmembrane helices are labelled individually. For some of the sodium coupled transporters, sodium is not directly involved in the coordination of the substrate. It should be noted that the MhsT structure is shown in an open conformation whereas the rest are illustrated in the occluded conformation.

In the outward-open conformation, these aromatic residues move aside to open the substrate binding site for a substrate molecule to enter (Zhao *et al.*, 2010, 2013). Interestingly, the cocrystal structure of AdiC with agmatine has also been solved, showing a similar coordination of agmatine as arginine although missing the hydrogen bonding with TM1 due to the lack of a carboxyl group on agmatine (Ilgü *et al.*, 2016). Other co-structures have been determined in the past few years, making it possible to compare the binding sites of different members to identify their conserved residues and properties. The available structures have revealed that the substrate binding site is located near the interface of unwound segments of transmembrane helices in the centre of the protein, which is consistent among all APC members. Usually, the carboxyl group interacts with TM1 via hydrogen bonding and includes an interaction with TM3 in some cases, such as MhsT (Figure 1-17C). The amine group interacts with TM6 via hydrogen bonds through side chain and backbone carbonyl hydrogen bonding. Again, in MhsT this interaction is different. There the amine group interacts with TM1 and not TM6 (Figure 1-17C). In some cases the substrate is also involved in ion coordination (Yamashita *et al.*, 2005; Screpanti and Hunte, 2007).

Given the different sizes and charges of amino acids and their precursors or derivatives, the transporters have evolved to accommodate a certain subset of them. Comparison between the binding site of LeuT and MhsT showed that steric restrictions prevented LeuT from binding tryptophan, as is the case in MhsT. Additionally, the specificity of AdiC for both arginine and agmatine results from the interaction of their side chain with the side chain of Ser357 and the carbonyl group of Ala96, providing the polar environment for the substrate (Ilgü *et al.*, 2016). Thus a combination of the electrostatic composition of the substrate binding site and the size restrictions determines the substrate specificity of the APC superfamily members.

1.3.4 Ion binding sites

The different composition of the aqueous soltution on the cell membrane creates a negative inward membrane potential of -40 mV – (- 80) mV (potential difference between cytoplasm and exterior). This potential is used by secondary active transporters to couple substrate transport, against their electrochemical gradient, to a positively charged ion (Jack *et al.*, 2000).

In LeuT, two sodium ion (Na1 and Na2) binding sites have been identified, whereas in Mhp1 only one was found (Shimamura *et al.*, 2010; Krishnamurthy and Gouaux, 2012). The Na1-site for both LeuT and MhsT reside in the substrate binding site interacting with the bound substrate (Figure 1-27), which leads to the conclusion that the binding of the ion precedes that of the substrate. The Na1 binding site is flanked by charged or polar residues near the unwound segment of TM1. For the LeuT-Glu290Ser mutant, the coupling mechanism changed to a dependence on chloride, as has been observed for the eukaryotic members of the NSS family (Forrest *et al.*, 2007; Zomot *et al.*, 2007; Tavoulari *et al.*, 2011). The second Na2 binding site of LeuT is located between TM1 and TM8 (Figure 1-18) and is not interacting with the substrate in this protein as is Na2 in MhsT (Figure 1-18). It might be possible that the second sodium ion is used to stabilise the protein for subsequent substrate binding.

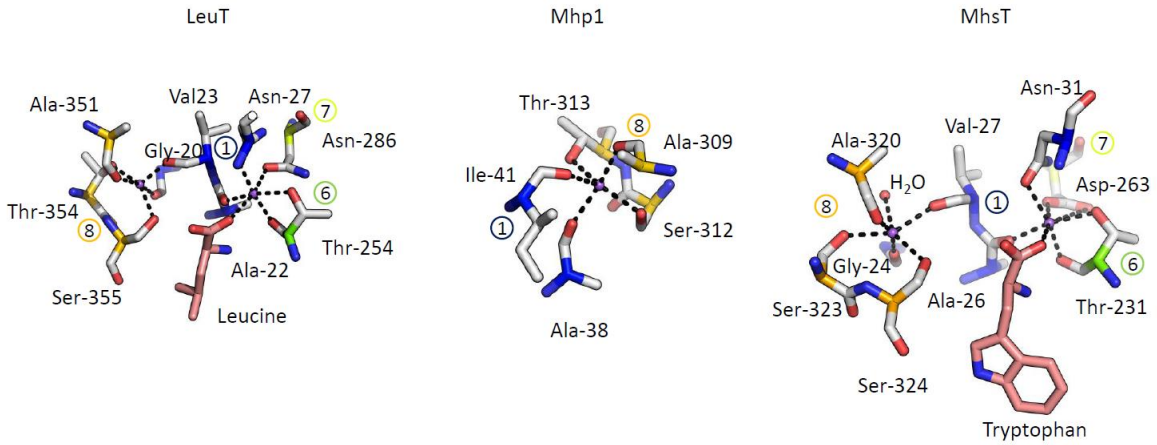


Figure 1-18: Examples of sodium binding sites of APC transporters

Here, selected examples of the ion binding sites of APC members are shown. In LeuT (2A65, 1.65Å) and MhsT (4US4, 2.6Å), two sodium binding sites are found, of which the Na1 site is directly involved in the interaction with the substrate. The second sodium binding site is thought to stabilise the adjacent helices. In Mhp1 (4D1B, 3.8Å) only one bound sodium has been observed in the solved crystal structures, and it does is not directly interact with the substrates. Although all three proteins share a common fold, similar binding modes for these proteins cannot be assumed. It is thus important to obtain crystal structures to correctly assign residues involved in ion and substrate binding.

However, much is known about the ion binding site of sodium-coupled transporters but the question arises as to how these transporters differ to the ones that do not use any ions.

A comparison of the occluded states of LeuT (PDB-ID: 2A65, 1.65Å) and the arginine/agmatine exchanger AdiC (PDB-ID: 3L1L, 3.0Å) shows that the ion binding sites of Na1 and Na2 of LeuT are obscured in AdiC (Figure 1-19). TM8 of AdiC is displaced sideways in respect to TM8 of LeuT (Figure 1-19), which moves the carboxyl group of Gly285 (AdiC) away from the cluster of Ser289,Gly21 and Met24. The increased distance disrupts the ion

binding site in comparison to LeuT. Additionally, coordination of Na1 is not possible in AdiC since the interacting polar side chains of LeuT, Asn27 and Asn286, are not present in AdiC (Figure 1-19). Furthermore, the slight displacement of the carboxyl group of arginine relative to the carboxyl group of leucine is observed, would fruther disrupt ion coordination in this region (Figure 1-19). Thus, AdiC is, based on this structural comparison, not providing the distinct environment relevant for ion coupling and shows how sodium-coupled transporters have adapted to a different coupling mechansim.

The structural insight into the differences between AdiC and LeuT in terms of ion binding have shown how LeuT evolved to be sodium-coupled and why AdiC cannot use this mechansims. Obtaining structural information about a proton-coupled APC superfamily members would further shed light into the evolution of different coupling mechanism. It would be interesting to perform a similar comparison to understand how, for example the SLC36 family differs in these regions, if the protons bind in a similar region as observed in their sodium-coupled ancestors, and if they are also involved in substrate coordination.

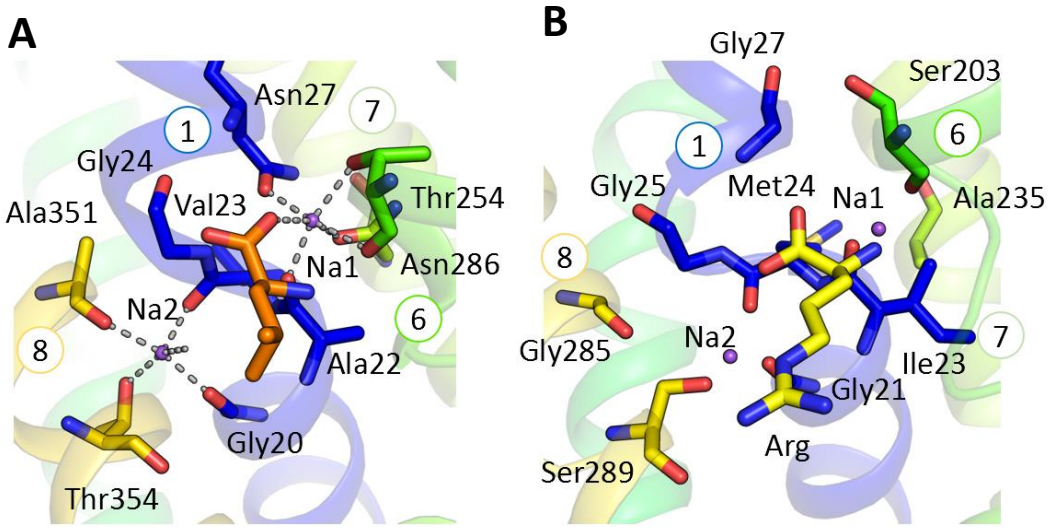


Figure 1-19: Comparison of sodium binding site of LeuT to the sodium-independent transporter AdiC

Comparison of LeuT (PDB-ID: 2A65, 1.65 Å) (A) and the exchanger AdiC (PDB-ID: 3L1L, 3.0 Å) (B) shows the difference in the sodium (purple spheres) binding site of LeuT and that of the Na^+-independent exchanger AdiC. The substrate leuine (orange) of LeuT and arginine (yellow) of AdiC are shown. In addition, the position of the two sodium ions (Na1 and Na2) from LeuT were added to the substrate binding site of AdiC (B).

1.3.5 Extracellular gates locking the substrate and ion binding site

Upon substrate and ion binding, the protein undergoes conformational changes towards the occluded state in which the binding site is shielded, preventing the substrate release to either side of the membrane (Figure 1-21). Published crystal structures of different APC members in the occluded state have identified residues involved in forming the extracellular and intracellular gate, and revealed that this arrangement differs between members of the APC superfamily. In LeuT, the extracellular gate consists of residues within TM1, TM3, TM6, and extracellular loop (EL) 4 (Figure 1-20), which acts as the 'plug' (Krishnamurthy and Gouaux, 2012). A similar arrangement to EL4 of LeuT is found for EL4 of *Mj*ApcT (Shaffer *et al.*, 2009). In the latter, EL4 lies on top of the closed putative binding site formed by TM3 and TM6, suggesting a common gating mechanism with LeuT. It has been proposed that the outward-open state might be the thermodynamically favoured state of the analysed transporters, LeuT and GltPh, probably to enhance the uptake of possible nutrition on the outside of the cell (Weyand et al., 2011).

Thus, a peptide or molecule mimicking the extracellular loops mentioned above could be a possible inhibitor for this transporter class, giving an opportunity to identify inhibitors for human homologues such as the NSS. Although Glt_{Ph}, an EAAT transporter, belongs to a different secondary-active superfamily, it is a good example of how, based on structure-based observations an inhibitor has been discovered: DL-threo-beta-benzyloxy aspartate (TBOA). TBOA's inhibitory effect on human EAAT transporters was described before any structural evidence was found (Shimamoto *et al.*, 1998).

The structure of the co-crystallised complex was solved, and revealed that the inhibitor forced the HP2 to adopt an open confirmation, resulting in the disruption of the sodium binding site (Reyes *et al.*, 2013). This emphasizes that a detailed understanding of the transporter and the mechanistic aspects are necessary in order to develop novel drugs for transporter homologues in humans. This will only be possible by using complementary information from both structural and biochemical data.

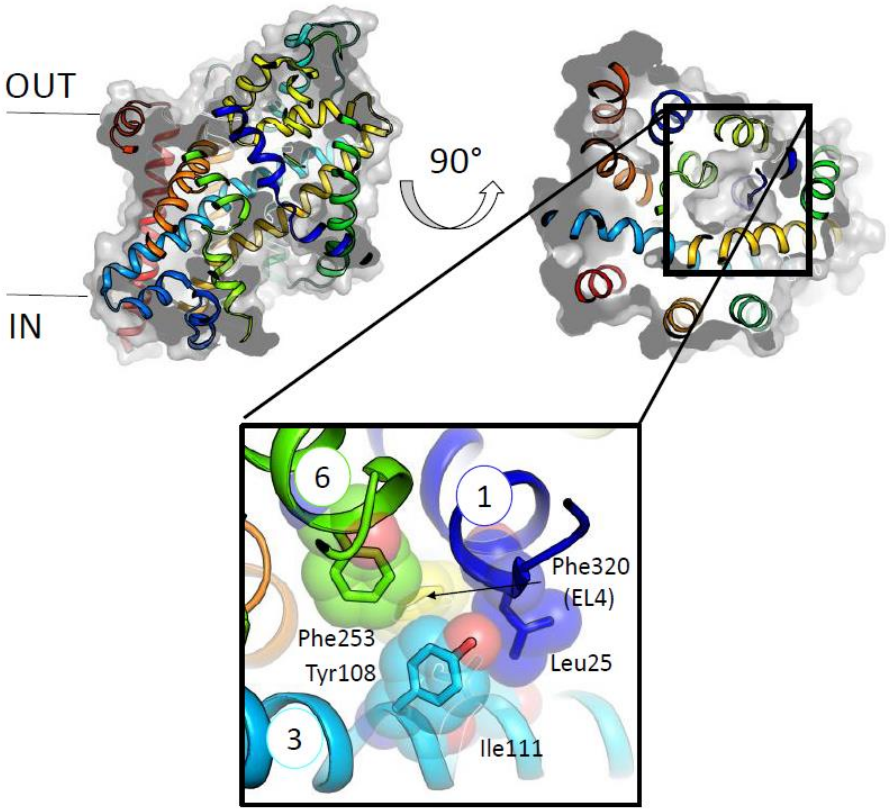


Figure 1-20: Extracellular gate of LeuT

The inward open structure of LeuT (3TT3, 3.12 Å) revealed the residues involved in forming the extracellular gate. Leu25 of TM1, Tyr108 and Ile11 of TM3, Phe253 of TM6 and Phe230 of the extracellular loop 4 (EL4) close the top of the binding site.

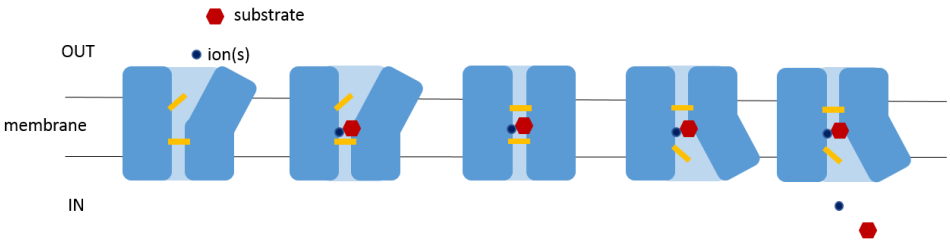


Figure 1-21: Schematic of the gating mechanism by secondary active transporters

Simplified model of the gating mechanism of secondary active transporters (blue). Gates on top and below (yellow bars) of the substrate binding site ensures that the substrate is not released during the translocation process. The gates close the substrate binding site when both ion(s) (dark blue circle) and substrate (red hexagon) are bound. When the transporter changes its conformation, the corresponding gate on the intracellular side opens to release the substrate and ion on the other side of the membrane.

1.4 SLC7 FAMILY OF AMINO ACID TRANSPORTER

The SLC7 family of amino acid transporters is a eukaryotic solute carrier family homologous to the APC family members *Mj*ApcT, AdiC and GadC. The SLC7 family is divided into two subgroups, the cationic amino acid transporters (CATs) and the heteromeric amino acid transporters (HATs). The latter consist of a light subunit, which is the actual transporter, and a heavy subunit belonging to the SLC3 family. Due to its close relation to bacterial APC transporters, it is thought that the core transporter structure of the CATs and the light subunit of the HATs resemble the LeuT fold. As described above, the LeuT fold is characterised by the pseudo-symmetrical relation between the first five TMs and the consecutive five TMs, also known as "5+5 – symmetry". This symmetry relation is found for the first 10 TM helices of the CATs and also for the light subunits of HATs (Bartoccioni *et al.*, 2010).

Another common feature of the LeuT fold transporters is the unwound segments of the first helix of each repeat, carrying a common GSG-motif, which is more pronounced in TM1 than in TM6 (Figure 1-22). The unwound segments provide a polar environment for ion and substrate binding, interacting strongly with the carboxyl moiety of the substrate. The CATs and LATs also carry this motif (Figure 1-22) which leads to the conclusion that the substrate binding site of these transporters is similarly located to that of the bacterial and archaeal APC transporters. This has been supported experimentally (Font, 2001; Bartoccioni *et al.*, 2010).

To date, there is no near-atomic structural data available for the SLC7 family and only a few mutational studies that locate the substrate binding site and explain the diverse substrate specificity of this family. Structural data of CATs and HATs would be valuable in understanding the binding mode and how the transporters differentiate between various amino acids, as well as understanding how the light and heavy subunits of the HATs interact with one another.

Table 1-5: Human SLC7 family

Human gene name	Protein name	Forms a heterodimer with	Substrates
SLC7A1	CAT1	-	cationic AAs
SLC7A2	CAT2	-	cationic AAs
SLC7A3	CAT3	-	cationic AAs
SLC7A4	CAT4	-	cationic AAs
SLC7A5	LAT1	4F2hc	large neutral AAs
SLC7A6	y+ LAT2	4F2hc	cationic AAs (Na-independent) large neutral AAs (Na-dependent)
SLC7A7	y+ LAT1	4F2hc	cationic AAs (Na-independent) large neutral AAs (Na-dependent)
SLC7A8	LAT2	4F2hc	neutral AAs
SLC7A9	b0 +AT	rBAT	cationic AAs, large neutral AAs
SLC7A10	Asc-1	4F2hc	small neutral AAs
SLC7A11	xCT	4F2hc	cysteine, glutamate
SLC7A13	AGT-1	rBAT	aspartate, glutamate, cysteine
SLC7A14	-	-	-

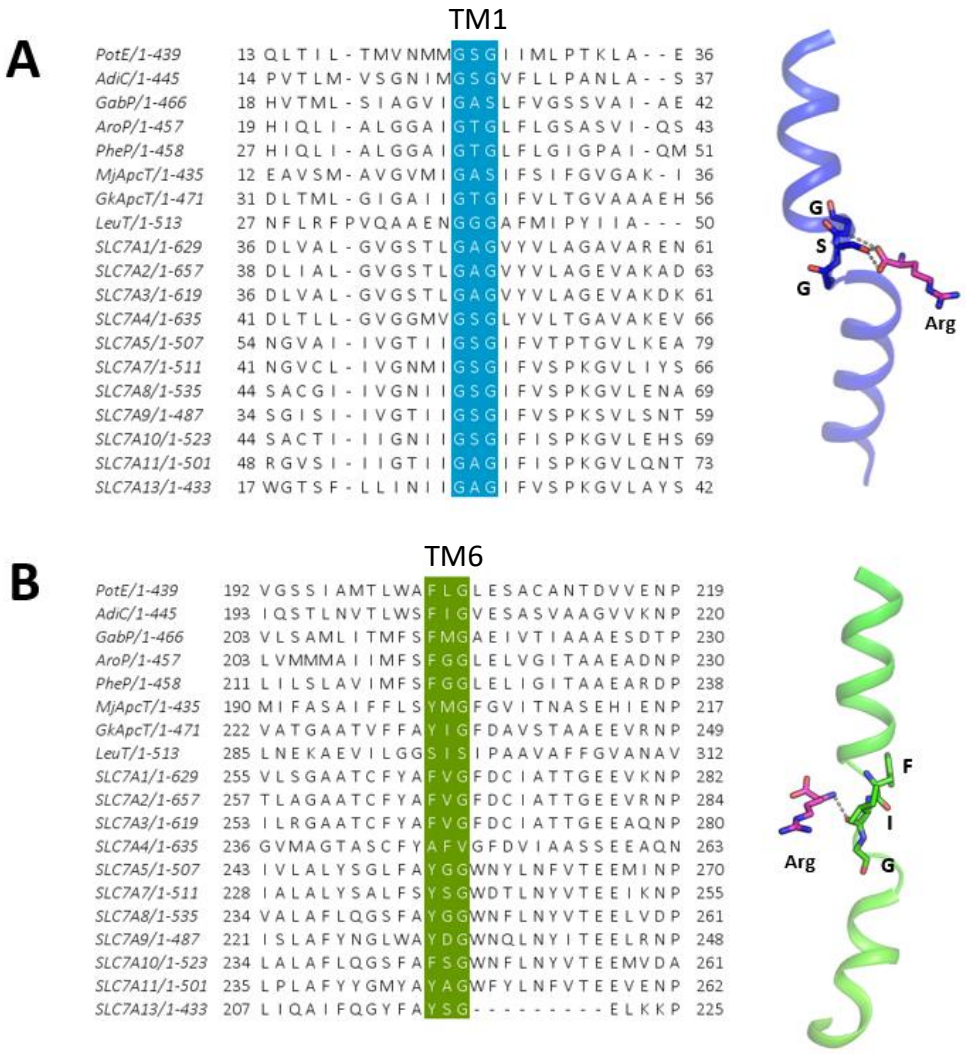


Figure 1-22: GSG motifs in helix 1 and helix 6 of APC transporters

Sequence alignment of prokaryotic APC members and members of the SLC7 family using CLUSTAL OMEGA (Sievers *et al.*, 2011). The GSG motif of the first discontinuous helix in those transporters, which is highly conserved among them is highlighted (A). The GSG motif interacts with the carbonyl moiety of the substrate (right, from AdiC; PDB-ID: 3OB6, 3.0Å) emphasising the importance of this region and its conservation. The symmetry related second discontinuous helix does not contain a highly conserved GSG motif (alignment to the left) (B). Only the glycine is conserved in this region, probably for flexibility reasons during the transport cycle. This region does not interact much with the substrate (right, from AdiC, 3OB6, 3.0Å).

1.4.1 Heteromeric amino acid transporters

The HATs, heteromeric amino acid transporters, belong to the SLC7 family and consist of a light and heavy subunit (Fotiadis *et al.*, 2013). The light subunit is associated with members of the SLC3 family (SLC3A2 and SLC3A1), consisting of a single transmembrane helix, that is bound via a disulphide bridge to a loop region between helix 4 and 5 of the HAT transporter (Figure 1-23). The SLC3 protein consists of the N-terminal single transmembrane helix followed by the heavy subunit which is located on the extracellular side of the membrane. The interaction between the LAT and SLC3 protein ensures the correct trafficking and localisation of the HAT protein to the cellular membrane (Sitte, 2004). Otherwise the protein remains in the membrane of the endoplasmic reticulum and cannot fulfil its function. The nine LAT proteins associate with only two of the SLC3 proteins, either with 4F2hc or rBAT. Both proteins show high similarity to bacterial α-amylases, but do not function as such any more (Gabriško and Janeček, 2009). The heavy subunit is covalently bound to the light subunit via a conserved disulphide bridge that resides between the loop region of helices 4 and 5 of the light subunit and a loop of the heavy subunit. There is no other information about how these two proteins interact within the membrane and with one another.

The light subunit, also called the LAT protein, is predicted to have 12 transmembrane helices (Gasol *et al.*, 2004). LATs are known to transport a broad variety of substrates, such as neutral, cationic, anionic and aromatic amino acids, in a sodium-independent manner. Transporters of neutral amino acids are the only sodium-dependent transporters of this family (Fotiadis *et al.*, 2013). Investigation of another bacterial homologue SteT, a serine threonine exchanger, suggested similarities in the location of the substrate binding site between other prokaryotic APC members and the LAT proteins. Here, mutation of Lys295 to cysteine, which is homologous to Trp293 in AdiC, broadened the substrate specificity of SteT (Bartoccioni *et al.*, 2010). Additionally, comparison to disease related mutations in the LATs supported a similar binding site to the APC superfamily (Mykkanen, 2000; Font, 2001). LAT1 (SLC7A5) has pronounced expression levels in the micro vessels of the central nervous system, where it is responsible for the uptake of L-3,4-dihydroxyphenylalanine (L-DOPA) across the blood-brain barrier. Furthermore, LAT1 is found in high levels in the placenta, the brain, the testes and the colon, as well as in the inner blood-retinal barrier for maintenance of neutral amino acids and neurotransmitters (Kanai *et al.*, 1998; Kageyama *et al.*, 2000). LAT2 (SLC7A8)

shows high expression levels in brain, liver, spleen, heart, kidney, lung, small intestine, skeletal muscle, testes, prostate, ovaries and foetal liver (Pineda, 1999). LAT dysfunction is associated with various diseases such as lysinuric protein intolerance or cystinuria (Mykkanen, 2000; Font, 2001). Furthermore, LATs are overexpressed in cancer cells. They are involved in the mTORC signalling pathway, regulating cell proliferation and additionally ensuring constant nutrition for the tumour cells (Nicklin *et al.*, 2009).

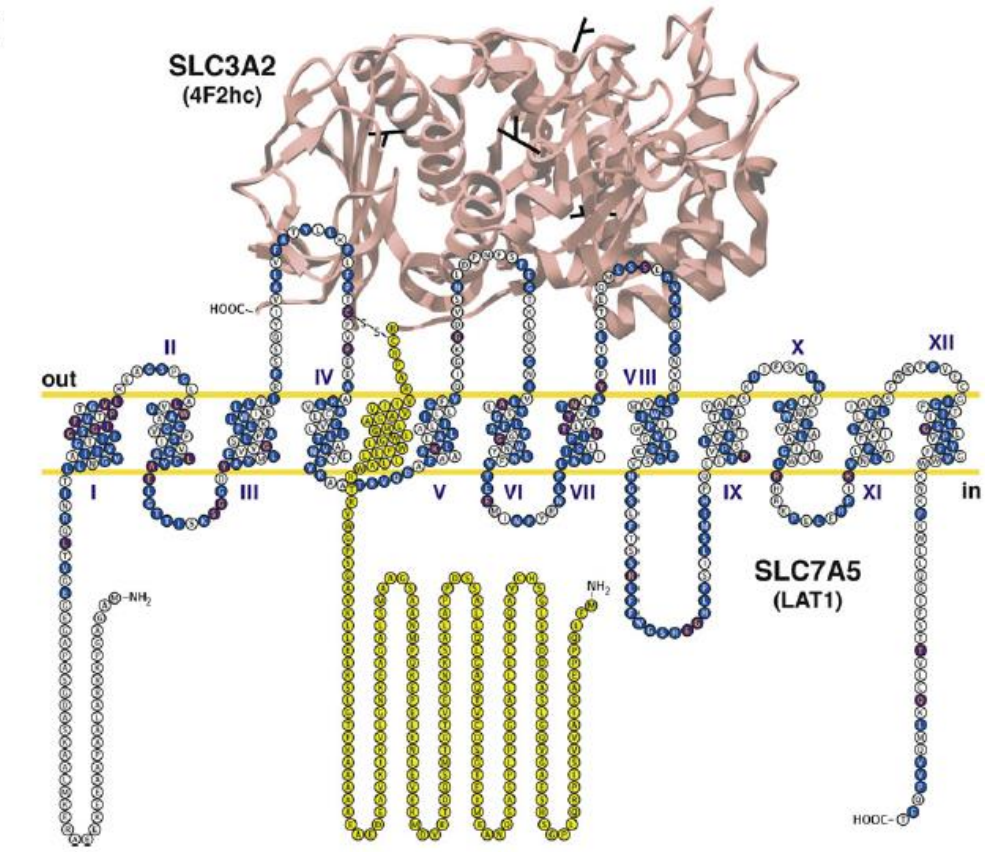


Figure 1-23: Topology of a HAT protein in the membrane

The human HAT proteins form heterodimers with a heavy subunit, responsible for transporter regulation and correct trafficking, and a light subunit, which performs the actual transport of substrate. The latter belongs to the SLC7 family with 12 TMs and intracellular N- and C-termini. The heavy subunit belongs to the SLC3 family and contains a single spanning TM (yellow). The interaction between the two proteins is partially mediated by a disulphide bond formed btween two conserved Cys residues. The extracellular domain of the SLC3 proteins has been solved to a resolution of 2.1 Å but it is unknown where the TM of this protein sits with respect to that of the SLC7 protein.

1.4.2 Cationic amino acid transporters

Due to their functions as arginine transporters and the location of the transporters in macrophages, the CATs are important proteins in the immune response pathway by providing the initial substrate for the signalling cascade.

The first cationic amino acid transporter (CAT) protein was described in 1989 and was associated with a receptor function for murine leukaemia virus, MuLV. By analysis of a hydropathy plot of the protein, Albritton *et al.* suggested that the 622 residue protein had 14 transmembrane helices (Albritton *et al.*, 1989). Insertion of the cDNA into a human cell line made them susceptible to a virus that usually only infects mouse cell lines. Two years later, the function of this receptor in mouse was reported to be an amino acid transporter selective for lysine, histidine and L-ornithine. This was the first time that the receptor was associated with a transporter, and in the same year a homologue in humans was found, namely H13 (Kim *et al.*, 1991).

Why the mouse receptor and not H13 recognises the MuLV virus was further investigated. Mutational studies revealed that the most critical residue was Tyr235, which was thought to reside in an extracellular loop of the mouse receptor. Reciprocal mutation in H13 at positions 240/242 or 242/244 on the human protein gained the function of the MuLV receptor from mouse (Yoshimoto *et al.*, 1993). Later, the H13 protein was named CAT1, short for cationic amino acid transporter 1 and over the following few years three more human CAT proteins were identified based on sequence similarity and function. Thus in human, there are four known CAT proteins, CAT1 to CAT4, although CAT4 is an orphan protein with unknown function (Wolf *et al.*, 2002).

CAT1 is ubiquitously expressed in all tissues except the liver, which is the organ where CAT2A is dominantly found. Furthermore, CAT2A is expressed in cardiomyocytes, cardiac microvascular endothelial cells, the pancreas, and skeletal and vascular smooth muscle although at lower levels. CAT2B is only found in lung tissue after treatment with bacterial lipopolysaccharide or pro-inflammatory cytokines, thus only after an inflammation. CAT3 is located to the plasma membrane and is tissue specific, expressed in thymus, Burkitt's lymphoma cells, uterus, testis, mammary gland and MOLT-4 leukemia cells. Additionally, CAT3 is localised especially in the adult-stage brain and the mesoderm in embryos. Since CAT3 is most abundant during embryogenesis compared to the other CAT proteins, which are only

found in adult tissues, CAT3 must play a major role in the development of the foetus (Ito and Groudine, 1997).

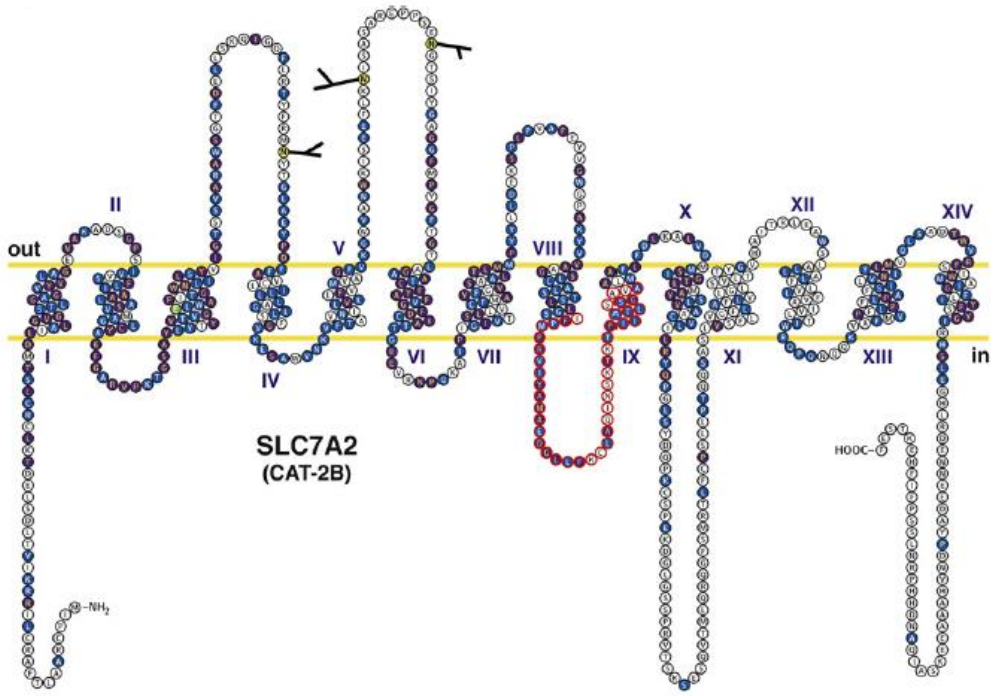


Figure 1-24: Topology of a CAT protein in the membrane

CATs consist of 14 TMs with intracellular N- and C-termini. CATs derive from an ancestral 12 TM protein of which the last two helices were duplicated. They have extracellular glycosylation sites (highlighted as black branches). CAT2 is the only CAT protein that exists as two variants which differ only in a 42 amino acid long region (highlighted in red). Figure adapted from Fotiadis *et. al*, 2013.

All CAT proteins are thought to have 12-14 transmembrane helices and perform sodium-independent transport of Arg, Lys, His with varying substrate affinities.

CAT2 exists as two splice variants in the human proteome which differ only in an intracellular loop region of 42 residues. Comparison with available structures of APC transporters do not explain the role of this amino acid segment. CAT proteins are predicted to consist of 14 transmembrane helices with an intracellular N- and C-terminus.

Early research described CATs as transporters for Lys, His, Arg and L-ornithine. No competition of transport was observed with D-Arg, GABA, L-Ser or L-Glu, which makes the transporter selective for cationic amino acids and stereospecific, suggesting that the substrate binding site is asymmetric and allows only one orientation of the substrate. All CATs are

sodium-independent transporters and no proton-coupling has been observed, which suggests that the transporters are uniporters or exchangers (Kim *et al.*, 1991; Ito and Groudine, 1997).

Some of the CATS are trans-stimulated, meaning that the transport rate increased upon presence of substrate on the intracellular side. This effect has been observed for CAT1, CAT2B and CAT3. CAT1, 2B and 3 are reported to be high affinity transporters for arginine, whereas CAT2A is a low affinity transporter with K_M values range from 40 µM to 1 mM (Fotiadis *et al.*, 2013). As CAT2A and CAT2B differ only in a small portion of their protein sequence, investigations on the 42 spanning residues have been performed and have identified Arg369 as the potential residue responsible for the different apparent K_M for arginine (Habermeier *et al.*, 2003). Arg369 does not reside close to the binding site of the protein, which means that it influences the substrate affinity only indirectly, probably by interference with the conformational changes during the transport cycle, thus slowing down the transport rate.

Initial experiments on CAT3 showed high conservation between hCAT3, rCAT3 and mCAT3 which only differ in 5 and 11 residues, respectively (Vékony *et al.*, 2001). Expression of CAT3 in oocytes and subsequent transport assays observed no transport of Ser, Leu, Gln, Pro, Phe, Met or citrulline but for Arg, Lys, ornithine (Vékony *et al.*, 2001). Thus, CAT3 was identified as being a member of the CAT family not only by its sequence similarity of 80 % to the other CAT members, but also due to its similar function, though no uptake of His was reported even when different pHs were tried in the experimental set up (Vékony *et al.*, 2001).

1.4.3 *Mj*ApcT – an archaeal SLC7 homologue

*Mj*ApcT, an archaeal amino acid transporter of the APC family within the APC superfamily, is the only distantly related SLC7 homologue that has been identified to date. The structure of *Mj*ApcT has been solved to 2.32 Å by X-ray crystallography in the occluded substrate-free state (Shaffer *et al.*, 2009). The protein consists of 12 TM helices and superimposition to LeuT and AdiC shows that the structure shares the LeuT-fold with a root mean square deviation (RMSD) of 5.4 Å and 6.1 Å, respectively (Figure 1-25). Structural alignment of TM1 and TM6 of *Mj*ApcT onto the corresponding helices of LeuT and AdiC shows that the unwound segments of *Mj*ApcT are conserved structural elements among these APC members (Figure

1-25). In the case of LeuT, the unwound segment of TM6 is more pronounced and results in a noticeable shift of the intracellular half of this helix (Figure 1-25). This phenomenon is not present in AdiC or *Mj*ApcT (Figure 1-25), showing the structural divergence of this conserved sequence motif.

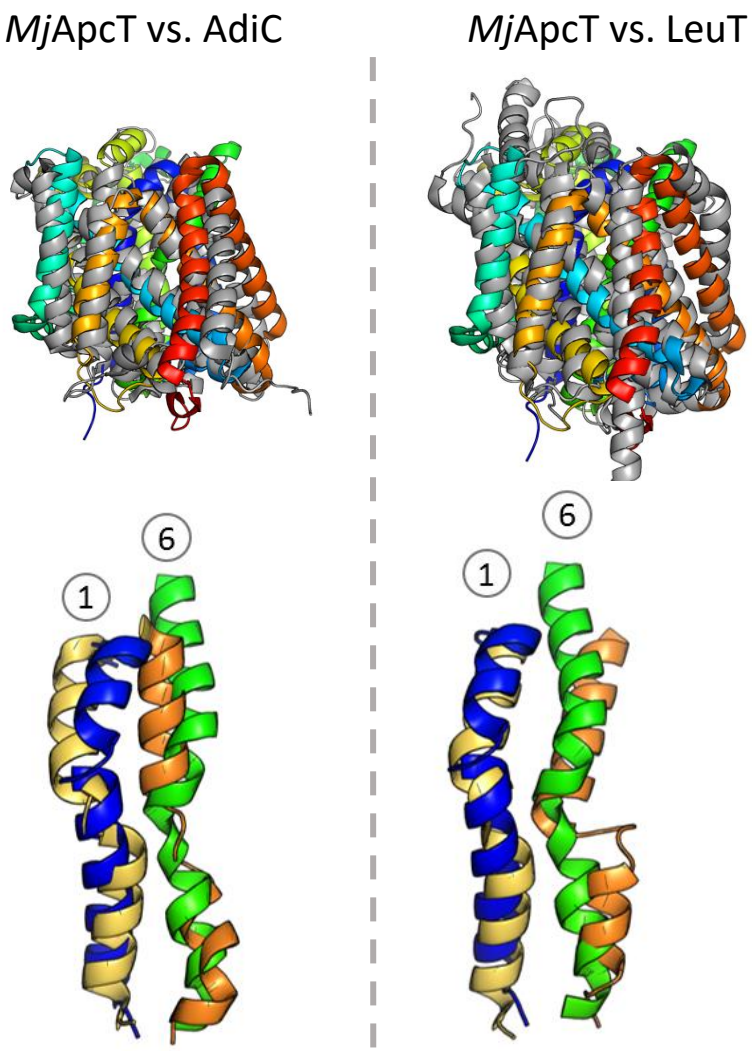


Figure 1-25: Comparison of MjApcT to LeuT and AdiC

Superposition of *Mj*ApcT (3GIA, 2.32 Å) onto the structure of AdicT (top left, 3L1L, 3.0 Å) resulted in an RMSD of 5.4 Å. Structural comparison of the discontinuous helices TM1 and TM6 (bottom left), shows that TM1 (blue) of *Mj*ApcT is moved further towards TM6 (green) than TM1 and TM6 of AdiC. Further comparison of *Mj*ApcT to LeuT (2A65, 1.65 Å) showed an overall RMSD of 6.2 Å (top right) and a close structural similarity between TM1 of the two proteins. However, TM6 of LeuT shows a sideways displacement at the unwound segment of TM6 (bottom right).

Although no substrate-bound structure was obtained, a water filled cavity was found in the centre of the protein which is assumed to be the binding site (Figure 1-26) (Shaffer *et al.*,

2009). Structural superimposition of *Mj*ApcT with substrate-bound structures of LeuT and the closely related AdiC shows that the water cavity is located below the substrate binding site of the compared APC transporters (Figure 1-26).

As mentioned previously, a conserved GSG motif of the unwound segments of TM1 is found for the APC superfamily (Figure 1-22). *Mj*ApcT has the sequence GAS at this position and amongst the aligned APC members, is the only one that shows such a high divergence of this motif. It remains to be determined what effect this difference might have on the orientation of the substrate and substrate binding.

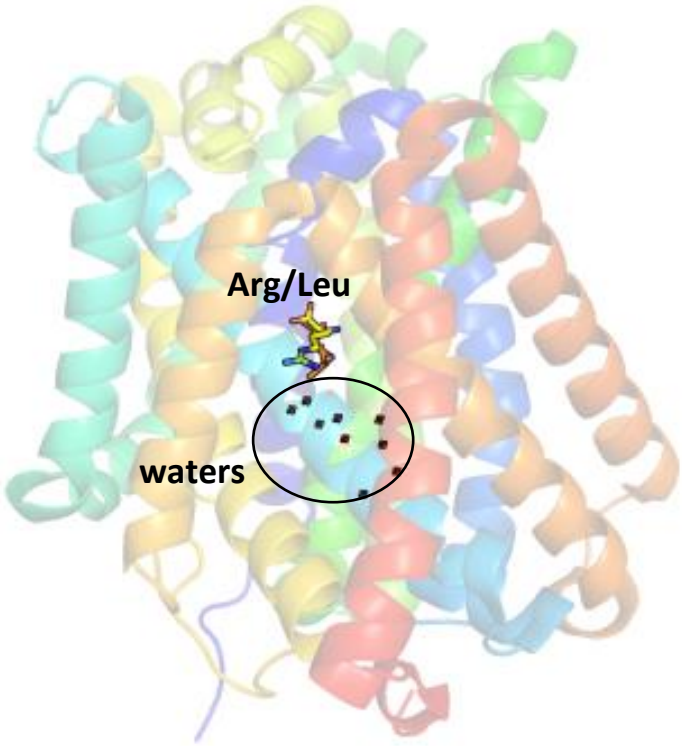


Figure 1-26: Water-filled cavity of MjApcT

*Mj*ApcT was solved in the occluded substrate-free state (rainbow, cartoon), but a water filled cavity (indicated by the circle) was found (Shaffer *et al.*, 2009). Comparison to the location of the substrates Leu of LeuT, and Arg of AdiC show that the cavity is located below the substrate binding site of the other APC transporters.

Further comparison of *Mj*ApcT (PDB-ID: 3GIA, 2.32 Å) with LeuT (PDB-ID: 2A65, 1.65 Å) revealed that a lysine accommodates a similar position to that of Na2 in LeuT. As it is the only charged residue that has been found within the transmembrane part of the protein, and given its analogous geometry to the Na2 site of LeuT (Figure 1-27), it was suggested that this

lysine is involved in the proton-coupling mechanism of *Mj*ApcT, since alanine transport was enhance in the presence of a pH gradient (Figure 1-28A) (Shaffer *et al.*, 2009). Mutation of this residue to an alanine resulted in the loss of alanine transport (Figure 1-28B) (Shaffer *et al.*, 2009). The same effect has been observed for the lysine-specific permease LysP (Kaur *et al.*, 2014). Here, mutation of Lys163 to alanine or leucine abolished transport, whereas the mutation to an arginine restored transport of lysine into liposomes (Kaur *et al.*, 2014). Interestingly, the antiporter CaiT has been discovered to carry an arginine residue at the same position as Lys158 of *Mj*ApcT (Figure 1-27). Mutation of Lys163 to alanine or glutamate at this location reduced the antiport activity of CaiT (Kalayil *et al.*, 2013), a similar result to that observed for *Mj*ApcT (Figure 1-28B) (Shaffer *et al.*, 2009). However, the authors questioned if the mutations might have an influence on CaiT's sensitivity towards sodium (Kalayil *et al.*, 2013). Therefore, the antiport measurements were repeated in the presence of sodium chloride, and a three- to twelve-fold increase in antiport activity was observed for mutants Arg262Glu and Arg262Ala, respectively (Figure 1-29) (Kalayil *et al.*, 2013). Still, the activity was only 25 % (Arg262Glu) and 10 % (Arg262Ala) of that of the wildtype (Figure 1-29) (Kalayil *et al.*, 2013).

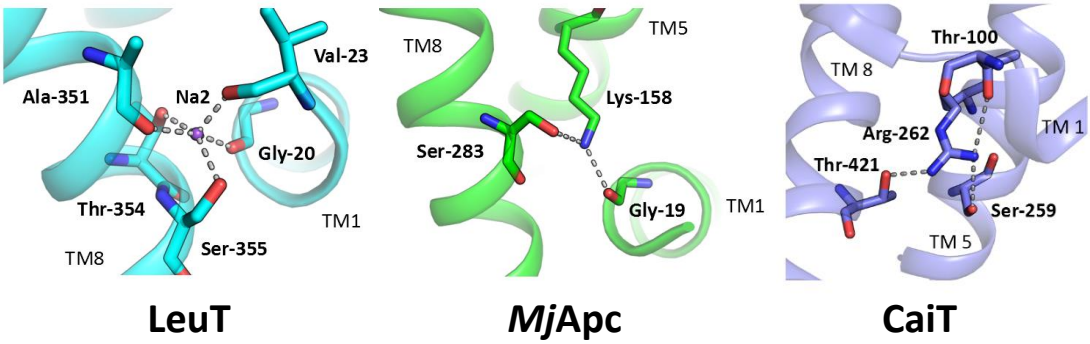


Figure 1-27: Comparison of the orientation of Na2 in LeuT, Lys158 in MjApcT and Arg262 of CaiT

The second sodium ion (Na2, purple sphere) in LeuT (PDB-ID:2A65, 1.65Å, blue) is coordinated between TM1 and TM8 via interaction with the backbone of the unwound helix segment in TM1 and side chain interactions from TM8 (left). Comparison of this ion binding site with the structure of *Mj*ApcT (PDB-ID: 3GIA, 2.32Å, green) revealed a similar coordination of Lys158, interacting with the backbone of TM1 and the side chain of Ser283 of TM8 (middle). Mutation of Lys158 to alanine resulted in the loss of transport activity and it is assumed that this is the main residue involved in the proton-coupling mechanism of *Mj*ApcT. Arg262 in the antiporter CaiT (PDB-ID: 2WSW, 2.29Å, purple) is also coordinated by residues in TM 1, 5 and 8 (right). Mutations on this residue converted CaiT into an Na^+-dependent transporter.

The uptake measurements on the *Mj*ApcT-Lys158Ala mutant were performed without a sodium chloride electrochemical gradient present. Thus, it would be of interest to investigate this mutant in respect to sodium-dependent alanine uptake similar to the research performed with the CaiT mutants.

However, it is still not known whether Lys158 of *Mj*ApcT is involved in a protonation-deprotonation cycle to drive substrate transport, since alanine transport by *Mj*ApcT has only been shown to be pH-dependent but the evidence of a proton-coupling mechanism is lacking. Therefore, we assume that Lys158 might stabilise the protein in a specific conformation during the transport cycle as proposed for Arg262 of CaiT. Both theories explain the loss of activity for *Mj*ApcT when this lysine is mutated to alanine. It remains to be determined whether the mutation of the lysine results in the transporter being locked into an occluded state, thereby preventing substrate binding, or if it stops the transporter from adopting other conformations.

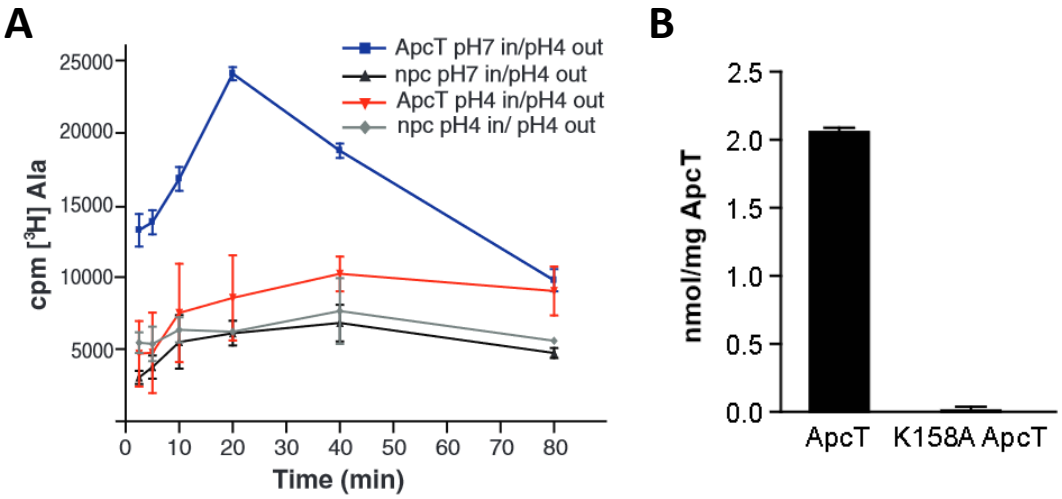


Figure 1-28: Alanine uptake by MjApcT and the influence of Lys158

Liposome-based assays of *Mj*ApcT were performed in presence and absence of a pH gradient (Shaffer *et al.*, 2009). The uptake of radiolabelled alanine was monitored over time for liposomes containing *Mj*ApcT (ApcT) or devoid of protein (npc). Enhanced alanine uptake was observed when a pH gradient of 3 pH units was applied (A). Since Lys158 was suggested to be responsible for a proton-coupling mechanism, mutation to alanine abolished alanine uptake (B). Figures adapted from Shaffer *et al.*, 2009.

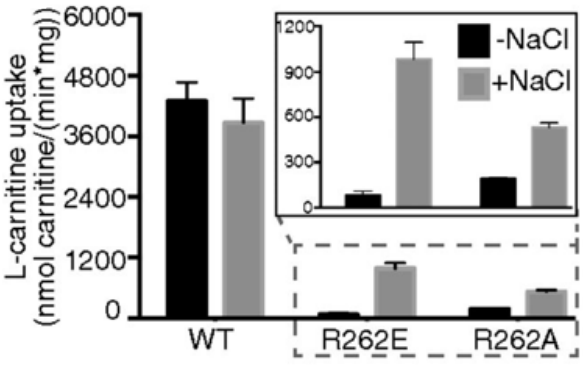


Figure 1-29: Mutational studies on CaiT modulating its sodium sensitivity

Proteoliposomes containing either CaiT-WT, CaiT-Arg262Glu, or CaiT-Arg262Ala were used for counterflow measurements with 10 mM unlabelled γ-butyrobetaine in the internal buffer and 40 µM [^{14}C] l-carnitine in the external buffer. Uptake was measured in absence (-NaCl) or presence (+NaCl) of 50 mM sodium chloride to investigate Na^+-dependent transport activity. Figure adapted from Kalayil *et al.*, 2013.

1.5 Advances in membrane protein expression, purification and crystallisation

Membrane proteins play a vital role in many cellular processes and are one of the main targets for drug development (Overington *et al.*, 2006; Bull and Doig, 2015). It took 25 years from the first crystal structure of a soluble protein (Kendrew et al., 1960) to the structure of a membrane protein (Deisenhofer *et al.*, 1985) being obtained. Ever since, the number of membrane proteins structures deposited into the protein data bank (PDB) has exponentially increased (Figure 1-30) but remains lower than that of soluble proteins (3227 membrane protein structures vs. 131411 protein structures in the PDB as of 16. June 2017, http://pdbtm.enzim.hu/?_=/statistics/growth). This is due to the challenges in over-expression, solubilisation, purification and crystallisation of membrane proteins.

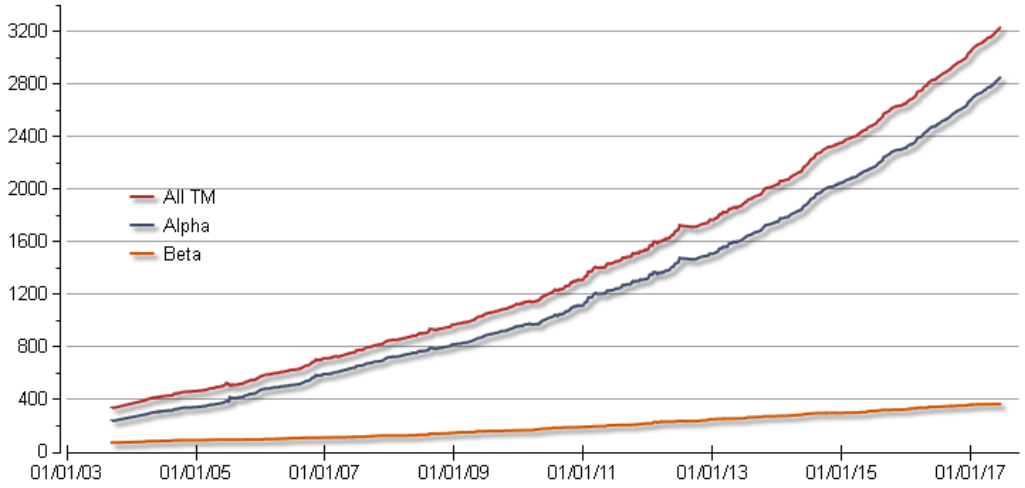


Figure 1-30: Growth of submitted structures of membrane proteins to the PDB based on www.tcdb.org

Correct folding and insertion of the protein into the membrane is a crucial step during protein expression. Once the protein has been successfully over-expressed, the next challenge is to isolate the membrane protein from its lipid environment using detergents.

Detergents are amphiphilic, possessing both hydrophilic and hydrophobic properties, molecules consisting of a hydrophilic head with different chemical properties and a hydrophobic tail of varying length (Iwata, 2003). They can be categorised based on their head group into ionic, non-ionic or zwitterionic detergents. Depending on the chemical nature of the head group, they have different stabilisation and destabilisation effects on the membrane protein, which are often referred to as a "mild" or "harsh" property (Iwata, 2003; Privé, 2007). "Harsh" detergents, mostly ionic detergents, have high solubilisation efficiencies but tend to denature or destabilise the membrane protein whereas "mild" detergents, such as sugar based detergents, are more likely to keep the protein stable in solution (Iwata, 2003; Privé, 2007). Nevertheless, pre-testing of suitable detergents to be used during the purification is a vital step. The most commonly used detergents for purification and crystallisation are dodecyl maltopyranoside (DDM), decyl maltopyranoside (DM), octyl glucopyranoside (OG) and n-lauryl dimethylamine n-oxide (LDAO). They have been the most successful in the crystallisation of numerous of membrane proteins (Parker and Newstead, 2012, 2016). Crystallising the proteins in a detergent can have the negative effect of diminishing the diffraction quality of the crystals, resulting from the formation of a detergent belt around the membrane protein, thus decreasing the surface area to form crystal contacts (Figure 1-31) (Iwata, 2003; Bill *et al.*, 2011). Nevertheless, detergents can in some cases be successful in providing crystals giving high resolution structures of membrane proteins (Parker and Newstead, 2016).

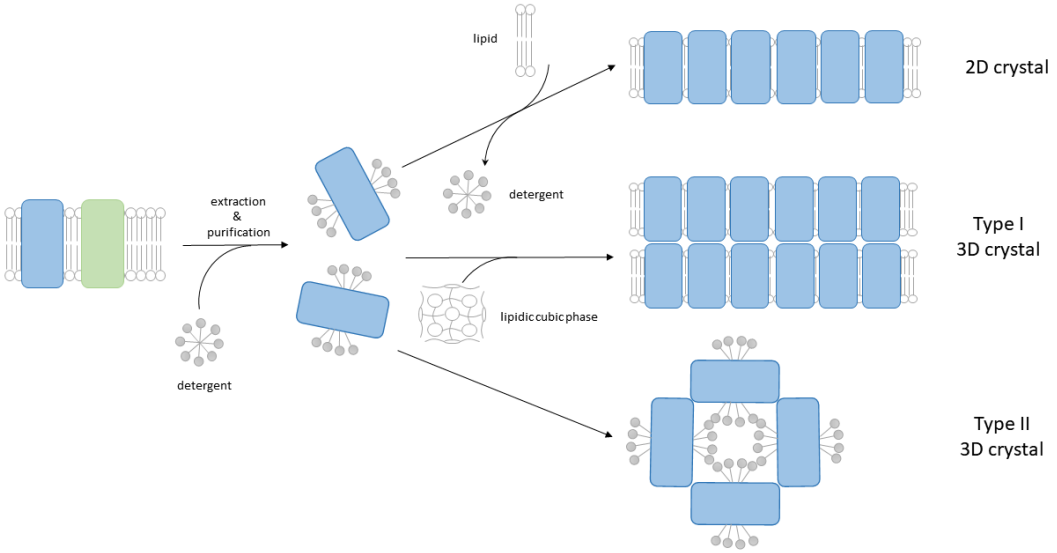


Figure 1-31: Influence of detergents and lipids on crystal formation

The membrane protein of choice (blue) is isolated from the lipid bilayer of the cell membrane with detergents. The detergent micelle forms a belt around the hydrophobic part of the membrane protein. The purification process includes the separation of the protein of choice from other membrane proteins (green). Depending on the crystallisation experiments and upon addition of lipids to the protein, different crystal forms can grow: 2D crystal, Type I and Type II 3D crystals.

Recent advances have resulted in protocols and methods tackling the hurdles in obtaining a sample of pure membrane protein of interest. Specific *E.coli* strains, such as the *E.coli* C41 and C43 strains (Miroux and Walker, 1996; Dumon-Seignovert *et al.*, 2004), contain mutations designed to increase the yields during membrane protein over-expression, that would otherwise be toxic to the cells. The addition of a GFP protein to the C-terminus of the membrane protein not only introduces a way to determine the protein yield through fluorescence measurements, but also acts as a reporter for correct membrane insertion and protein folding (Drew *et al.*, 2001, 2008). When using a membrane protein-GFP fusion construct, the solubilisation efficiency for a given detergent can be determined. Fluorescent size exclusion chromatography (FSEC) can also be performed in order to ensure that the protein is still properly folded and not aggregating when using a certain detergent (Kawate and Gouaux, 2006).

A monodisperse peak in the fluorescence trace indicates that the protein is stable and folded in the detergent, whereas a polydisperse peak is associated with unfolded or unstable protein (Kawate and Gouaux, 2006). The right choice of detergent for solubilisation and

purification ensures that the protein is properly folded, which increases the chances of crystallisation (Sonoda *et al.*, 2011; Deller *et al.*, 2016).

Over the past decades, several optimisation strategies for the crystallisation of membrane proteins have been developed, such as reintroducing lipids into the protein detergent mixture by adding phospholipids or reconstitution of the proteins into bicelles (Wang *et al.*, 2012; Tang *et al.*, 2016; Lenaeus *et al.*, 2017). Additionally, reconstitution of the membrane protein of choice into a lipid environment, the so-called lipidic cubic phase (LCP), was first described in 1996 (Landau and Rosenbusch, 1996). This method represents a complementary approach to vapour diffusion crystallisation of membrane proteins, and will be described in more detail elsewhere in this thesis. In short, the membrane protein is incorporated into a lipid bilayer that forms an overall three-dimensional network of lipid channels (i.e. the lipidic cubic phase) filled with an aqueous solution (the buffer or precipitant) (Rummel *et al.*, 1998). Through the addition of the precipitant solution, a phase transition is initiated, allowing the embedded proteins to make contacts through which protein crystals can form (Nollert, 2003).

Membrane proteins are dynamic entities and when extracted from the membrane can exist in different conformational states. Together with their limited contact surface area because of the detergent micelle around the membrane protein, this conformational heterogeneity can decrease the likelihood of crystal formation (Bill *et al.*, 2011). To support crystallisation, nanobodies and Fab-fragments can be used, so-called crystallisation scaffolds, to provide crystal contacts, or to capture the protein in a specific conformation, thus reducing the conformational heterogeneity (Tereshko *et al.*, 2008). This method has proven to be very valuable, especially in the studies of the different conformational states of LeuT (Krishnamurthy and Gouaux, 2012), but it has also facilitated the crystallisation of novel proteins, such as the serotonin transporter (Coleman et al., 2016) and of a fluoride channel (Stockbridge *et al.*, 2015). Another advantage of the crystallisation scaffolds is that, due to their low structural diversity, published structures of nanobodies and antibody fragments can be used to provide initial phase information in the structure solution process after X-ray diffraction data collection from the crystals. Though using crystallisation scaffolds is a beneficial method for crystallising challenging proteins, it is also cost- and time-intensive as there is limited availability of in-house facilities, and ordering specific antibodies or nanobodies from industrial companies can further increase the costs. In the case of GPCRs, the introduction

of a crystallisation scaffold protein into the membrane protein apocytochrome b562 (BRIL) has provided a common method for crystallising GPCRs (Chun *et al.*, 2012). Either a BRIL protein or T4 lysozyme is introduced into a loop region between two transmembrane helices or at the N-terminus, and has the additional benefit of providing a molecular replacement model, as structural data for both proteins are available (Engel *et al.*, 2002; Rosenbaum *et al.*, 2007). However, the insertion of the crystallisation scaffold can introduce complications in folding. Therefore, identifying the correct location for the scaffold into the protein is crucial and requires rational design (Chun *et al.*, 2012).

In conclusion, the challenge of working with membrane proteins has seen significant progress over the past years. Though the specific path to take towards a pure, homogenous protein and a protein crystal structure can still be a difficult decision, preliminary small-scale tests for expression and purification are invaluable in this process.

1.6 AIMS OF THIS THESIS

The aim of the work described in this thesis is to understand the structural basis of amino acid transport via the SLC7 family of cationic amino acid transporters. Functional data obtained by mutational studies and amino acid uptake into oocytes revealed some mechanistic aspects of this family such as their specificity for arginine, lysine and in some cases also histidine and that the affinity for the substrates can be modulated by point mutations in an intracellular part of the transporter, as reported for CAT2A and CAT2B. Nevertheless, to fully understand why this region has such an important effect, and how the transporters discriminate between their substrates and couple the transport, relies on structural data. To date, no close homologue in complex with an amino acid is available yet for this type of transporter. As for other eukaryotic transporters, studies on the prokaryotic homologues, which are usually easier to purify and are more stable, can shed light into the transport mechanism of a whole family of transporters as seen for LeuT.

A range of experimental procedures were necessary to understand this transporter family including cloning, small-scale and large-scale expression as well as further structural and functional studies on a suitable candidate (Chapter 2)

Extensive screening of prokaryotic *Mj*ApcT homologues (Chapter 3), which showed a different evolutionary relationship to the human SLC7 members, identified *Gk*ApcT as a suitable candidate for functional and structural studies to gain insight into the SLC7 family.

In general, there is little known about proton-coupled APC transporters, since most of the research has been focussed on the sodium-coupled neurotransmitter transport family and their related proteins. To date, only one structure of the occluded state of *Mj*ApcT from *Methanocaldococcus jannaschii* has been published, a transporter that is suggested to be proton-coupled although further evidence is necessary.

Thus, this thesis aims to understand the structural basis for amino acid transport via the SLC7 family, how different amino acids are recognised and how the proton-coupled transport of substrate is utilised in order to develop a platform for further studies on human CAT and HAT transporters in the future.

Initial functional studies performed on *Gk*ApcT identified the protein as a proton-coupled amino acid transporter specific for small neutral and polar amino acids (Chapter 4). The conditions under which substrate uptake of *Gk*ApcT can be monitored were established for further studies.

The thesis presents *Gk*ApcT as the first proton-coupled APC transporter studied of which a crystal structure in a substrate bound conformation has been obtained (Chapter 5). Additionally, the structure of a humanised *Gk*ApcT mutant with arginine bound was solved (Chapter 6). *In vitro* studies were performed to identify the coupling mechanism of *Gk*ApcT and investigate residues that might be important for substrate binding based on the structural data available (Chapters 6). Mutational studies were carried out to pinpoint residues involved in the coupling mechanism to investigate whether the mechanisms are independent of one another (Chapter 6). The knowledge obtained by the structural and functional studies of this closely related SLC7 homologue identified a single residue to influence the substrate specificity of the transporter (Chapter 6). Additional mutational studies guided by the research on the human CATs revealed that the residue modulating CAT2As substrate affinity has the same effect on *Gk*ApcT's affinity towards its substrate, making *Gk*ApcT a good model for understanding the human CAT proteins.

Additionally, serial crystallography experiments were performed on *Gk*ApcT investigating the possibly to obtain structural information with this method and further phasing information (Chapter 7).

2 Materials and methods

2.1 EQUIPMENT AND CHEMICALS

Chemicals used in this study were purchased from Sigma Aldrich, Fisher Scientific, Invitrogen and Avanti Polar Lipids unless specified otherwise in the text.

The equipment and instruments used are listed below:

- TL-100 ultracentrifuge (all from Beckman Coulter), 5430R (Eppendorf)
- Innova 44 shaker-incubators
- iCycler PCR thermal cycler (Bio-Rad)
- NanoDrop ND-1000 spectrometer
- SpectraMax M3 fluorescence plate reader (Molecular Devices)
- E1061 cell disruptor (Constant Systems) with Neslab ThermoFlex1400 cooling unit (Thermo Scientific)
- ÄKTA Purifier system (GE Healthcare)
- Mosquito® pipetting robot (TTPLabtech)
- 100 µl gas-tight syringes and syringe coupler (Art Robbins)
- Gryphon LCP robot (Art Robbins)
- Ministrel HTUV imager (Rigaku)
- Gallery incubator (Rigaku)
- Wallac 1409 DSA liquid scintillation counter
- Carry Eclipse Fluorescence Spectrophotometer (Agilent Technologies)
- Mini extruder (Avanti Polar Lipids)
- *in situ* plates (MiTeGen and Molecular Dimensions Ltd.)
- MSC plates (SwissCi)
- V-bottom mini plates (Fischerbrand)

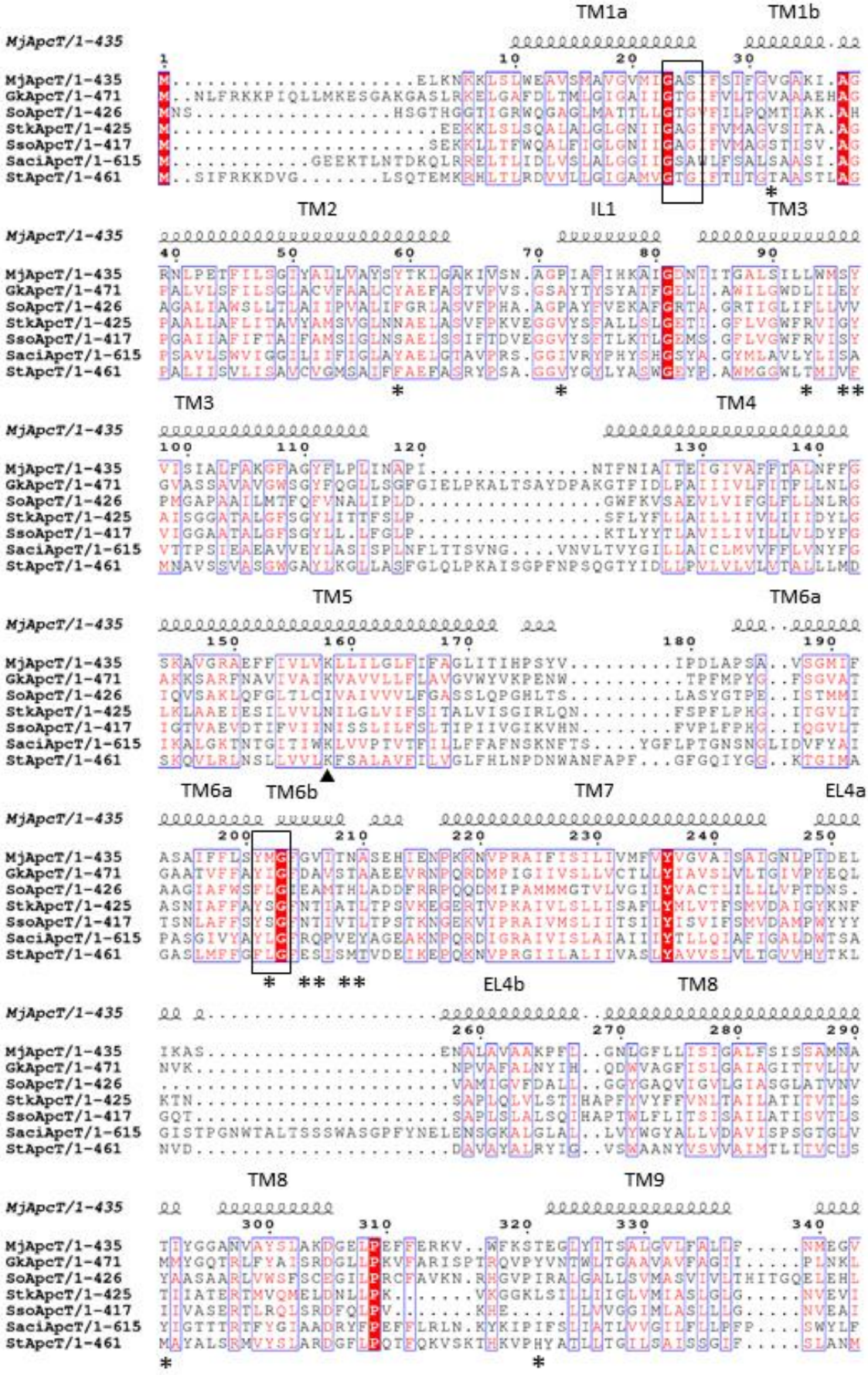
2.1.1 Screening for homologues of the archaeal transporter ApcT

One bottleneck in determining the structure of a membrane protein is to obtain high quantities of stable and folded protein for crystallisation trials. Therefore, screening a number of homologues has been a successful approach in order to increase the chances of obtaining sufficient quantities of pure protein (Robinson, 2011). Additionally, the use of homologues from thermophilic organisms or artificially thermostabilised proteins can have a positive effect on the ability of the protein to crystallise (Deller *et al.*, 2016).

Here, homologues of the archaeal *Mj*ApcT transporter were identified by searching with its sequence against the protein sequence entries of the UniProt library in BLAST. Archaeal and bacterial organisms for which the genomic DNA was available were used in the search. The homologues that were screened as suitable candidates for protein expression, protein purification and crystallisation are listed in Table 2-1.

Table 2-1: ApcT homologues for functional and structural studies

Protein	**Organism**	**Accession code**	**MW [kDa]**	**Seq. sim. [%] to *Mj*ApcT**
***Mj*ApcT**	*Methanocaldococcus jannaschii*	NP_247591.1	47.3	100
***Gk*ApcT**	*Geobacillus kautophilus*	YP_146783.1	50.8	43.4
***Saci*ApcT**	*Sulfolobus acidocaldium*	YP_256768.1	67.2	31.7
***St*ApcT**	*Streptococcus thermophilus*	YP_139794.1	49.8	43.8
***Stk*ApcT**	*Sulfolobus tokodaii*	NP_376756.1	44.8	43.1
***So*ApcT**	*Shewanella oneidensis*	NP_720082.2	45.1	40.5
***Sso*ApcT**	*Sulfolobus solfatorius*	NP_342901.1	45.1	42.2


TM1a
TM1b
MjApcT/1-435
MjApcT/1-435
GkApcT/1-471
SoApcT/1-426
StkApcT/1-425
SsoApcT/1-417
SaciApcT/1-615
StApcT/1-461
TM2
IL1
TM3
TM3
TM4
TM5
TM6a
TM6a
TM6b
TM7
EL4a
EL4b
TM8
TM8
TM9

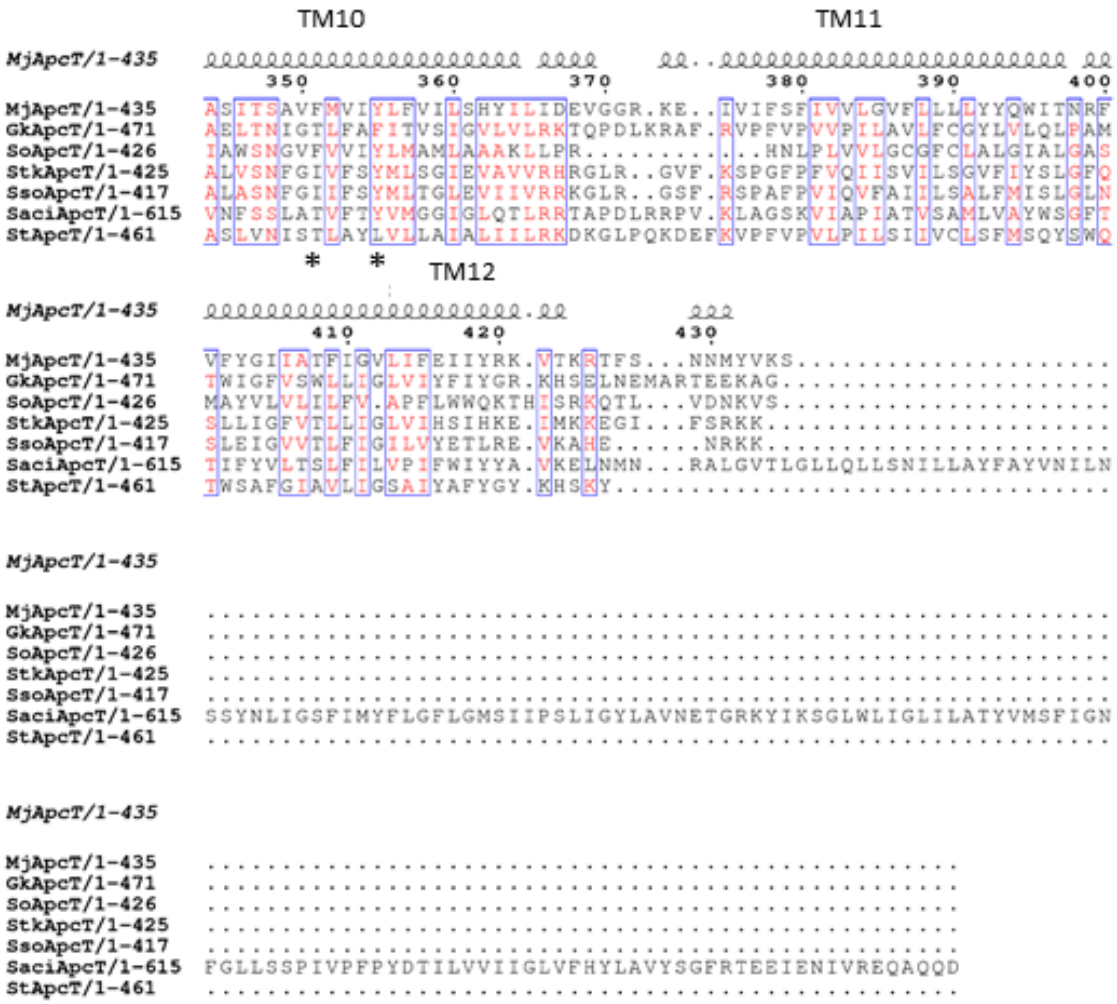


Figure 2-1: sequence alignment GkApcT, MjApcT, SaciApcT, SoApcT, StApcT, StkApcT and SsoApcT

The sequence alignment (accession codes are listed in Table 2-1was generated using PROMALS3D (Pei and Grishin, 2014). The alignment is coloured by conservation (high conservation – red background, low conservation – red letters, no conservation – white background and black letters Secondary structure elements, based on the published structure of *Mj*ApcT (3GIA, 2.32 Å), are given above the alignment and the transmembrane helices (TM), the intracellular loop (IL) and extracellular loop (EL) region are labelled. Additionally, the two GSG-motifs are highlighted by the black box and the residues involved in forming the waterfilled cavity of *Mj*ApcT (*) and the proposed residue for porton-coupling (▲) are indicated.

2.1.2 Preparation of competent cells

For cloning, the competent *E.coli* Omnimax cell line was used and for expression studies, the vector was transformed into the competent cell line C43(DE3). For both, cell cultures were grown in 100 ml lysogenic broth (LB) media and shaken at 37 °C until an OD_{550} of 0.4 was reached. The cell culture was evenly divided into 50 ml falcon tubes and cooled on ice

for 15 min. The cells were pelleted by centrifugation at 4°C and 3500 g for 8 min. The pellet was resuspended in 40 ml (0.4 x of initial culture volume) of TfbI (30 mM KAc, 100 mM RbCl, 10 mM $CaCl_2$, 50 mM $MnCl_2$ and 15 % glycerol, pH 5.8). Cells were pelleted as described above and resuspended in 4 ml (0.04 times initial culture volume) TfbII (10 mM MOPS, 75 mM $CaCl_2$, 10 mM RbCl and 15 % glycerol, pH 5.8). The solution was kept on ice for 15 min, divided into 100 µl aliquots, flash frozen and stored at -80 °C.

2.1.3 Transformation of the pWaldo-plasmid into competent cells

A 100 µl aliquot of competent cells was mixed with 0.5-2 µl of plasmid DNA and incubated on ice for 20 min followed by a heat shock treatment at 42 °C for 45 s. The sample was kept on ice for 2 min before 700 µl of LB were added and the cells were grown at 37°C while being shaken for 1-2 h. The cells were pelleted by centrifugation at 4000 g for 2 min, and the pellet was resuspended in 100 µl of LB and plated onto LB agar plates with 50 µg/ml kanamycin (KAN). The plate was kept at 37 °C overnight for colonies to grow. After this incubation period, the plate containing the bacterial colonies could be kept at 4 °C for up to 10 days.

2.1.4 Cloning and mutagenesis of ApcT homologues and YneM

The gene coding for the protein of interest was amplified from the genomic DNA by PCR. The conditions for the PCR run are given in Table 2-2. The success of the PCR was evaluated by running each sample on a 1% (1.5 % for YneM) agarose gel in 1x TAE buffer (40 mM Tris-HCl pH 7.6, 20 mM acetic acid, 1mM EDTA), at 110 V for 30 min. The band on the gel running at the size of the amplified gene of interest was purified from the gel using the GeneJET Gel extraction kit from Thermo scientific. For this, the band was cut out carefully from the agarose gel and transferred into a reaction tube. Binding buffer was added in a 1:1 ratio (w/w) and incubated at 50°C for 10 min until the gel piece was completely dissolved. The solution was transferred into a GeneJet purification column and centrifuged for 1 min at 12,000 g followed by a washing step with ethanol-containing washing buffer, again at 12,000 g for 1 min. Residual buffer was removed by another centrifugation step (12,000 g for 1 min). 50 µl

of water was added to the membrane of the GeneJet purification column and left for 5 min at room temperature.

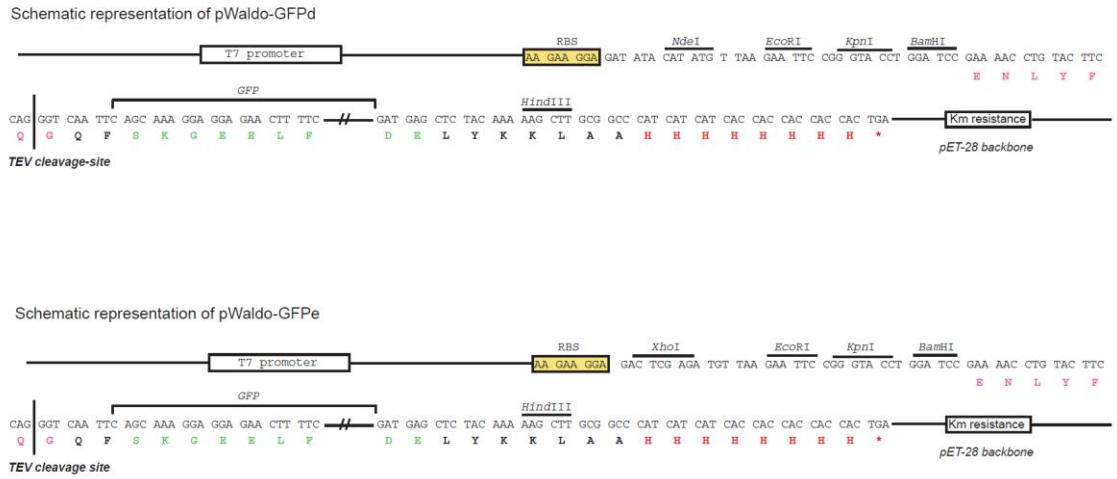


Figure 2-2: Representation of the pWaldo-GFP plasmids.

The gene of interest is inserted either into the pWaldo-GFPd plasmid (restriction sites: NdeI & BamHI) or the pWaldo-GFPe plasmid (restriction sites: XhoI & BamHI). Both plasmids are under the control of the T7-promotor. The promotor sequence is followed by a ribosomal binding site (RBS) and multiple cloning sites (MCS) into which the gene of interest is inserted. A TEV cleavage site lies between the gene of interest and a C-terminal bacterial GFP sequence, followed by an octa-histidine tag and a gene for kanamycin resistance.
Figure adapted from Drew *et al.*, 2006.

The DNA was eluted from the membrane by centrifugation at 12,000 g for 2 min. The concentration of the eluted sample was determined with a NanoDrop ND-1000 spectrometer. The sample was stored at -20°C until further use. For insertion of the gene of interest into the pWaldo-GFP vector (Figure 2-2), both the vector and target DNA were digested with the appropriate restriction enzymes for 2 h at 37 °C. As described above, the GeneJet gel extraction kit was used to extract the DNA from the reaction mixture. The digested DNA and vector were subsequently ligated using the Quick-Stick Ligase (Bioline).

The reaction was performed at room temperature for 20 min and 10 µl of the reaction mixture was transformed into OMNIMAX cells and plated on LB-agarose plates containing 50 µg/ml kanamycin as selective antibiotic. The plates were incubated overnight at 37 °C. The next day, single colonies were picked and inoculated into 10 ml LB (+ kanamycin), shaken at 37 °C overnight and harvested the following day (1000 x g, 10 min, 4 °C). Successful insertion

into the pWaldo-vector was tested by digesting the plasmid with the corresponding restriction enzymes and running the reaction on an agarose gel. Colonies that showed a positive digestion pattern were sequenced (using the Source Bioscience service).

Table 2-2: PCR conditions to amplify the gene of interest from a genomic DNA template

step	T [°C]	Time [min]	cycles
Denaturation	94	2	1
Denaturation	94	0.25	30
Annealing	55	0.15	
Elongation	70	1	
Elongation	70	7	
End	4	∞	1

2.1.5 Mutagenesis

Mutagenesis primers were designed and used for amplification of the whole pWaldo-plasmid using a KOD hot start kit (Novagen). The reaction mixture (50 µl final volume) contained 0.4 µM forward primer, 0.4 µM reverse primer and 200 µM of each of the dNTPs, 1.5 mM $MgSO_4$, 1x KOD buffer and 0.8 µl *Thermococcus kodakaraensis* KOD1 DNA polymerase. A standard PCR protocol (Table 2-3) for mutagenesis reactions was used. The final PCR product was supplemented with DpnI (DpnI gene from *Diplococcus pneumoniae* G41) as a restriction enzyme which specifically digests methylated DNA, which in this form is only present as the mother plasmid. This step ensures removal of the mother plasmid from the mutated plasmid, which is not methylated. After an incubation period of 2 h at 37 °C, 10-25 µl of the reaction solution was transformed into OMNIMAX cells as described above. Individual colonies were picked and grown overnight in 10 ml LB. The cells were pelleted for 10 min at 4000 x g and the plasmid harvested by using the mini prep kit (Thermo scientific). Successful insertion of the point mutation was verified by sequencing (using the Source Bioscience service: https://www.sourcebioscience.com/).

Table 2-3: PCR conditions to amplify the whole plasmid with mutagenesis primers

step	T [°C]	Time [min]	Cycles
Denaturation	94	2	1
Denaturation	94	0.25	17
Annealing	55	1	
Elongation	68	8	
Elongation	68	7	1
End	4	∞	1

2.1.6 Preparation of the C43Δ*yneM E.coli* strain

The protocol for the gene knock-out of YneM was provided by the group of Colin Kleanthous, Department of Biochemistry, University of Oxford, and was based on that described in Lee *et al.*, 2009. In short, 300 bp sections up- and downstream (H1 and H2, Figure 2-3) of the *yneM* gene were cloned into the pDOC-K vector into cloning region 1 and 2 (CR1 and CR2), respectively, flanking the kanamycin cassette (Figure 2-3). The pDOC-K vector containing H1 and H2 (pDOC-K-H1-H2, Figure 2-3) was transformed via electroporation into the C43(DE3) *E. coli* strain, used for expression of *Gk*ApcT.

Activation of the pABSCE vector, transformed into the same cells, induced recombination of the kanamycin cassette into the gene locus of *yneM* (Figure 2-3). Cells were plated onto LB agar plates containing kanamycin.

To verify for the loss of the pABSCE vector, a colony was spread onto LB agar plates with chloramphenicol (CAM), it was expected that no colonies would grow in the presence of this antibiotic. The KO strain was then cured from the kanamycin resistance according to a standard protocol (Baba *et al.*, 2006), since the pWaldo plasmid carrying the *Gk*ApcT gene is also kanamycin resistant. Successful gene knock-out was verified by PCR.

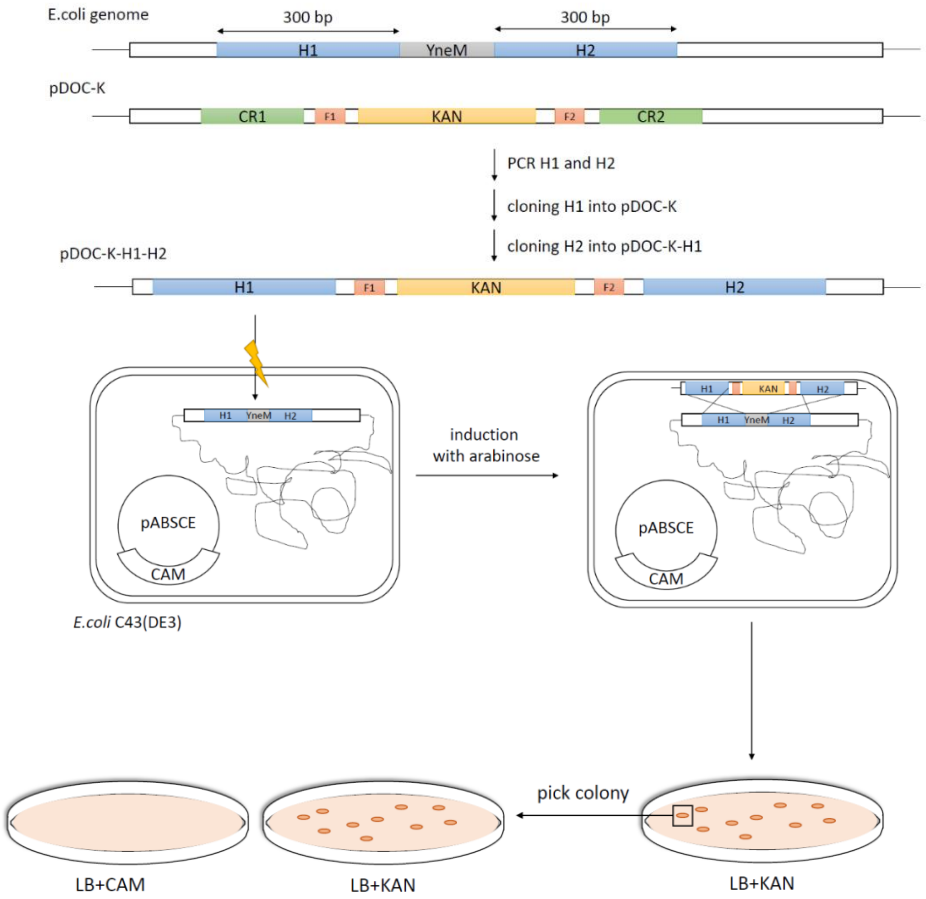


Figure 2-3: YneM gene knock out protocol schematic

2.1.7 Small-scale expression in *E.coli*

Colonies of transformed C43(DE3) *E.coli* cells were inoculated into 10 ml of LB (+KAN) and shaken overnight at 37 °C. The following day, 1 ml of this starting culture was inoculated into 100 ml of LB (+ KAN), again shaking at 37 °C constantly. The OD_{600} was measured and isopropyl β-D-1-thiogalactopyranoside (IPTG), to a final concentration of 400 µM, added to induce expression when the OD_{600} reached 0.6-0.8. The temperature was reduced to 25 °C and the culture was shaken overnight. 1 ml of the overnight culture was spun down and the pellet resuspended into 100 µl 1x phosphate buffered saline (PBS). The GFP content of the cell suspension was measured with a Spectra Max M3 fluorescent plate reader (Molecular Devices) at $\lambda_{excitation}$=490 nm; $\lambda_{emission}$=512 nm. The expression level of the desired protein was estimated based on the method described in Drew *et al.* 2008.

2.1.8 Fluorescent size exclusion chromatography

10 ml cultures from small scale expression experiments were harvested by centrifugation at 3900 g for 10 min at 4 °C, resuspended into 0.5-1 ml of 1xPBS and frozen at -80 °C to initially lyse the cells for subsequent cell disruption. The cell suspension was thawed and sonicated on ice four times with 1 min intervals by using VirSonic475 (Virtis) for which each cycle was programmed as a 1 s sonication time in 1s intervals for a total sonication time of 10 s. The lysed cells were centrifuged at 10,000 g for 15 min to pellet unbroken cells. The supernatant was transferred into a fresh tube and centrifuged at 200,000 g for 20 min to pellet the membranes. The membrane pellet was resuspended in 1xPBS, frozen in liquid nitrogen and stored at -80 °C or solubilised directly into buffer containing 1xPBS and 1 % n-Dodecyl β-D-maltoside (DDM). It was gently mixed at 4 °C for 1 h followed by a centrifugation step at 200 000 g for 20 min at 4 °C to separate the insolubilized fraction. To determine the solubilisation efficiency, the GFP fluorescence of samples before and after the centrifugation was measured. 200 µl of the solubilised fraction were loaded onto a Superose^{TM} 6 10/300 GL (GE Healthcare) column, pre-equilibrated in buffer containing 20 mM Tris-HCl pH 7.5, 150 mM NaCl, and 0.03% DDM. During the run, the GFP fluorescence was monitored ($\lambda_{excitation}$ = 488 nM, $\lambda_{emission}$ = 512 nM).

2.1.9 Large-scale expression in *E.coli* and cell harvest

Freshly transformed C43(DE3) *E.coli* cells were inoculated into 100 ml of LB media as starter culture and incubated overnight at 37 °C. The starter culture was then inoculated into 3-6 L of terrific broth (TB) media in a 1:100 ratio. The culture was induced with IPTG to a final concentration of 0.4 mM at an OD_{600} of 0.6-0.8 and the temperature was decreased to 25 °C for overnight expression. Cells were harvested and resuspended in 1xPBS and stored at -80 °C. To prepare the cell suspension for the purification process, the suspension was thawed and the cells disrupted with a TS series continuous cell disruptor (Constant Systems, UK) running at 30 kpsi and 4 °C. The cellular debris was removed by centrifugation at 20,000 g at 4 °C, and whole cell membranes were isolated (centrifugation at 45,000 g, 2 h, 4 °C). Membranes were resuspended in ice-cold 1xPBS and flash-frozen in 50 mL falcon tubes into liquid nitrogen. The presence of GFP-tagged membrane protein was monitored by taking

100 µl samples at each step and measuring the remaining GFP-fluorescence ($\lambda_{excitation}$ = 488 nM, $\lambda_{emission}$ = 512 nM).

2.2 PURIFICATION

The prepared membranes were thawed and kept on ice. NaCl and PBS were added to a final concentration of 150 mM NaCl and 1xPBS, respectively. To solubilise the membrane proteins, DDM (Glycon) was added to a final concentration of 1% (w/v). The solution was kept at 4 °C, stirred for 1 h and then ultra-centrifuged at 45,000 g, 4 °C for 1 h. 10 ml of nickel nitrilotriacetic acid (Ni-NTA) resin were washed with water and pre-equilibrated with 1xPBS. The resin and the appropriate concentration of imidazole (Incubation step, see Table 2-4), were added to the supernatant and the mixture was stirred for 2-3 h to allow protein binding to the resin. Samples were taken before and after the incubation with the resin, and the binding efficiency determined via measuring the GFP-fluorescence of both samples. If the binding efficiency was below 50 %, the incubation step was prolonged until the binding efficiency reached a plateau. The resin-protein mixture was poured into a glass column and washed with different imidazole concentrations to remove any other contaminants (Wash1 and Wash2 steps, see Table 2-4). The membrane protein-GFP-8xHis fusion was eluted off the resin with 250 mM Imidazole. The flow-through of each step was collected and samples taken to monitor the presence of the desired membrane protein via GFP-fluorescence. The total amount of GFP-containing protein in the final elution step was estimated, and tobacco etch virus (TEV) protease (1 mg/ml) was added in a 1:1 ratio. To remove any residual imidazole that could hinder the cleavage reaction and further purification steps, the sample was dialysed overnight in dialysis buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl and 0.03% DDM) in a Spectra dialysis membrane with a 3.5 kDa cut-off. The next day, a 5 ml HisTrap™ column (GE Healthcare) stored in 20% ethanol was washed with water and equilibrated with the dialysis buffer using a peristaltic pump running at 2 ml/min. The sample was filtered through a 0.22 µM MILLEX™ GP filter unit to remove any precipitated protein. The sample was then loaded onto a HisTrapTM column. During this step, any residual proteins interacting with the Ni-NTA resin and the His-tagged TEV and GFP would be bound by the resin. The untagged protein was collected in the flow-through and concentrated to 500 µl using an

Amicon centrifugal filter unit (cut-off 50 kDa). The HisTrapTM column was washed with 500 mM Imidazole to elute the residual proteins (His-tagged TEV and GFP).

The concentrated protein was loaded onto a SuperdexTM200 10/300 (S200) column (GE Healthcare) pre-equilibrated in either reconstitution buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.3 % (w/v) DM), or crystallisation buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, and 0.03 % (w/v) DDM). The flow rate for each run was 0.4 ml/min and the absorbance at 280 nm was monitored. Samples of the fractions expected to contain the protein were taken and loaded onto a 12% Tris-glycine gel to validate the purity of the protein in the final step. Additionally, samples taken during the purification were loaded onto the same gel to monitor the loss of protein during the purification process and if necessary, the protocol was altered accordingly. For example, imidazole was added to the sample before loading it onto the HisTrapTM column in subsequent purifications if the SDS gel showed that GFP-free protein remained bound to the column.

Fractions containing the protein of interest were pooled together and the concentration of the protein was estimated using a NanoDrop^{TM} spectrophotometer. For this, the specific extinction coefficient (ε) of the protein was calculated using ProtParam (http://web.expasy.org/protparam/) based on its amino acid sequence. The protein was concentrated to a final value of ~10 mg/ml or 20 mg/ml depending on the crystallisation trial.

Table 2-4: Different Imidazole concentrations (in mM) used for individual purification of ApcT homologues

Protein	Incubation	Wash 1	Wash 2	Elution
***Mj*ApcT**	15	15	30	250
***Gk*ApcT**	20	20	30	250
***St*ApcT**	30	30	40	250
***So*ApcT**	30	30	40	250

2.3 Reconstitution of membrane transporters into Liposomes

In order to be able to understand the transport mechanism of ApcT, the reconstitution of this protein into liposomes, lipid vesicles, was performed. This method has the advantage

that the outside and inside environment of the vesicles can be directly manipulated to identify the pH-dependence, electrogenicity or substrate specificity of the wildtype transporter and its mutants.

2.3.1 Preparation of lipids

1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine (POPE) and 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-(1'-*rac*-glycerol) (POPG), or *E.coli* polar lipids and phosphatidylcholine (PC) were mixed in a 3:1 (w/w) ratio in a round bottom flask under a fume hood. The mixture was dried in a rotary evaporator. The dried lipids were washed twice with 10 ml propane and dried again. The lipids were resuspended in degassed storage buffer (50 mM KPi, pH 7) to a final concentration of 20 mg/ml, aliquoted into 1 ml samples, and freeze thawed three times until snap frozen for final storage at -80 °C. Prior to reconstitution, lipids were thawed and diluted to a final concentration of 5 mg/ml in storage buffer (50 mM KPi, pH 7) and extruded 11 times through a 0.4 µM filter with a mini extruder kit (Avanti Polar Lipids).

2.3.2 Reconstitution into liposomes by rapid dilution

Purified protein in reconstitution buffer (150 mM NaCl, 20 mM Tris-HCl pH 7.5 and 0.3% DM) was reconstituted into POPE:POPG (3:1 ratio w/w) liposomes. Protein and lipids were mixed in a 1:60 ratio (protein:lipid) and incubated at room temperature for 15 min. The mixture was then diluted into storage buffer (50 mM KPi, pH 7) and pelleted by ultracentrifugation (4 °C, 45,000 g, 2 h). The pellet was resuspended into storage buffer and dialysed in 3 L storage buffer overnight with two changes of buffer in between to remove any residual detergent. The dialysed samples were harvested by centrifugation at 60,000 rpm (TLA120.2 rotor, benchtop ultra) at 4 °C for 25 min. The pellet was resuspended in storage buffer to a final concentration of 0.5 µg/µl of protein and freeze thawed three times in liquid nitrogen. The reconstitution efficiency was determined by running 2 µg of the sample together with a concentration standard of the protein on a 12% Tris-glycine SDS gel.

2.4 TRANSPORTER ASSAYS

On the day of the experiment, proteoliposomes were thawed at room temperature and the required volume for the assay removed from the proteoliposome stock. Unless indicated otherwise, 5 µg of proteoliposomes per experimental time point were used. In the case of experiments with multiple time points, the final volume was a multiple of the volume needed for one time point. The proteoliposomes were pelleted at 50,000 rpm at 4 °C and the pellet resuspended in inside buffer to a final volume of 450 µl, followed by five freeze thaw cycles to ensure the breakage of the liposomes, and thus the exchange of the buffer between the inside and the outside of the vesicles. To obtain evenly sized unilamellar vesicles, the samples were extruded using a mini-extruder set (Avanti) and 1 ml Hamilton syringes 11 times through a 0.4 µm filter, and again pelleted at 50, 000 rpm at 15 °C for 25 min. The supernatant was removed and the pellet resuspended into the required volume of inside buffer (2 µl/experimental time point for uptake of radiolabelled substrate, 4 µl/reaction for pyranine assays).

2.4.1 Time-dependent uptake of radio-labelled substrate

Internal and external buffer composition for $\Delta\Psi$-driven experiments is given below (Table 2-5). To monitor the transport of the substrate into the liposome and to investigate under what conditions uptake is favoured, radiolabelled ^{3}H-alanine (60 Ci/mmol) was used for this purpose. Unless stated otherwise, a mixture of 100 µM alanine and 250 nM ^{3}H-alanine was used for the transport assays. The outside buffer was preheated to 30 °C, and the reaction was started upon addition of the liposomes into the buffer. As a control, empty liposomes (containing no protein) were used under the same conditions.

Table 2-5: Buffer conditions for the ΔΨ-driven transport assays

internal buffer I	**external buffer I**
20 mM KPi pH 6.5	120 mM NaPi pH 6.5
100 mM KAc	2 mM $MgSO_4$
2 mM $MgSO_4$	10 µM valinomycin
	250 nM ^{3}H-alanine

For each time point, 50 µl of the reaction mixture was immediately diluted into 2 ml of 0.1 M LiCl and filtered over a 0.22 µM nitrocellulose filter under vacuum, and the filter washed twice with 2 ml ice cold 0.1 M LiCl. The liposomes stick to the filter via hydrophobic interactions and any residual radiolabelled substrate is washed off. The filter was removed and added to a vial containing 7.5 ml Ultima Gold^{TM} (Perkin Elmer) scintillation cocktail.

The remaining radioactivity level on the filters was determined with a scintillation counter. The signal was averaged from 300 measurements and resulted in a final signal in counts per minute (cpm). Based on a standard curve for the radiolabelled amino acid, this signal can then be converted into molar amounts. The final amount of alanine or arginine taken up into the liposomes was additionally corrected for the unlabelled substrate in the external buffer.

To investigate if the transport of alanine was electrogenic, assays were performed with proteoliposomes loaded with a potassium buffer and diluted into sodium buffer of similar ion concentration. Upon addition of the K^+-ionophore valinomycin, the potassium ions can permeate the membrane thus creating a negative inside membrane potential ($\Delta\Psi$) of -82.5 mV calculated with the Nernst equation. This membrane potential can then be used as a driving force for substrate uptake if the substrate transport is electrogenic and in this study, whether substrate uptake is coupled to a positively charged ion (Figure 2-4). As controls, assays were repeated with the internal buffer I on the outside, without any addition of valinomycin to the external buffer I and with addition of carbonyl cyanide m-chlorophenyl hydrazone (CCCP) to the external buffer I. CCCP is a proton ionophore which depletes any membrane potential through proton influx.

For competition assays, electrogenic uptake of ^{3}H-alanine using 5 µg protein was measured in the presence of 10 mM either of unlabelled alanine, D-alanine, glutamate, proline, arginine, lysine, ornithine or γ-aminobutyric acid (GABA). A single measurement was taken after 1 min.

To determine the IC_{50} values, proteoliposomes were assayed similarly but in ranging concentrations of unlabelled competitor. IC_{50} curves were fitted with a sigmoidal dose response method from PRISM GRAPH PAD (GraphPad Software, La Jolla California USA, www.graphpad.com)

Kinetic constants for WT *Gk*ApcT and the Met321Ser mutant were determined by time-course measurements of $\Delta\Psi$-driven uptake of ^{3}H-ala, and ^{3}H-arginine. The external buffer

contained a mixture of 250 nM ^{3}H-alanine or ^{3}H-arginine, and varying concentrations of unlabelled substrate. The linear part of the time-course measurement gave the initial transport rates for each concentration of substrate. Plotting the initial rate against the substrate concentration and fitting of the Michaelis-Menten equation, in PRISM, to the data to give the apparent K_M and v_{max} values for transport.

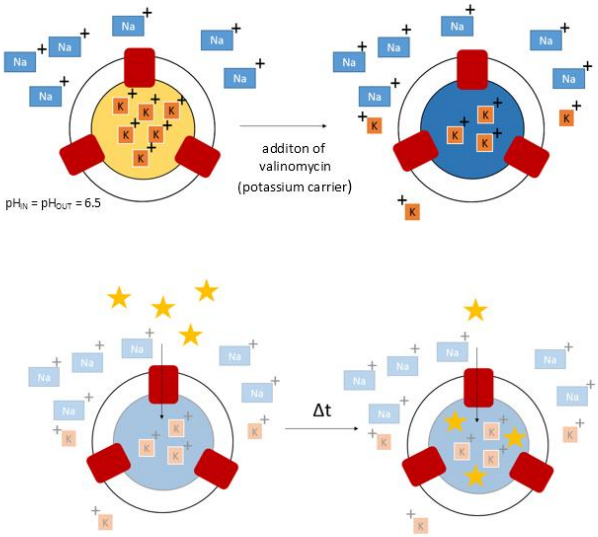


Figure 2-4: Schematic of electrogenic experiments

Proteoliposomes are loaded with potassium buffer (K^+) at pH 6.5 and diluted into buffer containing the same amount of sodium ions (Na^+). Upon addition of valinomycin a negative inward potential is created. Uptake of the radiolabelled substrate (yellow star) by the transporter (red) is monitored over time (Δt) if the negative potential is used for coupled substrate transport to a positively charged ion.

2.4.2 Counterflow experiments

To determine which amino acids are recognised by the membrane transporter under study, counterflow assays were performed. No pH or salt gradient was applied which means that the transport should only be driven by the substrate concentration. The liposomes were prepared with an excess of a specific amino acid in the internal buffer. The unlabelled substrate inside the liposome is transported if it is recognised by the transporter, resulting in an exchange of ^{3}H-alanine with the labelled amino acid in the external buffer system (Figure 2-5). Thus, the labelled amino acid accumulates in the lumen and the concentration can be determined. This kind of transport is independent from the cotransport of the coupled ion, as the transport mechanism is not dependent on any other external gradient than that of the substrate concentration on the inside and the outside if the proteoliposomes. To

achieve the transport, the transporter must undergo a full cycle of conformational changes which are driven by specific amino acids.

For the investigated transporter, the optimal parameters for counterflow were determined. For this, counterflow experiments were performed with an internal excess of a chosen compound at 2 mM (buffer conditions are given in Table 2-6). The liposomes for each experiment were prepared as described above.

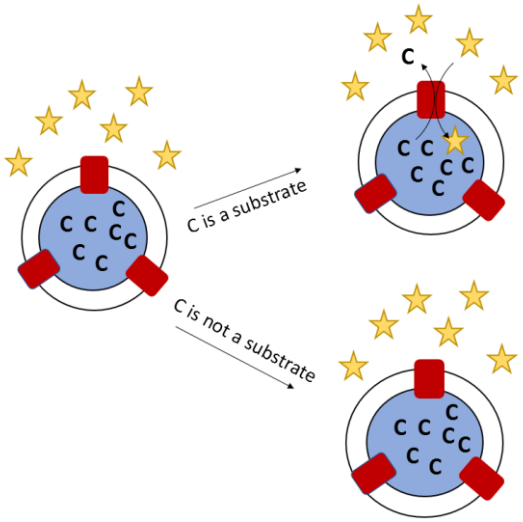


Figure 2-5: Schematic of counterflow exeperiments

Proteoliposomes are loaded with an excess of a chosen compound (C). As there are no other gradients applied (pH gradient, salt gradient) the transport is only substrate driven. If the amino acid is recognised by ApcT, in order to complete the transport cycle, alanine (as favoured amino acid) is bound and transported into the liposome. Radiolabelled alanine (yellow star) is used as a reporter amino acid to determine the amounts of alanine transported.

The proteoliposomes were loaded with internal buffer II and were diluted into external buffer II. The internal and external buffer had the same pH. To determine the pH under which counterflow was optimal, a range from pH 4 to pH 8 was screened for alanine uptake. Again, 5 µg of proteoliposomes were used for each experimental time point. The reaction was started upon addition of the liposomes to the external buffer, and was stopped by taking 50 µl samples with immediate filtration under vacuum followed by two washing steps with 2 ml ice cold 0.1 M LiCl. The radioactivity level was determined as already described before and corrected for the alanine concentration that was transported into the liposomes.

Table 2-6: Buffer conditions for counterflow experiments

internal buffer II	**external buffer II**
25 mM citrate phosphate buffer pH 6.0	25 mM citrate phosphate buffer pH 6.0
100 mM NaCl	100 mM NaCl
2 mM $MgSO_4$	2 mM $MgSO_4$
2 mM compound	250 nM ^{3}H-alanine
	100 µM alanine

2.4.3 Pyranine assays to monitor proton flux

To monitor the influx of protons into liposomes, the vesicles were loaded with pyranine, a fluorescent non-permeable dye. Pyranine (trisodium 8-hydroxypyrene-1,3,6-trisulfonate) changes its excitation maximum depending on its protonation state (Wolfbeis, 1983).
Proteoliposomes were prepared as usual. After the first harvesting step, the pellet was re-suspended into internal buffer III (Table 2-7) additionally containing 1 mM pyranine, followed by seven freeze-thaw cycles, extrusion through a 0.4 µm filter (11 times) and another harvesting step at 15 °C. The supernatant was discarded and the pellet washed with 100 µl inside buffer and resuspended in 80 µl internal buffer. A G-25 column was equilibrated with 100 µl inside buffer, and the resuspended pellet was transferred to the column and centrifuged at 800 g for 2 min to remove any residual pyranine. The recovered liposomes were harvested (50,000 rpm, 15 °C, 30 min.) and the pellet resuspended at 4 µl/ experiment in internal buffer III (Table 2-7). Liposomes were diluted into 300 µl external buffer III (Table 2-7) in a 0.85 ml cuvette containing a small magnetic flea under continual stirring. The fluorescent change of pyranine was measured with a Cary eclipse fluorimeter at $\lambda_{excitation}$=415 and 460 nm; $\lambda_{emission}$=510 nm. The reaction was initiated with addition of 300 µl external buffer III (Table 2-7) containing substrate and was measured over a period of 20 min at room temperature.
The data were transferred into Prism Graph Pad and the intensity values at 460 nm divided by the intensity values at 415 nm excitation wavelength to obtain $F_{R\,(460\,nm/415\,nm)}$. The values were normalised so that the starting signal had a value of one in order to compare individual experiments with each other.

Table 2-7: buffer conditions for pyranine assays

internal buffer III	external buffer III
5 mM HEPES-NaOH pH 6.8	5 mM HEPES-NaOH pH 6.8
120 mM KCl	120 mM NaCl
2 mM $MgSO_4$	2 mM $MgSO_4$

2.5 HOMOLOGY MODELLING

Homology models of *Gk*ApcT, *At*CAT2 and *h*PAT1 were generated using a multiple sequence alignments between the target and template protein (Figure A 1 and Figure A 2) generated with PROMALS3D with default parameters (Pei and Grishin, 2014) and used as an input for MODELLER (Webb and Sali, 2014).

For the repeat swap model of *Gk*ApcT, the modelling protocol reported by Vergara-Jaque *et al.*, 2015 was followed (Figure 2-6). First, the crystal structure of *Gk*ApcT was analysed in PYMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC) and divided into its symmetrical halves, named repeat unit 1 and repeat unit 2 (Figure 2-6, a). Both halves were then extracted as new objects in PYMOL and superimposed onto one another: first repeat unit 1 was superimposed with the *super* command in PYMOL, performing a sequence-independent structure based alignment, onto repeat unit 2, and then repeat unit 2 was separately superimposed onto repeat unit 1 to generate the alternate conformation of the protein (Figure 2-6, b). This was stored as a coordinate file and used as the template model (Figure 2-6, d). A sequence alignment with ALIGNME (Stamm *et al.*, 2014) (Figure A 3) was performed and the repeat units' sequences were then interchanged and merged to generate a swapped template model sequence (Figure 2-6, c). Any N-terminal or C-terminal overhangs that were not part of the symmetrical region were kept in their original location. This alignment was then used as a MODELLER (Webb and Sali, 2014) input to generate the initial repeat swap model (Figure 2-6, e). This model was then inspected and the alignment adjusted if any structural elements were not modelled correctly.

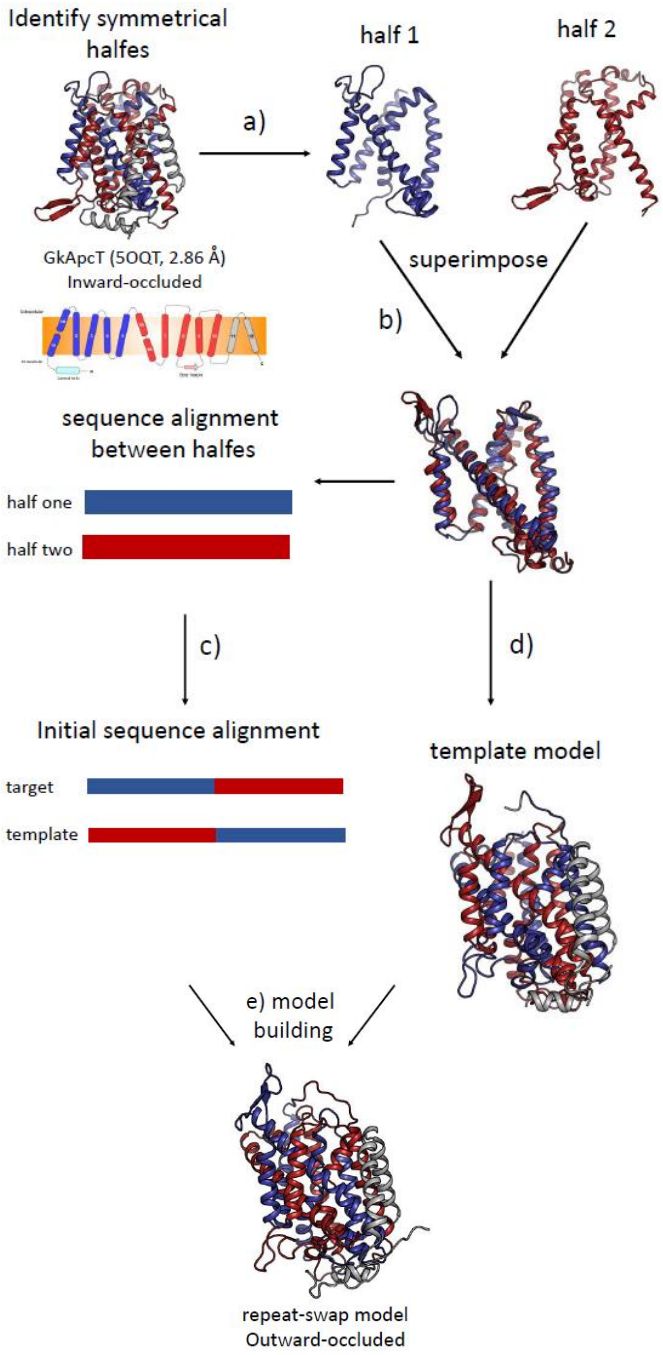


Figure 2-6: Schematic protocol for repeat-swap modelling

2.6 THERMAL STABILITY MEASUREMENTS

To evaluate the thermal stability, CD spectroscopy was performed using a Jasco J-815 Spectropolarimeter (JASCO, Japan). Concentrated protein was diluted 1:100 into CD buffer containing 10mM KPi pH 6.5 and 50 mM $NaSO_4$, supplemented with either 0.03 % DDM, unless specified otherwise, to reach a final protein concentration of 0.1 mg/ml. Absorption spectra were collected in a wavelength range from 200 to 240 nm and temperature range of 20°C-86°C. Melting curves were obtained by plotting the absorption value at λ=220 nm against the temperature. Graphs were plotted in PRISM (GraphPad) and analysed using a sigmoidal fitting curve.

Wildtype *Gk*ApcT and mutants Glu115Gln, Lys191Ala, Lys191Asn, and Lys191Arg were further investigated on their thermal stability in the presence and absence of 10 mM alanine using the Prometheus NT.48 (NanoTemper Technologies). The proteins were diluted to a final concentration of 0.5 mg/ml into buffer containing 25 mM citrate phosphate, pH 6.0, 100 mM NaCl, 2 mM $MgSO_4$ and 0.03% (w/V) DDM. Thermal measurements were carried out in a range from 20°C-95°C with 1°C/min steps. The resulting melting curves were generated with the internal software by plotting the $F_{330\,nm}/F_{350\,nm}$ signal against the temperature. The final graphs were plotted in PRISM (GraphPad) using the first derivative of the initial plots against the temperature.

2.7 CRYSTALLISATION

2.7.1 Vapour diffusion crystallisation

Crystallisation trials were carried out for *Mj*ApcT, *Gk*ApcT, StApcT and SoApcT using MemGold (Molecular Dimensions Ltd.) and MemGold2 (Molecular Dimensions Ltd.) crystallisation screens from Molecular Dimensions. Vapour diffusion crystallisation was set up in sitting drop 96-well plates (SwissCi). 70 µl of the reservoir solution was manually dispensed into the wells of the plates and 100 nl of reservoir solution and 100 µl of protein solution (10 mg/ml) were mixed by a Mosquito liquid handling robot. The plates were stored at either 4 °C or 20 °C and inspected using a Rigaku imaging system. Initial hits were optimised

using a sparse matrix optimisation screen with varying PEG, detergent and salt concentrations based on the initial crystallisation conditions. Additionally, hanging drop experiments were performed in a 24 well hanging drop plate with 500 µl reservoir solution and a drop by mixing of 1 µl of protein and 1 µl of the reservoir solution.
Crystals were harvested into cryo-loops (MiTeGen) and for cryoprotection were transferred into the reservoir solution supplemented with detergent and an increased concentration of PEG of 3 % more than in the crystallisation condition. The crystals were then cryo-cooled and stored in liquid nitrogen until screened at Diamond Light Source MX beamlines.

2.7.2 *In meso* crystallisation

The detergent-solubilised protein was mixed with either molten 1-(*cis*-9-Octadecenoyl)-rac-glycerol (monoolein, Sigma Aldrich), 1-(7Ztetradecenoyl)-rac-glycerol (7.7MAG, Avanti Polar Lipids Inc.) or 1-(7Z-pentadecenoyl)-rac-glycerol (7.8MAG, Avanti Polar Lipids Inc.) in a volume ratio of 60:40 (lipid:protein) for the monoolein and monoolein/cholesterol mixtures, and 50:50 for 7.7MAG and 7.8MAG. The lipid and protein were kept separately in 100 µl gas-tight syringes (Art Robbins Instruments). For mixing, the syringes were connected with a Teflon connector (Art Robbins Instruments) and both solutions mixed gently until a fully transparent and homogenous LCP solution was formed (Figure 2-7). As crystallisation screens, MemMeso (Molecular Dimensions Ltd.), diluted MemGold1 and MemGold2 (Molecular Dimensions Ltd.), and custom crystallisation screens were used. For further optimisation, the Hampton salt stock (Hampton Research) and MemAdvantage (Molecular Dimensions Ltd.) screens were employed. Additionally, screens were designed individually for optimisation of the initial hits found for each protein.

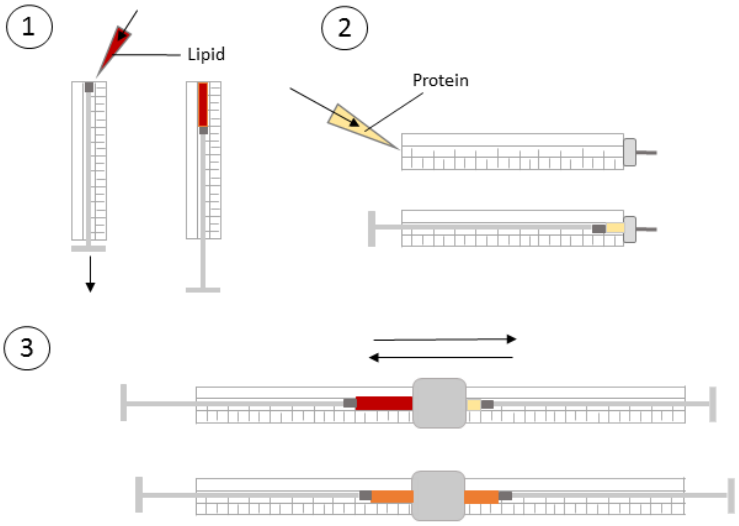


Figure 2-7: Preparation of LCP

(1) Molten lipid (red) is transferred with a pipette into a syringe. (2) The appropriate amount of detergent solubilised protein (yellow) is transferred into a second syringe from the back. The plunger is pushed into the syringe, moving up the protein solution to the end of the syringe, ensuring no bubbles appear. (3) The two syringes are coupled together and the contents mixed through alternate movement of the plungers until a transparent LCP (orange) has formed.

2.7.2.1 Mixing of lipids for optimised *in meso* crystallisation

It has previously been shown that the addition of cholesterol to the lipids used for LCP can increase the chances of obtaining crystals and increase crystal size (Cherezov *et al.*, 2007). 1g of monoolein (Avanti Polar Lipids Inc.) was supplemented with 100 mg of cholesterol (Sigma Aldrich) in a glass vial. To dissolve both lipids, 2 ml of chloroform was added. The solvent was removed by using a nitrogen stream, and the vial was heated in a water bath to prevent the lipids from freezing. To remove any remaining chloroform, the vials were covered with a perforated cap and kept under vacuum overnight (Virtis benchtop K lyophiliser). The dried lipid mixture was melted at 42 °C, divided into 50 µl aliquots and stored at -80 °C.

2.7.2.2 Crystallisation in *in meso* sandwich plates

50 nl of the protein-laden LCP were dispensed onto a siliconized microscope slide (85.5 mm x 127.8 mm x 1.0 mm, Thermo Scientific) with a sticky grid (96 grid points, 77 mm x 112 mm x 5 mm, Saunders) followed by the dispensation of 800 nl of precipitant, and sealing with a cover glass (77 mm x 112 mm, Thermo Scientific) (Figure 2-8). An LCP-handling robot (Gryphon LCP, Art Robbins Instruments) was used to prepare the LCP sandwich plates. The LCP plates were stored at either a constant 20 °C or 4 °C.

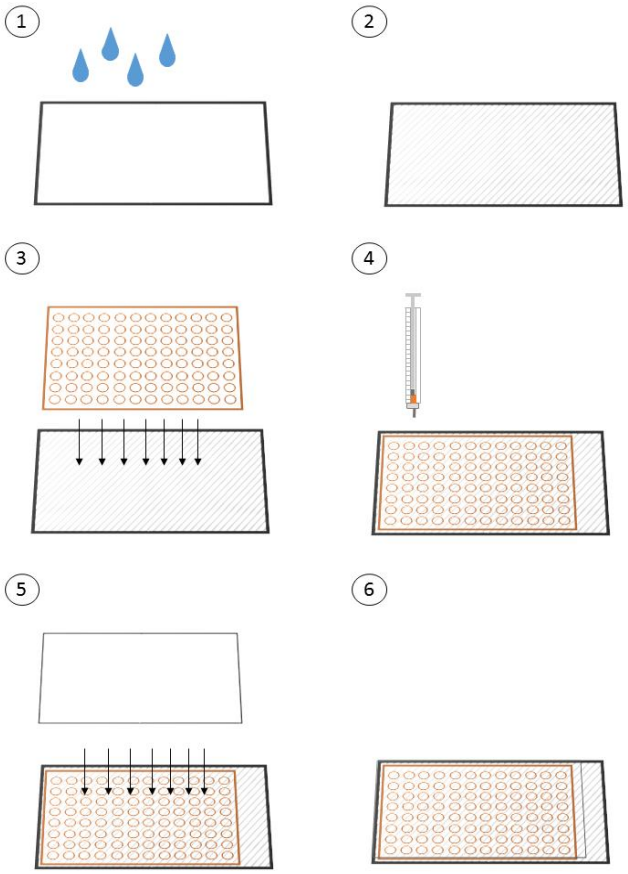


Figure 2-8: Preparation of LCP sandwich plates for in meso crystallisation
(1) The glass base plate is hydrophobically covered with a silanising reagent (RainX, Kraco Car Care International Ltd.). (2) Residual RainX is removed with a tissue until the glass surface is completely clean. (3) The cover slip from one side of the sticky gasket is removed and with the sticky side down, applied onto the coated glass base plate. (4) The second cover slip of the sticky gasket is removed. The plate is fixed into the LCP robot. One 50 nl drop of LCP is printed into each well, followed by the addition of 800 nl of precipitant solution on top. (5) To seal the crystallisation wells, a clean cover slip is applied to the plate. (6) The LCP sandwich plate is now ready to be stored to allow crystal growth.

2.7.2.3 Retrieving crystals from LCP sandwich plates

To extract the crystals from LCP sandwich plates, the cover glass around the drop containing the desired crystals was cut around the edges with a diamond cutter. The cover glass was then removed gently with tweezers which sometimes resulted in the breaking of the cover glass over the LCP drop. Once the LCP drop was uncovered in this way, a drop of the corresponding precipitant condition was applied to prevent the drop from dehydration. The original precipitant solution around the LCP bolus tends to move to the side once the cover glass is lifted, and LCP grown crystals can disappear quickly without the application of more precipitant solution.

2.7.2.4 Preparation of samples for continuous LCP injector

LCP was prepared as described above. A lipid mixture of monoolein + 10 % cholesterol was used for *Gk*ApcT, and 7.8MAG for $PepT_{St}$ and the $PepT_{St}$-SeMet derivative. 10 µl of LCP and

80 µl precipitant were used for each experimental setup. First, enough LCP was prepared for the required number of experiments. Thorough cleaning, and working on tin foil reduced the number of dust particles in the syringes and LCP. The dust, if present, would induce clogging of the nozzle during injection into the beam during data collection.

Once enough LCP was prepared, it was evenly distributed between 100 µl Hamilton syringes. Each syringe had the coupler applied so that the void in the coupler was fully filled with LCP (male syringe). Thus, bubbles were prevented from entering the precipitant and LCP. 80 µl of precipitant solution were aspirated into the 'female' syringe and connected to its male counterpart (Figure 2-9). By holding the syringe pair horizontally and slowly pushing the plunger of the 'male' syringe, LCP was injected into the precipitant (Figure 2-9). When all the LCP had been injected, both syringes stayed connected and parafilm was applied around the coupler area to prevent the mixture from drying out (Figure 2-9). The syringes were then stored at 20 °C.

Before the injection of the LCP which now contained crystals, the syringes were disconnected so that the coupler stayed on the syringe with the LCP and precipitant. The precipitant was removed by slowly pushing down the plunger of that syringe until almost all of the precipitant was extruded. To incorporate residual precipitant and to ensure an even distribution of crystals for the experiment, 2 µl of lipid was mixed with the residual LCP.

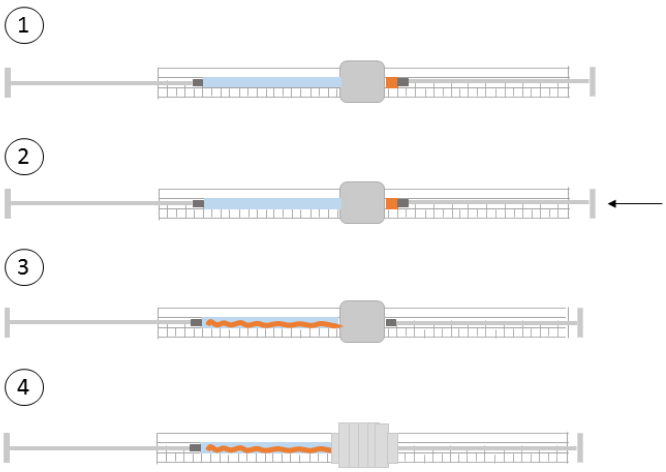


Figure 2-9: Preparation of LCP for crystallisation in a syringe
(1) A syringe containing the precipitant solution (blue) and a syringe with prepared LCP containing protein (orange) are coupled together. (2) The plunger of the LCP containing syringe is slowly pushed towards the other syringe. (3) The extruded LCP forms a curved tube surrounded by the precipitant. (4) The syringes stay connected to each other and the coupler region is sealed to keep the syringes air tight.

2.7.2.5 Continuous LCP injection - experimental setup

For the serial crystallography (SX) experiment at synchrotron beamline P11 at PETRA III, Hamburg, LCP nozzles and an experimental setup similar to that described by Nogly *et al.* (Nogly *et al.*, 2015) were used. The injector consisted of the nozzle body, an LCP reservoir, a hydraulic stage body, a hydraulic stage plunger and a gas line feedthrough. The LCP nozzles consisted of an inner and outer capillary. The inner capillary was conical shaped and extruded the LCP sample. The outer capillary had a melted aperture and was connected to an external helium gas line that surrounded the LCP for stabilisation.

The LCP injector used in this experiment was supplied by Henry Chapman's group at DESY CFEL, and was purchased from Arizona State University (Weierstall, 2014). The injector was positioned vertically downwards at an angle of 90° to the X-ray beam. An on-line microscope and a remotely controllable HPLC system and gas line were also installed. The 4 µm x 9µm full width at half maximum (FWHM) beam was aligned with the LCP injector with the aid of a fixed yttrium aluminium garnet (YAG) screen. The nozzle was then moved into place after the X-ray beam has been aligned on the nozzle position was focused. A falcon tube under the LCP nozzle was installed to catch any residual LCP. In case the experiment had to be interrupted leaving the sample exposed to the air in the experimental hutch, a humidifier was used to prevent the sample from drying out. Diffraction images were collected at ambient pressure and temperature using a Pilatus 6 M detector operating at 25 Hz. During data collection, no shutter was used and the sample was exposed for 20 ms at 1.3 x 10^{13} photons/s. In the second attempt to collect SX data on crystals of *Gk*ApcT and $PepT_{St}$ a rotating shutter was used and sample exposure was reduced to 8ms with a similar incident flux as before.

2.8 HEAVY ATOM SOAKING

2.8.1 Heavy atom pre-screening

Pre-screening with AuCN, $HgCl_2$, $PtCl_2$, HgAc, and CH_3HgCl using purified *Gk*ApcT was performed to identify heavy metals that might destabilise the protein and would thus not be used during derivatisation. 0.1 mg/ml of *Gk*ApcT was mixed with 3x this concentration of

heavy metal compound and left on ice for 2 h. To remove any precipitant, the solution was centrifuged for 30 min at 14 000 g at 4 °C. The supernatant was loaded onto a Superdex^{TM} 200 10/150 column.

2.8.2 Heavy atom soaking of crystals in LCP

The precipitant condition of the well containing crystals to be soaked was mixed with the heavy metal of choice to a final concentration of 100 µM or 1 mM. The LCP plate well was opened as usual and 1 µl of the heavy metal – precipitant mixture was added onto the LCP drop. One well of the sticky grid was cut out and stuck onto the open well. The adhesive cover was removed and the well closed with a glass cover slip. The crystals were left to soak for 1 h, 2 h and overnight, and harvested as usual for data collection.

2.8.3 Heavy atom pre-derivatisation

Purified protein was mixed with 3 x heavy metal to protein concentration and left on ice for 2h. To remove any precipitated protein, the solution was spun down at 14 000 g for 30 min at 4 °C. The supernatant was loaded onto a Superdex^{TM} 200 10/300 column and fractions containing protein were united and concentrated to the desired value for crystallisation in LCP.

2.9 DATA PROCESSING AND PHASING

Spot finding, indexing and integration of the collected diffraction images were performed with DIALS (Sauter *et al.*, 2013), followed by scaling and merging with AIMLESS (Evans and Murshudov, 2013). Initial phases were obtained by molecular replacement (MR) with PHASER (McCoy *et al.*, 2007) in the CCP4 suite. The template used for MR was generated using CHAINSAW (Stein, 2008) with the published *Mj*ApcT structure (PDB-ID: 3GIA, 2.32Å) and the protein sequence of *Gk*ApcT. The model was loaded into PYMOL and loops were removed manually.

2.10 SULPHUR SAD DATA COLLECTION

Due to difficulties placing the protein sequence into the electron density map generated by molecular replacement, anomalous data were collected at beamline I23, Diamond Light Source, UK and BL1A, Photon Factory, Japan. Inverse beam data collection was performed in both cases under vacuum at I23 and in a helium environment at BL1A, collecting at wavelengths of 2.75 Å and 2.7 Å, respectively.

The two data sets per inverse beam data collection were processed individually in XDS (Kabsch, 2010) and combined using XDS_SCALE (Kabsch, 2010), followed by MR. Initially, the MR model described above and at later stages, the refined model was used. To generate the difference Fourier map, the files were prepared with SHELXC (Sheldrick, 2008) and the maps calculated using ANODE (Thorn and Sheldrick, 2011).

2.11 MODEL BUILDING AND REFINEMENT

Model building into the electron density map was performed in COOT (Emsley *et al.*, 2010). The initial MR model, the maps from MR and the maps from the sulphur SAD data collection were used to guide the correct insertion of the protein sequence. Structure refinement was carried out in BUSTER (Bricogne *et al.*, 2016) to a final resolution of 2.86 Å with an R_{work}/R_{free} of 0.22/0.27. Geometry restraints for ligands were calculated with GRADE (http://grade.globalphasing.org).

3 Screening of homologues

3.1 Small and large-scale expression of ApcT homologues

To find suitable candidates for structural and biochemical studies of a potential proton-coupled APC transporter, homologues of *Mj*ApcT were identified from bacteria and archaea. A BLAST search (Altschul *et al.*, 1990) against the UniProt and non-redundant protein library was performed using the protein sequence of *Mj*ApcT as the template.

Table 3-1: Screened ApcT homologues

Protein name	***Organism***	***Optimal growth Temp.[1]***	***Optimal growth pH[1]***	
*Mj*ApcT	*Methanocaldococcus jannaschii*	48 °C-94 °C	5.2-7	Archaea
*Gk*ApcT	*Geobacillus kautophilus*	60 °C-74 °C	2-12	Bacteria, gram +
*Saci*ApcT	*Sulfolobus acidocaldarium*	55 °C-90 °C	1-9	Archaea
*St*ApcT	*Streptococcus thermophilus*	45 °C	6.5	Bacteria, gram +
*Stk*ApcT	*Sulfolobus tokodaii*	80 °C	2.5-3	Archaea
*So*ApcT	*Shewanella oneidensis*	22-35 °C	6-9	Bacteria, gram -
*Sso*ApcT	*Sulfolobus solfataricus*	80 °C	2-4	Archaea

[1] the optimal growth temperature and optimal growth pH of the organisms from which the proteins originate from

The selection of proteins identified as suitable from the BLAST search is given in Table 3-1. Additionally, information on the number of predicted transmembrane helices, and the optimal growth temperature and pH of the individual organisms, if this information was available in the literature, are given. All candidates are derived from thermophilic organisms that can survive in “harsh” environments with high temperatures and pHs. Proteins from thermophilic organisms tend to be more temperature stable, which is one fundamental property necessary for crystallisation (Deller *et al.*, 2016).

The protein sequences of the selected homologues were used to generate a phylogenic tree to elucidate the evolutionary relationship to members of the SLC7 and the APC family (Figure 3-1). *Gk*ApcT, *Sso*ApcT, *Stk*ApcT and *St*ApcT show a very close relationship to the CATs

of the SLC7 family. *So*ApcT is closely related to LeuT, *Saci*ApcT shares a close relationship with *Mj*ApcT and AdiC (Figure 3-1). Therefore, the pool of the selected *Mj*ApcT homologues might represent a range of proteins with different substrate specificities and coupling mechanisms.

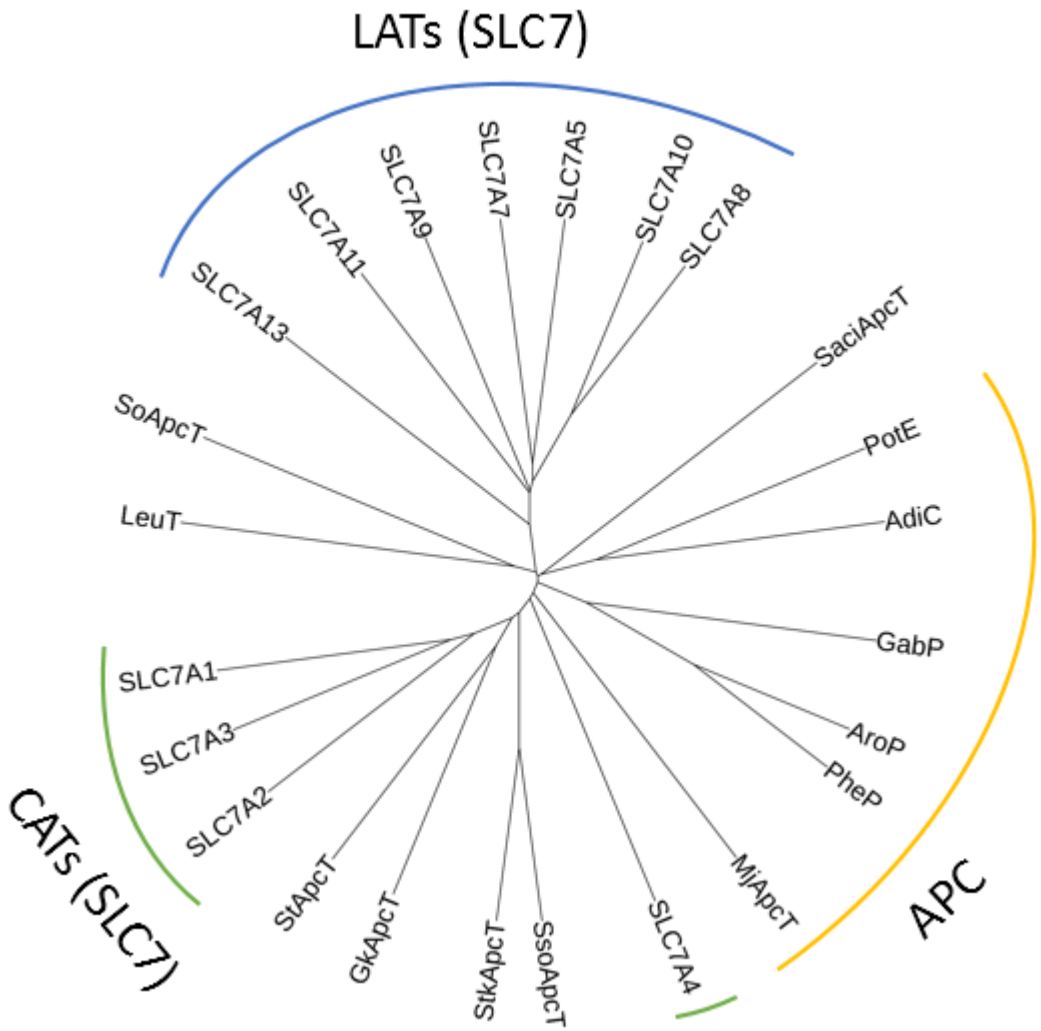


Figure 3-1: Phyolgenic tree of screen homologues and APC members

A phylogenic tree was generated using a multiple sequence alignment generated with PROMALS3D (Pei and Grishin, 2014). The screened homologues were *Saci*ApcT, *Gk*ApcT, *So*ApcT, *St*ApcT, *Stk*ApcT and *Sso*ApcT.

Genes encoding ApcT homologues of thermophilic bacteria and archaea were cloned into the pWaldo-GFP vector under the control of the T7 promoter (Drew *et al.*, 2001) through which the proteins are expressed with a C-terminal TEV-cleavage site, GFP-tag and an 8xHis-tag. The constructs were transformed and expressed in the C43(DE3) *E.coli* strain. The GFP-tag has several advantages; functional, and thus fluorescent, GFP is only available when the

fused membrane protein is properly folded, expression levels can be determined by GFP-fluorescence in the cell suspension and the tagged-protein can easily be followed throughout the purification process, though cleavage of the GFP can destabilise the protein of interest.

To identify potential candidates for large-scale expression and purification, a small-scale expression of each construct in 10 ml LB medium was carried out. Of the investigated proteins, five were found to be suitable: *Mj*ApcT, *Gk*ApcT, *So*ApcT, *Saci*ApcT and *St*ApcT (Figure 3-2).

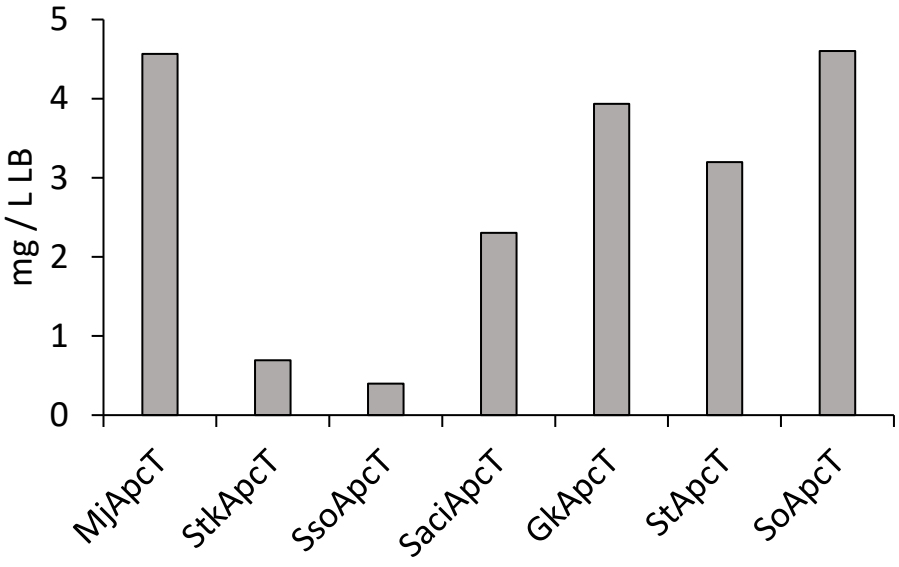
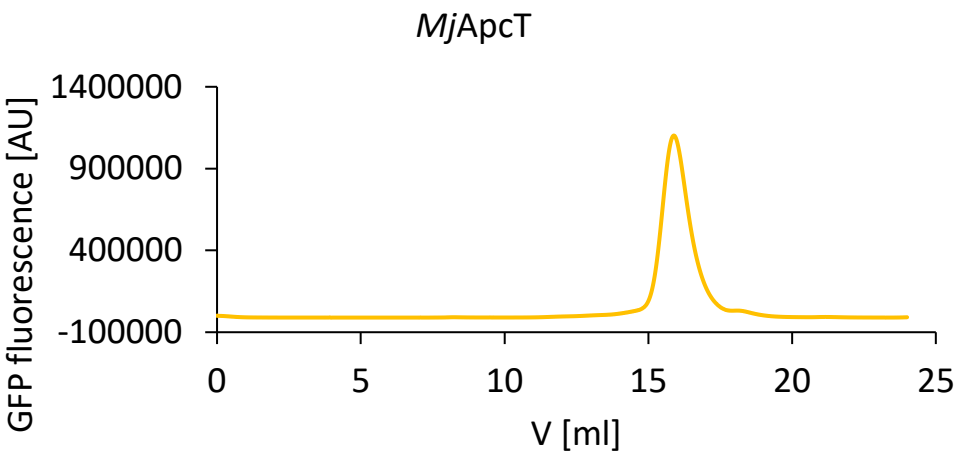
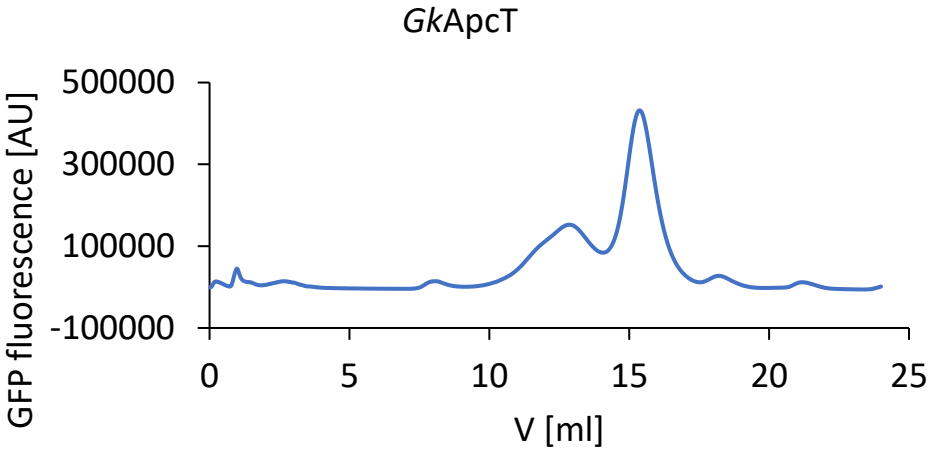
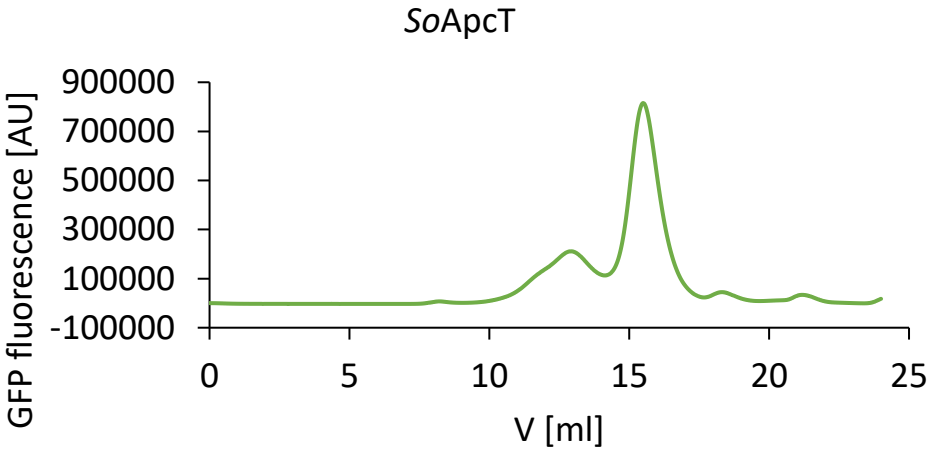


Figure 3-2: Expression levels of ApcT homologues in LB medium

*Sso*ApcT and *Stk*ApcT showed low expression levels in LB medium (0.7 mg/l and 0.4 mg/l respectively). Repetition of small scale expression for both constructs was performed in LB, TB and PASM5052 media but none of them increased the expression level. To determine the stability of the expressed proteins in detergent, FSECS were performed. DDM was used as a first choice for FSEC as is has proven successful for solubilisation for several membrane proteins and it is a relatively mild detergent. For FSECs, cells were grown in 10 ml TB cultures, and overnight protein expression was started over induction with IPTG, followed by a small-scale membrane preparation and solubilisation of the membranes in buffer containing 1 % DDM.


*Mj*ApcT
GFP fluorescence [AU]
1400000
900000
400000
-100000
0
5
10
15
20
25
V [ml]


*Gk*ApcT
GFP fluorescence [AU]
500000
300000
100000
-100000
0
5
10
15
20
25
V [ml]


*So*ApcT
GFP fluorescence [AU]
900000
700000
500000
300000
100000
-100000
0
5
10
15
20
25
V [ml]

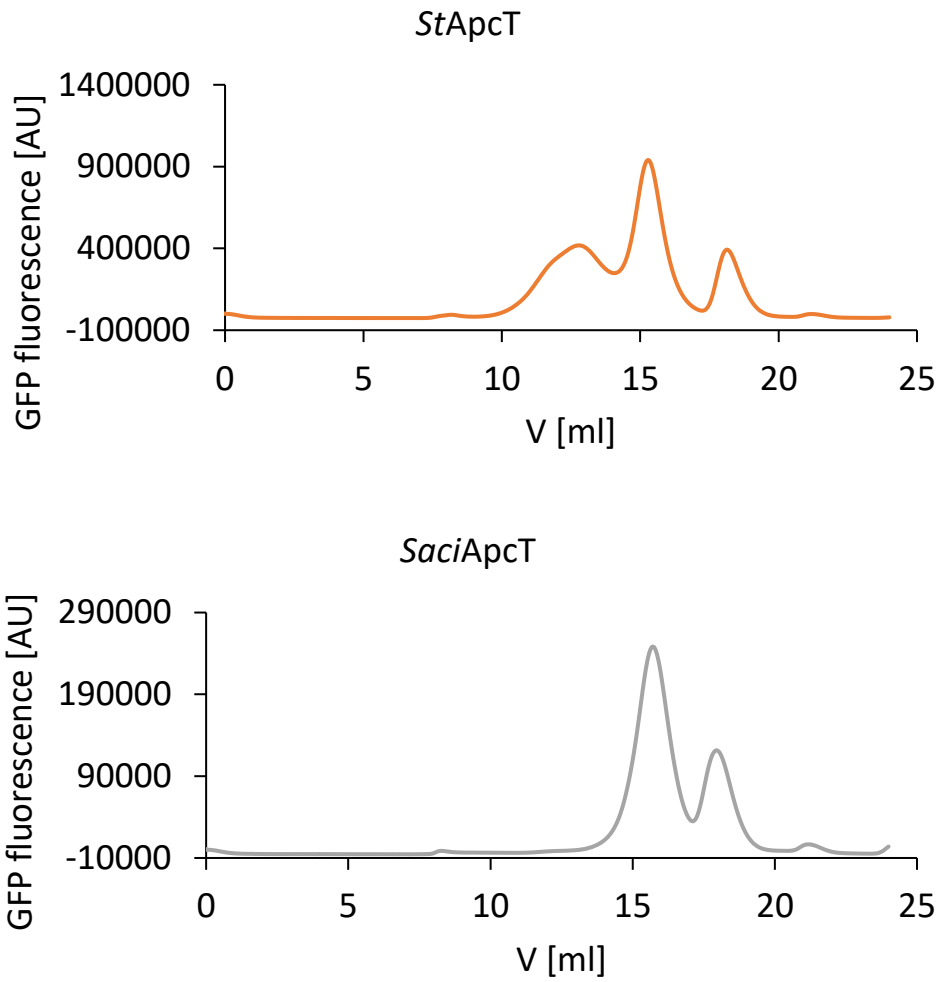


Figure 3-3: FSEC profiles of MjApcT, GkApcT, StApcT, SoApcT and SaciApcT in DDM

Gel filtration profiles for *Mj*ApcT, *Gk*ApcT, *St*ApcT, *So*ApcT and *Saci*ApcT after solubilisation in 1 % DDM. The supernatant of the solubilisation reaction was loaded onto a S200 10/300 column and the fluorescence of GFP in arbitrary units (AU) was monitored over the elution volume (V) in ml.

The soluble portion was separated from the non-solubilised protein by centrifugation. Size exclusion chromatography was performed by loading the supernatant onto a size exclusion column and following the fluorescence of GFP over time (Figure 3-3). The FSEC trace of the *Mj*ApcT-GFP fusion shows a monodisperse peak indicating a stable protein and no aggregation in DDM (Figure 3-3). The solubilisation efficiency was 100 % in DDM for all the constructs tested here.

For *Gk*ApcT-GFP, a peak at the elution volume of 12 ml is observed which indicates a possible dimer formation of the protein. A small peak observed at an elution volume of ~18 ml

corresponds to free GFP, whereas the peak at 16 ml corresponds to the GkApcT-GFP fusion protein (Figure 3-3). A similar profile is observed for *So*ApcT, *Saci*ApcT and *St*ApcT, although for the latter the peak of free GFP is much higher which might indicate proteolysis of the GFP off the fusion protein. However, the FSEC analysis identified DDM is a suitable candidate for the purification of the investigated proteins. It was noted that most of the protein was lost during the extraction from the membrane of the *Saci*ApcT sample, identified by determining the GFP fluorescence before and after cell lysis (approx. 80% loss of *Saci*ApcT-GFP after cell disruption).

Large-scale expressions of the selected constructs (*Mj*ApcT, *Saci*ApcT, *Gk*ApcT, *So*ApcT, and *St*ApcT) were carried out in two litres of TB media. The cells were harvested and the expression level determined by GFP-fluorescence measurements ($\lambda_{excitation}$=488 nm, $\lambda_{emission}$=512 nm). All investigated constructs showed similar levels of expression with a wet pellet weight of around 25-30 g per litre of TB.

3.2 Purification of ApcT homologues

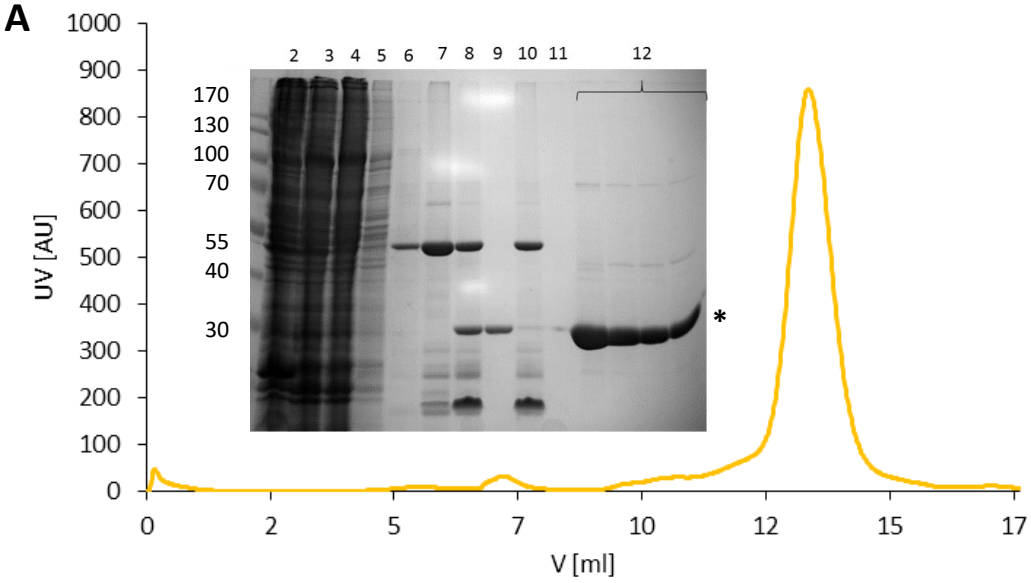
Based on the results from the pre-screening of the expression of each construct, *Mj*ApcT, *Saci*ApcT, *Gk*ApcT, *So*ApcT and *St*ApcT were further purified. Briefly, the purification involved the solubilisation in DDM, removal of non-solubilised proteins via ultra-centrifugation, purification over a Ni-NTA column, removal of the GFP/His-tag through TEV-cleavage overnight, purification of the cleavage reaction over a His-trap column, and size exclusion chromatography, here also referred to as gel filtration (Drew *et al.*, 2008).

Table 3-2: Purification specifications from ApcT homologues

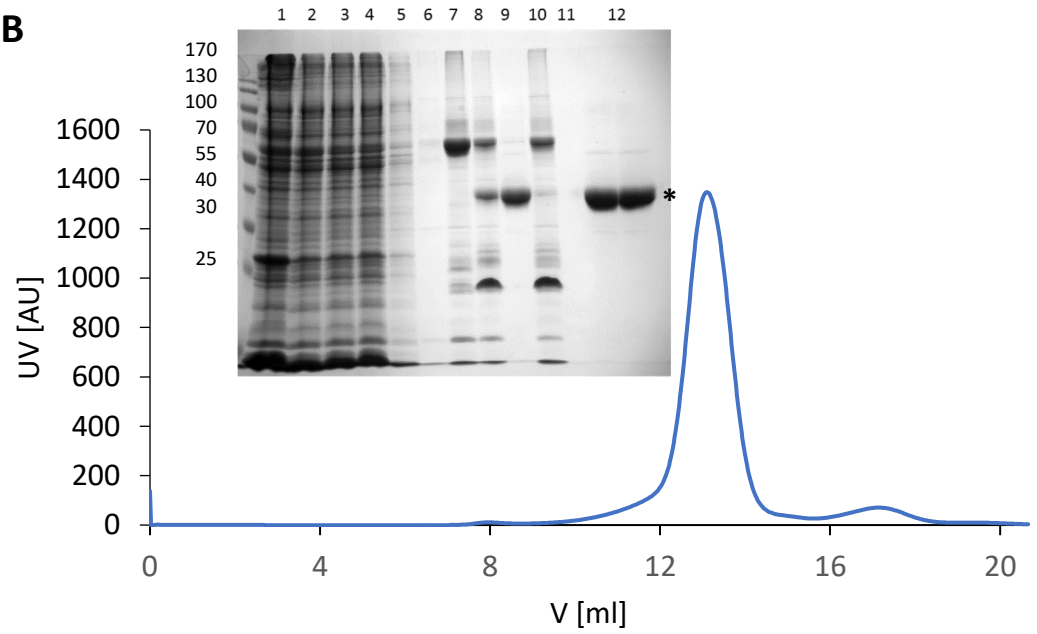
Protein	***Protein size [kDa]***	***Solubilisation efficiency [%]***	***Binding efficiency to resin [%]***	***Yield of purified protein from 1L TB [mg]***
***Mj*ApcT**	47.3	*90-95*	*60-70*	*1*
***Gk*ApcT**	50.8	*90-100*	*60-70*	*1-2*
***St*ApcT**	45.1	*80-90*	*50-60*	*0.5-1*
***So*ApcT**	49.8	*90-100*	*70-80*	*0.3-0.6*

The proteins showed similar and reproducible solubilisation efficiency in DDM ranging from 80 % to 100 %. Similar observations could be made for their binding efficiency to the Ni-NTA resin (50-80%, Table 3-2). Although the expression level of *Saci*ApcT was high (around 2.5 mg/l), it could not be purified successfully. As noted during the small-scale membrane extraction, most of the protein was lost during cell disruption which was also the case for the large-scale extraction from the membrane.

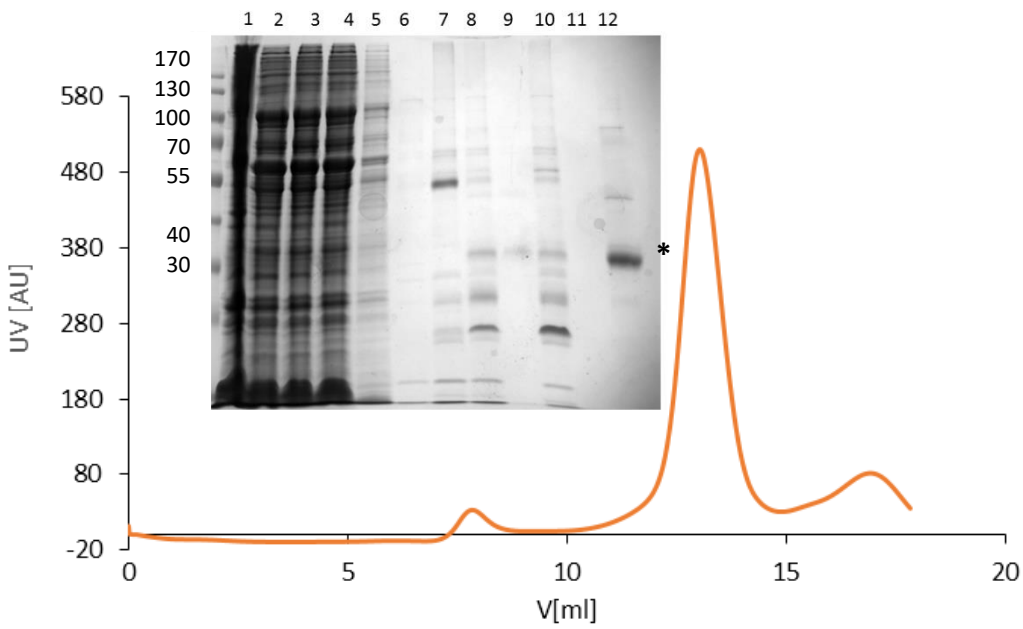
Analysis of the SDS-gels and SEC profiles (Figure 3-4) suggests homogeneous and pure protein samples for *Mj*ApcT, *Gk*ApcT, *So*ApcT and *St*ApcT. After TEV-cleavage, a shift of the band is observed which is in accordance with the loss of the GFP/His-tag from the target proteins (shift of around 28 kDa) as indicated by the gels shown in Figure 3-4. The cleaved protein is seen in the flow-through during purification over a Ni-NTA His-trap and thus it has been successfully separated from the GFP/His-tag and other non-specifically bound proteins that were co-eluted after the first affinity purification step. In the case of *Gk*ApcT and *St*ApcT (Figure 3-4B and D), the SDS-gel reveals minor contaminants that were reduced by altering the purification protocol, which includes the incubation of the protein solution and the resin in 20 mM imidazole, followed by a washing step with 40 mM imidazole. The purified proteins were concentrated to 10 mg/ml and further used for crystallisation trials and stability assays. It was not possible to purify *Saci*ApcT in high quantities, since 80% of the SaciApcT-GFP fusion were lost during cell disruption and preparation of the membranes and subsequent purification of the protein from the membranes revealed aggregation, as seen by the SEC profile and only small amounts of protein (Figure 3-4E).



B


1 2 3 4 5 6 7 8 9 10 11 12
170
130
100
70
55
40
30
25
*
1600
1400
1200
1000
800
600
400
200
0
UV [AU]
0
4
8
12
16
20
V [ml]

C


1 2 3 4 5 6 7 8 9 10 11 12
170
130
100
70
55
40
30
*
580
480
380
280
180
80
-20
UV [AU]
0
5
10
15
20
V[ml]

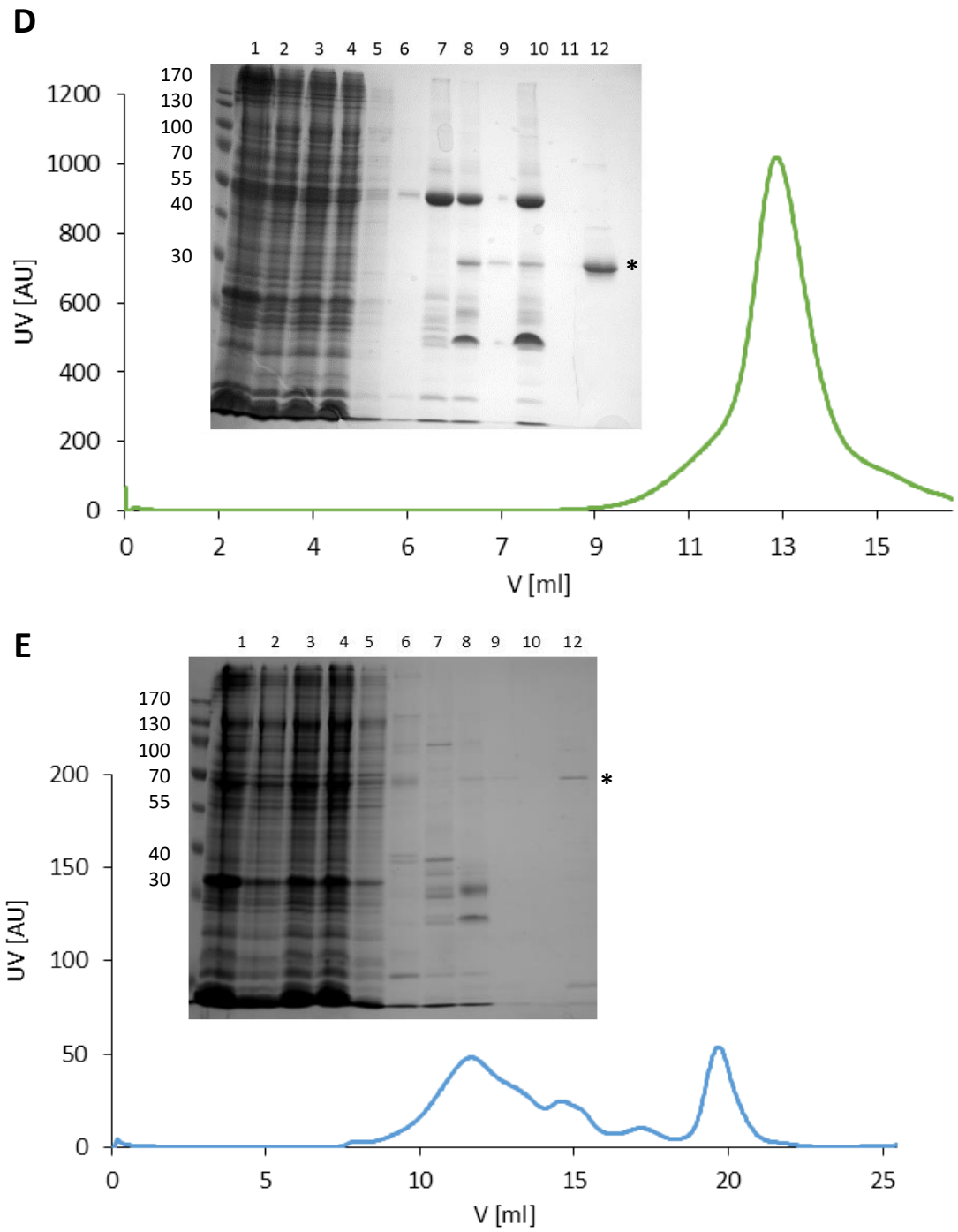


Figure 3-4: SDS gel analysis of purified ApcT homologues

12 % Tris-glycine gel and SEC profiles of the purification of *Mj*ApcT (A), *Gk*ApcT (B), *St*ApcT (C), *So*ApcT (D), *Saci*ApcT (E). During the purification process of the listed proteins, samples were taken and loaded on the 12 % Tris-glycine gel. The SEC profiles show the monitored absorption at 280 nm (UV) in arbitratry units (AU). 1: crude membrane sample; 2: solubilisation in DDM; 3: soluble fraction; 4: flow-through from Ni-NTA resin incubation; 5: flow-through of washing step with 15 mM imidazole; 6: flow-through of washing step with 30 mM imidazole; 7: flow-through of elution step with 250 mM imidazole; 8: overnight TEV cleavage and dialysis; 9: flow-through of reverse IMAC with His-trap column; 10: elution of His-trap with 250 mM imidazole; 11: flow-through of concentrator unit; 12: fraction of main peak from gel filtration run. The target protein is indicated by an asterisk.

3.3 THERMAL STABILITY OF SELECTED APCT HOMOLOGUES AND DETERGENT SCREENING

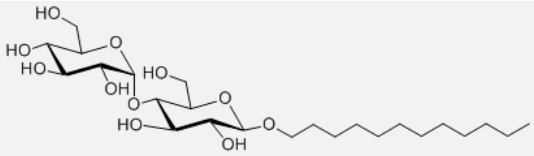
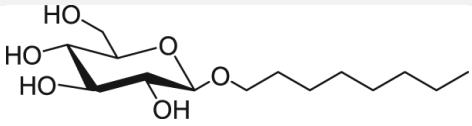
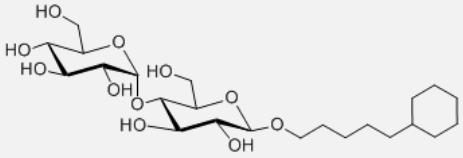
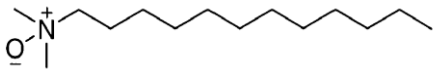
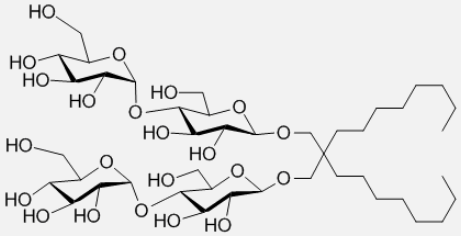
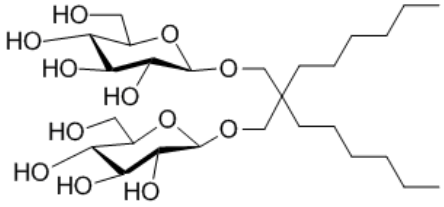
Protein stability has been associated with the formation of high-resolution X-ray diffraction data (Sonoda *et al.*, 2011). One factor that can influence the stability of a membrane protein is the detergent it is solubilised in. Therefore, to identify detergents suitable for crystallising the ApcT homologues *Gk*ApcT and *Mj*ApcT. The two transporters were chosen since crystallisation with *Gk*ApcT resulted in low diffracting crystals (Section 3.4.1), and it was anticipated to also obtain crystals of *Mj*ApcT to potentially identify the substrate binding site of this protein. Purified *Gk*ApcT and *Mj*ApcT were incubated with buffer containing different selected detergents (Table 3-3).

DDM, OG and LDAO were chosen, since they have been the most commonly used detergents for crystallisation (Parker and Newstead, 2012). Cymal-5 has been successful in crystallisation of the APC members BetP and CaiT (Schulze *et al.*, 2010; Perez *et al.*, 2011). DMNG and OGNG are examples of a new class of MNG detergents containing a central quaternary carbon and two lipophilic chains (Chae *et al.*, 2010).

Incubation overnight, followed by gel filtration, gave the stability of *Gk*ApcT and *Mj*ApcT in LDAO, OG, Cymal-5, DMNG and OGNG (using 3 x CMC of each detergent). *Mj*ApcT showed no precipitation at all but *Gk*ApcT precipitated in OG, LDAO and OGNG. For each sample, aggregated protein was separated from the residual solution via centrifugation. The supernatant was used for further analyses. To increase the detection sensitivity of the protein signal during the size exclusion experiment, the intrinsic tryptophan fluorescence ($\lambda_{excitation}$=280 nm; $\lambda_{emission}$=350 nm) was measured (Figure 3-5).

The elution profiles reveal that *Gk*ApcT is most stable in Cymal-5 and DMNG showing a monodisperse peak for both detergents (Figure 3-5). *Mj*ApcT shows a sharp peak in LDAO, OG and Cymal-5, which would suggest these detergents should be used for further experiments. Circular dichroism (CD) can be used as a method to calculate the nature of the secondary structure of a protein. CD measurements are based on the different absorption of right-handed and left-handed circularly polarised light when passing through a chiral molecule. The amide backbone of a protein provides such chirality and shows characteristic CD signals for wavelengths below 240 nm. When the amide bond is arranged in a specific array e.g. an alpha helix, this arrangement has a distinct CD profile (Figure 3-6).

Table 3-3: Selected detergents for stability assays

Detergent name	Chemical structure	Property	CMC * (% w/V)
DDM		non-ionic	0.0087
OG		non-ionic	0.53
Cymal-5		non-ionic	0.12
LDAO		zwitterionic	0.023
DMNG		non-ionic	0.0034
OGNG		non-ionic	0.058

*CMC – critical micelle concentration

Therefore, CD scans and melting curves were collected for *Mj*ApcT and *Gk*ApcT in the detergents that kept the protein stable during the SEC experiments. The results were compared with the results obtained using DDM, the detergent used for purification, to identify the stability and impact of the detergents on the secondary structure of the membrane protein. Thus, detergents that do not disrupt the secondary structure of the protein and improve its thermal stability determined via the CD melting curves, are good candidates as detergent to use in crystallisation trials.

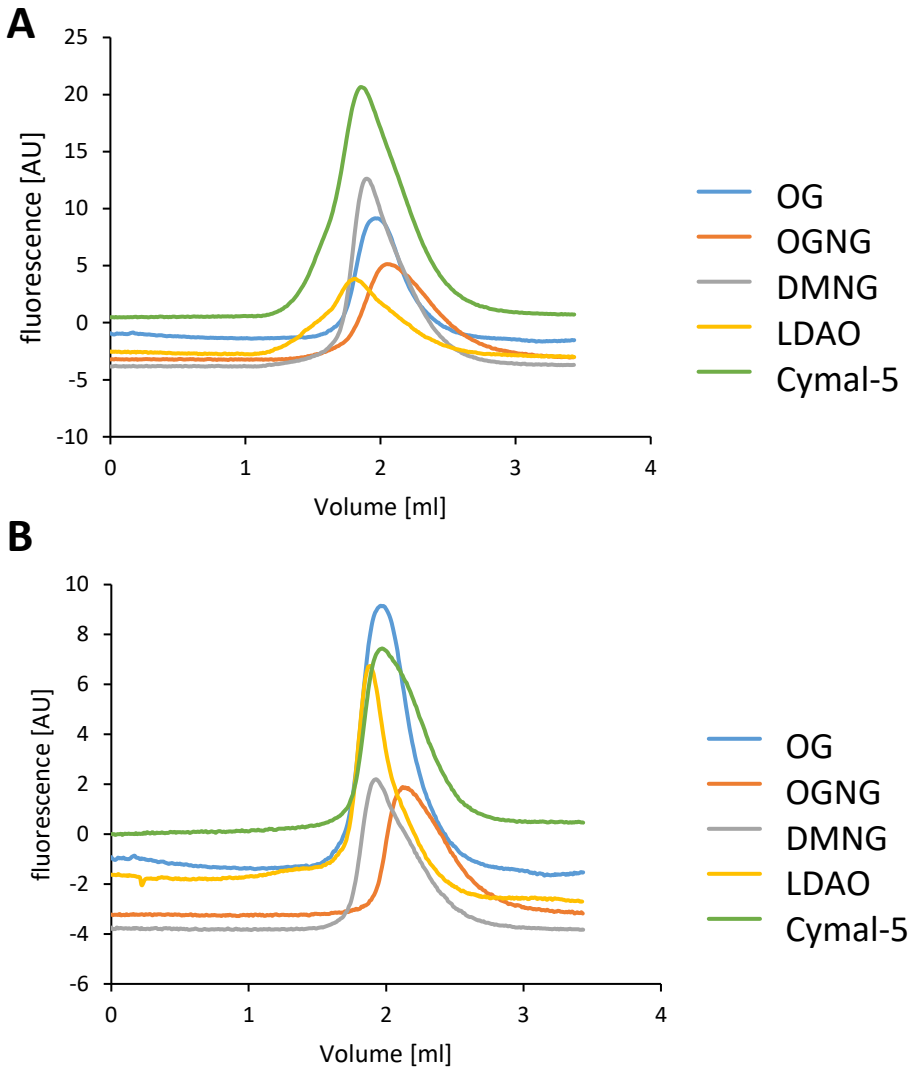


Figure 3-5 GkApcT and MjApcT size exclusion experiments in different detergents

FSEC profiles of *Gk*ApcT (A) and *Mj*ApcT (B) after incubation with OG, OGNG, DMNG, LDAO and Cymal-5. Samples were loaded onto a Superdex 200 5/150. The intrinsic tryptophan fluorescence (in arbitrary units, AU) was monitored over the eluted volume.

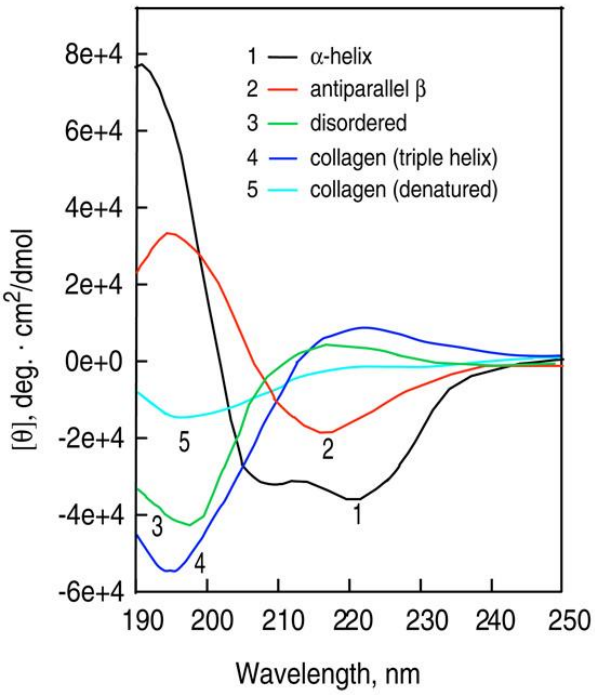


Figure 3-6: Characteristic CD spectra of different secondary structures

Figure adapted from Greenfield, 2007.

CD measurements revealed that the investigated proteins show the common CD profile for alpha helical proteins (Figure 3-6 and Figure 3-7) and share a common profile for the detergents that were tried (Figure 3-7). Thus, it appears that the protein secondary structure is unaffected and does not suffer from destabilisation at room temperature by using these detergents. The melting curves (Figure 3-7B, C) reveal that *Gk*ApcT in Cymal-5 shows no shift in its ellipticity at 220nm (which refers to the region in the CD diagram that characterises the alpha helical nature of a protein), suggesting that the protein is more stable in Cymal-5 than in DDM and DMNG. The melting temperature using DMNG is similar to that of DDM, but the ellipticity signal levels off at a lower value. This indicates that the protein is not fully degraded and that in turn the detergents partially stabilise the secondary structure at higher temperatures. Thus, both Cymal-5 and DMNG could be used for purifying and crystallising *Gk*ApcT in the future. *Mj*ApcT shows a higher temperature stability in DDM and Cymal-5.

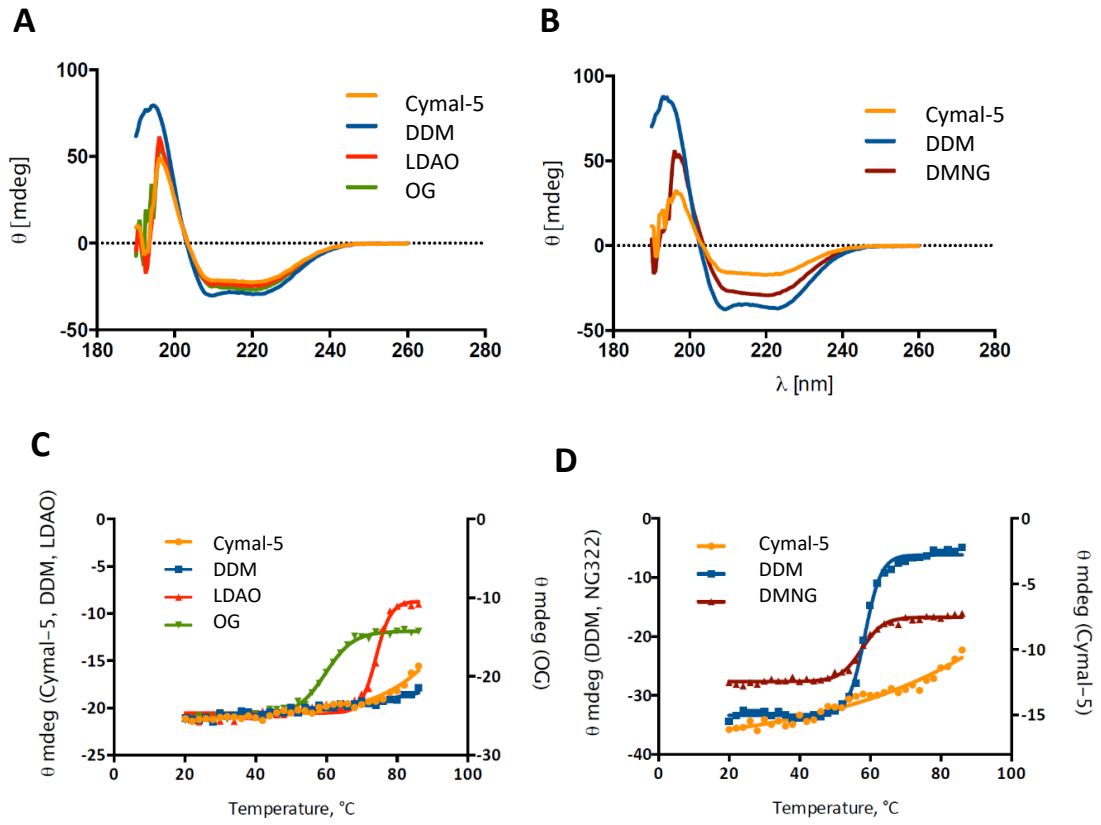


Figure 3-7: CD analysis of MjApcT and GkApcT

CD spectra of *Mj*ApcT (A) and *Gk*ApcT (B) in different detergents were taken over a wavelength range of 190-260 nm. Additionally, melting-curves were generated where a CD spectrum was taken at 2 °C temperature intervals over a range from 20-86 °C. Here, the ellipticity at 220 nm at each temperature point is plotted against the temperature for *Mj*ApcT (C) and *Gk*ApcT (D).

There is no shift in the ellipticity signal at 220 nm in DDM and only at very high temperatures for Cymal-5. Thus, Cymal-5 could be used for this protein for both purification and crystallisation trials.

Table 3-4: Melting temperatures (°C) of MjApcT and GkApcT in different detergents

Detergent	*Mj*ApcT	*Gk*ApcT
DDM	n.a.	58.5
Cymal-5	n.a.	n.a.
DMNG	-	57.6
LDAO	74	-
OG	59.8	-

3.4 CRYSTALLISATION TRIALS AND OPTIMISATION ON APCT HOMOLOGUES

3.4.1 Vapour diffusion

Vapour diffusion crystallisation experiments were carried out at 20 °C and 4 °C. No crystals were found for *Mj*ApcT, *St*ApcT, *So*ApcT and *Saci*ApcT. Crystals of *Gk*ApcT were observed in MemGold2 after eight days at 4 °C in two conditions. These conditions were optimised and used for sitting-drop and hanging-drop experiments. The optimisations involved gradients in salt and PEG concentration (Figure 3-8). Crystals appeared in the sitting drop experiments after one week but the crystals were too small for testing with normal synchrotron X-ray

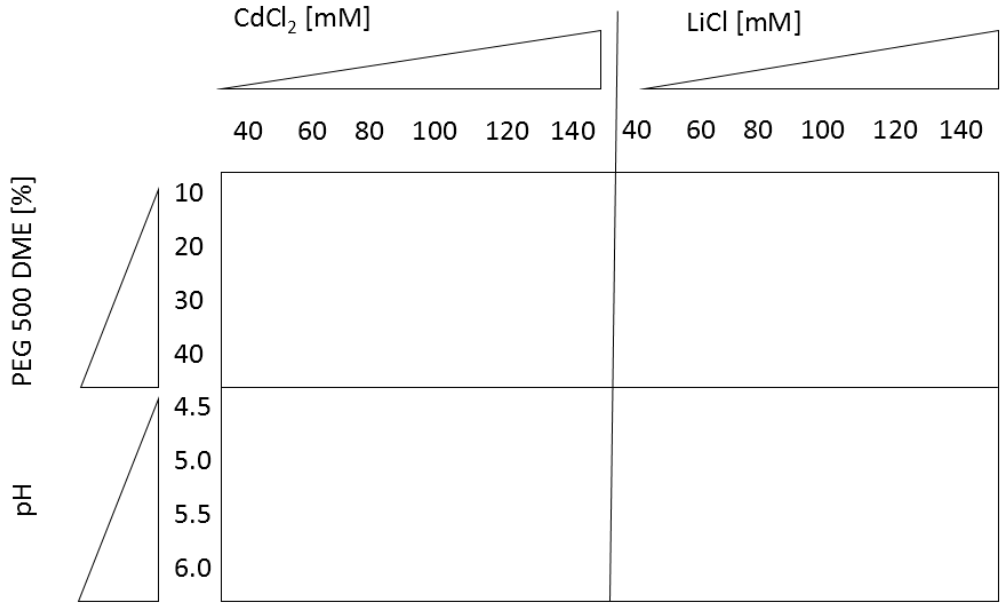


Figure 3-8: Optimisation screen of condition A8 of MemGold2 for vapour diffusion experiments with GkApcT:

The optimisation screen for vapour diffusion crystallisation trials with *Gk*ApcT is based on a crystallisation condition which contained 0.1 M LiCl, 0.1 M $CdCl_2$, 0.1 M NaAc pH 4.5, 30% (v/v) PEG400.

beam sizes.

After further extensive optimisation of the crystallisation conditions using vapour diffusion larger crystals of *Gk*ApcT were obtained (Figure 3-9) but their diffraction quality did not improve to resolutions higher than 10 Å.

Therefore, another optimisation trial was carried out. It showed that the crystals could be reproduced but they did not get sufficiently larger with varying PEG or salt concentrations.

A change of detergent for crystallisation was considered as an alternative strategy. The initial detergent screens performed for *Gk*ApcT suggested DMNG and Cymal-5 as potential alternatives to DDM for crystallisation trials. The protein was purified as usual in DDM but the final gel filtration step was performed in buffer containing 0.0102 % of DMNG and 0.36 % of Cymal-5. *Gk*ApcT was concentrated as usual to 10 mg/ml and the protein used for sitting drop crystallisation trials using MemGold1 and MemGold2, and then stored at 4 °C and 18 °C respectively, but the diffraction quality did not improve with the listed detergents.
In parallel with these experiments, crystallisation in a lipid environment, the lipidic cubic phase (LCP), was also performed.

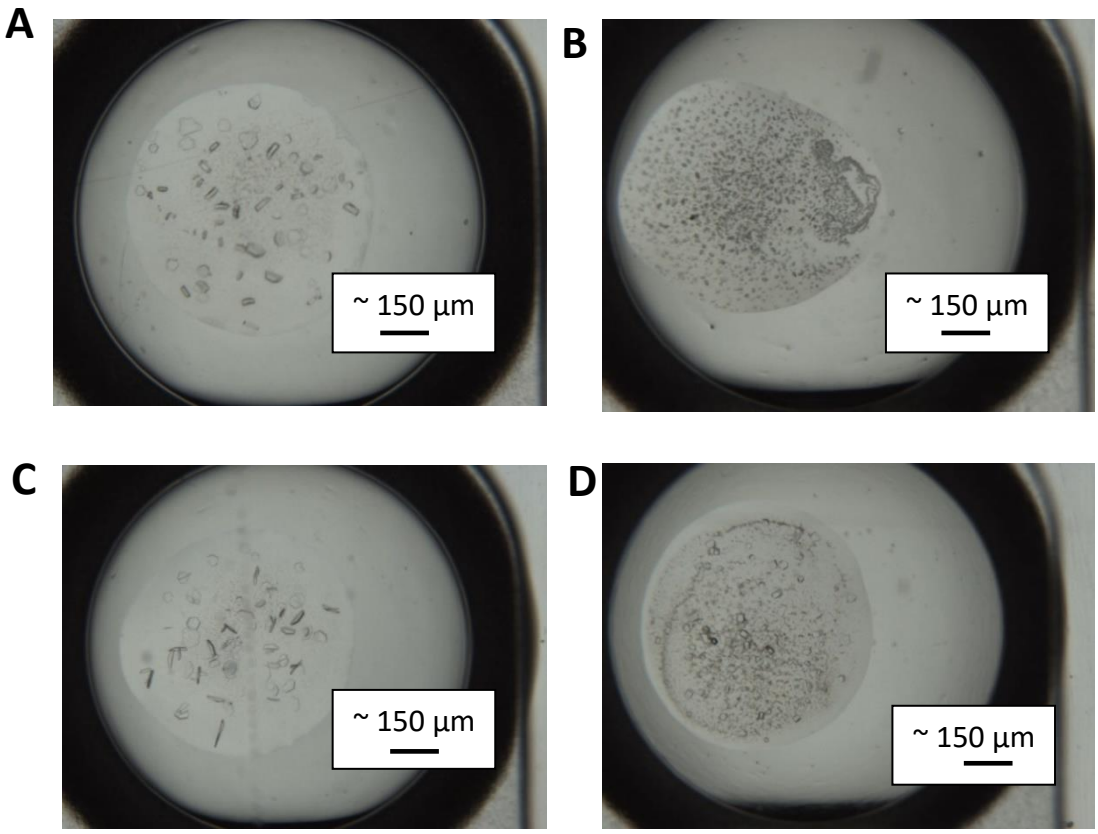


Figure 3-9: Examples of crystals of GkApcT in vapour diffusion sitting drops

*Gk*ApcT crystals grew in sitting drop plates after 7 days. Shown are crystals of *Gk*ApcT (purified in DDM) in (A) 0.1 M sodium chloride, 0.1 M MES pH 6.5, 36% v/v PEG 300, 4 °C; (B) 0.1 M Lithium Chloride, 0.1 M CdCl, 0.1 M NaAc pH 4.5, 30% (v/v) PEG400, 4 °C; (C) 0.2 M NaCl, 0.1 M HEPES-NaOH pH 7.0, 29 PEG400, 4 °C and (D) 29 % v/v PEG 400, 0.1 M HEPES-NaOH pH 7.0, 0.1 M NaCl, 0.1 M $LiSO_4$,0.012 M Cymal-2, 4 °C.

3.4.2 Crystallisation in the lipidic cubic phase

All screened APC homologues, *Mj*ApcT, *Gk*ApcT, *So*ApcT and *St*ApcT were subject to crystallisation in LCP. As a starting condition, 10 mg/ml of protein was reconstituted into monoolein, a commonly and successfully used lipid for this crystallisation technique. The concentration was chosen based on the published structures of other bacterial membrane proteins in LCP. In comparison, for eukaryotic membrane transporters, it has been recommended to use an increased protein concentration, as has been shown to be beneficial for crystallising GPCRs (concentration used: 40 mg/ml) (Cherezov *et al.*, 2007; Byrne *et al.*, 2016; Zhang *et al.*, 2017b).

3.4.2.1 Crystallisation of ***Mj***ApcT in LCP

Micro-crystals of *Mj*ApcT were observed in LCP and the conditions giving these crystals optimised to obtain crystal sizes suitable for the use in a synchrotron X-ray beam. A sparse matrix optimisation was performed, using gradients in salt and PEG concentrations (Figure A 4A). Bigger crystals (~ 25 µm in the longest dimension) could be observed with lower PEG concentrations. Crystals were directly mounted into cryo-loops, cryo-cooled into liquid nitrogen, without additional cryo-protectant, and screened at 100 K on beam lines I24 and I04 at the Diamond Light Source (DLS), but no protein diffraction patterns were observed. This can be due to different reasons. Firstly, this could have been because no crystals were present in the loops. It should be noted that the mounting of the crystals out of the LCP environment is a difficult task. When a well has been opened to harvest the crystal inside the LCP, the contact with air during this process cause the crystals to dissolve. Since the visibility of a crystal is lost once it is mounted into the cryo-loop, it is difficult to know if a crystal has been successfully mounted. Therefore, X-ray diffraction grid scans must be performed to identify the presence of a crystal in the cryo-loop (Cherezov *et al.*, 2009).
Secondly, the crystals screened at the beam lines might not have been protein crystals. The identification of LCP grown protein crystals relied on their birefringent properties but salt crystals can exhibit the same signal. Thus, the nature of the observed birefringent object can be misinterpreted. New techniques have been developed to identify protein crystals in

LCP. One option is the Second Order Nonlinear Imaging of Chiral Crystals (SONICC) system (Kissick *et al.*, 2010, 2011) which relies on the non-linear optical process of second-harmonic generation (SHG) by chiral objects. A disadvantage of this method is that not only protein crystals are visible, but other chiral objects such as salts, too. Another approach is the use of a UV microscope to help distinguish between salt and protein crystals, but it cannot be used for proteins that do not contain naturally occurring aromatic residues. To understand why some of the conditions did not give any crystals in LCP, LCP fluorescence recovery after photo-bleaching (LCP-FRAP) (Xu *et al.*, 2011) could be used to investigate the mobility of the protein in the LCP. It has been proposed that the higher the mobility, the more likely it will be to crystallise (Xu *et al.*, 2011).

Further optimisation (Figure A 4A & B) and repetition of the screens already used, resulted in crystals of *Mj*ApcT of ~ 30 µm in size (Figure 3-10).

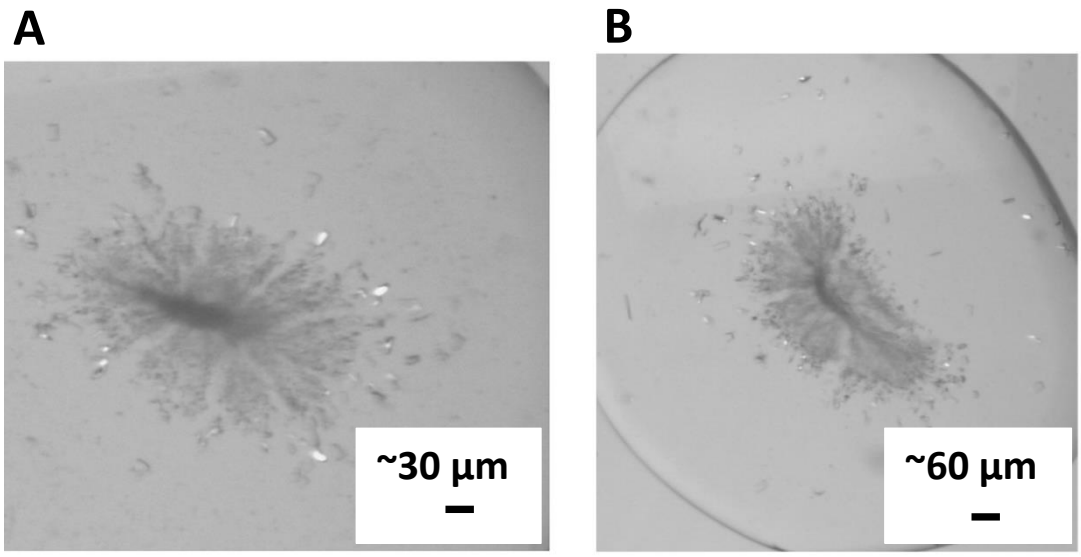


Figure 3-10: MjApcT crystals grown in LCP

LCP grown crystals of *Mj*ApcT were observed in conditions containing (A) 27 % PEG500 DME, 0.05 M $MgSO_4$, 0.1 M MES pH 6.0 and (B) 26 % PEG500 DME, 0.1 M $MgSO_4$, 0.1 M MES pH 6.0. The crystals are shown under cross-polarised light and are roughly 20 µm in size.

Again, crystals were mounted and screened at beamline I24 at DLS. This time, protein diffraction patterns were observed (Figure 3-11) and a complete data set was collected for *Mj*ApcT grown in 26 % PEG500 DME, 0.1 M $MgSO_4$, 0.15-0.18 M MES-NaOH pH 6.0. Auto processing with the XIA2 pipeline (Winter, 2010) using XDS (Kabsch, 2010) was performed and resulted in a data set up to 3.5 Å resolution (Table 3-5). The spacegroup was $P2_1$ with unit cell parameters of a = 54.14 Å, b = 222.97 Å, c = 92.05 Å and α = 90.00°, ß = 100.78°, γ

= 90.00°. Analysis of the Matthews coefficient suggested the presence of four *Mj*ApcT molecules per asymmetric unit.

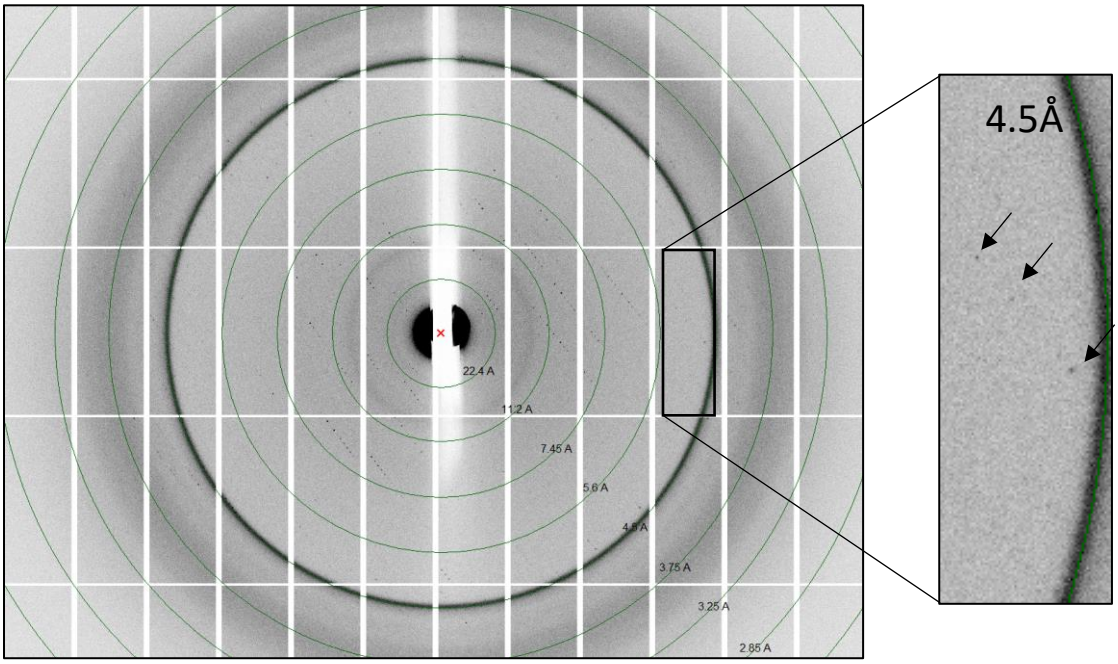


Figure 3-11: Diffraction pattern recorded for LCP grown crystals of MjApcT

Arrows indicate recorded reflections on the detector.

Molecular replacement with PHASER (McCoy *et al.*, 2007) using the published structure of *Mj*ApcT (3GIA, 2.32 Å), with deletion of flexible loop regions, provided initial phase information, followed by a refinement step in BUSTER (Bricogne *et al.*, 2016). Due to the low resolution (3.5 Å) of the maps, the position of side chains could not be verified with certainty, but the overall fold of the structure suggested a closed conformation identical to the published structure (Shaffer *et al.*, 2009) (Figure A 6). Thus, the conformation of the protein did not change in the lipid environment compared with the detergent environment. All other screening attempts to optimise the crystallisation conditions of the crystal by using other host lipids and additive screens did not result in an increase in the diffraction resolution.

Table 3-5: Data collection and refinement parameters of MjApcT

	***Mj*ApcT**
Beamline	I24, DLS
Detector	PILATUS 6M
Wavelength (Å)	0.9686
Temperature (K)	100
Exp. Time (s)	0.1
Osc. (°)	0.5
Total number of images	360
Mosaicity (°)	0.311
Wilson B-factor (Å²)	68.9
spacegroup	P 1 2_1 1
unit cell	54.14 Å, 222.96 Å, 92.05 Å
	90.00 °, 100.78 °, 90.00 °
Resolution (Å)[1]	90.43-3.5 (3.74-3.5)
R_{merge}	0.455 (2.011)
R_{meas}	0.544(2.394)
R_{pim}	0.295 (1.287)
Total number of observations	90518 (16572)
Total number unique	26998(4881)
I/σ(I)	2.1 (0.6)
$CC_{1/2}$	0.823 (0.428)
Completeness (%)	99.9 (99.9)
Multiplicity	3.4(3.4)

[1]Values for highest resolution shell are given in parentheses

3.4.2.2 Crystallisation of *Gk*ApcT in LCP

Initial crystals of *Gk*ApcT were found in the MemMeso screen using monoolein as the host lipid, but the crystals were very small (~ 5 µm). The hit condition contained 0.1M sodium citrate tribasic dihydrate pH 5.0 and 40 % v/v PEG 200. A sparse matrix optimisation screen with varying PEG concentrations and different pHs (Figure A 5A), increased the crystal size to ~10 µm. Those crystals were harvested and screened for diffraction at the I24 beamline at DLS. The maximal visible diffraction went to 4-5 Å but the crystals were too small and radiation sensitive to obtain a full data set. Therefore, further optimisation of the crystallisation conditions was indicated. The specific lipid used for crystallisation is another component that can influence the crystal size. Thus, different lipids, and a mixture of the lipids monoolein and cholesterol were tested. The latter is usually used for the crystallisation of GPCRs (Hanson *et al.*, 2008; Salom *et al.*, 2013; Fenalti *et al.*, 2015; Byrne *et al.*, 2016). The

monoolein-cholesterol composition and the addition of the substrate alanine to the crystallisation screen led to a slight increase in crystal size (Figure 3-12A) but did not improve the diffraction resolution. Using the salt screen from Hampton Research, needles with a maximum size of 20 µm in the longest dimension (Figure 3-12B & C) appeared after a few days and subsequent screening at the DLS I24 beamline revealed that those crystals diffracted to 3.5 Å visible resolution. A full data set was collected and processed using XDS. Unfortunately, the processing revealed that, due to radiation damage, viable data could only be obtained from the first 10°-20° of data collection. However, this was enough to determine the space group of the crystal lattice to be $P4_12_12_1$. The inclusion of additional data sets using additional crystals still did not result in a sufficiently complete data set for structure solution by MR.

The crystals obtained with the LCP method are of sizes smaller than 20 µm and hence the use of a micro-focus beam is indicated. In principle, crystals that small, and normally overlooked for the purposes of synchrotron data collection, could potentially be used for X-ray data collection at an X-ray Free-Electron Laser source (XFEL) via SX. The development of SX originated in macromolecular crystallography experiments using XFELs, during which each crystal is destroyed in one single shot and therefore a continuous replenishment of fresh crystals has to be provided (Weierstall, 2014). SX can be utilized if crystallization has resulted in many small crystals, deemed too small for collection of a full data set via the conventional goniometer-based rotation method.

The fundamental advantages of SX over conventional (i.e., goniometer-based) crystallography are twofold: first, time-resolved studies of fast (< ms), and, more importantly, irreversible processes become accessible. Secondly, in the case of serial crystallography using the femtosecond-short pulses of an XFEL, data collection can be performed at incident X-ray intensities beyond conventional radiation damage limits in a so-called “diffraction-before-destruction” approach during which a “snapshot” of the crystal is obtained (Barty *et al.*, 2011; Chapman *et al.*, 2011). SX pioneered at XFELs, is usually termed Serial Femtosecond Crystallography (SFX). It has, as well as advancing the data collection and data processing procedures themselves, resulted in significant developments in batch crystallization of the protein of interest in buffer media containing many small microcrystals (Wu *et al.*, 2015b), as well as microfluidic injection of these sample-saturated media in a continuous stream for rapid screening (Heymann *et al.*, 2014; Perry *et al.*, 2014; Pawate *et al.*, 2015) or

using LCP injection systems to deliver crystals of membrane protein crystals (DePonte *et al.*, 2008; Weierstall *et al.*, 2012, 2014).

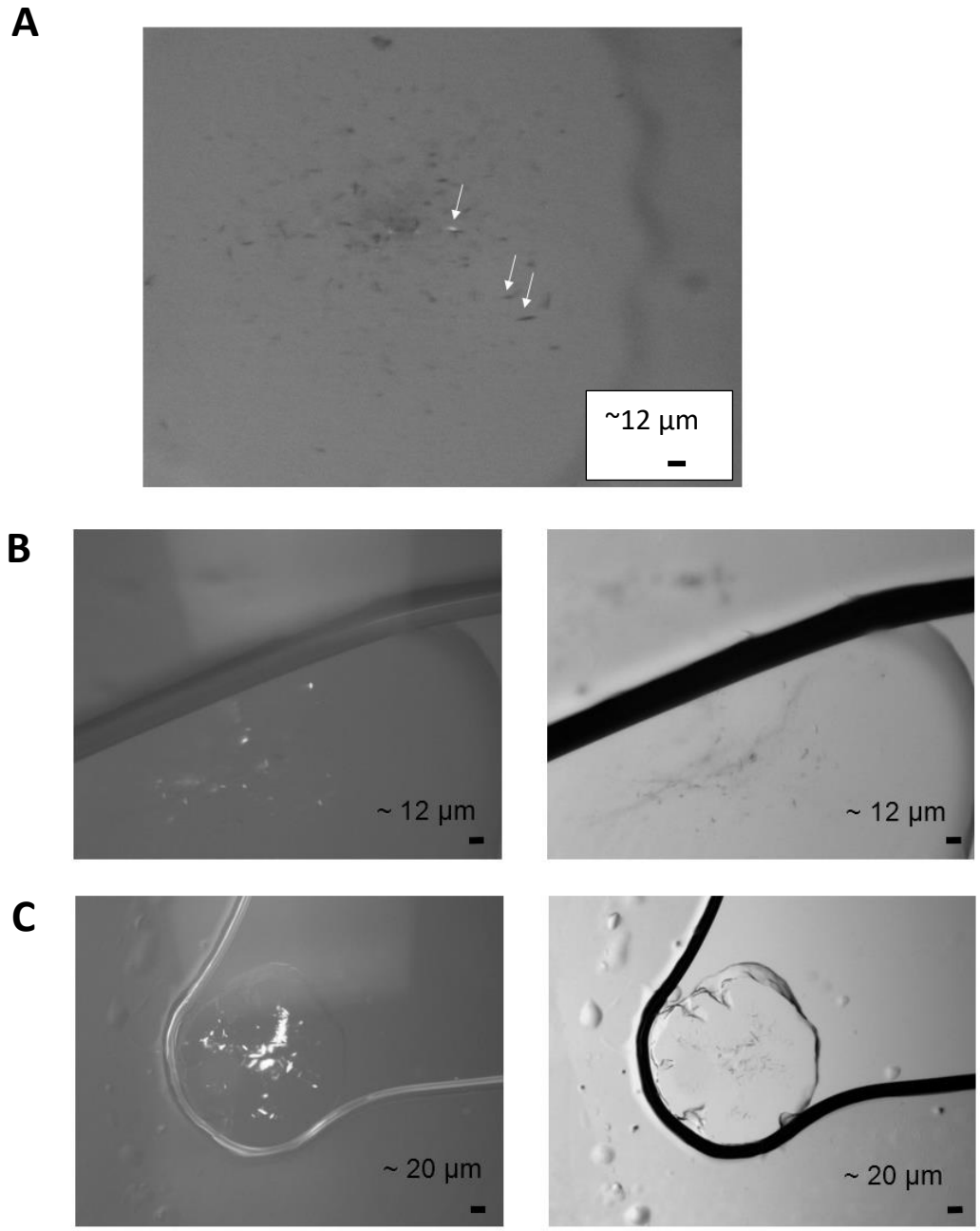


Figure 3-12: LCP crystals of GkApcT after optimisation

Initial hit of *Gk*ApcT crystals in 40 % PEG200, 0.1M NaCit pH 5.0, 10 mg/ml *Gk*ApcT grown at 20 °C using MO as the lipid matrix(A). Crystal hits after optimisation using the Hampton Research stock options screen(B). Crystals appeared after 1 day and were grown in 36 % PEG200, 0.1 M NaAc pH 4.0, 0.1M Knatart at 20 °C using MO:Chl as hosting lipid. The crystals were further optimised by changing the KNatart concentration to 0.05 M (C). Crystals are shown in cross polarised (left) and visible light (right). Arrows indicate individual crystals.

The method of SX has also been transferred to synchrotrons (Gati *et al.*, 2014) and even membrane protein crystals in LCP have been utilized, as was first demonstrated by experiments with bacteriorhodopsin at the ESRF (Nogly *et al.*, 2015).

For this thesis, a SX experiment at the P11 beamline at PETRA-III, Hamburg, has been conducted in collaboration with the group of Prof. Henry Chapman, CFEL, DESY, Hamburg, and will be described in Chapter 7. In the preparation process for batch crystallization of the proteins for the experiment at P11 the purification protocol of *Gk*ApcT was adapted so that instead of purifying from 2L of cell culture, 4 L was prepared and purified as batch and the protein was concentrated to 30 mg/ml instead of to 10 mg/ml, as a higher protein concentration gave a higher crystal density in the crystal growth set up used for the SX experiment. Simultaneously, the same protein batch of *Gk*ApcT was diluted from 30 mg/ml to 10 mg/ml with a gel filtration buffer containing DDM for the LCP plate set up, and then rescreened for crystals using the optimisation screens from past trials. Using the initial crystallisation condition from the MemMeso screen, together with the addition of the stock option salt screen from Hampton Research, new crystals appeared in different conditions than previously observed, such as the crystallisation condition containing 100 mM potassium fluoride (KF). In it, 20 µm diamond shaped crystals appeared after one week. Further optimisation (Figure A 5B) resulted in crystals that were either rod or cube shaped, and varied in size from 20-30 µm (Figure 3-13).

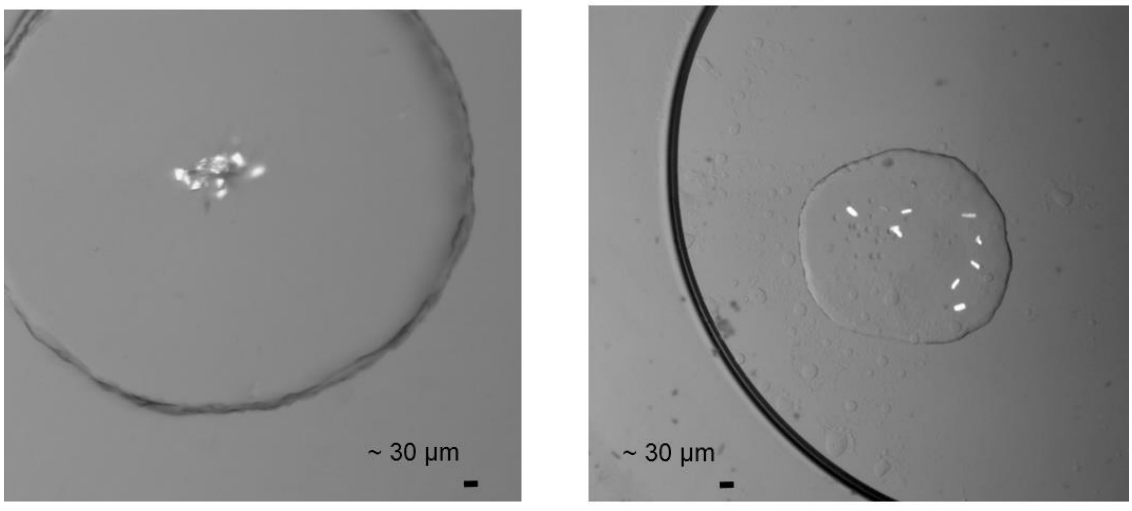


Figure 3-13: GkApcT crystals

Crystals of *Gk*ApcT appeared after 1 day and grew to their full size as shown above in 7-14 days. Crystals appeared in conditions containing 28-34 % PEG400, 0.1 M NaAc pH 4.0, 0.08-1.2 M KF. Monoolein with 10 % cholesterol was used as the host lipid.

3.5 Conclusion

To obtain a full picture of the function of a protein, biochemical and structural data are crucial. For membrane proteins, producing the required amounts of pure protein in sufficient quantities can be challenging, and common difficulties include low expression yields, as well as e.g. destabilisation of the protein when extracted from its membrane environment and during progression to a pure protein solution. Sampling over a selection of homologous proteins represents a viable way to increase the chances of successful protein expression, purification, and finally crystallisation.

Here, several candidates were assayed through this pipeline to obtain structural data of a substrate bound ApcT transporter, leading to novel insights into the transport mechanism of this family and its human homologues.

Five of the seven proposed candidates were overexpressed in high quantities in *E. coli*, FSECS was used to identify DDM as the detergent of choice for purification, and sufficient yields were obtained of pure *St*ApcT, *So*ApcT, *Mj*ApcT and *Gk*ApcT. Crystallisation trials with the four constructs resulted in data sets from LCP grown crystals of *Mj*ApcT and *Gk*ApcT.

4 The transporter function of *Gk*ApcT

To understand how a transporter works, which substrates it transports, and under which conditions this transport is facilitated, functional studies are vital. For this, an environment is necessary that can maintain electrochemical gradients, since this is the driving force of substrate translocation by the secondary active transporter studied here. Additionally, it is desirable to study the transporter as close as possible to its native membrane environment. Furthermore, it should be possible to easily change the experimental settings to enable screening a range of conditions under which the transporter can be analysed. The use of liposomes for transporter activity measurements provides the stated properties.

4.1 Liposome based assay to understand transport of amino acids across the membrane

Liposomes are lipid vesicles into which a membrane protein can be incorporated and that offer the possibility of manipulating the exterior and interior medium surrounding their membrane. Liposome-based methods are therefore feasible to use in the study of the mechanism of membrane transporters. The lipids that are used for liposome preparation depend on the protein that is reconstituted and should be tested in advance. *E. coli* polar lipids can be used for example for membrane proteins of bacterial origin. The lipid mixtures are derived from extraction from *E.coli* total lipids with acetone, separating the polar lipids from the nonpolar lipids and lipopolysaccharide. The *E.coli* polar lipid mixture is composed of the cationic phosphatidylethanolamine (PE), and the anionic phosphatidylglycerol (PG), and cardiolipin (CL) (Figure 4-1A) at 67 %, 23.2 % and 9.8 % of the lipid mixture, respectively (https://avantilipids.com/product/100600/). Usually, the *E.coli* polar lipids are supplemented with phosphatidylcholine (PC) for liposome based assays. Alternatively, a mixture of synthetic POPE:POPG in a 3:1 ratio, (Figure 4-1B) has been used for the arginine/agmatine antiporter AdiC (Tsai *et al.*, 2012). POPE and POPG are synthetic structurally similar lipids to PE and PG from *E.coli*, therefore mimicking the lipid composition of *E.coli*.

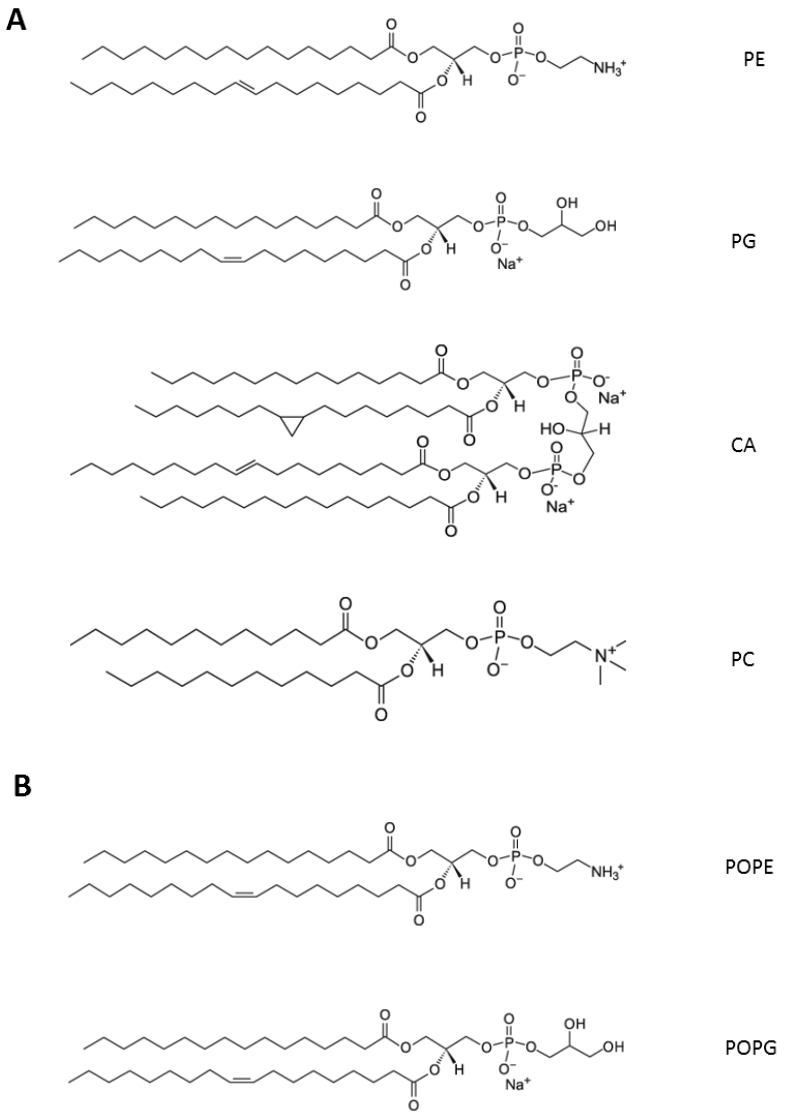


Figure 4-1: Chemical structures of lipids used for liposomes

E.coli:PC (3:1) lipid mixture: *E.coli* polar lipids consist of phosphatidylethanolamine (PE), phosphatidylglycerol (PG) and cardiolipin (CA). Here, the mixture was supplemented with phosphatidylcholine (PC) (A). Synthetic POPE:POPG (3:1) liposomes consisted of 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE) and 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoglycerol (POPG) mimicking a similar composition to that of *E.coli* polar lipids (B).

The membrane protein of choice is usually kept in solution with a detergent which results in the destabilisation of the liposomes during the reconstitution process. This makes the liposomes more ion-permeable during liposome-based assays, especially when gradients across the lipid membrane are applied. Consequently, removal of the detergent during the membrane protein reconstitution is desirable. For liposome preparation, the solvent that

keeps the lipid mixture in solution has to be removed first. The dried lipids are then rehydrated in an aqueous solution (here 50 mM KPi pH 7.0). In this process, the dried lipid film starts to swell and forms multilamellar vesicles (MLVs) in the aqueous solution (Figure 4-2). Before the membrane protein can be reconstituted into the lipid environment, the MLVs have to be pre-treated either by sonication or extrusion through a small filter unit to obtain unilamellar vesicles (ULVs), which consist of a single lipid bilayer.

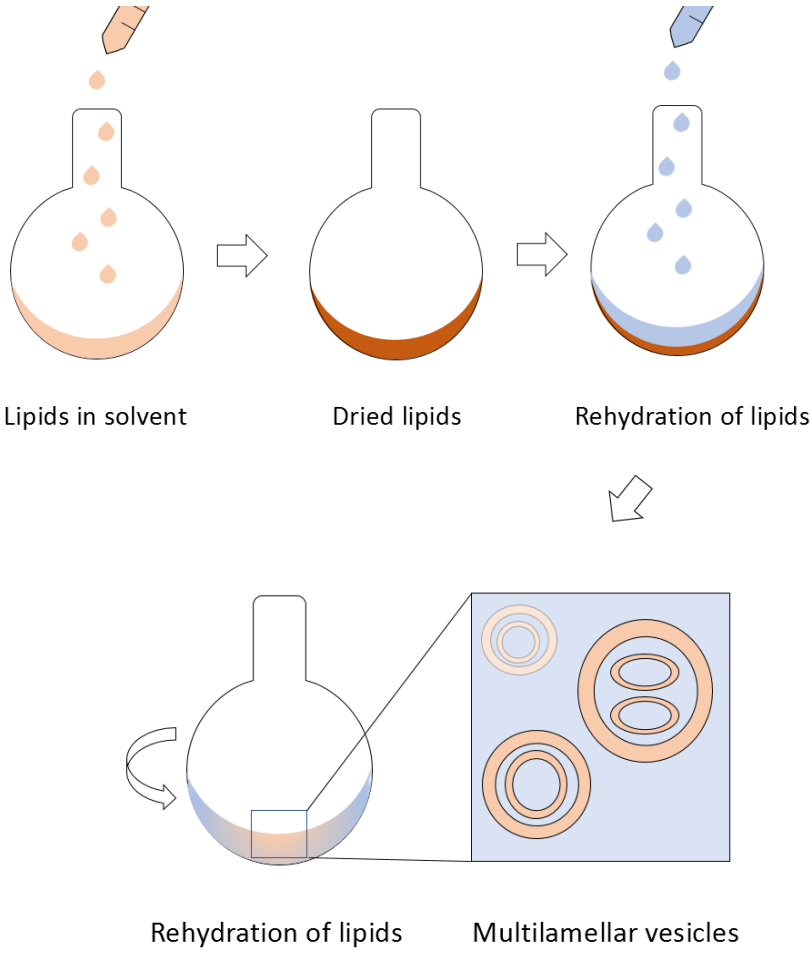


Figure 4-2: Preparation of lipid vesicles

Lipids are mixed in a round bottomed flask in the desired ratio. They are kept in solution using a solvent that is removed under vacuum. The dried lipid film is rehydrated and resuspended using a buffer solution, thereby initiating the swelling of the lipid bilayer towards the formation of multilamellar vesicles (MLVs). Unilamellar vesicles are obtained by sonication or extrusion of the MLV solution.

The membrane protein is then titrated into the ULV solution and the mixture is incubated to ensure an interaction of the membrane protein with the lipids. During this process, the

detergent also forms contacts with the lipids. To separate the detergent from the protein-micelle-lipid mixture, different removal methods have been introduced (Figure 4-3).

(1) Rapid dilution (Figure 4-3A)

The rapid dilution method, which was the method of choice in this thesis, can be applied to proteins that are stable in detergents with high critical micelle concentration (CMC), such as DM. The protein-micelle-lipid mixture is titrated into an excess of aqueous solution devoid of detergent, thereby reducing the detergent concentration below its critical micelle concentration. The reconstituted protein is then retrieved via centrifugation and resuspended into buffer. To remove any residual detergent, additional methods such as dialysis, or the addition of biobeads have to be carried out.

(2) Dialysis (Figure 4-3B)

When using high CMC detergents, dialysis of the protein-micelle-lipid mixture in non-detergent buffer can be carried out to lower the detergent concentration. However, for some detergents, this procedure takes 1-2 days of dialysis with additional buffer changes to increase detergent removal. The protein-micelle-lipid mixture is transferred into a dialysis tube and dialysed in a high volume (1-3 L) of non-detergent buffer. To enhance the removal rate of the detergent, biobeads can be added to the buffer.

(3) Biobeads (Figure 4-3C)

For proteins solubilised in low CMC detergents, impractical for detergent removal via dialysis or rapid dilution, but sometimes necessary to guarantee the stability of the membrane protein, biobeads can be used for detergent removal. After incubation of the lipid mixture with the membrane protein, the detergent can be removed via dialysis by adding biobeads to the dialysis buffer. The surface of the biobeads attracts the detergent, removing it from the aqueous solution of the buffer driving continuous dialysis. The biobeads need to be exchanged frequently for increased removal efficiency.

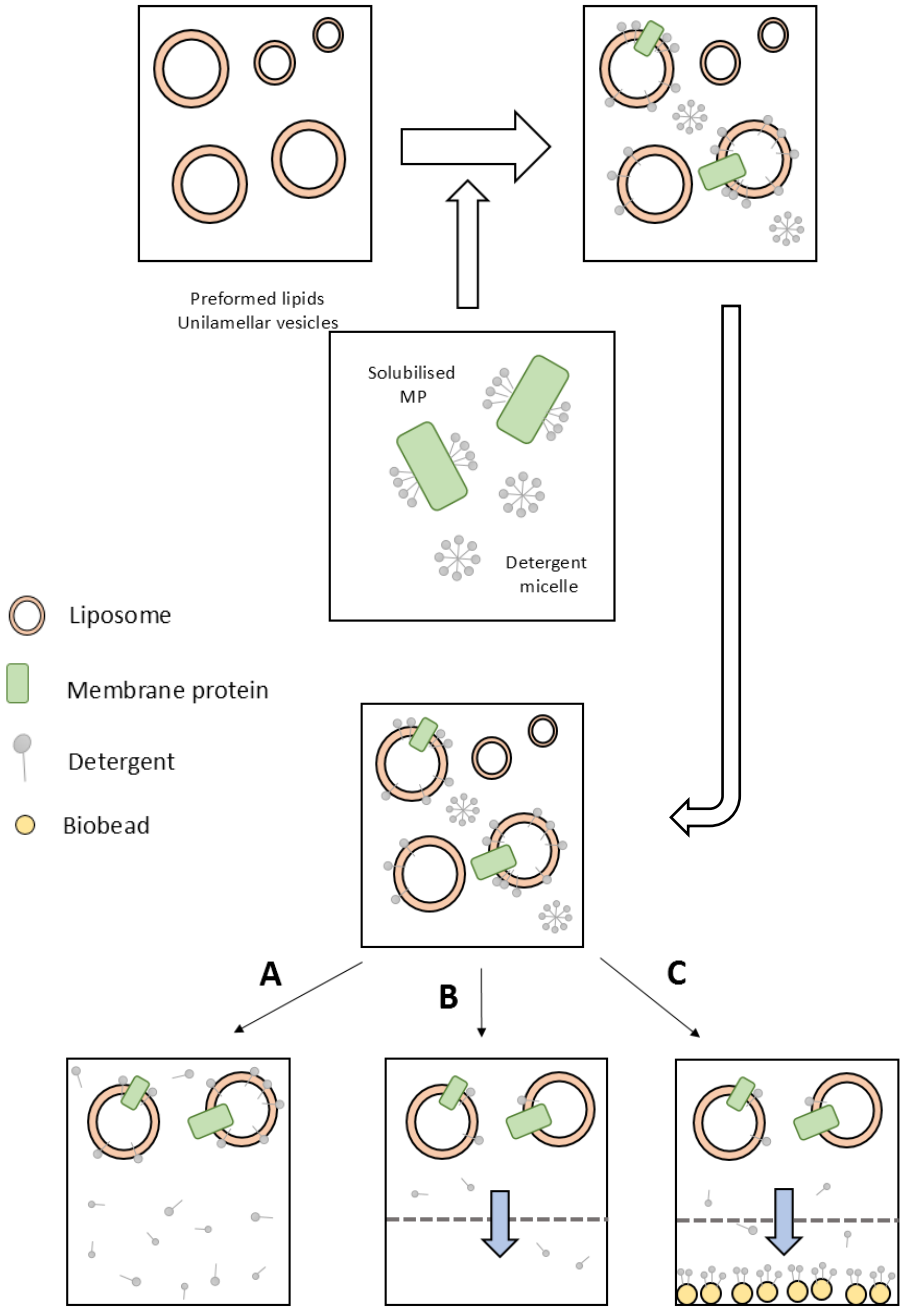


Figure 4-3: Preparation of proteoliposomes

Preformed lipid vesicles are incubated with the protein-micelle solution in the desired protein-lipid ratio. The detergent micelles and proteins interact with the lipid environment. Detergent removal is achieved by rapid dilution in detergent-free buffer (A), dialysis against detergent-free buffer (B), or using biobeads and dialysis (C).

After successful reconstitution, the membrane protein-lipid mixture can be concentrated via centrifugation, resuspended in buffer and stored at -80°C. The reconstitution efficiency

can be determined by loading each sample on an SDS gel together with a known amount of the same protein as a marker.

The different variables involved in the process can be adjusted to increase the reconstitution efficiency, for instance by using a different detergent, changing the detergent removal strategy, using a different lipid composition or a different lipid to protein ratio.

4.2 *IN VITRO* ASSAYS

To understand the function of *Gk*ApcT and identify its specific properties, liposome-based assays were performed. The host lipid and assay conditions were initially based on the experiments performed with *Mj*ApcT (Shaffer *et al.*, 2009). *Mj*ApcT showed transport activity for alanine in *E.coli* polar lipids:PC (3:1 w/w) liposomes (Figure 4-1A), which have been used successfully for activity measurements on other APC members (Fang *et al.*, 2007; Singh *et al.*, 2007; Shaffer *et al.*, 2009; Ma *et al.*, 2012a; Perez *et al.*, 2012). In the previously described experiments from Shaffer *et al.*, *Mj*ApcT was reconstituted in a 1:110 protein:lipid ratio and uptake of tritiated alanine was measured using an internal buffer containing 25 mM citrate phosphate buffer pH 7.0, 100 mM NaCl, and 4 mM Ala, and an external buffer with 25 mM citrate phosphate at the desired pH, 100 mM NaCl, and 500 nM ^{3}H-Ala (Shaffer *et al.*, 2009).

Before the liposome reconstitution of *Mj*ApcT and *Gk*ApcT and subsequent uptake measurements, it was determined whether the *E.coli*:PC (3:1) liposomes show any ion permeability when using the applied pH gradient (internal pH = 7.0, external pH = 4.0) used to measure pH-dependent uptake of alanine by *Mj*ApcT. Any leakage of the liposomes would equilibrate the pH-gradient during the measurements hindering any pH-dependent uptake, and thus obscure the experimental data. The time-course measurements for *Mj*ApcT were performed for a maximum of 30 min under the conditions stated above.

To investigate a potential proton leakage, liposomes devoid of protein were loaded with internal buffer containing 25 mM citrate phosphate buffer pH 7.0**,** 100 mM NaCl, 2 mM $MgSO_4$, and 1 mM pyranine, a pH-sensitive fluorescent dye. The liposomes were then diluted 1:100 into external buffer (25 mM citrate phosphate buffer pH 4.0 to pH 7.0**,** 100 mM NaCl and 2 mM $MgSO_4$). The time-course measurements show that a pH gradient of one or more pH units cannot be maintained by the liposomes (Figure 4-4A) making them unsuitable

for the anticipated experiments using a pH gradient of three pH units as described by Shaffer *et. al*. Therefore, the lipid composition for subsequent experiments was changed and checked again for proton-permeability using the pyranine assay. For this, a POPE:POPG mixture of 3:1 was used as it has been reported to withstand high pH gradients (Tsai *et al.*, 2012). The liposomes could hold a pH gradient of two pH units and showed only a slight proton-leakage over 9 min (Figure 4-4B), making it possible to perform alanine uptake experiments on *Mj*ApcT and *Gk*ApcT with this pH gradient.

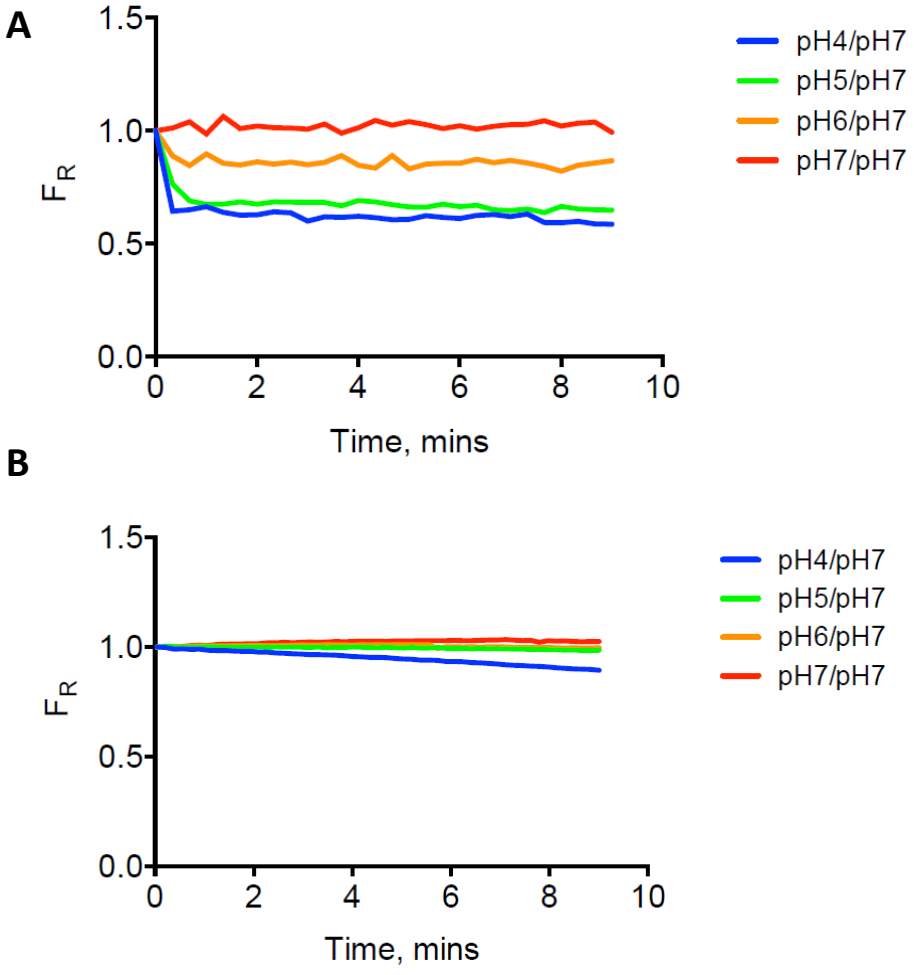


Figure 4-4: Proton leakage measurements of liposomes of different composition

Liposomes composed of a mixture *E.coli*:PC (A), and POPE:POPG (B) were loaded with 25 mM citrate phosphate pH 7.0, 100 mM NaCl, 2 mM $MgSO_4$ and 1 mM pyranine to monitor any proton influx. The liposomes were then diluted into 25 mM citrate phosphate pH 4.0 to pH 7.0 (blue to red), 100 mM NaCl and 2 mM $MgSO_4$. The relative fluorescence (FR), resulting from division of the fluorescent signal ($\lambda_{emission}$=510 nm and $\lambda_{excitation}$ = 460 nm or 415 nm) was monitored over time.

4.3 ESTABLISHING CONDITIONS FOR ACTIVITY MEASUREMENTS ON *GK*APCT

Similar conditions to those used to investigate the transport activity of *Mj*ApcT were used for *Gk*ApcT. As a control, *Mj*ApcT was similarly assayed. Both proteins were reconstituted into POPE:POPG (3:1) in a protein to lipid ratio of 1:60 was chosen. Similar reconstitution efficiencies were obtained for the two proteins (Figure 4-5).

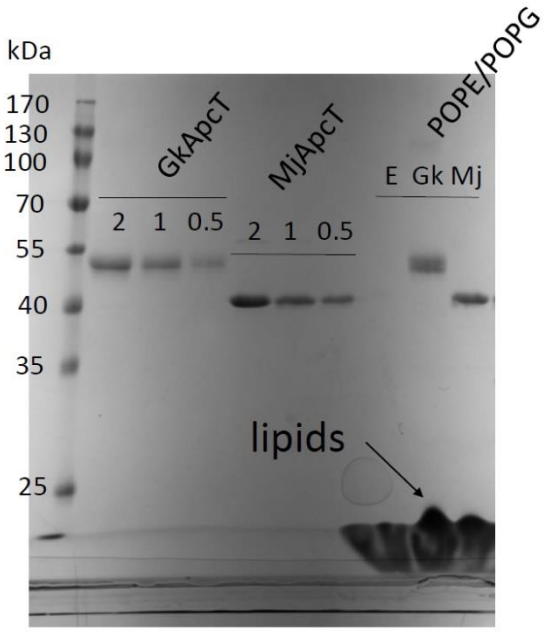


Figure 4-5: Reconstitution of MjApcT and GkApcT

*Mj*ApcT and *Gk*ApcT were reconstituted in a 1:60 (protein to lipid) ratio. The efficiency of the reconstitution was determined via 12 % SDS-glycine gel analysis by comparing the amount of protein in the proteoliposomes to known amounts of 2, 1 or 0.5 µg of the protein loaded onto the gel. Both samples of proteoliposomes contain similar amounts of protein and were reconstituted to ~100 %. E - empty liposomes, *Gk* – *Gk*ApcT, *Mj* - *Mj*ApcT

4.3.1 pH-dependent uptake of alanine

Time-dependent uptake of tritiated alanine was monitored by applying a pH gradient (internal pH = 7.0, external pH = 4.0), but no sodium or potassium electrochemical gradient, across the membrane using 5 µg (1.97 µM) of protein per experimental time point. The measurement results plotted in Figure 4-6 show that uptake of alanine by *Gk*ApcT was approximately six times higher than by *Mj*ApcT. Thus, *Gk*ApcT is an alanine transporter and radiolabelled alanine could be used for future experiments. In comparison, *Mj*ApcT showed specific transport of alanine, serine, glutamate and glutamine (Shaffer *et al.*, 2009).

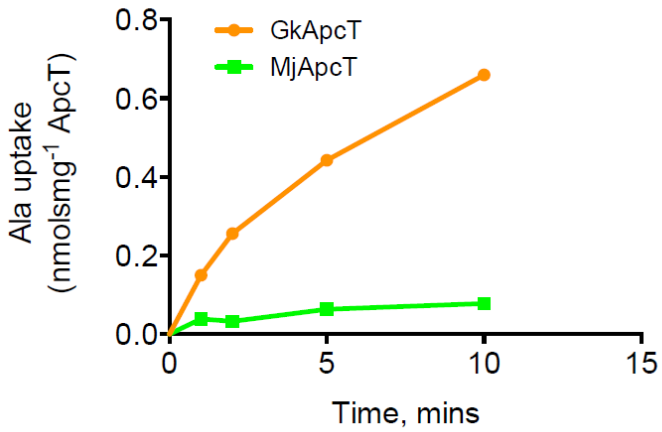


Figure 4-6: pH-dependent uptake of alanine by MjApcT and GkApcT

Proteolipsomes containing either *Mj*ApcT (green) or *Gk*ApcT (orange) loaded with internal buffer (25 mM citrate phosphate pH 7.0, 100 mM NaCl, 2 mM $MgSO_4$) were diluted 1:25 into external buffer (25 mM citrate phosphate pH 4.0, 100 mM NaCl, 2 mM $MgSO_4$, 500 nM ^{3}H-alanine). Data is plotted as uptake of alanine over time. 5 µg of protein were used per experimental time point.

4.4 The substrate specificity of *Gk*ApcT

Since it has been established that *Gk*ApcT is able to transport alanine under pH-dependent conditions it is of interest what other amino acids might be recognised by the transporter and how the protein compares to the substrate profile of *Mj*ApcT. One option would have been the repetition of the pH-dependent uptake assays with other radiolabelled amino acids. This would require 20 radiolabelled compounds to be tested. A second option was to perform counterflow assays. For this it is necessary to determine if an exchange between the unlabelled alanine and its tritiated counterpart is observed in the absence of an ion or proton gradient but driven by the difference in substrate concentration in the internal and external buffer. If *Gk*ApcT, for example, performed the bidirectional transport of alanine under substrate driven conditions, then the same could be tested for the other amino acids, by loading the proteoliposomes with an excess of other compounds. If the compound being tested resulted in a measured uptake of ^{3}H-Ala then it was a substrate (Kaback *et al.*, 2001). Initial time-course measurements were performed to validate that *Gk*ApcT shows counterflow activity under the used conditions. The experiment revealed that *Gk*ApcT showed substrate-driven uptake of alanine into the liposomes, indicating that the proposed buffer compositions were suitable for sustaining transport (Figure 4-7).

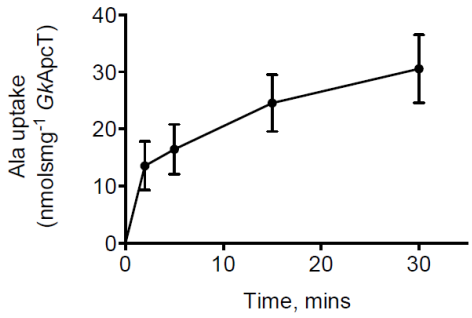


Figure 4-7: Time-course measurement of alanine uptake in counterflow conditions

*Gk*ApcT in POPE:POPG(3:1) liposomes was used for counterflow experiments. The internal and externa buffer conditions were identical (25 mM citrate phosphate pH 6.0, 100 mM NaCl, 2 mM $MgSO_4$) except for using 2 mM L-alanine in the internal buffer and 250 nM ^{3}H-alanine and 100 µM alanine in the external buffer. The uptake of alanine was followed over 30 min using 5 µg of protein per experimental time points. Mean values and their SEMs per experimental time point of three independent experimets are shown.

The substrate profile of *Gk*ApcT was determined under substrate-driven conditions using 2 mM of the substrates listed in Figure 4-8 in the internal buffer. Based on the results from these experiments, *Gk*ApcT can be classified as a transporter for small neutral and polar amino acids (Figure 4-8). Surprisingly, although *Gk*ApcT is closely related to the CAT transporters, the substrate specificity differs.

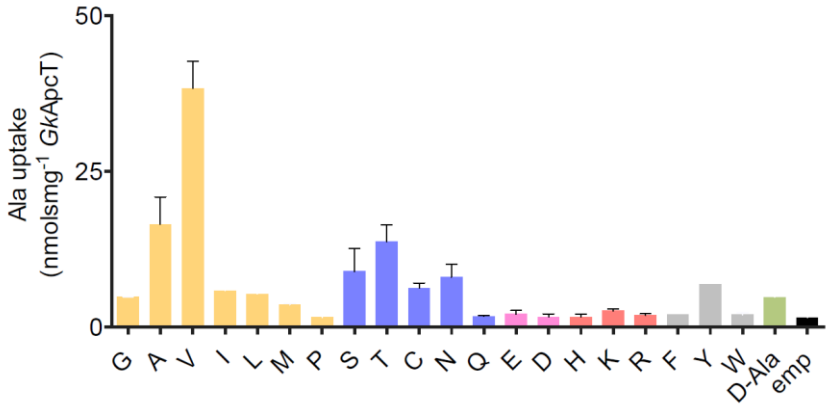


Figure 4-8: Substrate specificity of GkApcT

To determine the substrate profile for *Gk*ApcT, counterflow experiments with the internal buffer (25 mM citrate phosphate pH 6.0, 100 mM NaCl and 2 mM $MgSO_4$) containing an excess of the compound of choice at 2 mM concentration, were carried out. Proteoliposomes were diluted 1:25 into buffer containing 25 mM citrate phosphate pH 6.0, 100 mM NaCl, 2 mM $MgSO_4$, 100 µM alanine and 250 nM ^{3}H-alanine. Uptake after 5 min. is plotted for each compound. The compounds tested are shown as one-letter codes for the individual amino acids and the D-isomer of alanine (D-Ala). As negative control, the signal for liposomes containing no protein (emp). The compounds are grouped by their chemical properties: yellow – neutral, purple – polar, pink –acidic, red- basic, grey – aromatic, green – D-isomer, black – no protein in liposomes.

4.5 Is *Gk*ApcT a proton-coupled transporter?

As mentioned in the introduction of this thesis, based on the pH-dependent uptake of alanine into the liposomes, and the presence of the charged residue Lys158 in the transmembrane region of *Mj*ApcT, the transporter was proposed to be proton-coupled, although further evidence was not presented (Shaffer *et al.*, 2009). Therefore, *Gk*ApcT and *Mj*ApcT were investigated under electrogenic ($\Delta\Psi$-driven) conditions to verify this postulate. The conditions used for these experiments were changed to the buffer compositions used for studies on the proton-coupling mechanisms of the peptide transporter $PepT_{St}$ (Solcan *et al.*, 2012). For this, liposomes were loaded with potassium buffer (20 mM KPi pH 6.5, 100 mM KAc, 2 mM $MgSO_4$), and diluted 1:25 into buffer with equal numbers of sodium ions (120 mM NaPi pH 6.5, 2 mM $MgSO_4$). Upon addition of the potassium-carrier valinomycin to a final concentration of 10 µM (Figure 4-9), a negative inward membrane potential is generated and additionally a pH gradient through acetate diffusion gradients (Clerc and Barenholz, 1995). Using the equation (Viitanen *et al.*, 1986):

$$\Delta\Psi = -59 \log\left(\frac{[K^+]_{int}}{[K^+]_{ext}}\right)$$

With $[K^+]_{int}$ and $[K^+]_{ext}$ being the internal and external potassium concentration of the liposomes, respectively. The voltage of the membrane potential can be calculated. With an internal potassium concentration of 120 mM and a dilution of the liposomes of 1:25 into the external buffer, a membrane potential of -82.5 mV is generated.

A secondary active transporter harnesses the energy of the experimentally applied negative membrane potential, here -82.5 mV, to drive the ion^+-coupled uptake of the substrate into the liposome. To determine the coupling mechanism, four different experiments are performed. The first is to monitor alanine uptake in the presence of valinomycin which generates the negative inward potential. Assuming the transporter is proton-coupled, it would become protonated due to the positive charge build-up in the external buffer, which would initiate the transport, in this case, of alanine. The experiment is repeated but without the addition of valinomycin, which means that no membrane potential is generated and transport of alanine should be decreased under these conditions.

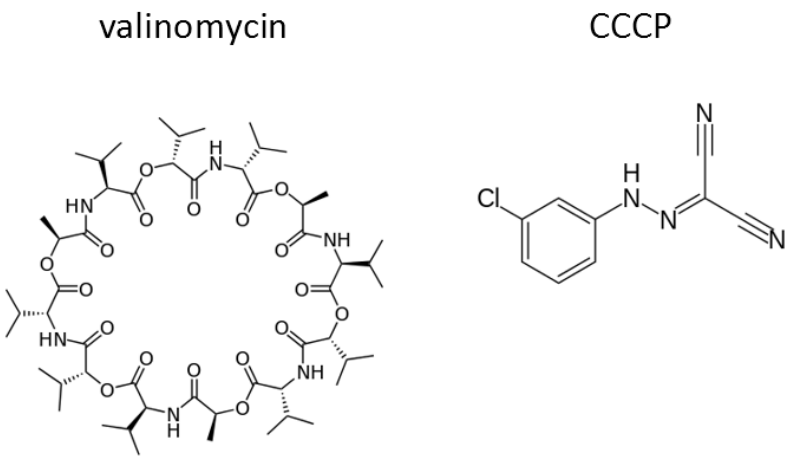


Figure 4-9: Chemical structures of valinomycin and CCCP

Valinomycin is a potassium carrier which can cross the lipid bilayer. Transport of the potassium ion is facilitated by coordinating it in the center of its ring structure, which can then be released on the other side of the lipid bilayer. CCCP works in a similar way but acts as a protonophore (proton-carrier). The protonated form (shown here) will cross the membrane towards the negatively charged or alkaline side, whereas the deprotonated form moves towards the positively charged or acidic side of the membrane. CCCP can thus dissipate the membrane potential or a pH gradient.

The third assay comprises the dilution of the liposomes into the same buffer that is on the inside of the membrane vesicle to establish if the observed transport activity was caused by the applied membrane potential. If the transporter is proton-coupled, no or little alanine uptake would be observed. The fourth experiment includes the addition of CCCP (Figure 4-9), an ionophore for protons, which dissipates any membrane potential created by valinomycin through proton-influx, and thus no alanine uptake dependent on proton-coupled transport would be detected.

Initial experiments with the addition of valinomycin were performed to identify if *Gk*ApcT and *Mj*ApcT show any alanine transport under these chosen buffer conditions. Again, *Gk*ApcT showed the highest uptake activity of the two proteins (Figure 4-10A & B) whereas no activity for *Mj*ApcT was detected, which did not come as a surprise since the of ^{3}H-Ala was entering the liposomes but also exiting them after 20 min in the experiments conducted by Shaffer *et. al*.

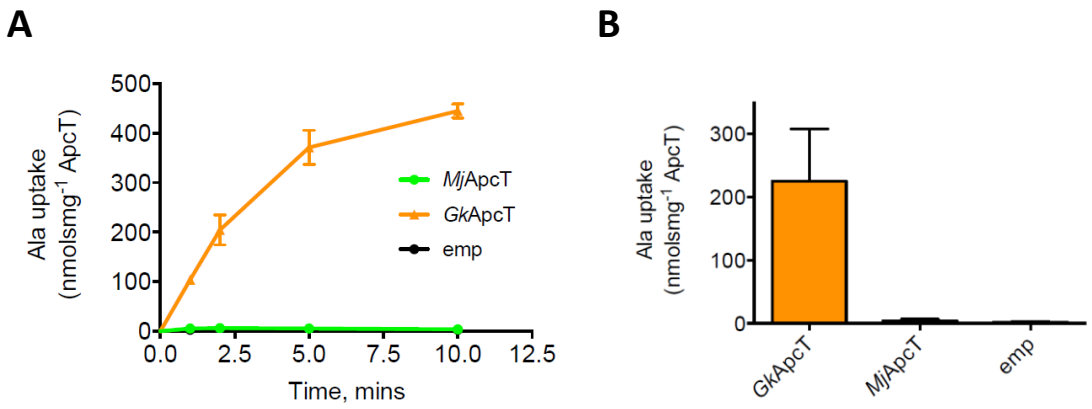


Figure 4-10: Electrogenic uptake of alanine by MjApcT and GkApcT

Time-course measurements of alanine uptake were performed on *Mj*ApcT and *Gk*ApcT under electrogenic conditions (inside buffer: 20 mM KPi pH 6.5, 100 mM KAc, 2 mM $MgSO_4$; external buffer: 120 mM NaPi pH 6.5, 2 mM $MgSO_4$, 100 µM alanine, 250 nM ^{3}H-alanine) (A). easier visualisation, the amount of alanine transported after 2 min is shown in the bar diagram (B).

Therefore, *Gk*ApcT was further investigated for the possibility that it is a proton-coupled transporter, by performing the control experiments as described previously. The time-course measurements show that only upon addition of valinomycin, increased alanine uptake is detected (Figure 4-11A). For simplicity, the values after one minute are plotted as a bar graph, showing the difference in uptake for each individual assay (Figure 4-11B). To rule out the possibility that *Gk*ApcT might be sodium-coupled, an experiment in the presence of valinomycin was performed with external buffer containing either sodium chloride or choline chloride (Figure 4-11B); the latter would not be able to act as a positive coupling ion. There is no difference in uptake between the two external buffer conditions, verifying that *Gk*ApcT is a proton-coupled amino acid transporter. In addition, pyranine assays were performed under high alanine concentrations. Upon addition of the substrate, proton-influx is detected (Figure 4-12), but not for liposomes devoid of protein.

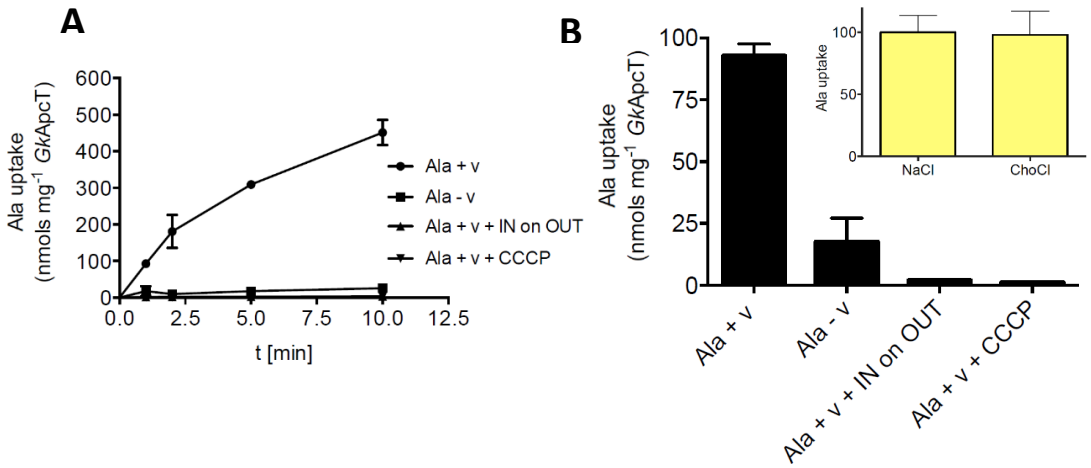


Figure 4-11: Electrogenic alanine uptake of GkApcT

Proteoliposomes containing 20 mM KPi, 100 mM KAc, 2 mM $MgSO_4$ were diluted 1:25 into external buffer (120 mM NaPi, 2 mM $MgSO_4$ and a mixture of 100 µM alanine and 250 nM ^{3}H-alanine) (A): in the presence of 10 µM valinomycin (Ala+v), in the absence of valinomycin (Ala-v), with 10 µM valinomycin and 10 µM CCCP (Ala+v+CCCP) and with the same potassium concentration in the internal and external buffer (Ala+v IN on OUT). Uptake values after 1 min for each experiment were plotted and additionally uptake of alanine after 1 min is plotted in the presence of either NaCl or ChoCl in the external buffer (B).

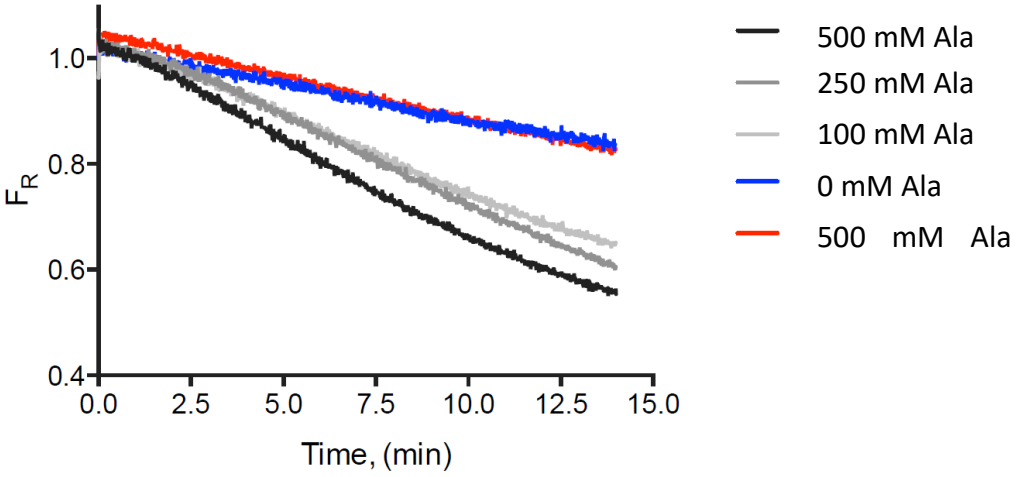


Figure 4-12: Pyranine assay monitoring substrate-driven proton-influx

Proteoliposomes containing *Gk*ApcT or no protein (e) were loaded with 5 mM HEPES pH 6.8, 120 mM KCl, 2 mM $MgSO_4$ and 1mM pyranine. Liposomes were then diluted into 5 mM HEPES pH 6.8, 120 mM NaCl, 2 mM $MgSO_4$ and varying concentrations of alanine. The relative fluorescence ($F_R = \lambda_{460\,nm}$ *-signal* / $\lambda_{415\,nm}$ *-signal)* of pyranine was monitored over time.

4.5.1 pH-dependence of alanine uptake

Since the uptake of alanine by *Gk*ApcT under substrate-driven and proton-driven conditions were now established, it was next investigated whether the experimental parameters were

optimum. Therefore, the pH profiles for counterflow and electrogenic uptake experiments were determined. Single measurements were performed for a given range of pHs (Figure 4-13).

Alanine uptake under electrogenic conditions after one minute was comparable for pH 5.5-7 but dropped at pH 7.5 (Figure 4-13A). The optimal pH for electrogenic alanine uptake was determined to be 6.5. A more pronounced pH profile was obtained under substrate driven conditions (Figure 4-13B), with an optimal pH of 6.0 and slightly less uptake at pH 5, but reduced outside this range.

An explanation why the pH-profiles were so different under the two experimental conditions originates from the transport mechanism. In electrogenic uptake, the generated negative inward membrane potential induces protonation of, the yet to be determined, site for proton-coupling, thus initiating the transport cycle. This mechanism seems to be independent of pH for *Gk*ApcT. However, substrate-driven conditions (counterflow) showed a distinct pH-profile, since the proton-coupling site of *Gk*ApcT needs to be protonated for the transporter to perform the exchange of unlabelled alanine with labelled alanine. The pH-optimum indicates that the pKa of that functional group is between 5.5 and 6.0.

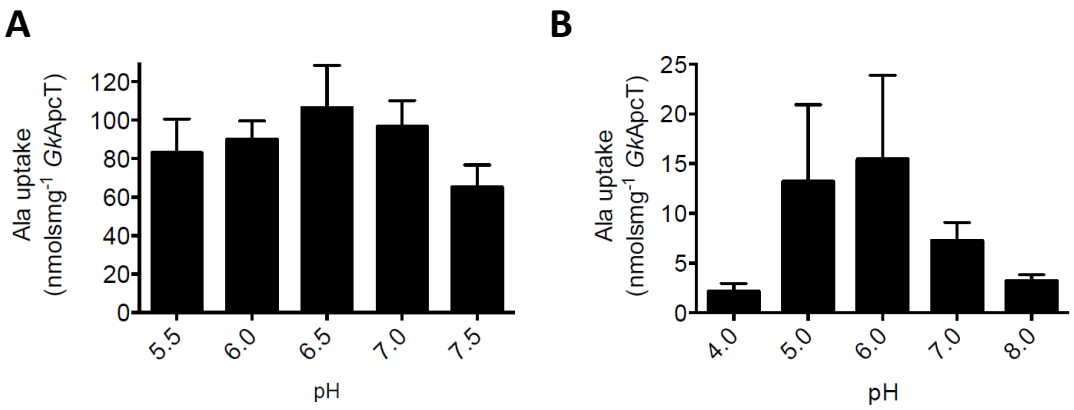


Figure 4-13: pH profile of alanine uptake

The pH profile of alanine uptake under electrogenic (A) and counterflow (B) conditions was determined for *Gk*ApcT. Proteoliposomes were loaded with KPi buffer at the given pH and diluted 1:25 into NaPi buffer of the same pH and a mixture of 100 µM alanine and 250 nM ^{3}H-alanine for the electrogenic uptake and for counterflow citrate phosphate buffer at the shown pHs were used. The uptake after 1 min (A) and 5 min (B) at 25°C are plotted.

4.6 TRANSPORT KINETICS OF ALANINE UPTAKE

Kinetic measurements of the uptake of alanine were carried out. For this, the protein concentration was reduced to 1.25 µg per experimental time point to obtain a linear uptake graph to be able determine the initial uptake rate (Figure A 9). Measurements under electrogenic conditions were performed with varying concentrations of alanine in the external buffer. The initial uptake velocities were plotted against the concentrations of alanine and the resulting graph was analysed using the Michaelis-Menten equation:

$$v = \frac{v_{max}[S]}{K_M + [S]}$$

where *v* is the rate, v_{max} the maximal rate, K_M the Michaelis-Menten constant, which refers to the substrate concentration *[S]* where v_{max} is half maximum.

Based on the kinetic measurements, a v_{max} of 2.38 nmols sec^{-1} mg^{-1} (0.12 alanine per sec. per transporter) and an apparent K_M value for alanine of 89 µM were determined. The apparent K_M is in accordance with the substrate specificities of the human CAT1, CAT2B and CAT3 that had measured K_M values for arginine in the range of 38-450 µM (Fotiadis *et al.*, 2013). These transporters are known as high-affinity transporters, whereas CAT2A is a low affinity transporter with K_M values for arginine of 2-5 mM (Fotiadis *et al.*, 2013). *Gk*ApcT is in good agreement with the determined nanomolar to millimolar range of affinities for other APC members such as LeuT (K_M(Ala) = 583 ± 28nM) (Singh *et al.*, 2008), the mutant $LeuT^{K+I359Q}$ (K_M (Trp) = 78.9±8.8 µM)(Piscitelli and Gouaux, 2012), BetP (K_M (betaine) = 3.5±0.6 µM) (Perez *et al.*, 2011), and SteT (K_M (Ser) = 1.2 ± 0.2 mM) (Reig *et al.*, 2007).

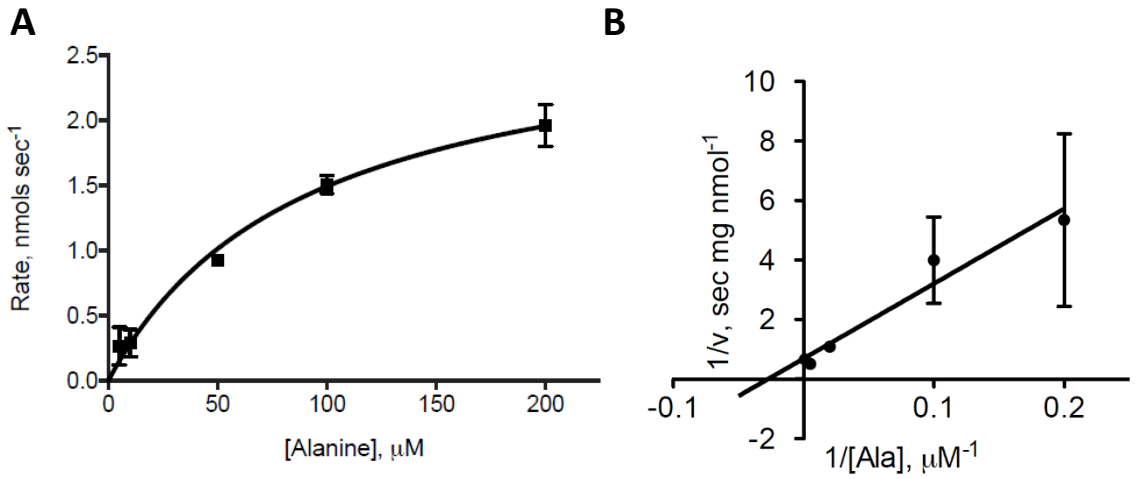


Figure 4-14: Kinetics of alanine uptake by GkApcT-WT

The transport kinetics of *Gk*ApcT for alanine transport were investigated under $\Delta\Psi$-driven conditions (A). Proteoliposomes containing internal buffer I (20 mM KPi pH 6.5, 100 mM KAc and 2 mM $MgSO_4$) were diluted 1:25 into external buffer I (20 mM NaPi pH 6.5, 100 mM NaCl, 2 mM Mg_2SO_4, 250 nM ^{3}H-alanine, 10 µM valinomycin and varying alanine concentrations). The initial uptake rates were plotted against the alanine concentrations used and kinetic parameters were determined in PRISM using the Michealis-Menten equation. The Lineweaver-Burk plot of the experimental data is shown as control (B).

4.7 COMPETITION WITH ALANINE UPTAKE

*Gk*ApcT was further subjected to competition assays to identify substrates able to compete with alanine binding and transport, verifying the results from the counterflow experiments but additionally identifying possible competitive inhibitors of *Gk*ApcT.

Competition assays were carried out under electrogenic conditions by adding 10 mM of unlabelled competitor to the external buffer. As a control, proline and glutamate were used (Figure 4-15) which did not induce transport in counterflow assays (Figure 4-7). Neither of the amino acids competed with alanine in the competition assays (Figure 4-15). Additionally, L-ornithine and γ-aminobutyric acid (GABA) were also tested for competition. L-ornithine is a known substrate for the SLC7 family (Kim *et al.*, 1991) whereas GABA had been identified as substrate for members of the proton-coupled SLC36 family (Thwaites *et al.*, 2000; Larsen *et al.*, 2008). At 10 mM of the two compounds, alanine uptake was reduced to 75 %, whereas the identified substrates of *Gk*ApcT, Ala, Thr and Val reduced the signal for alanine uptake close to zero.

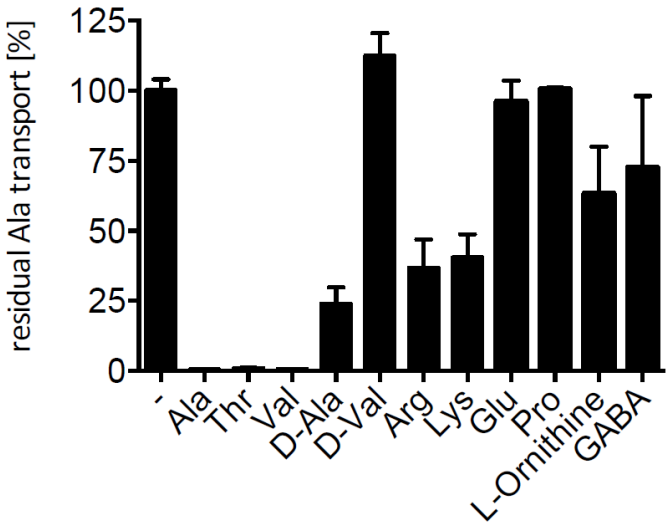


Figure 4-15: Competition experiments with GkApcT

Competition assays on *Gk*ApcT were carried out under *ΔΨ*-driven conditions. Proteiliposomes were diluted 1:25 into buffer containing 120 mM NaPi pH 6.6, 2 mM $MgSO_4$, 10 µM valinomycin and 250 nM ^{3}H-Ala with no competitor added (-) or in the presence of 10 mM competitor. The data shows the relative uptake of Ala after 1 min. to no competitor added.

Competition assays can give an initial indication of the range of compounds that can interact with the substrate binding site of the investigated transporter. A decreased signal implies that the compound is either a competitor or a potential inhibitor of the transporter. Therefore, together with the counterflow activity for a given compound the two possibilities can be distinguished from one another.

The assays conducted here helped identify alanine, threonine, valine, D-alanine, lysine and arginine as potential competitors (Figure 4-15). Alanine, threonine and valine were previously identified as potential substrates for *Gk*ApcT, but arginine, lysine and D-alanine did not induce high counterflow activity (Figure 4-8). Thus, D-alanine, lysine and arginine are potential inhibitors of *Gk*ApcT. Arginine and lysine might interact with the substrate binding site, but the cavity is sterically hindered for the transporter to fully close over these molecules, possibly locking the transporter in an open conformation. Therefore, the transporter cannot proceed with the transport cycle. Arginine is a likely candidate as an additive for crystallisation trials to obtain crystals in order to solve the outward-open conformation of *Gk*ApcT as previously described for LeuT (Singh *et al.*, 2008). The inhibitory effect of D-alanine indicates the stereospecificity of the substrate binding site. If the carboxyl group of the D-isomer of alanine interacts with TM1 similarly to that of L-alanine, then the amide group of D-alanine would point downwards. It would thus be unable to form interactions with the

water molecule and the backbone of TM6, and this might hinder the transport cycle. The competition assays performed only gave a crude indication of the affinity of the transporter to the investigated substrates and inhibitors. A more detailed approach is to determine the half maximum inhibitory concentration (IC_{50}) for each compound. This quantitative measurement enables a clear distinction to be made between the affinity of different inhibitors and substrates. Here, IC_{50} experiments were performed under electrogenic conditions with varying concentrations of the competitor (Figure A 10). The results are listed in Table 4-1. The lowest IC_{50} values, which indicated the highest affinity, were found for Ala (2.4 ± 0.6 µM) followed by Thr (16.1 ± 4.3 µM) and Val (23.5 ± 3.9 µM), which is consistent with the results from the counterflow experiments which indicated that these amino acids are substrates for *Gk*ApcT. D-Ala and Arg as potential inhibitors show higher IC_{50} values in the millimolar range, whereas L-ornithine and GABA are very poor inhibitors of *Gk*ApcT, indicating that these compounds barely form stable interaction with the substrate binding site.

Table 4-1: IC_{50} values for GkApcT

Competitor	***IC*$_{50}$**
L-Ala	2.4 ± 0.6 µM
D-Ala	954 ± 38 µM
L-Arg	6 ± 2.5 mM
L-Thr	16.1 ± 4.3 µM
L-Val	23.5 ± 3.9 µM
L-Lys	4.14 ± 1.5 µM
L-ornithine	26.2 ± 12.4 mM
GABA	31.4 ± 17.4 mM

4.8 CONCLUSION

The initial functional *in vitro* experiments presented here established the experimental conditions to study *Gk*ApcT's transport activity. First, two different lipid compositions were investigated for their proton permeability which revealed that the combination of POPE:POPG (3:1) was the best choice in maintaining a stable pH gradient. These experiments were necessary since *Gk*ApcT and its homologue *Mj*ApcT were investigated under

pH-dependent conditions using a pH gradient of three pH units to drive alanine uptake. These experiments revealed that *Gk*ApcT transports alanine and that the protein was more active than the archaeal *Mj*ApcT under the given experimental conditions. Given that the organism *Gk*ApcT originates from, *G. kaustophilus*, grows in environments with pHs 2-12, it is not surprising that the transporter utilises the pH gradient across the cellular membrane to drive uptake of nutrition.

Further determination of the substrate profile of *Gk*ApcT from counterflow experiments, identified the protein to be specific for small neutral and polar amino acids.

Alanine uptake measurements under electrogenic conditions, in combination with control experiments, identified *Gk*ApcT to be a proton-coupled amino acid transporter, which would make it the first APC family member which has been verified as using proton-symport for substrate translocation. As mentioned above, the use of protons for substrate transport is probably derived from the acidic environment that *G. kaustophilus* can grow in. The energy of the proton motive force across the cellular membrane is then used as the energy source for transport. Thus, the proton-coupling mechanism of *Gk*ApcT represents the adaption of this bacterial transporter to its surrounding environment. Consequently, *G. kaustophilus* must possess an acidic resistance pathway, that maintains the intracellular pH. No proton-coupling mechanism for *Mj*ApcT could be confirmed under the tested conditions, which strengthens the assumption that *Mj*ApcT might be a uniporter based on the experimental evidence given here and in previous studies.

Furthermore, it was possible to obtain the kinetic constants for alanine uptake by *Gk*ApcT and to determine possible inhibitors of *Gk*ApcT such as arginine and lysine, revealing that these substrates are able to interact with the substrate binding site. This makes them potential additives for future crystallisation of *Gk*ApcT to try and capture the protein in a different conformation.

5 Crystal structure of *Gk*ApcT

Structural data are important for the process of understanding a protein and its function, but obtaining this information is difficult. Testing different crystallisation techniques and extensive optimisation of the resulting conditions led to small crystals of *Gk*ApcT. Further optimisation resulted in well diffracting crystals grown in a lipid environment. The following chapter will focus on the path leading to the crystal structure of *Gk*ApcT and what insights the structure gave into the protein's function and mechanism.

5.1 Data collection

*Gk*ApcT crystals grew in a lipid mixture of monoolein and 10 % cholesterol. The LCP drop was surrounded by precipitant containing 34-38% PEG200, 0.1 M NaAc pH 4.0, 10 mM Ala, and 0.08-0.12 M potassium fluoride. The crystals were either diamond or rod shaped, with 20-30 µm length in the longest dimension and reached their full size after 2-4 weeks at 20 °C. No additional cryoprotectant was added prior to flash-cooling into liquid nitrogen. They diffracted to a resolution of 2.8-3.4 Å. A complete data set to 2.86 Å at 100 K from a single crystal was collected at the micro-focus beamline I24 at DLS using a 10 x 10 µm^2 beam after the crystal was localised in the cryo-loop using a grid scanning method (Figure 5-1).

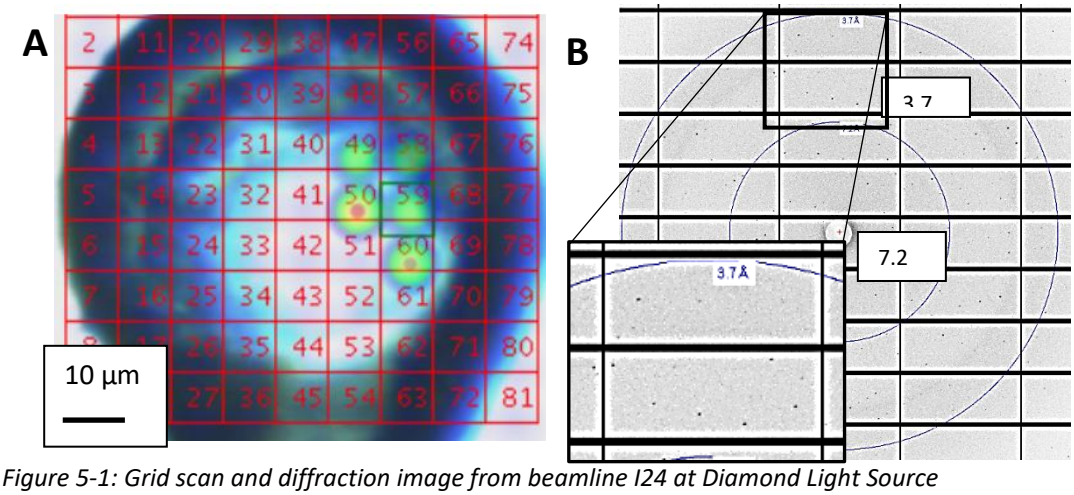


Figure 5-1: Grid scan and diffraction image from beamline I24 at Diamond Light Source

Once mounted, the visibility of LCP grown crystals is lost. Grid scans need to be performed to localise the crystal in the loop. At I24, a heat map is generated based on the observation of diffraction patterns in each position of the grid (10 x 10 µm^2 boxes) (A). The individual diffraction pattern for box 50 is shown in (B). The diffraction reached the edge of the detector.

The data set was processed with DIALS (Sauter *et al.*, 2013) (Table 5-1). The space group was $P2_12_12_1$ with the following unit cell dimensions: a=76.5Å, b=82.9Å, c=119.2Å, α=90.0°, β=90.0°, γ=90.0°. A Matthew's coefficient calculation suggested that there was one protein molecule per asymmetric unit (V_m = 3.77 A^3/Da).

Table 5-1: Merging statistics for data set collected from GkApcT crystal

	***Gk*ApcT-WT**
Beamline	I24, DLS
Detector	PILATUS 6M
Wavelength (Å)	0.9686
Temperature (K)	100
Osc. (°)/image	0.5
Exp. time (s)	0.1
Total number of images	360
Space group	$P2_12_12_1$
Unit cell a b c (Å)	76.9 83.15 119.74
α β γ (°)	90.0 90.0 90.0
Mosaicity (°)	0.23
Wilson B-factor (Å^2)	64.70
Resolution (Å)[1]	68.3 -2.86 (2.93-2.86)
R_{merge}	0.152 (1.789)
R_{meas}	0.172 (1.982)
R_{pim}	0.077 (0.842)
No. obs.	95402 (7084)
No. unique obs.	18347 (1319)
I/σ(I)	6.3 (1.1)
$CC_{1/2}$	0.834 (0.49)
Completeness (%)	99.9 (100)
Multiplicity	5.2 (5.4)

[1]Values for highest resolution shell are given in parentheses

5.2 PHASING AND REFINEMENT

5.2.1 Molecular replacement

The structure of *Mj*ApcT (PDB-ID: 3GIA, 2.32 Å resolution), homologous to *Gk*ApcT (~30% sequence identity to *Mj*ApcT) served as a molecular replacement (MR) template in PHASER. Initial attempts using the full structure or truncated versions, where flexible loops were deleted, did not lead to a solution. Therefore, a homology model of *Gk*ApcT based on the structure of *Mj*ApcT (PDB-ID: 3GIA, 2.32 Å) was used, generated with MODELLER (Webb and Sali, 2014) using a sequence alignment (Figure A 7) between the two proteins from CLUSTAL OMEGA (Sievers *et al.*, 2011). The models consisted of either the full structure, or the transmembrane domain without any loop regions. Again, attempts to obtain a MR solution did not succeed. Another strategy was the use of a *Gk*ApcT model generated in CHAINSAW (Stein, 2008). The program takes the alignment (Figure A 7) of *Gk*ApcT as the target and *Mj*ApcT as the template protein, and generates a homology model of which the non-conserved residues are pruned, dismissing any atoms that are not common to both target and template residues. The CHAINSAW model of *Gk*ApcT was used as input for PHASER. This time, a possible solution was found. Initial analysis of the calculated electron density did not show any obvious misplacement of the *Gk*ApcT protein, such as clashes between symmetry related molecules or parts of the protein that did not fit into the electron density. Surprisingly, the initial map revealed additional density at the N-terminal region (Figure 5-2), suggesting that the obtained phases via MR were correct and not biased by the initial search model.

Following inspection of the sequence alignment of *Mj*ApcT with *Gk*ApcT, it was suggested that the observed N-terminal extension might be an additional TM of *Gk*ApcT. However, the overall quality of the electron density map made it difficult to identify particular side chains. To overcome this obstacle, obtaining additional phase information from experimental phasing, or at least an indication where particular residues are located was crucial. Different approaches were followed in attempts to successfully build the crystal structure into the density.

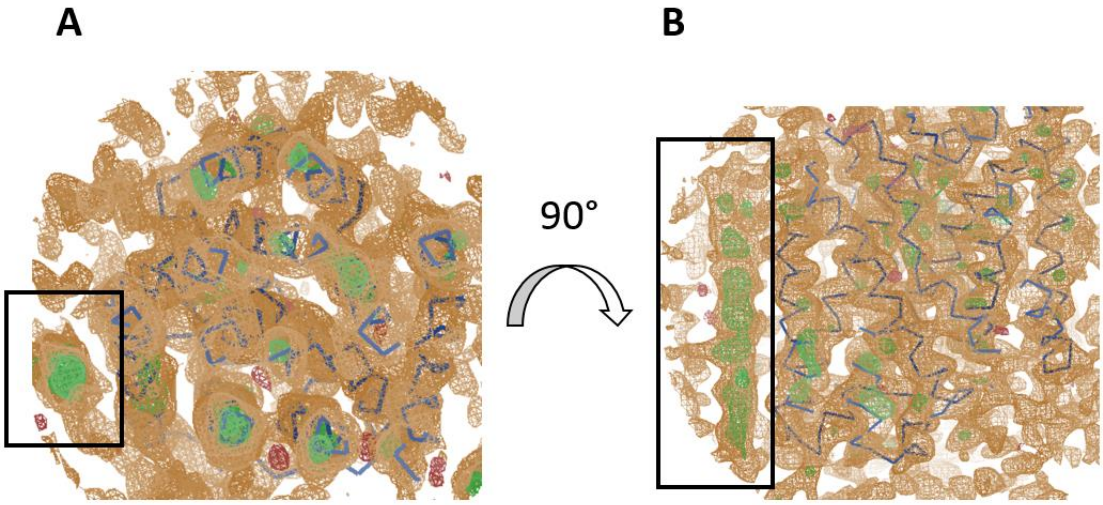


Figure 5-2: Additional density found close to the GkApcT structure

After molecular replacement, additional positive difference density (green, 3σ) was found adjacent to *Gk*ApcT's helical bundle (blue). The $2F_O$-F_C map is shown in orange contoured at 1σ.

5.2.2 SAD phasing using heavy atoms and SeMet incorporation

During the diffraction data collection process, a small proportion of the X-rays hitting the crystal are reflected off the planes formed by the regular array of the atoms in the crystal. The scattered waves are recorded during the experiment as distinct spots on the detector. The intensity of those spots, also known as reflections, contain information on the three-dimensional arrangement of the atoms in the crystal, as every atom contributes to the overall scattering factor of the crystal. The square root of the intensity for a given reflection hkl is proportional to the amplitude of the wave scattered off from the plane *hkl*. To calculate the electron density from the recorded intensity data collected from a crystal, the phase of this scattered wave is necessary, but is lost during the experiment. This is known as the phase problem.

Different methods have been introduced to solve the phase problem in order to be able to obtain a crystal structure of a given protein. For these methods to be used, normally prior structural knowledge is necessary. For instance, in MR, a homologous structure is used to provide initial phase information. The structure factor amplitudes from the collected data and the MR model are used to generate a Patterson map for each of the two. First, a rotation function is applied to the MR model Patterson map to find the correct orientation of

the MR model in the unit cell of the unknown structure. The second step is the application of a translation function to identify the correct position of the molecule. This is guided by iterative comparison of the Patterson maps of the known and unknown structures during rotation and translation of the known structure. However, if no or only poor models are available for phasing, marker atoms can be used to utilise prior structural information. In this project, the incorporation of heavy atoms via soaking into the protein crystal and pre-incubation of the protein with heavy atoms, followed by crystallisation of the derivative protein, were chosen as alternative methods to MR. Additionally, expression of the *Gk*ApcT-SeMet derivative, which is a well-established method for pre-derivatisation of bacterial expressed proteins, was tried. Both approaches can provide phase information via single-wavelength anomalous dispersion (SAD) experiments. The principle that underlies these phase determination methods is that the heavy atoms introduce changes in the intensity of the Bijvoet pairs ((hkl) and (-h-k-l)) that would normally share the same intensity value. The intensity difference, subsequently changing the amplitude of the structure factors of the Bijvoet pair, makes it possible to deduce the location of the heavy atoms (using the Patterson map) in the structure and the phases of the heavy atom substructure, φ_H. The phases of the native protein structure can be estimated using the following equation:

$$\varphi_P = \varphi_H \pm cos^{-1}[(F_{PH}^2 - F_P^2 - F_H^2)/2F_P F_H]$$

where F_P is the structure factor amplitude of a particular reflection for the native protein, F_{PH} is the structure factor amplitude of the same reflection for the derivative protein structure, and F_H is the structure factor amplitude of that reflection for the heavy atom substructure.

5.2.2.1 Heavy atom soaking

The binding of heavy atoms to the reactive sulphur group of free cysteines in the protein of interest is one option to try for solving the crystal structure of this protein via the SAD method. Most commonly, the protein crystal is soaked in its precipitant solution containing additionally a salt of the heavy atom of choice for a defined period, and then harvested, cryo-cooled and stored in liquid nitrogen until data collection. In this case, cryo-protection

was not necessary, since the lipid is already a cryo-protectant. The concentration of heavy atom added and the soak time for the crystal should be varied over different scales and should be determined experimentally. If no free cysteines are present in the protein sequence, then they can be introduced via mutation of the corresponding gene (Nagai *et al.*, 1990), keeping in mind however that the internal structure of the protein and possible crystal contacts might be disturbed.

*Gk*ApcT contains four cysteines that, based on the generated homology model, do not form any disulphide-bridges. Preliminary soaking experiments were performed directly on the crystals grown in LCP. For this, the LCP plate was opened carefully to expose an individual well. The remaining precipitant was wicked away and 2 µl of the precipitant condition containing 100 µM or 1 mM of the heavy atom compound ($Hg(Ac)_2$, $HgCl_2$, AuCN) was added on top of the LCP drop. The well was closed with a glass slide and after a few hours or up to a day of incubation, the crystals were harvested. Most of the crystals survived this procedure but due to contact with air, it was assumed that the crystal quality might suffer.

Harvested crystals were directly cryo-cooled, without adding any additional cryo-protectant, and stored in liquid nitrogen. The crystals were screened at the DLS beamline I24 with X-rays of wavelengths determined by fluorescent scans on the crystal which corresponded to the L_{III} edge of relevant heavy atoms (Table 5-2). However, no protein diffraction was observed when using mercury or gold, in which the crystals had been soaked overnight. The same effect was observed for soaks that lasted for three hours. It was uncertain as to whether the crystals had disappeared during the harvesting process or were harvested into the cryo-loops. During the harvesting process, the crystals are visualised using cross-polarised light but they occasionally move within the lipid environment making it difficult to be sure that the crystal was harvested into the cryo-loop. Additionally, it is not possible to check if the crystal was mounted since it becomes invisible during this process. There was one sample where the LCP was so thin that the outline of the crystal in the loop could be made out, but this was a rare occasion. A grid scan on this area did not show any diffraction, suggesting that the heavy atom had bound to the protein but might have disturbed the crystal lattice. It should be noted that for LCP grown crystals it is not possible to remove any residual heavy atoms from the precipitant solution, making these crystals more radiation sensitive. This is because of the increased X-ray absorption coefficient of the heavy atoms, adding

to the energy absorbed by the sample and thus increasing the dose and decreasing the crystal lifetime for data collection (Murray *et al.*, 2004).

Table 5-2: X-ray absorption edges for the heavy atoms used

Heavy atom	L_{III} Absorption edge (keV)
Au	11.92
Hg	12.28
Pt	11.56

Simultaneously with the direct soaking of the crystals, pre-derivatisation was performed using purified *Gk*ApcT incubated with the heavy atom of choice, followed by crystallisation trials in LCP. As a control experiment, 1.97 µM purified *Gk*ApcT was incubated with a three-times molar excess of heavy atom solution for three hours and loaded onto a S200 5/150 column to identify suitable candidates for pre-derivatisation experiments (Figure 5-3). None of the heavy atoms used resulted in destabilisation of the protein: the UV trace of the protein is wildtype-like and no precipitation was observed.

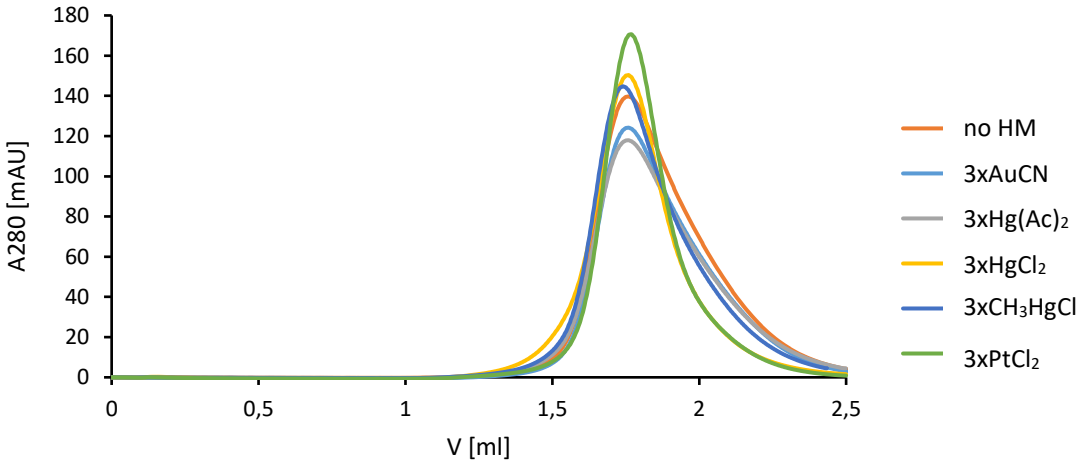


Figure 5-3: Size exclusion chromatography of heavy metal derivatives of GkApcT

Size exclusion profiles of *Gk*ApcT pre-derivatised with various heavy atom compounds at three-fold molar excess. HM- heavy metal compound, AuCN – gold cyanide, $Hg(Ac)_2$- $Hg(O_2CCH_3)_2$ mercury acetate, $HgCl_2$ – mercury chloride, CH_3HgCl- methyl mercury chloride, $PtCl_2$ – platinum chloride

10 mg/ml *Gk*ApcT was soaked for 3 h on ice with a three-fold molar excess of heavy atoms and then reconstituted into LCP. For crystal trials, both the original and optimisation screens were used in case the derivatised protein did not crystallise in the same condition as the wildtype. Rescreening did not lead to crystals of the derivative protein.

An alternative approach was to introduce selenomethionine (SeMet) instead of the natural methionine into the protein. *Gk*ApcT has nine methionines that could be replaced by this method.

5.2.2.2 SAD using SeMet incorporation

To replace SeMet with Met, the cell expression media is supplemented with an excess of SeMet. For this, PASM5052 minimal media was used, containing only little amounts (10 mg/L) of methionine. An excess of SeMet was added to the media to a final concentration of 100 mg/L. The cells were grown to an OD of 0.6-0.8, induced with 400 µM IPTG and kept shaking overnight at 25 °C.

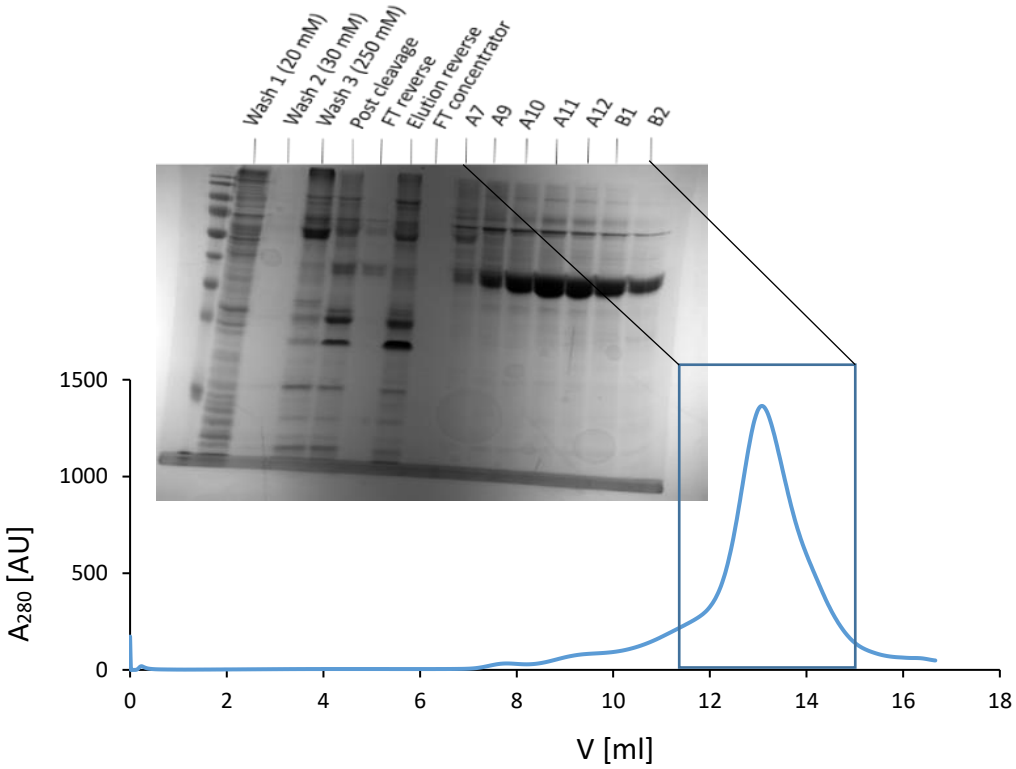


Figure 5-4: Purification of GkApcT-SeMet

SDS gel analysis shows the purification steps of *Gk*ApcT-SeMet derivative protein. Lanes are marked with the individual steps. Markings A7-B2 correspond to elution volumes 12-14.5 ml on the gel filtration profile (blue).

The absolute yield of wet cell pellet and protein was half of that compared with when using TB, which is not unusual given that SeMet is toxic for the cells. During the first purification, the protein did not solubilise properly in DDM and the binding efficiency to the nickel resin was very poor, ~30 % after 4 h vs. 60-70% for WT. In addition, re-expressing the protein did not lead to reproducible results. It was reasoned that the cells might be under stress during growth and expression, due to the SeMet being available from the beginning of the inoculation of the starting culture. To reduce the stress on the cells, SeMet was added just before induction. This led to similar yields as for previous expression trials for SeMet derivatisation, which was nearly half of the expression in TB medium, but the protein could be purified and was more stable during this process, although more contaminants were found in the purified sample (Figure 5-4). Crystal trials in LCP were performed using standard screens and the optimisation screens from previous trials, but no crystals were observed. Additionally, the concentration of the protein was increased to 25 mg/ml for crystallisation at 4 °C but again these crystal trials were unsuccessful.

5.2.3 Sulphur SAD

Instead of introducing anomalous scatterers into a protein, the anomalous signal of the naturally occurring sulphur can be used. The disadvantage is that sulphur is a weak scatterer when using wavelengths (1.5 – 1.9 Å) at conventional MX synchrotron beamlines. To obtain a reasonable anomalous signal from sulphur, data collection near its absorption edge at 5 Å (Figure 5-5) is recommended. However, this poses a problem, since at longer wavelengths the absorption of X-rays by the sample and air also increases significantly, making data collection at such wavelengths unfeasible. Instead, longer wavelengths (>2 Å) at dedicated beamlines can be used (Wagner *et al.*, 2016). At these long wavelengths, more of the X-ray beam is absorbed by the air in the sample chamber and also the anomalous signal of the sulphur is lower than that at its absorption edge, but still higher than at wavelengths lower than 2 Å.

To overcome the first problem, either a helium (Photon Factory BL1A, (Liebschner *et al.*, 2016)) or vacuum (DLS I23, (Wagner *et al.*, 2016)) sample chamber is used to reduce the

absorption of the X-ray beam. To accurately detect the low anomalous signal from the sulphur, data with high multiplicity must be measured. The beamline I23 at DLS and BL1A at the Photon Factory, Japan were both available for data collection on *Gk*ApcT crystals.

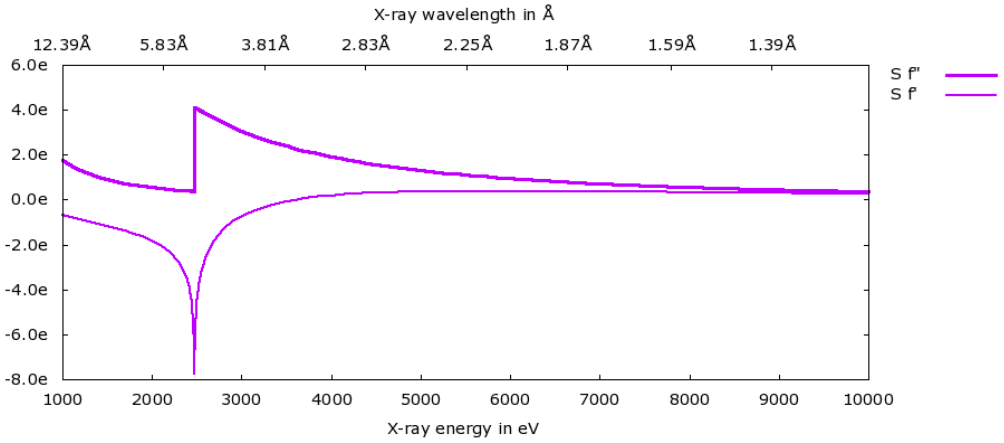


Figure 5-5: X-ray scattering plot for sulphur

The absorption edge for sulphur is at 2.4720 keV (5.01 Å). The values of f' (dispersive term) and *f*'' (absorption term), representing the anomalous scattering factor, change with the wavelength. The absorption term f'' increases at wavelengths above 2 Å, and is maximum at the *K* absorption edge of 5 Å.

At beamline I23, a curved PILATUS 12M and mini kappa goniometer reside in vacuum. The sample holder has to be cooled conductively under vacuum before the sample can be transferred onto the goniometer head. This step takes around 40 min until completion. The available loop size was 60 µm, double the size of the crystals. This results in more LCP being mounted which could increase the background under the reflection data. The beam size for data collection was 50 x 50 µm^2. BL1A has a different sample environment. There, the Eiger4M detector is inside a helium chamber to reduce the absorption of the X-rays in air. Data were collected at wavelengths of 2.75 Å (I23, square-shaped collimator with a 200 µm aperture, 50 x 50 µm2 FWHM Gaussian profile beam) and 2.7 Å (BL1A, capillary-shaped collimator with 150 µm aperture, 10 x 10 µm2 FWHM Gaussian profile beam) with an inverse beam geometry. The data were processed in XDS, followed by preparation of the files with SHELXC (Sheldrick, 2008) and generation of an anomalous difference Fourier map with ANODE

(Thorn and Sheldrick, 2011). Anomalous signal for the data collected at I23 and BL1-A was observed to resolution ranges of 4.79-5.16 Å (Table A 2) and 3.22-3.5 Å (
Table A 3), respectively, but was not sufficient for *ab initio* phasing. However the calculated anomalous difference Fourier map (Figure 5-6) in combination with the previously obtained MR (see Section 5.2.1) was used to identify anomalous difference peaks corresponding to the sulphur atoms that could guide the model building process, ensuring the correct register of the protein sequence could be modelled.

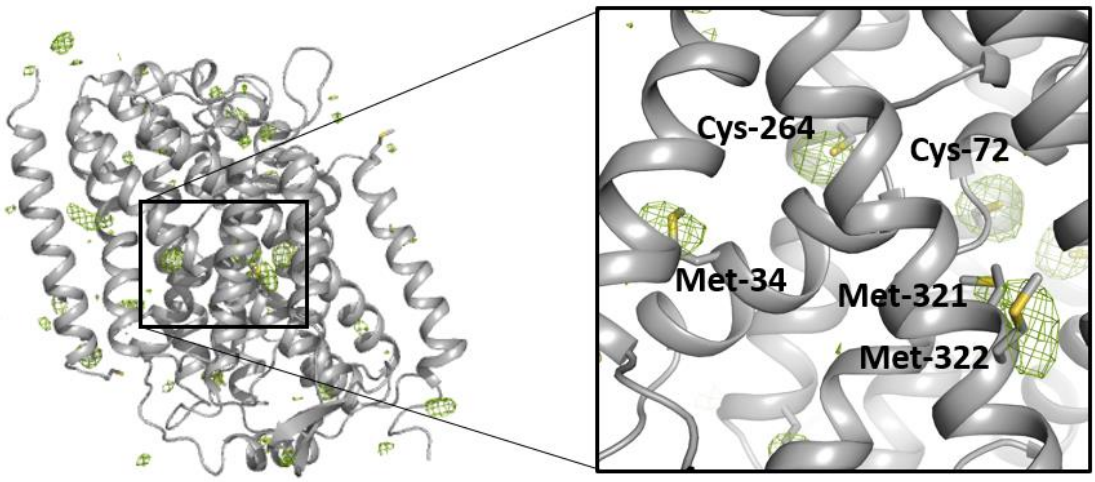


Figure 5-6: Anomalous difference Fourier electron density

Anomalous difference Fourier map, contoured at 3.0 σ (green mesh), for sulphur atoms in the asymmetric unit. Difference peaks can be clearly seen around two methionine residues of *Gk*ApcT that were used as guide points to build the protein sequence into the electron density map obtained by MR.

Two peaks on one helix could be made out, corresponding to Met321 and Met322 as predicted by the homology model but now with verified positions. Thus, data collection on beamlines I23 and Bl1A was beneficial for this project, as it enabled the correct placement of some residues which were used as guide points for model building and led to the successful solution of the crystal structure of *Gk*ApcT (Table 5-4). Model building into the $2F_o$-F_c map and corrections to the model based on the F_o-F_c map in COOT (Emsley *et al.*, 2010) was followed by refinement against the crystallographic data using BUSTER, using individual B-factor, and translation-libration-screw (TLS) refinement (Bricogne *et al.*, 2016). After each refinement cycle, the R_{work} and R_{free} values were analysed on their performance. If both

or only the R_{free} increased, then the previous model building step did not add any meaningful information. In this case, the model used for the refinement was carefully inspected in COOT together with the $2F_o$-F_c and F_o-F_c maps, adjusted and refined again. If these adjustments did not result in a decrease of the R-values, then this model was not used any further. Instead, the building/refinement process proceeded using a model from the previous refinement stages. In the final steps of model building and refinement, ligands and later water molecules were added (see Figure A 8 for final refinement step). In total over 300 refinement and modelling steps were performed. MOLPROBITY (Chen *et al.*, 2010) was used to validate the final model.

Table 5-3: Data collection and merging statistics for the long wavelength data collection

Beamline	**BL1A**	**I23**
Detector	EIGER4M	PILATUS 12M
Wavelength (Å)	2.7	2.75
Temperature (K)	100	100
Beam size (µm x µm FWHM)	10 x 10	50 x 50
Osc. (°)	0.2	0.1
Exp. time (s)	0.2	0.1
Sweeps	1 2	1 2
Number of images	1800 1800	1800 1800
Mosaicity (°)	0.22 0.20	0.12 0.13
Wilson B-factor (Å²)	46.7	118.8
Unit cell		
a b c (Å)	77.19, 83.41, 119.64	75.97, 82.55, 119.67
α ß γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Res. limit (Å)[1]	48.28-3.3 (3.56-3.3)	59.84-4.0 (4.47-4.0)
R_{merge}	0.3468 (3.179)	0.176(1.525)
R_{meas}	0.3542 (3.239)	0.186 (1.61)
R_{pim}	0.079 (0.856)	0.056 (0.5)
No. obs.	295781 (62962)	69717 (16454)
No. unique obs.	11734 (2328)	6512 (1711)
I/σ(I)	11.1 (1.1)	11.4 (1.5)
$CC_{1/2}$	0.996 (0.458)	0.999 (0.654)
Compl. (%)	98.9 (97.7)	96.8 (92)
Multiplicity	20.1 (21.4)	10.7 (9.6)
Anomalous compl. (%)	99.0 (97.8)	94.9 (87.8)
Anomalous multiplicity	10.9 (11.3)	5.7 (5.2)
Mid-slope of anomalou propablity[2]	1.112	1.017

[1]Values for highest resolution shell are given in parentheses
[2] Values close to one indicate strong anomalous signal

Table 5-4: Refinement statistics for the GkApcT structure solution

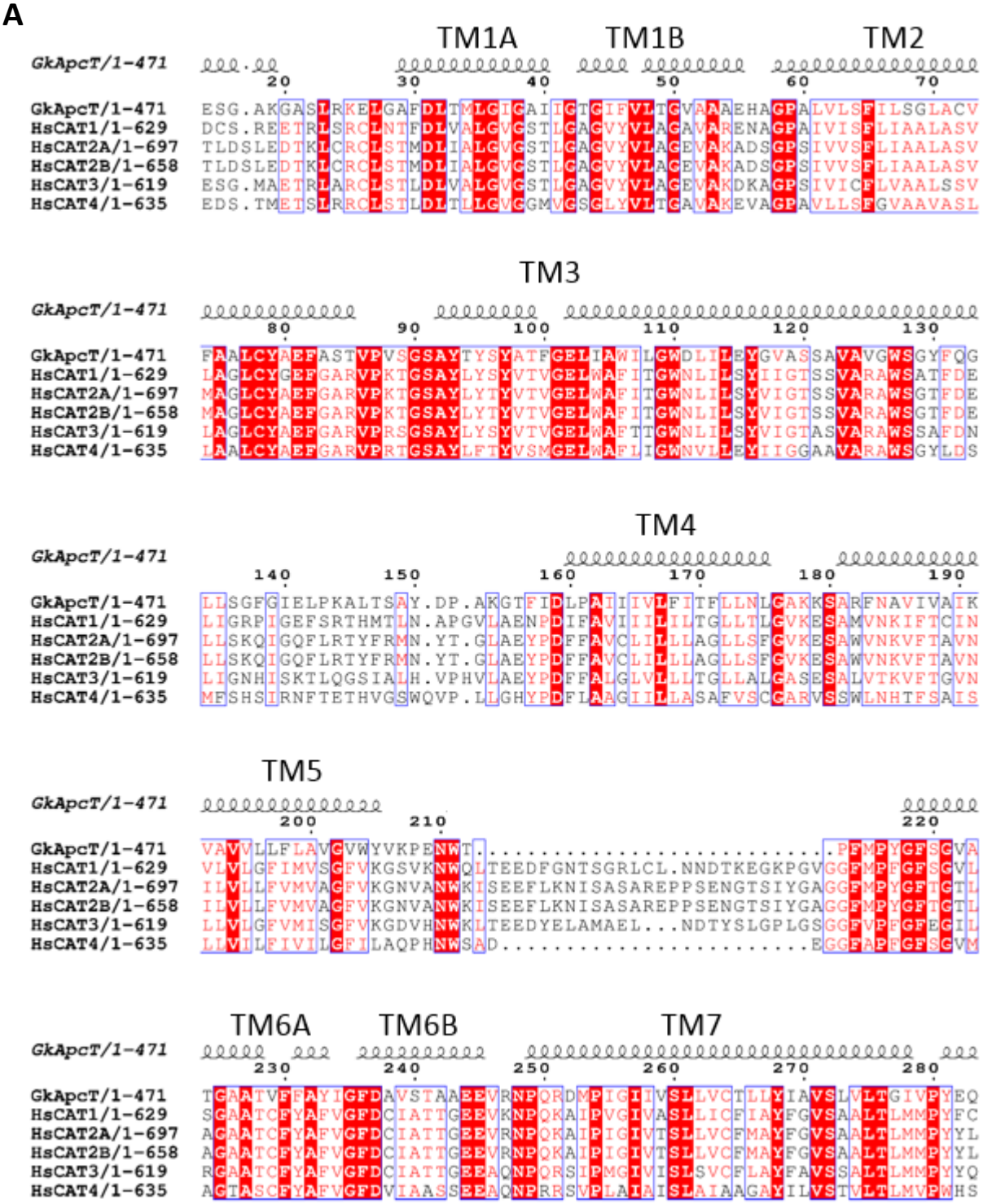
	***Gk*ApcT (PDB-ID: 5OQT)**
Refinement	
Refinement program	BUSTER 2.10.2
Reflections used in refinement	18180 (1770)
Reflections used for R_{free}	901 (81)
R_{work}/R_{free} before refinement	0.423/0.454
R_{work}	0.215 (0.263)
R_{free}	0.270 (0.315)
CC_{work}	0.800 (0.750)
CC_{free}	0.993 (0.705)
Number of non-hydrogen atoms	4132
In macromolecule	3695
In ligands	329
Protein residues	485
Water molecules	108
RMS_{bonds}	0.014
RMS_{angles}	2.08
Ramachandran favoured (%)	95
Ramachandran outliers (%)	0.63
Rotamer outliers (%)	2.1
Clashscore	5.72
Average B-factor (Å^2)	79.91

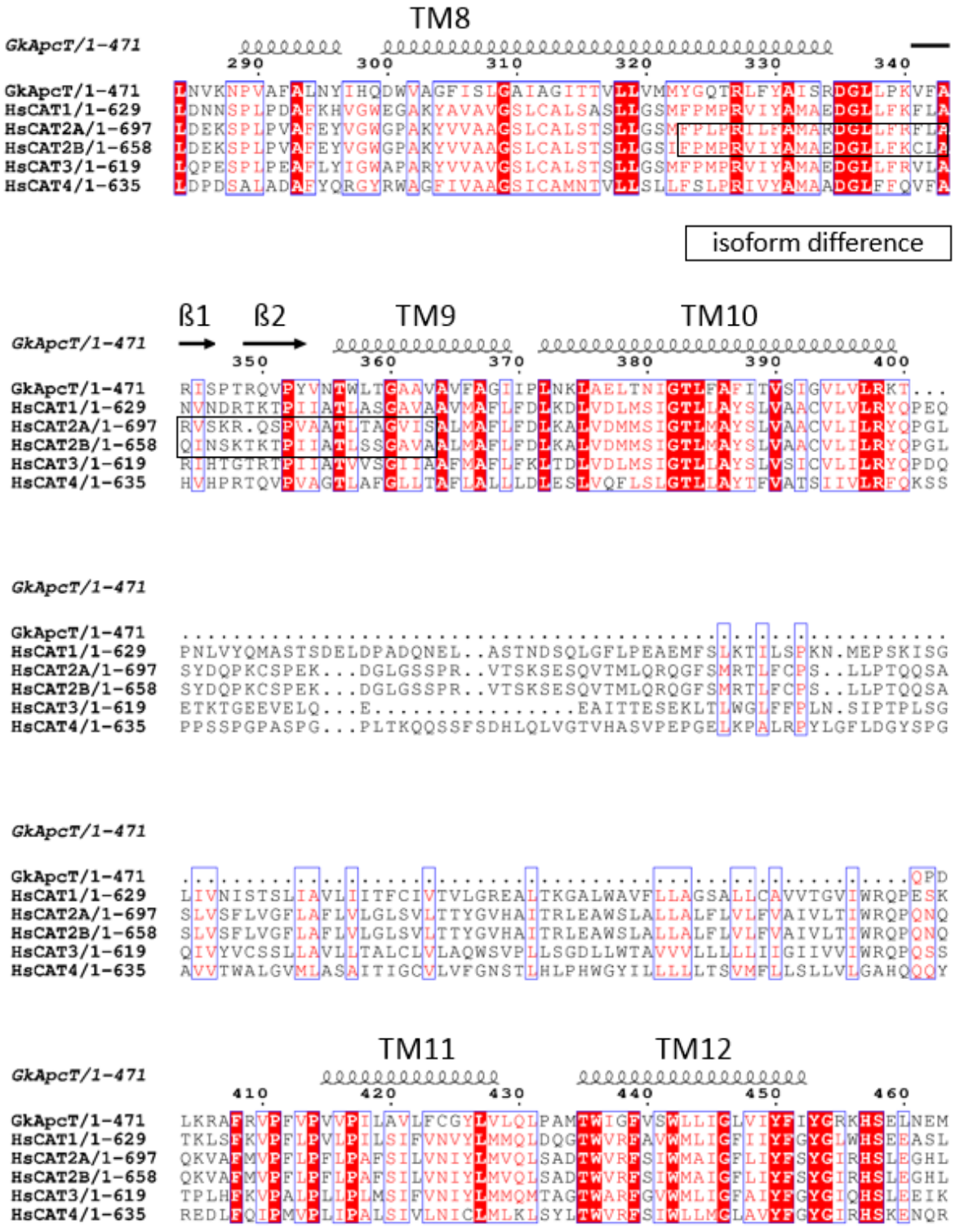
[1]Values for highest resolution shell are given in parentheses

5.3 Structure of *Gk*ApcT

The final model of *Gk*ApcT represents the structure of the closest SLC7 family homologue to date, sharing 40 % sequence identity with the human CATs (Figure 5-7).

A




TM8
GkApcT/1-471
290
300
310
320
330
340
GkApcT/1-471
HsCAT1/1-629
HsCAT2A/1-697
HsCAT2B/1-658
HsCAT3/1-619
HsCAT4/1-635
isoform difference
ß1
ß2
TM9
TM10
GkApcT/1-471
350
360
370
380
390
400
GkApcT/1-471
HsCAT1/1-629
HsCAT2A/1-697
HsCAT2B/1-658
HsCAT3/1-619
HsCAT4/1-635
GkApcT/1-471
GkApcT/1-471
HsCAT1/1-629
HsCAT2A/1-697
HsCAT2B/1-658
HsCAT3/1-619
HsCAT4/1-635
GkApcT/1-471
HsCAT1/1-629
HsCAT2A/1-697
HsCAT2B/1-658
HsCAT3/1-619
HsCAT4/1-635
TM11
TM12
GkApcT/1-471
410
420
430
440
450
460
GkApcT/1-471
HsCAT1/1-629
HsCAT2A/1-697
HsCAT2B/1-658
HsCAT3/1-619
HsCAT4/1-635

```
GkApcT/1-471
                470
GkApcT/1-471    ARTEEKAG.....................................
HsCAT1/1-629    DADQARTPDGNLD.QCK............................
HsCAT2A/1-697   RDENNEEDAYPDN.VHAAAE.EKSAIQANDHHPRNLSSPFIFHEKTSEF
HsCAT2B/1-658   RDENNEEDAYPDN.VHAAAE.EKSAIQANDHHPRNLSSPFIFHEKTSEF
HsCAT3/1-619    SNQPSRKSRAKTVDLDPGTL.YVHSV...................
HsCAT4/1-635    ELPGLNSTHYVVF.PRGSLEETVQAMQPPSQAPAQDPGHME.......
```

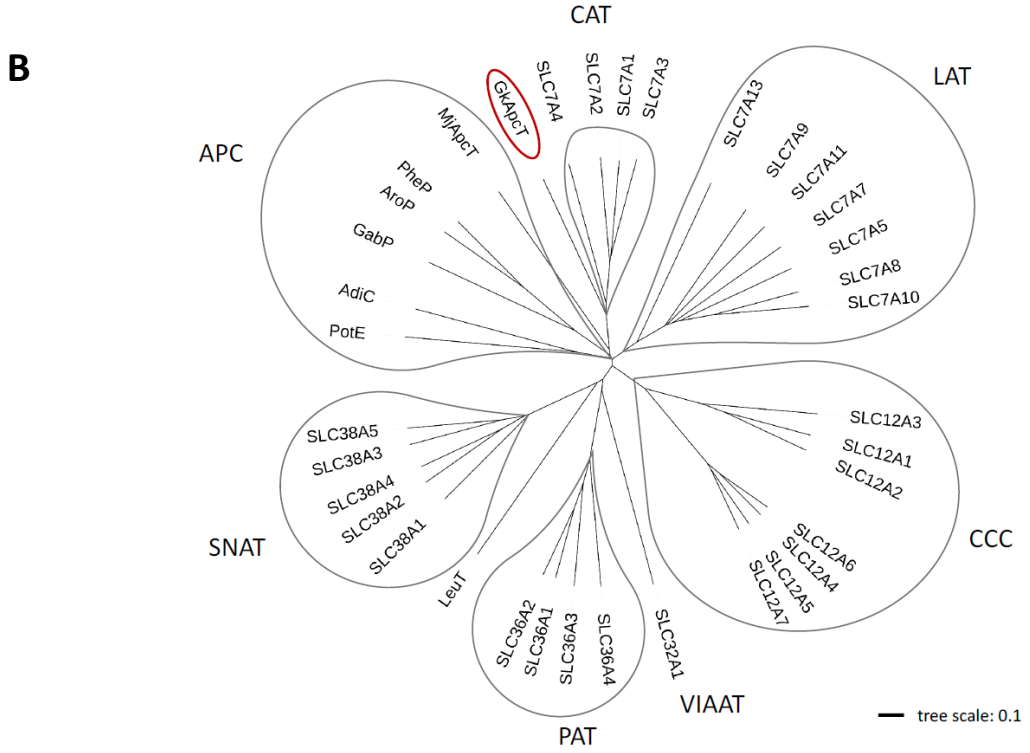


Figure 5-7: Sequence alignment to human SLC7/CAT family

Sequence alignment of *Gk*AcpcT (WP_042382604) and human CAT members: CAT1 (Uniprot: P30825), CAT2A (P52569-2), CAT2B (P52569-1), CAT3 (Q8WY07) and CAT4 (O43246) generated with PROMALS3D (Pei and Grishin, 2014) and JalView (Waterhouse *et al.*, 2009) (A). The alignment is coloured by conservation (high conservation – red background, low conservation – red letters, no conservation – white background and black letters) and the transmembrane helices of *Gk*ApcT are marked as coils above the alignment. Accession codes for the sequences used in the phylogenic tree (B) are given in the Appendix.

*Gk*ApcT consists of 12 transmembrane helices and shows the protein in the inward occluded substrate-bound conformation (Figure 5-8). The structure also revealed the association with another helical protein identified as YneM, which belonged to the additional density found during the initial phasing trials (see Section 5.2). The crystal structure of *Gk*ApcT shares a LeuT-fold as was expected due to its relation to other LeuT- fold like proteins in the APC family (Figure 5-8C). The N-terminus of *Gk*ApcT is extended forming a lateral helix on the

intracellular side of the protein (Figure 5-8A) and also shows an intracellular β-sheet, referred to as the β-hairpin, that is not observed for other APC transporters. The N-and C-terminii are located on the intracellular side (Figure 5-8A).

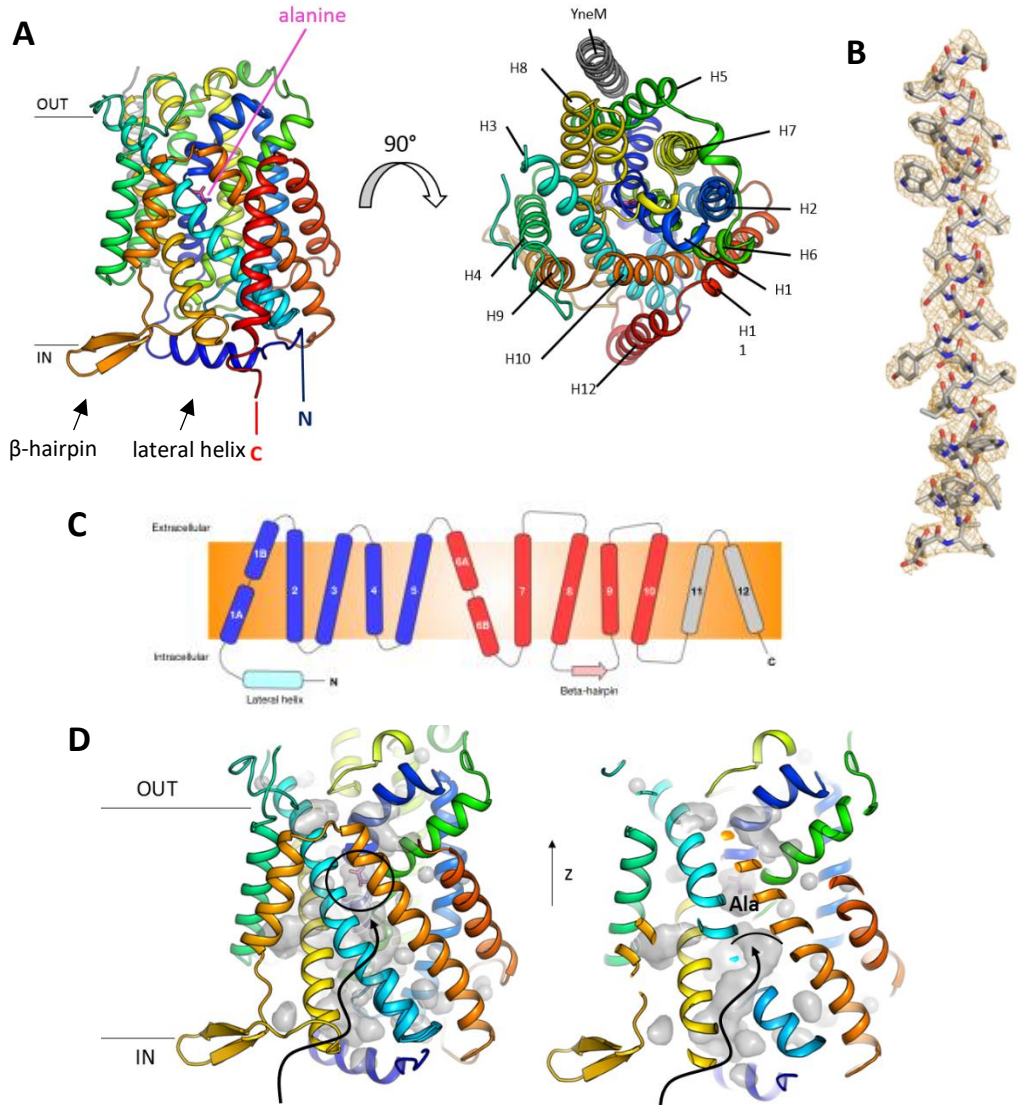


Figure 5-8: Crystal structure of GkApcT with topology diagram

*Gk*ApcT (rainbow, from N-terminus (blue) to C-terminus (red)) in side and top view (A). Helices are marked and the additional helix found in the structure (YneM) are labelled. Representative electron density map (*2Fo-Fc*, orange, 1*σ*) is shown for TM3 of *Gk*ApcT (B). The topology diagram of *Gk*ApcT shows the typical LeuT-fold topology with the symmetrical parts highlighted in blue and red (C). The substrate binding site is occluded but an intracellular cavity (indicated by the arrow) protrudes towards the substrate (D).

The crystal lattice of the *Gk*ApcT structure shows the common two-layer arrangement for LCP grown crystals of membrane proteins. In these type I crystals, the two layers consist of oppositely oriented molecules (Figure 5-9).

The crystal structure shows the protein with an alanine bound, identified based on prior information of its substrate specificity (Chapter 4) and used for co-crystallisation experiments. Although this protein is so closely related to the CATs, the substrate specificity differs noticeably, given that the CATs are transporters of cationic amino acids whereas *Gk*ApcT favours small neutral and polar amino acids. This raises the question: how did the CATs evolve to bind such a different class of substrates?

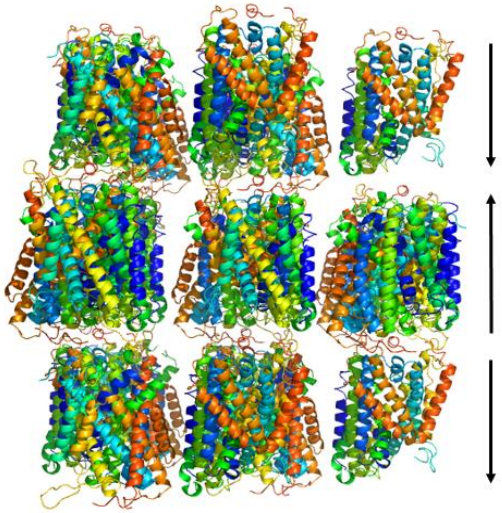


Figure 5-9: Crystal packing

LCP grown crystals tend to form type I crystals, forming layers of protein with inverted orientations. This was also the case for the crystals formed by *Gk*ApcT (rainbow N to C-terminal). Of the three layers illustrated, the middle one shows *Gk*ApcT in one orientation surrounded by layers of the protein oriented in the opposite direction (indicated by arrows).

5.4 Comparison to *Mj*ApcT, LeuT and AdiC

*Gk*ApcT is a member of the APC family showing the topology of the LeuT fold, the common fold of the APC family. Comparisons with LeuT, AdiC and the more closely related *Mj*ApcT were performed. Crystal structures of the three proteins are available in the occluded state of which only LeuT and AdiC have been solved with a substrate bound in this conformation. Superposition shows the highest RMSD of 18.3 Å to LeuT, followed by an RMSD of 3.8 Å to AdiC, and of 1.8 Å to *Mj*ApcT (Figure 5-10).

The alternating access mechanism of LeuT is described by the rocking-bundle model, where transporter consists of a scaffold, and the flexible core domain performing the main motions

during the transport cycle (Figure 5-11). A closer look at the core domain reveals additional differences between LeuT and *Gk*ApcT (Figure 5-11).

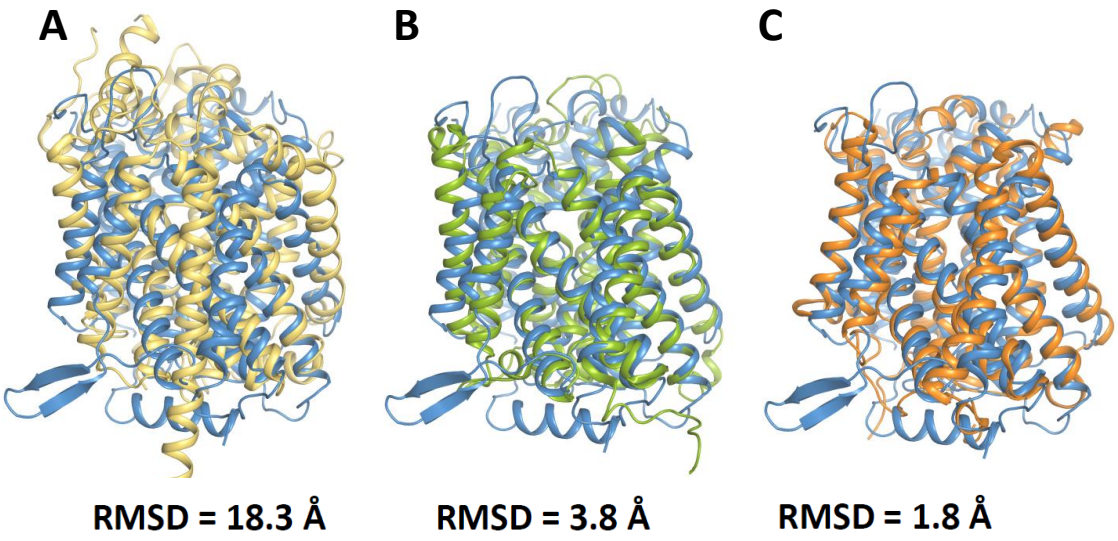


Figure 5-10: Superposition of GkApcT onto LeuT, AdiC and MjApcT

*Gk*ApcT (blues) was superimposed onto the occluded states of LeuT (yellow, PDB-ID: 2A65, 1.65 Å) (A), AdiC (green, PDB-ID:3L1L, 3.0 Å) (B) and *Mj*ApcT (orange, PDB-ID: 3GIA, 2.32 Å) (C). The RMSD of Cα is given underneath each superimposition.

The core domain of *Gk*ApcT consists, as in LeuT, of TM1, 2, 6 and 7 and the extracellular loop 4 between TM6 and TM7 (Figure 5-11). Between LeuT and *Gk*ApcT, TM1 and TM6 of *Gk*ApcT are located further towards the centre of the protein, thereby pushing EL4 upwards, which is pushed further into the binding site in LeuT. Additionally, TM2 shows an upward movement possibly forcing TM1 and TM6 further towards the binding site which then pushes TM7 away from it. Since *Gk*ApcT was crystallised in the inward-occluded state, it seems that the differences between *Gk*ApcT and LeuT give insight into the subtle conformational changes from the occluded to the inward-occluded state, the latter in which LeuT has not yet been crystallised.

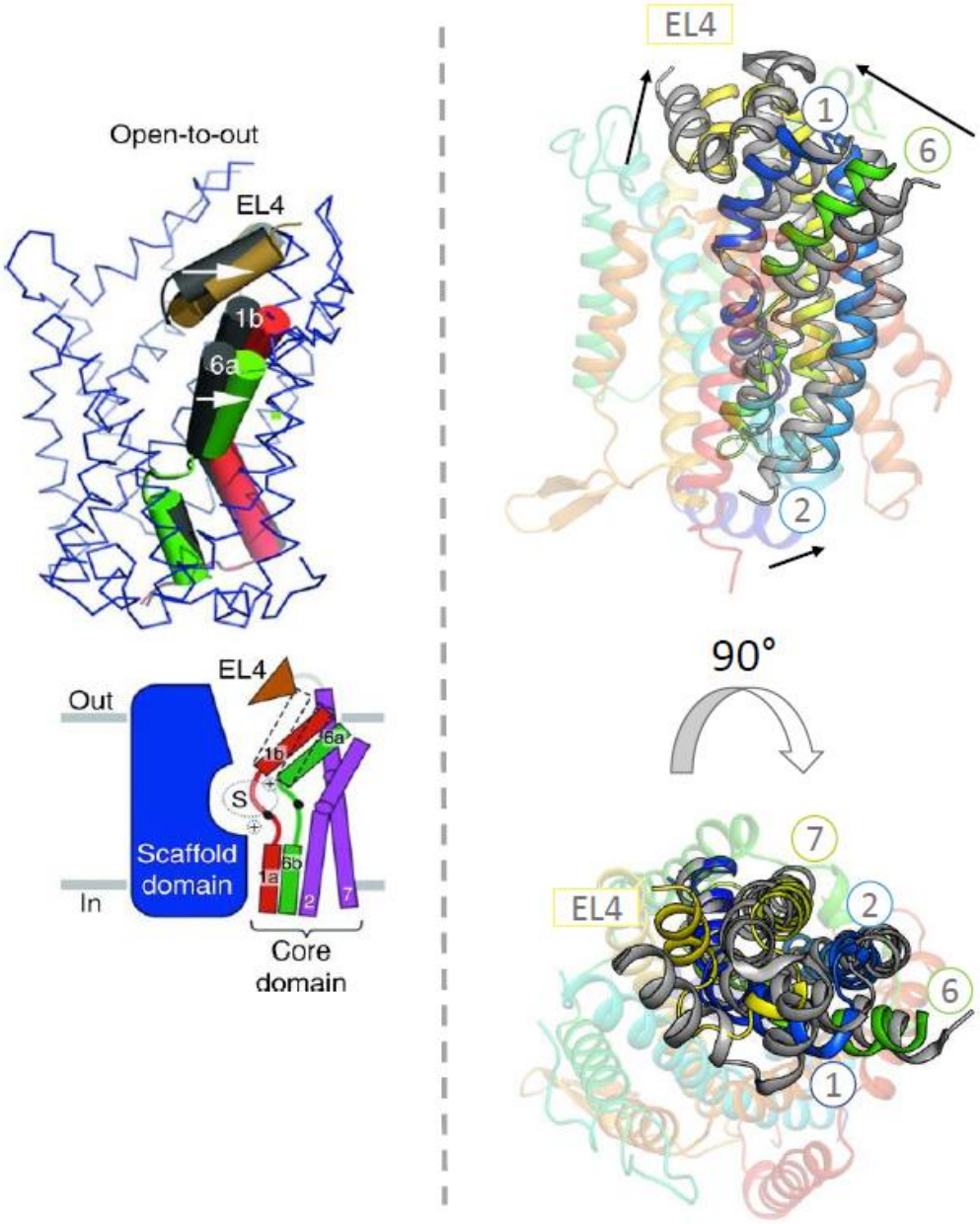


Figure 5-11: Structural differences in the core domain between GkApcT and LeuT

The rocking-bundle mechanism observed for LeuT divides the protein into a rigid part, the scaffold domain (left, dark blue) and the flexible core domain, comprised of TM1 (left, red), TM6 (left, green), EL4 (brown) and TMs 2 and 7 (left, purple), as illustrated by Krishnamurthy and Gouaux, 2012 (left). The core domain performs the main conformational changes for substrate transport. The structural arrangement of the core domain helices (labelled) of *Gk*ApcT (right, coloured) was compared to that of the core domain of LeuT (right, grey). Both structures are shown in their occluded, substrate bound conformation.

5.5 MECHANISTIC ASPECTS OF SUBSTRATE TRANSLOCATION BASED ON *GK*APCT STRUCTURE

The alternating access mechanism has been described for many members of the APC family based on structural evidence (Gao *et al.*, 2010; Weyand *et al.*, 2011; Krishnamurthy and Gouaux, 2012; Perez *et al.*, 2012). In LeuT, the substrate translocation is performed by movement of one part of the transmembrane domains (the core domain), while the remaining part remains rigid (the scaffold domain) (Krishnamurthy and Gouaux, 2012; Penmatsa and Gouaux, 2014). The substrate binding site is opened to the extracellular side by the simultaneous movement of TM1 and TM6, combined with the lifting motion of EL4 (Krishnamurthy and Gouaux, 2012; Penmatsa and Gouaux, 2014). Upon substrate and ion binding, the protein closes over the substrate binding site, and the substrate release into the intracellular side is achieved by a kink movement of TM1 (Krishnamurthy and Gouaux, 2012; Penmatsa and Gouaux, 2014).

Although the crystal structure of *Gk*ApcT represents the inward-occluded conformation, the structure might still give information how the release of the substrate could be achieved. Here, the movement of TM1 would be sterically hindered by the presence of the YneM protein and additionally by the N-terminal lateral helix to which TM1 is attached (Figure 5-12).

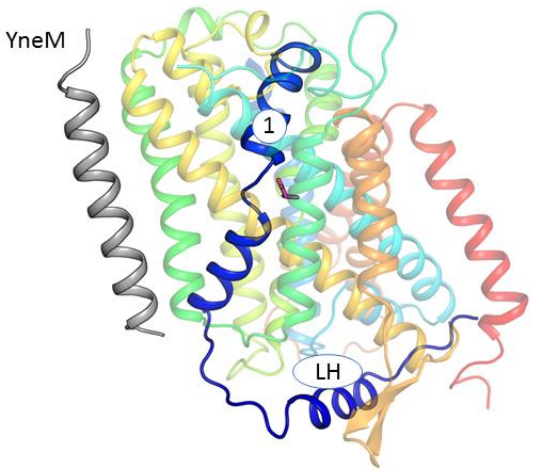


Figure 5-12: Relative position of YneM to TM1 of GkApcT

GkApcT's (coloured) TM1 and the N-terminal lateral helix (both labelled, dark blue) are highlighted together with YneM which resides closely to the N-terminal part of TM1.

Based on the symmetrical domain arrangement ('5+5 symmetry') it is possible, if either an outward or inward-open/occluded conformation is present, to model the opposite conformation of the protein. This method is called repeat-swap modelling (Vergara-Jaque *et al.*, 2015). It has proven to be valuable in predicting the helical movement during the transport

cycle of several proteins, such as LeuT (Forrest *et al.*, 2008), the glutamate transporter Glt_{Ph} (Crisman *et al.*, 2009), or the sodium/calcium exchanger NCX_*Mj* (Liao *et al.*, 2012). The modelling method has also been successful in predicting the elevator mechanism of the sodium/proton exchanger NhaA (Schushan *et al.*, 2012).

Using this method, the outward-occluded state of *Gk*ApcT was generated using MODELLER (Webb and Sali, 2014). Superimposition of the original structure onto the repeat-swapped model of *Gk*ApcT revealed that during the transition from the outward-occluded to the inward-occluded state, TM1 moves towards the binding site to close it from the extracellular side and TM6 kinks out to open on the intracellular side, opening the substrate cavity (Figure 5-13). This observation is in accordance with our assumption that it might not be possible for TM1 to perform such a movement on the intracellular side of the protein due to its restriction by the lateral helix at the N-terminus.

To summarise, TM1 kinks away on the extracellular side when the transporter changes from the inward to the outward-occluded state, and TM6 performs a kink-movement when the transporter transitions the other way. These helical movements are opposite to those observed for LeuT and AdiC. The transition from the occluded to the outward open states of LeuT and AdiC require TM6 to move aside to open the substrate binding site (Kowalczyk *et al.*, 2011; Krishnamurthy and Gouaux, 2012). Since no inward-occluded state of AdiC has been solved yet, but is available for LeuT, the transition from occluded to inward-open in LeuT revealed the 'dog-leg' movement of TM1 (Krishnamurthy and Gouaux, 2012) and not TM6 as we propose for *Gk*ApcT. Thus, *Gk*ApcT could be the first example to use TM6 for intracellular substrate release rather than the symmetry-related TM1. However, this would be consistent with the repeat swap methodology, the transporter can also change the movement of the helices between the symmetrical repeats. Unless the transporter is captured in an inward-open conformation these statements remain speculative.

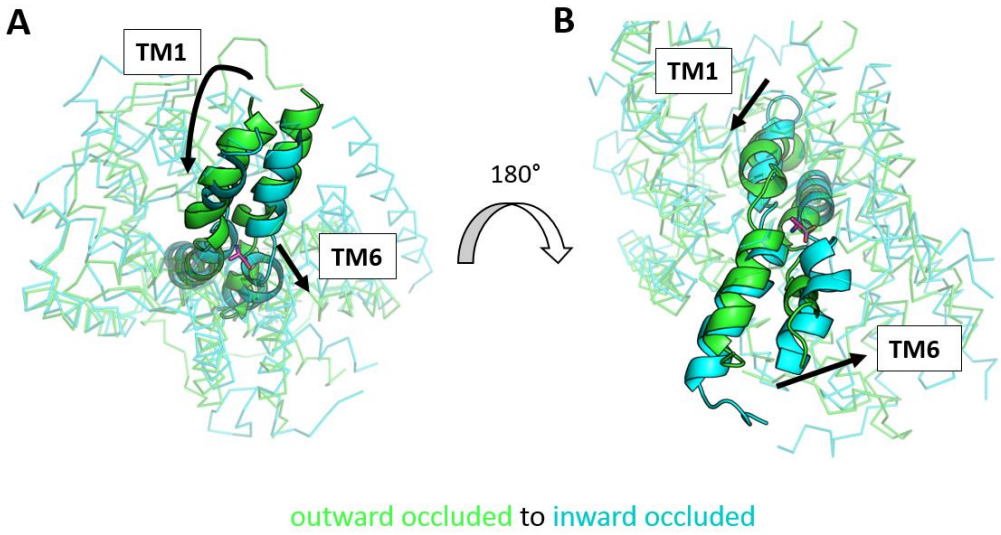


Figure 5-13: Major structural changes suggested by the repeat swap model of GkApcT

Comparison of the repeat-swap model, showing *Gk*ApcT in the outward-occluded form (green) and the original structure (blue) in the inward-occluded conformation, gives information on the helical movement during the transport cycle. To switch from the outward to the inward-occluded state, the bottom part of TM6 swings aside, whereas the extracellular half of TM1 moves towards the centre of the protein.

5.6 Summary

This chapter presented the workflow from the collection of a full data set on a single crystal of *Gk*ApcT at 2.86 Å to the final model of the protein. Though MR provided initial phases, further information on the positions of the sulphur atoms in the protein enabled the building of the final model.

The different methods for obtaining additional phase information have been described and the difficulties that can be faced in the process highlighted. Although some are commonly used, they did not prove fruitful in this case. Thus, each method should to be tested in parallel. The novel approach of using long-wavelength data collections at DLS I23, which has recently been made available to users, and at the Photon Factory BL1-A beamline were the key in this case. Although direct sulphur SAD phasing was not successful, the combination with the initial MR phases enabled the calculation of a difference Fourier map showing the sulphur peaks relative to the MR model. This allowed the correct positioning of the protein sequence into the electron density. In total over 300 model and refinement cycles were performed until the final model of *Gk*ApcT was obtained.

The crystal structure presented here shows *Gk*ApcT in the inward-occluded substrate-bound conformation. The structure revealed a surprising interaction with another protein, YneM. Additionally, the structure verifies *Gk*ApcT's relationship with the LeuT-fold family by comparison with other available structures of APC members.
Based on *Gk*ApcT's inward-occluded conformation, a repeat-swapped model could be generated to model the outward-occluded conformation. This model suggested that the helical movement from one to the other conformation was different to that found for other APC members, although based on the internal symmetry it had already been suggested that the helix performing the kink movement might be interchangeable.

The structure of *Gk*ApcT was further analysed and combined with functional data to gain insight into the substrate binding and proton-coupling mechanism. The studies also aimed to identify the effect of YneM on this protein and the role of the other novel secondary structure features, which will be discussed in detail in the next chapter.

6 Insights into the transport mechanism of an ApcT homologue

The structural data obtained for *Gk*ApcT helps to explain the substrate specificity of the protein but also offers the possibility of investigating the residues that are involved in substrate binding, proton-coupling and the translocation mechanism. The following chapter focuses on the mutational studies that were performed based on the available structural information, in order to gain a deeper understanding of the mechanism of action of *Gk*ApcT and its human homologues, the SLC7 family.

6.1 ALANINE BINDING SITE

Based on previous experiments that determined the substrate specificity of *Gk*ApcT, and comparison with the homologue *Mj*ApcT, alanine was suggested as one of the main substrates of *Gk*ApcT and was therefore used for co-crystallisation experiments, resulting in final structure containing a bound alanine molecule in the centre of the protein.

However, before the crystal structure of *Gk*ApcT was available, important residues in substrate recognition were identified based on the comparison of mutational studies on other APC members shown to have an effect on substrate translocation (Kowalczyk *et al.*, 2011; Shang *et al.*, 2013; Kaur *et al.*, 2014) (Figure 6-1A). Therefore, Thr43, Leu112, Tyr116, Val123, Phe231, Met321 were mutated and assayed under substrate-driven conditions to determine the effect of those mutations on alanine transport. The residues chosen for mutagenesis were located to be in the centre of the preliminary homology model of *Gk*ApcT (Figure 6-1A). The highlighted residues were substituted to alanine and all mutants could be purified and reconstituted into liposomes for further analysis (Figure 6-1A). The mutants *Gk*ApcT-Leu112Ala and *Gk*ApcT-Tyr116Ala showed reduced counterflow activity, indicating that these residues are important for substrate binding or for the translocation mechanism (Figure 6-1B).

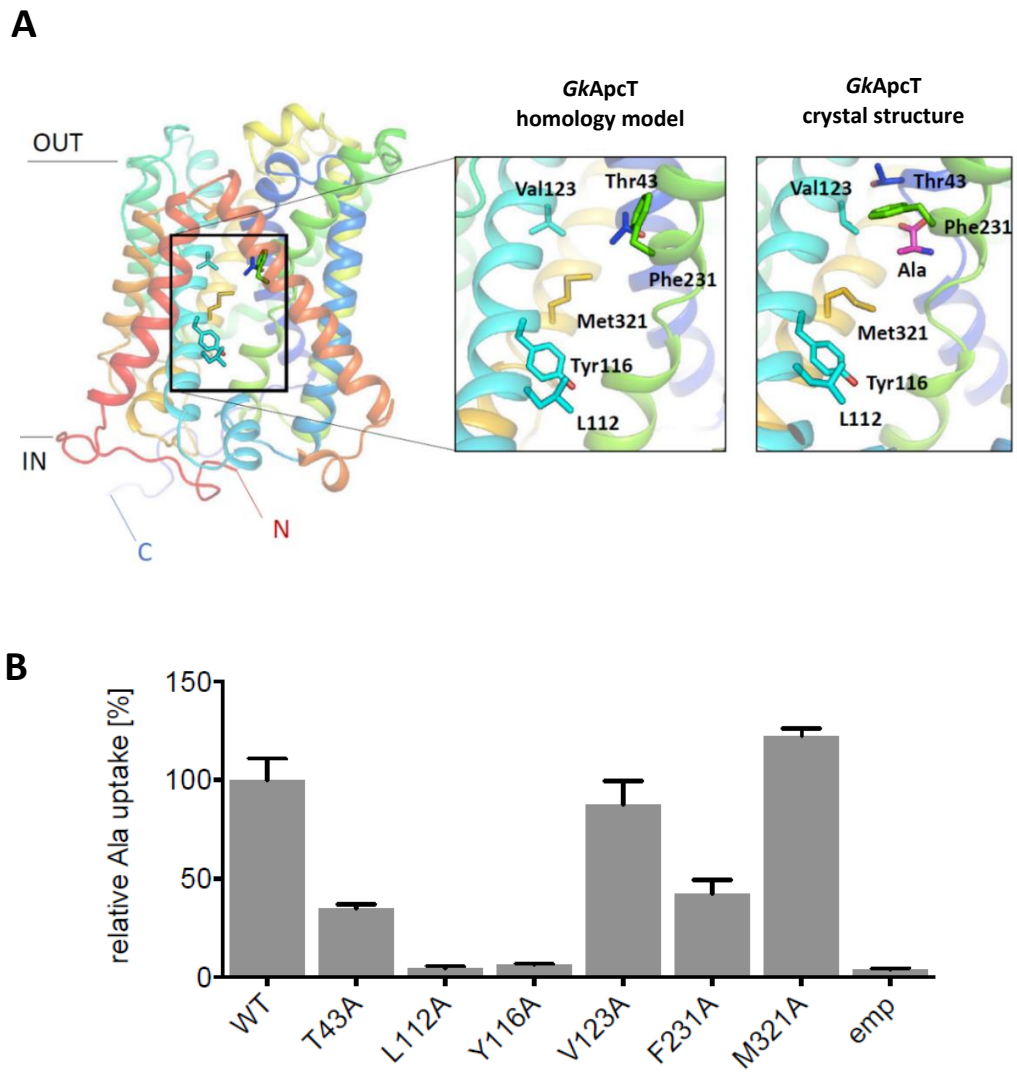


Figure 6-1: Proposed binding site residues

Based on mutational studies on other APC family members and comparison of the location of those mutant residues with a homology model of *Gk*ApcT, several residues were identified that were proposed to be involved in the substrate translocation mechanism or be important for the function of *Gk*ApcT. Additionally, comparison with the crystal structure of *Gk*ApcT revealed that the homology model suggested the correct location of these residues (A). Counterflow experiments with the proposed mutants were carried (B). Proteoliposomes containing 25 mM citrate phophsate pH 6.0, 100 mM NaCl, 2mM $MgSO_4$ and 2 mM alanine were diluted 1:25 into buffer with 25 mM citrate phosphate pH 6.0, 100 mM NaCl, 2mM $MgSO_4$,250 nM ^{3}H-alanine and 100 µM alanine. As control, liposomes devoid of protein (emp) were also tested. Uptake levels after 5 min are plotted.

Reduced activity was also observed for mutants Thr43Ala and Phe231Ala but no effect was observed for mutations Met321Ala and Val123Ala (Figure 6-1B), possibly because the latter maintained the hydrophobic environment necessary for alanine binding and transport. Therefore, Met321 and Val123 are not directly involved in substrate binding, but Thr43, Leu112, Tyr116 and Phe231 are playing a role in the translocation or binding mechanism of alanine. Superposition of the homology model and the crystal structure reveals that the

overall position of the mutants was correct in the model, although slight differences in the position of the side chains were observed (Figure 6-1A), for example Thr43 of the homology model sits where the alanine is located in the crystal structure (Figure 6-1A). Nevertheless, the homology model provided valuable initial insight into the structure of *Gk*ApcT.

During the model refinement of *Gk*AcpT, difference density was observed between the discontinuous parts of TMs 1 and 6 (Figure 6-2D), as observed for other APC transporters (Yamashita *et al.*, 2005; Gao *et al.*, 2009; Shimamura *et al.*, 2010). The mutational data on residues assumed to form the substrate binding site, supported the assumption that the difference density found in the crystal structure belonged to a bound alanine molecule. As seen in Figure 6-2 the carboxyl and amine groups of alanine interact only with the backbone of TMs 1 and 6, where most of the interactions are made with TM1 which contains, along with other APC family members, a conserved GlySerGly (GSG) motif.

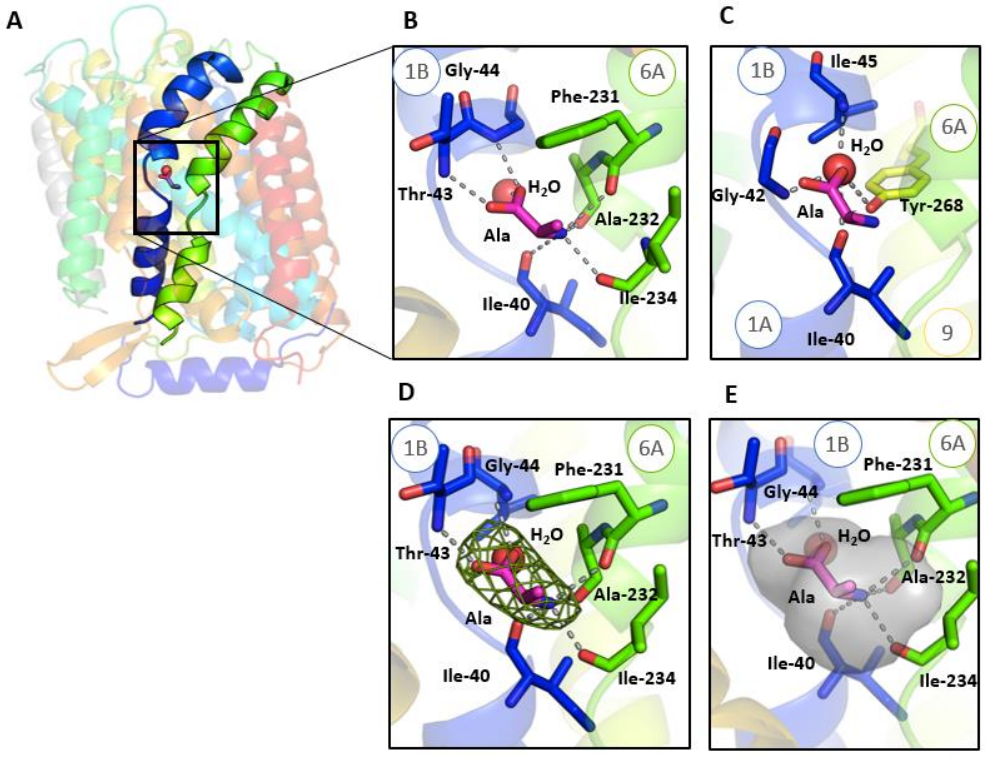


Figure 6-2: Substrate binding site of GkApcT

Interaction of alanine (pink) in the binding site of *Gk*ApcT mostly formed by TM1 (blue) and TM6 (green) (A&B). A water molecule interacts with alanine, Tyr268 and TM1(C). Positive difference density (F_o-F_c, 3 σ, green mesh) was apparent in the structure during refinement into which the alanine molecule fit (D). The small cavity in the binding site explains the specificity of *Gk*ApcT for small molecules (E).

The substrate binding cavity is formed by residues Ala39, Ile40, Ile41, Gly42, Thr43, Gly44, Ala119, Val123, Phe231, Ala232, Ile234, Ile314, Val317, Met321, and Thr379. The carboxyl group of alanine interacts via hydrogen bonds with the backbone amines of Thr43 and Ala42, and with a water molecule (Figure 6-2A). The amino group of alanine interacts with the backbone of residues Phe231, Ile234 and Ala232 (Figure 6-2B). The water molecule mostly forms contacts to TM1, via the backbone nitrogens of Gly42 and Ile45, the backbone carbonyl group of Ile40 and the side chain hydroxyl group of Tyr268 (Figure 6-2C). The volume of the substrate binding site cavity is sterically limited to a volume of ~290 $Å^3$ (Figure 6-2E). This results in the substrate specificity of *Gk*ApcT for small non-polar and polar residues as identified in the counterflow experiments (see Section 4.4), since the binding site would sterically occlude any larger amino acids. The substrate binding site is mainly limited by residue Met321 on the intracellular side.

6.2 STRUCTURAL COMPARISON OF THE AMINO ACID BINDING SITE

Comparison of the location of alanine in *Gk*ApcT to other substrate bound structures of the APC superfamily shows a similar orientation of the substrate between TM1 and TM6 (Figure 6-3). The carboxyl groups of all the shown substrates point towards the unwound segment of TM1 whereas the amine group is oriented towards TM6 and the side chains face away from the discontinuous helices (Figure 6-3).

For the arginine/agmatine antiporter, AdiC, the transporter has been solved in different conformations with the substrate bound (Gao *et al.*, 2009; Kowalczyk *et al.*, 2011; Ilgü *et al.*, 2016). The arginine ligand was found to be enclosed by two aromatic residues (Trp202 and Trp293) stacking above and below the substrate in the occluded state (Figure 6-4). In the outward-open state, Trp202 is rotated aside, opening the substrate binding site for the substrate to enter the cavity (Figure 6-4). It is likely that the same rotational movement is performed by Trp293 of AdiC, although the protein structure has not been solved in the inward-open conformation to confirm this theory. In *Gk*ApcT, Phe231 and Met321 are in positions equivalent to the residues in AdiC, forming the same 'gate', although there is no direct interaction with Met321 found (Figure 6-4).

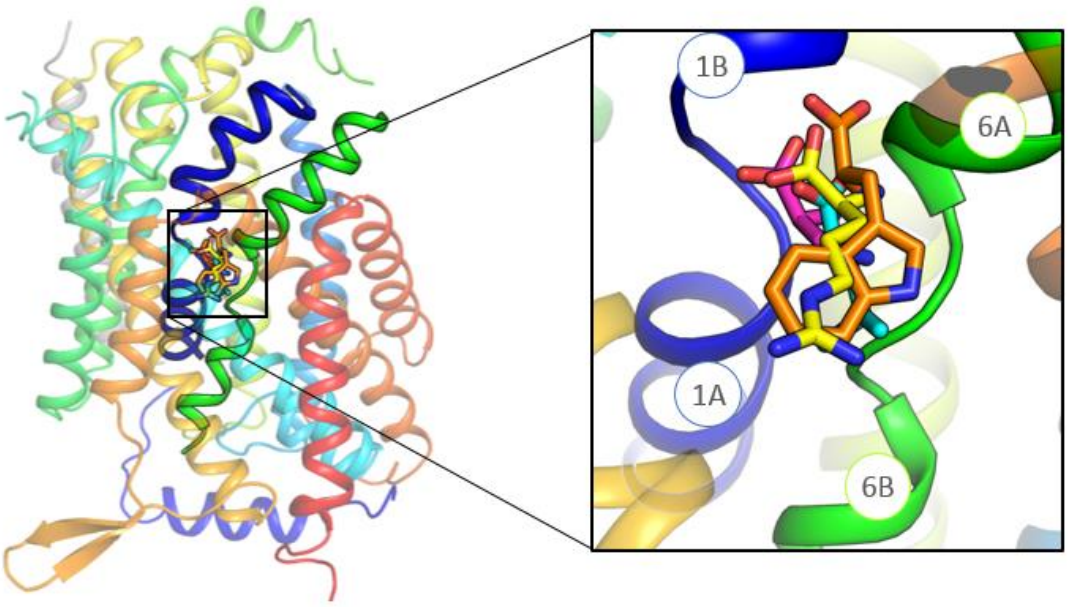


Figure 6-3: Substrate binding site of GkApcT aligns with that of other APC transporters

Alanine in the binding site of *Gk*ApcT aligns with the substrates bound to other APC members confirming its position and orientation in the crystal structure of *Gk*ApcT. The substrates sit between TM1 (red) and TM6 (green). Arginine bound in AdiC (PDB-ID:3L1L, 3.0 Å, neon green) and leucine and tryptophan bound to LeuT (PDB-ID: 2A65, 1.65 Å, light blue; PDB-ID: 3F3A, 2.0 Å orange) and alanine to *Gk*ApcT (PDB-ID: 5OQT, 2.86 Å, pink) are shown.

Based on the comparison to AdiC, it was suggested that Phe231 of *Gk*ApcT might perform a similar rotation movement to open and close the substrate binding during the transport cycle. This model would suggest that TM6 bends aside on the extracellular side in the open-outward conformation, and bends towards the centre of the protein once the substrate is bound, which contradicts the findings from the repeat-swapped model discussed in Section 5.5 where it was proposed that TM1 performs this movement. The conformational changes occurring during the transport cycle can only be speculated until more structural and biochemical data have been obtained to gain a deeper insight into those events. The data presented here on the substrate specificity of *Gk*ApcT illustrate its ability to accommodate a variety of substrates.

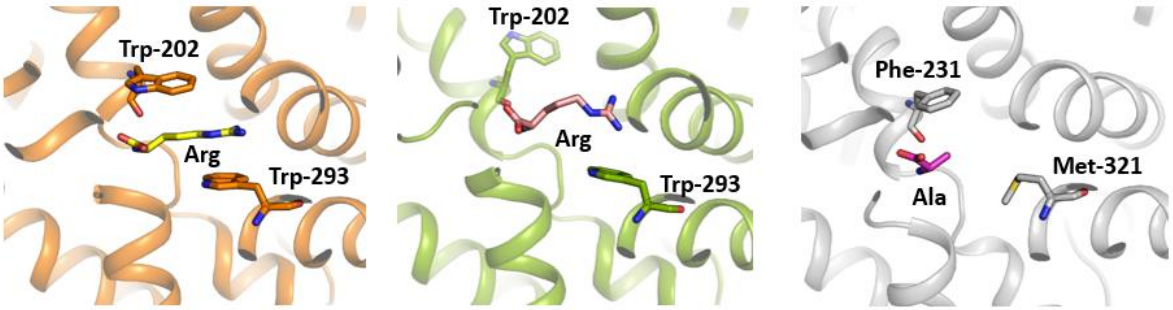


Figure 6-4: Comparison of GkApcT to AdiC

AdiC in the substrate-bound occluded conformation (PDB-ID:3L1L, 3.0Å) (left), the substrate-bound outward open conformation (PDB-ID:3OB6, 3.0Å) (middle), and *Gk*ApcT (PDB-ID: 5OQT, 2.86Å) (right) for comparison. The aromatic residues of AdiC, capturing arginine, are replaced by Phe231 and Met321 in *Gk*ApcT which show a similar arrangement to the equivalent residues in AdiC.

Additionally, the structure of *Mj*ApcT was solved in an occluded substrate-free state but a water cavity was found within the protein that was thought to be the potential binding site (Shaffer *et al.*, 2009). The water cavity is located on the intracellular half of the protein as shown in Figure 6-5. The cavity is mostly delimited by polar residues which would explain *Mj*ApcT's specificity towards alanine, serine, glutamate and glutamine (Shaffer *et al.*, 2009). Superimposition of the current *Gk*ApcT onto the structure of *Mj*ApcT reveals that the substrate binding pocket of *Gk*ApcT sits above the water cavity (Figure 6-5) as seen in the comparison of *Mj*ApcT with LeuT and AdiC (see Section 1.4.3). Thus, it seems possible that the water cavity of *Mj*ApcT was misinterpreted as the substrate binding site since the structures of substrate bound APC transporters consistently have their substrate binding site above that cavity. However, since the otherwise conserved GlySerGly motif of APC superfamily members is not conserved in MjApcT, there being GlyAlaSer, is might be plausible that the substrate binding site is displaced relative to the other transporters. Mutational studies based on a homology model fo *Mj*ApcT using *Gk*ApcT as template or a co-crystal structure of MjApcT with one of its substrates might give insight into these assumptions.

The crystal structure of *Gk*ApcT provided additional sites for mutational studies to determine how the substrate specificity is achieved. It also allowed differences from the SLC7 family to be identified and showed which residues are invaluable for the proton-coupling mechanism. Studies based on those mutations might shed light into the mechanistic aspects of the membrane transporter *Gk*ApcT.

Further investigation of the substrate binding site showed that a water molecule interacts with the bound alanine and that this water molecule is in a similar position to the first sodium ion found in LeuT (Figure 6-6). Additionally, Tyr268 participates in the coordination of the water molecule and its hydroxyl group could substitute for the Na1 of LeuT, as it also has a similarly orientation to Na1 of LeuT. The location of the second sodium ion of LeuT is occupied by Lys191, which is the corresponding residue to Lys158 of *Mj*ApcT (Section 1.4.3).

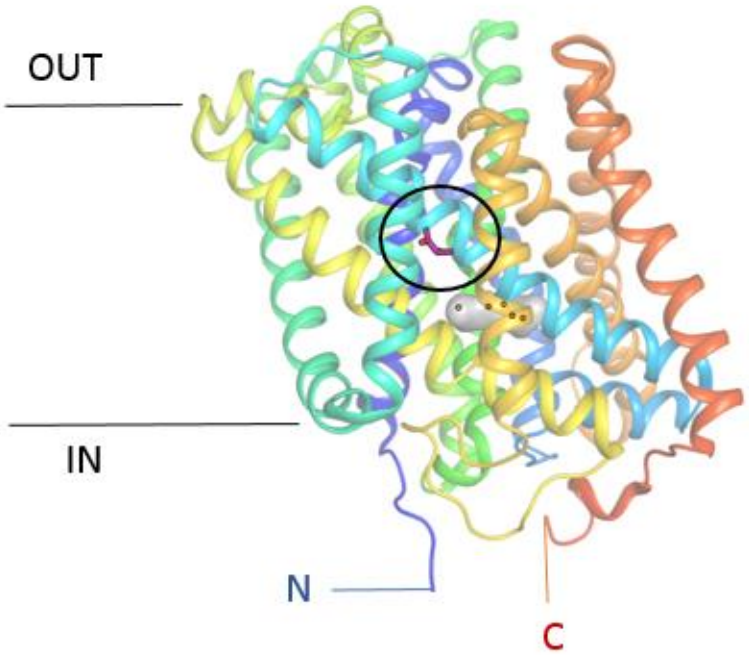


Figure 6-5: Comparison of location of alanine binding site of GkApcT and water cavity of MjApcT

The water cavity (grey surface, water molecules as red spheres) in *Mj*ApcT (PDB-ID:3GIA, 2.32 Å, cartoon in rainbow colour) resides below the substrate binding site found in *Gk*ApcT (alanine shown in pink sticks and with black oval around it).

As described in section 1.4.3, previous research on *Mj*ApcT suggested that Lys158 was involved in the proposed proton-coupling mechanism due to its similar orientation and location as Na2 bound to LeuT (Yamashita *et al.*, 2005). Since Lys191 of *Gk*ApcT occupies a similar position (Figure 6-6) interacting with residues of TM8 and TM5 (Figure 6-7A), its possible involvement in the proton-coupling mechanism of *Gk*ApcT was investigated.

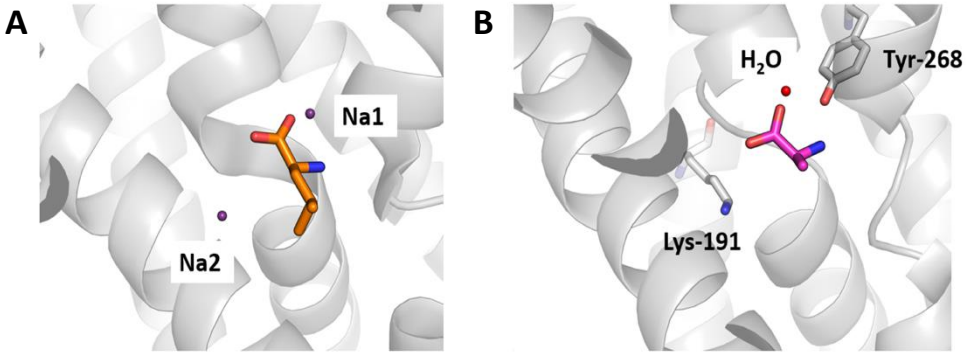


Figure 6-6: Comparison of binding mode of GkApcT and LeuT

Comparison of the binding sites of *Gk*ApcT and LeuT (PDB-ID: 2A65, 1.65Å) (A) revealed that the two sodium ions (Na1 and Na2, purple spheres) are in similar positions to residues Lys191, Tyr216 and a water molecule in *Gk*ApcT (PDB-ID: 5OQT, 2.86Å) (B).

For this, three different mutants were prepared: the residue was mutated to arginine to preserve the positive charge, to asparagine to preserve a polar environment, and to alanine to abolish any charge. The mutants were reconstituted into liposomes and electrogenic uptake of alanine was measured and compared to the wildtype activity. Only *Gk*ApcT-Lys191Asn weakly (~ 8% of WT activity) recovered alanine transport activity, whereas the other mutants showed a complete loss of their transporter function (Figure 6-7B). To exclude that the loss of activity resulted from the inability of the mutants to bind alanine, thermal stability measurements were performed. The detergent-solubilised mutants were diluted into counterflow buffer supplemented with 0.03% DDM. Temperature-intervall measurements in presence and absence of 10 mM alanine in the buffer are plotted (Figure 6-7C). The wildtype-protein and the mutants show a clear shift of their melting temperature (peak of the curve) when alanine was added to the proteins, indicating that a substrate-induced stabilisation of the protein occurs which can be assumed to indicate substrate binding in this case. The measurements also show that the *Gk*ApcT mutants have the same stability under substrate-occluded conditions as the wildtype protein, confirming that no structural stabilisation has occurred which could obscure the liposome-based assays.

It is assumed that the lysine at this position is not directly involved in the coupling of the proton, since the mutation to asparagine was able to restore the proton-coupled uptake of alanine, although to a small extent. Additionally, unlike CaiT, the mutant *Gk*ApcT-Lys191A did not show any sodium-dependent uptake, since the generated inwardly directed mem-

brane potential under the $\Delta\Psi$-driven conditions used in the experiments could also be harnessed for sodium-coupled symport. However this was not the case for the conducted experiments which implies that the observed modulated sodium-dependence of CaiT might have been specific for this protein (Kalayil *et al.*, 2013). We therefore suggest that the lysine is involved in a stabilisation mechanism via formation of hydrogen bonds to surrounding helices (Figure 6-7A) suggesting a similar mechanism in the sodium-coupled transporters, since Na2 found in LeuT, given its similar position and interaction. This might explain the loss of function for most of the mutants. Probably the charge and size of the residue are crucial to maintain the protein's function.

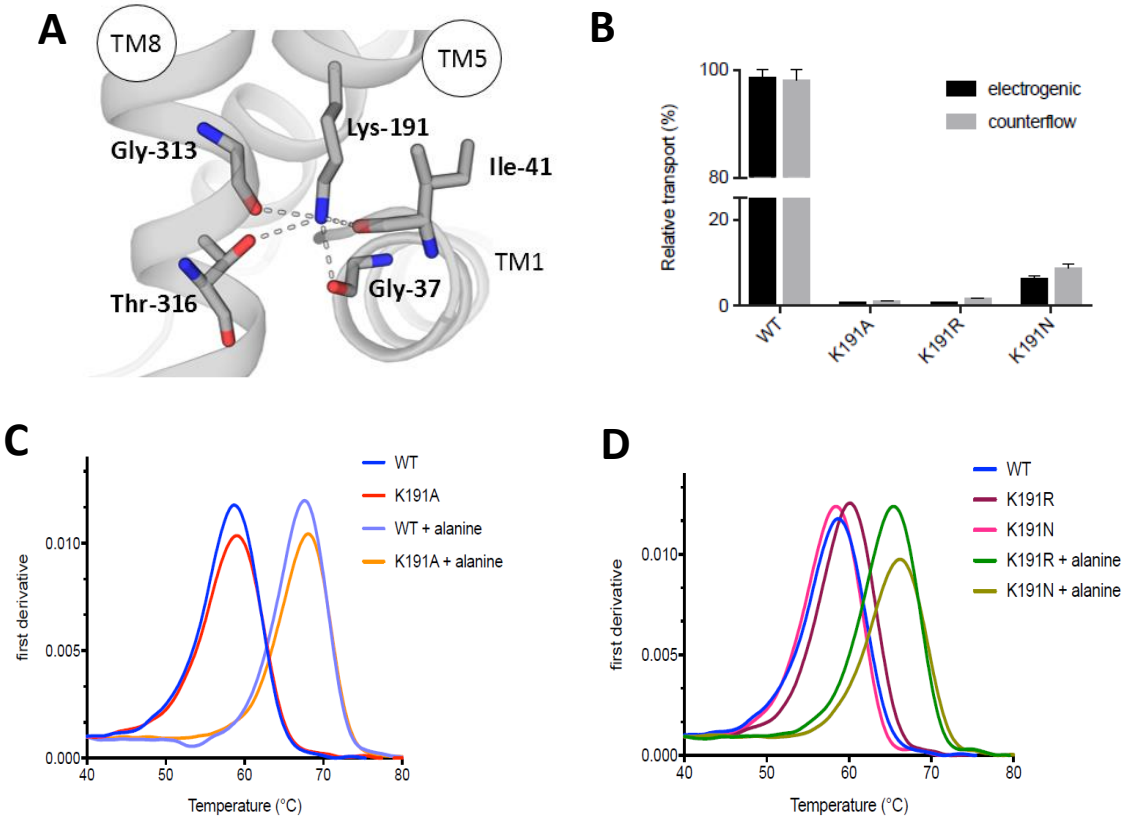


Figure 6-7: Influence of Lys191 on GkApcT transport activity

Since Lys191 of *Gk*ApcT (PDB-ID: 5OQT, 2.86Å) occupies a similar position and coordination to that of Lys158 in *Mj*ApcT (A). Electrogenic (relative uptake after 1 min) and counterflow (relative uptake after 5 min) experiments were performed fo WT and mutants Lys191Ala, Lys191Arg, and Lys191Asn using 5 µg of protein per experimental time point (B). Additionally, thermal stability measurements for WT and the listed mutants were performed. For this, the proteins were diluted 1:50 into buffer containing 25 mM citrate phosphate pH 6.0, 100 mM NaCl, 2 mM $MgSO_4$ and 0.03 % DDM to a final concentration of 0.5 mg/ml (C). Additionally, 10 mM Ala (+alaninen) were present in the buffer.

6.3 BINDING SITE MUTATIONS BASED ON THE SLC7 FAMILY

Using a similar approach to that of the modulation of LeuT to transport substrates of the closely related TnaT tryptophan transporter (Piscitelli and Gouaux, 2012), residues potentially influencing the substrate specificity of *Gk*ApcT (Figure 6-8A) were compared with the corresponding residues of the human CATs. Based on the sequence alignment a subset of them was found to differ (Figure 6-8). Single-mutants of *Gk*ApcT were generated to match the equivalent residues of the human CATs. The mutants were cloned into C43(DE3) *E. coli* cells, over-expressed in TB and purified as described for *Gk*ApcT (Section 3.2). They were subjected to reconstitution into liposomes with subsequent determination of their substrate specificity towards arginine, alanine and lysine. Counterflow assays using radiolabelled alanine as the reporter amino acid were performed with 10 mM alanine, arginine or lysine in the internal liposome buffer and the results compared to wildtype *Gk*ApcT. The counterflow experiments reveal that only mutant *Gk*ApcT-Met321Ser showed an ability to transport arginine and slight activity for lysine (Figure 6-8C).

The IC_{50} of arginine changed from the 6 mM for *Gk*ApcT-WT to 11.4 µM for the mutant (Figure 6-9A). The results point to the change of the WT towards being an arginine transporter, indicating the importance of the residue in substrate selectivity of *Gk*ApcT as a neutral amino acid transporter and the cationic CAT transporter in humans. Additionally, Met321Ser is still a proton-coupled transporter for arginine (Figure 6-9B).

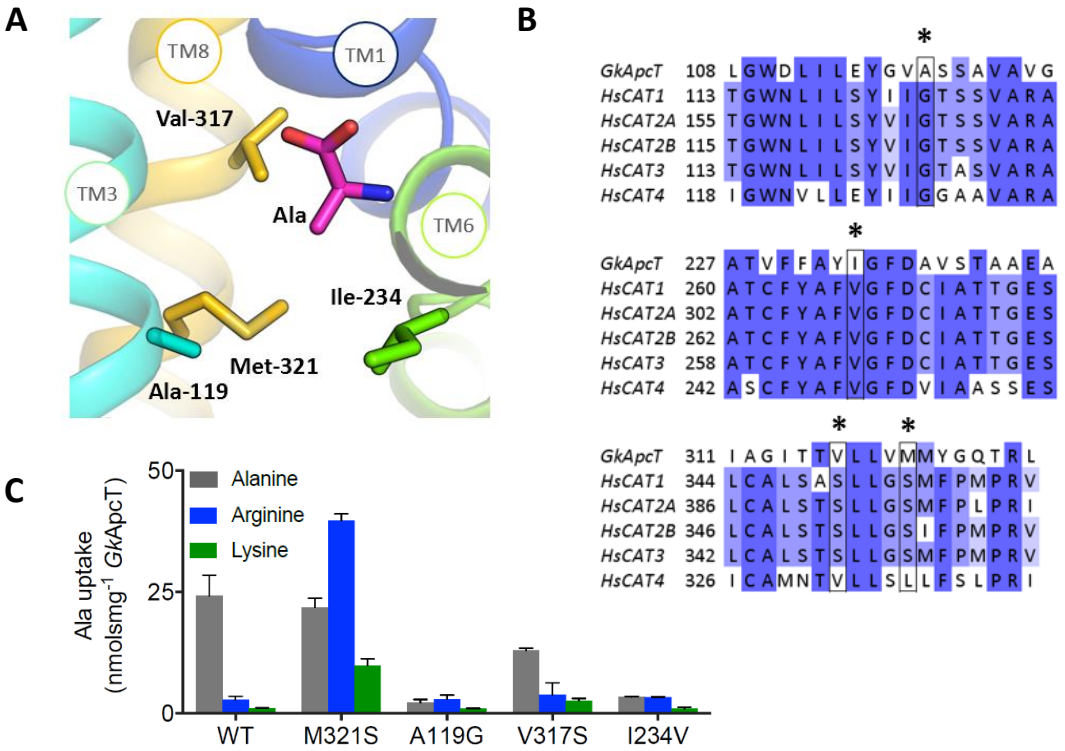


Figure 6-8: Residues involved in shaping the substrate specificity of GkApcT and the human CATs

Inspection of the substrate binding site of *Gk*ApcT suggested residues that might determine its substrate specificity. These residues were compared with the human CATs to identify any differences that might cause the change in specificity. Five residues that showed differences between *Gk*ApcT and the CATs were identified (A & B). Mutations of these residues in *Gk*ApcT to the ones found in the human CATs were used for counterflow experiments using 10 mM alanine, arginine or lysine in the internal buffer (25 mM citrate phosphate pH 6.0, 100 mM NaCl, 2 mM $MgSO_4$) (C). The bar graph shows uptake levels after 5 min. The alignment in (B) is coloured by conservation from no conservation (white) to 100 % conservation (dark purple).

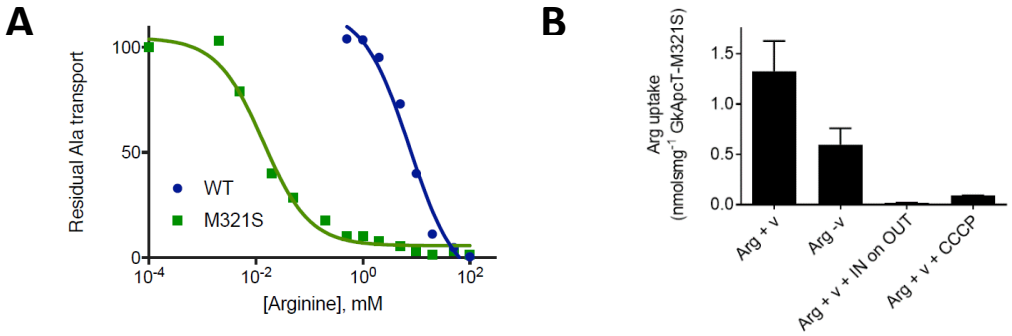


Figure 6-9: GkApcT-Met321Ser

IC_{50} experiments with *Gk*ApcT-WT and the mutant *Gk*ApcT-Met321Ser using arginine as the competitor under electrogenic conditions using 250 nM ^{3}H-alanine are plotted (A). The IC_{50} for arginine changed from 6 mM (WT) to 11.4 µM (Met321Ser). Additionally, electrogenic uptake assays using a mixture of 1 µM arginine and 250 nM ^{3}H-arginine were performed with mutant Met321Ser (B). Uptake after 1 min. is plotted.

6.4 CHANGING *GK*APCT TO AN ARGININE TRANSPORTER

The mutation of Met321 to serine resulted in a change towards increasing sensitivity to arginine in counterflow and competition assays, and also showed electrogenic uptake activity for arginine (Section 6.3). The Michaelis-Menten kinetics of arginine uptake of *Gk*ApcT-Met321Ser were determined (Figure 6-10) with a K_M value for arginine of 97.7 ± 36.1 µM and v_{max} of 0.2633 ± 0.22 nmols mg^{-1} sec^{-1}.

The mutant shows a similar apparent K_M to arginine as the wildtype protein does to alanine (Section 4.6) showing that the mutation did not influence the substrate affinity but altered the maximum transport velocity v_{max}. Since v_{max} is proportional to k_{cat}, the turnover rate and rate limiting step of transport, and to the protein concentration. The same amount of protein was used for K_M determination for *Gk*ApcT-WT and *Gk*ApcTMet321Ser, which means that the decrease in v_{max} resulted from a decrease in k_{cat}.

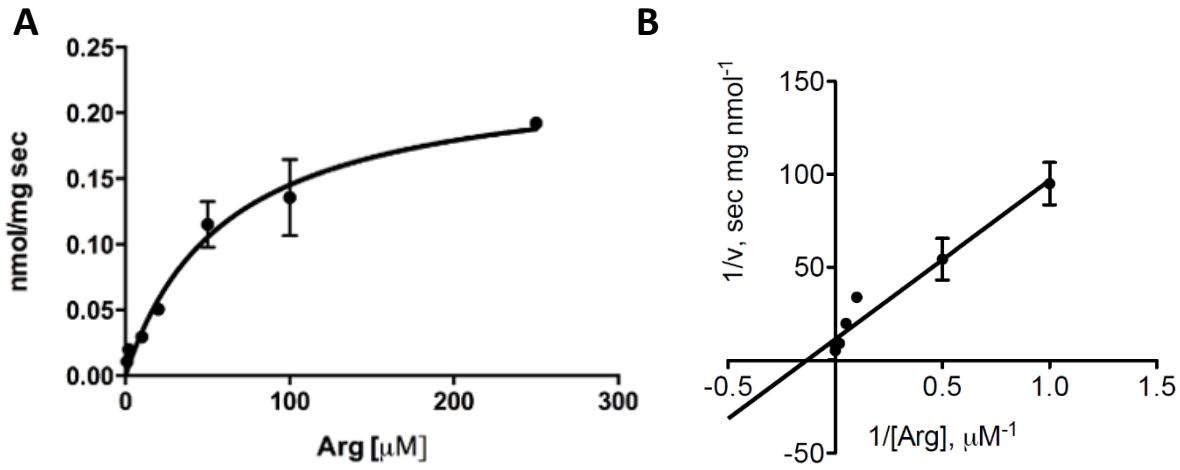


Figure 6-10: Uptake of arginine by GkApcT-M321S

The transport kinetics of *Gk*ApcT-Met321Ser for arginine transport were investigated under $\Delta\Psi$-driven conditions (A). Proteoliposomes containing internal buffer (20 mM KPi pH 6.5, 100 mM KAc and 2 mM $MgSO_4$) were diluted 1:25 into external buffer (20 mM NaPi pH 6.5, 100 mM NaCl, 2 mM Mg_2SO_4, 250 nM ^{3}H-arginine, 10 µM valinomycin and varying arginine concentrations). The initial uptake rates were plotted against the arginine concentrations used and kinetic parameters were determined in PRISM using the Michaelis-Menten equation. The Lineweaver-Burk plot of the experimental data is shown as control (B).

Similarly, this means that the substrate affinity, K_M, of *Gk*ApcTMet321Ser stayed alike to the WT affinity due to changes in the on and off rates (k_{on} and k_{off}) of substrate binding due to the relation:

$$K_M = \frac{k_{cat} + k_{off}}{k_{on}}$$

If k_{cat} decreases, either k_{off} increased or k_{on} decreased, which in turn means that the binding site mutation did alter the rates of arginine binding and also influenced the transporter to perform a slower transport turnover than the wildtype protein.

To identify the binding mode of arginine in the *Gk*ApcT-Met321Ser mutant, crystallisation trials were performed. The mutant was incubated with arginine on ice for 2 h before reconstitution into a MO:Chl mixture. Crystals appeared in similar conditions to those of the wildtype protein. The diamond-shaped crystals of *Gk*ApcT-Met321Ser grew to full size (20 µm in the longest dimension) within two weeks. They were harvested and cryocooled (no cryoprotectant was necessary), and stored in liquid nitrogen until X-ray data collection. A full data set was collected at the I24 beamline, DLS, to a resolution of 3.13 Å (Table A 4). MR using the WT structure without any ligands gave a solution with a TFZ score of 25.8.

Difference density was found at the site where the alanine was bound to *Gk*ApcT (Figure 6-11). After further refinement with BUSTER (Bricogne *et al.*, 2016), model inspection, and adjustment in COOT (Emsley *et al.*, 2010), the difference density in the binding site persisted, with a size similar to that of an arginine molecule (Figure 6-11).

Though the data were of lower resolution (3.13 Å) than the wildtype data set with alanine (2.86 Å), an arginine molecule was modelled into the site of the observed difference density, followed by refinement in BUSTER (Bricogne *et al.*, 2016) after which no negative F_o-F_c density was found in that region giving a good indication that no misplacement of the ligand had been made, and that the difference density had not been misinterpreted. The final structure had an R_{work} and R_{free} of 23.1% and 29.5 %, respectively. Arginine is similarly oriented to the bound alanine in the *Gk*ApcT-WT structure (Figure 6-12A) but with the side chain extending into the protein. Comparison of arginine bound to AdiC (PDB-ID: 3L1L) and *Gk*ApcT-Met321Ser shows two different orientations for the same substrate (Figure 6-12B).

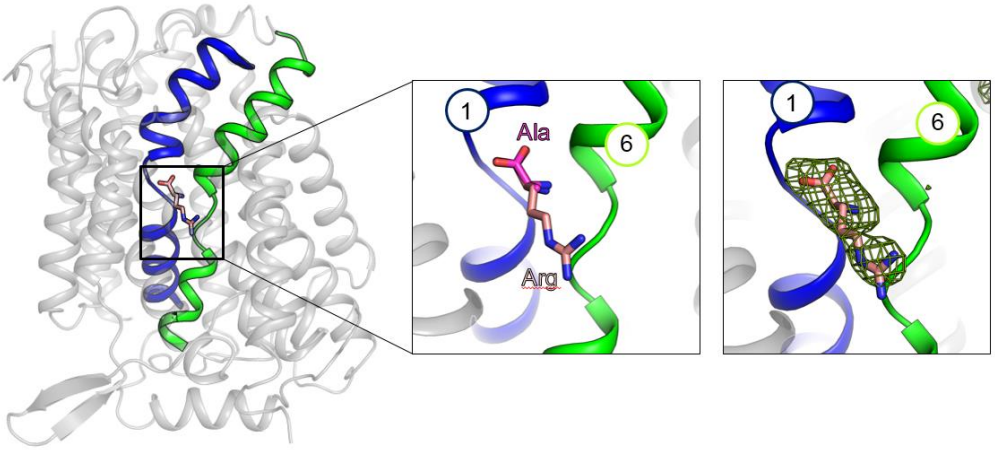


Figure 6-11: Arginine binding to GkApcT-Met321Ser

Location of the arginine bound to *Gk*ApcT-Met321Ser (PDB-ID: 6F34, 3.13 Å) between the two unwound helices TM1 (blue) and TM6 (green) (left). The binding mode of arginine was compared to the alanine molecule of the *Gk*ApcT-WT crystal structure (PDB-ID: 5OQT, 2.86 Å) (middle). During refinement, positive difference density appeared in the binding site (F_o-F_c map in green contoured at 1 σ), which corresponded to a bound arginine molecule (right).

In AdiC, the arginine molecule is captured parallel to Trp202 and Trp293, whereas in *Gk*ApcT-Met321Ser, the arginine side chain is more diagonally arranged. Here, Ser321 sits at a similar position to Trp293 of AdiC, leaving enough space for the side chain to occupy the opened cavity. Additionally, a vertical orientation of arginine as seen in AdiC would not be possible, since this arrangement would be sterically hindered by TM3 (Figure 6-12B).

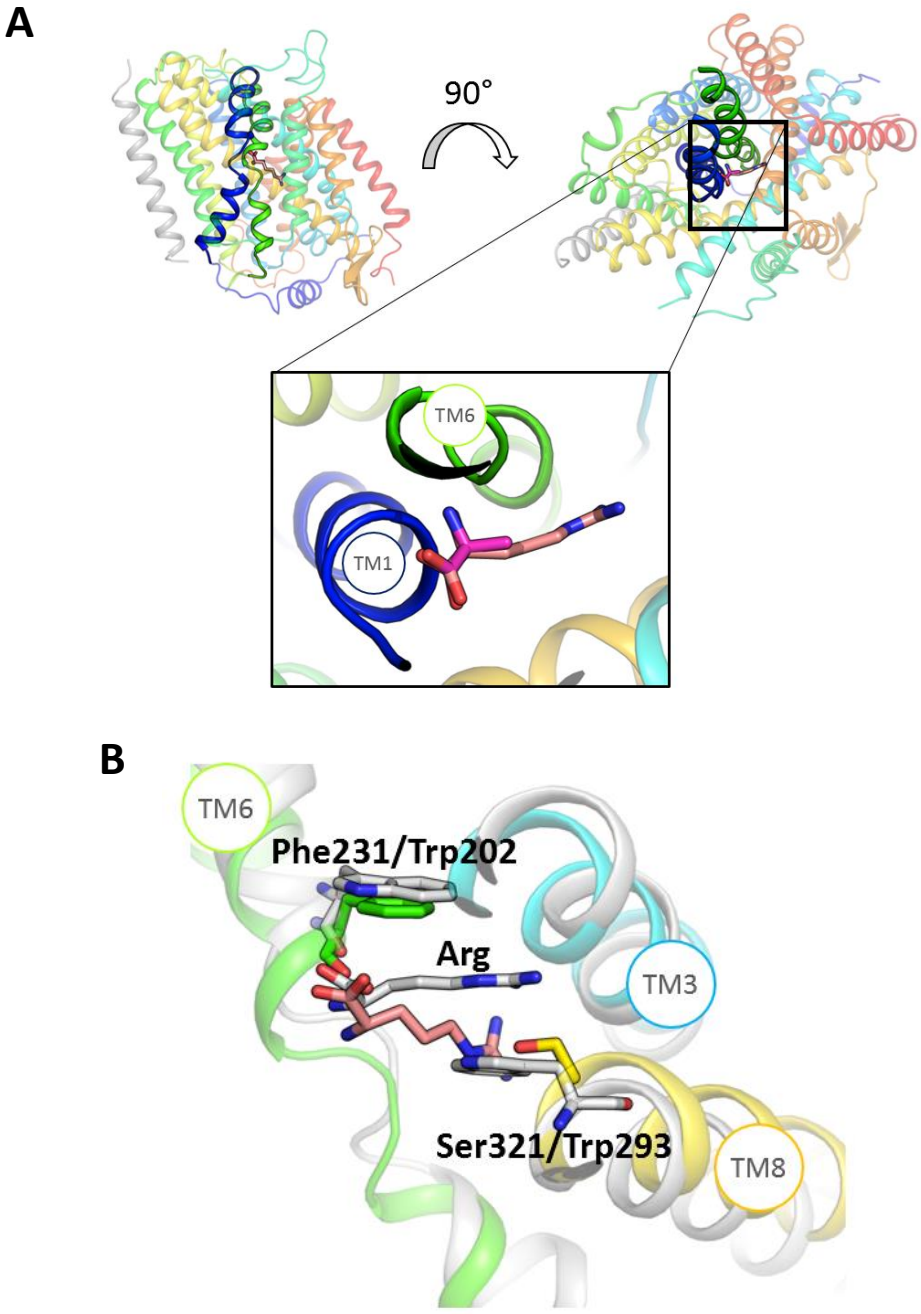


Figure 6-12: Structure of the arginine bound GkApcT-Met321Ser mutant

Comparison between the *Gk*ApcT-WT (PDB-ID: 5OQT, 2.86Å) and *Gk*ApcT-Met321Ser (PDB-ID: 6F34, 3.13 Å) structures shows that L-Ala (pink) and L-Arg (light pink) are similarly bound, and coordinated by TM1 and TM6 (A). AdiC, the arginine-agmatine exchanger, was solved with an arginine in the occluded state (PDB-ID: 3L1L, 3.0Å, grey). The arginine molecule in this structure is aligned parallel to Trp202 and Trp293 whereas the arginine in the *Gk*ApcT-Met321Ser structure shows a diagonal arrangement, due to steric restriction posed by TM3(light blue) (B). This helix is pushed further into the binding site when compared to the structure of AdiC.

6.5 PROTON-COUPLING MECHANISM OF *Gk*APCT

We have identified *Gk*ApcT as a proton-coupled amino acid transporter, which could not be verified for *Mj*ApcT, although proposed to be proton-coupled.
Mutational studies based on the structure of *Gk*ApcT were performed to identify the residues involved in the coupling mechanism, which is discussed below.

6.5.1 The intracellular Glu-Asp pair

Following the observation of the potential helical rearrangement during the transport cycle of *Gk*ApcT, a closer look at intracellular cavity identified two charged residues near the substrate binding site, Asp237 and Glu115. Among the CAT family, Asp237 is highly conserved. Since TM6 is the symmetry-related partner of TM1, we thought it is likely that this helix may be responsible for the release of the substrate into the cell. To investigate the role of the two residues on the transport activity of *Gk*ApcT, Glu115 and Asp237 were mutated to alanine. Additionally, Glu115 was changed to glutamine, and Asp237 to asparagine, and a Glu115Asp/Asp237Glu mutant was designed. Only for *Gk*ApcT-Glu115Gln, could the transport function be restored partially in counterflow experiments and also showed increased thermal stability in presence of alanine (Figure A 11). Together with the inability of this mutant to transport alanine under electrogenic conditions, the experiment suggested that Glu115 is an important residue in the proton-coupling mechanism of *Gk*ApcT. The loss of transport activity for Asp237Asn suggests that Asp237 is important for the function of the transporter. It was surprising that the Glu115Asn/Asp237Glu did not recover the WT function and this indicates that the transporter is sensitive to change in this region or that the interaction between these two residues involves other partners and not only the water molecule as mediator. The interaction with Asp237 through a single water molecule coordinated between both residues (Figure 6-13C) seems to be additionally important for the translocation mechanism of the transporter.
The pKa of Glu115 was predicted to be 8.18 using the PROPKA server (http://nbcr-222.ucsd.edu/pdb2pqr_2.0.0/). Thus, Glu115 is likely to be protonated in the crystal struc-

ture, and a possible mechanism is that the residue becomes deprotonated through exposure to the cytosol on the intracellular side due to the kink movement of TM6 (Figure 6-13B) which probably releases the substrate on the intracellular side.

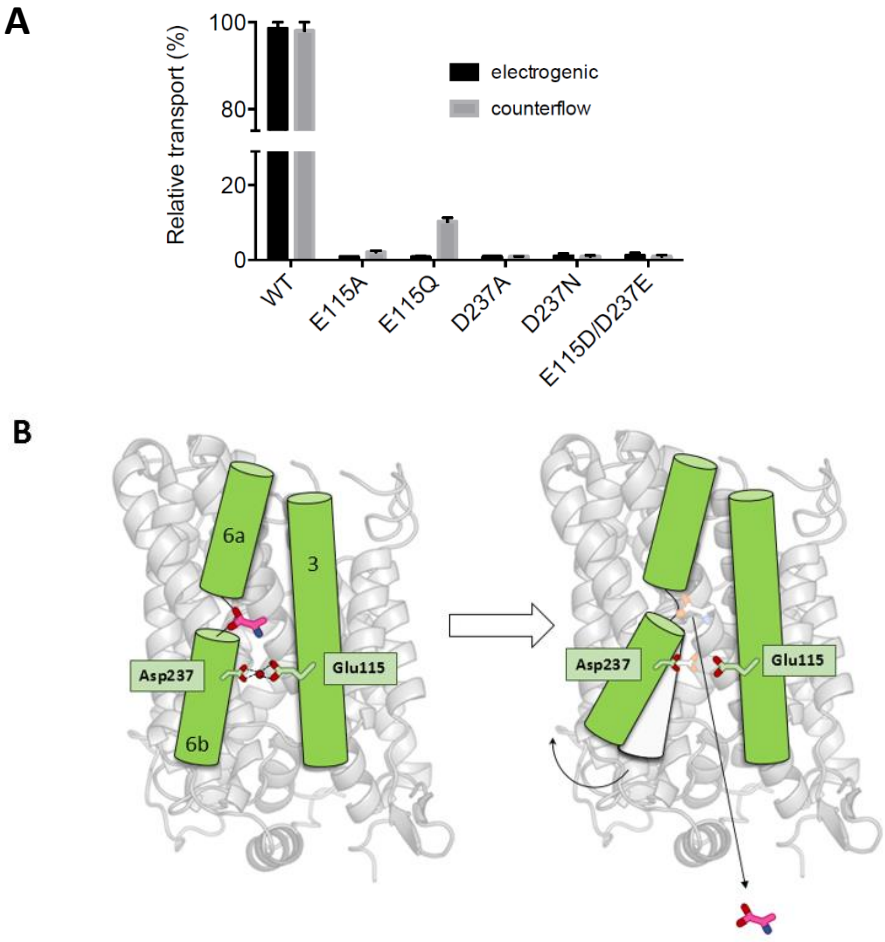


Figure 6-13: Identification of a proton-coupling pathway in GkApcT

Mutational studies on *Gk*ApcT targeting residues proposed to be involved in the proton-coupling mechanism were carried out. Single time-point measurements under proton-driven (uptake after 1 min shown) and counterflow (uptake afer 5 min shown) conditions were performed for each mutant and compared with the wildtype protein (A). Based on the results from the activity measurements, a model for substrate release on the intracellular side via a kink movement of TM6 is proposed (B).

6.5.2 Water filled cavity of *Gk*ApcT

The water molecule interacting with Glu115 and Asp237 on the intracellular side of the substrate binding site forms part of a water network (Figure 6-14) that stretches into the cytoplasm. The water molecules form hydrogen bonds with surrounding charged and polar residues on TM1, 2, 3 and 6 but also with residues of the intracellular loop 1 (IL1). Given the arrangement of the water molecules and the interacting residues in the form of a transient

'zig-zag' water molecule network, proton-release might be facilitated via this proton-relay system (Figure 6-14). The network does not seem to form a continuous stream of water molecules. However, the network can be completed if the polar and charged side chains of the residues highlighted in Figure 6-14A are also involved in this hypothetical proton-relay system (Figure 6-14B).

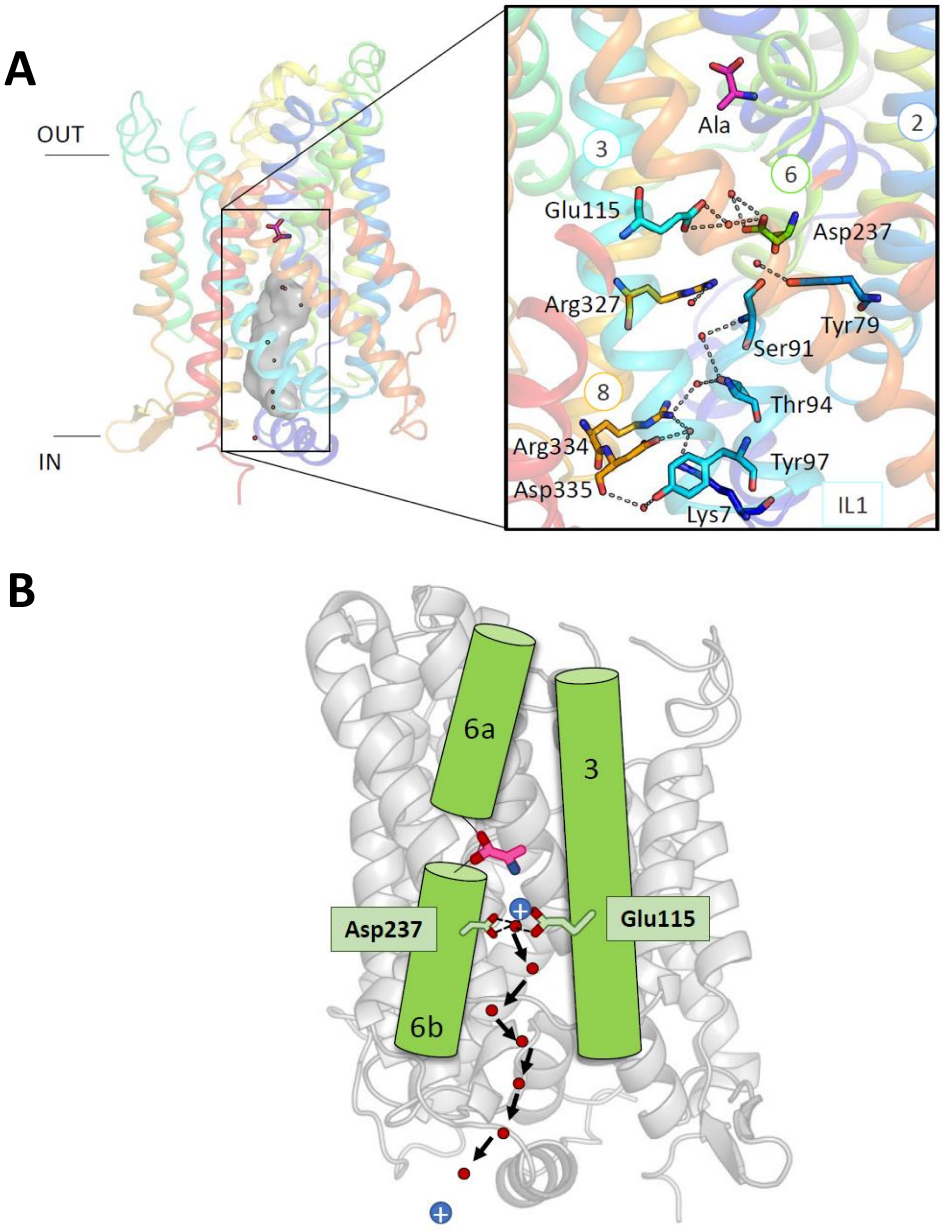


Figure 6-14: Water-filled cytoplasmic cavity

A cytoplasmic water cavity was found in the structure of *Gk*ApcT (PDB-ID: 5OQT, 2.86Å) stretching up to the substrate binding site. The water molecules are coordinated by charged and polar residues surrounding the cavity (A). This arrangement gives rise to the speculation that proton-release is facilitated by this water network (B).

6.5.3 A network of intracellular charged residues influencing the transport mechanism

On the cytoplasmic surface of *Gk*ApcT, in close proximity to the water filled channel, a network of charged residues was found (Figure 6-15). Lys179 and Glu245, as well as Glu244 and Arg327 could form salt bridges with one another, possibly stabilising the protein in one of the transport cycle steps or participating in the proton if the mechanism hypothesised in the previous section holds true. Arg327 has been particularly observed to interact with one of the water molecules in that region (Figure 6-14).

To investigate the role of each residue, alanine mutations were carried out and in addition, the proposed salt bridge residues were swapped. Electrogenic uptake was lost or reduced for most of the mutations. Lys179Ala was the only mutant that retained high transport activity (~ 65% of WT) showing that this residue is not strongly involved in the transport mechanism of *Gk*ApcT. Reduction or inactivation of the transport mechanism by altering residues Glu244, Glu245 and Arg327 indicated that although they are not close to the substrate binding site, they are involved in the function of *Gk*ApcT.

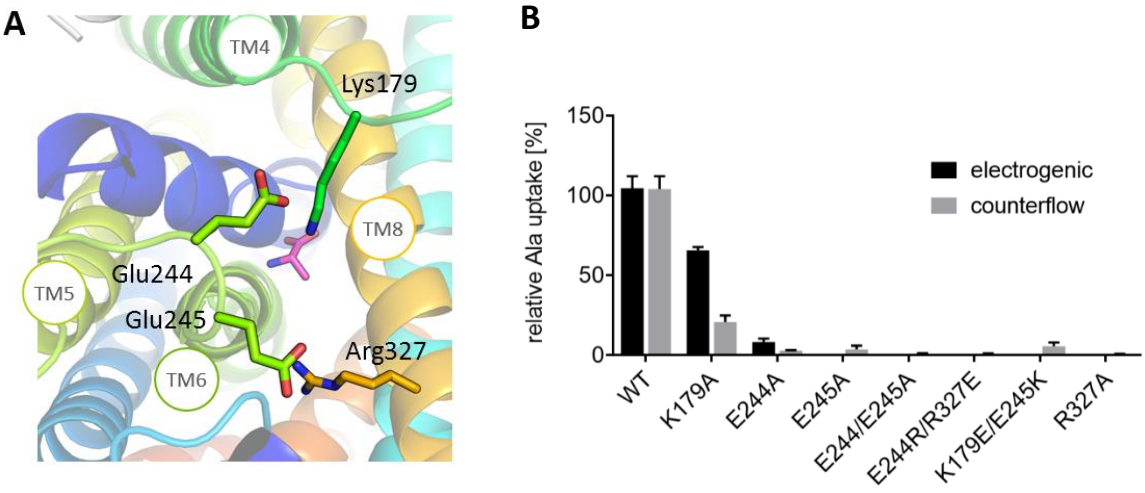


Figure 6-15: Intracellular charged residues

A network of charged residues was found on the intracellular side of *Gk*ApcT (PDB-ID: 5OQT, 2.86Å) that might form salt bridges when the transporter changes its conformational state (A).
Mutational studies on these residues revealed their influence on alanine uptake under electrogenic (internal buffer: 20 mM KPi pH 6.5, 100 mM KAc and 2 mM $MgSO_4$; external buffer: 20 mM NaPi pH 6.5, 100 mM NaCl, 2 mM Mg_2SO_4, 250 nM ^{3}H-alanine, 10 µM valinomycin) or counterflow conditions (internal buffer: 25 mM citrate phosphate pH 6.0, 100 mM NaCl, 2 mM $MgSO_4$, 10 mM alanine; external buffer: 25 mM citrate phosphate pH 6.0, 100 mM NaCl, 2 mM Mg_2SO_4, 250 nM ^{3}H-alanine, 100 µM alanine) using 5 µg of protein per experimental time point (B).

6.6 THE B-HAIRPIN DOMAIN

The CAT2 transporter exists in two isoforms in humans, CAT2A and CAT2B. The two splice variants are identical except for a 42 amino acid stretch that is thought to indirectly influence the substrate affinity and transport rate (Habermeier *et al.*, 2003). Highlighting this region in the sequence alignment between *Gk*ApcT and human CATs, including the two splice variants of CAT2, reveals the homologous region in *Gk*ApcT (Figure 6-16C). In its structure, this stretch partially belongs to a loop region on the intracellular side of *Gk*ApcT. This loop contains an antiparallel β-sheet, situated between TM8 and TM9, and the endpoints of the loop each contain a proline residue acting as 'hinge' for the loop. Comparison of this region between *Gk*ApcT and *Mj*ApcT shows that in *Mj*ApcT the loop region closes the binding site on the intracellular side, whereas in *Gk*ApcT an open conformation is found for the loop (Figure 6-16D). Given the different pHs at which the proteins were crystallised, it could be suggested that the different conformations of the loop result from a protonation/deprotonation process. Another function of this loop region could be the interaction with the lipid bilayer and thus the stabilisation of the transporter in the membrane, given that the loop contains several positively charged residues (Figure 6-16C). Mutations of the prolines on each end of the loop were performed, either to glycine or valine, to disrupt the possible 'hinge' movement of the loop. Both double-mutants were less stable during purification but could be reconstituted. The activity in electrogenic uptake assay was reduced, indicating that the transport mechanism was disrupted by these mutations (Figure 6-16A).

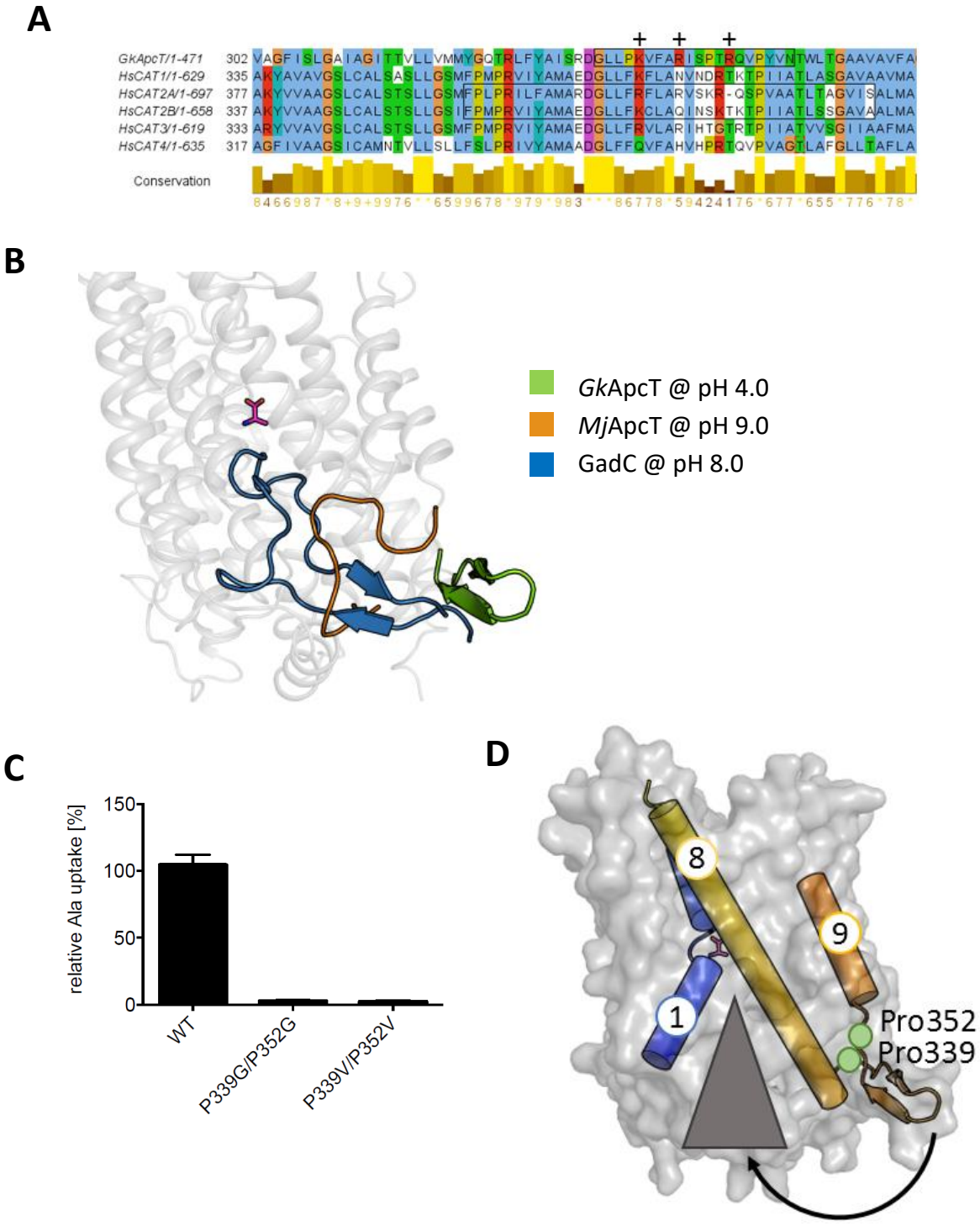


Figure 6-16: Mutational studies on β-hairpin domain

The region where the two splice variants of CAT2 differ partially contains a stretch corresponding to the β-hairpin domain of *Gk*ApcT (A). it might be possible that the hairpin domain interacts with the lipid bilayer, as it contains a pattern of positively charged residues. In comparison to *Gk*ApcT (PDB-ID: 5OQT, 2.86Å, grey and green)), *Mj*ApcT (PDB-ID: 3GIA, 2.32Å, orange) contains an unwound loop region that closes part of the exit path of the substrate, whereas in GadC (PDB-ID: 4DJI, 3.19Å, blue) a C-terminal 'plug' has been found closing the binding site (B). Alteration of the proline residues at the hinge points of the hairpin domain reduces the transport activity under electrogenic conditions, suggesting an involvement of this domain in the transport mechanism (C). It is proposed that this domain might close over the binding site when the transporter changes from the inward-open form to the outward-open confirmation (D).

6.7 INFLUENCE ON SUBSTRATE AFFINITY OF R334

As mentioned in the previous section, the human CAT2 transporter exists as two splice variants. These only differ in a small region of the protein sequence, which indirectly influences the substrate affinity of both proteins (Habermeier *et al.*, 2003). One residue that differs between them is an arginine (CAT2A) and a glutamate (CAT2B) in an otherwise highly conserved region (Figure 6-17B). *Gk*ApcT has an arginine residue at position 334. The structural data reveal that the residue sits on the intracellular end of TM8 forming interactions with water molecules and the adjacent Asp335 (Figure 6-17A), which is highly conserved among the CATs (Figure 6-17B). The change to glutamate is thought to increase the substrate affinity of the transporter (Habermeier *et al.*, 2003), probably through interference of the interaction with Asp. Therefore, Arg334 was mutated to glutamate to investigate if this statement holds true for *Gk*ApcT.

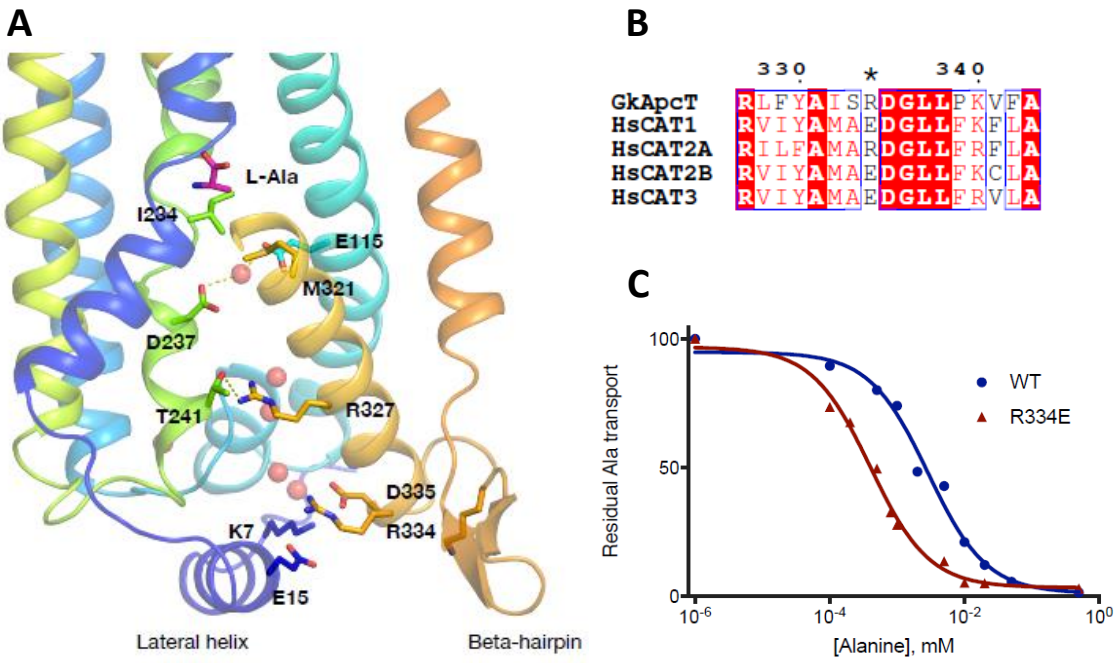


Figure 6-17: R334 residue as possible regulator of GkApcT's activity

The residue Arg334 is located on the intracellular side of the substrate binding site interacting with water molecules in this area, which might form a proton-relay system (A). Additionally, Arg334 forms hydrogen bonds to residues of the lateral helix as well as to the adjacent Asp335 (A) which is conserved among *Gk*ApcT and the human CATs, though Arg334 is either an Arg or Glu in the human CATs (indicated by star above alignment), influencing the substrate specificity of the transporter (B). The IC_{50} curves for alanine were determined for *Gk*ApcT-WT and the Arg334Glu mutant (C).

Indeed, an increased sensitivity to alanine is observed in IC_{50} experiments. The IC_{50} for alanine dropped from 2 µM to 0.4 µM from the wildtype to the mutant protein (Figure 6-17C). This indicates that the change to glutamate influences the affinity for alanine, verifying the results from mutational studies on the human CATs, therefore supporting *Gk*ApcT as a valid model for the CAT transport mechanism.

6.8 The lateral N-terminal helix of *Gk*ApcT

Early stages of analysing the protein sequence of *Gk*ApcT showed that it has an N-terminal extension of 22 residues which is not present in the sequence of *Mj*ApcT (Figure 6-18A).

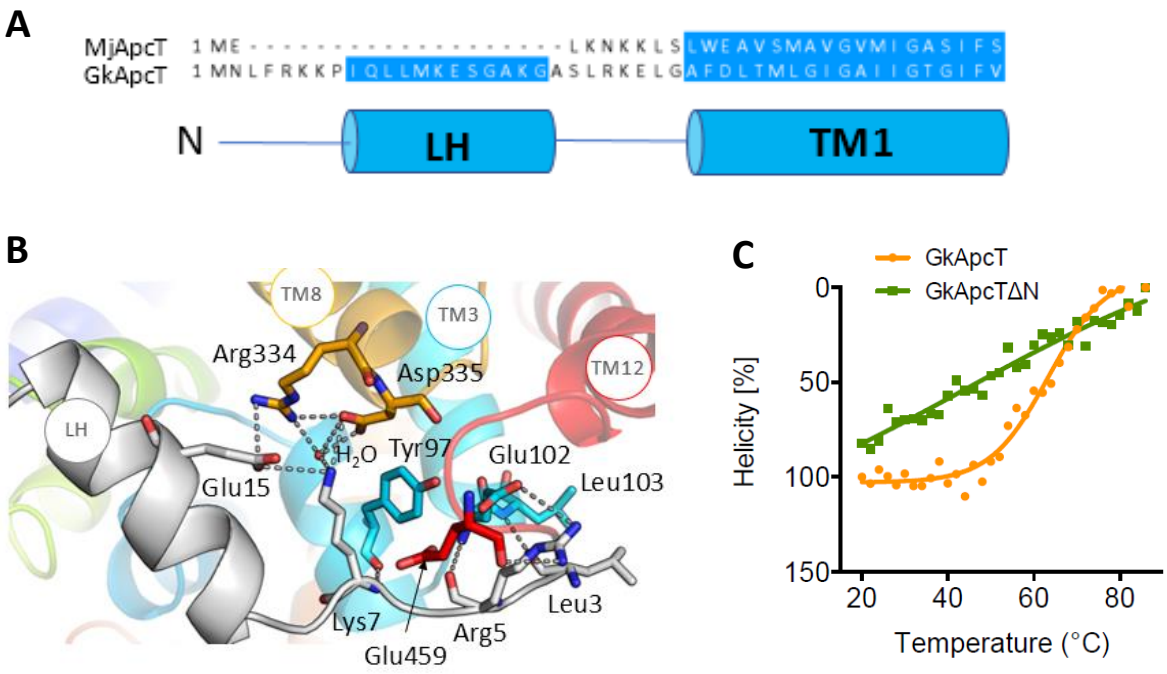


Figure 6-18: Effect on GkApcT N-terminal deletion

The lateral helix derives from the N-terminal extension found in *Gk*ApcT but not in *Mj*ApcT (A) and interacts with residues of TM3, TM8 and TM12 on the intracellular side of the protein (B). An N-terminal deletion variant of GkApcT-WT was purified and the thermal stability of *Gk*ApcT-WT and its N-terminal deleted variant, GkApcTΔN, at 0.1 mg/ml were determined using CD measurements at 220 nm (C).

The N-terminal extension forms an intracellular lateral helix in the structure of *Gk*ApcT, mainly interacting with TM3 and TM8 via hydrogen bonds formed between Arg5 and Glu102, Lys7 and Asp335, Glu15 and Arg334 (Figure 6-18B).

Additionally, a water molecule is coordinated by Lys7, Asp335, R334 and Glu15 (Figure 6-18B). Deletion of the helix caused instability of the protein structure (Figure 6-18C). Due to its connection to TM1, the lateral helix might prevent the 'kink' movement that is observed in LeuT (Krishnamurthy and Gouaux, 2012) which is compensated by the similar movement of the symmetrical related TM6. Thus, this structural element is an important part of the transporter, influencing the overall protein stability and possibly the transporter function.

6.9 Cholesterol interaction

During the optimisation process of the crystallisation conditions for *Gk*ApcT, it was suggested monoolein supplemented with cholesterol might be used as an alternative lipid environment. This composition has mostly been used for crystallising GPCRs and other membrane proteins of higher eukaryotes, as the membrane in which they usually reside contains cholesterol (Dufourc, 2008). Cholesterol is a small molecule that assists in rigidifying the membrane and can help to stabilise proteins in the membrane (Song *et al.*, 2014).

In the case of *Gk*ApcT, the use of cholesterol improved the crystal size, so that the crystals grown in the LCP were bigger with (~20-30 µm in the longest dimension) than without cholesterol (~10-15 µm in the longest dimension), and the overall crystal quality improved (from ~5Å to ~3 Å). During the refinement process, difference density was found between YneM and *Gk*ApcT that appeared to be a bound cholesterol molecule, explaining the improvement of the LCP grown crystals by helping the proteins to maintain contact with each other and in forming crystal contacts. The cholesterol molecule is in a hydrophobic pocket formed by residues YneM-Phe21, Ala24, and *Gk*ApcT-Ala200, and Val201, and is stabilised via pi-pi-stacking to aromatic residues Trp204 and Tyr205 (Figure 6-19). The polar head group faces towards the exterior environment. Interestingly, compounds structurally similar to cholesterol are found in the bacterial but not in the archaeal membranes, so-called hopenes or hopanoids (Welander *et al.*, 2009). They play a vital role in the regulation of cell membrane fluidity and rigidity (Welander *et al.*, 2009; Wu *et al.*, 2015a). These compounds are found

to be additionally important for the membrane to withstand high temperatures, low pHs and resistance to antibiotics (Welander *et al.*, 2009; Malott *et al.*, 2012, 2014). It may explain the interaction of *Gk*ApcT with this molecule, as *G. kaustophilus* is found in these harsh environments and hopanoids would prevent leakage of the membrane. To verify the stabilising effect of cholesterol on *Gk*ApcT, CD-melts were performed with and without the supplementation of cholesterol to the detergent solubilised protein. In addition, *Mj*ApcT and $PepT_{St}$ were used as negative controls.

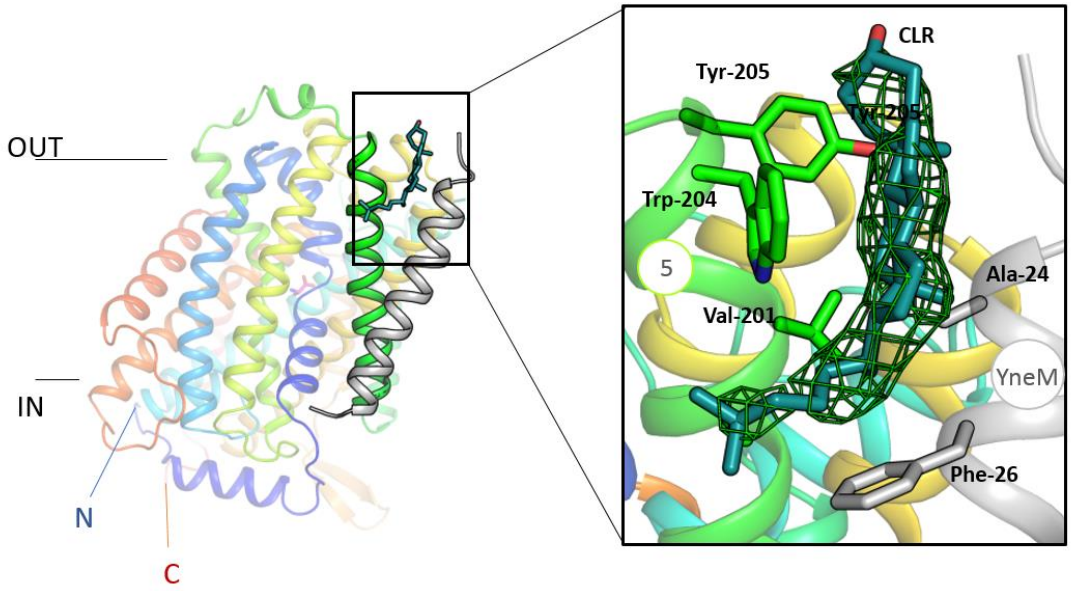


Figure 6-19: Cholesterol binding pocket

Cholesterol is located between TM5 (green) of *Gk*ApcT (PDB-ID: 5OQT, 2.86Å, rainbow) and YneM (grey) in a hydrophobic pocket between both helices. The cholesterol interacts with Tyr206 via Π-Π-stacking. The positive F_o-F_c map contoured at 2.8 σ is shown around the cholesterol molecule.

The melting curves show a noticeable shift of ~15 °C in the melting temperature of *Gk*ApcT when 0.006 % (w/V) cholesterol hemisuccinate is added to the protein solution (57.4°C (DDM) to 72°C (DDM+ CHS)). This result could arise from the cholesterol supporting the interaction between *Gk*ApcT and YneM, and thus influencing its stability. No effect in the presence of DMPC:CHAPSO bicelles was observed.

Interestingly, there was no effect on *Mj*ApcT (T_m(DDM)= 61.9°C, Tm(DDM+CHS) = 62.2°C), which could be a result of using OG as detergent and not DDM, or that CHS does not interact with *Mj*ApcT. OG was chosen in this case to enable a melting curve to be collected, since in previous experiments, *Mj*ApcT did not show a suitable melting profile in DDM (Figure 6-21). Another explanation for the lack of a stabilising effect of CHS on the *Mj*ApcT could be that

it is reported that archaeal organisms, to which *Methanocaldococcus jannaschii* belongs, do not contain hopanoids or similar compounds in their membrane. Therefore, it is possible that cholesterol simply has no effect.

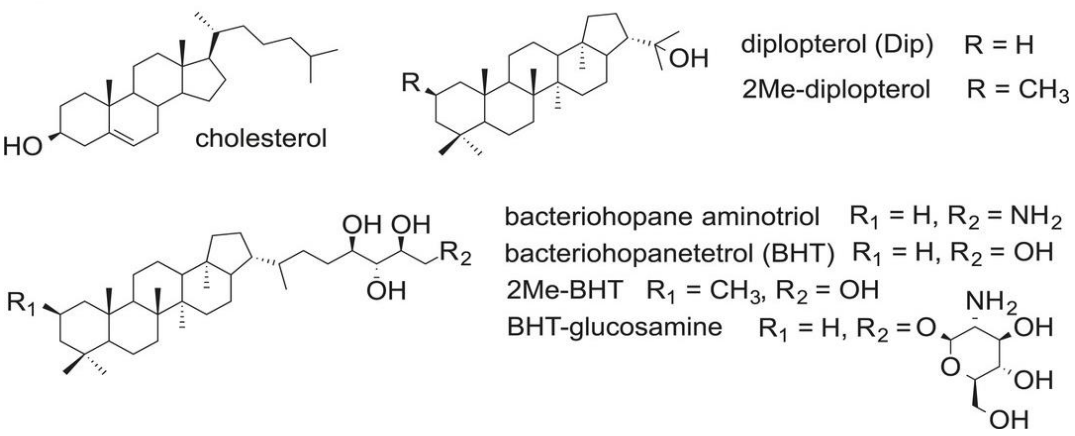


Figure 6-20: Structures of cholesterol and hopanoids

Figure adapted from Wu et al., (2015)

Cholesterol also had only little effect on the stability of $PepT_{St}$, where a shift of 5°C (Tm (DDM) = 54.9°C, Tm (DDM+CHS) = 61°C) was observed. $PepT_{St}$ belongs to the MFS family and has a different overall fold, and has not been reported to interact with cholesterol or similar compounds. Nevertheless, the slight increase in the melting temperature suggests that the protein is also able to interact with cholesterol but without such a significant effect as seen for *Gk*ApcT. It might therefore be of interest to test other APC family homologues on their possible interaction with cholesterol. Furthermore, the use of cholesterol supplemented monoolein for LCP crystallisation trials should be an additional tool in testing different lipids and *Gk*ApcT has set an example that the use of cholesterol in LCP crystallisation should not only be restricted to eukaryotic proteins.

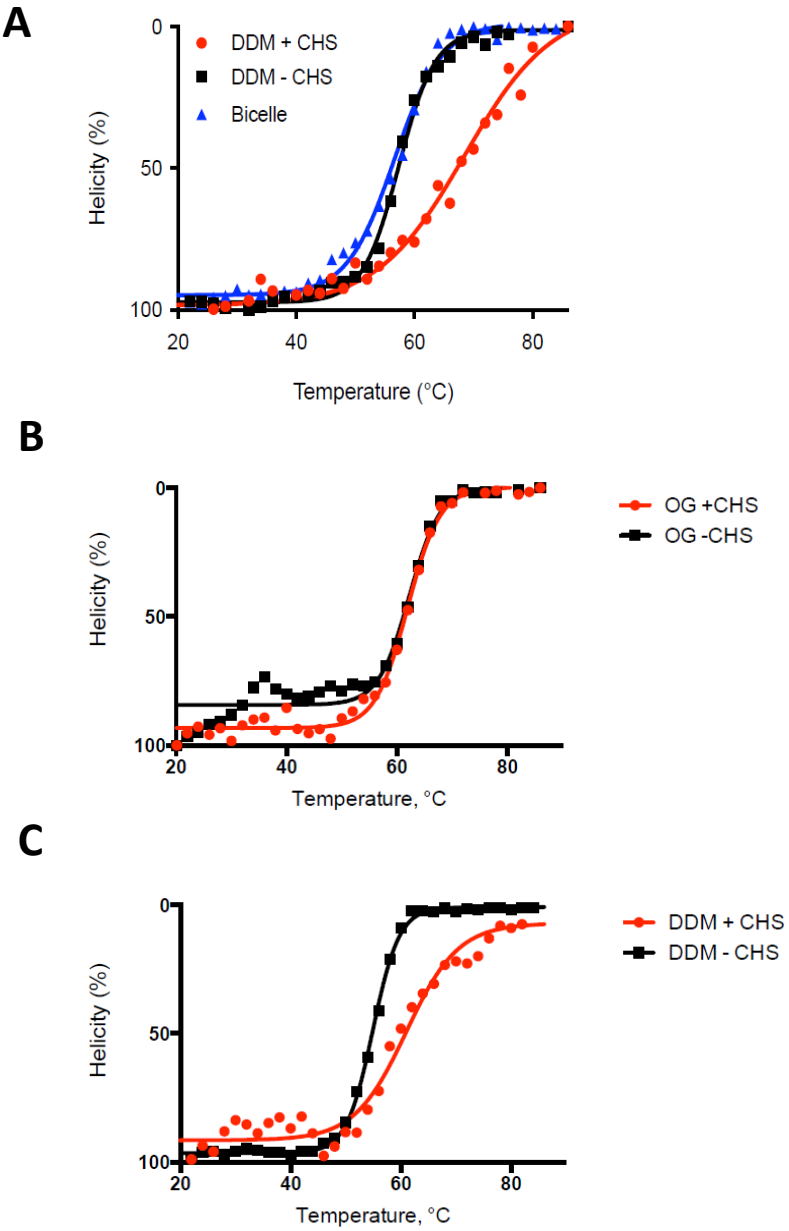


Figure 6-21: The influence of cholesterol on thermal stability of membrane proteins

CD melts from 20-86 °C were performed for *Gk*ApcT (A), *Mj*ApcT (B) and $PepT_{St}$ (C) in detergent without any additives or with cholesterol hemisuccinate (CHS) or in the presence of DMPC:CHAPSO bicelles in the case of *Gk*ApcT. The proteins were diluted 1:100 into buffer (10mM KPi pH 6.5, 50 mM $NaSO_4$ and either 0.03 % DDM (DDM-CHS), 0.03 % DDM/0.006% CHS (DDM+CHS), or 0.03 % DDM/0.07 % DMPC:CHAPSO (2.8:1) bicelles (Bicelle)) to a final concentration of 0.1 mg/ml.

6.10 YneM association

As stated previously, the crystal structure of *Gk*ApcT revealed an association to an additional protein, YneM. During initial analysis of the difference electron density close to helix one, it was assumed that this helix belonged to the N-terminal extension of 23 additional amino acids compared with the *Mj*ApcT protein sequence. Secondary structure prediction identified this N-terminal extension to be a helix (Section 6.8). Modelling of the sequence into the density and subsequent refinement did decrease the *R*-factors during refinement, but detailed inspection of the density around this region revealed that the sequence was not correctly built. It was then concluded that the orientation of the helix had to be altered such that the N-terminus of *Gk*ApcT would have been in the same position as the N-terminus of the additional helix. Therefore, the additional helix could not be a part of *Gk*ApcT. Given that *Gk*ApcT was recombinantly expressed in *E. coli*, and that the pWaldo plasmid only carried *Gk*ApcT, it was concluded that the helical protein was derived from *E. coli,* and was co-purified and co-crystallised with *Gk*ApcT.

A poly-Ala model was initially fitted into the additional density and after subsequent refinement cycles, features of the side chains of aromatic residues appeared. From initial observations and the anomalous difference Fourier map revealing the sulphur atoms, it was possible to build a partial model of the helix and use the sequence of this model for a BLAST search against the proteome of *E. coli*. YneM was identified as a hit. Secondary structure prediction of the 31 amino acid long protein revealed that it contains a transmembrane helix identical to the length of the helix built into the *Gk*ApcT density. The sequence of the helix was completed and the refined *Gk*ApcT model showed no errors for the model at the YneM helix. It can thus be assumed with great confidence that YneM is the protein that is attached to *Gk*ApcT.

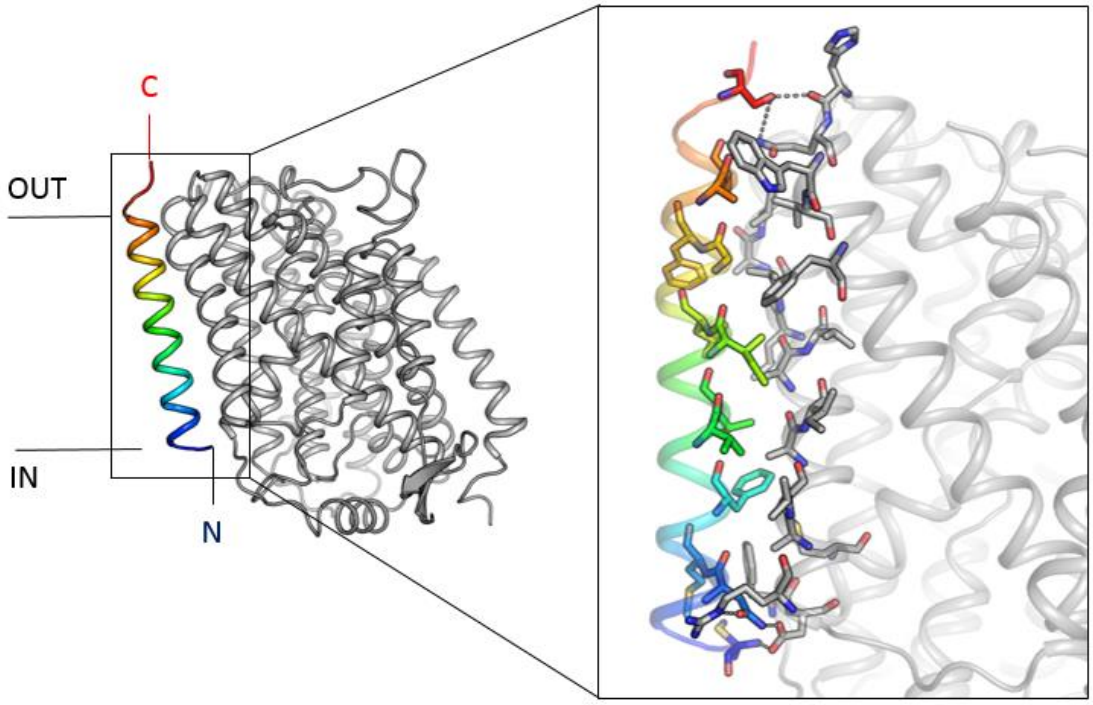


Figure 6-22: YneM is associated to GkApcT

Left: The structure of *Gk*ApcT (PDB-ID: 5OQT, 2.86 Å, rainbow) and YneM (grey) were solved by crystallisation in LCP. The N- and C-terminus of *Gk*ApcT are intracellular as well as the N-terminal of YneM. YneM contains a single membrane spanning helix (coloured from N- to C-terminus) while *Gk*ApcT consists of 12 transmembrane helices. Right: YneM (rainbow) interacts via hydrophobic interactions with *Gk*ApcT (grey). Hydrogen bonds (grey dashed lines) between both proteins are formed at the N- and C-terminus of YneM.

There is very little known about YneM. This is mainly due to the failure to annotate small peptides and proteins in *E.coli* and also of other organisms in the past. In 2009, more coding regions of *E.coli* were starting to be annotated for these small proteins (Alix and Blanc-Potard, 2009). In *E.coli*, over 30 proteins smaller than 35 residues have been identified, whereas in *G. kaustophilus*, only 5 such proteins are annotated. Further investigation of these small proteins revealed their participation in response to temperature stress on the organism (Alix and Blanc-Potard, 2009; Hemm *et al.*, 2010). When the cell is stressed, these small helical proteins interact with membrane proteins and stabilise them. Additionally, some of the genes encoding for the small transmembrane proteins are close to the genes coding for membrane transporters. It is assumed that YneM fulfils a stabilising function for *Gk*ApcT. A BLAST search with the YneM sequence using the proteome library of *G. kaustophilus* was not able to identify a related protein, but this could have been due to the limited

annotation of such proteins in this organism. Furthermore, the sequence of YneM is characterised by the hydrophobic residues that build the helix. One could imagine that the helix could undergo many mutations to other non-polar residues without disruption of the helical structure, preventing the identification of other small similar helical proteins. The density at this region was very well resolved and no other density near the protein was observed, which would suggest that YneM specifically interacts with *Gk*ApcT in this region. It has been reported that the LAT transporters of the SLC7 family interact with SLC3 proteins, consisting of a soluble heavy subunit and one transmembrane helix, via a disulphide bond between the loop region of the LATs between helices 3 and 4, and the extracellular loop between the transmembrane helix and the heavy subunit of SLC3 (Figure 6-23B). Based on docking studies using a homology model of LAT2, derived from AdiC, it was suggested that the single transmembrane helix is located close to TM4 of LAT2 (Rosell *et al.*, 2014). Given the close relation of *Gk*ApcT to the SLC7 family, we propose that the region where YneM interacts with *Gk*ApcT might be the interaction zone of the LATs with the SLC3 proteins. Based on a homology model of LAT1 generated with the crystal structure of *Gk*ApcT, the corresponding Cys of the SLC7 protein forming the disulphide bridge is located in a long extracellular loop. Although this loop is distant to the helix of 4F2hc (SLC3 family member, based on YneM here), it might be long and flexible enough to form the disulphide bond as we propose in the LAT1/4F2hc model (Figure 6-23B).

Interestingly, while the sequence alignment between YneM and the single TM helices of rBAT and 4F2hc (SLC3 family) shows very little residue-specific conservation between them, the electrostatics are, however, somewhat similar (Figure 6-24). These results might give an indication that the position of YneM relates to that of the helices of rBAT or 4F2hc in respect to the HAT transporter. Still, further experimental data are required to verify this theory.

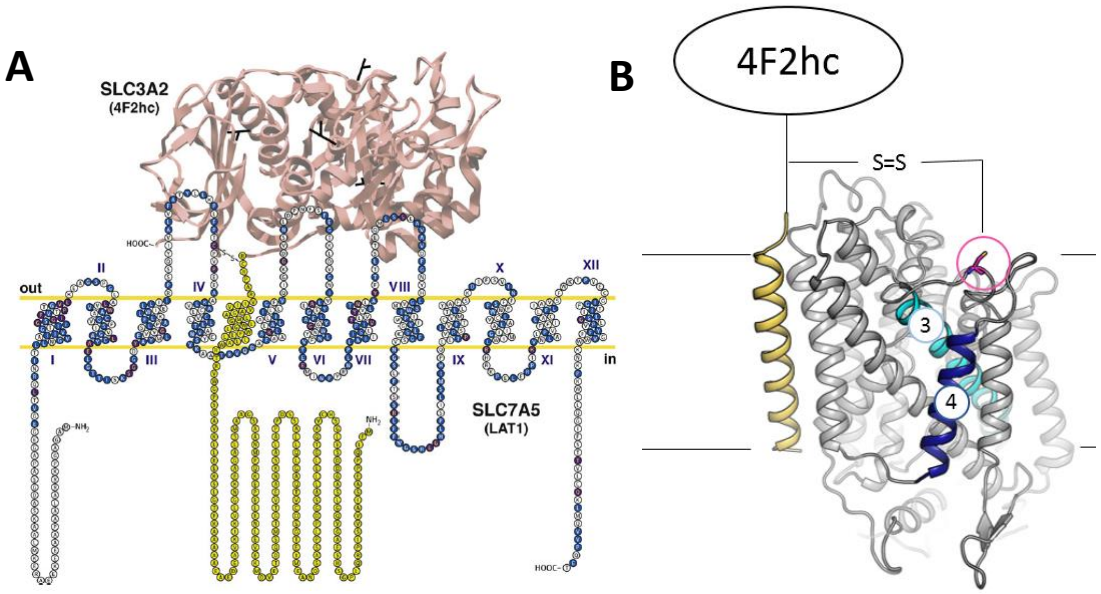


Figure 6-23: Possible model for 4F2hc interaction with LAT1

The topology model of LAT1 shows the interaction region of 4F2hc and LAT1 and the proposed location of the single TM of 4F2hc relative to LAT1 (A). Using YneM as a template, the transmembrane helix of 4F2hc (yellow) was generated using MODELLER and a model of LAT1 (grey) based on the *Gk*ApcT structure (PDB-ID: 5OQT, 2.86Å) (B). The interaction between 4F2hc and LAT1 is formed by a disulphide bridge (marked with S=S). The cysteine residues (pink) of LAT1 involved in this interaction sits in a loop between helix 8 (light blue) and helix 9 (dark blue). The cysteine residue of 4Fh2c resides in an extracellular loop region between the single TM and the extracellular domain and is not marked in the figure.

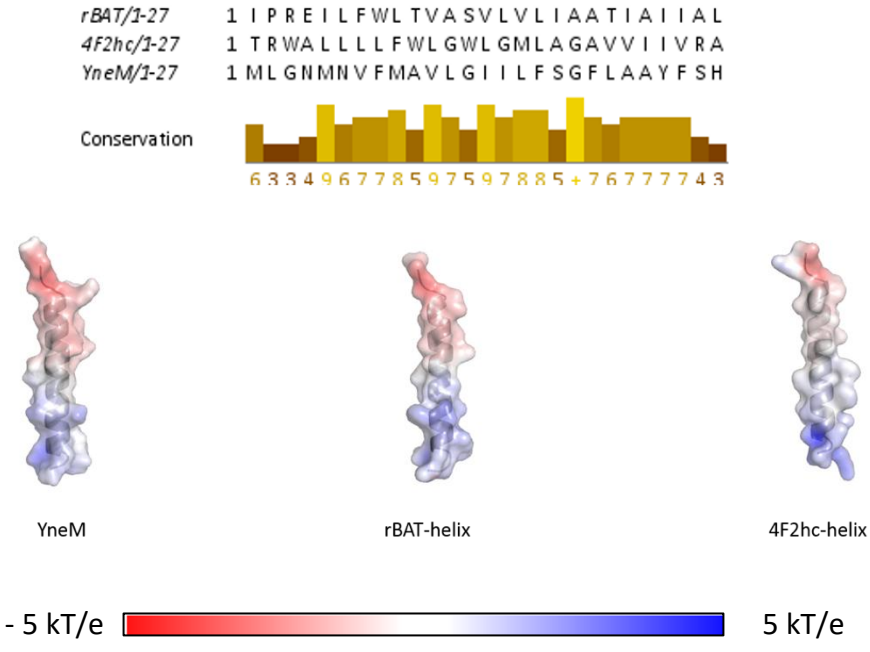


Figure 6-24: Similarities between YneM and the single transmembrane helices of rBAT and 4F2hc

The alignment shows very little conservation but shows that the amino acid composition between YneM and the single TMs of rBAT and 4F2hc are similar. Electropotential calculations using the APBS plugin in PYMOL show similar isosurface (coloured from red (-5 kT/e) to blue (5 kT/e)) for all three helices.

6.10.1 Impact of YneM on the function and stability of *Gk*ApcT

To investigate the effect of the interaction of YneM on the function and stability of *Gk*ApcT, an *E. coli yneM* knock-out (KO) C43(DE) strain was generated. *Gk*ApcT was expressed in the C43(DE3)Δ*yneM* strain. The quantity of over-expressed protein was similar to expression trials in WT C43(DE3) cells and the protein could be purified (Figure A 13).

Additionally, YneM was cloned into the pWaldo-GFP plasmid and successfully purified by Margaret Young (OxIon DPhil rotation student) (Figure A 14). Pure *Gk*ApcT from C43(DE3)Δ*yneM* was concentrated to 20 mg/ml and diluted into CD buffer supplemented with equimolar amounts of YneM or without the addition of YneM. Thermal stability assays using CD were performed in buffer containing either DDM or DDM+CHS, with a total protein concentration of 0.1 mg/ml for thermal stability measurements (Figure 6-25).

The CD measurements revealed that the loss of YneM had no effect on the thermal stability of *Gk*ApcT in DDM (T_m (DDM) = 57°C) and that cholesterol also had no further effect on the thermal stability (T_m (DDM+CHS) = 57°C). However, the stabilising effect observed in the protein purified from the WT C43(DE3) cells could be recovered when YneM was added to the protein, increasing the melting temperature from 57°C to 80°C. Based on the crystal structure of *Gk*ApcT, YneM interacts with *Gk*ApcT mainly via hydrophobic interactions and via a cholesterol molecule between the two proteins. Thus, the observed effect of cholesterol on the stability of *Gk*ApcT might originate from an enhanced interaction between *Gk*ApcT and YneM.

*Gk*ApcT from C43(DE3)Δ*yneM* was further reconstituted into POPE:POPG (3:1) liposomes for functional studies. The proteoliposomes were used for electrogenic assays to compare the activity of this protein with *Gk*ApcT (Figure 6-26).

The protein showed a lower transport activity for alanine than *Gk*ApcT from WT C43(DE3) cells under electrogenic conditions (Figure 6-26 A) although initial rates of transport were similar (Figure 6-26 B). The reduced activity might come from a stabilisation effect of YneM mediated through lipids between *Gk*ApcT and YneM in the liposomes that influence the activity of the transporter during the proton-coupling transport, but further investigation of this will have to be performed in the future.

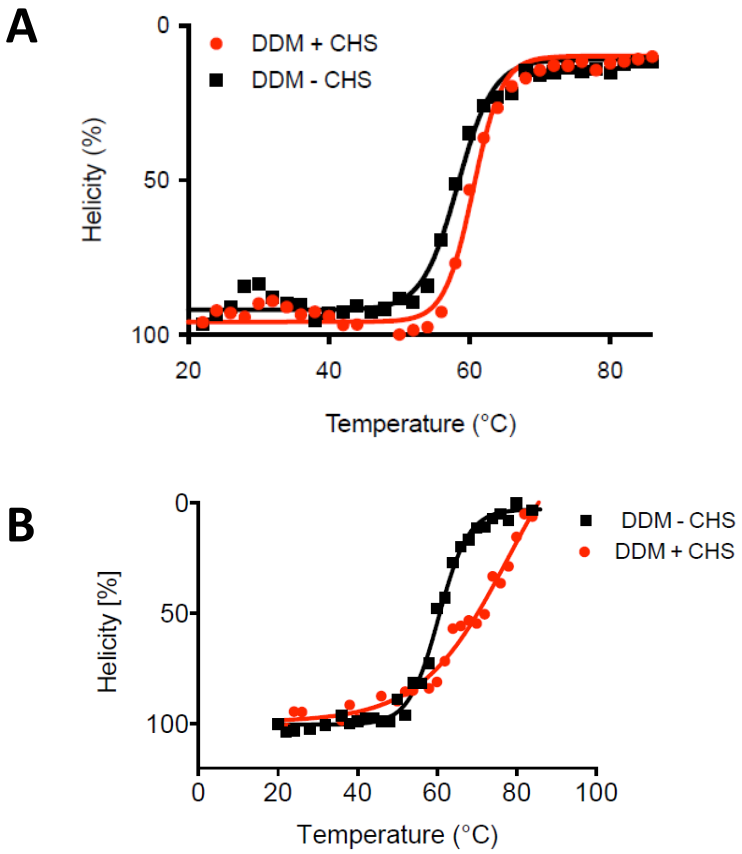


Figure 6-25: Thermal stability of GkApcT-C43 and GkApcT-C43ΔYneM

CD measurements at temperature intervals from 20-86°C were carried out for *Gk*ApcT-C43Δ*yneM* (A) and *Gk*ApcT-C43Δ*yneM* + YneM at 0.1 mg/ml final protein concentration (B). The proteins were diluted into buffer containing 10mM KPi pH 6.5, 50 mM $NaSO_4$ and either 0.03 % DDM (DDM-CHS), or 0.03 % DDM/0.006% CHS (DDM+CHS).

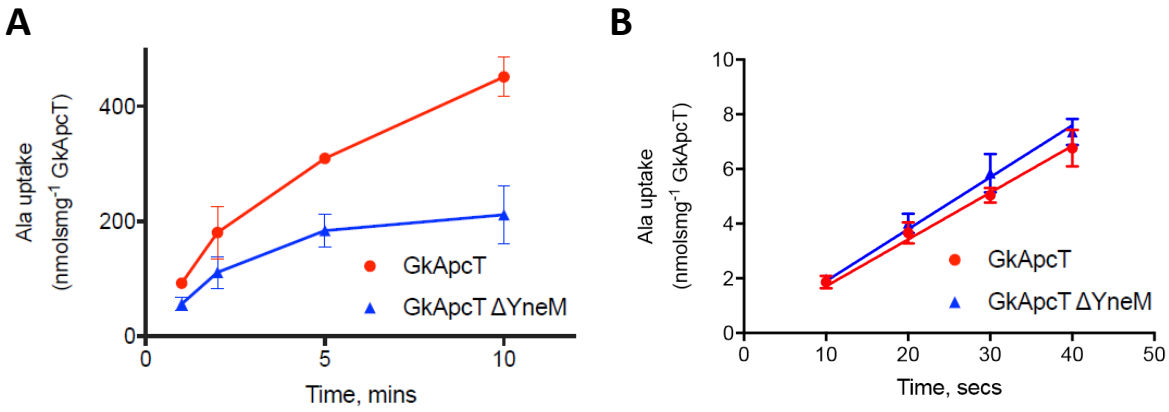


Figure 6-26: Influence of YneM on the transport activity of GkApcT

Alanine uptake under electrogenic conditions was monitored over time for *Gk*ApcT-C43 and *Gk*ApcT-C43Δ*yneM* using 5 µg of protein per experimental time point (A). Initial rates of transport were determined for 1.25 µg of protein per experimental time point (B) under the same experimental conditions used in (A).

6.11 Implications for eukaryotic CATs and PATs based on *Gk*ApcT

The mutational studies presented here showed that *Gk*ApcT is a reasonable structural and functional model for the human CATs.

The sequence alignment between *Gk*ApcT and the CATs from *Arabidopsis thaliana* revealed the conservation of the Glu-Asp pair, thought to be involved in the proton-coupling mechanism of *Gk*ApcT (Figure 6-13). However, the human CATs are not known to be proton-coupled, although their ancestors in *A. thaliana*, *At*CAT6 and *At*CAT1 have been proposed to use protons to drive substrate uptake (Frommer *et al.*, 1995; Hammes *et al.*, 2006). *At*CATs also contain this Glu-Asp pair, or an Asp-Asp pair for *At*CAT2 and *At*CAT6 (Figure 6-27). *At*CAT8 which has been revealed to be proton-independent also carries this residue pair (Figure 6-27).

A more distantly related family to *Gk*ApcT known to be proton-coupled are the human PATs of the SLC36 family. Sequence alignment between *Gk*ApcT and the four *h*PATs showed that the Glu-Asp is switched to an Asp-Glu pair where the glutamate is located three residues upstream from *Gk*ApcT's Asp237 (Figure 6-27) which corresponds to one helix turn. The homology model of *h*PAT1 based on the structure of *Gk*ApcT shows that the Asp-Glu pair are still in proximity to each other (Figure 6-27).

Further investigation of the *At*CATs revealed some differences in substrate specificity, *At*CAT1 and *At*CAT5 are specific for Arg and Lys (Su, 2004), whereas *At*CAT6 recognises large uncharged amino acids (Hammes *et al.*, 2006) and *At*CAT8 transports Gln (Yang *et al.*, 2010). Based on the finding that the mutation of Met321 in *Gk*ApcT to serine in CATs was responsible for modulating the substrate specificity towards arginine, an attempt was made to correlate the variation at this position in the *At*CATs to their substrate profile. *At*CAT1, *At*CAT5 and *At*CAT8 carry a glycine at this location, which would extend the substrate binding site to accommodate arginine and lysine and still provide a polar environment for substrate coordination as serine did in the *Gk*ApcT-Met321Ser mutant. In *At*CAT6 an alanine is located in this region, leaving the substrate binding site hydrophobic but still increasing its size, which would result in the recognition of larger amino acids. However, *At*CAT4, as in the human CATs, has a serine where Met321 resides in *Gk*ApcT, but has shown no recognition of arginine or lysine in experiments (Su, 2004).

Therefore, it seems as if Met321 is a key determinant of the substrate specificity of *Gk*ApcT and the homologous CATs. However, some differences cannot be explained by this residue alone, meaning that other residues in the substrate binding site are involved in this, too. More experimental data such as mutational studies and functional *in vitro* assays are necessary to understand how the substrate variation between the *At*CATs and *h*CATs is achieved but also the function of the Asp-Glu pair in *At*CATs and PATs requires more experimental attention.

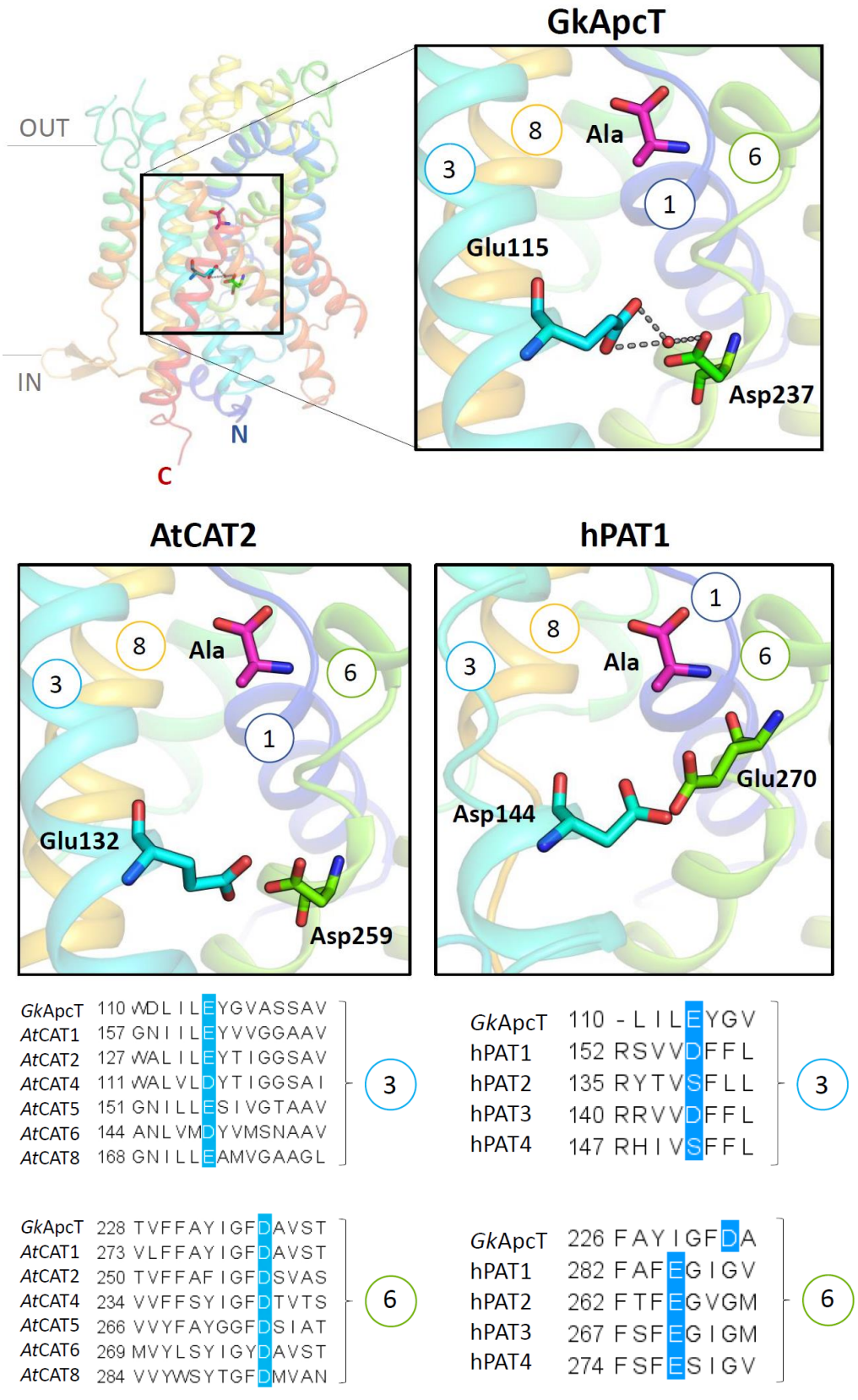


Figure 6-27: Glu-Asp-pair is conserved in other proton-coupled APC transporters

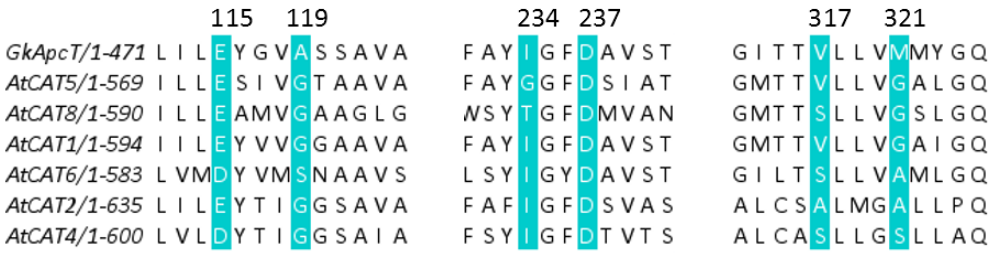


Figure 6-28: Substrate specificity of AtCATs

In the sequence alignment between *Gk*ApcT and *At*CATs a selection of the investigated residues of this thesis are highlighted to investigate if the difference in substrate specificity of the *At*CATs can be derived from the findings on *Gk*ApcT.

6.12 CONCLUSIONS

This Chapter has given an overview about the functional studies performed on *Gk*ApcT based on its structure and comparison to the SLC7 family. It enabled insight into its transport mechanism although the information on the transporter still leaves many questions unanswered. Together with the structural data, it was possible to identify possible residues important for the function of this transporter and to verify the hypotheses using liposome-based assays. Although showing high sequence identity with the human CAT transporters (~40 %), the substrate profile differed. Based on the structure of *Gk*ApcT and comparison to the sequences with the human CATs, various mutations in the binding site were proposed and assayed. A single *Gk*ApcT variant of Met321Ser switched *Gk*ApcT's specificity towards arginine, the major substrate of the human CATs. Taking the research on this mutant further, a crystal structure with a bound arginine was solved. Still, it was possible to verify that the serine at this location allowed the modification of the substrate binding site to fit the arginine molecule and to make the protein into an arginine transporter. This result shows that only minor changes in the binding site have a big influence on the transporter which is probably a reason why the sequence identity between *Gk*ApcT and the human CATs shows only a few variations.

Further comparison with the human CATs verified the influence of Arg334 on the transport function although this residue is distantly located from the substrate binding site. Mutation

of the equivalent residue in the human CATs resulted in a change of substrate affinity (Habermeier *et al.*, 2003) which was also the case in the studies here. Changing the arginine to a glutamate, as found in the high affinity CAT transporters CAT1, CAT2B and CAT3 (Figure 6-17B), increased *Gk*ApcT's affinity to alanine. It is not yet known how this residue influences the affinity of the transporter but we propose that the translocation mechanism is affected by this mutation and thus changes the transport kinetics.

Furthermore, the proton-coupling mechanism of *Gk*ApcT was investigated resulting in the identification of an important Glu-Asp pair below the substrate binding site. The mutational studies proposed Glu115 to be the key component for proton-coupling of *Gk*ApcT. Additionally, a water network forming a connection into the cytoplasm suggested a proton-relay system to participate in proton-release.

The structure of *Gk*ApcT additionally revealed several secondary structure features such as the β-hairpin domain and the lateral helix, as well as the associated protein YneM. Investigations on the influence of these on *Gk*ApcT's activity were performed and showed that the β-hairpin and the lateral helix are vital for the stability of *Gk*ApcT. YneM only showed a stabilising effect when cholesterol was present in the detergent buffer. Since *Gk*ApcT and YneM coordinate a cholesterol molecule between them, the results verified that this interaction is necessary for both proteins to interact closely with each other. Based on these results, we propose that the single TM of the SLC3 proteins interaction with the human LAT transporters is enhanced with cholesterol present. It keeps both proteins in close proximity for disulphide-bridge formation, necessary for correct localisation of the LAT protein. Nevertheless, the structure of *Gk*ApcT not only added a piece to the puzzle of where the substrate binding site might be located for proton-coupled amino acid transporters of the APC family, but also provided a structural template for the SLC7 family of APC transporters to which *Gk*ApcT shows a high sequence identity. Although they differ in their substrate specificity, it is now possible to identify the residues important for this difference and to understand why the two splice variants of CAT2 differ in their substrate affinity.

7 Serial crystallography on LCP grown crystals

7.1 SERIAL CRYSTALLOGRAPHY

Data collection from a series of different crystals is termed serial crystallography (SX). The general idea behind this method is to obtain the maximum diffraction signal from the crystal before the diffraction quality is diminished by radiation damage. This becomes even more important when collecting data on crystals that are very radiation sensitive or contain metal clusters that are readily reduced or disordered even at low X-ray doses (Garman, 2010). A continuous supply of new crystals during data collection is realised by novel injection-based or fixed-target sample delivery. The method of SX has become widely used at the new X-ray sources, the X-ray free electron laser (XFEL) facilities, where serial femtosecond X-ray crystallography (SFX) experiments on protein crystals are performed. At these, short X-ray pulses in the femtosecond range (~10-150 fs) at high brilliance (10^{12} photons/pulse) enable the collection of data on single crystals before their complete destruction (Chapman *et al.*, 2006; Barty *et al.*, 2011). SFX experiments have become popular for systems where only very small crystals (nm-µm range) can be obtained during crystallisation trials and for which conventional data collection is thus impractical. Before the new XFEL sources made such experiments feasible, usually intensive optimisation was carried out to obtain larger crystals for data collection. Interestingly, in some of those cases it was possible to extend the resolution of the initial small crystals (1-5 µm in the longest dimension) beyond the resolution reached on the optimised larger crystals (50-400 µm in the longest dimension) (Boutet *et al.*, 2012; Fenalti *et al.*, 2015; Zhang *et al.*, 2015). The impact of the short X-ray pulses on the protein crystals has been simulated, analysing the ionisation effect on the sample after the incident X-ray pulse. A pulse of 10 fs was found to be the minimum pulse length for collection of a diffraction pattern before structural or global radiation damage can occur and be recorded, referred to as 'diffraction-before-destruction' since the samples vaporises due to the high intensity of the X-ray pulse (Chapman *et al.*, 2006; Barty *et al.*, 2011).

7.1.1 Comparison of SX with conventional synchrotron data collection

Conventional data collection at a synchrotron source is based on collecting diffraction data at cryotemperatures (100 K) on a single rotating crystal, usually using rotation angles in the range of 0.05-1° per image. Through combining these images, Bragg peaks are obtained, making it possible to estimate the structure factor amplitudes accurately. The development of brighter synchrotron X-ray sources and smaller focal spots result in higher flux densities in which the crystal can suffer serious radiation damage during data collection (Schulze-Briese *et al.*, 2005) especially if, due to low symmetry, more data have to be collected from the crystal in order to acquire a complete data set. Initial information such as the crystal symmetry and the unit cell dimensions of the crystals can be used to design optimum data collection strategies (Popov and Bourenkov, 2003; Bourenkov and Popov, 2006).

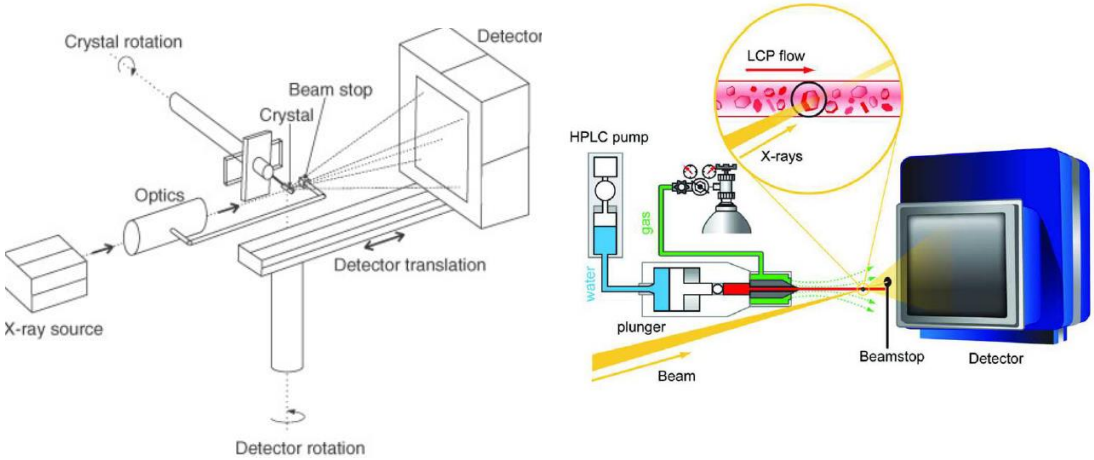


Figure 7-1: Comparison between conventional X-ray crystallography and serial crystallography data collection

Data collection in conventional X-ray crystallography records a rotation series which means several diffraction images are collected for a single crystal at 100 K(left) while it is rotated. However, during a SX experiment, a stream of crystals is continuously injected into the X-ray beam which results in the collection of a single diffraction image at RT for a single crystal (right). Figures adapted from Messerschmidt and Huber, 2000 (left), and Nogly *et al.*,2015 (right).

It is also increasingly common to combine data from several crystals of the same protein in this case, thus decreasing the loss of valuable information due to radiation damage.

Data collection using SX and conventional methods differ in several aspects. During data collection on a single crystal, the crystal can be manipulated in its orientation in the X-ray

beam and data can be collected on different sections of the crystal. In SX, when using streams of crystals, it is not possible to position the crystals individually and due to requiring fresh sample, the previously irradiated crystal cannot be restored for further measurements if they have not already been destroyed by the beam. Using fixed-target devices on the other hand makes it possible to orientate the crystal or reuse the same crystal.
During data collection in SX, usually only one or a few diffraction patterns per crystal are recorded. In SFX experiments, the sample is completely vaporised during data collection when hit by the high intensity X-ray pulse. Only partial Bragg peaks are recorded and the crystals are randomly orientated when the X-ray beam irradiates them. These three features necessitate that hundreds of thousands of diffraction patterns have to be gathered during data collection. Data collection in SX is usally performed at room temperatures and in SFX in vacuum.

7.1.2 Sample delivery methods for SX experiments

As the protein crystals are damaged during the data collection or even damaged to destruction in the case of using an XFEL, the sample needs to be replenished continuously before the next X-ray pulse arrives at the sample area. Various sample delivery systems have been designed to allow data collection from samples with different properties, and have been employed either at an XFEL source or a synchrotron source.

7.1.2.1 Gas dynamic virtual nozzle

One of the first sample delivery devices used at an XFEL was the gas dynamic virtual nozzle (GDVN) (DePonte *et al.*, 2008; Weierstall *et al.*, 2012). This device continuously injects a stream of fully hydrated crystals from an aqueous reservoir solution. It consists of two fused capillaries, an inner silica optical fibre (360 µm OD, 20–50 µm ID) and an outer borosilicate glass capillary (1.2 mm OD, 0.9 mm ID). The sample is injected through the inner capillary whereas the outer capillary is connected to a helium or nitrogen gas line and focusses the sample stream to a thin liquid jet (Figure 7-2). The outer capillary is ground to obtain a pencil-shaped end, preventing any background scatter and accumulation of sample around

the outer capillary. The velocity of the sample stream is regulated by a remotely controllable HPLC. Depending on the viscosity of the sample, flow rates ranging from 5-50 µl/min can be achieved, which in turn means that most of the sample is lost during data collection without any diffraction pattern being recorded. Roughly 1.6 nl of sample is wasted between two X-ray pulses in a common SFX experiment at the LCLS (Weierstall, 2014). Therefore, high crystal density has to be ensured for a high recording rate (7.8 % in Weierstall, 2014), the 'hit' rate of the SX experiment. For this, the crystallisation of the protein has to be scaled up to have enough sample (> 10 ml of 1 mg/ml of protein for (Chapman *et al.*, 2011)) for the experiment, which can be problematic for samples for which only small amounts of protein can be obtained. The high crystal density in turn has another disadvantage. The crystals are prone to settle, but anti-settling devices have been developed (Lomb *et al.*, 2012). Additionally, the crystallisation precipitant, especially at high salt or high PEG concentration, increases the viscosity of the sample which can cause difficulties during the injection such as clogging of the nozzle. Dust, salt crystals and protein crystals are factors that lead to nozzle clogging. Protein crystals that are too big to fit through the nozzle can be filtered before injection via an inline filter unit. If clogging occurs, the whole nozzle needs to be replaced and the flow of the sample out of the injector has to be adjusted again since each nozzle is unique and hand-crafted.

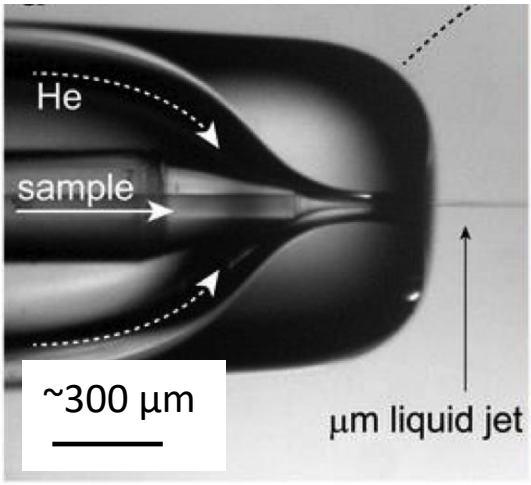


Figure 7-2: GDVN nozzle tip

The liquid sample containing the crystals is injected through the inner capillary and focussed to a liquid jet by the helium gas stream from the outer capillary. Figure adapted from https://portal.slac.stanford.edu/sites/lcls_public/instruments/sed/Pages/Equipment.aspx

Recently a double flow nozzle has been designed. In this, an inert solvent focusses the sample stream, thereby reducing sample consumption and ice formation at the tip of the nozzle that can lead to clogging (Oberthuer *et al.*, 2017). Furthermore, if the pressure of the surrounding solvent/solution is adjusted differently, then a mixing of this solution with the sample can occur. This can provide the basis of novel time resolved studies, where for example a substrate can be mixed with the protein crystals to potentially observe substrate binding events (Oberthuer *et al.*, 2017). This can also be performed by a mixing process before the injection of the sample (Stagno *et al.*, 2016).

Using the GDVN, time resolved studies on light-sensitive proteins have been performed. The samples were illuminated with a laser of the specific wavelength that initiated light-induced processes in the proteins, prior to data collection. Normally, light and dark state data were collected but also intermediate states could be resolved with this method such as for the photo reaction centre (Wohri *et al.*, 2010), photosystem I (Aquila *et al.*, 2012), photosystem II (Kupitz *et al.*, 2014), PYP (Tenboer *et al.*, 2014; Pande *et al.*, 2016) and myoglobin through photolysis of the Fe-CO complex (Barends *et al.*, 2015).

7.1.2.2 High viscosity injectors

Membrane proteins can either be crystallised in detergent or in a lipid environment. Crystallisation using vapour diffusion of detergent-solubilised membrane proteins provides an aqueous environment with low viscosity, an environment fitting for the GDVN, as does crystallisation in the more viscous lipidic sponge phase. For LCP grown crystals on the other hand, the GDVN is not suitable. Thus, the GDVN was modified in order to be used as an injection device for more viscous crystallisation media (Weierstall *et al.*, 2014). The setup is similar to the original GDVN, except that the inner capillary extends beyond the outer capillary. The latter stabilises the extruded LCP using a helium or nitrogen gas stream. In comparison to the GDVN, higher background scattering is observed, as the diameter of the extruded sample ranges from 30-60 µm. Due to the high viscosity, higher pressures are achieved during the extrusion. The sample reservoir is attached to an HPLC and the sample is pushed into the inner capillary via two Teflon balls sitting behind the sample (Figure 7-3). An advantage of the high viscosity is the lower injection velocities, resulting in lower sample

consumption. Given that the purification of some membrane proteins is very challenging in order to get sufficient quantities for crystallisation experiments from the experimenter's point of view, the importance of reducing sample consumption cannot be overemphasised. The commonly used lipid monoolein (9.9MAG) poses some difficulties when extruded into a vacuum chamber. The Joule-Kelvin effect results in a temperature change of the extruded material, leading to the freezing of the LCP sample. Therefore, when using monoolein for LCP crystallisation and subsequent SFX experiments, the LCP has to be doped with 7.9 MAG before the extrusion and data collection, to prevent the freezing of the sample. However, this doping can pose the problem of sample destabilisation due to the addition of another lipid and thus should be tested in advance.

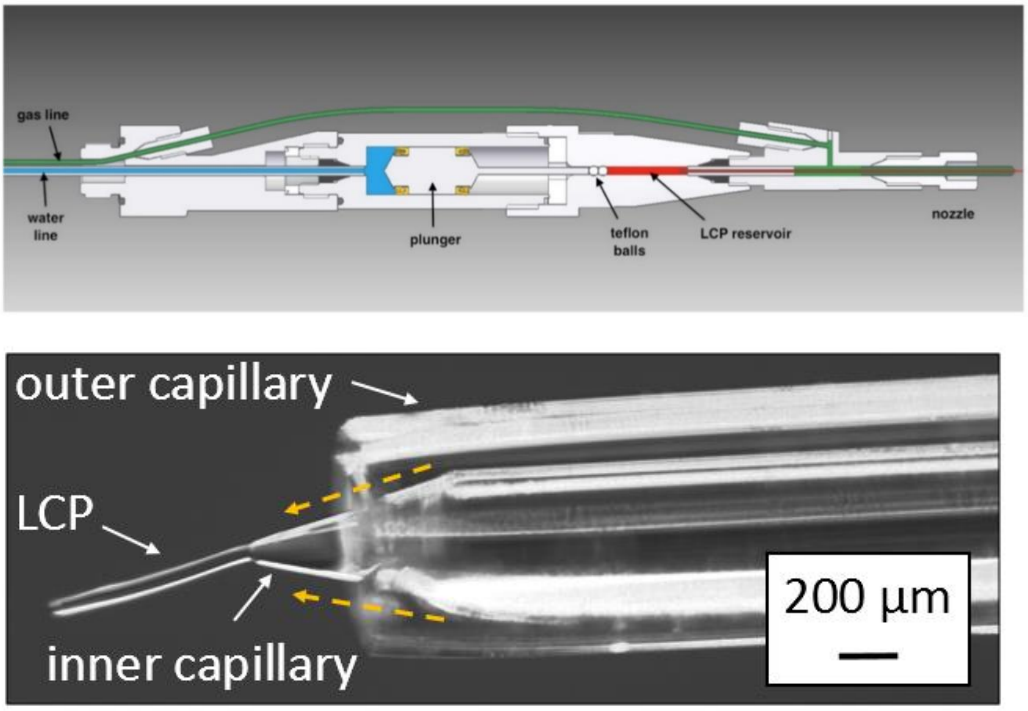


Figure 7-3: LCP injector and injector nozzle

The LCP injector consists of a reservoir that is loaded with the sample containing the crystals. The sample is extruded through the inner capillary of the nozzle via a water line that is connected to an HPLC, and the sample jet is stabilised with nitrogen or helium gas (yellow dotted arrows) that exits in the sample region of the outer capillary. Top picture is adapted from Weierstall *et al.* 2014.

LCP has not only been used for SFX and SX studies of membrane proteins crystallised in this environment, but also for soluble proteins to reduce sample consumption during data col-

lection. In addition, other viscous materials have been tested successfully for serial crystallography data collection on soluble proteins embedded into those materials. This includes the use of low melting point agarose and grease media (Sugahara et al., 2014, 2016; Conrad et al., 2015; Fromme et al., 2015). The LCP injector has ben succesfully used for the RT structure determination of an opioid-receptor (Fenalti *et al.*, 2015), rhodopsin (Kang *et al.*, 2015), GPCRs (Liu *et al.*, 2013; Zhang *et al.*, 2015, 2017a, 2017a, 2017b), diacylglycerol kinase (Caffrey *et al.*, 2014) and bacteriorhodopsin (Nogly *et al.*, 2015).

Serial crystallography using the LCP injector on a synchrotron source using bacteria-rhodopsin as a model membrane protein has been performed successfully with a hit rate of 0.5-2 % and a 43 % indexing rate (Nogly *et al.*, 2015). There were only minor structural differences, with a C_α RMSD of 0.54 Å, to the one solved via SFX at an XFEL (Nogly *et al.*, 2015). Using the LCP injector it was also possible to perform time-resolved studies on light-sensitive membrane proteins using SFX on diacylglyverol kinase (Li *et al.*, 2015) and at a synchrotron source on bacteriorhodopsin (Nogly *et al.*, 2016).

7.1.2.3 Alternative sample delivery methods

Instead of using continuous streams of crystals, the crystals can be stationary by being fixed on a chip that is either rotating or is raster-scanned to collect the diffraction data (Hirata *et al.*, 2014; Suga *et al.*, 2014; Coquelle *et al.*, 2015; Hunter *et al.*, 2015; Roedig *et al.*, 2015; Sherrell *et al.*, 2015; Oghbaey *et al.*, 2016; Schubert *et al.*, 2016). This fixed-target approach has the advantage that the sample consumption is very low, as most to all the crystals fixed on the device are irradiated during the experiment (Figure 7-4C). Several on-chip devices for data collection at 100 K have been developed which usually consist of a low-background mesh or matrix that contain small wells into which the crystals are embedded (Roedig *et al.*, 2015; Sherrell *et al.*, 2015). Again, scaling up the crystallisation conditions is necessary in this case to obtain a solution saturated with crystals that is dispensed onto the chip. Another method is the crystallisation of the protein in microfluidic devices that can be used for SX data collection, reducing the waste of the sample to practically zero (Heymann *et al.*, 2014; Perry *et al.*, 2014; Pawate *et al.*, 2015). Both methods, on-chip and microfluidics are restricted to aqueous crystalline solutions and are thus not suitable for LCP grown crystals.

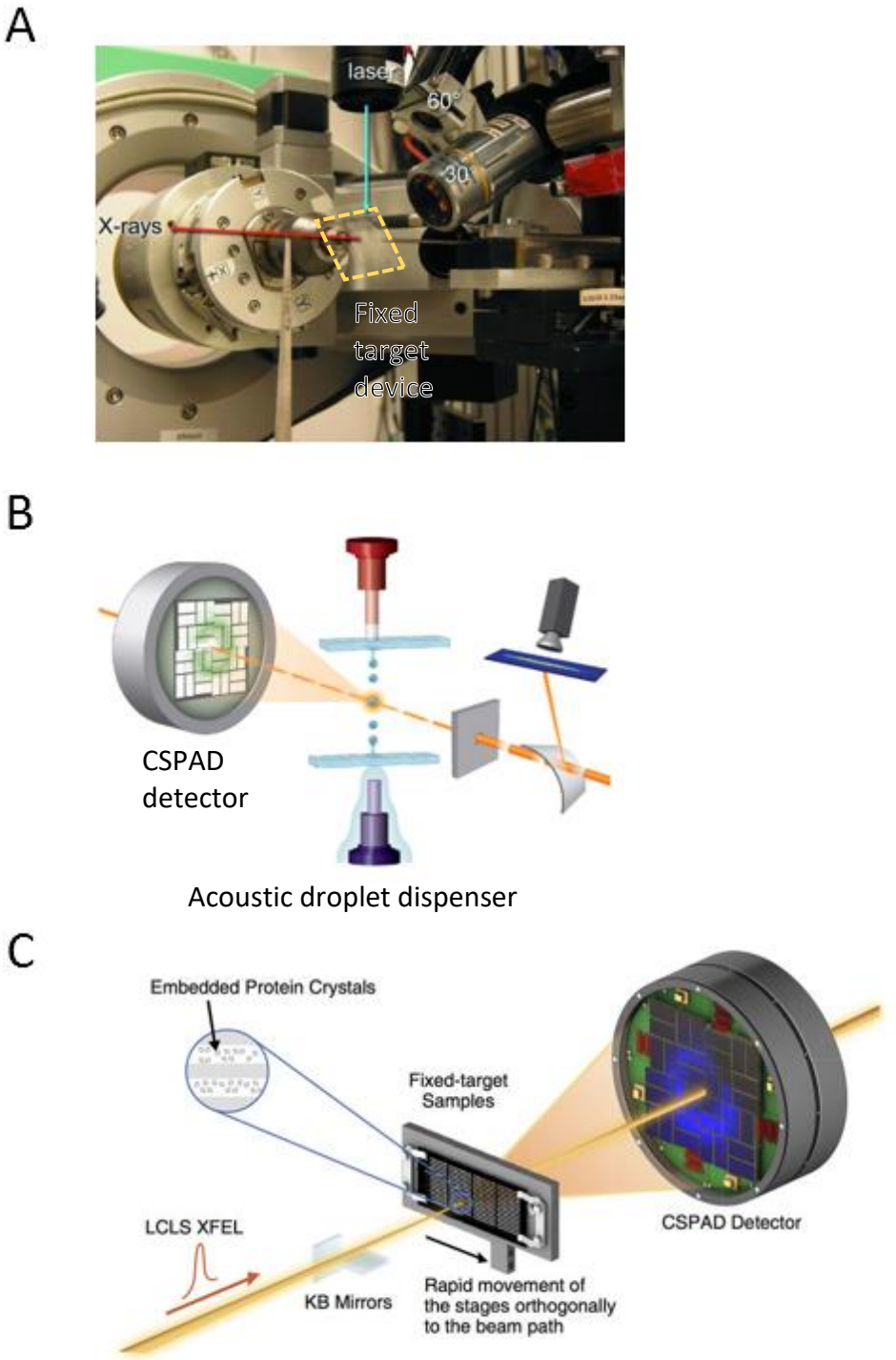


Figure 7-4: Alternative sample delivery methods

Alternative sample delivery methods other than the liquid jet and the LCP jet have been introduced and include: microfluidics device that can be rotated in the X-ray beam (Figure adapted from Pawate *et al.,* 2015) (A), acoustic droplet dispersion. The sample can either be delivered from above or below the X-ray interaction region (Figure adapted from Roessler *et al.*, 2016) (B), a fixed target chip is mounted on a goniometer stage. The mesh contains crystals in each well. The device can be raster scanned and rotated relative to the X-ray beam (Figure adapted from Hunter *et al.,* 2015) (C).

Acoustic droplet dispersion is another way to precisely deliver the protein crystals into the X-ray beam (Roessler *et al.*, 2016). The frequency of dispensing the drops containing the crystals can be synchronised with the pulse frequency of the X-ray beam. This method is applicable for aqueous to slightly viscous solutions such as precipitants containing high PEG concentrations, but not for LCP based crystallisation techniques. Furthermore, the helical scanning method used on several *in vivo* grown crystals in one cryo-loop at 100 K (Gati *et al.*, 2014), normally performed on a single crystal at cryo-temperatures, initially revealed the feasibility of SX at synchrotron sources and led to further SX experiments using different sample delivery methods and also time resolved studies at RT on light-sensitive proteins (Heymann *et al.*, 2014).

7.2 SX EXPERIMENTS AT P11, PETRA III, ON LCP GROWN CRYSTALS

Setting up crystallisation trials in LCP for serial crystallography experiments requires that instead of using LCP sandwich plates (Figure 7-5A), the membrane protein is crystallised in a syringe to obtain a bulk phase containing enough crystals for the SX experiment. The LCP is extruded as a tube from one syringe into the precipitant solution held in another syringe (Figure 7-5B).

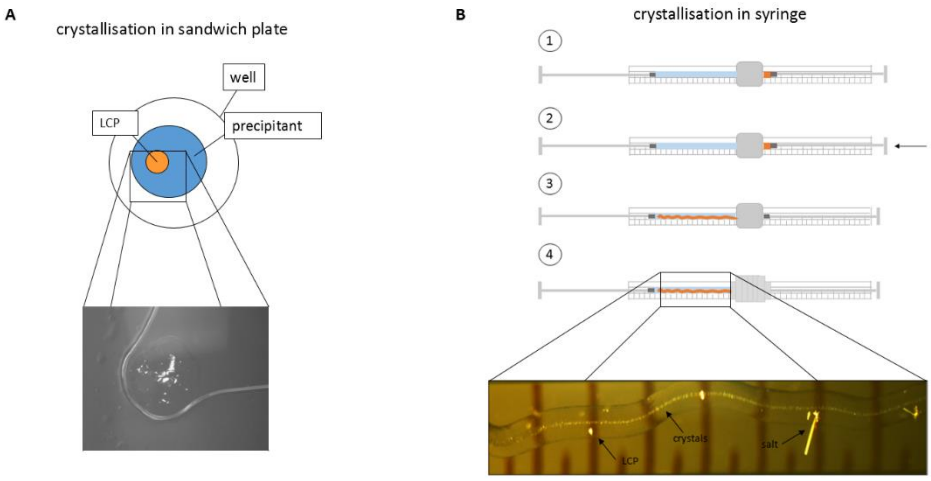


Figure 7-5: Differences between common LCP crystallisation and the crystallisation set up for SX experiments

Conventional crystallisation setup of LCP crystallisation. The LCP drop (orange) is surrounded by the precipitant drop (blue) in a well, sealed by two glass slides (A). LCP crystallisation in a syringe setup (B). The LCP is mixed beforehand and kept in one syringe. Coupling this syringe to one that contains the precipitant (B1), the LCP is carefully extruded into it (B2&3). The syringes are kept joined together and sealed for storage (B4).

To obtain crystals by the syringe methods, adjusting the concentrations of the precipitant components from that of the sandwich plate setup for syringe crystallisation is not always straightforward.

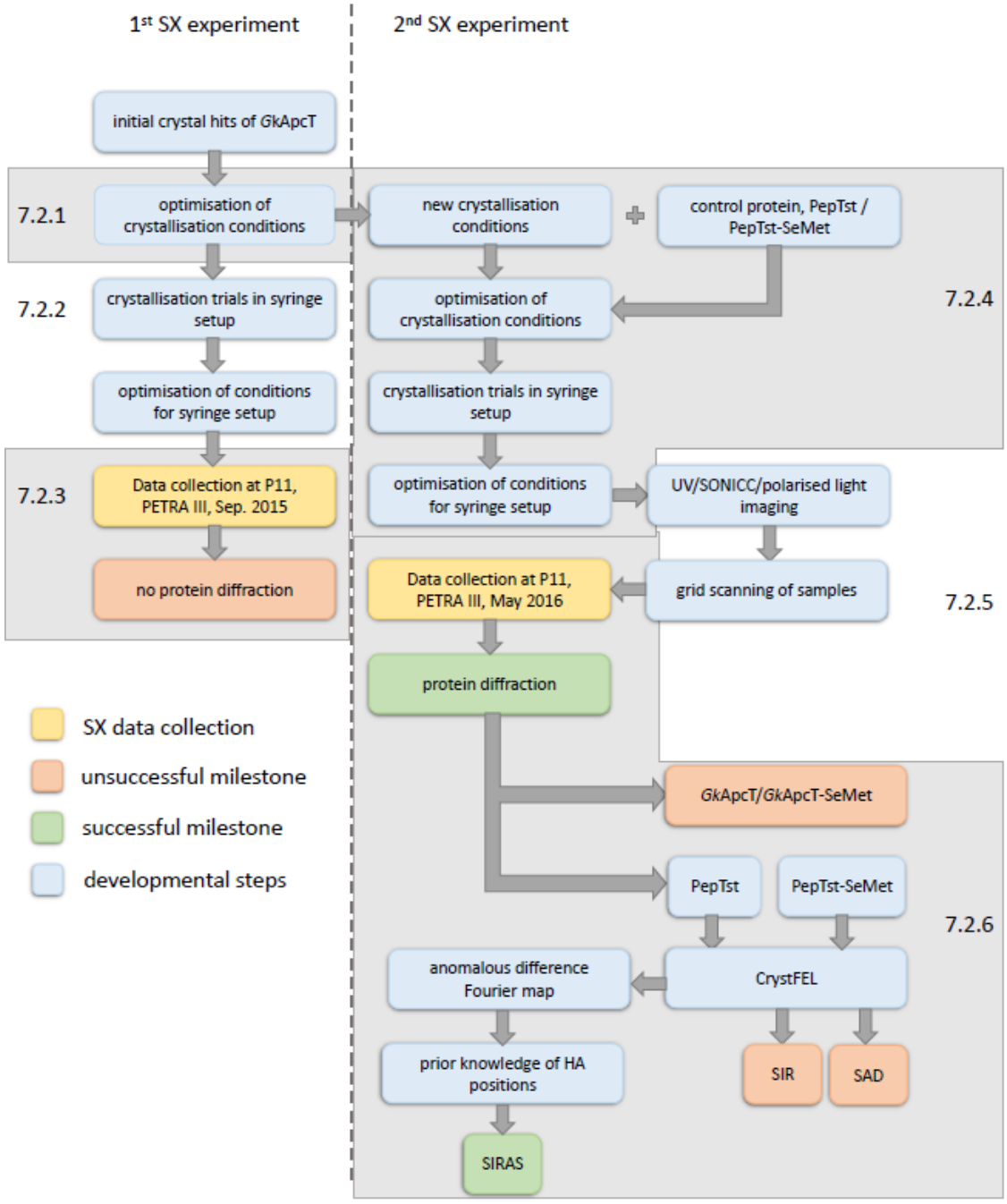


Figure 7-6: Work flowchart for the two SX experiments

Schematic showing the individual steps taken of the work undertaken in this Chapter. The numbers of the sections explaining the experiments in more detail are given next to the corresponding boxes in the flowchart.

It requires several trials to obtain a suitable crystal density which includes changing the PEG or salt concentration, or even changing the protein concentration (Liu *et al.*, 2014). The process of optimising the conditions for SX experiments on LCP grown crystals will be described for the examples of *Gk*ApcT and $PepT_{St}$, as well as trials to collect SX data at a synchrotron source for both samples. An overview of the steps taken to achieve the final results are illustrated in Figure 7-6.

7.2.1 Crystallisation trials for first SX experiment

Initial crystallisation trials on *Gk*ApcT using the conventional LCP setup resulted in crystals diffracting to 3.5 Å in an optimised condition containing: 38 % PEG 200, 100 mM NaAc pH 4.0, 100 mM potassium sodium tartrate (KNatart) and 10 mM alanine. Collection of a full data set to the initially observed resolution of 3.5 Å was restricted due to the high radiation sensitivity of the crystals at cryogenic temperatures. The unit cell of those crystals was determined (a=b=156 Å, c=57 Å, α=ß=γ=90 °) and the space group was P422. Therefore, it was possible to merge the first 10-20 frames (1° oscillation per frame) of a rotation series of four crystals, resulting in a complete data set to 4.2 Å resolution.

Based on these results, it was reasoned that from merging several data sets of 10 frames or less collected from these crystals with subsequent merging would resemble the strategy of a serial crystallography experiment. One option, would be to harvest hundreds if not thousands of crystals and take them to a micro-focus beamline and collect at 100 K only the first few rotational diffraction patterns from each crystal and combine them to one data set. This in turn would be restricted by the time dedicated to the user at such a beamline, considering the time spent mounting the cryo-loop and raster scanning it to identify the position of the crystal, followed by data collection.

Another option would be the continuous injection of LCP grown crystals directly into the X-ray beam as is the case for SFX experiments at an XFEL. The restriction there would also be the limited time, and expense of such an experiment at an XFEL facility.

In 2015, Nogly *et al*. transferred the setup used for SFX experiments to a synchrotron and were able to collect a full data set at room temperature and ambient pressure for bacteriorhodopsin at the ESRF (Nogly *et al.*, 2015). For this, the experimental end station had to

be changed to accommodate the same injector as used for SFX experiments, to extrude the LCP tube into the beam (Nogly *et al.*, 2015).

The use of the LCP injector at P11, PETRAIII, Germany, was part of a long-term proposal of the group of Henry Chapman. Crystals of *Gk*ApcT, in addition to other samples were used to test the injector and data collection at this beamline. The optimisation of the crystallisation of *Gk*ApcT towards using it for SX experiments, and the outcome of this experiment is described in the following sections.

7.2.2 Scaling up from LCP sandwich plates

For the first SX experiment carried out with *Gk*ApcT, an initial concentration of 10 mg/ml of protein was used for crystallisation in a syringe. The protein was reconstituted into LCP as described in Section 3.4.2.2 using the MO:Chl mixture. 10 µl of LCP were injected into 80 µl of precipitant containing 38% PEG200, 100 mM NaAc pH 4.0, 10 mM alanine and 40-100 mM potassium sodium tartrate. The composition of the precipitant was similar to the one in which the crystals of *Gk*ApcT grew in in the LCP sandwich plate setup. In none of the conditions did protein crystals grow in the LCP tube.

Additional optimisation trials were performed with varying salt and PEG concentrations as suggested elsewhere (Liu *et al.*, 2014). Using a precipitant containing 46 % w/V PEG, 100 mM NaAc pH 4.0, 100 mM potassium sodium tartrate, 10 mM alanine, crystals were formed but so did salt crystals, possibly of the potassium sodium tartrate. Knowing that salt crystals can block the injection nozzle during data collection, lower salt concentrations were tried.

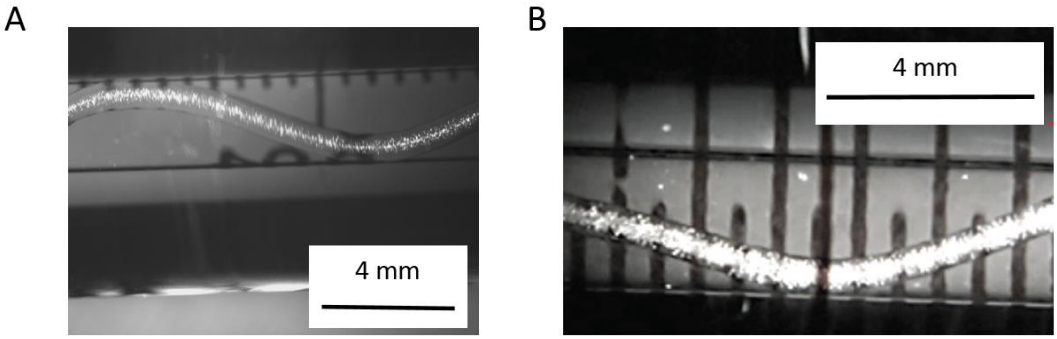


Figure 7-7: Comparison of crystal densities in the LCP syringes

Crystallisation of *Gk*ApcT (A) compared to crystallisation of a GPCR in a syringe (B). (B) was adapted from Liu *et al.* 2014.

Furthermore, the syringes were thoroughly cleaned before the crystallisation experiments

to prevent any dust particles entering them and forming potential nucleation points for the salt crystals. The crystal density obtained in the optimised conditions was low compared with previous experiments using this setup and could not be increased any further before the experiment commenced (Figure 7-7).

For SFX experiments on GPCRs, the crystallisation conditions in the plate setup could be used directly for crystallisation of the protein in the syringe (Liu *et al.*, 2013, 2014). The case of *Gk*ApcT highlights the fact that a considerable amount of time can and should be spent on the optimisation of the crystallisation conditions. Crystals of *Gk*ApcT formed after 1 day and could be kept in the syringe for over a month without visual degradation of the crystals. Salt crystals formed occasionally, probably nucleated by residual dust particles in the solutions or the syringe.

7.2.3 Experimental setup at beamline P11, PETRAIII

The experimental setup used for SX data collection at the P11 beamline at PETRAIII did not differ significantly from the one described by Nogly *et al.* 2015 and is shown in Figure 7-8.

A collimator was used to block the incident radiation that is not parallel to the optical axis before it reaches the focal region. The X-ray beam was focused using bendable Kirkpatrick-Baez (KB) mirrors to produce a 5 µm (vertical) x 9 µm (horizontal) FWHM focal size confirmed visibly using the inline microscope and a fluorescent yttrium aluminium garnet (YAG) screen mounted temporarily on the injector nozzle tip. When irradiated with X-rays, the YAG screen emits yellow light with an emission maximum at a wavelength of 574 nm (Blasse and Bril, 1967). The X-ray flux in the focal spot was measured using a calibrated diode by the beamline scientists and determined to be ~ 10^{13} photons/s.

Using the LCP injector (Weierstall *et al.*, 2014), the crystal laden LCP was continuously extruded into the X-ray focus from a reservoir containing 20 µl of the LCP-crystal mixture. The thickness of the LCP stream depends on the inner capillary diameter of the nozzle. Here, inner diameters of 30, 50, and 75 µm were tested, of which 50 µm was the preferred diameter used during data collection. In general, smaller diameters are better, as they result in reduced LCP background scattering, though if clogging of the nozzle is observed repeatedly, a change to a bigger diameter should be considered. The injector was mounted on an *x*-*y*-*z* stage and the inline- and an additional sideway microscope were used to aid alignment of

the injection nozzle to the beam and provide feedback to the user during alignment. Below the X-ray focal region, a small tube ("sample catcher tube", Figure 7-8) was mounted, collecting the LCP stream after it was exposed to X-rays, thereby preventing unnecessary contamination of the beamline instrumentation.

To record the diffraction patterns from the protein crystals, a PILATUS 6M detector with a maximum frame rate of 25 Hz was positioned 0.3 m from the sample. During data collection, the sample exposure time was typically 8 ms per frame.

Before the injection experiment could commence, the precipitant solution needed to be removed and the remaining LCP mixed with more lipid to form LCP with the residual precipitant. In the latter step, the protein crystals might have dissolved through the addition of more lipid or might be damaged by the mixing process. This was tested with the syringe set ups. After mixing with additional MO:Chl, small crystals were still present in the lipid, indicating that the protein crystals survived the procedure. For the experiment, several syringes were prepared for crystallisation of *Gk*ApcT.

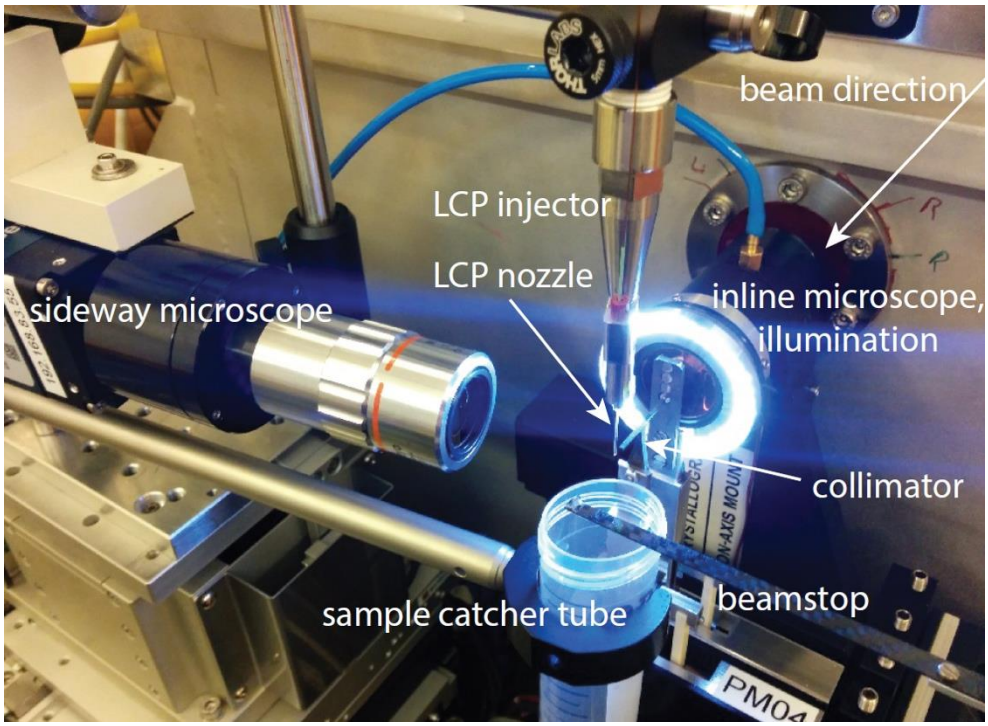


Figure 7-8: Experimental setup at beamline P11

Photograph of the experimental setup in the focal region, looking from the Pilatus 6M detector upstream.

During the experiment, the crystal-laden LCP stream flowed (60-300 nl/min using a pressure of 240-300 psi) continuously past the focal region. The hit rate depends mostly on the crystal concentration in the LCP stream. Diffraction data frames were acquired continuously (with a rate close to 25 Hz) while the LCP injector was running. Online data analysis of individual diffraction images for quick feedback during the beamtime and identification of diffraction patterns containing a crystal 'hit' was performed utilizing the software ONDA (Mariani *et al.*, 2016). Peak-finding and hit-finding parameters in ONDA were adjusted in the initial stages of the experiment, aided by visual inspection of individual diffraction patterns, until a sufficient set of conditions was found. Diffraction patterns classified as 'hits' were further analysed with CRYSTFEL (White *et al.*, 2012).

During the data collection, no protein diffraction was observed and only diffraction patterns of what seemed to be salt crystals were collected (Figure 7-9).

There are several explanations for this result:

(1) Since the maximum frame rate of the Pilatus 6M is limited to 25 Hz, for exposure times smaller than 40 ms there is a "dead time" window during which no data can be recorded (i.e., a 10 ms exposure is followed by a 30 ms dead time before the next data frame can be recorded). During the dead time, the LCP stream is exposed to X-rays, causing severe heating of the LCP stream (visible in the inline microscope as bubble formation in the LCP stream, reminiscent of "boiling") and thus damaging the protein crystals in the LCP.

(2) As mentioned above, the potassium sodium tartrate used in the precipitant was likely to form salt crystals. Thus, it was assumed that during the extrusion of the precipitant, the salt concentration locally increased in the residual LCP tube, and during mixing with additional lipid, the salt crystals had damaged the protein crystal in the mixing process. The sample was checked afterwards for protein crystals via cross-polarised light under the microscope, but salt crystals can also give a positive signal and be misinterpreted as protein crystals.

(3) The protein crystals might have dissolved during the addition of the extra lipid, although this scenario would usually happen during addition of a different lipid than the hosting lipid. This is normally carried out when monoolein is used during crystallisation. In this case, 7.9MAG should be added before the data collection in vacuum to prevent freezing of the sample as it enters the environment.

(4) The crystal density was perhaps too low when compared to other SX experiments on LCP grown membrane protein crystals.

Based on this knowledge, another attempt was made to collect SX data on *Gk*ApcT to solve its structure. To overcome the possible problems stated above, more control and optimisation experiments were conducted.

The use of a rotating shutter, to synchronise irradiation and data collection on the detector and to reduce the radiation damage on the whole sample was suggested. Additionally, *Gk*ApcT was subject to further crystallisation trials in order to identify other conditions that do not form salt crystals as easily, and optimisation of the scaled-up syringe setup was attempted in order to reduce the amount of PEG, which also affects salt crystal formation. For the steps of precipitant extrusion and lipid mixing, further pre-screening methods were introduced such as using SONICC and UV to monitor protein crystals in the sample during each step, and using the single wells of *in situ* LCP plates to screen for protein diffraction in the sample as a verification of their crystal density.

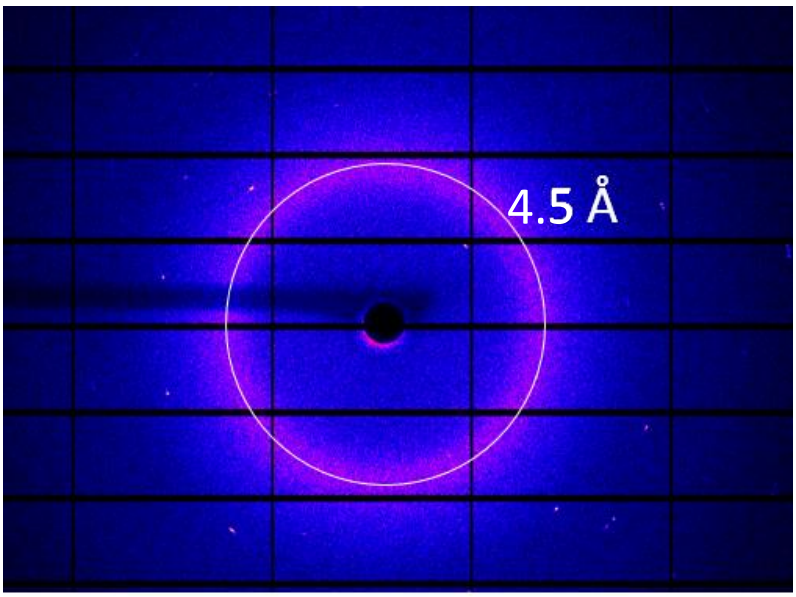


Figure 7-9: Example of diffraction pattern of GkApcT sample

The diffraction pattern shows distinct powder diffraction ring at 4.5 Å which corresponds to the LCP present in the sample. The LCP ring is a good indicator that the X-ray beam is penetrating the sample. There are some sharp spots at higher resolution but no diffraction characteristic of protein crystals is observed.

7.2.4 SX experiments on *Gk*ApcT and $PepT_{St}$ – towards phasing

From the results of the previous SX experiment at beamline P11, it was concluded that by solving the problems that occurred during data collection and introducing further control experiments on the sample, as well as using a model protein as an additional control, another SX experiment could be conducted. A proposal for such an experiment, again at P11, PETRAIII, and in collaboration with Henry Chapman's group at CFEL, Hamburg, was submitted and approved. Purification, crystallisation and sample preparation were conducted by myself. Data collection was carried out jointly by me, Dr. Ed Lowe, members of Henry Chapman's group, and by the beamline scientists on P11. All online data analysis (peak- and hit finding with ONDA) and indexing and peak integration with CRYSTFEL was performed by Dr. Stephan Stern (CFEL, Hamburg). Further data analysis, starting with MTZ-files, and phasing was performed by Dr. Ed Lowe (Dept. of Biochemistry, University of Oxford).

7.2.4.1 Changes to crystallization protocol of *Gk*ApcT

The initial optimisation trials for the large-scale crystallisation of *Gk*ApcT had aimed to obtain a higher crystal density within the LCP than in the previous experiment. From previous experiments, it was known that higher PEG concentrations increased the crystal density slightly, but that when the PEG concentration increased further, potassium sodium tartrate formed salt crystals. Therefore, *Gk*ApcT was used in another screening process to identify additional conditions which excluded the use of KNatart from the precipitant, and thus reduced the risk of salt crystal formation in either of the steps required for data collection. Another component that had not been changed or optimised until then was the protein concentration. It was reasoned that an increase of protein in the LCP could possibly improve crystal formation, because the mobility of the protein in the bigger LCP tube would be reduced. Therefore, *Gk*ApcT expression and purification was scaled up to obtain more protein and thus a higher concentration for the crystallisation trials. Instead of 2 L cell culture, the protein was purified from 4 L to a final concentration of 30-40 mg/ml of protein.
For the rescreening process, *Gk*ApcT was diluted to 10 mg/ml using the gel filtration buffer used for conventional LCP crystallisation trials, rescreening the conditions and optimisation

screens used for the crystallisation of *Gk*ApcT. New hits appeared, which were also used for crystallisation trials in the syringe but without any success. However, these hits were further optimised and finally a condition was identified that produced *Gk*ApcT crystals diffracting to 2.8 Å. Thus, due to the alteration in the purification process driven by the preparation for the SX experiments, the final data set of *Gk*ApcT was collected on a single crystal at 100 K (see Chapter 4).

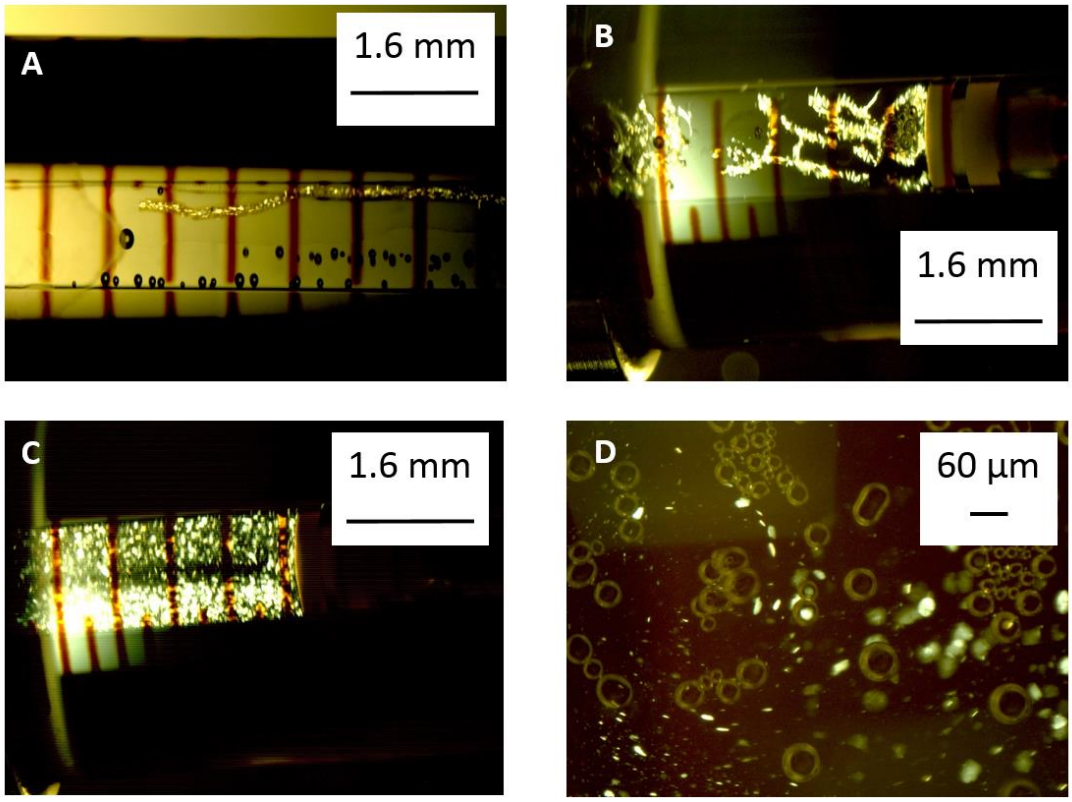


Figure 7-10: Crystallisation of GkApcT in syringe setup

*Gk*ApcT crystallised in high density in LCP using 100 mM NaAc pH 4.0, 10 mM alanine, 100 mM ammonium citrate, and 46 % (V/V) PEG200 as precipitant and a protein concentration of 20 mg/ml (A). The protein crystals are still present after extrusion of the precipitant(B) and additional lipid mixing (C and D).

Initially, *Gk*ApcT was crystallised in the syringe using the LCP plate crystallisation condition that had resulted in higher resolution diffracting crystals of *Gk*ApcT, but using 20-30 mg/ml of protein instead of 10 mg/ml. No crystals appeared even after several optimisation attempts. Based on the rescreening process, other conditions resulting in crystals of *Gk*ApcT were tested using the syringe setup. Using ammonium citrate instead of potassium sodium tartrate resulted in a high crystal density and this time no salt crystals formed even at high PEG concentrations (Figure 7-10). Based on these conditions, *Gk*ApcT-SeMet (30 mg/ml)

was also subjected to crystallisation trials, and crystals appeared after 1 d of incubation at 20 °C using the precipitant containing 46 % (V/V) PEG200, 100 mM NaAc pH 4.0, 10 mM alanine and 100 mM ammonium citrate (Figure A 15). In the conventional crystallisation setup, no crystals of *Gk*ApcT-SeMet were observed in the sandwich plates. Given the difficulties in obtaining crystals of a *Gk*ApcT derivative with conventional LCP crystallisation, it was proposed to use SX data collection to perform SAD experiments using the SeMet derivative of *Gk*ApcT.

Obtaining the *ab initio* crystal structure of a protein using SFX or SX has only recently been achieved, with lysozyme being the first enzyme to be solved via SAD using SFX (Barends *et al.*, 2013). An attempt to solve the structure of luciferins using a similar approach was unsuccessful, but the collection of data on native luciferin crystals grown to perform SIRAS was possible instead (Yamashita *et al.*, 2015). Despite using heavy atoms, anomalous data can also be collected from the naturally occurring sulphur in cysteines and methionines. The anomalous sulphur signal was successfully detected from crystals of a GPCR using SFX and used to solve the structure of this protein (Batyuk *et al.*, 2016), and a more extensive study on lysozyme derivatives investigated the amount of data necessary to perform SAD, SIR or SIRAS (Nakane *et al.*, 2016). All these studies show that it is possible to perform phasing experiments using serial crystallography data collection strategies. Although the studies are based on data collection at an XFEL, the feasibility of performing serial crystallography experiments at a synchrotron source can also be extended to phasing experiments.

7.2.4.2 Crystallisation of $PepT_{St}$

Unsure if the SAD SX phasing approach for *Gk*ApcT-SeMet would be successful, a control protein was used as a sample alongside it, $PepT_{St}$. This protein is a well-known proton-coupled peptide transporter of the MFS family (Solcan *et al.*, 2012; Lyons *et al.*, 2014). The crystallisation conditions were already established in our research group and it was known that the $PepT_{St}$-SeMet derivative can be expressed, purified and crystallised. Crystal trials of the wildtype protein and the SeMet-derivative were performed. $PepT_{St}$ crystallises in 7.8MAG as host lipid and precipitant containing 16-23 % (V/V) PEG 400, 0.1 M HEPES-NaOH

pH 7.0 and 0.15-0.55 M $NH_4H_2PO_4$ using 10 mg/ml of protein. This concentration was initially used for the scaling up process for the SX experiment. Given the experience with *Gk*ApcT, a range of increasing PEG concentrations were tried but did not lead to successful crystallisation. As for *Gk*ApcT, a higher protein concentration was tested for $PepT_{St}$. At 20 mg/ml the protein crystallised in 23 % (V/V) PEG 400, 0.1 M HEPES-NaOH pH 7.0 and 0.15M $NH_4H_2PO_4$. Due to the higher PEG concentration, the sponge phase formed. To test the handling of this different lipid phase for the final experiment, all preparation steps before the transfer of the sample into the injector sample reservoir were carried out. The precipitant had to be extruded very carefully so as not to lose the protein crystals, and could not be fully extruded, so that a few microliters of the precipitant remained in the syringe. Additional 7.8MAG was added to form LCP with the residual precipitant. The crystals survived the process and still showed a high crystal density after the mixing process (Figure 7-11).

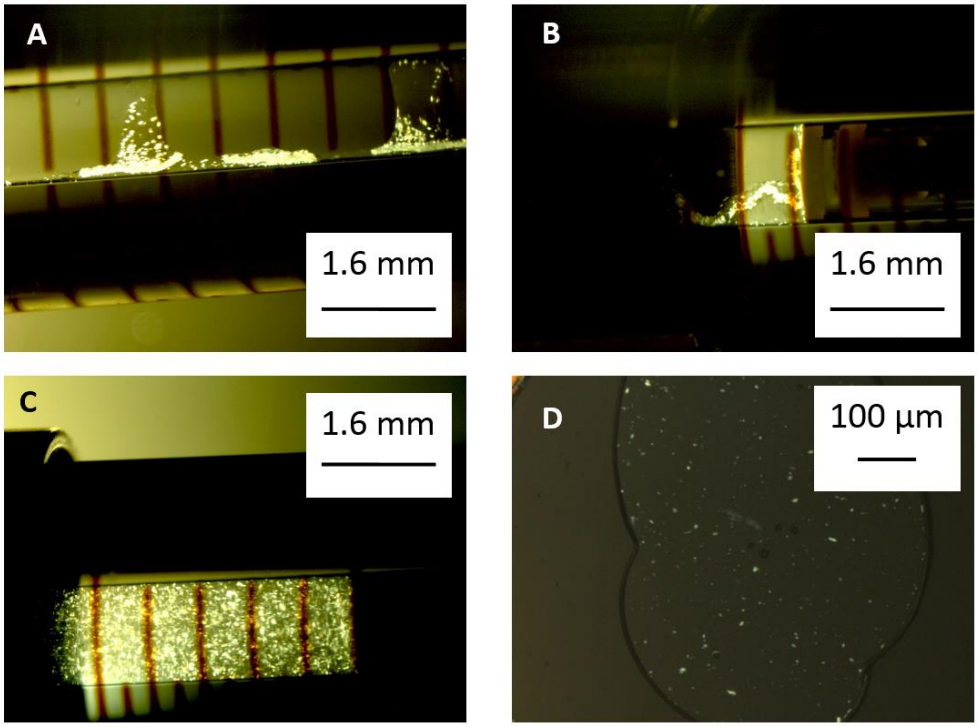


Figure 7-11: Crystallisation of $Pept_{St}$ in syringe setup

$PepT_{St}$-WT crystallised at 30 mg/ml protein concentration using 7.8MAG as host lipid and precipitant containing 23 % (V/V) PEG 400, 0.1 M HEPES-NaOH pH 7.0 and 0.15M $NH_4H_2PO_4$ (A). Due to the formation of the sponge phase, extrusion of the precipitant was difficult but the crystals survived the procedure (B). More 7.8MAG had to be added to mix with the remaining precipitant in the syringe. The crystals were still present after lipid mixing (C and D).

The $PepT_{St}$-lipid-crystal mixture had a lower viscosity than that for *Gk*ApcT. Due to the high PEG concentration used here, and the different lipid, instead of LCP the more liquid sponge phase was formed, visually seen as lipid drops in the precipitant rather than a lipid tube (Figure 7-11). Thus, it was possible that the sample in the reservoir would extrude more quickly due to its more liquid like nature, which meant that more sample was prepared for the data collection.

7.2.5 Pre-screening approaches

Unmixed and premixed samples were shipped to Hamburg and additionally stored for 3 weeks at 20°C. On both occasions, the samples still contained crystals and the phase of the LCP/sponge phase did not change due to shipping or storage.

The samples were tested for crystal survival using cross-polarisation and UV microscopy as well as SONICC (Haupert and Simpson, 2011). The samples were prepared, and then dispensed onto UV transparent LCP plates and screened under the microscope. The SONICC and UV images suggested a high crystal density for both proteins, although it was difficult to overlay the images for UV and SONICC (Figure 7-12 and Figure A 16).

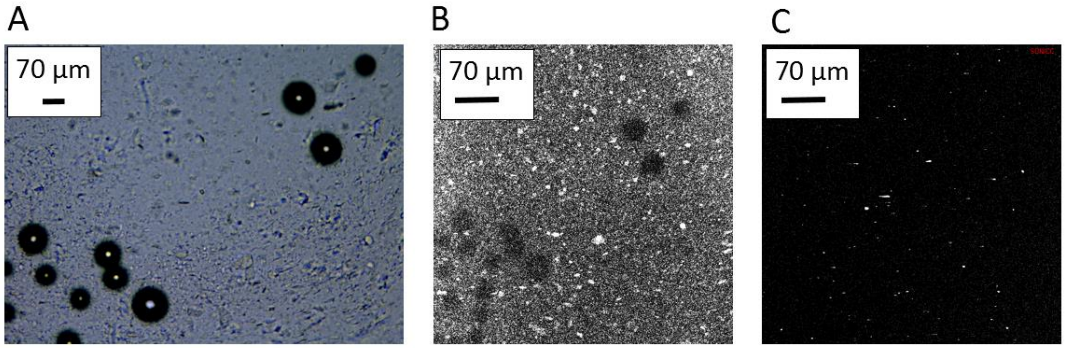


Figure 7-12: confirmation of GkApcT crystals by UV and SONICC

Crystal lipid mixture was dispensed onto single wells of an LCP plate and the mixture inspected by bright field microscopy (A), UV (B) and SONICC (C) confirming that the crystals observed via the cross polariser are protein crystals.

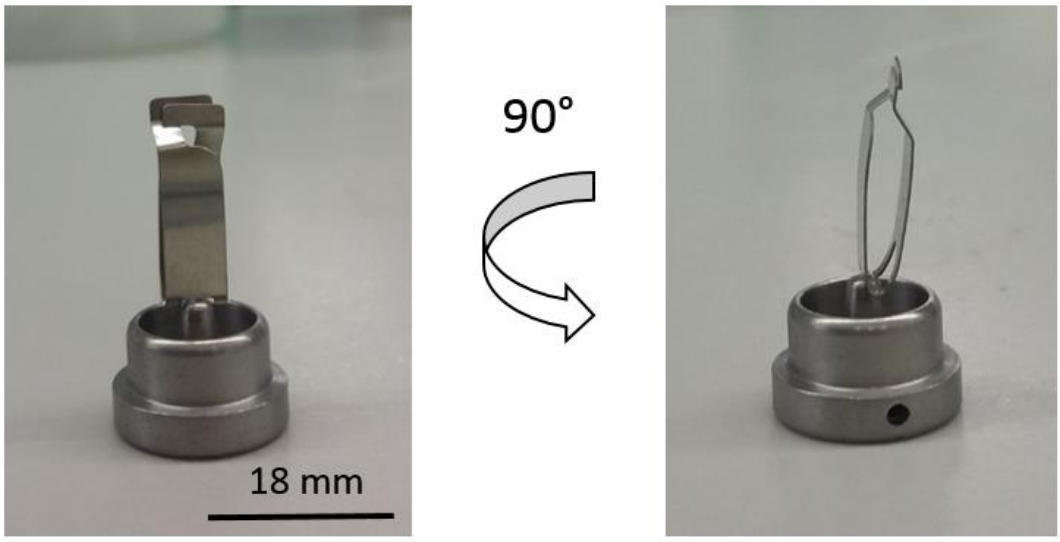


Figure 7-13: single well Diffrax™ sample holder

In addition to these steps, the Diffrax™ plates from Molecular Dimensions Ltd. were used for grid scanning X-ray diffraction experiments. The mixed LCP samples were dispensed into a single well, sealed and mounted onto the well-holding device (Figure 7-13). The grid scan was performed at beamline P11, PETRA III, Hamburg at ambient pressure and room temperature, mimicking the experimental conditions during the SX beam time. For all the samples, protein diffraction was observed (Figure 7-14) using the raster scanning approach, indicating that the protein crystals survived the mixing process with additional lipid. The visible resolution achieved for *Gk*ApcT was around 7 Å using the raster-scanning, which suggests that the crystal quality might have suffered and that it might only be possible to obtain low resolution data for both the WT and SeMet derivative. Therefore, data on *Gk*ApcT was initially collected using the fixed-target approach on the *in situ* well to obtain more data on the possible resolution achievable and to gain more information about the possible hit rate.

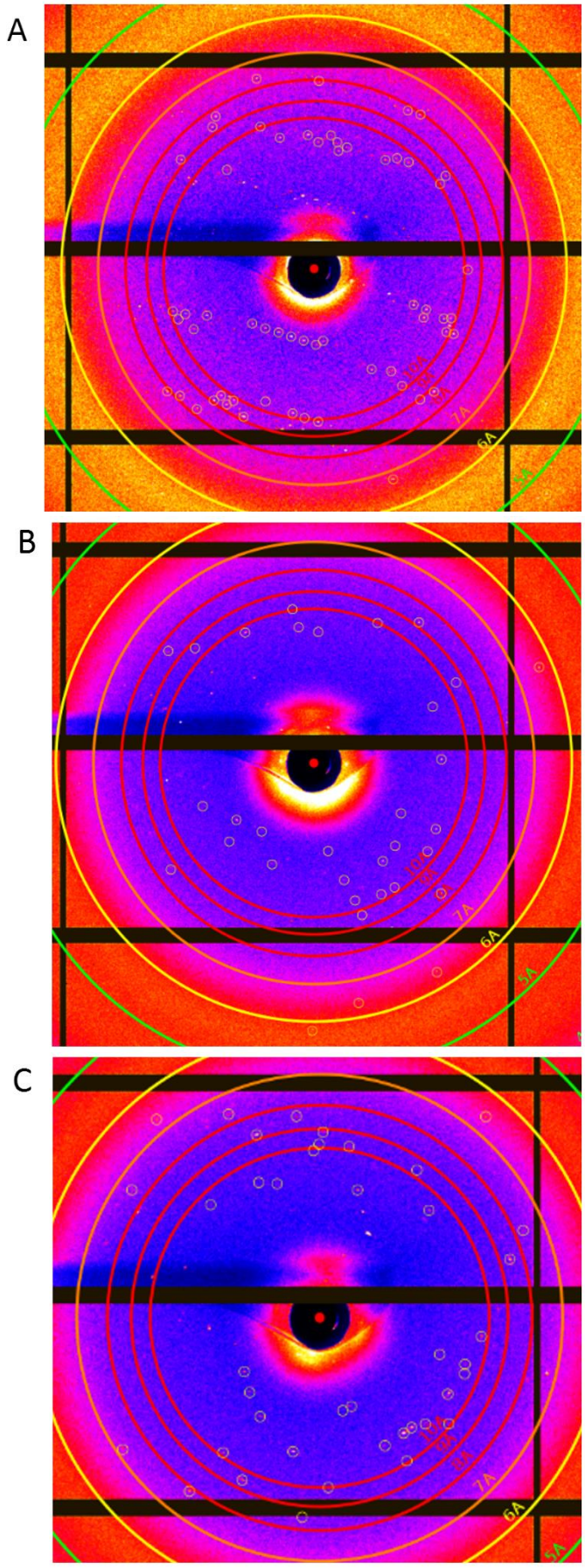


Figure 7-14: Prescreening results from raster-scans

Room-temperature raster-scanning of the Diffrax™ well containing mixed crystal samples of GkApcT-SeMet (A), $PepT_{St}$-WT (B), $PepT_{St}$-SeMet (C). Protein diffraction was observed from all three samples scanned, verifying that the protein crystals survived the mixing process and that the samples could be used for SX experiments.

7.2.6 Second SX experiment using the LCP injector: May 2016

The X-ray wavelength was set to be 12.8 keV (0.9686 Å) which is close to the selenium edge (12.651 keV ≈ 0.98 Å) but was chosen to be above it to avoid missing any anomalous signal. This precaution not only accounts for uncertainty of the oxidation state of Selenium, but also for potential uncertainties in the photon energy/wavelength calibration of the beamline.

Based on observations from the SX experiment in September 2015, that due to the continuous irradiation there was considerable radiation damage to the sample, a shutter/chopper wheel blocking X-rays in the detector dead time window was used. The chopper wheel consisted of a rotating metal disk with off-centred holes. During the experiment, the chopper disk rotated, driven by a stepper motor, with its rotation speed controlled by the user. When X-rays passed through the holes, data are acquired with the Pilatus detector. The signal-trigger to the Pilatus detector is sent by a photodiode behind the chopper wheel, thereby synchronizing the passage of X-rays to the data acquisition, while the chopper wheel blocked the X-rays during the detector dead time window. The hole size, in combination with the rotation speed, determine the effective exposure time and the data rate. Different hole sizes provide the option of finding the best compromise of data collection rate and exposure time. The rotation speed was optimized to be as close to a 25 Hz data acquisition rate as possible. An increase of sample flow rate and the application of chopper wheel were utilised to avoid sample deterioration.

7.2.6.1 SX data sets and data processing

Fixed-target data collection on LCP *in situ* wells on *Gk*ApcT-WT and *Gk*ApcT-SeMet revealed that only low resolution data to about 7 Å, but in some cases to 4 Å, could be collected and that the hit rates of the data collection were very low using the grid scanning. Both samples were also used for direct injection into the beam using the LCP injector, but again only low hit rates (<0.7 %) could be achieved and only a small portion of the images could be indexed. 127 indexed patterns were obtained for *Gk*ApcT-WT, with which it was possible to determine the unit cell parameters to be a= 55.7 Å, b=159.6 Å, c=160.0 Å, α=β=γ=90°. Not

enough indexable patterns were obtained to determine the unit cell for *Gk*ApcT-SeMet. Therefore, data collection on $PepT_{St}$ continued for the remaining allocated time in order to obtain enough data to perform phasing experiments as a proof-of-principle approach on this sample.

In total, for the $PepT_{St}$-SeMet and $PepT_{St}$-WT diffraction data, 824,858 and 535,900 images were recorded, respectively. There were 136,637 ($PepT_{St}$-SeMet) and 61,446 ($PepT_{St}$-WT) hits found, resulting in "hit rates" of 16.5 % ($PepT_{St}$-SeMet, during an individual run even 40 % hit rate) and 11.4 % ($PepT_{St}$-WT). Diffraction patterns classified as a "hit" were further processed with the software CRYSTFEL.

The workflow conducted in CRYSTFEL resembled very closely the one outlined in the CRYSTFEL tutorial (http://www.desy.de/~twhite/crystfel/tutorial.html);

After the application of an initial geometry file (containing the wavelength and detector geometry, i.e., location of pixels with respect to the beam) and initial indexing, a starting set of unit cell parameters, the geometry and unit-cell parameters were adjusted. Five refinement cycles of geometry and unit cell were performed iteratively until the number of indexed patterns no longer changed and the distribution of unit cell parameters exhibited sharp peaks. Eventually, to obtain the final set of indexed diffraction intensities, the unit cell parameters were deliberately set to the determined peak values in the previous refinement cycles (a 5% tolerance of variation was still possible). The final numbers of consistently-indexed diffraction patterns considered for peak integration and post refinement were 37,476 ($PepT_{St}$-SeMet) and 5,670 ($PepT_{St}$-WT) respectively, hence the indexing rate (no. of indexed patterns/no. of hits) was 27.4 % (9.2 %), the combined hit-and-indexing rate (no. of indexed patterns / total patterns) was 4.5 % (~1%) for the SeMet (wildtype) dataset. The considerably worse hit rate for wildtype data compared to the SeMet data can be explained by the higher crystal density of in the SeMet sample-laden LCP. However, the reason for the considerable worse indexing success for the wildtype dataset (9.2 % vs. 27.4%) currently cannot be explained.

During indexing of diffraction patterns classified as 'hits', it was found that the unit cell parameters differed to the unit cell parameters of published $PepT_{St}$ structures (a=102.06 Å, b=110.33 Å, c=110.74 Å, α=β=γ=90°). For the SX experiment here, the unit cell was orthorhombic, a=105.32 Å, b=113.01 Å, c=108.68 Å, α=β=γ=90°. There are only 2 possible space

groups (C222 and $C222_1$) and neither of them is merohedral, so the indexing could be continued without the need to resolve a possible indexing ambiguity (Brehm and Diederichs, 2014; White *et al.*, 2016).

The final list of indexed (partial) diffraction intensities was used for peak integration/merging of intensities in order to produce full diffraction intensities for the WT and SeMet datasets. For merging, the symmetry (point group) applied was "222" in order to preserve the Bijvoet differences and new CRYSTFEL features including scaling, partiality prediction and post-refinement (White *et al.*, 2016) were applied to improve the quality of the merged intensities. The merging statistics for the native and derivative data are given in Table 7-1. Both data sets were complete and had a final resolution of 3.0 Å ($PepT_{St}$-WT) and 2.8 Å ($PepT_{St}$-SeMet). Differences in the resolution of the data sets could be due to preparational differences between the purification of each of the proteins or different stabilities of the crystals during their preparation for the SX experiment.

Figure 7-15A shows an example of an indexed diffraction pattern, with circles surrounding predicted peak positions. The LCP ring between 4 and 5 Å is clearly visible. Figure 7-14 B and C show the distribution of unit cell parameters for $PepT_{St}$-SeMet (B) and $PepT_{St}$-WT (C), including data from 37,476 ($PepT_{St}$-SeMet) and 5,670 ($PepT_{St}$-WT) indexed diffraction patterns, used for further analysis (peak integration and post-refinement). There is a distribution of parameters although these are quite narrow, signifying that all these tens of thousands of crystals were not completely isomorphous. The exact influence of the non-isomorphism of different crystals, inherent to serial crystallography, on data quality and further steps such as phasing, electron density maps, and model building, is not yet fully understood. Currently, a very narrow distribution of unit cell parameters is usually accepted.

Table 7-1: Merging statistics for $PepT_{St}$-WT and $PepT_{St}$-SeMet

	$PepT_{St}$-WT	**$PepT_{St}$-SeMet**
X-ray source	P11, PETRA III	
Detector	Pilatus 6M	
Temperature (K)	291	
Wavelength (Å)	0.9686	
Beam size (hor. x vert. in µm)	9 x 5 FWHM	
Flux (photons/sec)	~ 10^{13}	
Exposure time (ms)	8	
Crystal size	~10 x 10 x 10 μm^3	
Space group	C222	C222
Unit cell parameters		
a, b, c (Å)	105.0, 112.7, 109.1	105.3, 113.0, 108.7
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Number of collected images	535900	824858
Number of hits	61446	136637
Number of indexed images	5670	37476
Resolution (Å)[1]	44.41-3.0 (3.18-3.0)	47.39-2.8 (2.95-2.8)
Total number of observations	26662 (4232)	32664 (4718)
Total number unique	13331 (2116)	16332 (2359)
I/σ(I)	2.6 (0.7)	6 (1.2)
$CC_{1/2}$	0.961 (0.137)	0.993 (0.379)
Completeness (%)	100 (100)	100 (100)
Multiplicity	58.7 (42.2)	266.4 (187.4)
R_{split} (%)	26.29 (192.80)	10.52 (92.64)

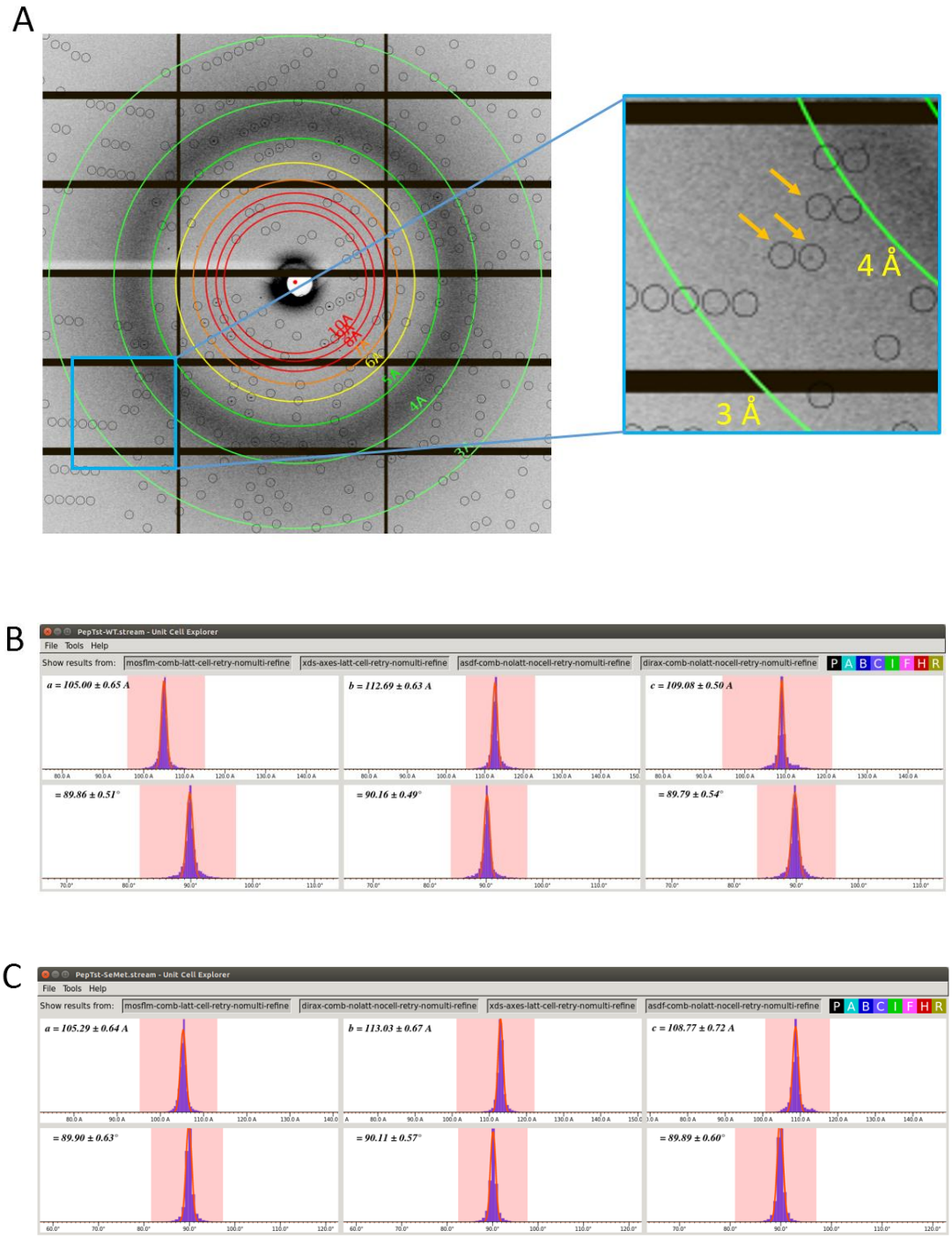


Figure 7-15: Data collection on PepT$_{St}$-WT and PepT$_{St}$-SeMet

Example diffraction pattern from the PepT$_{St}$-SeMet dataset with circles surrounding predicted peak positions and measured reflections (A), with the yellow arrows indicating the position of some of them. Output of CELL_EXPLORER of CRYSTFEL, showing the distribution of unit cell parameters for PepT$_{St}$-SeMet (B) and PepT$_{St}$-WT (C).

7.2.6.2 Phasing of $PepT_{St}$ data set

The structure factors of the native $PepT_{St}$ data set were used for MR in PHASER (McCoy *et al.*, 2007) using an available structure of $PepT_{St}$ (PDB-ID: 5D58, 2.4 Å). A TFZ score of 9.8 was found; a value over 8 indicates a definite solution has been located. The output MTZ-file contained the merged data and phases from the native data set. Initial attempts to use the derivative data to carry out either SIR or SAD failed. Since both data sets show isomorphism when comparing the unit cell parameters (Figure 7-15B and C), it was surprising that isomorphous replacement did not succeed as a phasing method in this case. This could be due to the fact that the data sets for both native and derivative data contain data from thousands of individual crystals, possibly resulting in significant non-isomorphism within the datasets themselves. This can be seen as the spread of their Gaussian distribution of unit cell parameters for each protein in Figure 7-15.

The MR solution structure was used as a guide to verify the presence and position of the selenium atoms in the derivative data. For this, the data collected from $PepT_{St}$-SeMet crystals was processed with CRYSTFEL and merged using the implemented software PARTILATO. The output MTZ-file was converted for subsequent use in the *SHELX* suite, since CRYSTFEL provided a list of reflection but without sorting and separating the intensities for the Bijvoet pair (I+ and I-). Therefore, SHELXC was used to sort the reflections accordingly, which is a vital step towards solving the structure via experimental phasing using SAD.

SHELXC was additionally used to identify the extent of the anomalous signal present in the data set and to generate the extra columns containing the intensities for the Bijvoet pairs. Anomalous signal was found to a resolution between 3.75 Å and 3.45 Å for the derivative data set processed here (Table 7-2), when using the heuristic approach of d''/sig (strength of anomalous signal d'') converging to a value of 0.8.

Using the phases from the MR solution together with the amplitudes of the anomalous differences, generated using CTRUNCATE (French and Wilson, 1978; Padilla and Yeates, 2003) the phased anomalous Fourier map was calculated. The positions of each sulphur atom in the MR output PDB-file containing the structure of $PepT_{St}$ were extracted and used to compare them to the peaks found in the anomalous Fourier map. Atom positions that did not correspond to any signal were removed from the coordinate file. Both the native and de-

rivative data set were scaled and combined into one MTZ-file. Together with the information on the heavy atom positions, the MTZ-file was used as input to perform SIRAS using SHARP (Bricogne *et al.*, 2003) with the native data set as reference.

Table 7-2: Anomalous signal statistics from SHELXC for $PepT_{St}$

Resl. (Å)	**16.07**	**9.89**	**7.45**	**6.09**	**5.21**	**4.59**	**4.12**	**3.75**	**3.45**	**3.21**	**3.0**
N(data)	104	314	527	740	968	1181	1432	1669	1908	2079	2410
<I/sig>	14.3	23.3	21.2	16.1	13.5	15.1	13.8	9	6.5	4.1	2.7
Compl. (%)	93.7	100	100	100	100	100	100	100	100	100	100
Multipl.	1.5	1.7	1.8	1.8	1.9	1.9	1.9	1.9	1.9	1.9	1.9
R_{pim} (%)	5.98	4.45	4.41	5.99	6.73	5.5	5.4	8.17	10.5	17.97	28.14
R_{anom} (%)	11.97	8.9	8.81	11.97	13.46	11.01	10.81	16.33	21.04	35.95	56.27
<d"/sig>	1.09	1.65	1.48	1.35	1.18	1.06	0.94	0.91	0.77	0.82	0.71

Resl – resolution; N(data)-number of reflections; <I/sig> - average intensity over sigma; Compl. – completeness; Multpl. – multiplicity; R_{pim} - $\boldsymbol{R_{pim}} = \frac{\sum_{hkl}\sqrt{\frac{1}{n-1}}\sum_{j=1}^{n}|\boldsymbol{I}_{hkl,j}-\langle\boldsymbol{I}_{hkl}\rangle|}{\sum_{hkl}\sum_{j}\boldsymbol{I}_{hkl,j}}$, ; *Ranom* - $\boldsymbol{R_{anom}} = \frac{\sum_{h}|\boldsymbol{I}_{(-h)}-\boldsymbol{I}_{(+h)}|}{\sum_{h}\overline{\boldsymbol{I}_{(h)}}}$; <d''/sig> - strength of anomalous signal d''

SIRAS combines SAD and SIR phasing methods, overcoming the phase ambiguity (Figure 7-16). For SIR experiments, knowing the amplitude of each structure factor of the native protein structure (F_P) using the Harker diagram, a circle (here for the structure factor of a single reflection) having the radius of the amplitude of F_P can be drawn (Figure 7-16, green circle). The aim is to obtain the phase angle (φ_P) for F_P. Knowledge of the phase and amplitude of a heavy atom (F_A) incorporated into the native structure can be used. Using this vector as an offset, a second circle (Figure 7-16, orange circle) can be drawn with a radius of the amplitude of the derivative protein structure (F_{PA}). This circle intersects the circle of F_P at two points giving F_P two possible phase angles as the solution. This is known as the phase ambiguity.

Adding the information from a SAD experiment can help to overcome the phase ambiguity. Selenomethionine for example can be used for SAD and SIRAS given that the native and derivative structure are still isomorphous. By changing the data collection wavelength, anomalous scattering from the selenium can be detected which results in an intensity difference between the Bijvoet pairs. Instead of one F_{PA} amplitudes for F_{PA+} and F_{PA-} are recorded. Extending the information from the SIR experiment using this information, two circles with radii of the amplitudes of F_{PA+} and F_{PA-}, are added to the Harker diagram (Figure 7-16, red dotted circles). Now the circles intersect only at one point and reveal the phase

angle of the reflection, thereby solving the phase ambiguity that one would have faced using SIR or SAD alone.

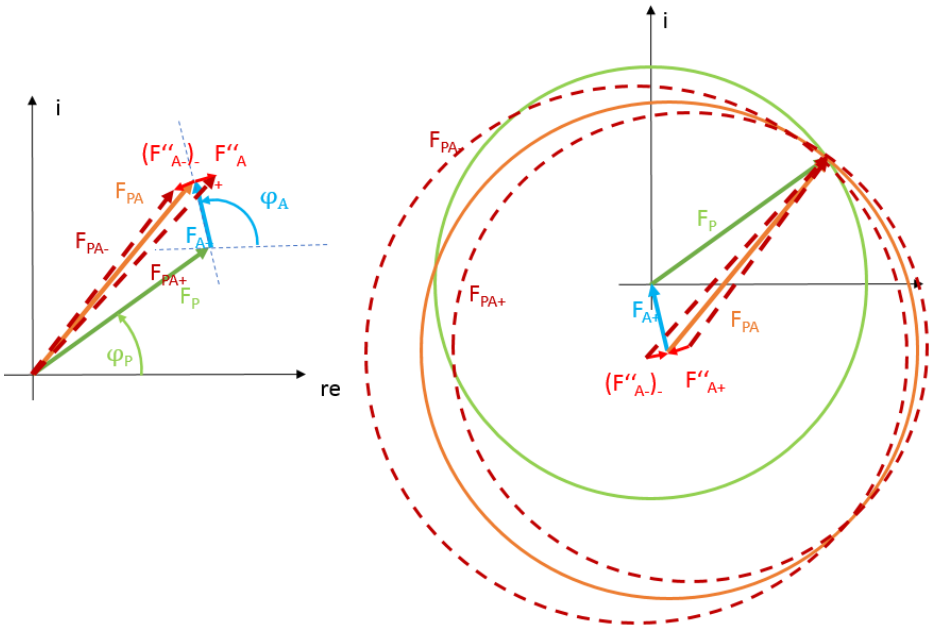


Figure 7-16: SIRAS

Combination of SIR and SAD experiments overcomes the phase ambiguity, resulting in a single solution of the phase information for the native protein structure. Determining the phase angle can be solved geometrically as shown here. The length of the vectors in the left diagram represent the amplitude of the structure factor *F*. For SIRAS, the structure factor for the derivative (F_{PA}) and native (F_{PA}) protein structure are recorded and in addition, so are the anomalous differences for the derivative protein (F_{PA-}, F_{PA+}).

It was not possible to identify the selenium sites in the derivative data solely from the data collected during the SX experiment. Instead, using the prior information on the positions of the selenium in the structure using the anomalous difference Fourier map, it was possible to generate an electron density map with interpretable character. Thus, the maps verify that the data indeed had the potential to solve the structure via experimental phasing, but also showed that more data should have been collected to reach this goal. There is no definitive answer as to the amount of data that needed to be further collected in order to obtain phase information via first principle methods.

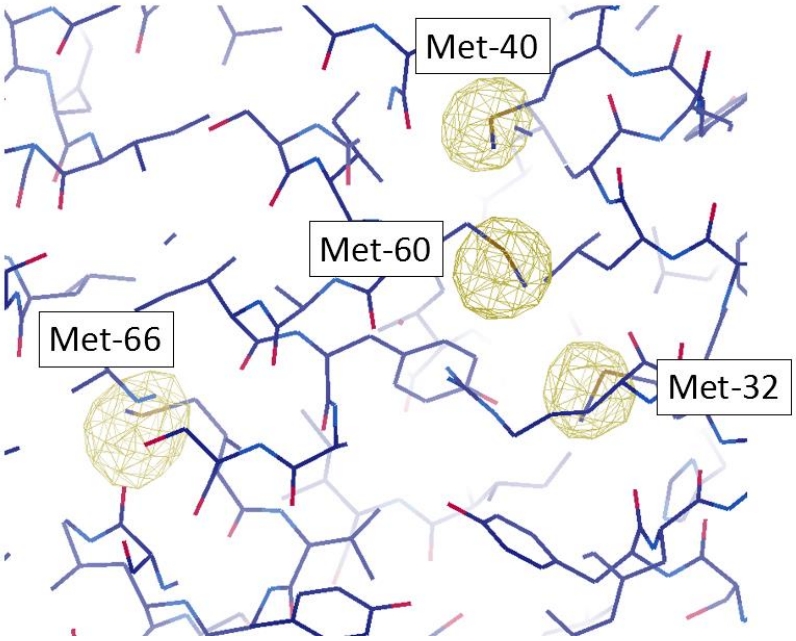


Figure 7-17: Anomalous peaks around methionines of $PepT_{St}$

Peaks at 4.5 σ in anomalous Fourier map were found around methionine residues indicating the positions of the SeMet in the derivative data.

7.2.6.3 Summary

The serial crystallography experiment conducted at beamline P11, PETRAII, resulted in complete data sets from crystals of $PepT_{St}$-WT and its SeMet derivative. Only diffraction at low resolution was observed for *Gk*ApcT crystals during this experiment, and the crystal density and indexed hits were not sufficient to generate a complete data set. Thus, further data collection on those crystals was aborted.

Nevertheless, the $PepT_{St}$ data sets were used for experimental phasing from SX experiments using a synchrotron source. Difficulties were experienced when using first principle methods and although the unit cell parameters for both data sets suggested isomorphism, the variety of unit cell dimension within each data set made single isomorphous replacement impossible. Additionally, SAD phasing also led to no results, raising the question of whether any anomalous signal was collected at all. Subsequent analysis of the data and molecular replacement with a known structure of $PepT_{St}$ revealed difference density around the sulphur atoms of various methionines in the structure, verifying that anomalous signal was present at the

correct positions required for SIR or SAD phasing. Due to low anomalous signal at low resolution, the phasing programs were unable to locate the substructure sites. Therefore, the use of prior knowledge on the positions of the selenium atoms was used, which succeeded in the generation of an electron density map in which helices of the protein could be made out, showing that experimental phasing with LCP grown crystals using SX at synchrotron sources can be successful. Although the experiment here did not lead to phasing using first principle methods, it could nevertheless be used to guide model building if an MR model is available and if verification on the correct position of the model built is needed. This is the case if the resolution of the data set is not high enough or the electron density map after MR does not show characteristic side chain features which could help in placing the protein structure. The experiment conducted here additionally verifies that the collection of SX data at a synchrotron source is feasible and can lead to a full data set being obtained from the protein crystals.

8 Final remarks and future perspectives

When this research project started in 2013, only one structure of a distantly related bacterial SLC7 transporter was known, *Mj*ApcT (Shaffer *et al.*, 2009) and it was proposed to be a proton-coupled amino acid transporter. This protein belongs to the APC family of amino acid transporters among which several structures have been solved, sharing a similar overall 3D-fold, namely the LeuT-fold, but which differ in their substrate specificity and coupling mechanism. The crystal structure of *Mj*ApcT was solved in the absence of a substrate molecule, but a water cavity in the centre of the protein indicated the location of the substrate binding site. Mutational studies on *Mj*ApcT suggested a lysine residue at position 158 to be involved in the proton-coupling mechanism, being the only charged residue that was found around the central cavity of the transporter. Additionally, mutation of this residue to alanine abolished alanine uptake (Shaffer *et al.*, 2009).

Initial homologue screening of different thermophilic bacterial and archaeal transporters identified *Gk*ApcT as a suitable candidate for crystallisation and functional studies, as outlined in Chapter 3. Given its close relationship to the SLC7 family, it was thought that this transporter could give insight into the transport mechanism of the human CAT proteins, as well as elucidate the function of the archaeal protein *Mj*ApcT.

Extensive optimisation and screening of crystallisation conditions resulted in crystals of *Gk*ApcT using the lipidic cubic phase method. Addition of cholesterol to the lipid mixture increased the crystal size and diffraction quality of the protein crystals, although this approach has mainly been used for eukaryotic membrane proteins such as GPCRs. X-ray diffraction experiments at micro-focus beamlines allowed the collection of a full data set to 2.86 Å and solution of the structure of *Gk*ApcT.

The resulting structure shows the protein in the substrate-bound inward-occluded state. Comparison with other substrate-bound APC transporters revealed a similar binding mode of the alanine molecule bound between the unwound segments of TM1 and TM6, which are conserved structural features in the APC family. Additionally, the alanine molecule coordinates a water molecule with the backbone of TM1, similar to the Na1 of LeuT, therefore

suggesting that the ion-binding site is conserved evolutionarily, and that the ion is exchangeable in this region between different transporters. It was thus verified that *Gk*ApcT belongs to the APC family sharing the LeuT-fold topology, although sharing only low sequence identity with *Mj*ApcT and LeuT.

The structure of *Gk*ApcT provided insight into the binding site and showed that the steric limitations of the binding site cavity are the main determinants of its substrate specificity for small neutral and polar amino acids. Furthermore, additional structural features were observed, such as the interaction with a single TM protein, YneM, and the interaction with cholesterol. This protein originates from the expression host *E.coli*, and was unexpectedly co-purified with *Gk*ApcT but it was shown that the specific interaction was enhanced when a cholesterol molecule was between them. Furthermore, thermal stability data suggested that cholesterol had a stabilising effect on *Gk*ApcT. Purification of *Gk*ApcT from an *E.coli* YneM knock-out strain and repetition of the thermal stability assays revealed that the stabilising effect of cholesterol originated from the interaction of YneM and *Gk*ApcT via this molecule. It was also surprising to see this cholesterol interacting with *Gk*ApcT. Similar molecules, called hopanoids, are found in some bacteria and fungi, increasing the rigidity of the cell membrane to withstand extreme environments. This interaction has not been reported for any other bacterial transporter and may represent a novel, and as yet undiscovered role for cholesterol-like molecules in prokaryotic membrane protein function.

The structure of *Gk*ApcT also presented several novel structural features for the LeuT-like fold, which included a lateral intracellular helix and a β-hairpin domain. The first of these derived from an N-terminal extension that was found in *Gk*ApcT and not *Mj*ApcT. This helix interacts with other helices in *Gk*ApcT and likely stabilises the protein, as suggested from the decreased thermal stability of the N-terminal deletion.

The β-hairpin domain is an antiparallel β-sheet found between TM8 and TM9, protruding from the protein in an angle of around 90° and leaving the water filled cavity below the substrate binding site open. In *Mj*ApcT an unstructured loop which appears to close the intracellular cavity is found in a similar position. Since both structures were solved at different pHs, it is suggested that this domain occludes the protein from the intracellular side and is involved in the substrate release via a protonation/deprotonation mechanism. The loop region also contains a pattern of positively charged residues, which raised the question of whether

this domain might be involved in the interaction with the lipid headgroups. Mutational studies targeting the proline residues at the end of the β-sheet drastically reduced the uptake of alanine, indicating that this domain has an important effect on the function of *Gk*ApcT. Similar to GadC's C-terminal plug involved in closing the substrate binding site, a comparable mechanism might occur for *Gk*ApcT.

To further understand the membrane transporter on the molecular level, *in vitro* assays were performed. Liposome-based assays provide a system where the buffer compositions on both sides of the vesicle membrane can be easily altered to examine the uptake of substrate under different conditions.
Functional studies revealed that *Gk*ApcT is a proton-coupled amino acid transporter specific for small neutral and polar amino acids, although no proton-coupling mechanism could be verified for *Mj*ApcT. This makes the structure of *Gk*ApcT the first of a proton-coupled APC transporter to date. Initially lacking further structural information, investigations made at the early stages of the project identified residues that were thought to be important for alanine binding. The structural data provided verification of the mutational studies, and residues proposed to be important for the function of the transporter were investigated.

In order to understand how the transporter showed a different substrate spectrum but high sequence identity with the human CAT transporters, comparison of the binding site residues between them enabled the identification of residues that might determine the substrate specificity and subsequent mutational studies were performed. These experiments showed that a single point mutation at position Met321 to serine gave *Gk*ApcT the opportunity to transport arginine. Additionally, the mutant was used for crystal trials using arginine as a substrate. A crystal structure with a bound arginine molecule was obtained to a resolution f 3.13 Å. The arginine molecule was found in a similar location as to alanine in the wildtype, and its location verified the assumption that the serine at position 321 would open the cavity so that the arginine could fit into the binding site to form interactions with Glu115 and Asp237. Thus, *Gk*ApcT is a valid model for understanding the human proteins and verifies a similar location of their binding site until more structural data are available.

Further inspection of the structure revealed several interesting structural features, such as the N-terminal lateral helix which is not present in *Mj*ApcT and which would hinder the proposed kink movement of TM1, as seen for LeuT (Krishnamurthy and Gouaux, 2012). Therefore, a different translocation mechanism was proposed for *Gk*ApcT, based also on the results from the generated repeat-swapped model, in which TM6 is the helix performing the 'kink' movement. This is supported by mutational studies that showed that Asp237 is a conserved residue between *Gk*ApcT and the SLC7 family, forming a hydrogen bonding network downstream of the substrate binding site, although Glu115 might also be involved in this mechanism. It is proposed that Glu115 undergoes a deprotonation on the intracellular side of the membrane either inducing the TM6 movement or reversing the transporter into its initial state. Further studies need to be performed to identify the protonation/deprotonation mechanism of this residue and its influence on the transport cycle.

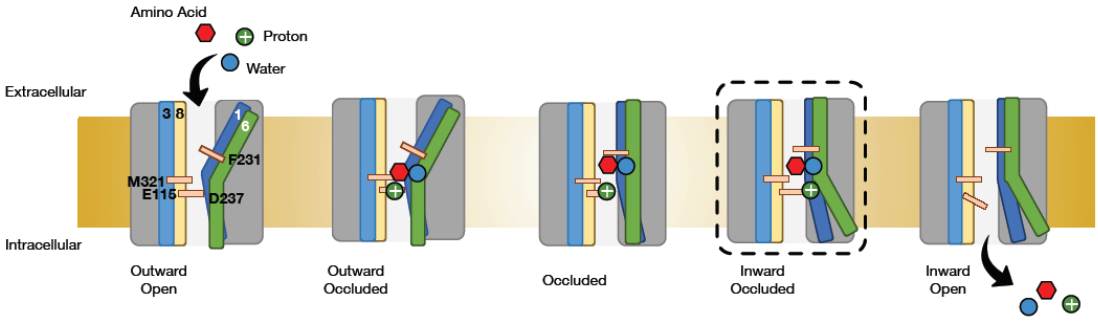


Figure 8-1: Proposed alternating access mechanism of GkApcT

Important residues are labelled and shown as light orange rectangles. Helices involved in the translocation mechanism and carrying the labelled residues are coloured and indexed. Arrows indicate the movement of the ion, water and substrate. The different conformational states of the transporter are given underneath each illustration.

The proposed mechanism illustrated in Figure 8-1 follows the following steps:

- Phe231 is flipped aside through the bending of TM6 to open the substrate binding cavity, as reported for AdiC with a kink movement of TM1
- Substrate and water bind in the SBS and Glu115 is protonated.
- TM1 and TM6 change their conformation and close the binding site at which Phe231 is sealing the SBS (occluded state)
- TM6b kinks aside, disrupting the interaction of Glu115-water-Asp237, causing the deprotonation of Glu115 and substrate release into the intracellular environment.

The work presented here provides the foundation for further work to understand the transport cycle of *Gk*ApcT. In particular, this would mean investigating of the protonation states of Glu115 and Asp237 and other charged residues that might be involved in proton-coupling. Further studies might reveal how the release of the substrate to the intracellular side is facilitated, but also allow an understanding of the helical movement and if they differ from the results from work on other APC members. Additionally, the results of this project delivered new information on the SLC7 family. The interaction of *Gk*ApcT with YneM might give the location of the single helix of the SLC3 family with which the human HATs interact and understanding of how these small proteins interact with other transmembrane proteins to stabilise them. Perhaps a possible application can be found for difficult targets in terms of stabilisation in solution or during expression trials. The interaction of YneM with *Gk*ApcT via cholesterol also revealed a novel and unexpected interaction of a bacterial membrane protein with a single TM protein. Therefore, the interactions of bacterial transporters with hopenes are another path worth investigating, as well as the interaction via cholesterol between the SLC7 and SLC3 family. The effect of sterols on the activity and stability of the LAT1-4F2hc heterodimer has been reported recently (Dickens *et al.*, 2017).

A deeper understanding of the different substrate affinities of the CAT2 proteins is necessary to understand how the single mutation of Arg334 to glutamate has its effect. Although located downstream of the substrate binding site, the mutation changed *Gk*ApcT's substrate affinity. Thus, the residues must be involved in the translocation mechanism by not directly influencing substrate binding and its affinity but by indirectly affecting the conformational changes of the transporter. How this is achieved remains to be determined. One way might be the use of electro paramagnetic resonance (EPR) studies to resolve the conformational dynamics of the protein for the wildtype and Arg334Glu mutant, to investigate conformational differences or differences in the velocity of the helical movements.

In general, it would be interesting to unravel the dynamics of the helical movements and the transport cycle, in particular, the mechanism of substrate release on the intracellular side which was proposed here to depend on the kink movement of TM6. This would be different to the observations made for other APC transporters where normally TM1 kinks away to release the substrate and ions.

A more detailed insight into the transport mechanism of APC transporters could potentially be gained using SX as a tool. Although, the method of time-resolved studies using SX or SFX is restricted to photosensitive proteins, the introduction of site-specific unnatural amino acids might overcome this obstacle. They could be used to cage the transporter and to initiate conformational changes that could be triggered by a laser at a distinct wavelength. The use of unnatural amino acids, introduced at specific positions in the protein sequence, additionally offers the possibility of other methods to being used, such as them acting as a probe for NMR or determining the protonation state of a specific residue (Noren *et al.*, 1989; Dougherty, 2000; Groff *et al.*, 2010).

*Gk*ApcT and $PepT_{St}$ were used for SX experiments at a synchrotron source showing the possibilities and difficulties that can be faced when preparing for these kinds of experiments. The two proteins behaved differently in the optimisation process for the SX experiment and upscaling from the conventional LCP plates took several attempts before it was successful.
In the end, it was not possible to collect sufficient data for *Gk*ApcT, but phasing of $PepT_{St}$ using SIRAS was possible although not directly. SX can be a powerful tool if only small crystals for a protein are available and a reasonable density of them can be obtained for data collection. As unconventional crystallography, a bit of luck is necessary to obtain the desired results.

The combination of structural studies and biochemical assays revealed the transport mechanism of *Gk*ApcT and showed how substrate recognition by this transporter is facilitated and can be altered. The information gathered for *Gk*ApcT additionally shed light into the mode of action of the CAT transporters.

Appendix

APPENDIX CHAPTER 1

Table A 1: Published crystal structures of APC superfamily members available in the PDB

Family	Name	Organism	PDB-ID-code (resolution in Å)	Associated publication
NSS[1]	LeuT	*Aquifex aeolicus*	4MM4 (2.89), 4MM5 (3.2), 4MM6 (3.1), 4MM7 (2.85), 4MM8 (3.31), 4MM9 (2.9), 4MMA (3.3, 4MMB (2.25), 4MMC (2.3), 4MMD (2.3), 4MME (2.5), 4MMF (2.7)	(Wang *et al.*, 2013)
			4HOD (3.3), 4HMK (3.0)	(Kantcheva *et al.*, 2013)
			4FXZ (2.6), 4FY0 (3.0)	(Wang and Gouaux, 2012)
			3USG (2.5), 3USI (3.11), 3USJ (3.5), 3USK (4.5), 3USL (2.71), 3USM (3.01), 3USO (4.5), 3USP (2.1)	(Wang *et al.*, 2012)
			3TT1 (3.1), 3TT3 (3.22), 3TU0 (2.99)	(Krishnamurthy and Gouaux, 2012)
			3QS4 (2.63), 3QS5 (2.6), 3QS6 (2.8)	(Piscitelli and Gouaux, 2012)
			3MPN (2.25), 3MPQ (2.25)	(Kroncke *et al.*, 2010)
			3GWU (2.14), 3GWV (2.35), 3GWW (2.46)	(Zhou *et al.*, 2013)
			3GJC (2.8), 3GJD (2.0)	(Quick *et al.*, 2009)
			3F3A (2.0), 3F3C (2.1), 3F3D (2.3), 3F3E (1.8), 3F48 (1.9), 3F4I (1.95), 3F4J (2.15)	(Singh *et al.*, 2008)
			2Q6H (1.85), 2Q72 (1.7), 2QB4 (1.9), 2QEI (1.85)	(Singh *et al.*, 2007)
			2QJU (2.9)	(Zhou *et al.*, 2007)
			2A65 (1.65)	(Yamashita *et al.*, 2005)
NSS[1]	MhsT	*Bacillus halodurans*	4US3 (2.1), 4US4 (2.6)	(Malinauskaite *et al.*, 2014)
APC[2]	AdiC	*Escherichia coli*	5J4I (2.21), 5J4N (2.59)	(Ilgü *et al.*, 2016)
			3OB6 (3.0)	(Kowalczyk *et al.*, 2011)
			3LRB (3.61), 3LRC (4.0), 3H5M (3.61), 3H6B (4.0)	(Gao *et al.*, 2009)
			3L1L (3.0)	(Gao *et al.*, 2010)
		Salmonella enterica	3NCY (3.2), 3HQK (3.2)	(Fang *et al.*, 2009)

APC[2]	ApcT	*Methanocaldococcus jannaschii*	3GIA (2.32), 3GI8 (2.59), 3GI9 (2.48)	(Shaffer *et al.*, 2009)
NSS[1]	vSGLT	*Vibrio parahaemolyticus*	2XQ2 (2.73)	(Watanabe *et al.*, 2010)
			3DH4 (2.7)	(Faham *et al.*, 2008)
NCS1[3]	Mhp1	*Microbacterium liquefaciens*	4D1A (3.4), 4D1B (3.8), 4D1C (3.7), 4D1D (3.7)	(Simmons *et al.*, 2014)
			2JLN (2.85), 2JLO (3.8)	(Weyand *et al.*, 2008)
			2X79 (3.8)	(Shimamura *et al.*, 2010)
BCCT[4]	BetP	*Corynebacterium glutamicum*	4LLH (2.8)	(Perez *et al.*, 2014)
			3P03 (3.35)	(Perez *et al.*, 2011)
			2WIT (3.35)	(Ressl *et al.*, 2009)
			4C7R (2.7)	(Koshy *et al.*, 2013)
			4DOJ (3.25), 4AIN (3.1)	(Perez *et al.*, 2012)
BCCT[4]	CaiT	*Proteus mirabilis*	4M8J (3.29)	(Kalayil *et al.*, 2013)
			2WSX (3.5), 2WSW (2.29)	(Schulze *et al.*, 2010)
			3HFX (3.15)	(Tang *et al.*, 2010)
APC[2]	GadC	*Escherichia coli*	4DJI (3.19), 4DJK (3.1)	(Ma *et al.*, 2012b)
EAAT[5]	GltPh	*Pyrococcus horikoshii*	1XFH (3.5)	(Yernool *et al.*, 2004)
			2NWL (2.96), 2NWW (3.2), 2NWX (3.29)	(Boudker *et al.*, 2007)
			3V8F (3.8), 3V8G (4.66)	(Verdon and Boudker, 2012)
			4IZM (4.5)	(Reyes *et al.*, 2013)
			4OYE (4.0), 4OYF (3.41), 4OYG (3.5), 4P19 (3.25), 4P1A (3.75), 4P3J (3.5), 4P6H (4.08), 5CFY (3.5)	(Verdon *et al.*, 2014)
			3KBC (3.51)	(Reyes *et al.*, 2009)
	EAAT1	*Homo sapiens*	5LLM (3.25), 5LLU (3.32), 5LM4 (3.1), 5*MJ*U (3.71)	(Canul-Tec *et al.*, 2017)
NCS2[6]	UraA	*Escherichia coli*	3QE7 (2.78)	(Lu *et al.*, 2011)
			5IC6 (3.7)	(Alguel *et al.*, 2016)
NSS[1]	DAT	*Drosophila melanogaster*	4M48 (2.95)	(Penmatsa *et al.*, 2013)
			4XP1 (2.89), 4XP4 (2.8), 4XP5 (3.3), 4XP6 (3.1), 4XP9 (2.8), 4XPA (2.95), 4XPB (3.05), 4XPF (3.27), 4XPG (3.21), 4XPH (2.9), 4XPT (3.36)	(Wang *et al.*, 2015)
			4XNU (2.98)	(Penmatsa *et al.*, 2015)
	SERT	*Homo sapiens*	5I6X (3.14), 5I6Z (4.53), 5I71 (3.15), 5I73 (3.24), 5I74 (3.39), 5I75 (3.49)	(Coleman *et al.*, 2016)

NRAMP[7]	ScaDMT	*Staphylococcus capitis*	4WGV (3.1)	(Ehrnstorfer *et al.*, 2014)
	DraNramp	*Deinococcus radiodurans*	5KTE (3.94)	(Bozzi *et al.*, 2016)
SulP/SLC26[8]	SLC26Dg	*D. geothermalis*	5DA0 (3.2)	(Geertsma *et al.*, 2015)
AE	AE_{CTD}	*Homo sapiens*	4YZF (3.5)	(Arakawa *et al.*, 2015)

[1]NSS - Neurotransmitter:Sodium Symporter Family, [2] APC - Amino Acid-Polyamine-Organocation Family, [3] NCS1 - Nucleobase:Cation Symporter-1 Family, [4]BCCT- Betaine/Carnitine/Choline Transporter Family, [5] EAAT- Excitatory Amino Acid Transporter, [6] NCS2 - Nucleobase:Cation Symporter-2 Family, [7] NRAMP - ''natural resistance-associated'' macrophage protein (Metal Ion Transporter) Family, [8] SulP – Sulfate Permease Family, [9] AE – Anion Exchanger family

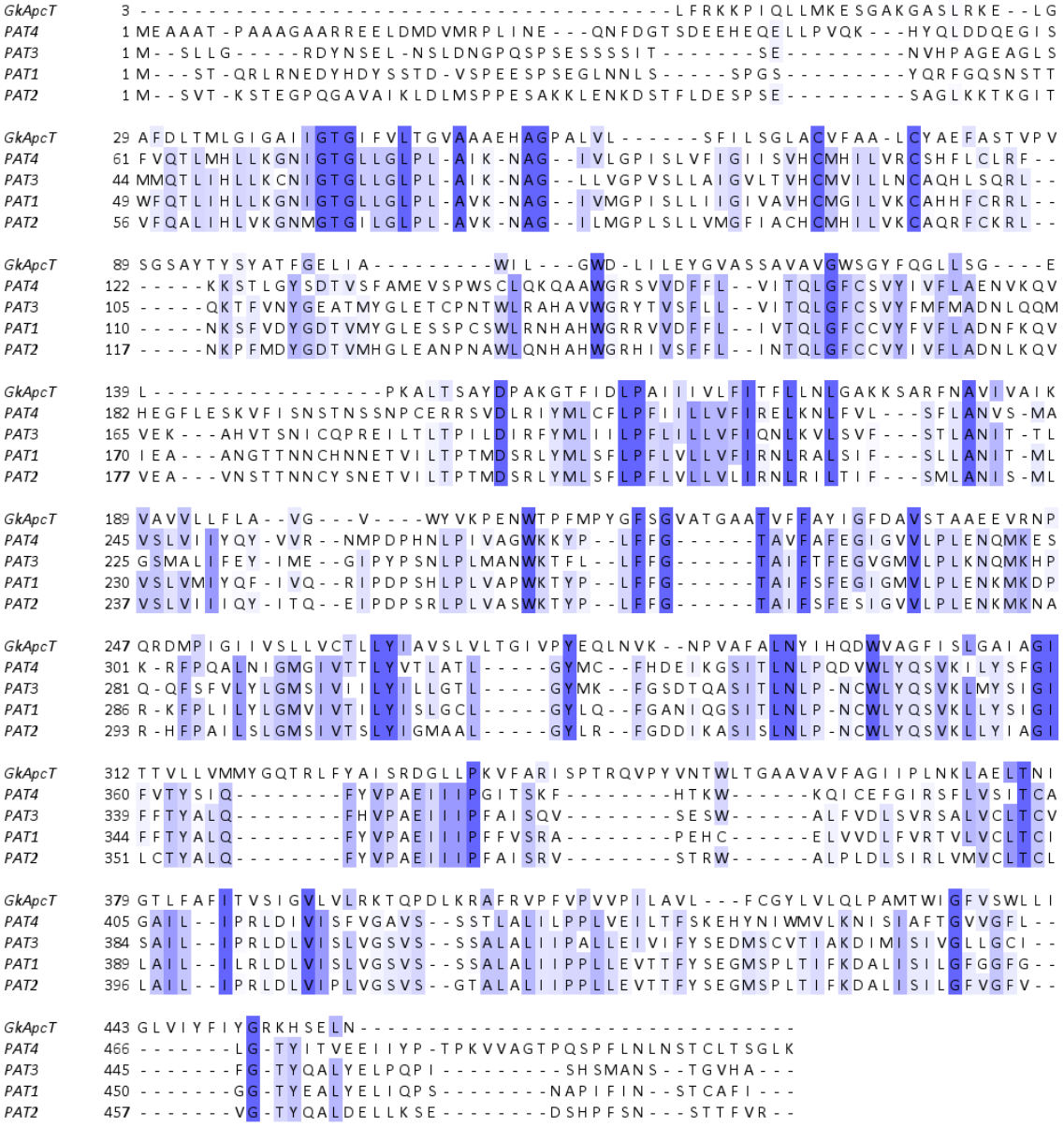


Figure A 1: Sequence alignment of GkApcT and hPATs for homology modelling

Sequence alignment of *Gk*ApcT with *h*PATs were aligned using PROMALS3D (Pei and Grishin, 2014). The alignment is coloured by conservation from high (deep purple) to no conservation (white).

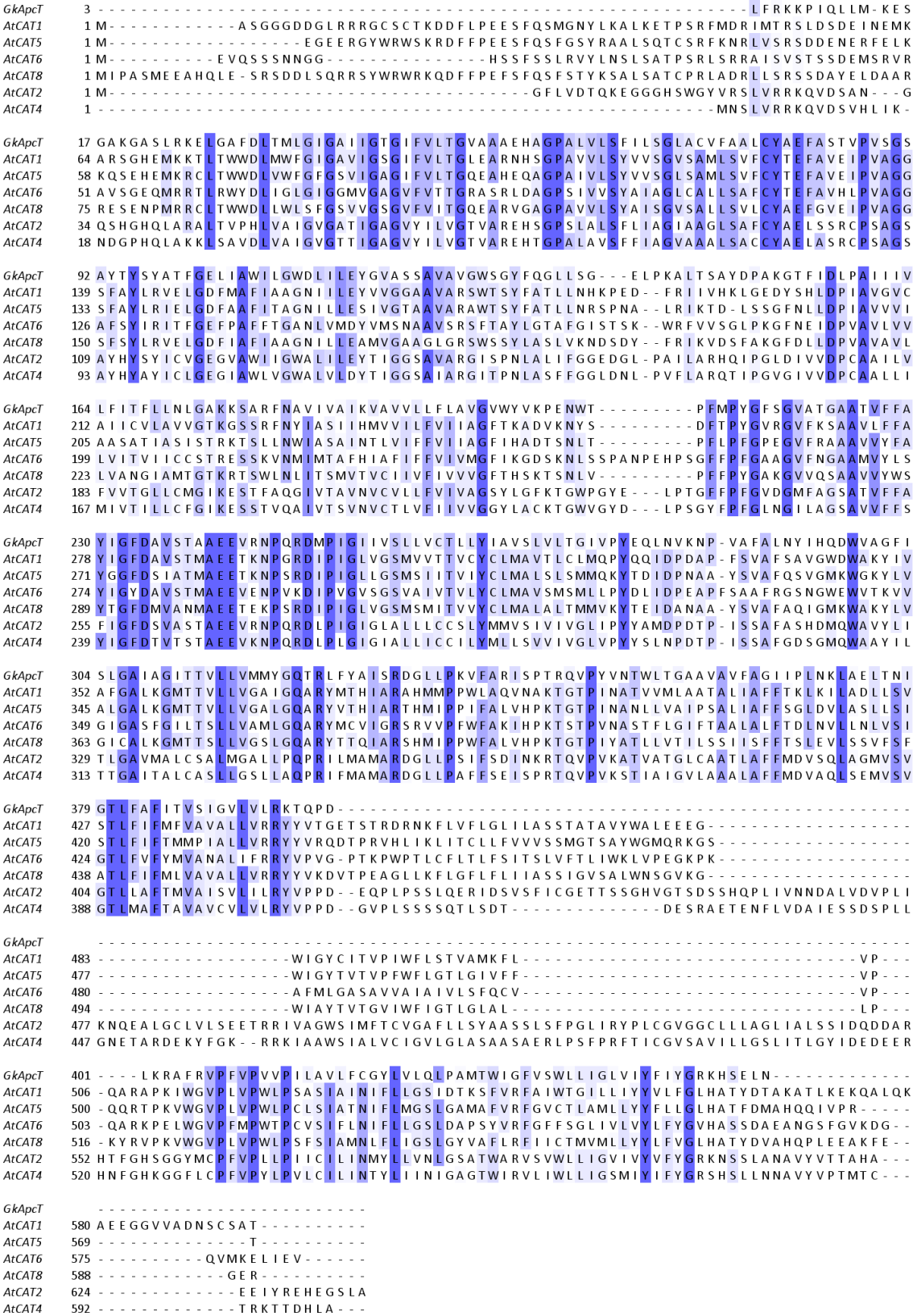


Figure A 2: Sequence alignment of GkApcT and AtCATs for homology modelling

Sequence alignment of *Gk*ApcT with *At*CATS were aligned using PROMALS3D (Pei and Grishin, 2014). The alignment is coloured by conservation from high (deep purple) to no conservation (white).

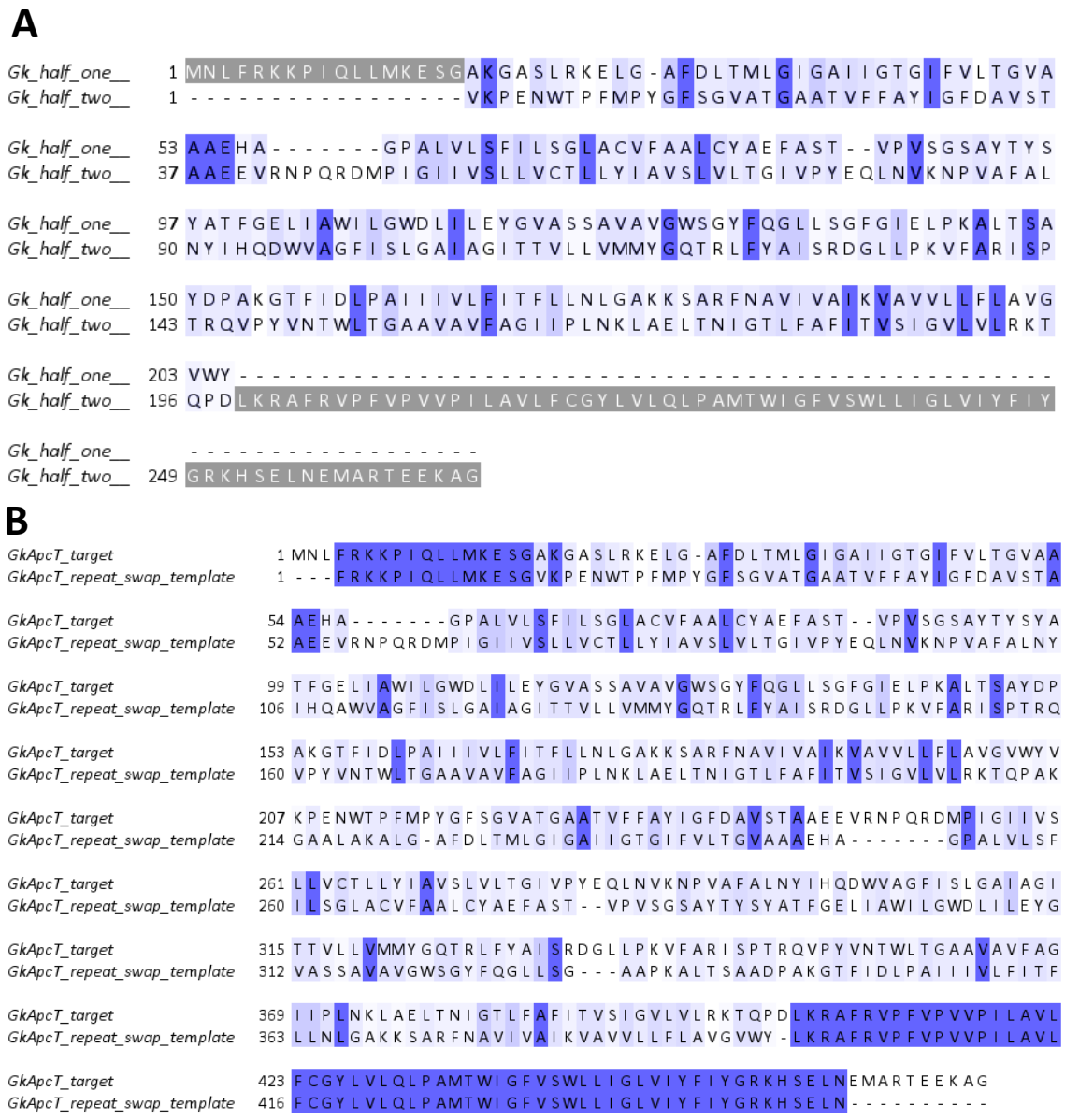


Figure A 3: Sequence alignments for repeat-swap modelling

Sequence alignment of the two symmetrical halves of *Gk*ApcT were aligned using ALIGNME(Stamm *et al.*, 2014) (A). The sequence of GkApcT as target was then aligned to the repeat-swap model using the alignment from the symmetrical repeats (B). The alignments are coloured by conservation from high (deep purple) to no conservation (white). The grey coloured sequence parts in (A) correspond to the lateral helix, and TM11 and 12.

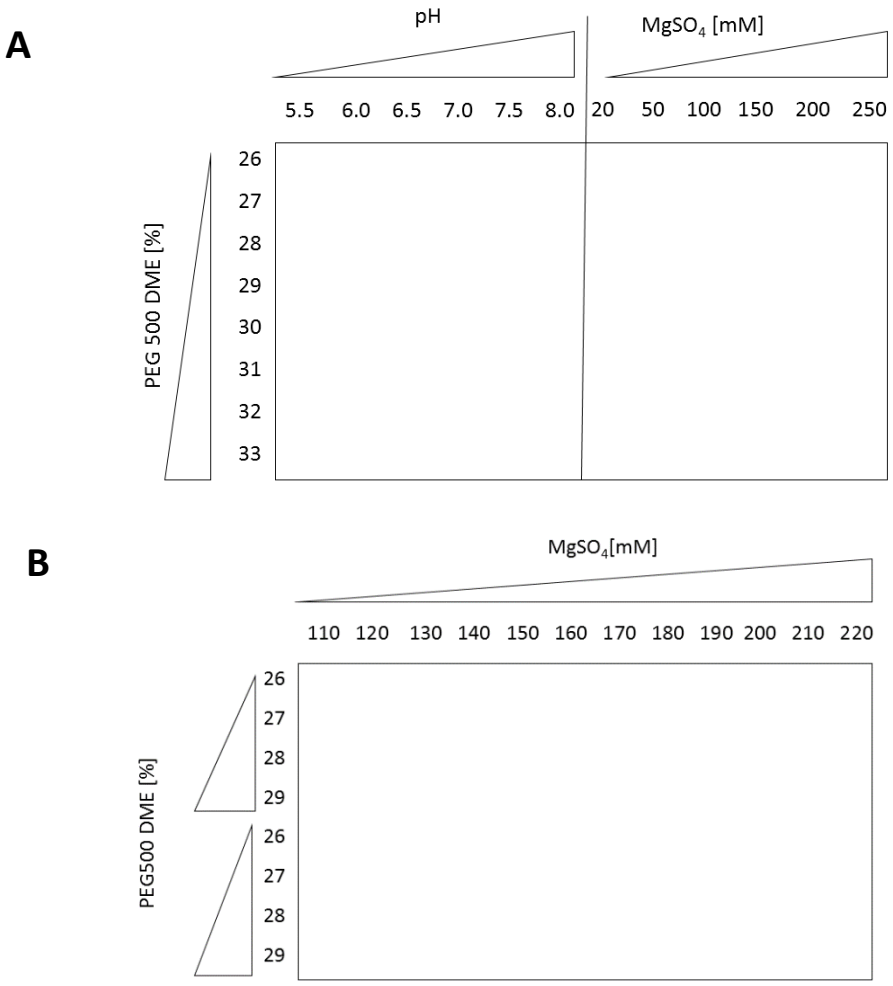


Figure A 4: Optimisation screens for crystallisation of MjApcT in LCP

Optimisation screens (A and B) for LCP crystallisation trials with *Mj*ApcT were designed based on the initial crystallisation condition containing 30 % (v/v) PEG500 DME, 100 mM $MgSO_4$, 100 mM MES pH 6.0. Buffers for the two screens had a final concentration of 100 mM. For screening the different pHs the following buffers were used: sodium citrate pH 5.5, MES pH 6.0 and pH 6.5, HEPES-NaOH pH 7.0, Tris-HCl pH 7.5 and pH 8.0.

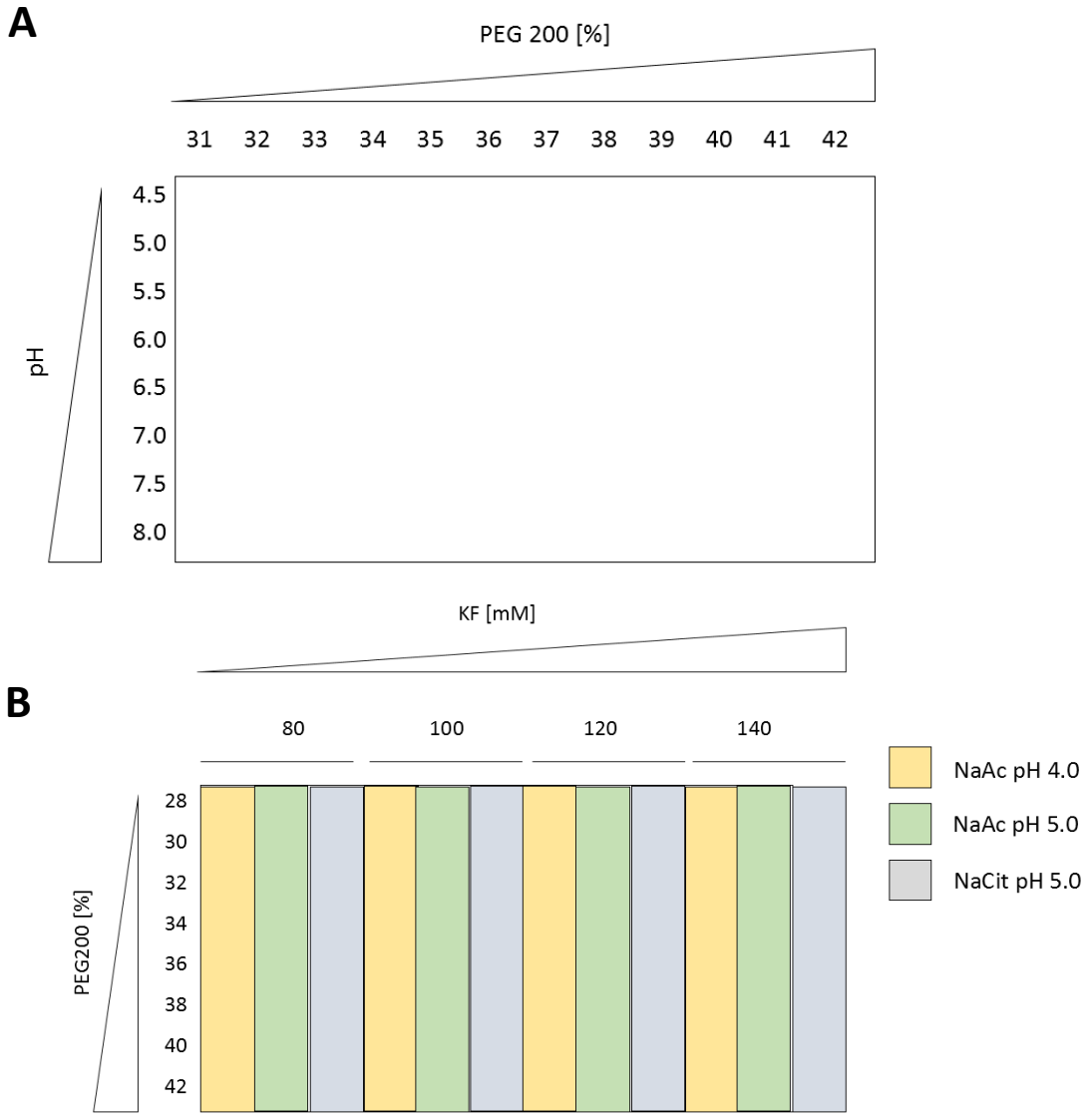


Figure A 5: Optimisation screen for crystallisation of GkApcT in LCP

An optimisation screen based on condition A8 of MemMeso (0.1 M NaAc pH 4.5, 30 % (v/v) PEG200) was designed with varying PEG concentrations and different pH (A). Crystals from this optimisation screen diffracted to a maximum of 4.5 Å. Therefore, the additional use of the Hampton stock options screen identified potassium fluoride (KF) as suitable additive. Additionally, another crystallisation screen was designed with varying salt and PEG concentrations (B). The screen contained 10 mM alanine and the buffer concentration was 100 mM.

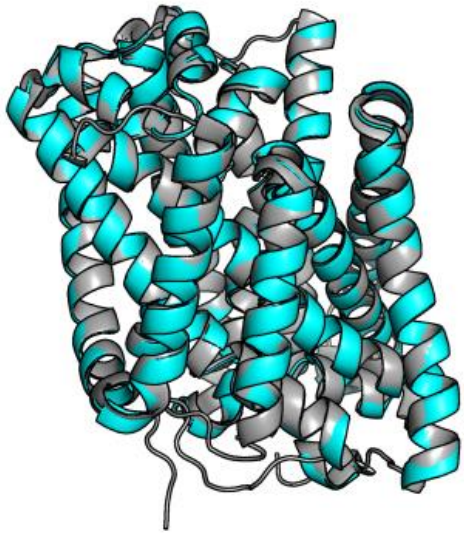


Figure A 6: Structural alignment of MjApcT structure of LCP grown crystals and PDB 3GIA

Structural data obtained from LCP grown crystals of *Mj*ApcT to a resolution of 3.5 Å resulted in a structural model (blue) which closely resembles (RMSD: 0.36 Å) the published structure of *Mj*ApcT from vapour diffusion experiments (grey, PDB-ID: 3GIA)

Table A 2: Anomalous signal statistics from SHELXC for I23 data

Resl. (Å)	**24**	**14.78**	**11.12**	**9.1**	**7.78**	**6.85**	**6.15**	**5.6**	**5.16**
N(data)	19	120	199	270	356	438	521	591	677
<I/sig>	9.7	10.4	9.6	9.4	7.7	5	2.7	2.1	1.5
Compl. (%)	40.4	96.8	99	98.5	99.4	99.3	99.6	99.7	99.7
Multipl.	8.2	11	11.7	12	12.2	12.3	12.4	12.3	12.6
R_{pim} (%)	5.67	5.41	5.29	5.49	5.93	11.89	27.37	44.91	55.51
R_{anom} (%)	7.66	6.12	9.11	8.41	10.46	22.18	54.43	93.9	123.4
<d"/sig>	1	0.54	0.64	0.63	0.69	0.76	0.72	0.77	0.84

Resl – resolution; N(data)-number of reflections; <I/sig> - average intensity over sigma; Compl. – completeness; Multpl. - multiplicity; R_{pim} - $\boldsymbol{R_{pim}} = \frac{\sum_{hkl}\sqrt{\frac{1}{n-1}}\sum_{j=1}^{n}|I_{hkl,j}-\langle I_{hkl}\rangle|}{\sum_{hkl}\sum_{j} I_{hkl,j}}$, ; Ranom - $\boldsymbol{R_{anom}} = \frac{\sum_{h}|I_{(-h)}-I_{(+h)}|}{\sum_{h}\overline{I_{(h)}}}$; <d''/sig> - strength of anomalou signal d''

Table A 3: Anomalous signal statistics from SHELXC for BL1A-data

Resl. (Å)	**15**	**9.23**	**6.95**	**5.68**	**4.86**	**4.28**	**3.84**	**3.5**	**3.22**	**2.99**
N(data)	157	459	766	1085	1394	1714	2040	2331	2665	2945
<I/sig>	32.9	36.5	31.2	19.1	16.4	13.1	5.6	1.8	1	0.9
Compl. (%)	95.2	100	100	99.9	99.4	99.1	98.4	97.7	96.7	94.4
Multipl.	18.7	24.9	26.3	25.8	24.5	22.8	24.2	26.5	27.1	23.4
Rpim (%)	2.6	1.54	1.89	3.73	4.48	6.22	16.16	42.36	60.94	52.23
Ranom (%)	11.24	3.78	3.8	7.01	9.36	13.66	34.8	88.3	143.4	150.7
<d"/sig>	1.4	1.01	0.83	0.89	0.91	0.94	1.12	1.01	0.67	0.45

Resl – resolution; N(data)-number of reflections; <I/sig> - average intensity over sigma; Compl. – completeness; Multpl. – multiplicity; Rpim - $\boldsymbol{R_{pim}} = \frac{\sum_{hkl}\sqrt{\frac{1}{n-1}}\sum_{j=1}^{n}|I_{hkl,j}-\langle I_{hkl}\rangle|}{\sum_{hkl}\sum_{j} I_{hkl,j}}$, ; Ranom - $\boldsymbol{R_{anom}} = \frac{\sum_{h}|I_{(-h)}-I_{(+h)}|}{\sum_{h}\overline{I_{(h)}}}$; <d''/sig> - strength of anomalous signal d''

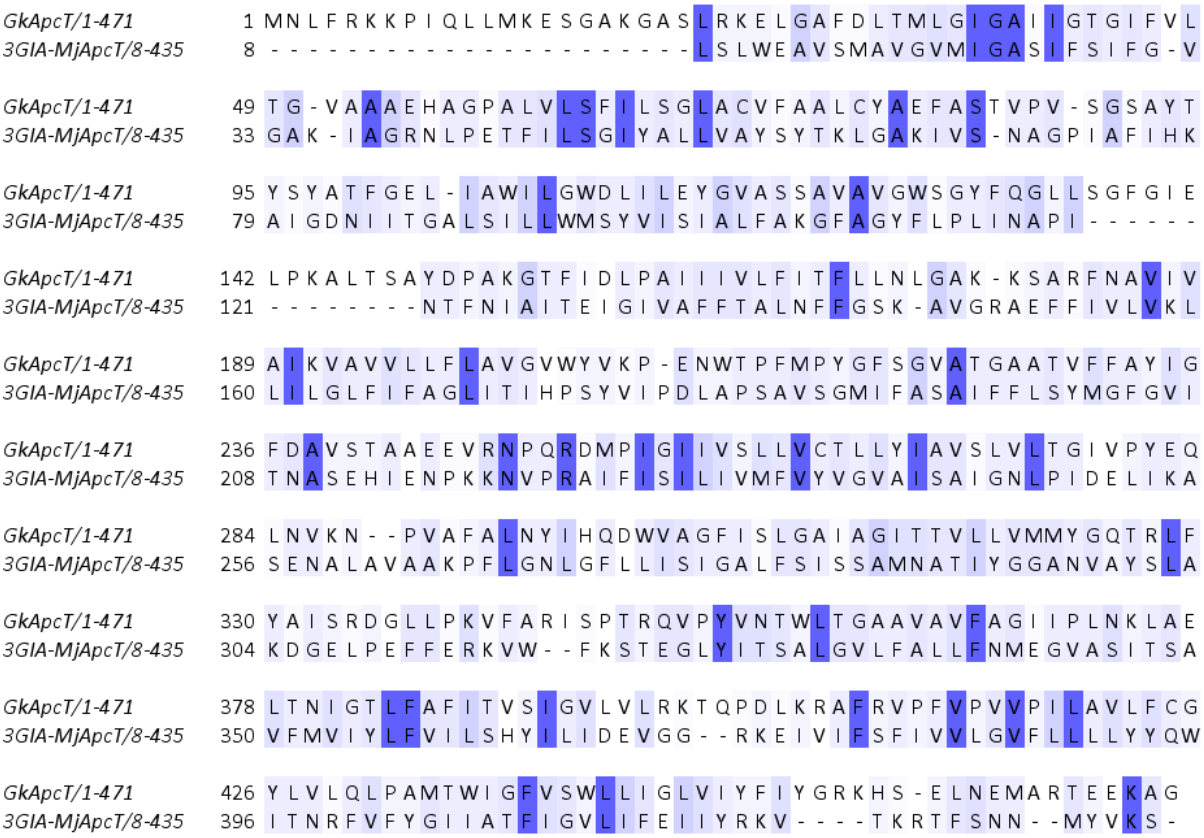


Figure A 7: Sequence alignment between MjApcT and GkApcT for CHAINSAW

To generate the MR with CHAINSAW (Stein, 2008) a sequence alignment between *Gk*ApcT and the structure *Mj*ApcT (PDB-ID:3GIA, 2.32Å) was generated with PROMALS3D(Pei and Grishin, 2014). The sequence alignment coloured by conservation (from high (deep purple) to low (white)).

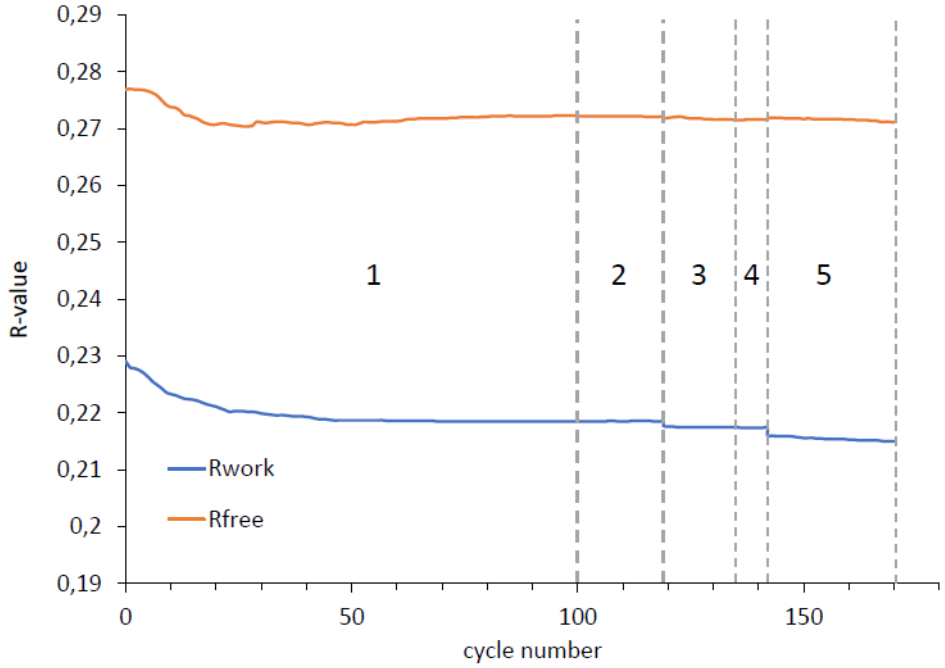


Figure A 8: Rwork/Rfree vs. cycle number of final refinement step

R_{work} and R_{free} were plotted against the cycle number as output by BUSTER. The TLS refinement in BUSTER includes the following 5 subcycles: 1 - the atoms are fixed and the isotropic B-factors are refined under standard B-correlation restraints; 2 - the atoms are allowed to move, and the TLS model refined in the previous cycle is refining the anisotropic and isotropic component of the B-factors; 3 - the atoms are fixed where they left the second cycle, and the TLS model parameters and isotropic B-factors are refined; 4 & 5 - atoms are allowed to move, with their total B-factors described via the same combination of the new TLS model parameters and the individual isotropic atomic B-factors

Accession codes for Phylogenic tree in Figure 5-7

Sequences of *Gk*ApcT (Q5L1G5), *Mj*ApcT (Q58026), PotE (P0AAF1), AdiC(P60061), LysP (P25737), GabP (P25527), AroP (P15993), PheP (P24207), LeuT (O67854), GNP1 (P48813), CAN1 (P04817), SLC7A1 (P30825), SLC7A2 (P52569), SLC7A3 (Q8WY07), SLC7A4 (O43246), SLC7A5(Q01650), SLC7A6 (A0AUM0), SLC7A7 (Q9UM01), SLC7A8 (Q9UHI5), SLC7A9 (P82251), SLC7A10 (Q9NS82), SLC7A11 (Q9UPY5), SLC7A13 (Q8TCU3), SLC12A1 (Q13621), SLC12A2 (P55011), SLC12A3 (P55017), SLC12A4 (Q9UP95), SLC12A5 (Q9H2X9), SLC12A6 (Q9UHW9), SLC12A7 (Q9Y666), SLC32A1 (Q9H598), SLC36A1 (Q7Z2H8), SLC36A2 (Q495M3), SLC36A3 (Q495N2), SLC36A4 (Q6YBV0), SLC38A1 (Q9H2H9), SLC38A2 (Q96QD8), SLC38A3 (Q99624), SLC38A4 (Q8R1S9) and SLC38A5 (Q8WUX1) were used for an alignment using PROMALS3D (Pei and Grishin, 2014).

Table A 4: Data collection and refinement parameters of GkApcT-Met321Ser

	GkApcT-Met321Ser (PDB-ID: 6F34)
Beamline	Proxima2A, Soleil
Detector	EIGER X 9M
Wavelength (Å)	0.98
Temperature (K)	100
Osc. (°)	0.1
Exp. Time (s)	0.1
Total number of images	360
Space group	$P2_12_12_1$
Unit cell a b c (Å)	77.04, 82.70, 118.78
α β γ (°)	90.0, 90.0, 90.0
Mosaicity (°)	0.211
Wilson B-factor (Å²)	81.70
Resolution (Å)[1]	64.63-3.13 (3.21-3.13)
R_{merge}	0.173(1.452)
R_{meas}	0.188 (1.581)
R_{pim}	0.074 (0.618)
Total number of observations	89902 (6552)
Unique observations	13951 (1032)
I/σ(I)	5.5 (1.1)
$CC_{1/2}$	0.998 (0.679)
Completeness (%)	99.9 (99.8)
Multiplicity	6.4 (6.3)
Refinement	
Refinement program	BUSTER
Resolution (Å)	64.58 - 3.13(3.24-3.13)
CC*	0.999 (0.915)
Reflections used in refinement	13760 (1340)
Reflections used for R-free	666 (37)
R_{work}	0.2309 (0.3414)
R_{free}	0.2951 (0.4146)
Number of non-hydrogen atoms	3802
macromolecules	3711
ligands	53
Protein residues	486
RMS_{bonds}	0.014
RMS_{angles}	1.73
Ramachandran favored (%)	92
Ramachandran outliers (%)	1.3
Rotamer outliers (%)	5.4
Clashscore	4.92
Average B-factor (Å²)	89.19

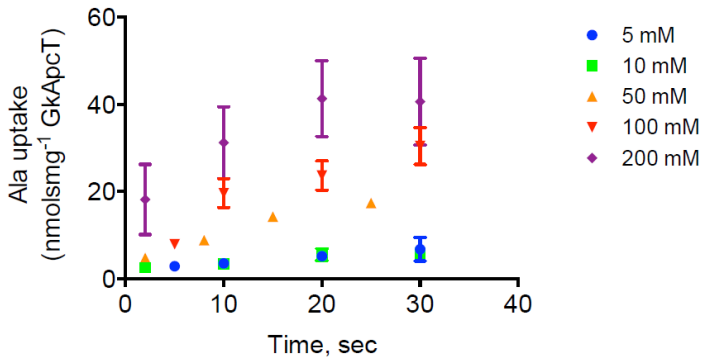


Figure A 9: Kinetics of GkApcT

Time course measurements of alanine uptake under $\Delta\Psi$-driven conditions and varying alanine concentrations.

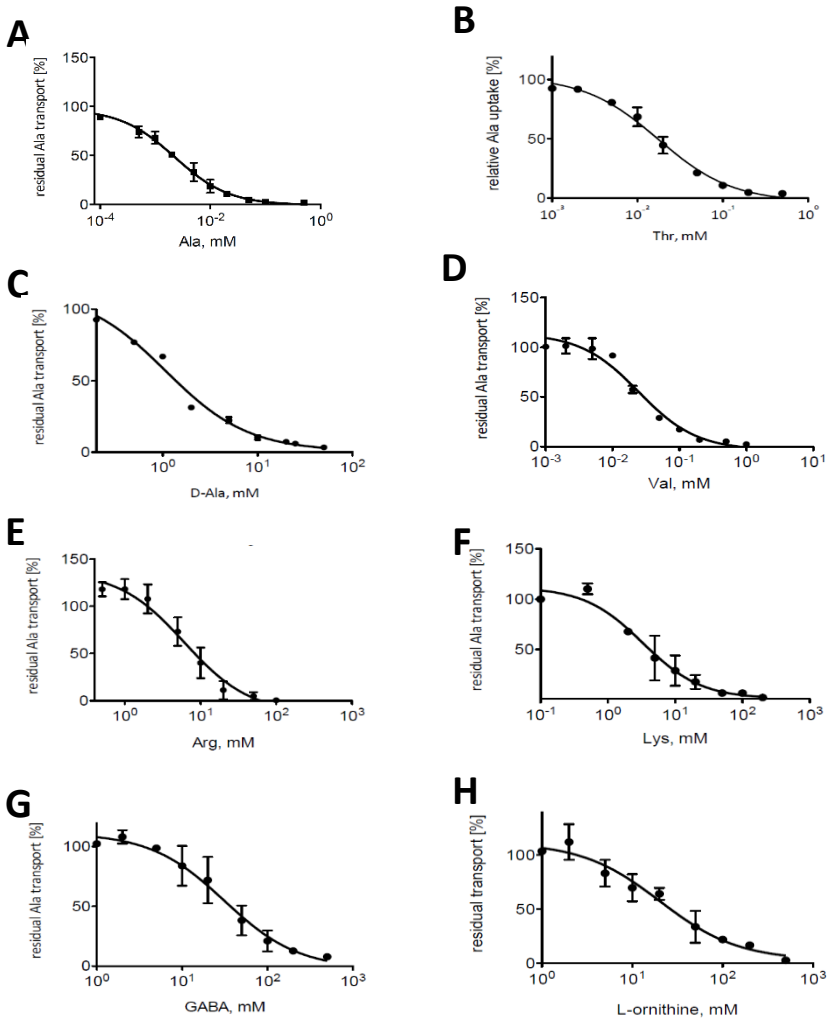


Figure A 10 IC_{50} experiments for GkApcT

IC_{50} measurements were carried out under electrogenic conditions. Varying concentrations of competitor were added to 250 nM ^{3}H-Ala in the external buffer and the reaction was stopped after 1 min. The resulting IC_{50} curves were analysed in PRISM using a sigmoidal curve fitting method.

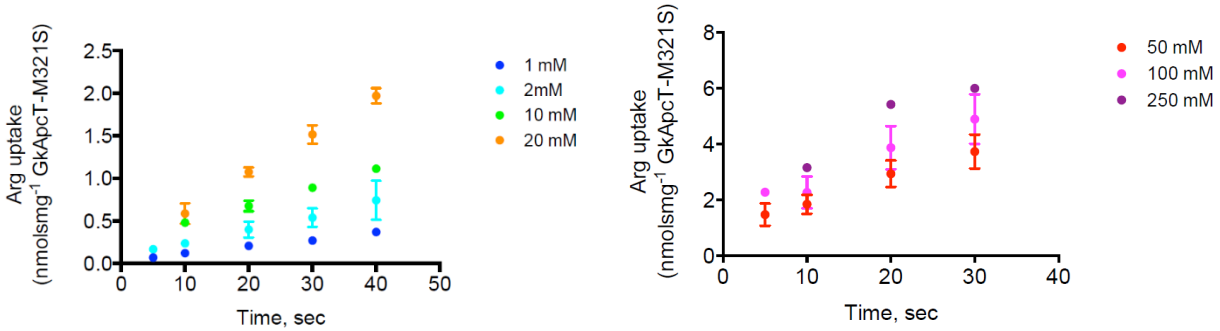


Figure A 12: Arginine transport kinetic for GkApcT-Met321Ser

Time course measurements of alanine uptake under $\Delta\Psi$-driven conditions and varying arginine concentrations.

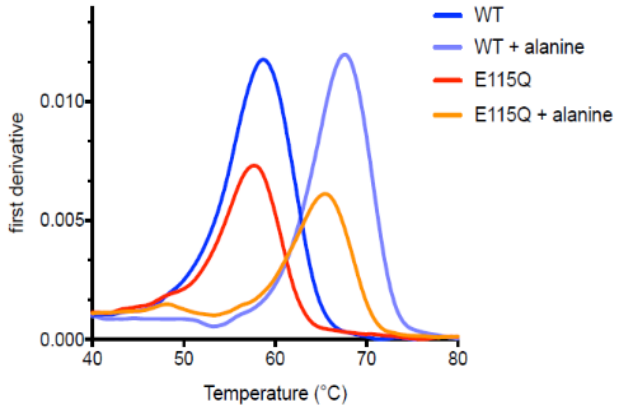


Figure A 11: Thermal stability of GLu115Gln in presence of the substrate alanine

0.5 mg/ml of protein in 25 mM citrate phosphate pH 6.0, 100 mM NaCl, 2 mM $MgSO_4$, 0.03 % DDM and with (+ alanine) or without 10 mM alanine were used for thermal fluorimetry using the PROMETHEUS NT.48 (NanoTemper). The first derivative of the initial $F_{330\,nm}/F_{350\,nm}$ plot against the temperature is shown.

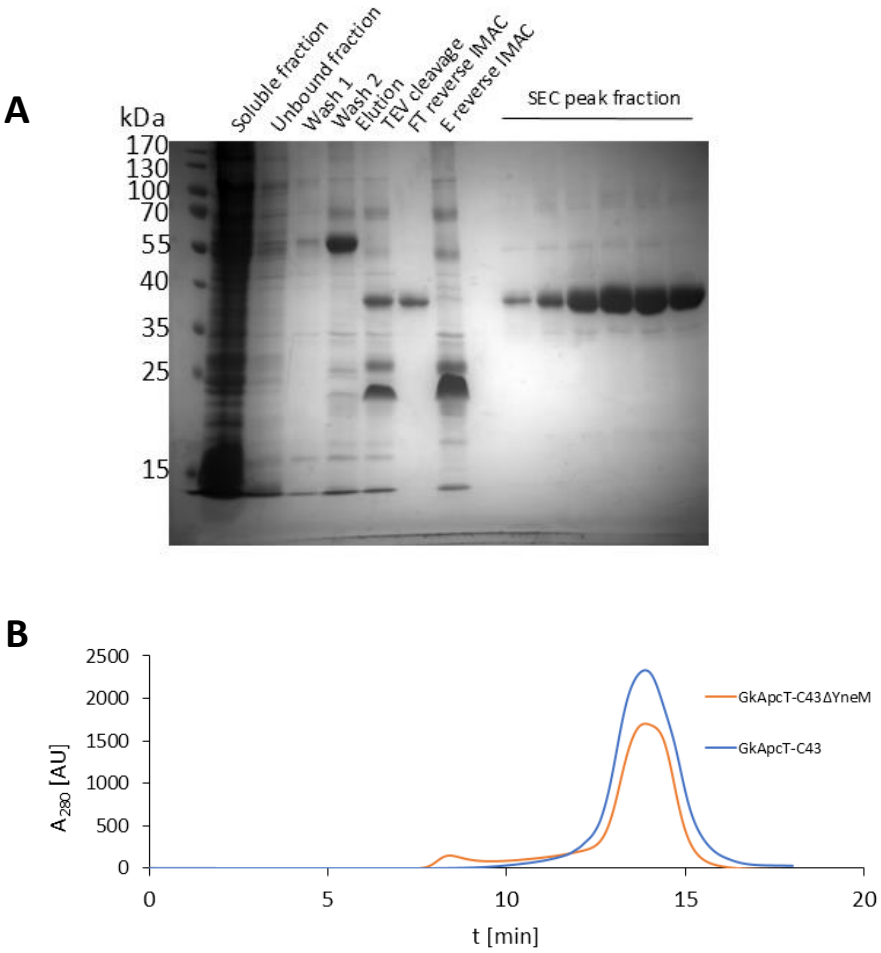


Figure A 13: Purification of GkApcT-C43ΔYneM

*Gk*ApcT was expressed in the C43(DE3)Δ*yneM* strain and purified as previously described. The gel shows that the protein could be successfully purified to high purity (A) and the size exclusion shows a monodisperse peak after similar elution time as GkApcT from common C43(DE3) cells (B).

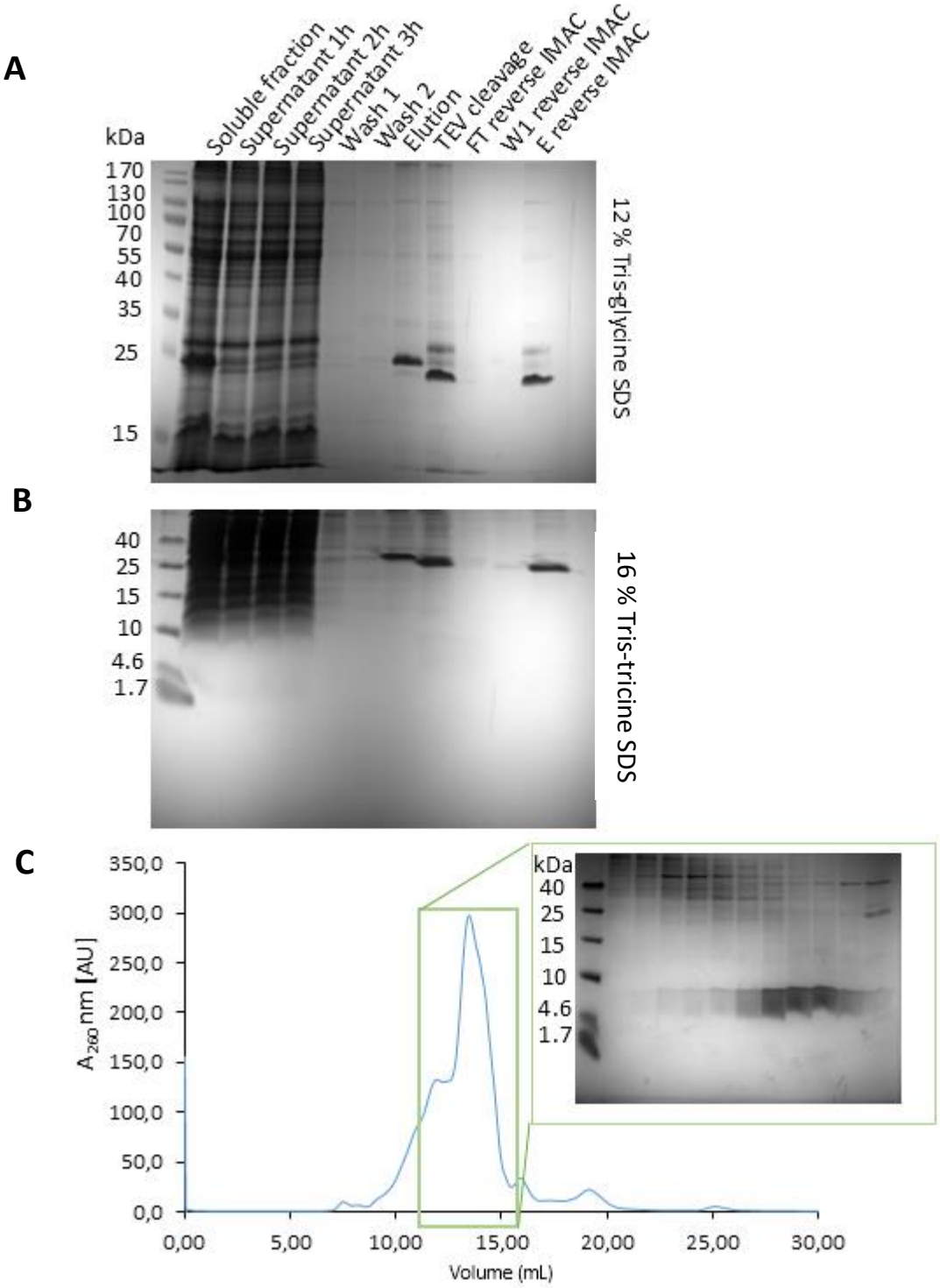


Figure A 14: Purification of full-length YneM

YneM was successfully purified. The 4.5 kDa protein was difficult to visualise on the 12 % Tris-glycine gel (A). Therefore, a 16% Tris-tricine SDS gel was used as well for each sample (B) but the protein could not be observed. Only after size exclusion when the protein was more concentration in the individual peak fractions, could YneM be seen on the Tris-tricine gel (C). The protein shows a size of ~ 9 kDa which might indicate a dimerization.

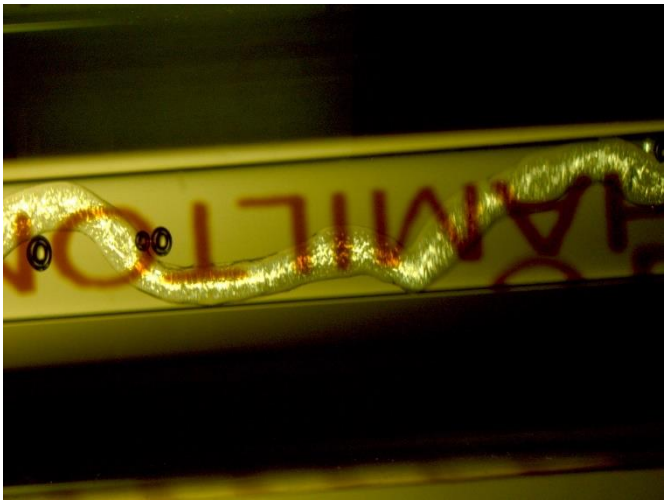


Figure A 15: Crystallisation of GkApcT-SeMet in the syringe setup

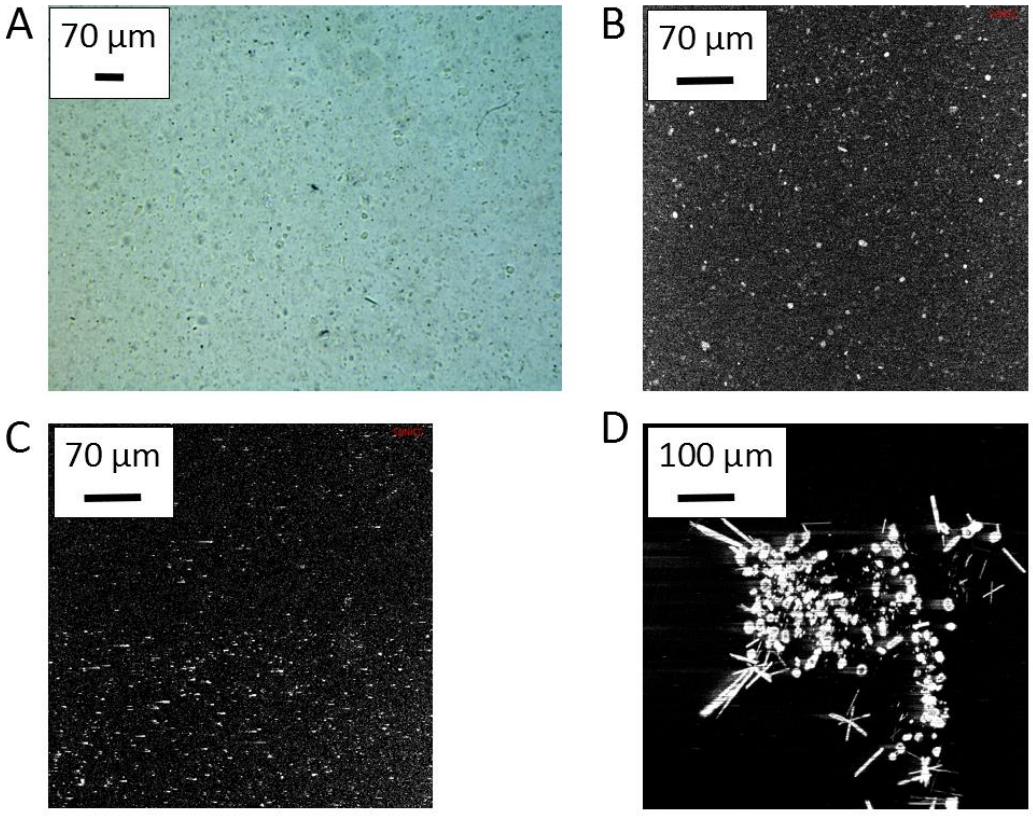


Figure A 16: Confirmation of $PepT_{St}$ crystals by UV and SONICC

Crystal lipid mixture was dispensed onto single wells of an LCP plate and the mixture inspected by bright field microscopy (A), UV (B) and SONICC (C) confirming that the crystals observed via the cross polariser are protein crystals. During image analysis salt crystals formed in one sample of PepTSt, that showed a positive signal when using SONICC, showing that the method can in some cases not be used to distinguish protein from salt crystals (D).

Publications

Jungnickel, K. E. J., Parker, J., and Newstead, S. 2018. Structural basis for amino acid transport by the CAT family of SLC7 transporters. *Nature communications* 9:550

Bibliography

Abramson, J. 2003. Structure and Mechanism of the Lactose Permease of *Escherichia coli*. *Science* 301: 610–615.

Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H. and Adams, P.D. 2012. Towards automated crystallographic structure refinement with *phenix.refine*. *Acta Crystallographica Section D Biological Crystallography* 68: 352–367.

Alberts, B. 2002. *Molecular Biology of the Cell*, 4th ed. Garland Science, New York .

Albritton, L.M., Tseng, L., Scadden, D. and Cunningham, J.M. 1989. A putative murine ecotropic retrovirus receptor gene encodes a multiple membrane-spanning protein and confers susceptibility to virus infection. *Cell* 57: 659–666.

Alguel, Y., Amillis, S., Leung, J., Lambrinidis, G., Capaldi, S., Scull, N.J., Craven, G., Iwata, S., Armstrong, A., Mikros, E., Diallinas, G., Cameron, A.D. and Byrne, B. 2016. Structure of eukaryotic purine/H^+ symporter UapA suggests a role for homodimerization in transport activity. *Nature Communications* 7: 11336.

Alix, E. and Blanc-Potard, A.-B. 2009. Hydrophobic peptides: novel regulators within bacterial membrane: Regulatory membrane peptides in bacteria. *Molecular Microbiology* 72: 5–11.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215: 403–410.

Androutsellis-Theotokis, A., Goldberg, N.R., Ueda, K., Beppu, T., Beckman, M.L., Das, S., Javitch, J.A. and Rudnick, G. 2003. Characterization of a functional bacterial homologue of sodium-dependent neurotransmitter transporters. *The Journal of Biological Chemistry* 278: 12703–12709.

Aquila, A., Hunter, M.S., Doak, R.B., Kirian, R.A., Fromme, P., White, T.A., Andreasson, J., Arnlund, D., Bajt, S., Barends, T.R.M., Barthelmess, M., Bogan, M.J., Bostedt, C., Bottin, H., Bozek, J.D., Caleman, C., Coppola, N., Davidsson, J., DePonte, D.P., Elser, V., Epp, S.W., Erk, B., Fleckenstein, H., Foucar, L., Frank, M., Fromme, R., Graafsma, H., Grotjohann, I., Gumprecht, L., Hajdu, J., Hampton, C.Y., Hartmann, A., Hartmann, R., Hau-Riege, S., Hauser, G., Hirsemann, H., Holl, P., Holton, J.M., Hömke, A., Johansson, L., Kimmel, N., Kassemeyer, S., Krasniqi, F., Kühnel, K.-U., Liang, M., Lomb, L., Malmerberg, E., Marchesini, S., Martin, A.V., Maia, F.R.N.C., Messerschmidt, M., Nass, K., Reich, C., Neutze, R., Rolles, D., Rudek, B., Rudenko, A., Schlichting, I., Schmidt, C., Schmidt, K.E., Schulz, J., Seibert, M.M., Shoeman, R.L., Sierra, R., Soltau, H., Starodub, D., Stellato, F., Stern, S., Strüder, L., Timneanu, N., Ullrich, J., Wang, X., Williams, G.J., Weidenspointner, G., Weierstall, U., Wunderer, C., Barty, A., Spence, J.C.H. and Chapman, H.N. 2012. Time-resolved protein nanocrystallography using an X-ray free-electron laser. *Optics Express* 20: 2706–2716.

Arakawa, T., Kobayashi-Yurugi, T., Alguel, Y., Iwanari, H., Hatae, H., Iwata, M., Abe, Y., Hino, T., Ikeda-Suno, C., Kuma, H., Kang, D., Murata, T., Hamakubo, T., Cameron, A.D., Kobayashi, T., Hamasaki, N. and Iwata, S. 2015. Crystal structure of the anion exchanger domain of human erythrocyte band 3. *Science* 350: 680–684.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L. and Mori, H. 2006. Construction of *Escherichia coli* K-12 in-

frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology* 2.

Baker-Austin, C. and Dopson, M. 2007. Life in acid: pH homeostasis in acidophiles. *Trends in Microbiology* 15: 165–171.

Barbeau, A. 1970. Dopamine and disease. 103: 824–832.

Barends, T.R.M., Foucar, L., Botha, S., Doak, R.B., Shoeman, R.L., Nass, K., Koglin, J.E., Williams, G.J., Boutet, S., Messerschmidt, M. and Schlichting, I. 2013. De novo protein crystal structure determination from X-ray free-electron laser data. *Nature* 505: 244–247.

Barends, T.R.M., Foucar, L., Ardevol, A., Nass, K., Aquila, A., Botha, S., Doak, R.B., Falahati, K., Hartmann, E., Hilpert, M., Heinz, M., Hoffmann, M.C., Kofinger, J., Koglin, J.E., Kovacsova, G., Liang, M., Milathianaki, D., Lemke, H.T., Reinstein, J., Roome, C.M., Shoeman, R.L., Williams, G.J., Burghardt, I., Hummer, G., Boutet, S. and Schlichting, I. 2015. Direct observation of ultrafast collective motions in CO myoglobin upon ligand dissociation. *Science* 350: 445–450.

Bartoccioni, P., Del Rio, C., Ratera, M., Kowalczyk, L., Baldwin, J.M., Zorzano, A., Quick, M., Baldwin, S.A., Vázquez-Ibar, J.L. and Palacín, M. 2010. Role of transmembrane domain 8 in substrate selectivity and translocation of SteT, a member of the L-amino acid transporter (LAT) family. *The Journal of Biological Chemistry* 285: 28764–28776.

Barty, A., Caleman, C., Aquila, A., Timneanu, N., Lomb, L., White, T.A., Andreasson, J., Arnlund, D., Bajt, S., Barends, T.R.M., Barthelmess, M., Bogan, M.J., Bostedt, C., Bozek, J.D., Coffee, R., Coppola, N., Davidsson, J., DePonte, D.P., Doak, R.B., Ekeberg, T., Elser, V., Epp, S.W., Erk, B., Fleckenstein, H., Foucar, L., Fromme, P., Graafsma, H., Gumprecht, L., Hajdu, J., Hampton, C.Y., Hartmann, R., Hartmann, A., Hauser, G., Hirsemann, H., Holl, P., Hunter, M.S., Johansson, L., Kassemeyer, S., Kimmel, N., Kirian, R.A., Liang, M., Maia, F.R.N.C., Malmerberg, E., Marchesini, S., Martin, A.V., Nass, K., Neutze, R., Reich, C., Rolles, D., Rudek, B., Rudenko, A., Scott, H., Schlichting, I., Schulz, J., Seibert, M.M., Shoeman, R.L., Sierra, R.G., Soltau, H., Spence, J.C.H., Stellato, F., Stern, S., Strüder, L., Ullrich, J., Wang, X., Weidenspointner, G., Weierstall, U., Wunderer, C.B. and Chapman, H.N. 2011. Self-terminating diffraction gates femtosecond X-ray nanocrystallography measurements. *Nature Photonics* 6: 35–40.

Batyuk, A., Galli, L., Ishchenko, A., Han, G.W., Gati, C., Popov, P.A., Lee, M.-Y., Stauch, B., White, T.A., Barty, A., Aquila, A., Hunter, M.S., Liang, M., Boutet, S., Pu, M., Liu, Z. -j., Nelson, G., James, D., Li, C., Zhao, Y., Spence, J.C.H., Liu, W., Fromme, P., Katritch, V., Weierstall, U., Stevens, R.C. and Cherezov, V. 2016. Native phasing of X-ray free-electron laser data for a G protein-coupled receptor. *Science Advances* 2: e1600292.

Bill, R.M., Henderson, P.J.F., Iwata, S., Kunji, E.R.S., Michel, H., Neutze, R., Newstead, S., Poolman, B., Tate, C.G. and Vogel, H. 2011. Overcoming barriers to membrane protein structure determination. *Nature Biotechnology* 29: 335–340.

Blasse, G. and Bril, A. 1967. A new phophor for flying-spot cathode-ray tubes for color television: yellow-emitting $Y_3Al_5O_{12}$–Ce^{3+}. *Applied Physics Letters* 11: 53–55.

Bogdan, C. 2001. Nitric oxide and the immune response. *Nature Immunology* 2: 907–916.

Boll, M., Daniel, H. and Gasnier, B. 2004. The SLC36 family: proton-coupled transporters for the absorption of selected amino acids from extracellular and intracellular proteolysis. *Pflügers Archiv : European Journal of Physiology* 447: 776–779.

Boll, M., Foltz, M., Anderson, C.M.H., Oechsler, C., Kottra, G., Thwaites, D.T. and Daniel, H. 2003. Substrate recognition by the mammalian proton-dependent amino acid transporter PAT1. *Molecular Membrane Biology* 20: 261–269.

Boudker, O. and Verdon, G. 2010. Structural perspectives on secondary active transporters. *Trends in Pharmacological Sciences* 31: 418–426.

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

Bourenkov, G.P. and Popov, A.N. 2006. A quantitative approach to data-collection strategies. *Acta Crystallographica Section D Biological Crystallography* 62: 58–64.

Boutet, S., Lomb, L., Williams, G.J., Barends, T.R.M., Aquila, A., Doak, R.B., Weierstall, U., DePonte, D.P., Steinbrener, J., Shoeman, R.L., Messerschmidt, M., Barty, A., White, T.A., Kassemeyer, S., Kirian, R.A., Seibert, M.M., Montanez, P.A., Kenney, C., Herbst, R., Hart, P., Pines, J., Haller, G., Gruner, S.M., Philipp, H.T., Tate, M.W., Hromalik, M., Koerner, L.J., van Bakel, N., Morse, J., Ghonsalves, W., Arnlund, D., Bogan, M.J., Caleman, C., Fromme, R., Hampton, C.Y., Hunter, M.S., Johansson, L.C., Katona, G., Kupitz, C., Liang, M., Martin, A.V., Nass, K., Redecke, L., Stellato, F., Timneanu, N., Wang, D., Zatsepin, N.A., Schafer, D., Defever, J., Neutze, R., Fromme, P., Spence, J.C.H., Chapman, H.N. and Schlichting, I. 2012. High-Resolution Protein Structure Determination by Serial Femtosecond Crystallography. *Science* 337: 362–364.

Bozzi, A.T., Bane, L.B., Weihofen, W.A., Singharoy, A., Guillen, E.R., Ploegh, H.L., Schulten, K. and Gaudet, R. 2016. Crystal Structure and Conformational Change Mechanism of a Bacterial Nramp-Family Divalent Metal Transporter. *Structure* 24: 2102–2114.

Brehm, W. and Diederichs, K. 2014. Breaking the indexing ambiguity in serial crystallography. *Acta Crystallographica Section D Biological Crystallography* 70: 101–109.

Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M. and Paciorek, W. 2003. Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallographica. Section D, Biological Crystallography* 59: 2023–2030.

Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O.S., Vonrhein, C. and Womack, T.O. 2016. *BUSTER*. Cambridge, United Kingdom, .

Bull, S.C. and Doig, A.J. 2015. Properties of Protein Drug Target Classes. *PLOS ONE* 10: e0117955.

Byrne, E.F.X., Sircar, R., Miller, P.S., Hedger, G., Luchetti, G., Nachtergaele, S., Tully, M.D., Mydock-McGrane, L., Covey, D.F., Rambo, R.P., Sansom, M.S.P., Newstead, S., Rohatgi, R. and Siebold, C. 2016. Structural basis of Smoothened regulation by its extracellular domains. *Nature* 535: 517–522.

Caffrey, M., Li, D., Howe, N. and Shah, S.T.A. 2014. 'Hit and run' serial femtosecond crystallography of a membrane kinase in the lipid cubic phase. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369: 20130621–20130621.

Camacho, J.A., Obie, C., Biery, B., Goodman, B.K., Hu, C.A., Almashanu, S., Steel, G., Casey, R., Lambert, M., Mitchell, G.A. and Valle, D. 1999. Hyperornithinaemia-hyperammonaemia-homocitrullinuria syndrome is caused by mutations in a gene encoding a mitochondrial ornithine transporter. *Nature Genetics* 22: 151–158.

Canul-Tec, J.C., Assal, R., Cirri, E., Legrand, P., Brier, S., Chamot-Rooke, J. and Reyes, N. 2017. Structure and allosteric inhibition of excitatory amino acid transporter 1. *Nature* 544: 446–451.

Casey, J.R., Grinstein, S. and Orlowski, J. 2010. Sensors and regulators of intracellular pH. *Nature Reviews Molecular Cell Biology* 11: 50–61.

Chae, P.S., Rasmussen, S.G.F., Rana, R.R., Gotfryd, K., Chandra, R., Goren, M.A., Kruse, A.C., Nurva, S., Loland, C.J., Pierre, Y., Drew, D., Popot, J.-L., Picot, D., Fox, B.G., Guan, L., Gether, U., Byrne, B., Kobilka, B. and Gellman, S.H. 2010. Maltose–neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins. *Nature Methods* 7: 1003–1008.

Chapman, H.N., Barty, A., Bogan, M.J., Boutet, S., Frank, M., Hau-Riege, S.P., Marchesini, S., Woods, B.W., Bajt, S., Benner, W.H., London, R.A., Plönjes, E., Kuhlmann, M., Treusch, R., Düsterer, S., Tschentscher, T., Schneider, J.R., Spiller, E., Möller, T., Bostedt, C., Hoener, M., Shapiro, D.A., Hodgson, K.O., van der Spoel, D., Burmeister, F., Bergh, M., Caleman, C., Huldt, G., Seibert, M.M., Maia, F.R.N.C., Lee, R.W., Szöke, A., Timneanu, N. and Hajdu, J. 2006. Femtosecond diffractive imaging with a soft-X-ray free-electron laser. *Nature Physics* 2: 839–843.

Chapman, H.N., Fromme, P., Barty, A., White, T.A., Kirian, R.A., Aquila, A., Hunter, M.S., Schulz, J., DePonte, D.P., Weierstall, U., Doak, R.B., Maia, F.R.N.C., Martin, A.V., Schlichting, I., Lomb, L., Coppola, N., Shoeman, R.L., Epp, S.W., Hartmann, R., Rolles, D., Rudenko, A., Foucar, L., Kimmel, N., Weidenspointner, G., Holl, P., Liang, M., Barthelmess, M., Caleman, C., Boutet, S., Bogan, M.J., Krzywinski, J., Bostedt, C., Bajt, S., Gumprecht, L., Rudek, B., Erk, B., Schmidt, C., Hömke, A., Reich, C., Pietschner, D., Strüder, L., Hauser, G., Gorke, H., Ullrich, J., Herrmann, S., Schaller, G., Schopper, F., Soltau, H., Kühnel, K.-U., Messerschmidt, M., Bozek, J.D., Hau-Riege, S.P., Frank, M., Hampton, C.Y., Sierra, R.G., Starodub, D., Williams, G.J., Hajdu, J., Timneanu, N., Seibert, M.M., Andreasson, J., Rocker, A., Jönsson, O., Svenda, M., Stern, S., Nass, K., Andritschke, R., Schröter, C.-D., Krasniqi, F., Bott, M., Schmidt, K.E., Wang, X., Grotjohann, I., Holton, J.M., Barends, T.R.M., Neutze, R., Marchesini, S., Fromme, R., Schorb, S., Rupp, D., Adolph, M., Gorkhover, T., Andersson, I., Hirsemann, H., Potdevin, G., Graafsma, H., Nilsson, B. and Spence, J.C.H. 2011. Femtosecond X-ray protein nanocrystallography. *Nature* 470: 73–77.

Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S. and Richardson, D.C. 2010. *MolProbity* : all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D Biological Crystallography* 66: 12–21.

Cherezov, V., Hanson, M.A., Griffith, M.T., Hilgart, M.C., Sanishvili, R., Nagarajan, V., Stepanov, S., Fischetti, R.F., Kuhn, P. and Stevens, R.C. 2009. Rastering strategy for screening and centring of microcrystal samples of human membrane proteins with a sub-10 m size X-ray synchrotron beam. *Journal of The Royal Society Interface* 6: S587–S597.

Cherezov, V., Rosenbaum, D.M., Hanson, M.A., Rasmussen, S.G.F., Thian, F.S., Kobilka, T.S., Choi, H.-J., Kuhn, P., Weis, W.I., Kobilka, B.K. and Stevens, R.C. 2007. High-Resolution Crystal Structure of an Engineered Human 2-Adrenergic G Protein-Coupled Receptor. *Science* 318: 1258–1265.

Chun, E., Thompson, A.A., Liu, W., Roth, C.B., Griffith, M.T., Katritch, V., Kunken, J., Xu, F., Cherezov, V., Hanson, M.A. and Stevens, R.C. 2012. Fusion Partner Toolchest for the Stabilization and Crystallization of G Protein-Coupled Receptors. *Structure* 20: 967–976.

Clerc, S. and Barenholz, Y. 1995. Loading of amphipathic weak acids into liposomes in response to transmembrane calcium acetate gradients. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1240: 257–265.

Coleman, J.A., Green, E.M. and Gouaux, E. 2016. X-ray structures and mechanism of the human serotonin transporter. *Nature* 532: 334–339.

Colman, R.J., Anderson, R.M., Johnson, S.C., Kastman, E.K., Kosmatka, K.J., Beasley, T.M., Allison, D.B., Cruzen, C., Simmons, H.A., Kemnitz, J.W. and Weindruch, R. 2009. Caloric Restriction Delays Disease Onset and Mortality in Rhesus Monkeys. *Science* 325: 201–204.

Coquelle, N., Brewster, A.S., Kapp, U., Shilova, A., Weinhausen, B., Burghammer, M. and Colletier, J.P. 2015. Raster-scanning serial protein crystallography using micro- and nano-focused synchrotron beams. *Acta Crystallographica. Section D, Biological Crystallography* 71: 1184–1196.

Crisman, T.J., Qu, S., Kanner, B.I. and Forrest, L.R. 2009. Inward-facing conformation of glutamate transporters as revealed by their inverted-topology structural repeats. *Proceedings of the National Academy of Sciences* 106: 20752–20757.

Deisenhofer, J., Epp, O., Miki, K., Huber, R. and Michel, H. 1985. Structure of the protein subunits in the photosynthetic reaction centre of *Rhodopseudomonas viridis* at 3Å resolution. *Nature* 318: 618–624.

Deller, M.C., Kong, L. and Rupp, B. 2016. Protein stability: a crystallographer's perspective. *Acta Crystallographica Section F Structural Biology Communications* 72: 72–95.

DePonte, D.P., Weierstall, U., Schmidt, K., Warner, J., Starodub, D., Spence, J.C.H. and Doak, R.B. 2008. Gas dynamic virtual nozzle for generation of microscopic droplet streams. *Journal of Physics D: Applied Physics* 41: 195505.

Dickens, D., Chiduza, G.N., Wright, G.S.A., Pirmohamed, M., Antonyuk, S.V. and Hasnain, S.S. 2017. Modulation of LAT1 (SLC7A5) transporter activity and stability by membrane cholesterol. *Scientific Reports* 7: 43580.

Doig, A.J. 2017. Frozen, but no accident - why the 20 standard amino acids were selected. *The FEBS Journal* 284: 1296–1305.

Dougherty, D. 2000. Unnatural amino acids as probes of protein structure and function. *Current Opinion in Chemical Biology* 4: 645–652.

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

Drew, D.E., Heijne, G. von, Nordlund, P. and de Gier, J.W. 2001. Green fluorescent protein as an indicator to monitor membrane protein overexpression in *Escherichia coli*. *FEBS Letters* 507: 220–224.

Dufourc, E.J. 2008. Sterols and membrane dynamics. *Journal of Chemical Biology* 1: 63–77.

Dumon-Seignovert, L., Cariot, G. and Vuillard, L. 2004. The toxicity of recombinant proteins in *Escherichia coli*: a comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3). *Protein Expression and Purification* 37: 203–206.

de Duve, C. and Wattiaux, R. 1966. Functions of Lysosomes. *Annual Review of Physiology* 28: 435–492.

Efeyan, A., Zoncu, R. and Sabatini, D.M. 2012. Amino acids and mTORC1: from lysosomes to disease. *Trends in Molecular Medicine* 18: 524–533.

Efeyan, A., Comb, W.C. and Sabatini, D.M. 2015. Nutrient-sensing mechanisms and pathways. *Nature* 517: 302–310.

Ehrnstorfer, I.A., Geertsma, E.R., Pardon, E., Steyaert, J. and Dutzler, R. 2014. Crystal structure of a SLC11 (NRAMP) transporter reveals the basis for transition-metal ion transport. *Nature Structural & Molecular Biology* 21: 990–996.

Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. 2010. Features and development of *Coot*. *Acta Crystallographica Section D Biological Crystallography* 66: 486–501.

Engel, C.K., Chen, L. and Privé, G.G. 2002. Insertion of carrier proteins into hydrophilic loops of the Escherichia coli lactose permease. *Biochimica Et Biophysica Acta* 1564: 38–46.

Eun, J.S., Suh, Y.H., Kim, D.K. and Jeon, H. 2000. Regulation of cytokine production by exogenous nitric oxide in murine splenocyte and peritoneal macrophage. *Archives of Pharmacal Research* 23: 531–534.

Evans, P.R. and Murshudov, G.N. 2013. How good are my data and what is the resolution? *Acta Crystallographica Section D Biological Crystallography* 69: 1204–1214.

Faham, S., Watanabe, A., Besserer, G.M., Cascio, D., Specht, A., Hirayama, B.A., Wright, E.M. and Abramson, J. 2008. The Crystal Structure of a Sodium Galactose Transporter Reveals Mechanistic Insights into Na^+/Sugar Symport. *Science* 321: 810–814.

Fairman, W.A., Vandenberg, R.J., Arriza, J.L., Kavanaught, M.P. and Amara, S.G. 1995. An excitatory amino-acid transporter with properties of a ligand-gated chloride channel. *Nature* 375: 599–603.

Fang, Y., Kolmakova-Partensky, L. and Miller, C. 2007. A bacterial arginine-agmatine exchange transporter involved in extreme acid resistance. *The Journal of Biological Chemistry* 282: 176–182.

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

Farwick, M., Siewe, R.M. and Krämer, R. 1995. Glycine betaine uptake after hyperosmotic shift in Corynebacterium glutamicum. *Journal of Bacteriology* 177: 4690–4695.

Fenalti, G., Zatsepin, N.A., Betti, C., Giguere, P., Han, G.W., Ishchenko, A., Liu, W., Guillemyn, K., Zhang, H., James, D., Wang, D., Weierstall, U., Spence, J.C.H., Boutet, S., Messerschmidt, M., Williams, G.J., Gati, C., Yefanov, O.M., White, T.A., Oberthuer, D., Metz, M., Yoon, C.H., Barty, A., Chapman, H.N., Basu, S., Coe, J., Conrad, C.E., Fromme, R., Fromme, P., Tourwé, D., Schiller, P.W., Roth, B.L., Ballet, S., Katritch, V., Stevens, R.C. and Cherezov, V. 2015. Structural basis for bifunctional

peptide recognition at human δ-opioid receptor. *Nature Structural & Molecular Biology* 22: 265–268.

Font, M. 2001. Functional analysis of mutations in SLC7A9, and genotype-phenotype correlation in non-Type I cystinuria. *Human Molecular Genetics* 10: 305–316.

Forrest, L.R., Krämer, R. and Ziegler, C. 2011. The structural basis of secondary active transport mechanisms. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 1807: 167–188.

Forrest, L.R., Tavoulari, S., Zhang, Y.-W., Rudnick, G. and Honig, B. 2007. Identification of a chloride ion binding site in Na^+/Cl^-dependent transporters. *Proceedings of the National Academy of Sciences* 104: 12761–12766.

Forrest, L.R., Zhang, Y.-W., Jacobs, M.T., Gesmonde, J., Xie, L., Honig, B.H. and Rudnick, G. 2008. Mechanism for alternating access in neurotransmitter transporters. *Proceedings of the National Academy of Sciences of the United States of America* 105: 10338–10343.

Forstermann, U. and Sessa, W.C. 2012. Nitric oxide synthases: regulation and function. *European Heart Journal* 33: 829–837.

Fotiadis, D., Kanai, Y. and Palacín, M. 2013. The SLC3 and SLC7 families of amino acid transporters. *Molecular Aspects of Medicine* 34: 139–158.

French, S. and Wilson, K. 1978. On the treatment of negative intensity observations. *Acta Crystallographica Section A* 34: 517–525.

Frommer, W.B., Hummel, S., Unseld, M. and Ninnemann, O. 1995. Seed and vascular expression of a high-affinity transporter for cationic amino acids in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America* 92: 12036–12040.

Gabriško, M. and Janeček, Š. 2009. Looking for the ancestry of the heavy-chain subunits of heteromeric amino acid transporters rBAT and 4F2hc within the GH13 α-amylase family: Origin of rBAT and 4F2hc within the GH13 α-amylase family. *FEBS Journal* 276: 7265–7278.

Gao, X., Lu, F., Zhou, L., Dang, S., Sun, L., Li, X., Wang, J. and Shi, Y. 2009. Structure and mechanism of an amino acid antiporter. *Science* 324: 1565–1568.

Gao, X., Zhou, L., Jiao, X., Lu, F., Yan, C., Zeng, X., Wang, J. and Shi, Y. 2010. Mechanism of substrate recognition and transport by an amino acid antiporter. *Nature* 463: 828–832.

Garman, E.F. 2010. Radiation damage in macromolecular crystallography: what is it and why should we care? *Acta Crystallographica Section D Biological Crystallography* 66: 339–351.

Gasol, E., Jiménez-Vidal, M., Chillarón, J., Zorzano, A. and Palacín, M. 2004. Membrane Topology of System Xc^- Light Subunit Reveals a Re-entrant Loop with Substrate-restricted Accessibility. *Journal of Biological Chemistry* 279: 31228–31236.

Gati, C., Bourenkov, G., Klinge, M., Rehders, D., Stellato, F., Oberthür, D., Yefanov, O., Sommer, B.P., Mogk, S., Duszenko, M., Betzel, C., Schneider, T.R., Chapman, H.N. and Redecke, L. 2014. Serial crystallography on *in vivo* grown microcrystals using synchrotron radiation. *IUCrJ* 1: 87–94.

Geertsma, E.R., Chang, Y.-N., Shaik, F.R., Neldner, Y., Pardon, E., Steyaert, J. and Dutzler, R. 2015. Structure of a prokaryotic fumarate transporter reveals the architecture of the SLC26 family. *Nature Structural & Molecular Biology* 22: 803–808.

Groff, D., Wang, F., Jockusch, S., Turro, N.J. and Schultz, P.G. 2010. A New Strategy to Photoactivate Green Fluorescent Protein. *Angewandte Chemie International Edition* 49: 7677–7679.

Habermeier, A., Wolf, S., Martine, U., Graf, P. and Closs, E.I. 2003. Two Amino Acid Residues Determine the Low Substrate Affinity of Human Cationic Amino Acid Transporter-2A. *Journal of Biological Chemistry* 278: 19492–19499.

Hammes, U.Z., Nielsen, E., Honaas, L.A., Taylor, C.G. and Schachtman, D.P. 2006. AtCAT6, a sink-tissue-localized transporter for essential amino acids in Arabidopsis. *The Plant Journal* 48: 414–426.

Hanson, M.A., Cherezov, V., Griffith, M.T., Roth, C.B., Jaakola, V.-P., Chien, E.Y.T., Velasquez, J., Kuhn, P. and Stevens, R.C. 2008. A Specific Cholesterol Binding Site Is Established by the 2.8 Å Structure of the Human β2-Adrenergic Receptor. *Structure* 16: 897–905.

Hara, K., Yonezawa, K., Weng, Q.-P., Kozlowski, M.T., Belham, C. and Avruch, J. 1998. Amino Acid Sufficiency and mTOR Regulate p70 S6 Kinase and eIF-4E BP1 through a Common Effector Mechanism. *Journal of Biological Chemistry* 273: 14484–14494.

Harrison, D.E., Strong, R., Sharp, Z.D., Nelson, J.F., Astle, C.M., Flurkey, K., Nadon, N.L., Wilkinson, J.E., Frenkel, K., Carter, C.S., Pahor, M., Javors, M.A., Fernandez, E. and Miller, R.A. 2009. Rapamycin fed late in life extends lifespan in genetically heterogeneous mice. *Nature* 460: 392–395.

Haupert, L.M. and Simpson, G.J. 2011. Screening of protein crystallization trials by second order nonlinear optical imaging of chiral crystals (SONICC). *Methods* 55: 379–386.

Hediger, M.A., Romero, M.F., Peng, J.-B., Rolfs, A., Takanaga, H. and Bruford, E.A. 2004. The ABCs of solute carriers: physiological, pathological and therapeutic implications of human membrane transport proteins. *Pflügers Archiv European Journal of Physiology* 447: 465–468.

Hemm, M.R., Paul, B.J., Miranda-Rios, J., Zhang, A., Soltanzad, N. and Storz, G. 2010. Small Stress Response Proteins in Escherichia coli: Proteins Missed by Classical Proteomic Studies. *Journal of Bacteriology* 192: 46–58.

Henley, J.M. 2001. Amino Acid Neurotransmitters. In: John Wiley & Sons, Ltd (Ed.) , *Encyclopedia of Life Sciences,* John Wiley & Sons, Ltd, .

Hersh, B.M., Farooq, F.T., Barstad, D.N., Blankenhorn, D.L. and Slonczewski, J.L. 1996. A glutamate-dependent acid resistance gene in *Escherichia coli*. *Journal of Bacteriology* 178: 3978–3981.

Heublein, S., Kazi, S., Ögmundsdóttir, M.H., Attwood, E.V., Kala, S., Boyd, C.A.R., Wilson, C. and Goberdhan, D.C.I. 2010. Proton-assisted amino-acid transporters are conserved regulators of proliferation and amino-acid-dependent mTORC1 activation. *Oncogene* 29: 4068–4079.

Heymann, M., Opthalage, A., Wierman, J.L., Akella, S., Szebenyi, D.M.E., Gruner, S.M. and Fraden, S. 2014. Room-temperature serial crystallography using a kinetically optimized microfluidic device for protein crystallization and on-chip X-ray diffraction. *IUCrJ* 1: 349–360.

Hildebrand, D. 2016. Lecture Twenty-four Nitrogen Metabolism -- Essential and non-protein amino acid biosynthesis; other N compounds. *BCH/PPA/PLS 609, Plant Biochemistry*. Available from: http://www.uky.edu/~dhild/biochem/24/lect24.html [Accessed: December 2017]

Hille, B. 2001. *Ion Channels of Excitable Membranes*, 3rd ed. Sinauer, Sunderland, Mass.

Hirata, K., Shinzawa-Itoh, K., Yano, N., Takemura, S., Kato, K., Hatanaka, M., Muramoto, K., Kawahara, T., Tsukihara, T., Yamashita, E., Tono, K., Ueno, G., Hikima, T., Murakami, H., Inubushi, Y., Yabashi, M., Ishikawa, T., Yamamoto, M., Ogura, T., Sugimoto, H., Shen, J.-R., Yoshikawa, S. and Ago, H. 2014. Determination of damage-free crystal structure of an X-ray-sensitive protein using an XFEL. *Nature Methods* 11: 734–736.

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

Hunter, M.S., Segelke, B., Messerschmidt, M., Williams, G.J., Zatsepin, N.A., Barty, A., Benner, W.H., Carlson, D.B., Coleman, M., Graf, A., Hau-Riege, S.P., Pardini, T., Seibert, M.M., Evans, J., Boutet, S. and Frank, M. 2015. Fixed-target protein serial microcrystallography with an X-ray free electron laser. *Scientific Reports* 4.

Ilgü, H., Jeckelmann, J.-M., Gapsys, V., Ucurum, Z., de Groot, B.L. and Fotiadis, D. 2016. Insights into the molecular basis for substrate binding and specificity of the wild-type L-arginine/agmatine antiporter AdiC. *Proceedings of the National Academy of Sciences* 113: 10358–10363.

Ito, K. and Groudine, M. 1997. A New Member of the Cationic Amino Acid Transporter Family Is Preferentially Expressed in Adult Mouse Brain. *Journal of Biological Chemistry* 272: 26780–26786.

Iversen, L. 2009. Neurotransmitter transporters and their impact on the development of psychopharmacology: Neurotransmitter transporters. *British Journal of Pharmacology* 147: 82–88.

Iwata, S. 2003. *Methods and Results in Crystallization of Membrane Proteins*. International University Line, La Jolla, Calif, .

Iyer, R., Williams, C. and Miller, C. 2003. Arginine-Agmatine Antiporter in Extreme Acid Resistance in Escherichia coli. *Journal of Bacteriology* 185: 6556–6561.

Jack, D.L., Paulsen, I.T. and Saier, M.H. 2000. The amino acid/polyamine/organocation (APC) superfamily of transporters specific for amino acids, polyamines and organocations. *Microbiology (Reading, England)* 146 (Pt 8): 1797–1814.

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

Jessell, T.M. and Kandel, E.R. 1993. Synaptic transmission: a bidirectional and self-modifiable form of cell-cell communication. *Cell* 72: 1–30.

Jung, H., Buchholz, M., Clausen, J., Nietschke, M., Revermann, A., Schmid, R. and Jung, K. 2002. CaiT of *Escherichia coli* , a New Transporter Catalyzing l-Carnitine/γ-Butyrobetaine Exchange. *Journal of Biological Chemistry* 277: 39251–39258.

Kaback, H.R., Sahin-Tóth, M. and Weinglass, A.B. 2001. The kamikaze approach to membrane transport. *Nature Reviews Molecular Cell Biology* 2: 610–620.

Kaback, H.R., Smirnova, I., Kasho, V., Nie, Y. and Zhou, Y. 2011. The alternating access transport mechanism in LacY. *The Journal of Membrane Biology* 239: 85–93.

Kabsch, W. 2010. *XDS*. *Acta Crystallographica Section D Biological Crystallography* 66: 125–132.

Kageyama, T., Nakamura, M., Matsuo, A., Yamasaki, Y., Takakura, Y., Hashida, M., Kanai, Y., Naito, M., Tsuruo, T., Minato, N. and Shimohama, S. 2000. The 4F2hc/LAT1 complex transports L-DOPA across the blood–brain barrier. *Brain Research* 879: 115–121.

Kakuda, D.K., Sweet, M.J., Macleod, C.L., Hume, D.A. and Markovich, D. 1999. CAT2-mediated l-arginine transport and nitric oxide production in activated macrophages. *Biochemical Journal* 340: 549–553.

Kalayil, S., Schulze, S. and Kühlbrandt, W. 2013. Arginine oscillation explains Na+ independence in the substrate/product antiporter CaiT. *Proceedings of the National Academy of Sciences of the United States of America* 110: 17296–17301.

Kanai, Y., Segawa, H., Miyamoto, K. -i., Uchino, H., Takeda, E. and Endou, H. 1998. Expression Cloning and Characterization of a Transporter for Large Neutral Amino Acids Activated by the Heavy Chain of 4F2 Antigen (CD98). *Journal of Biological Chemistry* 273: 23629–23632.

Kang, Y., Zhou, X.E., Gao, X., He, Y., Liu, W., Ishchenko, A., Barty, A., White, T.A., Yefanov, O., Han, G.W., Xu, Q., de Waal, P.W., Ke, J., Tan, M.H.E., Zhang, C., Moeller, A., West, G.M., Pascal, B.D., Van Eps, N., Caro, L.N., Vishnivetskiy, S.A., Lee, R.J., Suino-Powell, K.M., Gu, X., Pal, K., Ma, J., Zhi, X., Boutet, S., Williams, G.J., Messerschmidt, M., Gati, C., Zatsepin, N.A., Wang, D., James, D., Basu, S., Roy-Chowdhury, S., Conrad, C.E., Coe, J., Liu, H., Lisova, S., Kupitz, C., Grotjohann, I., Fromme, R., Jiang, Y., Tan, M., Yang, H., Li, J., Wang, M., Zheng, Z., Li, D., Howe, N., Zhao, Y., Standfuss, J., Diederichs, K., Dong, Y., Potter, C.S., Carragher, B., Caffrey, M., Jiang, H., Chapman, H.N., Spence, J.C.H., Fromme, P., Weierstall, U., Ernst, O.P., Katritch, V., Gurevich, V.V., Griffin, P.R., Hubbell, W.L., Stevens, R.C., Cherezov, V., Melcher, K. and Xu, H.E. 2015. Crystal structure of rhodopsin bound to arrestin by femtosecond X-ray laser. *Nature* 523: 561–567.

Kantcheva, A.K., Quick, M., Shi, L., Winther, A.-M.L., Stolzenberg, S., Weinstein, H., Javitch, J.A. and Nissen, P. 2013. Chloride binding site of neurotransmitter sodium symporters. *Proceedings of the National Academy of Sciences* 110: 8489–8494.

Kaur, J., Olkhova, E., Malviya, V.N., Grell, E. and Michel, H. 2014. A L-lysine transporter of high stereoselectivity of the amino acid-polyamine-organocation (APC) superfamily: production, functional characterization, and structure modeling. *The Journal of Biological Chemistry* 289: 1377–1387.

Keller, R., Ziegler, C. and Schneider, D. 2014. When Two Turn Into One: Evolution of Membrane Transporters From Half Modules. *Biological Chemistry* 395: 1379–88.

Khafizov, K., Staritzbichler, R., Stamm, M. and Forrest, L.R. 2010. A study of the evolution of inverted-topology repeats from LeuT-fold transporters using AlignMe. *Biochemistry* 49: 10702–10713.

Kim, J.W., Closs, E.I., Albritton, L.M. and Cunningham, J.M. 1991. Transport of cationic amino acids by the mouse ecotropic retrovirus receptor. *Nature* 352: 725–728.

King, D.E., Mainous, A.G. and Geesey, M.E. 2008. Variation in L-arginine intake follow demographics and lifestyle factors that may impact cardiovascular disease risk. *Nutrition Research* 28: 21–24.

Kissick, D.J., Wanapun, D. and Simpson, G.J. 2011. Second-Order Nonlinear Optical Imaging of Chiral Crystals. *Annual Review of Analytical Chemistry* 4: 419–437.

Kissick, D.J., Gualtieri, E.J., Simpson, G.J. and Cherezov, V. 2010. Nonlinear Optical Imaging of Integral Membrane Protein Crystals in Lipidic Mesophases. *Analytical Chemistry* 82: 491–497.

Koshy, C., Schweikhard, E.S., Gärtner, R.M., Perez, C., Yildiz, O. and Ziegler, C. 2013. Structural evidence for functional lipid interactions in the betaine transporter BetP. *The EMBO Journal* 32: 3096–3105.

Kowalczyk, L., Ratera, M., Paladino, A., Bartoccioni, P., Errasti-Murugarren, E., Valencia, E., Portella, G., Bial, S., Zorzano, A., Fita, I., Orozco, M., Carpena, X., Vázquez-Ibar, J.L. and Palacín, M. 2011. Molecular basis of substrate-induced permeation by an amino acid antiporter. *Proceedings of the National Academy of Sciences of the United States of America* 108: 3935–3940.

Krishnamurthy, H. and Gouaux, E. 2012. X-ray structures of LeuT in substrate-free outward-open and apo inward-open states. *Nature* 481: 469–474.

Kroncke, B.M., Horanyi, P.S. and Columbus, L. 2010. Structural Origins of Nitroxide Side Chain Dynamics on Membrane Protein α-Helical Sites. *Biochemistry* 49: 10045–10060.

Krulwich, T.A., Sachs, G. and Padan, E. 2011. Molecular aspects of bacterial pH sensing and homeostasis. *Nature Reviews. Microbiology* 9: 330–343.

Kupitz, C., Basu, S., Grotjohann, I., Fromme, R., Zatsepin, N.A., Rendek, K.N., Hunter, M.S., Shoeman, R.L., White, T.A., Wang, D., James, D., Yang, J.-H., Cobb, D.E., Reeder, B., Sierra, R.G., Liu, H., Barty, A., Aquila, A.L., Deponte, D., Kirian, R.A., Bari, S., Bergkamp, J.J., Beyerlein, K.R., Bogan, M.J., Caleman, C., Chao, T.-C., Conrad, C.E., Davis, K.M., Fleckenstein, H., Galli, L., Hau-Riege, S.P., Kassemeyer, S., Laksmono, H., Liang, M., Lomb, L., Marchesini, S., Martin, A.V., Messerschmidt, M., Milathianaki, D., Nass, K., Ros, A., Roy-Chowdhury, S., Schmidt, K., Seibert, M., Steinbrener, J., Stellato, F., Yan, L., Yoon, C., Moore, T.A., Moore, A.L., Pushkar, Y., Williams, G.J., Boutet, S., Doak, R.B., Weierstall, U., Frank, M., Chapman, H.N., Spence, J.C.H. and Fromme, P. 2014. Serial time-resolved crystallography of photosystem II using a femtosecond X-ray laser. *Nature* 513: 261–265.

Landau, E.M. and Rosenbusch, J.P. 1996. Lipidic cubic phases: a novel concept for the crystallization of membrane proteins. *Proceedings of the National Academy of Sciences* 93: 14532–14535.

Larsen, M., Larsen, B.B., Frølund, B. and Nielsen, C.U. 2008. Transport of amino acids and GABA analogues via the human proton-coupled amino acid transporter, hPAT1: Characterization of conditions for affinity and transport experiments in Caco-2 cells. *European Journal of Pharmaceutical Sciences* 35: 86–95.

Lecleire, S., Coeffier, M., Leblond, J., Hubert, A., Lemoulan, S., Petit, A., Ducrotte, P., Dechelotte, P. and Marion, R. 2005. Modulation of nitric oxide and cytokines production by l-arginine in human gut mucosa. *Clinical Nutrition* 24: 353–359.

Lee, C., Kang, H.J., von Ballmoos, C., Newstead, S., Uzdavinys, P., Dotson, D.L., Iwata, S., Beckstein, O., Cameron, A.D. and Drew, D. 2013. A two-domain elevator mechanism for sodium/proton antiport. *Nature* 501: 573–577.

Lee, D.J., Bingle, L.E., Heurlier, K., Pallen, M.J., Penn, C.W., Busby, S.J. and Hobman, J.L. 2009. Gene doctoring: a method for recombineering in laboratory and pathogenic Escherichia coli strains. *BMC Microbiology* 9: 252–265.

Lenaeus, M.J., Gamal El-Din, T.M., Ing, C., Ramanadane, K., Pomès, R., Zheng, N. and Catterall, W.A. 2017. Structures of closed and open states of a voltage-gated sodium channel. *Proceedings of the National Academy of Sciences* 114: 3051–3060.

Li, D., Stansfeld, P.J., Sansom, M.S.P., Keogh, A., Vogeley, L., Howe, N., Lyons, J.A., Aragao, D., Fromme, P., Fromme, R., Basu, S., Grotjohann, I., Kupitz, C., Rendek, K., Weierstall, U., Zatsepin, N.A., Cherezov, V., Liu, W., Bandaru, S., English, N.J., Gati, C., Barty, A., Yefanov, O., Chapman, H.N., Diederichs, K., Messerschmidt, M., Boutet, S., Williams, G.J., Marvin Seibert, M. and Caffrey, M. 2015. Ternary structure reveals mechanism of a membrane diacylglycerol kinase. *Nature Communications* 6: article number 10140.

Liao, J., Li, H., Zeng, W., Sauer, D.B., Belmares, R. and Jiang, Y. 2012. Structural Insight into the Ion-Exchange Mechanism of the Sodium/Calcium Exchanger. *Science* 335: 686–690.

Liebschner, D., Yamada, Y., Matsugaki, N., Senda, M. and Senda, T. 2016. On the influence of crystal size and wavelength on native SAD phasing. *Acta Crystallographica Section D Structural Biology* 72: 728–741.

Liu, B., Du, H., Rutkowski, R., Gartner, A. and Wang, X. 2012. LAAT-1 Is the Lysosomal Lysine/Arginine Transporter That Maintains Amino Acid Homeostasis. *Science* 337: 351–354.

Liu, W., Ishchenko, A. and Cherezov, V. 2014. Preparation of microcrystals in lipidic cubic phase for serial femtosecond crystallography. *Nature Protocols* 9: 2123–2134.

Liu, W., Wacker, D., Gati, C., Han, G.W., James, D., Wang, D., Nelson, G., Weierstall, U., Katritch, V., Barty, A., Zatsepin, N.A., Li, D., Messerschmidt, M., Boutet, S., Williams, G.J., Koglin, J.E., Seibert, M.M., Wang, C., Shah, S.T.A., Basu, S., Fromme, R., Kupitz, C., Rendek, K.N., Grotjohann, I., Fromme, P., Kirian, R.A., Beyerlein, K.R., White, T.A., Chapman, H.N., Caffrey, M., Spence, J.C.H., Stevens, R.C. and Cherezov, V. 2013. Serial femtosecond crystallography of G protein-coupled receptors. *Science* 342: 1521–1524.

Loewith, R., Jacinto, E., Wullschleger, S., Lorberg, A., Crespo, J.L., Bonenfant, D., Oppliger, W., Jenoe, P. and Hall, M.N. 2002. Two TOR Complexes, Only One of which Is Rapamycin Sensitive, Have Distinct Roles in Cell Growth Control. *Molecular Cell* 10: 457–468.

Lomb, L., Steinbrener, J., Bari, S., Beisel, D., Berndt, D., Kieser, C., Lukat, M., Neef, N. and Shoeman, R.L. 2012. An anti-settling sample delivery instrument for serial femtosecond crystallography. *Journal of Applied Crystallography* 45: 674–678.

Lu, F., Li, S., Jiang, Y., Jiang, J., Fan, H., Lu, G., Deng, D., Dang, S., Zhang, X., Wang, J. and Yan, N. 2011. Structure and mechanism of the uracil transporter UraA. *Nature* 472: 243–246.

Lyons, J.A., Parker, J.L., Solcan, N., Brinth, A., Li, D., Shah, S.T., Caffrey, M. and Newstead, S. 2014. Structural basis for polyspecificity in the POT family of proton-coupled oligopeptide transporters. *EMBO Reports* 15: 886–893.

Ma, D., Lu, P., Yan, C., Fan, C., Yin, P., Wang, J. and Shi, Y. 2012a. Structure and mechanism of a glutamate–GABA antiporter. *Nature* 483: 632–636.

Ma, D., Lu, P., Yan, C., Fan, C., Yin, P., Wang, J. and Shi, Y. 2012b. Structure and mechanism of a glutamate–GABA antiporter. *Nature* 483: 632–636.

Macháček, T., Panská, L., Dvořáková, H. and Horák, P. 2016. Nitric oxide and cytokine production by glial cells exposed in vitro to neuropathogenic schistosome Trichobilharzia regenti. *Parasites & Vectors* 9: 579–588.

Malinauskaite, L., Quick, M., Reinhard, L., Lyons, J.A., Yano, H., Javitch, J.A. and Nissen, P. 2014. A mechanism for intracellular release of Na(+) by neurotransmitter/sodium symporters. *Nature Structural & Molecular Biology* 21: 1006–1012.

Malott, R.J., Steen-Kinnaird, B.R., Lee, T.D. and Speert, D.P. 2012. Identification of Hopanoid Biosynthesis Genes Involved in Polymyxin Resistance in Burkholderia multivorans. *Antimicrobial Agents and Chemotherapy* 56: 464–471.

Malott, R.J., Wu, C.-H., Lee, T.D., Hird, T.J., Dalleska, N.F., Zlosnik, J.E.A., Newman, D.K. and Speert, D.P. 2014. Fosmidomycin Decreases Membrane Hopanoids and Potentiates the Effects of Colistin on Burkholderia multivorans Clinical Isolates. *Antimicrobial Agents and Chemotherapy* 58: 5211–5219.

Mantovani, A., Sica, A., Sozzani, S., Allavena, P., Vecchi, A. and Locati, M. 2004. The chemokine system in diverse forms of macrophage activation and polarization. *Trends in Immunology* 25: 677–686.

Mariani, V., Morgan, A., Yoon, C.H., Lane, T.J., White, T.A., O'Grady, C., Kuhn, M., Aplin, S., Koglin, J., Barty, A. and Chapman, H.N. 2016. *OnDA* : online data analysis and feedback for serial X-ray imaging. *Journal of Applied Crystallography* 49: 1073–1080.

Masson, J., Sagné, C., Hamon, M. and El Mestikawy, S. 1999. Neurotransmitter transporters in the central nervous system. *Pharmacological Reviews* 51: 439–464.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. and Read, R.J. 2007. Phaser crystallographic software. *Journal of Applied Crystallography* 40: 658–674.

Mills, C.D., Kincaid, K., Alt, J.M., Heilman, M.J. and Hill, A.M. 2000. M-1/M-2 Macrophages and the Th1/Th2 Paradigm. *The Journal of Immunology* 164: 6166–6173.

Miroux, B. and Walker, J.E. 1996. Over-production of Proteins inEscherichia coli: Mutant Hosts that Allow Synthesis of some Membrane Proteins and Globular Proteins at High Levels. *Journal of Molecular Biology* 260: 289–298.

Mitchell, P. 1985. Molecular mechanics of protonmotive F_0F_1 ATPases: Rolling well and turnstile hypothesis. *FEBS Letters* 182: 1–7.

Morris, S.M. 2002. Regulation of enzymes of the urea cycle and arginine metabolism. *Annual Review of Nutrition* 22: 87–105.

Morris, S.M. 2007. Arginine metabolism: boundaries of our knowledge. *The Journal of Nutrition* 137: 1602–1609.

Murray, J.W., Garman, E.F. and Ravelli, R.B.G. 2004. X-ray absorption by macromolecular crystals: the effects of wavelength and crystal composition on absorbed dose. *Journal of Applied Crystallography* 37: 513–522.

Mykkanen, J. 2000. Functional analysis of novel mutations in y+LAT-1 amino acid transporter gene causing lysinuric protein intolerance (LPI). *Human Molecular Genetics* 9: 431–438.

Nagai, K., Oubridge, C., Jessen, T.H., Li, J. and Evans, P.R. 1990. Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature* 348: 515–520.

Nakane, T., Hanashima, S., Suzuki, M., Saiki, H., Hayashi, T., Kakinouchi, K., Sugiyama, S., Kawatake, S., Matsuoka, S., Matsumori, N., Nango, E., Kobayashi, J., Shimamura,

T., Kimura, K., Mori, C., Kunishima, N., Sugahara, M., Takakyu, Y., Inoue, S., Masuda, T., Hosaka, T., Tono, K., Joti, Y., Kameshima, T., Hatsui, T., Yabashi, M., Inoue, T., Nureki, O., Iwata, S., Murata, M. and Mizohata, E. 2016. Membrane protein structure determination by SAD, SIR, or SIRAS phasing in serial femtosecond crystallography using an iododetergent. *Proceedings of the National Academy of Sciences* 113: 13039–13044.

Newsholme, P., Lima, M.M.R., Procopio, J., Pithon-Curi, T.C., Doi, S.Q., Bazotte, R.B. and Curi, R. 2003. Glutamine and glutamate as vital metabolites. *Brazilian Journal of Medical and Biological Research* 36: 153–163.

Nicklin, P., Bergman, P., Zhang, B., Triantafellow, E., Wang, H., Nyfeler, B., Yang, H., Hild, M., Kung, C., Wilson, C., Myer, V.E., MacKeigan, J.P., Porter, J.A., Wang, Y.K., Cantley, L.C., Finan, P.M. and Murphy, L.O. 2009. Bidirectional Transport of Amino Acids Regulates mTOR and Autophagy. *Cell* 136: 521–534.

Nogly, P., James, D., Wang, D., White, T.A., Zatsepin, N., Shilova, A., Nelson, G., Liu, H., Johansson, L., Heymann, M., Jaeger, K., Metz, M., Wickstrand, C., Wu, W., Båth, P., Berntsen, P., Oberthuer, D., Panneels, V., Cherezov, V., Chapman, H., Schertler, G., Neutze, R., Spence, J., Moraes, I., Burghammer, M., Standfuss, J. and Weierstall, U. 2015. Lipidic cubic phase serial millisecond crystallography using synchrotron radiation. *IUCrJ* 2: 168–176.

Nogly, P., Panneels, V., Nelson, G., Gati, C., Kimura, T., Milne, C., Milathianaki, D., Kubo, M., Wu, W., Conrad, C., Coe, J., Bean, R., Zhao, Y., Båth, P., Dods, R., Harimoorthy, R., Beyerlein, K.R., Rheinberger, J., James, D., DePonte, D., Li, C., Sala, L., Williams, G.J., Hunter, M.S., Koglin, J.E., Berntsen, P., Nango, E., Iwata, S., Chapman, H.N., Fromme, P., Frank, M., Abela, R., Boutet, S., Barty, A., White, T.A., Weierstall, U., Spence, J., Neutze, R., Schertler, G. and Standfuss, J. 2016. Lipidic cubic phase injector is a viable crystal delivery system for time-resolved serial crystallography. *Nature Communications* 7: article number 12314.

Nollert, P. 2003. Microscope detection options for colorless protein crystals grown in lipidic cubic phases. *Journal of Applied Crystallography* 36: 1295–1296.

Noren, C.J., Anthony-Cahill, S.J., Griffith, M.C. and Schultz, P.G. 1989. A general method for site-specific incorporation of unnatural amino acids into proteins. *Science (New York, N.Y.)* 244: 182–188.

Oberthuer, D., Knoška, J., Wiedorn, M.O., Beyerlein, K.R., Bushnell, D.A., Kovaleva, E.G., Heymann, M., Gumprecht, L., Kirian, R.A., Barty, A., Mariani, V., Tolstikova, A., Adriano, L., Awel, S., Barthelmess, M., Dörner, K., Xavier, P.L., Yefanov, O., James, D.R., Nelson, G., Wang, D., Calvey, G., Chen, Y., Schmidt, A., Szczepek, M., Frielingsdorf, S., Lenz, O., Snell, E., Robinson, P.J., Šarler, B., Belšak, G., Maček, M., Wilde, F., Aquila, A., Boutet, S., Liang, M., Hunter, M.S., Scheerer, P., Lipscomb, J.D., Weierstall, U., Kornberg, R.D., Spence, J.C.H., Pollack, L., Chapman, H.N. and Bajt, S. 2017. Double-flow focused liquid injector for efficient serial femtosecond crystallography. *Scientific Reports* 7: article number 44628.

Oghbaey, S., Sarracini, A., Ginn, H.M., Pare-Labrosse, O., Kuo, A., Marx, A., Epp, S.W., Sherrell, D.A., Eger, B.T., Zhong, Y., Loch, R., Mariani, V., Alonso-Mori, R., Nelson, S., Lemke, H.T., Owen, R.L., Pearson, A.R., Stuart, D.I., Ernst, O.P., Mueller-Werkmeister, H.M. and Miller, R.J.D. 2016. Fixed target combined with spectral mapping: approaching 100% hit rates for serial crystallography. *Acta Crystallographica Section D Structural Biology* 72: 944–955.

Ögmundsdóttir, M.H., Heublein, S., Kazi, S., Reynolds, B., Visvalingam, S.M., Shaw, M.K., Goberdhan, D.C.I. and Hansen, I.A. 2012. Proton-Assisted Amino Acid Transporter PAT1 Complexes with Rag GTPases and Activates TORC1 on Late Endosomal and Lysosomal Membranes. *PLoS ONE* 7: e36616.

Overington, J.P., Al-Lazikani, B. and Hopkins, A.L. 2006. How many drug targets are there? *Nature Reviews Drug Discovery* 5: 993–996.

Padilla, J.E. and Yeates, T.O. 2003. A statistic for local intensity differences: robustness to anisotropy and pseudo-centering and utility for detecting twinning. *Acta Crystallographica Section D Biological Crystallography* 59: 1124–1130.

Pande, K., Hutchison, C.D.M., Groenhof, G., Aquila, A., Robinson, J.S., Tenboer, J., Basu, S., Boutet, S., DePonte, D.P., Liang, M., White, T.A., Zatsepin, N.A., Yefanov, O., Morozov, D., Oberthuer, D., Gati, C., Subramanian, G., James, D., Zhao, Y., Koralek, J., Brayshaw, J., Kupitz, C., Conrad, C., Roy-Chowdhury, S., Coe, J.D., Metz, M., Xavier, P.L., Grant, T.D., Koglin, J.E., Ketawala, G., Fromme, R., rajer, V., Henning, R., Spence, J.C.H., Ourmazd, A., Schwander, P., Weierstall, U., Frank, M., Fromme, P., Barty, A., Chapman, H.N., Moffat, K., van Thor, J.J. and Schmidt, M. 2016. Femtosecond structural dynamics drives the trans/cis isomerization in photoactive yellow protein. *Science* 352: 725–729.

Parker, J.L. and Newstead, S. 2012. Current trends in α-helical membrane protein crystallization: an update. *Protein Science : A Publication of the Protein Society* 21: 1358–1365.

Parker, J.L. and Newstead, S. 2016. Membrane Protein Crystallisation: Current Trends and Future Perspectives. In: Moraes, I. (Ed.) , *The Next Generation in Membrane Protein Structure Determination,* Springer International Publishing, Cham, vol. 922. pp. 61–72.

Pawate, A.S., Šrajer, V., Schieferstein, J., Guha, S., Henning, R., Kosheleva, I., Schmidt, M., Ren, Z., Kenis, P.J.A. and Perry, S.L. 2015. Towards time-resolved serial crystallography in a microfluidic device. *Acta Crystallographica. Section F, Structural Biology Communications* 71: 823–830.

Pebay-Peyroula, E., Dahout-Gonzalez, C., Kahn, R., Trézéguet, V., Lauquin, G.J.-M. and Brandolin, G. 2003. Structure of mitochondrial ADP/ATP carrier in complex with carboxyatractyloside. *Nature* 426: 39–44.

Pei, J. and Grishin, N.V. 2014. PROMALS3D: Multiple Protein Sequence Alignment Enhanced with Evolutionary and Three-Dimensional Structural Information. In: Russell, D.J. (Ed.) , *Multiple Sequence Alignment Methods,* Humana Press, Totowa, NJ, vol. 1079. pp. 263–271.

Penmatsa, A. and Gouaux, E. 2014. How LeuT shapes our understanding of the mechanisms of sodium-coupled neurotransmitter transporters: Role of LeuT in understanding neurotransmitter transporters. *The Journal of Physiology* 592: 863–869.

Penmatsa, A., Wang, K.H. and Gouaux, E. 2013. X-ray structure of dopamine transporter elucidates antidepressant mechanism. *Nature* 503: 85–90.

Penmatsa, A., Wang, K.H. and Gouaux, E. 2015. X-ray structures of Drosophila dopamine transporter in complex with nisoxetine and reboxetine. *Nature Structural & Molecular Biology* 22: 506–508.

Perez, C., Koshy, C., Yildiz, O. and Ziegler, C. 2012. Alternating-access mechanism in conformationally asymmetric trimers of the betaine transporter BetP. *Nature* 490: 126–130.

Perez, C., Koshy, C., Ressl, S., Nicklisch, S., Krämer, R. and Ziegler, C. 2011. Substrate specificity and ion coupling in the Na^+ /betaine symporter BetP: Structure and function of BetP-G153D. *The EMBO Journal* 30: 1221–1229.

Perez, C., Faust, B., Mehdipour, A.R., Francesconi, K.A., Forrest, L.R. and Ziegler, C. 2014. Substrate-bound outward-open state of the betaine transporter BetP provides insights into Na+ coupling. *Nature Communications* 5: 4231–42.

Perry, S.L., Guha, S., Pawate, A.S., Henning, R., Kosheleva, I., Srajer, V., Kenis, P.J.A. and Ren, Z. 2014. *In situ* serial Laue diffraction on a microfluidic crystallization device. *Journal of Applied Crystallography* 47: 1975–1982.

Pineda, M. 1999. Identification of a Membrane Protein, LAT-2, that Co-expresses with 4F2 Heavy Chain, an L-type Amino Acid Transport Activity with Broad Specificity for Small and Large Zwitterionic Amino Acids. *Journal of Biological Chemistry* 274: 19738–19744.

Piscitelli, C.L. and Gouaux, E. 2012. Insights into transport mechanism from LeuT engineered to transport tryptophan: Insights into transport mechanism. *The EMBO Journal* 31: 228–235.

Popov, A.N. and Bourenkov, G.P. 2003. Choice of data-collection parameters based on statistic modelling. *Acta Crystallographica. Section D, Biological Crystallography* 59: 1145–1153.

Privé, G.G. 2007. Detergents for the stabilization and crystallization of membrane proteins. *Methods* 41: 388–397.

Quick, M. and Javitch, J.A. 2007. Monitoring the function of membrane transport proteins in detergent-solubilized form. *Proceedings of the National Academy of Sciences of the United States of America* 104: 3603–3608.

Quick, M., Winther, A.-M.L., Shi, L., Nissen, P., Weinstein, H. and Javitch, J.A. 2009. Binding of an octylglucoside detergent molecule in the second substrate (S2) site of LeuT establishes an inhibitor-bound conformation. *Proceedings of the National Academy of Sciences* 106: 5563–5568.

Rebouche, C.J. and Seim, H. 1998. Carnitine metabolism and its regulation in microorganisms and mammals. *Annual Review of Nutrition* 18: 39–61.

Rebsamen, M., Pochini, L., Stasyk, T., de Araújo, M.E.G., Galluccio, M., Kandasamy, R.K., Snijder, B., Fauster, A., Rudashevskaya, E.L., Bruckner, M., Scorzoni, S., Filipek, P.A., Huber, K.V.M., Bigenzahn, J.W., Heinz, L.X., Kraft, C., Bennett, K.L., Indiveri, C., Huber, L.A. and Superti-Furga, G. 2015. SLC38A9 is a component of the lysosomal amino acid sensing machinery that controls mTORC1. *Nature* 519: 477–481.

Reddy, V.S., Shlykov, M.A., Castillo, R., Sun, E.I. and Saier, M.H. 2012. The major facilitator superfamily (MFS) revisited: The major facilitator family revisited. *FEBS Journal* 279: 2022–2035.

Reig, N., del Rio, C., Casagrande, F., Ratera, M., Gelpi, J.L., Torrents, D., Henderson, P.J.F., Xie, H., Baldwin, S.A., Zorzano, A., Fotiadis, D. and Palacin, M. 2007. Functional and structural characterization of the first prokaryotic member of the L-amino acid transporter (LAT) family: a model for APC transporters. *Journal of Biological Chemistry* 282: 13270–13281.

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

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

Reyes, N., Oh, S. and Boudker, O. 2013. Binding thermodynamics of a glutamate transporter homolog. *Nature Structural & Molecular Biology* 20: 634–640.

Robinson, A.S. 2011. *Production of Membrane Proteins: Strategies for Expression and Isolation*. Wiley-VCH, Weinheim, .

Roedig, P., Vartiainen, I., Duman, R., Panneerselvam, S., Stübe, N., Lorbeer, O., Warmer, M., Sutton, G., Stuart, D.I., Weckert, E., David, C., Wagner, A. and Meents, A. 2015. A micro-patterned silicon chip as sample holder for macromolecular crystallography experiments with minimal background scattering. *Scientific Reports* 5: article number 10451.

Roessler, C.G., Agarwal, R., Allaire, M., Alonso-Mori, R., Andi, B., Bachega, J.F.R., Bommer, M., Brewster, A.S., Browne, M.C., Chatterjee, R., Cho, E., Cohen, A.E., Cowan, M., Datwani, S., Davidson, V.L., Defever, J., Eaton, B., Ellson, R., Feng, Y., Ghislain, L.P., Glownia, J.M., Han, G., Hattne, J., Hellmich, J., Héroux, A., Ibrahim, M., Kern, J., Kuczewski, A., Lemke, H.T., Liu, P., Majlof, L., McClintock, W.M., Myers, S., Nelsen, S., Olechno, J., Orville, A.M., Sauter, N.K., Soares, A.S., Soltis, S.M., Song, H., Stearns, R.G., Tran, R., Tsai, Y., Uervirojnangkoorn, M., Wilmot, C.M., Yachandra, V., Yano, J., Yukl, E.T., Zhu, D. and Zouni, A. 2016. Acoustic Injectors for Drop-On-Demand Serial Femtosecond Crystallography. *Structure* 24: 631–640.

Rosell, A., Meury, M., Alvarez-Marimon, E., Costa, M., Perez-Cano, L., Zorzano, A., Fernandez-Recio, J., Palacin, M. and Fotiadis, D. 2014. Structural bases for the interaction and stabilization of the human amino acid transporter LAT2 with its ancillary protein 4F2hc. *Proceedings of the National Academy of Sciences* 111: 2966–2971.

Rosenbaum, D.M., Cherezov, V., Hanson, M.A., Rasmussen, S.G.F., Thian, F.S., Kobilka, T.S., Choi, H.-J., Yao, X.-J., Weis, W.I., Stevens, R.C. and Kobilka, B.K. 2007. GPCR Engineering Yields High-Resolution Structural Insights into 2-Adrenergic Receptor Function. *Science* 318: 1266–1273.

Ruffin, V.A., Salameh, A.I., Boron, W.F. and Parker, M.D. 2014. Intracellular pH regulation by acid-base transporters in mammalian neurons. *Frontiers in Physiology* 5: 1–11.

Rummel, Hardmeyer, Widmer, Chiu, Nollert, Locher, Pedruzzi, Landau and Rosenbusch 1998. Lipidic Cubic Phases: New Matrices for the Three-Dimensional Crystallization of Membrane Proteins. *Journal of Structural Biology* 121: 82–91.

Sagne, C., Agulhon, C., Ravassard, P., Darmon, M., Hamon, M., El Mestikawy, S., Gasnier, B. and Giros, B. 2001. Identification and characterization of a lysosomal transporter for small neutral amino acids. *Proceedings of the National Academy of Sciences* 98: 7206–7211.

Saier, M.H. 2000. A functional-phylogenetic classification system for transmembrane solute transporters. *Microbiology and Molecular Biology Reviews: MMBR* 64: 354–411.

Salom, D., Padayatti, P.S. and Palczewski, K. 2013. Crystallization of G Protein-Coupled Receptors. In: *Methods in Cell Biology,* Elsevier, vol. 117. pp. 451–468.

Sancak, Y., Peterson, T.R., Shaul, Y.D., Lindquist, R.A., Thoreen, C.C., Bar-Peled, L. and Sabatini, D.M. 2008. The Rag GTPases Bind Raptor and Mediate Amino Acid Signaling to mTORC1. *Science* 320: 1496–1501.

Sauter, N.K., Hattne, J., Grosse-Kunstleve, R.W. and Echols, N. 2013. New Python-based methods for data processing. *Acta Crystallographica Section D Biological Crystallography* 69: 1274–1282.

Schloss, P., Mayser, W. and Betz, H. 1992. Neurotransmitter transporters A novel family of integral plasma membrane proteins. *FEBS Letters* 307: 76–80.

Schubert, R., Kapis, S., Gicquel, Y., Bourenkov, G., Schneider, T.R., Heymann, M., Betzel, C. and Perbandt, M. 2016. A multicrystal diffraction data-collection approach for studying structural dynamics with millisecond temporal resolution. *IUCrJ* 3: 393–401.

Schulze, S., Köster, S., Geldmacher, U., Terwisscha van Scheltinga, A.C. and Kühlbrandt, W. 2010. Structural basis of Na^+-independent and cooperative substrate/product antiport in CaiT. *Nature* 467: 233–236.

Schulze-Briese, C., Wagner, A., Tomizaki, T. and Oetiker, M. 2005. Beam-size effects in radiation damage in insulin and thaumatin crystals. *Journal of Synchrotron Radiation* 12: 261–267.

Schushan, M., Rimon, A., Haliloglu, T., Forrest, L.R., Padan, E. and Ben-Tal, N. 2012. A Model-Structure of a Periplasm-facing State of the NhaA Antiporter Suggests the Molecular Underpinnings of pH-induced Conformational Changes. *Journal of Biological Chemistry* 287: 18249–18261.

Screpanti, E. and Hunte, C. 2007. Discontinuous membrane helices in transport proteins and their correlation with function. *Journal of Structural Biology* 159: 261–267.

Shaffer, P.L., Goehring, A., Shankaranarayanan, A. and Gouaux, E. 2009. Structure and Mechanism of a Na^+-Independent Amino Acid Transporter. *Science* 325: 1010–1014.

Shang, X., Zhang, Y., Zhang, G., Chai, X., Deng, A., Liang, Y. and Wen, T. 2013. Characterization and molecular mechanism of AroP as an aromatic amino acid and histidine transporter in *Corynebacterium glutamicum*. *Journal of Bacteriology* 195: 5334–5342.

Sheldrick, G.M. 2008. A short history of *SHELX*. *Acta Crystallographica Section A Foundations of Crystallography* 64: 112–122.

Shen, J., Zeng, Y., Zhuang, X., Sun, L., Yao, X., Pimpl, P. and Jiang, L. 2013. Organelle pH in the Arabidopsis Endomembrane System. *Molecular Plant* 6: 1419–1437.

Sherrell, D.A., Foster, A.J., Hudson, L., Nutter, B., O'Hea, J., Nelson, S., Paré-Labrosse, O., Oghbaey, S., Miller, R.J.D. and Owen, R.L. 2015. A modular and compact portable mini-endstation for high-precision, high-speed fixed target serial crystallography at FEL and synchrotron sources. *Journal of Synchrotron Radiation* 22: 1372–1378.

Shimamoto, K., Lebrun, B., Yasuda-Kamatani, Y., Sakaitani, M., Shigeri, Y., Yumoto, N. and Nakajima, T. 1998. DL-threo-beta-benzyloxyaspartate, a potent blocker of excitatory amino acid transporters. *Molecular Pharmacology* 53: 195–201.

Shimamura, T., Weyand, S., Beckstein, O., Rutherford, N.G., Hadden, J.M., Sharples, D., Sansom, M.S.P., Iwata, S., Henderson, P.J.F. and Cameron, A.D. 2010. Molecular basis of alternating access membrane transport by the sodium-hydantoin transporter Mhp1. *Science* 328: 470–473.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D. and Higgins, D.G. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* 7: 539–544.

Simmons, K.J., Jackson, S.M., Brueckner, F., Patching, S.G., Beckstein, O., Ivanova, E., Geng, T., Weyand, S., Drew, D., Lanigan, J., Sharples, D.J., Sansom, M.S., Iwata, S., Fishwick, C.W., Johnson, A.P., Cameron, A.D. and Henderson, P.J. 2014. Molecular mechanism of ligand recognition by membrane transport protein, Mhp1. *The EMBO Journal* 33: 1831–1844.

Singh, S.K., Yamashita, A. and Gouaux, E. 2007. Antidepressant binding site in a bacterial homologue of neurotransmitter transporters. *Nature* 448: 952–956.

Singh, S.K., Piscitelli, C.L., Yamashita, A. and Gouaux, E. 2008. A Competitive Inhibitor Traps LeuT in an Open-to-Out Conformation. *Science* 322: 1655–1661.

Sitte, H.H. 2004. Sodium-Dependent Neurotransmitter Transporters: Oligomerization as a Determinant of Transporter Function and Trafficking. *Molecular Interventions* 4: 38–47.

Skou, J.C. 1957. The influence of some cations on an adenosine triphosphatase from peripheral nerves. *Biochimica Et Biophysica Acta* 23: 394–401.

Slonczewski, J.L., Fujisawa, M., Dopson, M. and Krulwich, T.A. 2009. Cytoplasmic pH Measurement and Homeostasis in Bacteria and Archaea. In: *Advances in Microbial Physiology,* Elsevier, vol. 55. pp. 1–317.

Slotboom, D.J. 2013. Structural and mechanistic insights into prokaryotic energy-coupling factor transporters. *Nature Reviews Microbiology* 12: 79–87.

Slotboom, D.J., Konings, W.N. and Lolkema, J.S. 1999. Structural features of the glutamate transporter family. *Microbiology and Molecular Biology Reviews: MMBR* 63: 293–307.

Solcan, N., Kwok, J., Fowler, P.W., Cameron, A.D., Drew, D., Iwata, S. and Newstead, S. 2012. Alternating access mechanism in the POT family of oligopeptide transporters. *The EMBO Journal* 31: 3411–3421.

Song, Y., Kenworthy, A.K. and Sanders, C.R. 2014. Cholesterol as a co-solvent and a ligand for membrane proteins: Cholesterol-Protein Interactions. *Protein Science* 23: 1–22.

Sonoda, Y., Newstead, S., Hu, N.-J., Alguel, Y., Nji, E., Beis, K., Yashiro, S., Lee, C., Leung, J., Cameron, A.D., Byrne, B., Iwata, S. and Drew, D. 2011. Benchmarking Membrane Protein Detergent Stability for Improving Throughput of High-Resolution X-ray Structures. *Structure* 19: 17–25.

Spencer, R.F. and Baker, R. 1992. GABA and Glycine as Inhibitory Neurotransmitters in the Vestibuloocular Reflex. *Annals of the New York Academy of Sciences* 656: 602–611.

Stagno, J.R., Liu, Y., Bhandari, Y.R., Conrad, C.E., Panja, S., Swain, M., Fan, L., Nelson, G., Li, C., Wendel, D.R., White, T.A., Coe, J.D., Wiedorn, M.O., Knoska, J., Oberthuer, D., Tuckey, R.A., Yu, P., Dyba, M., Tarasov, S.G., Weierstall, U., Grant, T.D., Schwieters, C.D., Zhang, J., Ferré-D'Amaré, A.R., Fromme, P., Draper, D.E., Liang, M., Hunter, M.S., Boutet, S., Tan, K., Zuo, X., Ji, X., Barty, A., Zatsepin, N.A., Chapman, H.N., Spence, J.C.H., Woodson, S.A. and Wang, Y.-X. 2016. Structures of riboswitch RNA reaction states by mix-and-inject XFEL serial crystallography. *Nature* 541: 242–246.

Stamm, M., Staritzbichler, R., Khafizov, K. and Forrest, L.R. 2014. AlignMe--a membrane protein sequence alignment web server. *Nucleic Acids Research* 42: 246–251.

Stein, N. 2008. *CHAINSAW* : a program for mutating pdb files used as templates in molecular replacement. *Journal of Applied Crystallography* 41: 641–643.

Stockbridge, R.B., Kolmakova-Partensky, L., Shane, T., Koide, A., Koide, S., Miller, C. and Newstead, S. 2015. Crystal structures of a double-barrelled fluoride ion channel. *Nature* 525: 548–551.

Su, Y.-H. 2004. Molecular and Functional Characterization of a Family of Amino Acid Transporters from Arabidopsis. *Plant Physiology* 136: 3104–3113.

Suga, M., Akita, F., Hirata, K., Ueno, G., Murakami, H., Nakajima, Y., Shimizu, T., Yamashita, K., Yamamoto, M., Ago, H. and Shen, J.-R. 2014. Native structure of photosystem II at 1.95 Å resolution viewed by femtosecond X-ray pulses. *Nature* 517: 99–103.

Suzuki, S. and Henderson, P.J.F. 2006. The Hydantoin Transport Protein from Microbacterium liquefaciens. *Journal of Bacteriology* 188: 3329–3336.

Swier, L.J.Y.M., Guskov, A. and Slotboom, D.J. 2016. Structural insight in the toppling mechanism of an energy-coupling factor transporter. *Nature Communications* 7: atricle number 11072.

Szabó, C., Southan, G.J., Wood, E., Thiemermann, C. and Vane, J.R. 1994. Inhibition by spermine of the induction of nitric oxide synthase in J774.2 macrophages: requirement of a serum factor. *British Journal of Pharmacology* 112: 355–356.

Tang, L., Bai, L., Wang, W. and Jiang, T. 2010. Crystal structure of the carnitine transporter and insights into the antiport mechanism. *Nature Structural & Molecular Biology* 17: 492–496.

Tang, L., El-Din, T.M.G., Swanson, T.M., Pryde, D.C., Scheuer, T., Zheng, N. and Catterall, W.A. 2016. Structural basis for inhibition of a voltage-gated $Ca2^+$ channel by $Ca2^+$ antagonist drugs. *Nature* 537: 117–121.

Tavoulari, S., Rizwan, A.N., Forrest, L.R. and Rudnick, G. 2011. Reconstructing a Chloride-binding Site in a Bacterial Neurotransmitter Transporter Homologue. *Journal of Biological Chemistry* 286: 2834–2842.

Tenboer, J., Basu, S., Zatsepin, N., Pande, K., Milathianaki, D., Frank, M., Hunter, M., Boutet, S., Williams, G.J., Koglin, J.E., Oberthuer, D., Heymann, M., Kupitz, C., Conrad, C., Coe, J., Roy-Chowdhury, S., Weierstall, U., James, D., Wang, D., Grant, T., Barty, A., Yefanov, O., Scales, J., Gati, C., Seuring, C., Srajer, V., Henning, R., Schwander, P., Fromme, R., Ourmazd, A., Moffat, K., Van Thor, J.J., Spence, J.C.H., Fromme, P., Chapman, H.N. and Schmidt, M. 2014. Time-resolved serial crystallography captures high-resolution intermediates of photoactive yellow protein. *Science* 346: 1242–1246.

Tereshko, V., Uysal, S., Koide, A., Margalef, K., Koide, S. and Kossiakoff, A.A. 2008. Toward chaperone-assisted crystallography: protein engineering enhancement of crystal packing and X-ray phasing capabilities of a camelid single-domain antibody (VHH) scaffold. *Protein Science : A Publication of the Protein Society* 17: 1175–1187.

Thorn, A. and Sheldrick, G.M. 2011. *ANODE* : anomalous and heavy-atom density calculation. *Journal of Applied Crystallography* 44: 1285–1287.

Thwaites, D.T., McEwan, G.T., Brown, C.D., Hirst, B.H. and Simmons, N.L. 1993. Na^+-independent, H^+-coupled transepithelial beta-alanine absorption by human intestinal Caco-2 cell monolayers. *The Journal of Biological Chemistry* 268: 18438–18441.

Thwaites, D.T., Basterfield, L., McCleave, P.M.J., Carter, S.M. and Simmons, N.L. 2000. Gamma-aminobutyric acid (GABA) transport across human intestinal epithelial (Caco-2) cell monolayers. *British Journal of Pharmacology* 129: 457–464.

Trimmer, B.A. 2001. Nitric Oxide and the Control of Firefly Flashing. *Science* 292: 2486–2488.

Tsai, M.-F., Fang, Y. and Miller, C. 2012. Sided functions of an arginine-agmatine antiporter oriented in liposomes. *Biochemistry* 51: 1577–1585.

Västermark, A. and Saier, M.H. 2014. Evolutionary relationship between 5+5 and 7+7 inverted repeat folds within the amino acid-polyamine-organocation superfamily. *Proteins* 82: 336–346.

Vastermark, A., Wollwage, S., Houle, M.E., Rio, R. and Saier, M.H. 2014. Expansion of the APC superfamily of secondary carriers. *Proteins* 82: 2797–2811.

Vavilin, D.V. and Vermaas, W.F.J. 2002. Regulation of the tetrapyrrole biosynthetic pathway leading to heme and chlorophyll in plants and cyanobacteria. *Physiologia Plantarum* 115: 9–24.

Vékony, N., Wolf, S., Boissel, J.-P., Gnauert, K. and Closs, E.I. 2001. Human Cationic Amino Acid Transporter hCAT-3 Is Preferentially Expressed in Peripheral Tissues †. *Biochemistry* 40: 12387–12394.

Verdon, G. and Boudker, O. 2012. Crystal structure of an asymmetric trimer of a bacterial glutamate transporter homolog. *Nature Structural & Molecular Biology* 19: 355–357.

Verdon, G., Oh, S., Serio, R.N. and Boudker, O. 2014. Coupled ion binding and structural transitions along the transport cycle of glutamate transporters. *eLife* 3.

Vergara-Jaque, A., Fenollar-Ferrer, C., Kaufmann, D. and Forrest, L.R. 2015. Repeat-swap homology modeling of secondary active transporters: updated protocol and prediction of elevator-type mechanisms. *Frontiers in Pharmacology* 6: 1–12.

Viitanen, P., Newman, M.J., Foster, D.L., Wilson, T.H. and Kaback, H.R. 1986. Purification, reconstitution, and characterization of the lac permease of *Escherichia coli*. *Methods in Enzymology* 125: 429–452.

Voglis, G. and Tavernarakis, N. 2006. The role of synaptic ion channels in synaptic plasticity. *EMBO Reports* 7: 1104–1110.

Wagner, A., Duman, R., Henderson, K. and Mykhaylyk, V. 2016. In-vacuum long-wavelength macromolecular crystallography. *Acta Crystallographica Section D Structural Biology* 72: 430–439.

Wang, H. and Gouaux, E. 2012. Substrate binds in the S1 site of the F253A mutant of LeuT, a neurotransmitter sodium symporter homologue. *EMBO Reports* 13: 861–866.

Wang, H., Elferich, J. and Gouaux, E. 2012. Structures of LeuT in bicelles define conformation and substrate binding in a membrane-like context. *Nature Structural & Molecular Biology* 19: 212–219.

Wang, H., Goehring, A., Wang, K.H., Penmatsa, A., Ressler, R. and Gouaux, E. 2013. Structural basis for action by diverse antidepressants on biogenic amine transporters. *Nature* 503: 141–145.

Wang, K.H., Penmatsa, A. and Gouaux, E. 2015. Neurotransmitter and psychostimulant recognition by the dopamine transporter. *Nature* 521: 322–327.

Wang, S., Yan, R., Zhang, X., Chu, Q. and Shi, Y. 2014. Molecular mechanism of pH-dependent substrate transport by an arginine-agmatine antiporter. *Proceedings of the National Academy of Sciences of the United States of America* 111: 12734–12739.

Watanabe, A., Choe, S., Chaptal, V., Rosenberg, J.M., Wright, E.M., Grabe, M. and Abramson, J. 2010. The mechanism of sodium and substrate release from the binding pocket of vSGLT. *Nature* 468: 988–991.

Webb, B. and Sali, A. 2014. Comparative Protein Structure Modeling Using MODELLER: Comparative Protein Structure Modeling Using MODELLER. In: Bateman, A., W.R. Pearson, L.D. Stein, G.D. Stormo and J.R. Yates (Eds.) , *Current Protocols in Bioinformatics,* John Wiley & Sons, Inc., Hoboken, NJ, USA, pp. 5.6.1-5.6.32.

Weierstall, U. 2014. Liquid sample delivery techniques for serial femtosecond crystallography. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369: 20130337–20130337.

Weierstall, U., Spence, J.C.H. and Doak, R.B. 2012. Injector for scattering measurements on fully solvated biospecies. *The Review of Scientific Instruments* 83: article number 035108.

Weierstall, U., James, D., Wang, C., White, T.A., Wang, D., Liu, W., Spence, J.C.H., Bruce Doak, R., Nelson, G., Fromme, P., Fromme, R., Grotjohann, I., Kupitz, C., Zatsepin, N.A., Liu, H., Basu, S., Wacker, D., Won Han, G., Katritch, V., Boutet, S., Messerschmidt, M., Williams, G.J., Koglin, J.E., Marvin Seibert, M., Klinker, M., Gati, C., Shoeman, R.L., Barty, A., Chapman, H.N., Kirian, R.A., Beyerlein, K.R., Stevens, R.C., Li, D., Shah, S.T.A., Howe, N., Caffrey, M. and Cherezov, V. 2014. Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography. *Nature Communications* 5: article number 3309.

Welander, P.V., Hunter, R.C., Zhang, L., Sessions, A.L., Summons, R.E. and Newman, D.K. 2009. Hopanoids Play a Role in Membrane Integrity and pH Homeostasis in Rhodopseudomonas palustris TIE-1. *Journal of Bacteriology* 191: 6145–6156.

Weyand, S., Shimamura, T., Beckstein, O., Sansom, M.S.P., Iwata, S., Henderson, P.J.F. and Cameron, A.D. 2011. The alternating access mechanism of transport as observed in the sodium-hydantoin transporter Mhp1. *Journal of Synchrotron Radiation* 18: 20–23.

Weyand, S., Shimamura, T., Yajima, S., Suzuki, S., Mirza, O., Krusong, K., Carpenter, E.P., Rutherford, N.G., Hadden, J.M., O'Reilly, J., Ma, P., Saidijam, M., Patching, S.G., Hope, R.J., Norbertczak, H.T., Roach, P.C.J., Iwata, S., Henderson, P.J.F. and Cameron, A.D. 2008. Structure and Molecular Mechanism of a Nucleobase-Cation-Symport-1 Family Transporter. *Science* 322: 709–713.

White, T.A., Kirian, R.A., Martin, A.V., Aquila, A., Nass, K., Barty, A. and Chapman, H.N. 2012. CrystFEL: A software suite for snapshot serial crystallography. *Journal of Applied Crystallography* 45: 335–341.

White, T.A., Mariani, V., Brehm, W., Yefanov, O., Barty, A., Beyerlein, K.R., Chervinskii, F., Galli, L., Gati, C., Nakane, T., Tolstikova, A., Yamashita, K., Yoon, C.H., Diederichs, K. and Chapman, H.N. 2016. Recent developments in *CrystFEL*. *Journal of Applied Crystallography* 49: 680–689.

Winter, G. 2010. *xia2* : an expert system for macromolecular crystallography data reduction. *Journal of Applied Crystallography* 43: 186–190.

Wise, R.A. 1996. Neurobiology of addiction. *Current Opinion in Neurobiology* 6: 243–251.

Wohri, A.B., Katona, G., Johansson, L.C., Fritz, E., Malmerberg, E., Andersson, M., Vincent, J., Eklund, M., Cammarata, M., Wulff, M., Davidsson, J., Groenhof, G. and Neutze, R. 2010. Light-Induced Structural Changes in a Photosynthetic Reaction Center Caught by Laue Diffraction. *Science* 328: 630–633.

Wolf, S., Janzen, A., VéKony, N., Martiné, U., Strand, D. and Closs, E.I. 2002. Expression of solute carrier 7A4 (SLC7A4) in the plasma membrane is not sufficient to mediate amino acid transport activity. *Biochemical Journal* 364: 767–775.

Wolfbeis, O.S. 1983. Fluorimetric Analysis 1. A Study on Fluorescent Indicators for Measuring Near Neutral Fluorimetric analysis: I. A study on fluorescent indicators for measuring near neutral ('physiological') pH values. *Fresenius Zeitschrift Fuer Analytische Chemie* 314: 119–124.

Wong, F.H., Chen, J.S., Reddy, V., Day, J.L., Shlykov, M.A., Wakabayashi, S.T. and Saier, J. 2012. The Amino Acid-Polyamine-Organocation Superfamily. *Journal of Molecular Microbiology and Biotechnology* 22: 105–113.

Wright, E.M. 2004. Surprising Versatility of Na^+-Glucose Cotransporters: SLC5. *Physiology* 19: 370–376.

Wu, C.-H., Bialecka-Fornal, M. and Newman, D.K. 2015a. Methylation at the C-2 position of hopanoids increases rigidity in native bacterial membranes. *eLife* 4.

Wu, G. 2009. Amino acids: metabolism, functions, and nutrition. *Amino Acids* 37: 1–17.

Wu, W., Nogly, P., Rheinberger, J., Kick, L.M., Gati, C., Nelson, G., Deupi, X., Standfuss, J., Schertler, G. and Panneels, V. 2015b. Batch crystallization of rhodopsin for structural dynamics using an X-ray free-electron laser. *Acta Crystallographica. Section F, Structural Biology Communications* 71: 856–860.

Xu, F., Liu, W., Hanson, M.A., Stevens, R.C. and Cherezov, V. 2011. Development of an Automated High Throughput LCP-FRAP Assay to Guide Membrane Protein Crystallization in Lipid Mesophases. *Crystal Growth & Design* 11: 1193–1201.

Xu, W., Liu, L.Z., Loizidou, M., Ahmed, M. and Charles, I.G. 2002. The role of nitric oxide in cancer. *Cell Research* 12: 311–320.

Yamashita, A., Singh, S.K., Kawate, T., Jin, Y. and Gouaux, E. 2005. Crystal structure of a bacterial homologue of Na+/Cl--dependent neurotransmitter transporters. *Nature* 437: 215–223.

Yamashita, K., Pan, D., Okuda, T., Sugahara, M., Kodan, A., Yamaguchi, T., Murai, T., Gomi, K., Kajiyama, N., Mizohata, E., Suzuki, M., Nango, E., Tono, K., Joti, Y., Kameshima, T., Park, J., Song, C., Hatsui, T., Yabashi, M., Iwata, S., Kato, H., Ago, H., Yamamoto, M. and Nakatsu, T. 2015. An isomorphous replacement method for efficient de novo phasing for serial femtosecond crystallography. *Scientific Reports* 5: article number 14017.

Yang, H., Bogner, M., Stierhof, Y.-D. and Ludewig, U. 2010. H^+-Independent Glutamine Transport in Plant Root Tips. *PLoS ONE* 5: e8917.

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

Yoshimoto, T., Yoshimoto, E. and Meruelo, D. 1993. Identification of amino acid residues critical for infection with ecotropic murine leukemia retrovirus. *Journal of Virology* 67: 1310–1314.

Zhang, H., Unal, H., Gati, C., Han, G.W., Liu, W., Zatsepin, N.A., James, D., Wang, D., Nelson, G., Weierstall, U., Sawaya, M.R., Xu, Q., Messerschmidt, M., Williams, G.J., Boutet, S., Yefanov, O.M., White, T.A., Wang, C., Ishchenko, A., Tirupula, K.C., Desnoyer, R., Coe, J., Conrad, C.E., Fromme, P., Stevens, R.C., Katritch, V., Karnik, S.S. and Cherezov, V. 2015. Structure of the Angiotensin receptor revealed by serial femtosecond crystallography. *Cell* 161: 833–844.

Zhang, H., Qiao, A., Yang, D., Yang, L., Dai, A., de Graaf, C., Reedtz-Runge, S., Dharmarajan, V., Zhang, H., Han, G.W., Grant, T.D., Sierra, R.G., Weierstall, U., Nelson, G., Liu, W., Wu, Y., Ma, L., Cai, X., Lin, G., Wu, X., Geng, Z., Dong, Y., Song, G., Griffin, P.R., Lau, J., Cherezov, V., Yang, H., Hanson, M.A., Stevens, R.C., Zhao, Q., Jiang, H., Wang, M.-W. and Wu, B. 2017a. Structure of the full-length glucagon class B G-protein-coupled receptor. *Nature* 546: 259–264.

Zhang, X., Zhao, F., Wu, Y., Yang, J., Han, G.W., Zhao, S., Ishchenko, A., Ye, L., Lin, X., Ding, K., Dharmarajan, V., Griffin, P.R., Gati, C., Nelson, G., Hunter, M.S., Hanson, M.A., Cherezov, V., Stevens, R.C., Tan, W., Tao, H. and Xu, F. 2017b. Crystal structure of a multi-domain human smoothened receptor in complex with a super stabilizing ligand. *Nature Communications* 8: article number 15383.

Zhao, C., Noskov, S.Y. and Livesay, D.R. 2013. The Molecular Mechanism of Ion-Dependent Gating in Secondary Transporters. *PLoS Computational Biology* 9: e1003296.

Zhao, Y., Terry, D., Shi, L., Weinstein, H., Blanchard, S.C. and Javitch, J.A. 2010. Single-molecule dynamics of gating in a neurotransmitter transporter homologue. *Nature* 465: 188–193.

Zhou, X., Levin, E.J., Pan, Y., McCoy, J.G., Sharma, R., Kloss, B., Bruni, R., Quick, M. and Zhou, M. 2013. Structural basis of the alternating-access mechanism in a bile acid transporter. *Nature* 505: 569–573.

Zhou, Z., Zhen, J., Karpowich, N.K., Goetz, R.M., Law, C.J., Reith, M.E.A. and Wang, D.-N. 2007. LeuT-Desipramine Structure Reveals How Antidepressants Block Neurotransmitter Reuptake. *Science* 317: 1390–1393.

Zomot, E., Bendahan, A., Quick, M., Zhao, Y., Javitch, J.A. and Kanner, B.I. 2007. Mechanism of chloride interaction with neurotransmitter:sodium symporters. *Nature* 449: 726–730.