Biochemical and biophysical studies of the prokaryotic proton dependent oligopeptide transporters

Nicolae Solcan

Department of Biochemistry
and
Linacre College

A thesis submitted for the degree of *Doctor of Philosophy*,
December 2013

Supervisor: Dr Simon Newstead

Biochemical and biophysical studies of the prokaryotic proton dependent oligopeptide transporters

Nicolae Solcan

Department of Biochemistry and Linacre College, University of Oxford

Submitted in partial fulfillment of the requirements for the degree of *Doctor of Philosophy*

Trinity Term 2013

Abstract

The proton dependent oligopeptide transporters (POT family) are members of the Major Facilitator Superfamily of secondary active transporter proteins. They use the transmembrane proton gradient to drive the uptake of di- and tripeptides into the cytoplasm. Members of the family are highly conserved in pro- and eukaryotic genomes, and in humans they are responsible for the oral absorption of many drug families, including β-lactam antibiotics. Recently, the crystal structures of $PepT_{So}$ and $PepT_{St}$, two prokaryotic homologues of the human proteins PepT1 and PepT2, captured the proteins in two distinct conformations, providing insight into the structural aspects of the transport mechanism. A protocol was designed for functional liposome reconstitution of POT proteins, and transport assays were conducted to characterise their substrate specificity, pH dependence and kinetic properties. Using site-directed mutagenesis, we identified binding site residues involved in peptide recognition and proton translocation, and distinguished between the two roles by comparing protein activity in proton- and peptide-driven conditions. We also investigated the roles of key residues in the conformational transitions that accompany the transport cycle, using data from biochemical assays, molecular dynamics simulations and modeling, as well as electron paramagnetic resonance measurements. In addition, several bacterial POT members were screened for crystallisation, in order to assess their stability and crystal diffraction quality in different detergents. Further work was performed with bacterial POT homologues YdgR and GkPOT, including binding studies using NMR spectroscopy and assaying drug transport *in vivo* and *in vitro*. Together, the data establish bacterial POTs as model systems for studying the mammalian oligopeptide transporters, and a mechanistic model for peptide transport is proposed.

I dedicate this work to my mother, who has been my best friend and greatest inspiration throughout the years. She has taught me to love freedom and beauty, and how to recognise them in the many shapes and colours of the world. Today, she continues to remind me that everything is possible given the necessary amount of love and confidence; no lesson has served me better, and for that I thank her.

Acknowledgements

I would first like to thank Simon, my supervisor, for being extremely supportive and helpful during the last three and a half years. His ideas, the constant feedback, and his contagious enthusiasm for the project have all been crucial elements to a rewarding DPhil experience. It took me a couple of years in the lab to get over doubts about my career choices and finally feel comfortable as a scientist, and to a great extent I owe that transformation to his confidence and support.

Just as importantly, I am glad to have been surrounded by wonderful lab mates. I want to thank Jo Parker for having answers to every science-related question I have ever asked, and generally for being awesome. I also thank Sonya Hanson and Johnny Beale, long-standing members of the Newstead Team, for all the fun times in and outside the lab – research is nice, but it's even nicer when you're working side by side with people you like. Great times were had with past lab members as well: working with Jane, Daphne and Marko has been a pleasure, and I am happy to have them as friends.

Many other people have helped me during my DPhil. I am grateful to Jason Schnell, Jo Claridge and Christina Redfield for their invaluable expertise and support with the NMR experiments. I am indebted to Marcella Orwick for introducing me to the world of liposomes, and to Patricia Dijkman for her work and the useful discussions about the DEER project. Many thanks to Kathryn Scott for her help and counsel concerning the radiation work. I thank Chris Mulligan for his advice on setting up the transport assays, and Phil Fowler for the insightful discussion on the transport mechanism. I also thank David Staunton, Ed Lowe and Matthew Jennions for helping with various experiments. More thanks go out to our other collaborators and the people who were involved in aspects of the peptide transporter project at different stages: Lucy Forrest and Sebastian Radestock, Ryuichiro Ishitani and Hideaki Kato, and Mohd Samsudin. Their contributions have made this work richer and more meaningful and are greatly appreciated.

Lastly, and most importantly, I would like to thank Jess for standing by me when I was overcome by uncertainty and needed a friend. Just one touch of your hand makes everything better. Meeting you has brought me joy and strength; I can only hope I never take that for granted, and offer you the same in return.

Table of contents

List of abbreviations

ACE	angiotensin enzyme
AMP-PNP	adenylyl-imidodiphosphate
APS	ammonium persulfate
ATP	adenosine triphosphate
ATR	attentuated total reflectance
$C_{12}E_8$	octaethylene glycol monododecyl ether
CCCP	carbonyl cyanide m-chlorophenylhydrazone
CL	cardiolipin
CMC	critical micelle concentration
CPM	(radiation) counts per minute
CV	column volume
DDM	n-dodecyl β-D-maltoside
DEER	double electron-electron resonance
DHPC	1,2-diheptanoyl-*sn*-glycero-3-phosphocholine
DM	n-decyl β-D-maltoside
DMPC	1,2-dimyristoyl-*sn*-glycero-3-phosphocholine
DMSO	dimethylsulfoxide
dNTP	deoxyribonucleotides
DPPC	1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine
DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EM	electron microscopy
EPL	*E. coli* polar lipids
EPR	electron paramagnetic resonance

ETC	electron transport chain
FRET	Förster resonance energy transfer
FSEC	fluorescence size exclusion chromatography
FTIR	Fourier transform infrared spectroscopy
GFP	green fluorescent protein
GlySar	glycylsarcosine
GPCR	G-protein coupled receptor
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
HMQC	heteronuclear multiple quantum correlation
I3C	5-amino-2,4,6-triiodoisophthalic acid
IC_{50}	half-maximal inhibitory concentration
IPTG	isopropyl β-D-1-thiogalactopyranoside
ITC	isothermal titration calorimetry
KAc	potassium acetate
Kan	kanamycin
K_D	dissociation constant
K_i	inhibition constant
K_m^{app}	apparent K_m of transport
KPi	potassium phosphate
L-DOPA	L-3,4-dihydroxyphenylalanine
LDAO	lauryldimethylamine oxide
LMV	large multilamellar vesicle
LUV	large unilamellar vesicle
MCS	multiple cloning site
MD	molecular dynamics
MES	2-(*N*-morpholino)ethanesulfonic acid
MFS	Major Facilitator Superfamily
MMC	methyl mercury chloride
MOPS	3-(*N*-morpholino)propanesulfonic acid

MTSL	1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl
MW	molecular weight
NADH	nicotinamide adenine nucleotide
NaPi	sodium phosphate
NM	n-nonyl β-D-maltoside
NMR	nuclear magnetic resonance
NSS	neurotransmitter:sodium symporter
OG	n-octyl β-D-glucopyranoside
OPT	oligopeptide transporter (not to be confused with POT/PTR)
PBS	phosphate buffered saline
PC	phosphatidylcholine
PCR	polymerase chain reaction
PE	phosphatidylethanolamine
PEG	polyethylene glycol
PET	positron emission tomography
PG	phosphatidylglycerol
PHT	peptide/histidine transporter
PIPES	piperazine-*N*,*N*′-bis(2-ethanesulfonic acid)
PMF	proton motive force
POT/PTR	proton-dependent oligopeptide transporter/peptide transporter
RALS	right angle light scattering
RFU	relative fluorescence units
RI	refractive index
RMSD	root-mean-square deviation
RT	room temperature
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SeMet	selenomethionine
SPA	scintillation proximity assay
SUV	small unilamellar vesicle

TDG	(β-D-galactopyranosyl-1-thio-β-D-galactopyranoside
TEMED	tetramethylethylenediamine
TEV	tobacco etch virus
TM	transmembrane (helix)
Tris	2-amino-2-hydroxymethyl-propane-1,3-diol
β-Ala-Lys-AMCA	β-L-alanyl-L-lysyl-7-amino-4-methylcoumarin-3-acetic acid
$\Delta\Psi$	membrane potential

"Now the nonsense must stop, and the serious business must begin."

(Scott Walker, 1968)

Chapter 1: Introduction

1.1. Bacterial membrane biology

Viewed in the context of evolution, the advent of the cell as a structural and functional unit of life was a crucial event. Whether or not its original purpose was to concentrate the genetic material into a protective structure, cellular organisation has since become a ubiquitous feature of all life forms, arguably with the exception of viruses. The defining characteristic of all living cells, prokaryotes and eukaryotes alike, is the presence of a lipid membrane. Broadly speaking, the membrane has a double function: it acts as a barrier separating the intracellular compartment from the external environment, but it also enables communication between the two. The components of the membrane by and large dictate how the two functions are balanced.

1.1.1. Lipid composition

Apart from delineating the internal compartment (protoplasm) and defining the cell's shape, the plasma membrane is responsible for the passage of ions and molecules into or out of the cell. This can occur either via passive diffusion (as in the case of oxygen or carbon dioxide), or through selective transport, the latter achieved through the activity of channels and transporters. The lipid component of the membrane consists mainly of phospholipids

(amphipathic molecules with a hydrophilic head group and a hydrophobic tail), which self-assemble to form a continuous three-dimensional bilayer (Finean, 1969; Bretscher, 1972). The hydrophobic effect causes the tails (two long chain fatty acids ester-linked to a diacylglycerol backbone) to self-associate in solution forming the core of the membrane, while the head groups align at the internal and external sides, free to form polar contacts with water (Vance and Vance, 2002). The hydrophobic core contains variable amounts of unsaturated and/or branched alkyl chains, which can disrupt the tight packing contributed by the saturated chains, increasing membrane fluidity (fig. 1-1).

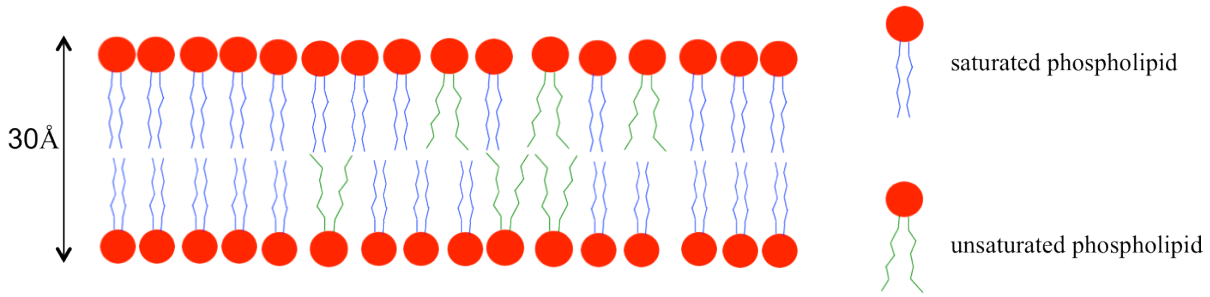


Fig. 1-1: General structure of a phospholipid bilayer.

The basic bilayer architecture is shared by bacterial, archaeal and eukaryotic plasma membranes, but Gram negative bacteria have an additional, outer membrane, with a different lipid composition (Zhang and Rock, 2008), separated from the inner membrane by the periplasmic space. In addition, all bacteria as well as plants, fungi and some archaea are further enclosed by a peptidoglycan cell wall.

The phospholipid content of the bilayer varies with the cell type and can change in response to internal or external factors. Bacterial membranes consist almost exclusively of three phospholipid types: phosphatidylethanolamine (PE), phosphatidylglycerol (PG) and cardiolipin (CL). In *E. coli*, a typical example, they occur with average frequencies of 75% PE, 20% PG and 5% CL (Goldfine, 1984) (fig. 1-2). PE and PG are synthesised in the inner leaflet of the plasma membrane from glycerol-3-phosphate and two acyl chains to yield

phosphatidic acid, to which the polar head group (ethanolamine or glycerol) is attached via an enzymatic reaction (Cronan, 2003). CL is typically obtained from two PG molecules, but an alternative pathway in *E. coli* produces CL using both PE and PG as substrates (Tan *et al*, 2012). The length of the hydrophobic acyl chains is variable, although studies have shown that Gram-positive bacteria feature predominantly C_{15} branched chains, while Gram-negative ones (such as *E. coli*) prefer C_{16} straight chain fatty acids for their inner membrane phospholipids (Cho and Salton, 1966).

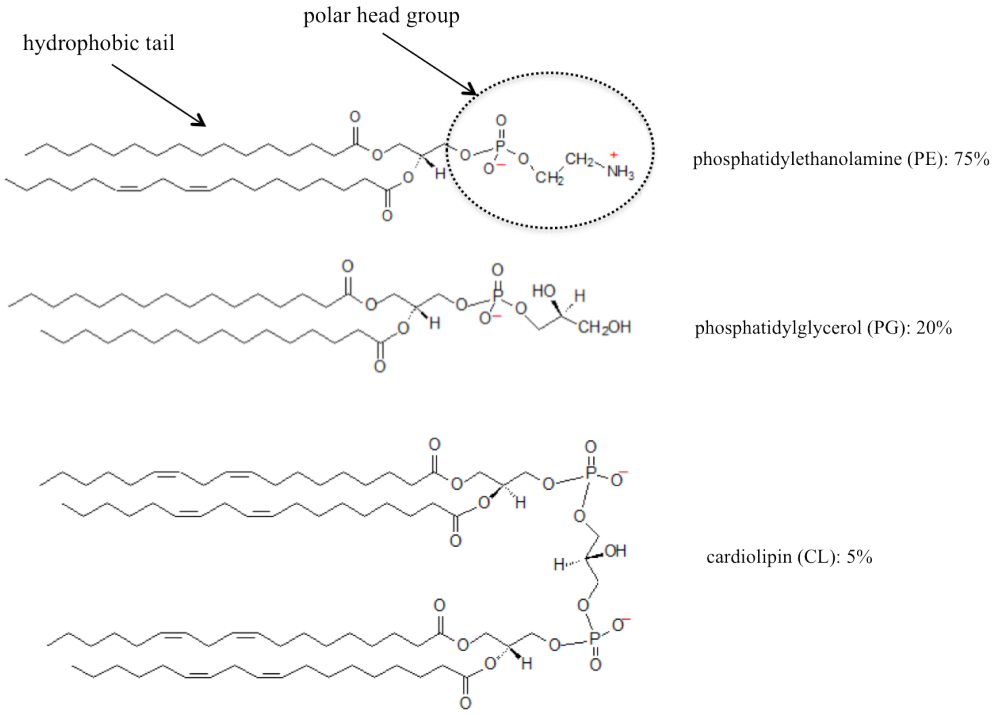


Fig. 1-2: Phospholipids and their preponderance in the *E. coli* inner membrane.

Changing the length and degree of saturation of existing acyl chains is an important strategy used by prokaryotic cells to adapt to changes in the environment such as temperature changes (Farrell and Rose, 1967) or acidic stress (Jordan *et al*, 1999). Phospholipid synthesis and modification is a tightly regulated process, that allows cells to control membrane fluidity/permeability, curvature and, perhaps most crucially, their interaction with integral membrane proteins.

The specifics of lipid-protein interactions in biological membranes are currently not well understood. However, a wealth of data from both *in vivo* and *in vitro* experiments suggests the role of lipids in membrane protein topology and function is crucial. For example, the lack of PE from liposomes (Chen *et al*, 1984; Wang *et al*, 2002) as well as *E. coli* cells (Bogdanov *et al*, 2002) causes lactose permease (LacY) to adopt an inverted orientation in the membrane and abolishes lactose transport. Similarly, proton-coupled multidrug transporter LmrP depends on the zwitterionic head group of PE to facilitate key conformational changes *in vitro*, by changing the pK_a of ionisable side chains (Gbaguidi, 2007). The use of native *E. coli* lipids in liposomes has been shown to increase activity of transporter EmrE three-fold (Curnow *et al*, 2004), while a more recent study has proposed a lipid-dependent gating mechanism by a voltage-dependent potassium channel (Schmidt *et al*, 2009). Furthermore, PG has been found to play a role in membrane insertion of other MPs (Cronan, 2003). CL itself is not a vital membrane component, but it is always in a protonated state at physiological pH and can therefore act as a "proton sink" in transport processes driven by a proton gradient (Vance and Vance, 2002). Whether acting as chaperones or being involved in complex mechanisms such as proton translocation, lipids greatly influence membrane protein function, and this should be taken into consideration whenever one sets out to assay a protein's activity in a membrane environment.

1.1.2. The electrochemical gradient

Over half a century ago, the chemiosmotic theory offered an explanation of the coupling mechanism between the generation of energy and the downhill movement of ions (more specifically protons) across a biological membrane (Mitchell, 1961). A crucial aspect of the

chemiosmotic model is the generation and maintenance of a transmembrane proton gradient (by proton pumps localised in the membrane), which is utilised to drive ATP synthesis. Bacteria depend on this ability, as unicellular systems for which internal pH homeostasis is vital to their adapting to a variable external environment (Krulwich *et al*, 2011). It is at the level of the bacterial plasma membrane (as well as the archaeal and the inner membrane of mitochondria and chloroplasts) that chemical energy is generated in the form of ATP, setting in motion active transport and complex metabolic pathways.

The transmembrane proton gradient is generated through electron transport chains. Examples of these are the mitochondrial respiratory complexes, of which complexes I, III and IV are also proton pumps, and their equivalent redox systems in the inner membrane of respiring bacteria (Berrisford *et al*, 2009). They transfer the electrons derived from the oxidation of NADH to quinone, coupling this to the translocation of protons from the cytosol into the periplasmic space (or from the matrix into the intermembrane space in the case of mitochondria). Alternatively, light energy can be used to generate proton gradients in the absence of electron transfer (by bacteriorhodopsin in archaea) or by photosynthetic electron transport chains in the thylakoidal membrane of chloroplasts, where photosystems I and II and the cytochrome b_6f complex contribute to the transfer of protons into the thylakoid lumen to fuel ATP synthesis (Nicholls and Ferguson, 2002). Several structures of the individual components of the photosynthetic ETC's have been solved to date, such as those of the reaction centres of photosynthetic bacteria (Deisenhofer *et al*, 1985; Roszak *et al*, 2003) and those of the photosystems and the cyt b_6f complex in cyanobacteria and plants (Yamashita *et al*, 2007; Guskov *et al*, 2009; Amunts *et al*, 2010); furthermore, the structure of respiratory complex I was recently solved in its entirety (Efremov *et* al, 2010; Efremov and Sazanov, 2011; Baradaran *et al*, 2013), providing insight into the mechanism of electron transfer and the associated proton translocation.

In the case of complex I for example, the process is driven by conformational changes upon the transfer of two electrons from cytoplasmic NADH to quinone in the binding site, causing the transporter subunits in the membrane to activate and pump out four protons into the periplasm, thereby generating the gradient. The reaction can be summarised as:

$$NADH + H^+ + Q + 4H^+_{in} \rightarrow NAD^+ + QH_2 + 4H^+_{out}$$

, where "in" and "out" denote the intracellular and periplasmic spaces. The resulting proton gradient can be represented as ΔpH (equal to the difference between pH_{in} and pH_{out}). For biological membranes, the cumulated effects of ΔpH and the membrane potential (ΔΨ) add up to what is generally called the electrochemical potential difference, measured in kJ/mol (Rottenberg, 1979), or its equivalent, the proton motive force, measured in mV (Δp). Specifically, the Δp is calculated as follows:

$$\Delta p\ (mV) = \Delta\Psi - 2.3 \cdot (RT/F) \cdot \Delta pH$$

, where R is the universal gas constant, T the absolute temperature and F the Faraday constant (ΔΨ is the potential difference in mV between the inside and outside of the membrane, and ΔpH is the corresponding difference in pH units). ΔΨ is the electrical component of Δp, and ΔpH is the chemical one.

The two components contribute to the electrochemical gradient to different extents in different cells. In most cases, multicellular eukaryotes experience a negligible pH difference across their plasma membranes, as the extracellular pH of tissues is maintained close to physiological values at all times (Krulwich *et al*, 2011). In contrast, ΔpH in bacteria can be quite pronounced, depending on the extracellular environment. The electrochemical gradient in prokaryotes is typically called the proton motive force (Δp) (Rottenberg, 1979). As the ΔpH component of the Δp varies, the bacterial membrane becomes more or less permeable to

protons, which calls for efficient regulation of the ΔΨ (actively modifying the distribution of cations and anions on both sides of the membrane via transporters and channels to compensate for either alkaline or acidic stress). Regulation of ΔpH and ΔΨ in bacteria is tightly interconnected, an example of its use being the adaptation mechanisms of acidophiles and alkaliphiles (Krulwich *et al*, 2011). While neutralophiles like *E. coli* have a negligible ΔpH, they maintain a constant positive (out)/negative (in) charge distribution, adding up to an overall moderately negative value for the Δp (or PMF), which favours metabolic processes (table 1-1). In adapting to a low pH environment, acidophiles counterbalance the effect of the strong pH gradient by creating an inverted membrane potential (- out/+ in), resulting in a similar PMF. In comparison, bacteria living in highly alkaline environments attempt to maintain a strongly negative membrane potential such that the PMF remains negative.

Parameter/Organism	Acidophile	Neutralophile	Alkaliphile
ΔpH	- 4.5	- 0.6	+ 2.2
ΔΨ	+ 10 mV	- 135 mV	- 180 mV
PMF	- 256 mV	- 170 mV	- 50 mV

Table 1-1: Regulation of the proton motive force in prokaryotes, illustrated by examples from different environments (adapted from Krulwich *et al*, 2011).

It is the PMF (or the electrochemical gradient as a more general concept pertaining to lipid membranes) that has the potential to drive processes like the synthesis and transport of molecules at the plasma membrane, through both its electrical and chemical components. As noted earlier, one immediate consequence of this is the production of ATP by ATP synthase through oxidative phosphorylation. ATP can then be utilised to drive primary active transport, while the PMF itself can further drive the uptake of solutes by secondary transporters.

1.2. Active transport in bacterial membranes

The main passage routes for materials into and out of the cell are membrane proteins like transporters and ion channels. Channels can only translocate ions down their concentration gradients, whereas ion transporters (ion pumps) can, in principle, drive ion flow against the gradient, performing "active" transport (Gadsby, 2009). The latter is true for transporters in general, whether their substrates are ions or molecules. The channel/transporter delineation can sometimes be difficult to make in the absence of protein structures and detailed mechanistic data, as has been the case with certain members of the ClC family of proton-coupled chloride transporters, some of which were until recently thought to be ion channels (Miller, 2006). Active transporters themselves are important components of the bacterial membrane, comprising between 3 and 16% of the entire proteome of prokaryotes (Ren and Paulsen, 2007). They can be classified into one of two types, depending on the energy source used to drive the transport.

1.2.1. Primary active transporters

Primary transport systems use a primary energy source such as ATP hydrolysis to drive the transport of solutes across a membrane (Davidson *et al*, 2008). In a broader sense, complex systems like electron transport chains can also be classified as primary active transporters, as

they use primary energy in the form of NADH or photons. Proteins that use ATP hydrolysis are ATP-binding cassette (ABC) transporters and ATP-driven ion pumps.

ATP-driven ion pumps, also known as P-type ATPases, break down ATP to power the transport of cations such as H^+, Na^+, K^+ or Ca^{2+}. Na^+/K^+ ATPases play a central role in establishing the membrane charge potential, while the H^+ ATPase pumps one proton per ATP molecule, replenishing the proton motive force needed for further ATP synthesis (Morth *et al*, 2011).

ABC transporters form the largest protein family in prokaryotes (Ren and Paulsen, 2007); for example, ABC transporter genes make up over 5% of the *E. coli* genome (Blattner *et al*, 1997). They transport molecules of various sizes, from amino acids and sugars to peptides and vitamins, from the cytoplasm into the periplasmic/extracellular space. The transport mechanism is the so-called "ATP switch", involving dimerisation of the cytoplasmic nulceotide binding domains (NBDs) upon binding of two ATP molecules, followed by ATP hydrolysis and subsequent conformational changes in the two transmembrane domains (TMDs) that enable substrate translocation (Higgins and Linton, 2004).

1.2.2. Secondary active transporters: structural and mechanistic aspects

In contrast to ABC transporters and P-type ATPases, secondary active transporters are energised in an indirect fashion, as they use the potential energy contained in a transmembrane ion gradient in order to power the uptake of substrates. They can be

subdivided further, depending on whether the transported substrate and the coupling ion are transported across the membrane in the same direction (symporters), one is exchanged for the other (antiporters), or the substrate is transported without the aid of ion coupling (uniporters) (fig. 1-3).

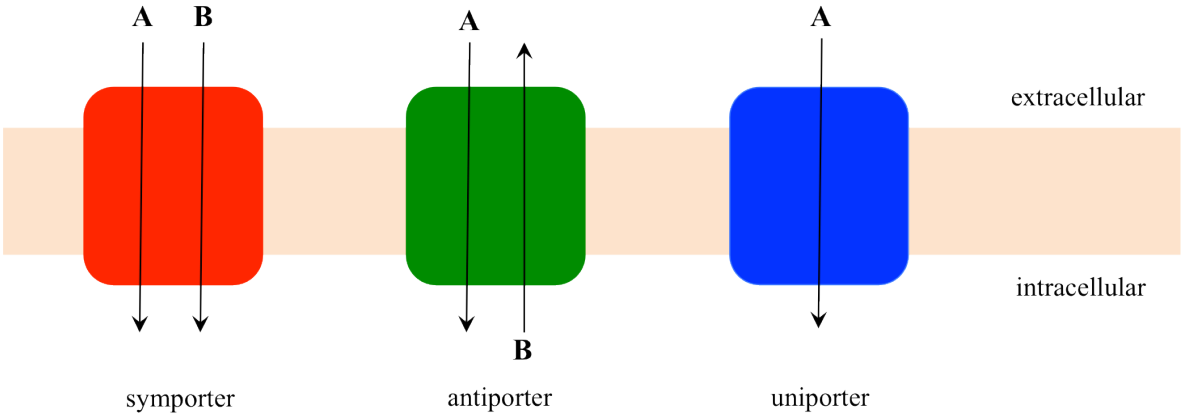


Fig. 1-3: Schematic representation of symport, antiport and uniport showing direction of transport for substrates A and B.

The gradients most commonly used for coupling are those of protons and sodium ions. Proton-coupled transporters are relatively less prominent in eukaryotes compared to sodium-coupled ones, whereas bacteria feature both in similar numbers (Saier, 2000). The tendency of eukaryotic transporters to be sodium-coupled is likely a consequence of the weak transmembrane proton gradients found in most tissues, as discussed in the previous subchapter. Not surprisingly, this tendency is shared with pathogenic bacteria, as well as those living under high-salt conditions. On the contrary, autotrophic prokaryotes tend to have mainly proton-coupled transporters as the PMF provides them with a constant source of energy while other ion gradients are rarely encountered (Ren and Paulsen, 2007).

The current mechanistic model for secondary transporters builds on the hypothesis formulated by Oleg Jardetzky in the 1960s explaining how membrane "pumps" are able to transport substrates against their concentration gradient. Jardetzky hypothesised that a pump should have a membrane-embedded substrate binding site, and be able to assume two conformations: outward-open (exposed to the extracellular medium) and inward-open, each conformation displaying different affinities to allow for binding and release of the substrate (Jardetzky, 1966). The hypothesis has since come to be known as the alternating access model, referring to the binding site being either open or closed to one side of the membrane at any given point (fig. 1-4).

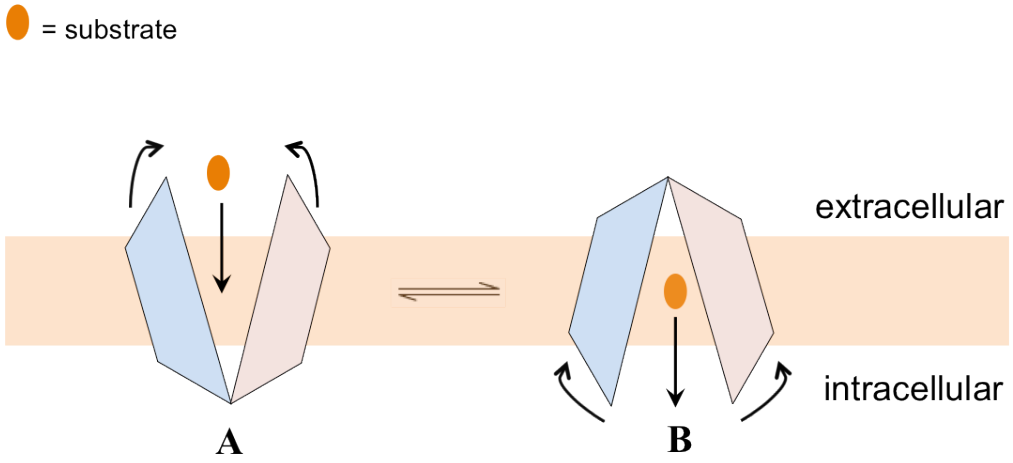


Fig. 1-4: Early model for the alternating access mechanism of transport, featuring the outward-open (A) and inward-open (B) states.

Confirmation of the model came around 40 years later with the first X-ray crystal structures of bacterial secondary active transporters, namely the H^+/lactose symporter lactose permease/LacY (Abramson *et al*, 2003) and the phosphate/glycerol-3-phosphate antiporter GlpT (Huang *et al*, 2003), followed shortly after by the Na^+/leucine symporter LeuT (Yamashita *et al*, 2005). As previously predicted (Goldkorn *et al*, 1983), LacY was shown to be an alpha-helical protein with a core structure of 12 transmembrane helices. The LacY fold is shared by GlpT, while LeuT is also alpha-helical but with a 10TM core. The structures of

LacY and LeuT with bound substrate molecules revealed a substrate binding site localised in the centre of the membrane, a highly conserved region accessible to either the intracellular side (LacY and GlpT) or (partially) the extracellular side (LeuT). The LeuT structure became the blueprint for what would be known as the LeuT fold, with LacY and GlpT representing the MFS fold (Major Facilitator Superfamily) (fig. 1-5). A third fold was later identified in the structure of Na^+/H^+ exchanger NhaA (Hunte *et* al, 2005), and the ClC family of proton-coupled chloride exchangers represents a distinct fold from all three (Miller, 2006).

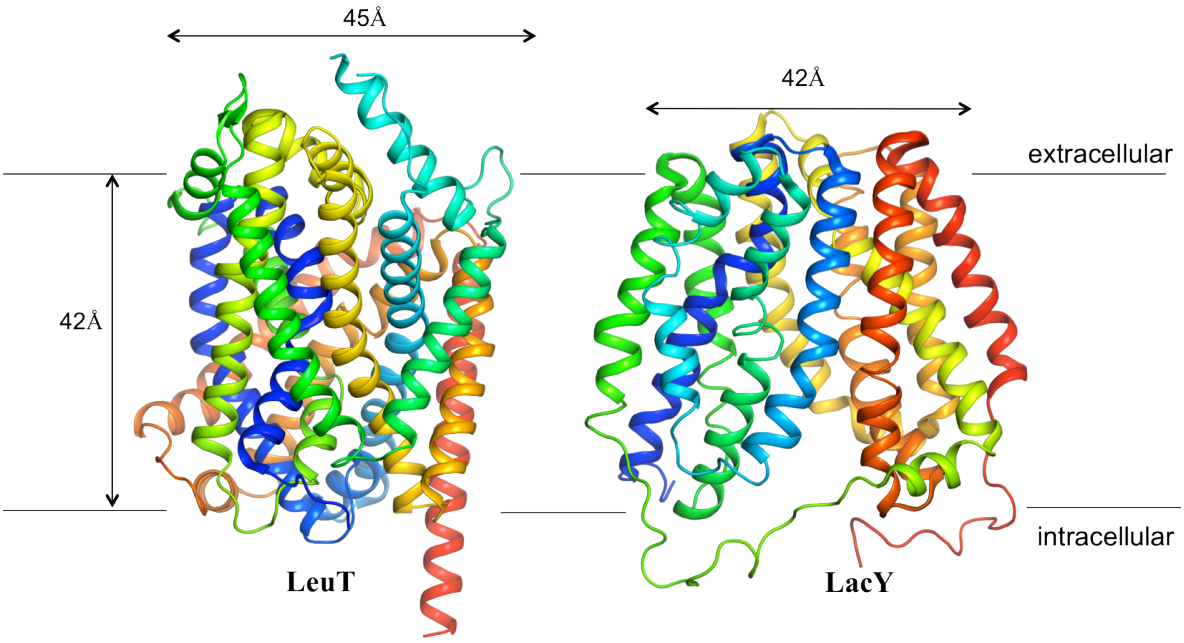


Fig. 1-5: Crystal structures of LeuT from *Aquifex aeolicus* (PDB:2A65) and LacY from *Escherichia coli* (PDB:2V8N), illustrating the two main structural folds of secondary active transporters.

The first published secondary active transporter structures, LacY and GlpT, revealed a striking common feature: the N-terminal domain (TMs 1-6) and the C-terminal domain (TMs 7-12) are related by a two-fold pseudosymmetry, such that the two can typically be superimposed with a low RMSD. This observation prompted a mechanistic model first formulated for GlpT (Huang *et al*, 2003), in which alternating access was tentatively explained through a rigid-body movement of the two domains, this “rocker-switch” motion

exposing the binding site to one side of the membrane at a time to allow passage of the substrate(s).

Later, the LeuT structure revealed the two-fold pseudosymmetry in the context of a different fold (TMs 1-5 symmetrical with respect to 6-10), and EmrD provided the third example of an MFS transporter potentially operating via the same mechanism. In all published structures, the two domains were symmetrical in relation to an axis perpendicular to the plane of the membrane, the central cavity being made up of side chains from both domains (Krishnamurthy *et al*, 2009). Upon further analysis of the available structural and functional data, the rocker-switch motion was proposed to provide alternating access to the binding site via the opening and closing of intra- and extracellular "gates" (Yamashita *et* al, 2005; Gouaux, 2009). Transporters would thus go through three distinct conformational states, as exemplified by the crystal structures: inward-open (LacY/GlpT), occluded (EmrD) and outward-open (LeuT), the three states completing the transport cycle (fig. 1-6).

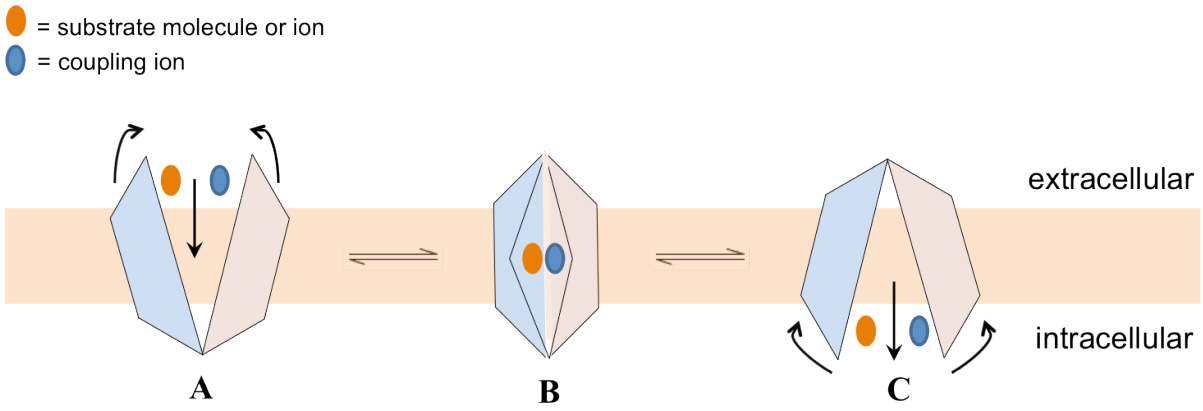


Fig. 1-6: Rocker-switch model for alternating access during coupled transport (cca. 2003), depicted here for a symporter. The three conformational states are (A) outward-open; (B) occluded; and (C) inward-open.

Additional structural data followed, providing more insight into the details of the alternating access mechanism. LeuT was crystallised in the outward- and inward-open states, completing the cycle of conformations (Krishnamurthy *et al*, 2012). Na^+-coupled transporters

from different families were found to adopt the LeuT fold: nucleobase transporter Mhp1, crystallised in all three conformations with and without substrate (Weyand *et al*, 2008; Shimamura *et al*, 2010); Na^+/galactose symporter vSGLT (Faham *et al*, 2008); Na^+/betaine symporter BetP (Ressl *et al*, 2009). Moreover, the LeuT core domain (pseudosymmetrical 5+5 TMs) is shared with sodium-independent transporters such as ApcT (Shaffer *et al*, 2009) and AdiC (Fang *et al*, 2009; Gao *et al*, 2010; Kowalczyk *et al*, 2011), both members of the APC family of amino acid transporters. Consequently, LeuT-like transporters from disparate families were brought together under the common umbrella of the APC superfamily, proposed to operate via a similar mechanism of alternating access (Abramson and Wright, 2009; Shi, 2013).

Besides the striking two-fold pseudosymmetry, transporters of the LeuT fold share what is perhaps a key feature to understanding the alternating access mechanism: inverted repeat topology. First observed in the LeuT structure, inverted repeats describe a different type of symmetry between the two 5-helix domains forming the 5+5 LeuT core. More precisely, each helix in the first N-terminal domain (TMs 1-5) is structurally related to its equivalent in the C-terminal domain (TMs 6-10) by a topological inversion, despite the lack of sequence conservation between the two (fig. 1-7, A). For example, helices 1 and 6 in the LeuT fold are both discontinuous (they break in the middle to form an intrahelical loop), mirroring each other's topologies in the membrane; helices 2 and 7 similarly share a structural feature and so on. This observation led to the hypothesis formulated by Lucy Forrest in 2008, according to which the structural features of a helix in the inward-open state are related to those of its inverted repeat in the outward-open state. The theory was tested, modeling the inward-facing structure of LeuT by "swapping" the inverted repeats in its outward-facing crystal structure (Forrest *et al*, 2008), the model being later confirmed by crystallographic data (Krishnamurthy *et al*, 2012).

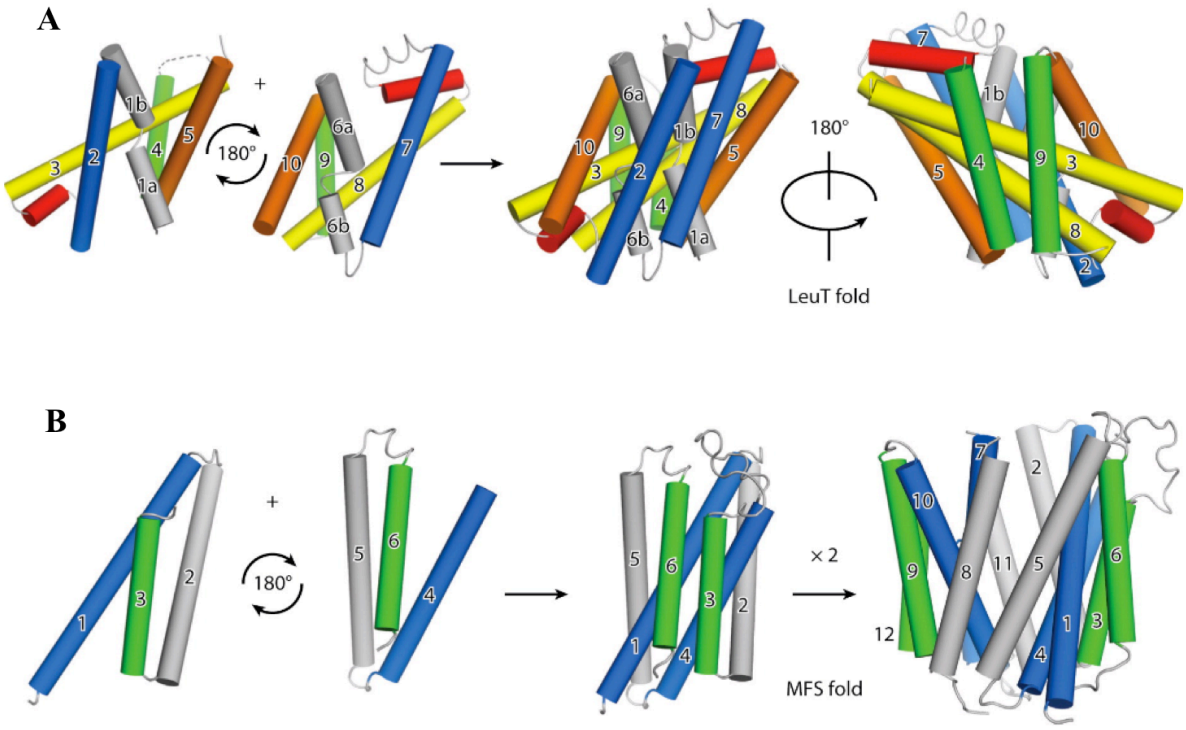


Fig. 1-7: Inverted repeat topology in secondary active transporters: the LeuT fold (A) and the MFS fold (B) consist of repeat subunits related by an inverted topological symmetry (Shi, 2013).

LeuT-like inverted repeats were identified in all APC superfamily members (Abramson & Wright, 2009), but also in transporters of different folds such as the MFS, NhaA and Glt_{Ph} (Forrest & Rudnick, 2009), proving to be a universal feature of secondary active transporters. MFS transporters have a 12TM core (6+6), with each 6-helix domain containing two inverted repeat elements (helices 1-3 and 4-6 in the N-terminal domain, 7-9 and 10-12 in the C-terminal). The 3-helix repeat elements assemble around an axis parallel to the plane of the membrane (fig. 1-7, B). Like LeuT, the NhaA fold features 5-helix repeats, but with additional discontinuous helices at positions 4 and 9 (Shi, 2013).

The inverted repeat topology model creates a framework for investigating coupled transport in a unified manner. It is consistent with the alternating access model, as the "repeat swapped models" always show opening or closure of the periplasmic/cytoplasmic cavities when compared to the crystal structure template in the opposite conformation (Forrest &

Rudnick, 2009; Radestock & Forrest, 2012). However, the model argues against rigid-body motion of the N- and C-terminal domains, instead suggesting a mechanism involving helical rearrangements between structurally related elements between and within the domains. Together with emerging structural data on different transporters, it contributes to a refinement of the alternating access model for coupled transport. Instead of three distinct conformations (outward-open, occluded and inward-open), the refined model includes a series of subtle structural changes initiated by the binding event(s), the transporter adopting transient, partially-occluded conformations in which intra- and extracellular gates open and close gradually before the release of the substrate(s) into the cytoplasm and recycling of the transporter back to the initial outward-open state (fig. 1-8).

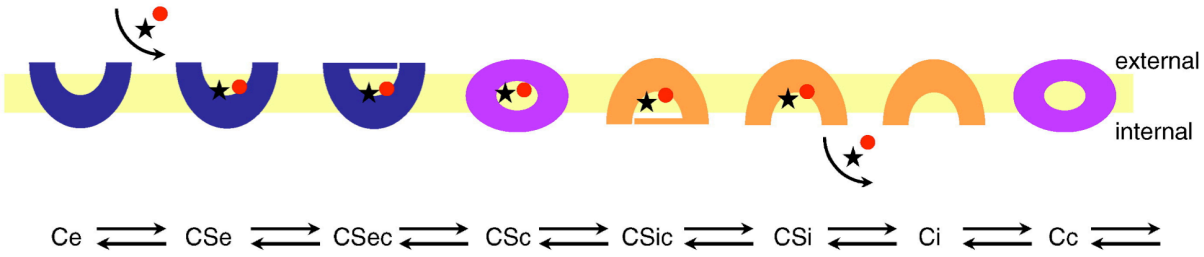


Fig. 1-8: The cycle of conformational states during coupled transport. Binding of substrates (S, star and circle) to the apo external-open carrier (Ce) causes a series of conformational changes: external-open (CSe); partially external-closed (CSec); fully closed/occluded (CSc); partially internal-closed (CSic); internal-open (CSi); apo internal-open (Ci). After the release of substrates, the carrier is recycled through a fully occluded apo state (Cc) and back to the outward-open state (Ce) (Forrest *et al*, 2010).

Inverted repeats are also found in the Na^+-coupled glutamate transporter Glt_{Ph}, believed to function not by typical alternating access, but by a "carrier transfer" mechanism, in which the binding site is translocated along with the substrate, as the transporter itself rotates in the membrane upon substrate binding, in agreement with an older hypothesis for solute transport across membranes (Widdas, 1952). Nevertheless, Glt_{Ph} displays inverted repeat symmetry like the other secondary transporters, and a repeat swapped model of its inward-open conformation was generated and confirmed by both biochemical data (Crisman

et al, 2009) and the eventual 3D structure obtained by crystallography (Reyes *et al*, 2009). Thus far, inverted repeat topology has proven a useful tool for investigating structural and mechanistic aspects across the spectrum of secondary active transporters.

1.3. The Major Facilitator Superfamily

Along with the ABC transporters, MFS proteins are ubiquitous among living organisms, accounting for over 50% of all transporters encoded by the bacterial genome, but being similarly abundant in archaea, yeast and higher eukaryotes (Ren and Paulsen, 1997; Pao *et al*, 1998). Highly diverse in terms of substrates, coupling mechanism and physiological function, MFS transporters have been further divided into 18 families based on similarities of sequence and predicted topology. Most are H^+-coupled, with only a minority being Na^+-dependent (Shi, 2013). Monosaccharides, peptides and nucleosides count among MFS substrates, as do metabolites involved in the Krebs cycle, nitrogen metabolism or energy generation in anaerobes. Many MFS proteins act as drug efflux pumps. Some families feature both symporters and antiporters, while others are restricted to one type of transport (Pao *et al*, 1998).

1.3.1. Common characteristics of MFS transporters

Even before any structural information was available, it was predicted and experimentally tested that members of the MFS have a core structure of 12 transmembrane-spanning regions,

thought to have evolved from a 6-helix unit, with two additional helices added to the 12TM core in some cases (Pao *et al*, 1998). In the light of the inverted repeat findings, more recent analyses using bioinformatics have proposed that the 3+3 intradomain repeat elements in the MFS may have arisen by internal gene duplication from smaller proteins (3 helices) that were able to adopt topologically inverted conformations with respect to the membrane. In support of this, the last three helices in MFS transporters (TMs 10-12) are structurally related to a trihelical voltage-gated Na^+ channel (Forrest and Rudnick, 2009). By forming the 6-helix fold, such proteins will have acquired the advantage of having a symmetrical pathway connecting their binding site to both sides of the membrane. Indeed, a handful of MFS members are predicted to have just one 6-helix bundle (Pao *et al*, 1998), the 12TM fold of the majority of MFS proteins likely the consequence of a later gene duplication event.

In addition to the overall architecture and inverted repeat organisation, major facilitators are recognised through the signature motif Gxxx(D/E)(R/K)xG(R/K)(R/K), found in the region between TMs 2 and 3 (fig. 1-9), present in even the more divergent members (Pao *et al*, 1998). The motif may play a role in the interaction with negatively charged lipidic head groups in the membrane (via the cationic side chains), and has been shown to be critical for the structure and function of the *E. coli* lactose permease (Jessen-Marshall *et al*, 1995).

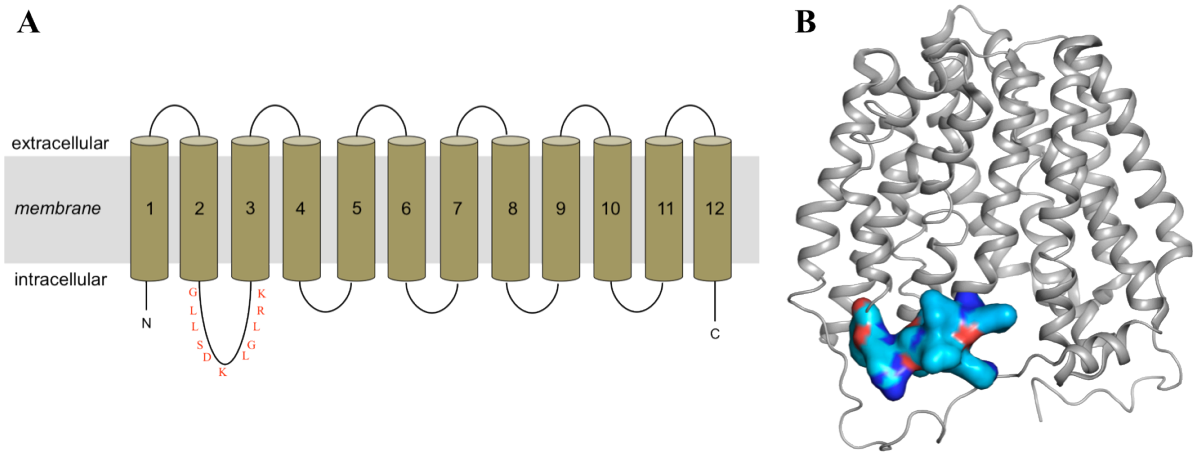


Fig. 1-9: Localisation of the MFS signature motif in LacY.
(A) Topology diagram with the sequence (GLLSDKLGLRK) printed in red between TMs 2 and 3.
(B) Surface representation of the motif mapped onto the LacY crystal structure (PDB:2V8N).

1.3.2. Early functional studies

The most well characterised MFS transporter is the H^+/galactoside symporter LacY, also known as the lactose permease. A part of the *lac* operon, the *E. coli* LacY gene was the first transporter gene to be cloned into a plasmid for overexpression (Teather *et al*, 1978), then purified in detergent micelles and functionally reconstituted into liposomes (Newman *et al*, 1981). Subsequent studies on liposome-reconstituted LacY have set the standard for assaying MFS transporter function *in vitro*.

It was found in the early 1980's that, although facilitated diffusion of lactose across the liposome membrane was detectable, transport was greatly energised by an artificially imposed electrochemical gradient, in the form of either a pH difference (ΔpH) or a "+ out/- in" electrical potential generated by controlled distribution of cations like K^+ and Na^+ in the buffers inside and outside the vesicles (ΔΨ). Electrical or pH gradients matching those encountered *in* vivo were generated in order to confirm that LacY was proton-coupled, as well as to assay substrate specificity and the kinetics of transport using radiolabeled lactose (Viitanen *et al*, 1984; Viitanen *et al*, 1986; Lolkema *et al*, 1991). Moreover, LacY was proven capable of transporting lactose down its concentration gradient, in the absence of the ΔΨ or ΔpH, as well as transporting lactose in both directions (into and out of the liposomes). These observations argued for the structural symmetry of the transporter and its dependence primarily on substrate availability, rather than a strict reliance on the electrochemical gradient (it does not transport protons in the absence of substrate for example).

Mutagenesis studies followed, and a putative binding site was identified through transport assays in liposomes and whole cells. Several, mostly conserved residues proved important for function, such as tyrosines and several charged residues involved in either

substrate binding, proton coupling or both (Roepe and Kaback, 1989; Frillingos and Kaback, 1996; He and Kaback, 1997). A transport mechanism was proposed for LacY, involving an initial, protonated ground state for the protein, which is disrupted upon substrate binding with high affinity at the periplasmic side, causing a conformational change that leads the transporter into the inward-open state, at the same time enabling translocation of the proton to the intracellular side where deprotonation lowers substrate affinity, allowing lactose to be released into the cytoplasm (Sahin-Tóth *et al*, 2000; Kaback *et al*, 2001). The crystal structures of LacY with and without bound substrate largely confirmed the postulated binding site interactions and validated the model (Abramson *et al*, 2003; Guan *et al*, 2007).

Similar *in vitro* studies were performed with lactose permeases from different organisms (Foucaud and Poolman, 1992; In't Veld *et al*, 1992; Knol *et al*, 1996), and other MFS transporters like GlpT (Fann *et al*, 2003), oxalate transporter OxlT (Yang *et al*, 2005), xylose transporter XylE (Sun *et al*, 2012) and more. Few eukaryotic MFS members have been successfully purified and reconstituted, a notable exception being the *A. thaliana* H^+/sugar symporter STP1, which was assayed for function in cytochrome C oxidase-energised vesicles (Stolz *et al*, 1994). In most cases, such as the human sugar transporters OAT1 (Perry *et al*, 2006) and Glut1 (Mueckler and Makepeace, 2009), homologues of GlpT and XylE respectively, or oligopeptide transporters PepT1 and PepT2, transport assays on eukaryotic MFS proteins have been confined to *in vivo* systems such as oocytes and mammalian cell lines.

1.3.3. Crystal structures

The structures of LacY and GlpT revealed the MFS fold and laid the foundation for subsequent work on related transporters. LacY and GlpT were captured in inward-facing conformations (GlpT fully open, LacY partially occluded with substrate bound), but it was with the structure of multidrug transporter EmrD that an MFS binding site was trapped in a fully occluded state, corresponding to the proposed intermediate state (Yin *et al*, 2006). The wild type LacY itself was then crystallised in a substrate-free, fully inward-open state (the first structure had been the conformationally constrained mutant Cys154Gly), prompting the assumption that the outward-open conformation of MFS proteins corresponds to a high energy state and is thus unfavourable for crystallisation (Guan *et al*, 2007).

The first outward-open MFS structure published was that of H^+/fucose symporter FucP (Dang *et al*, 2010). Additional information arrived with transporter structures from different families. Oligopeptide transporter $PepT_{So}$ was crystallised in the occluded conformation (Newstead *et al*, 2011), while the partially outward-open structure of *E. coli* xylose transporter XylE (Sun *et al*, 2012) completed the cycle of conformations (fig. 1-10).

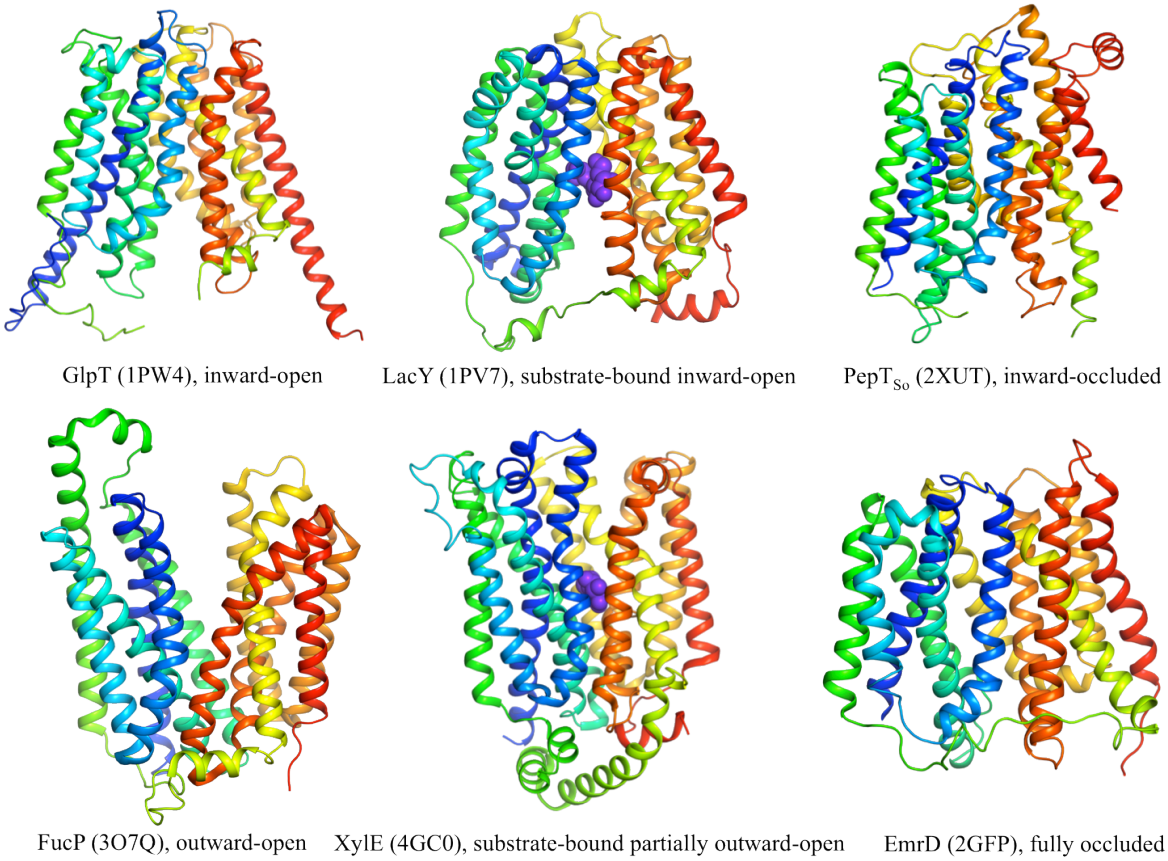


Fig. 1-10: Crystal structures of MFS transporters in six distinct conformational states, coloured blue to red from the N- to the C-terminus (PDB codes in brackets, substrates in purple).

Structures of XylE in the occluded and inward-open states were solved shortly after (Quistgaard *et al*, 2013), making XylE the first MFS member to be crystallised in all three main conformations of the transport cycle. Nitrate/nitrite antiporter NarU followed, its structure solved in both the occluded and inward-facing states (Yan *et al*, 2013). The inward-open structure of $PepT_{St}$ added another conformer to the oligopeptide transporter family (Solcan *et al*, 2012). Furthermore, the first structure of a eukaryotic MFS transporter arrived in 2013 with fungal H^+/phosphate transporter PipT in an inward-open, partially occluded substrate-bound form (Pedersen *et al*, 2013). Finally, 2013 also brought the first MFS structure solved by *in meso* crystallography: inward-open peptide transporter GkPOT (Doki *et al*, 2013). In total, crystal structures of 10 membrane transporters from different MFS families have been published to date (table 1-2). All of them share the canonical MFS fold, and showcase the various conformations adopted along the transport cycle.

Protein name	Organism	Direction of transport	Substrates	Conformation(s)
LacY	*Escherichia coli*	symporter	monosaccharides, H^+	inward-open (substrate-bound) (1PV7); inward-open (2V8N)
GlpT	*Escherichia coli*	antiporter	glycerol-3-phosphate, inorganic phosphate	inward-open (1PW4)
EmrD	*Escherichia coli*	antiporter	drugs, H^+	occluded (2GFP)
$PepT_{So}$	*Shewanella oneidensis*	symporter	di/tripeptides, H^+	occluded inward-facing (2XUT)
FucP	*Escherichia coli*	symporter	fucose, H^+	outward-open (3O7Q)
$PepT_{St}$	*Streptococcus thermophilus*	symporter	di/tripeptides, H^+	inward-open (2APS)
XylE	*Escherichia coli*	symporter	monosaccharides/H^+	partially outward-open (substrate-bound) (4GC0); partially inward-open (4JA3); fully inward-open (4JA4)
PipT	*Piriformospora indica (eukaryotic)*	symporter	phosphate/H^+	inward-facing occluded (substrate-bound) (4J05)
NarU	*Escherichia coli*	antiporter	nitrate/nitrite	occluded and partially inward-open (substrate-bound) (4IU9)
GkPOT	*Geobacillus kaustophilus*	symporter	di/tripeptides/H^+	inward-open (with and without substrate) (41KV; 4IKX; 4IKW; 4IKY; 4IKZ)
$PepT_{So2}$	*Shewanella oneidensis*	symporter	Di/tripeptides/H^+	inward-open (substrate-bound) (4LEP)

Table 1-2: Conformational states of MFS transporters as described by X-ray crystal structures (PDB codes for each structure in brackets).

1.3.4. Features of MFS binding sites

With the wide variety of transport substrates among major facilitators, one would expect the binding sites of these proteins to be equally diverse. Nevertheless, some common features can be identified. The most important of these is the presence of aromatic amino acids – a common feature of secondary active transporters in general, as many of them transport hydrophobic molecules, easily accommodated by π-π stacking interactions (Shi, 2013). The role of aromatic residues in sugar transport was proposed early on when essential tyrosines in LacY were identified by site-directed mutagenesis (Roepe and Kaback, 1989). Later, a model of the binding site with bound disaccharide showed a tryptophan stacking against the galactopyranosyl ring of the substrate, with additional hydrophobic interactions from neighbouring side chains (Guan *et al*, 2003).

The crystal structure of LacY with bound TDG (β-D-galactopyranosyl-1-thio-β-D-galactopyranoside) (Abramson *et al*, 2003) revealed the binding site was located centrally, shared between residues from both helical bundles (1, 4 and 5 from the N-terminal bundle, 7 and 11 from the C-terminal), a feature found in all MFS structures to date, consistent with the alternating access model. The structure also supported the earlier model, with Trp151 at the centre of the binding site, its indole side chain stacked against the galactoside ring of TDG (fig. 1-11). The position of the tryptophan is further supported by a cation-π interaction with adjacent Arg144, as is frequently encountered in the binding sites of proteins requiring stabilisation of a particular secondary structure element (Dougherty, 2007). The arginine also binds the substrate on the N-terminal side, being held in place by a hydrogen bond to Glu269; the opposite end of the substrate is similarly coordinated by Lys358 and Asp237. The residues,

together with Glu126, Asp240 and Gln359, form a polar surface to accommodate the polar substrate.

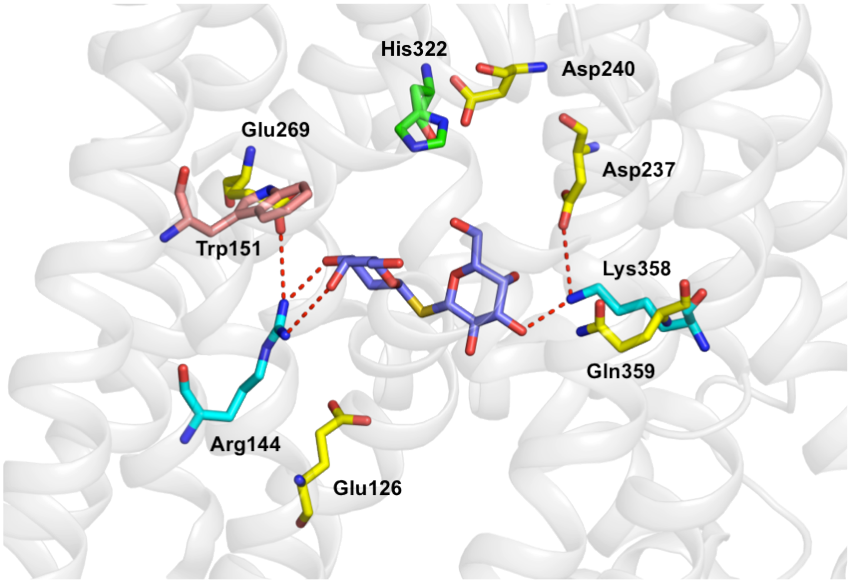


Fig. 1-11: LacY binding site in the substrate-bound, inward-facing conformation (PDB:1PV7). The disaccharide TDG (centre) is accommodated via interactions with polar and hydrophobic side chains, typical of MFS binding sites.

Binding sites of other sugar transporters show similar traits. The central cavity of FucP is shared between helices 1 and 4 (N-terminal) and 7 and 10 (C-terminal), with bulky tryptophan and tyrosine side chains positioning the fucose ring and charged residues on both sides (Dang *et al*, 2010). Likewise, XylE uses polar and aromatic groups to coordinate D-xylose and prevent it from exiting the binding site respectively; the roles of each residue are confirmed by transport assays (Sun *et al*, 2012).

Overall, a combination of aromatic/hydrophobic residues (for positioning the substrate) and polar/charged side chains (for more specific interactions) seems to be the key to an efficient binding site for sugar transporters, but this also applies to other MFS proteins with polar and/or charged substrates. The NarU structure reveals a layer of hydrophobic residues closing the periplasmic gate, and aromatic groups sealing the transporter at the cytoplasmic side, with polar residues at the centre of the cavity for binding nitrate/nitrite (Yan *et al*, 2013).

In PipT, phosphate is coordinated by a conserved group of residues with polar and aromatic side chains (Pedersen *et al*, 2013). GlpT employs positively charged side chains to increase the pK_a of a histidine, which in turn binds the oxyanionic glycerol-3-phosphate, while the complex is stabilised by tyrosines and salt bridge contacts (Law *et al*, 2008a). GlpT-related oxalate transporter OxlT, of which the structure had been solved by EM crystallography before the other MFS proteins (Hirai *et al*, 2002), requires a similar positioning of cationic side chains to preserve the basicity of the binding site (Yang *et al*, 2005).

Not surprisingly, proteins with highly hydrophobic substrates like the promiscuous multidrug resistance transporters, feature a modified binding site. EmrD is an example, its central cavity occupied almost entirely by aromatic (Tyr, Trp, Phe) and generally hydrophobic residues (Val, Leu, Ile), stacking against each other. Most of these residues are conserved among MDR transporters and serve to efficiently accommodate aromatic drugs (Yin *et al*, 2006).

1.3.5. Alternating access mechanism in the MFS

Early studies on LacY using mutagenesis and *in vitro* assays suggested that interactions between certain amino acids mediate the conformational changes required for alternating access to the binding site during the transport cycle. Specifically, Glu269 and His322 were predicted to interact and stabilise the protonated form of the transporter (He and Kaback, 1997). Later, Glu325 and Arg302 were identified as H^+ coupling residues along with the histidine, and a transport mechanism was proposed (Sahin-Tóth *et al*, 2000). The first LacY crystal structure revealed a putative pathway for proton translocation and coupling. Trapped in the inward-open state, the transporter appeared to be protonated at the carboxyl group of

Glu325 at the intracellular side (as there were no hydrogen bond donors in its vicinity), its side chain stabilised by surrounding aromatic groups (Abramson *et al*, 2003). Data from structure and simulations were corroborated to advance a model for transport, in which helix movements and salt bridge rearrangements between the glutamates, arginine and histidine orchestrated alternating access. The model was refined as more experimental data became available.

Comparing reactivity of cysteine side chains at the periplasmic and intracellular sides of wild type LacY in site-directed alkylation led to the observation that outward-facing and occluded conformations were favoured upon TDG binding, whereas the apo state reverted to the low energy, inward-facing state seen in the crystal structure (Kaback *et al*, 2007). A similar study found that LacY adopted all three conformations in membranes, the outward-open one being naturally favoured (Jiang *et al*, 2012). Single molecule FRET measurements yielded similar findings (Majumdar *et al*, 2007). Double electron-electron resonance (DEER) experiments showed the same three main populations of conformers in detergent solution, shifting towards the outward-facing state upon glucose binding (Smirnova *et al*, 2007).

Ultimately, the existing structural and functional data converged towards a detailed model of the transport cycle (fig. 1-12). In it, the transporter in its low energy, inward-open state (green) is deprotonated at Glu325 at the intracellular side (orange), transitioning to an occluded, substrate-free state. Reprotonation at the other side is immediate and opens the periplasmic cavity (blue). Substrate (S) binds with high affinity to this low energy state (grey), causing a conformational change that transfers the proton from the Glu269-His322 pair (periplasmic side) back to Glu325 (intracellular side), and induces the substrate-bound occluded state. Finally, the cytoplasmic cavity opens and the substrate is released, while the protonated transporter is back to the inward-open conformation. According to the model and supporting experimental data, both inward- and outward-open states are energetically

favoured in the membrane, while occluded states are unfavourable but can be induced by initial deprotonation (apo occluded) or substrate binding (substrate-bound occluded) (Madej *et al*, 2012). Initial deprotonation at the intracellular side can be a rate-limiting step in the absence of a transmembrane electrochemical gradient, while in the presence of the gradient the bottleneck is the release of the substrate into the cytosol against its concentration gradient (Guan and Kaback, 2006).

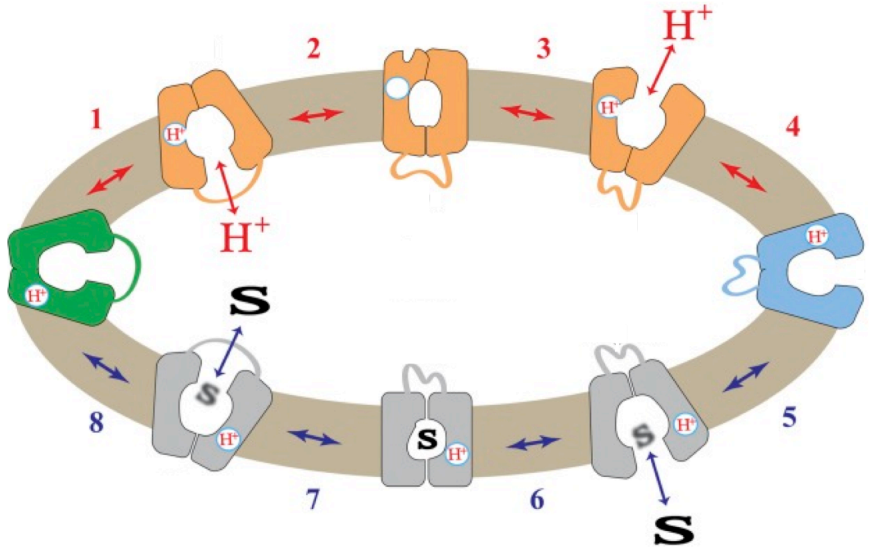


Fig. 1-12: The LacY transport cycle: overview of the proposed steps in the transport model (Madej *et al*, 2012).

Proton and substrate translocation within the binding site are necessarily accompanied by helical rearrangements. Early computer simulations of the outward-open state of LacY based on the inward-open structure postulated the opening of the periplasmic gate through the movement of TMs 1 and 7, synchronised with rearrangements in the cytoplasmic halves of TMs 2, 4, 5 (N-terminal) and 8, 10, 11 (C-terminal) (Abramson *et al*, 2003). A more recent outward-open LacY model, based on swapping inverted topology repeats, confirmed these movements, as well as the predicted salt bridges associated with them, such as the one between Glu325 and Arg302 at the intracellular side, sealing the cytoplasmic cavity after the

initial deprotonation event (Radestock and Forrest, 2011). The swapped model also outlines an important kink in TM10 topologically matching the one occurring in TM4 at a conserved proline residue, suggesting that rearrangement of the helices relies on internal topological symmetry relations as well as sequence conservation.

Similar helical movements and salt bridge reconfigurations have been suggested for the other MFS proteins. Movement of TMs 1 and 7 is believed to open the periplasmic gate in GlpT (Huang *et al*, 2003), while the main H^+-coupling residue (a histidine) coordinates interhelical salt bridges acting as springs providing alternating access to the binding site (Law *et al*, 2008b). Three salt bridges between the N- and C-terminal domains also form the basis for alternating access in XylE, while helices 5 and 10 move towards and away from the central cavity in synchronised fashion (Sun *et al*, 2012). A similar proton translocation path to that of LacY was identified in FucP, in which the protonation of an aspartate side chain at the periplasmic side enables fucose to bind, which in turn causes a conformational change and transfer of the proton to a glutamate as the protein goes into the inward-open state (Dang *et al*, 2010). Finally, an aspartate is also identified as the protonation site in PipT, its protonation thought to make phosphate binding favourable, inducing a structural rearrangement that opens a cytosolic tunnel (Pedersen *et al*, 2013). Common threads therefore appear to run through the current models of MFS transport, as increasing data from crystallography, transport assays, modeling and simulations provide a better understanding of the alternating access mechanism.

1.4. Peptide transport

Over the course of evolution, multicellular organisation has brought with it increasingly complex ways of adapting to environmental conditions and utilising natural resources. A good example of such adaptations in higher eukaryotes is the digestive tract of animals, (im)perfectly fitted to ingest and digest nutrients, to break them down into molecular building blocks ready to fuel the ceaseless metabolic undertakings of the body.

One aspect of digestion involves the metabolisation of dietary proteins, which begins in the oral cavity with the action of salivary enzymes. Proteins are then broken down in the stomach by proteases and peptidases, resulting in peptides and individual amino acids. The process continues in the upper intestine, where specialised transporters take up the small molecules from the intestinal lumen into the epithelial cells of the brush border membrane, to be further metabolised and finally absorbed into the blood stream. It is here that the peptide transport story begins.

1.4.1. Classification and conservation of peptide transporters

Peptide transport across biological membranes occurs in all living organisms. Several transporter families are responsible for the process. Prokaryotes have the greatest diversity, with a total of three peptide transport systems (Daniel *et al*, 2006). The first two of these, the Dpp and Opp families, are ABC transporters. Dpp stands for dipeptide permeases (they

transport dipeptides exclusively), while the Opp proteins (oligopeptide permeases) transport peptides between 4 and 18 residues in length (Detmers *et al*, 1998). The third transport system has been studied extensively, and is the only one conserved across phyla, from prokaryotes to humans: the proton-coupled oligopeptide transporter (POT) family, also known in literature as the PTR (peptide transporter) family (Daniel *et al*, 2006). As the name suggests, they are secondary active transporters that use the transmembrane proton gradient to drive the uptake of small peptides, more specifically di- and tripeptides.

A smaller family – the OPT transporters – are found in yeast, as well as bacteria, archaea, plants and some invertebrates. The OPT are also H^+-coupled symporters, but bear no sequence similarity to the POTs and seem to transport a wider variety of substrates such as peptides, glutathione and metabolites involved in regulating iron homeostasis (Lubkowitz, 2011). Higher eukaryotes however rely exclusively on the POT system for transporting peptides. In plants, they are described as the PTR family and play a major role in nitrogen mobilisation (Steiner *et al*, 1994); they are well represented in the plant world, counting up to 53 isoforms in *Arabidopsis thaliana* alone (Nour-Eldin *et al*, 2012). In contrast, vertebrates have fewer POT isoforms, which in humans are grouped into the SLC15 family, named after the corresponding gene nomenclature (Daniel and Kottra, 2004).

Despite these differences, POT family members across kingdoms show a high degree of sequence similarity, suggesting conservation of structure and function (Steiner *et al*, 1995; Daniel *et al*, 2006). Their status as MFS transporters has been debated due to their divergence from the sequences of other, established MFS families (Pao *et al*, 1998), but was recently confirmed when the structure of POT family member $PepT_{So}$ displayed the canonical MFS fold (Newstead *et al*, 2011).

1.4.2. The SLC15 family: PepT1 and PepT2

The human SLC15 family has four members: PepT1 (SLC15A1), PepT2 (SLC15A2), PHT1 (SLC15A4) and PHT2 (SLC15A3). The latter two are found in lysosomal membranes, while PepT1 and PepT2 are located in the plasma membrane. PHT1 and PHT2 proteins transport some di- and tripeptides along with free histidine, and they are more similar from the phylogenetic point of view to plant peptide and nitrate transporters, whereas sequence similarity to the PepT proteins is 20-25% (Daniel and Kottra, 2004). While little is known about the PHT system, PepT1 and PepT2 have been the objects of extended biochemical research over the past decades. Although their structures remain unsolved, the predicted topology is typical of MFS transporters: both N- and C-termini in the cytoplasm, with 12 transmembrane segments and an extracellular domain (of yet unknown function) between TMs 9 and 10 (Liang *et al*, 1995). Both are symporters driven by an inward proton gradient (fig. 1-13). PepT1 and PepT2 share 50% sequence identity and 70% similarity in humans (Daniel, 2004) as well as across species (Döring *et al*, 2002). Sequence identity between the same isoform in different species can be very high, such as 81% between the human and rabbit PepT1 (Mackenzie *et al*, 1996). Thus, the two isoforms are both closely related and highly conserved.

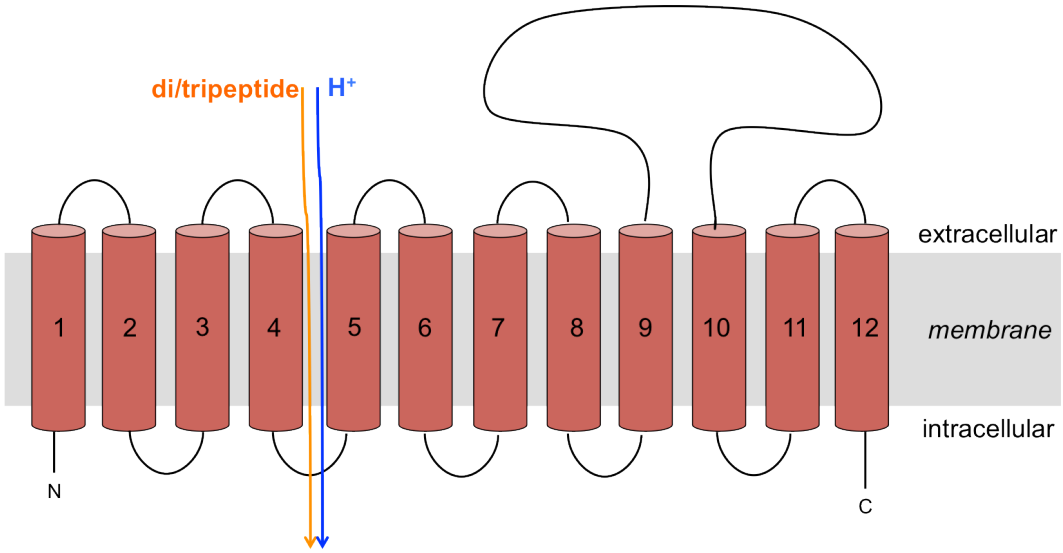


Fig. 1-13: PepT1/PepT2 topology diagram.

Intestinal absorption of dietary protein was long thought to occur exclusively in the form of single amino acids, until experiments with brush border membrane extracts in the 1970's and 1980's showed high intracellular peptidase activity specifically targeted towards certain di- and tripeptides, indicating that some such oligopeptides were not hydrolysed in the lumen but transported into cells directly (Webb *et al*, 1992). More specifically, dipeptide transport was first observed in the human intestine (Hellier *et al*, 1972), and later a similar process was discovered in rat kidneys, as dipeptides not hydrolysed in the plasma were retained in the proximal renal tubules (Adibi, 1977). Direct transport of oligopeptides and other molecules was later measured in isolated brush border membranes from kidney (Ganapathy *et al*, 1981) and intestine (Kato *et al*, 1989).

With the advent of recombinant DNA technology, the first mammalian peptide transporters were cloned using cDNA from rabbit enterocytes (Fei *et al*, 1994) and human kidney cells (Liu *et al*, 1995). The intestinal isoform was named PepT1, and the renal isoform PepT2 (fig. 1-14). PepT1 is also expressed in the kidney (Saito *et al*, 1995), but the two proteins differ in their localisation: PepT1 is found at the apical end of brush border cells (in

the kidney as well as the intestine), while PepT2 is localised mainly in the outer medulla of the proximal renal tubule (Shen *et al*, 1999). PepT2 mRNA has also been identified in epithelial tissues of lung alveolae (Groneberg *et al*, 2001), mammary gland (Groneberg *et al*, 2002) and in cardiomyocytes (Lin and King, 2007). In the nervous system, PepT2 is expressed in astrocytes, where it is believed to play a role in glutathione metabolism by taking up dipeptide precursor CysGly (Dringen *et al*, 1998), and in the choroid plexus (blood-brain barrier), where it may serve to release peptides from the brain into the cerebrospinal fluid (Shu *et al*, 2002).

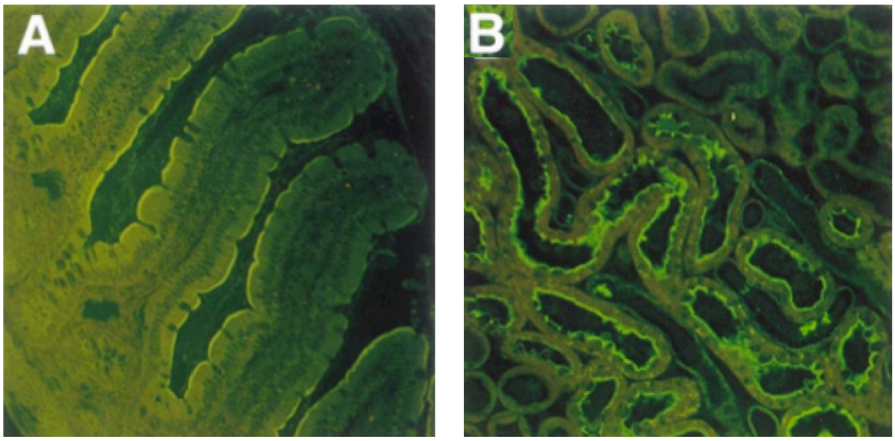


Fig. 1-14: Immunofluorescence labeling in rat cell cultures identifies peptide transporters in brush border membranes: PepT1 in the jejunum (A) and PepT2 in the proximal renal tubule (B) (Shen *et al*, 1999).

The coupling mechanism for di/tripeptide transport was also long debated. Sodium coupling was first advanced, as amino acid transport was known to be sodium-driven (Daniel, 2004). Despite initial inconclusive data, proton coupling was demonstrated in the 1980's in both intestinal and renal tissues, and the effects of both the pH gradient and the membrane potential were outlined (Ganapathy and Leibach, 1985; 1986). Generally, extracellular pH influences transport, unless substrate is available at saturating concentrations, in which case the membrane potential becomes the main driving force (Daniel and Kottra, 2004). The current model for di/tripeptide transport involves the coordinated activity of several

transporters and channels, working to maintain physiological ion gradients while distributing nutrients to the cells. For example, peptide transport at the apical side of enterocytes depends on the activity of the Na^+/H^+ antiporter NHE-3, which counteracts the intracellular acidification caused by proton influx, and restores the transmembrane proton gradient (Kennedy *et al*, 2002). In parallel, the Na^+/K^+ ATPase pumps out excess Na^+ ions at the basolateral end, taking up K^+ ions that can then exit the cell through K^+ channels. Meanwhile, cytosolic hydrolases break down the peptides, and the resulting amino acids are transported into the blood stream via (to date unidentified) amino acid transporters (fig. 1-15).

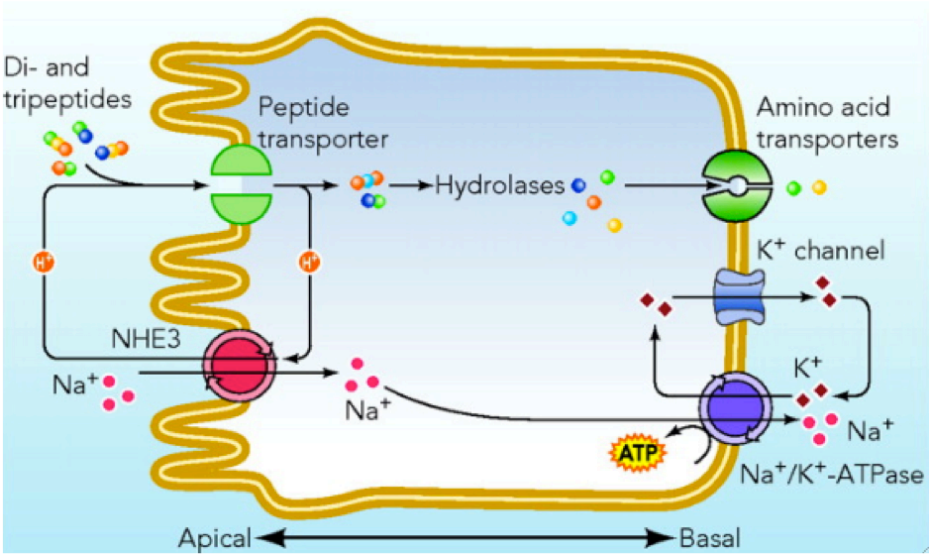


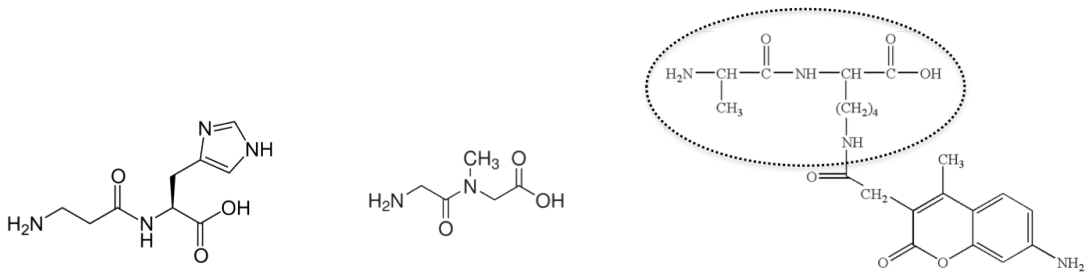
Fig. 1-15: Schematic representation of a eukaryotic intestinal cell. At the apical pole, the peptide transporter (PepT1) is functionally coupled to the Na^+/H^+ antiporter NHE3; at the basolateral pole, a similar feedback loop connects K^+ channels and the Na^+/K^+ ATPase (Daniel *et al*, 2006).

1.4.3. Substrate specificity of PepT1 and PepT2

The natural substrates of PepT1 and PepT2 are virtually identical, consisting of peptides made up of seemingly any combination of two and three amino acids. However, functional studies have revealed PepT1 to be the "high capacity, low affinity" isoform, with PepT2 being "high

affinity, low capacity", based on comparing the relations between their individual kinetic parameters such as the K_M of transport (affinity) and reaction rate/V_{max} (capacity), as measured for different substrates (Daniel and Kottra, 2004). Patch clamp measurements in oocytes have shown that peptide transporters, like LacY, transport substrates both into and out of cells at the same rate (but with lower affinity at the cytoplasmic side), and can do so in the absence of an electrochemical gradient (Kottra and Daniel, 2001).

It was established in early studies that single amino acids as well as tetra- and pentapeptides are not substrates (Ganapathy *et al*, 1981; Daniel *et al*, 1992), and D-isomers of dipeptides like D-Ala-D-Ala and D-Leu-D-Leu are also not transported (Thwaites *et al*, 1994; Lin and King, 2007). However, a D-amino acid can be accepted at the N-terminus, as in the case of substrates D-Phe-L-Gln and D-Phe-L-Ala (Meredith *et al*, 2000; Theis *et al*, 2002a). Non-hydrolysable peptide analogue glycylsarcosine (GlySar) (fig. 1-16) is typically used in its radiolabeled form for assaying PepT1 and PepT2 activity in apical intestinal (Caco-2) or basolateral renal (SKPT/MDCK) cell lines, as well as overexpressed in *Xenopus laevis* oocytes – either for direct uptake measurements or as a reporter in competition assays with other substrates (Thwaites *et al*, 1994; Liu *et al*, 1995; Terada *et al*, 2000; Shu *et al*, 2001). Other reporter substrates are fluorescent peptide derivative β-Ala-Lys-N_ε-AMCA (Dieck *et al*, 1999) and non-hydrolysable dipeptide L-4,4'-biphenylalanyl-L-proline (BipPro) (Knütter *et al*, 2007). Using these substrates allows for accurate quantification of uptake levels, as physiological di/tripeptides would be metabolised in the cell and become substrates for other transporters, potentially creating false negatives.



L-Ala-L-His (carnosine) glycylsarcosine (GlySar) beta-L-Ala-L-Lys-7-amino-4-methylcoumarin-3-acetic acid (beta-Ala-Lys-AMCA)

Fig. 1-16: Chemical structures of three hPepT1 substrates: dipeptide carnosine, peptide analogue GlySar and fluorescent peptide beta-Ala-Lys-AMCA (peptide component encircled).

Although all tested di- and tripeptides were transported, affinities vary greatly (from low micromolar to milimolar). Overall, both transporters have a preference for peptides with neutral and/or hydrophobic side chains (Eddy *et al*, 1995; Biegel *et al*, 2006). Affinity for charged peptides is enhanced by the presence of the N-terminal amino group or that of a backbone carbonyl at the N-terminal residue, and diminished by having a basic side chain (lysine, arginine) at the C-terminal position (Kottra *et al*, 2002; Theis *et* al, 2002b; Biegel *et al*, 2006). The binding site is thus believed to be asymmetrical and accommodate substrates stereospecifically. However, none of the intrinsic features of di/tripeptides are essential for transport, including the peptide bond and the terminal amino and carboxyl groups – although the absence of both termini is not competible with transport (Daniel and Kottra, 2004).

PepT1 and PepT2 display higher affinity for anionic peptides at lower pH than high pH, making it likely that they are transported in their protonated form (the opposite is true for cationic peptides). Indeed, inward electrical current measurements in oocytes have suggested that stoichiometry of PepT1-mediated transport is 1:1 (proton:peptide) for neutral and cationic peptides, but 2:1 for anionic ones – one proton required for coupling and one to neutralise the negative charge on the peptide (Steel *et al*, 1997; Kottra *et al*, 2002). In contrast, PepT2 exhibits a stoichiometry of 2:1 (neutral and cationic) and 3:1 (anionic), one additional proton translocated per cycle (Chen *et al*, 1999). Consistent with the mechanisms proposed for other

MFS members, electrophysiology data on both proteins indicate that protonation of the transporter precedes substrate binding (Mackenzie *et al*, 1996; Chen *et al*, 1999).

The structural origin of these differences between PepT1 and PepT2 is unclear. However, experiments using chimeric constructs have shed some light on the issue. Fusing the first six transmembrane segments in PepT2 with the last six in PepT1 (fig. 1-17) results in a functional transporter that retains all the main characteristics of PepT2 (substrate specificity, pH dependence, electrophysiological parameters), suggesting these traits are determined by essential residues found in the N-terminal domain – as well as pointing towards a different function for the extracellular domain between TMs 9 and 10, likely not involved in transport (Döring *et al*, 1996). Later studies using different PepT1/PepT2 chimeras expanded upon this initial observation and suggested the region responsible for substrate binding and transport includes residues from TMs 1, 4, 5 (N-terminal) but also 7-9 (C-terminal) (Döring *et al*, 2002). This is largely in agreement with the localisation of binding sites in other MFS transporters, as well as mutagenesis studies and homology models of PepT1 and PepT2 based on the LacY and GlpT structures (Meredith and Price, 2006; Pedretti *et al*, 2008; 2011).

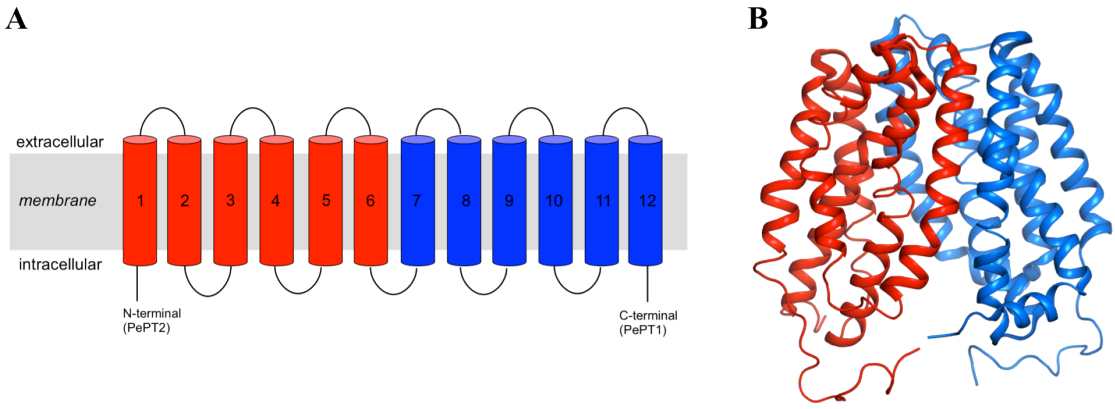
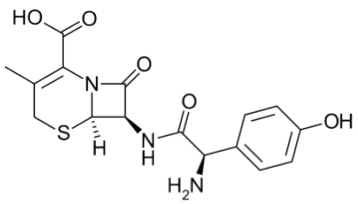


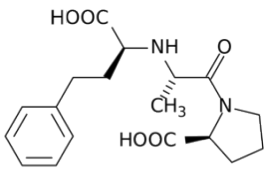
Fig. 1-17: Chimeric construct of PepT1 and PepT2.
(A) Topology diagram of the chimeric protein, containing the N- and C-terminal domains from PepT2 and PepT1 respectively.
(B) The equivalent two domains in the LacY crystal structure (PDB:2V8N).

1.4.4. Peptide transporters as drug targets

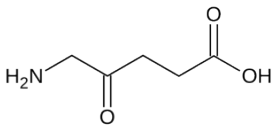
There is currently no data that links peptide transport to disease. Peptide transporters are known to compensate for lacking amino acid transport in cystinuric patients (Asatoor *et al*, 1970), and hPepT1 expression is upregulated in the colon of patients with short bowel syndrome, to make up for poor oligopeptide absorption in the small intestine (Ziegler *et al*, 2002). PepT1 expression also increases after fasting periods (Thamotharan *et al*, 1999). However, a recent study on PepT1-deficient mice found neither a pathological phenotype, nor any upregulation of amino acid transport, suggesting that PepT1 is not essential (Nässl *et al*, 2011). Comparably, PepT2 knockout mice were not affected despite significantly reduced peptide absorption in the kidneys (Rubio-Aliaga *et al*, 2003). Nevertheless, the clinical relevance of peptide transporters, like that of many transporter proteins (Giacomini *et al*, 2010), lies in their ability to recognise and transport a wide range of drugs. β-lactam antibiotics, antiviral compounds and imaging substances count among peptide transporter substrates (fig. 1-18).



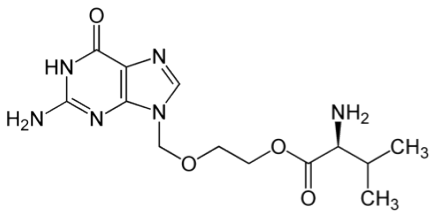
cefadroxil



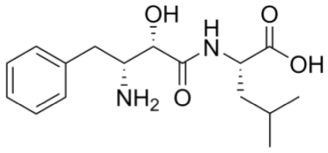
enalapril



5-aminolevulinic acid



valacyclovir



bestatin

Fig. 1-18: Molecular structures of selected pharmacologically important compounds transported by PepT1.

Cephalosporin β-lactams were the first substrates identified, in the late 1980's and early 1990's. Evidence of cefalexin, cefadroxil and cephradine transport in intestinal epithelia (Kato *et al*, 1989; Thwaites *et al*, 1994) was confirmed by measuring H^+coupled transport in PepT1-overexpressing oocytes (Fei *et al*, 1994; Boll *et al*, 1994). Ceftibuten and cyclacillin were added to the list as high affinity PepT1 substrates (Saito *et al*, 1995; Brentschneider *et al*, 1999), while penicillin derivatives (ampicillin and benzylpenicillin/penicillin G) were found to competitively inhibit dipeptide transport (Eddy *et al*, 1995). A later study confirmed that zwitterionic β-lactams (cefadroxil, cefalexin etc.) are preferred, but anionic ones (benzypenicillin, cefixime) become better substrates at low pH (Wenzel *et al*, 1996). PepT2 transports all of these as well; however, marked differences exist between PepT2 and PepT1 affinities for particular antibiotics (Ganapathy *et al*, 1995; Dantzig, 1997). For example, it has been proposed that PepT1 recognises cephalosporins with a stereochemistry matching the L-isomer of tripeptides (Bailey *et al*, 2000), while PepT2 prefers those with an amino group at the alpha position (Luckner and Brandsch, 2005). Based on such considerations, computational algorithms can be used to predict other putative substrates and inhibitors of peptide transporters (Ekins *et al*, 2005).

In addition to β-lactam antibiotics, peptide transporter substrates include antiviral drug valacyclovir (Faria *et al*, 2004), ACE inhibitors enalapril and fosinopril (Zhu *et al*, 2000; Shu *et al*, 2001), tumour supressor bestatin (Tomita *et al*, 1990; Nakanishi *et al*, 2000), and δ-aminolevulinic acid (imaging substance used in photodynamic therapy) (Döring *et al*, 1998; Neumann *et al*, 2003). Furthermore, GlySar itself has been successfully used as a PET tracer in mice, as it accumulates in large quantities in tumour xenografts (Mitsuoka *et al*, 2008). More generally, oligopeptide and drug transport could be used as a means to distinguish between tumour and normal cells, as the extracellular pH of tumours is lower than normal,

creating more prominent electrochemical gradients that favour H^+-coupled transport of solutes (Anderson *et al*, 2010).

For many of these compounds, PepT1 is the main entry route into the organism, while PepT2 is believed to be the sole point of re-entry for plasma-borne drug molecules after initial filtering in the kidney, contributing to better overall retention (Takahashi *et al*, 1998). Therefore, efforts are being made to understand how known drug substrates are recognised, and use that information to synthesise new ones. The prodrug approach (fusing active core drugs to amino acids to create a peptide-like substrate) can be particularly useful, as illustrated by the synthesis of L-DOPA-L-Phe (targeting PepT1) to improve intestinal absorption of Parkinson's disease therapeutic drug L-DOPA (Tamai *et al*, 1998). Designing molecules to specifically target peptide transporters is a challenging task in the absence of a structure, but attempts at pharmacophore mapping have been made using homology models (Pedretti *et al*, 2008; 2011).

1.4.5. Peptide transport in prokaryotes

As mentioned in the beginning of the chapter, POT proteins are conserved in all phyla. Remarkably, pro- and eukaryotic POT family members exhibit highly conserved sequences including the universally conserved family signature motifs: ExxERFxYY (at the periplasmic end of TM1) and FYxxINxG, also referred to as the PTR2_2 motif (in TM5). A third motif (PTR2_2), GxxxADxxxGKxxTI (between TMs 2 and 3), is less conserved, but mutations in all three regions result in loss of function (Daniel *et al*, 2006). The latter motif is a variation of the MFS signature sequence Gxxx(D/E)(R/K)xG(R/K)(R/K) (fig. 1-9), albeit allowing for

greater variability compared to most other MFS families (Pao *et al*, 1998). The high degree of sequence conservation suggests a similar topology and transport mechanism for POT members across kingdoms. Notable differences that set bacterial POTs apart are the absence of the extracellular domain between TMs 9 and 10, and the presence of two additional transmembrane segments in some proteins (to a total of 14), the latter believed to have evolved from 12TM proteins through the membrane insertion of a central loop (Saier, 2003).

The first bacterial POT was identified in *Salmonella typhimurium*, where the *tppB* (tripeptide permease) gene locus was shown to encode a transporter protein with features similar to those of the mammalian PepT system (Jamieson and Higgins, 1984). Homologous gene products were later described in many prokaryotes. The first POT to be purified was DtpT from *Lactococcus lactis*, which had been shown to perform H^+-driven transport of di/tripeptides, independent from the ATP-driven Opp system (Smid *et al*, 1989; Kunji *et al*, 1993). Functional characterisation of liposome-reconstituted DtpT confirmed the influence of the electrochemical gradient, the preference for hydrophobic side chains and N-terminal amino groups, and provided for the first time an unbiased, *in vitro* system for transport assays, using physiological peptide substrates (Hagting *et al*, 1997; Fang *et al*, 1999; 2000). In addition, DNA footprinting and macroarray profiling revealed peptide transport in *L. lactis* is tightly regulated as part of the response mechanism to thermic, acidic and osmotic shock (Xie *et al*, 2004; Goh *et al*, 2004).

E. coli has four POTs (YdgR, YhiP, YjdL and YbgH), three of which have been functionally characterised mainly in whole cell assays, revealing common traits (like the overall substrate specificity shared with PepT1 and PepT2), as well as prominent differences. For instance, YdgR has higher affinity for cephalosporin antibiotics than YhiP in competition assays, while ampicillin/amoxicillin have no effect on either (Weitz *et al*, 2007; Harder *et al*, 2008). At the same time, YjdL has an altogether different substrate specificity, showing very

low affinity for tripeptides (Ernst *et al*, 2009). Although a projection structure of YbgH was assembled to 19Å using 2D crystallisation (Casagrande *et al*, 2009), high-resolution structures were needed for explaining the particularities of different POT proteins.

1.4.6. $PepT_{So}$ crystal structure: snapshot of the peptide transport mechanism

The X-ray crystal structure of a bacterial POT family peptide transporter from Gram-negative bactetrium *Shewanella oneidensis* ($PepT_{So}$) was recently solved to 3.3Å resolution (Newstead *et al*, 2011). $PepT_{So}$ has approximately 30% sequence identity compared to hPepT1 and PepT2, with many of the central cavity residues being conserved in all three proteins (fig. 4-1).

The structure confirmed that POT family transporters belong to the Major Facilitator Superfamily, as they showcase the core MFS fold of 12 TMs made up of two 6-helix domains (N- and C-terminal bundles), with a central, membrane-embedded hydrophilic cavity and two additional helices between H6 and H7 (A and B) (fig. 1-19, A). The latter two are grouped together laterally, away from the core structure (fig. 1-19, C). In the structure, the central cavity is "occluded" (closed to both sides of the membrane), and, although not co-crystallised with a substrate, contains electron density for an unidentified ligand of the approximate size of a dipeptide (fig. 1-19, B). Due to its position within the structure and the high degree of sequence conservation, it was proposed that the residues outlining the central cavity constituted the peptide binding site.

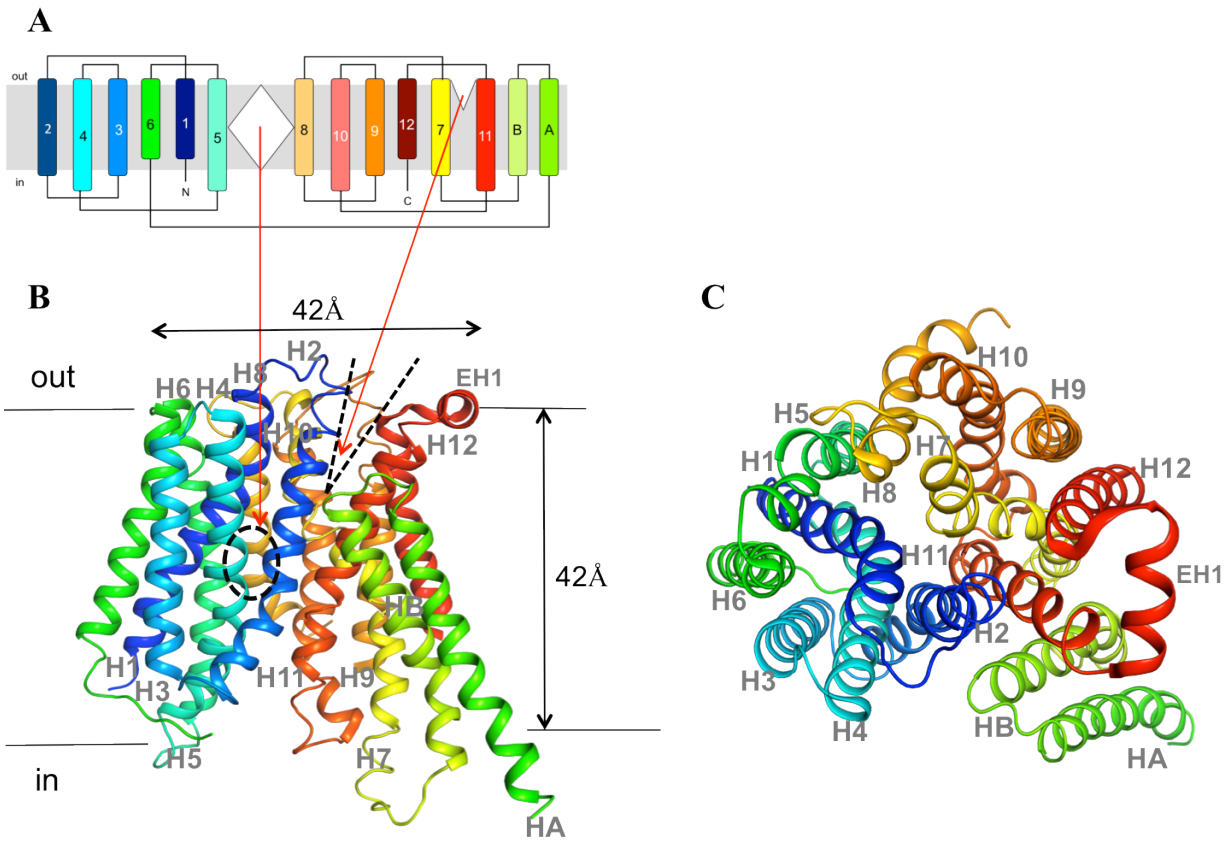


Fig. 1-19: Structure of $PepT_{So}$.
(A)Topology diagram of $PepT_{So}$. showing the central and extracellular cavities.
(B) Crystal structure of $PepT_{So}$ (PDB:2XUT), viewed in the plane of the membrane. Transmembrane helices coloured blue to red from the N to the C terminus. Red arrows indicate the central and extracellular cavities, marked with dashed circle and lines respectively.
(C) View of the $PepT_{So}$ structure from the extracellular side of the membrane, showing the 14 transmembrane helices.

An additional, extracellular cavity is observed between H7 and H11. The central cavity is separated from the extracellular one by a putative "extracellular gate", made up of helices 1 and 2 (N-terminal) stacking against 7 and 8 (C-terminal), similar to the one seen in the LacY structure (Abramson *et al*, 2003) and consistent with the gating mechanism for coupled transport first proposed for LeuT (Yamashita *et al*, 2005). At the intracellular side, the binding site is sealed by H4/H5 on one side stacking against H10/H11 on the other. The movement of C-terminal helices H7 and H10-12 is proposed to mediate the opening of this "intracellular gate", causing the transporter to switch to the inward-open conformation.

$PepT_{So}$ is the first example of a POT crystal structure and showcases an "inward-occluded" conformation for an MFS protein, in contrast to the fully occluded EmrD structure (Yin *et al*, 2006) and similar to the more recent structure of NarU (Yan *et al*, 2013). Shortly after $PepT_{So}$, another POT family homologue ($PepT_{St}$) was crystallised and revealed a similar fold but in an inward-open conformation (Solcan *et al*, 2012). As will be detailed in this thesis, the two structures reinforced the alternating access model for MFS transport and formed the basis of an in-depth study on both proteins.

1.5. Aims of this work

The structures of $PepT_{So}$ and $PepT_{St}$ in different conformations provide two snapshots of the peptide transport cycle in prokaryotes. Based on their close relatedness to the human peptide transporters PepT1 and PepT2, I aimed to functionally characterise the bacterial proteins using the structures as templates, and extend the use of these as model systems for the human peptide transporters.

My main approach was to characterise $PepT_{So}$ and $PepT_{St}$ by designing an efficient reconstitution protocol and a liposome-based assay to determine the substrate specificity, pH dependence and kinetic properties of peptide transporters. I aimed to investigate how substrate specificity is regulated by site-directed mutagenesis of the binding site followed by activity assays in different conditions, identifying residues involved in binding and proton coupling. I also tried to assess whether bacterial POTs can be used to study PepT-mediated

drug uptake. Additionally, using the structural information gained from crystal structures, electron paramagnetic resonance (EPR) experiments, molecular dynamics simulations and modeling in combination with the biochemical data, I aimed to describe in detail the conserved conformational changes that underlie the alternating access mechanism for the POT family of peptide transporters.

Other aspects of my project involved optimising crystallisation of bacterial peptide transporters by screening several homologues for expression and purification, and assessing their stability and crystal diffraction quality in different detergents. Further work on bacterial POT homologue GkPOT (recently crystallised to high resolution in an inward-open conformation with substrate bound) added more valuable data to the puzzle of substrate recognition and proton coupling. Finally, I aimed to characterise peptide binding in detergent solution and isotropic bicelles using isothermal titration calorimetry, scintillation proximity assays and NMR spectroscopy.

Chapter 2: Materials and methods

All chemicals were purchased from Sigma-Aldrich, VWR, Invitrogen and Fisher Scientific unless specified in the text. Detergents and lipids were purchased from Affymetrix and Avanti Polar Lipids (used for purification and crystallisation, respectively). The instruments used in the experiments presented here are listed below:

- centrifuges: Allegra X-22R, Microfuge 16, Avanti^{TM} J-20 and J-25, Optima L-100K and 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);
- ME2 NT diaphragm pump (Vaccubrand);
- Wallac 1409 DSA liquid scintillation counter;
- Bruker Elexsys 680 EPR spectrometer (94 GHz);
- MicroBeta TriLux microplate scintillation counter;
- ITC_{200} microcalorimeter;
- Omega NMR spectrometers (650 and 950 MHz).

2.1. Molecular biology methods

2.1.1. Preparation and transformation of competent cells

Chemically competent *E. coli* cells of the strains C43(DE3), BL21(DE3) and OmniMAX^{TM} (purchased from Life Technologies and Lucigen) were prepared by a method based on a previously described protocol (Hanahan, 1983). 100 ml LB (lysogeny broth) cultures were grown at 37°C to an OD_{550} of 0.48. The cultures were then cooled on ice for 15 min and distributed to 50 ml Falcon tubes. The cells were centrifuged at 4°C and 3,000 g for 15 min, and pellets were resuspended in a total of 40 ml transformation buffer I (30 mM KAc, 100 mM RbCl, 10 mM $CaCl_2$, 50 mM $MnCl_2$ and 15% glycerol, pH 5.8). After 15 min on ice, the centrifugation step was repeated and pellets resuspended in 4 ml transformation buffer II (10 mM MOPS, 75 mM $CaCl_2$, 10 mM RbCl and 15% glycerol, pH 6.5). The cells were distributed into 100 µl aliquots, flash-frozen in liquid N_2 and stored at -80°C.

For the transformation, an aliquot volume of competent cells was typically transformed with 1-5 µl plasmid DNA in sterile conditions, kept on ice for 15 min and heat shocked at 42°C for 40 s. After a further 10 min on ice, LB was added to the transformed cells to 1 ml, and the cells were grown at 37°C for 50 min then centrifuged at 7,000 g for 2-3 min. Pellets were resuspended in 100 µl LB and plated onto LB agar plates with the appropriate antibiotic selection (50 µg/ml for kanamycin or 100 µg/ml ampicillin/carbenicillin). Colonies were grown overnight at 37°C.

2.1.2. Site-directed mutagenesis and cloning of bacterial POT constructs

Primer pairs were designed containing the desired mutation flanked by ~10-40 complementary bases on each side. Primers were synthesised by Eurofins. Primer stocks (100 μM) were stored at -20°C in doubly distilled (dd) H_2O. The mutated sequences were incorporated into pWaldo-GFPd vectors, which are modified pET28(a+) plasmids with a C-terminal GFP tag (Waldo *et al*, 1999), and a tobacco etch virus (TEV) protease cleavage site between the multiple cloning site (MCS) and the GFP (Drew *et al*, 2001) (fig. 2-1).

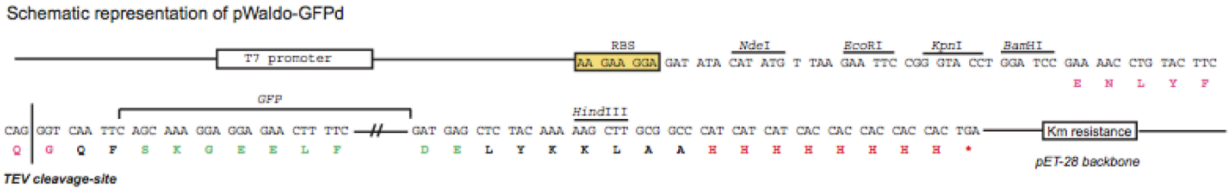


Fig. 2-1: Schematic representation of the pWaldo-GFPd plasmid. The T7 promoter region is followed by the RBS (ribosomal binding site) and the multiple cloning site (MCS) with recognition sequences for restriction enzymes. The TEV protease cleavage is next, followed by the C-terminal bacterial GFP sequence, an additional restriction site, the octahistidine (His_8) tag and the kanamycin resistance gene.

To incorporate the mutated sequences, plasmids were amplified by PCR using the KOD HotStart DNA polymerase kit (Novagen). A typical reaction mix contained: 1x KOD buffer, 0.5-2 ng/μl template DNA (pWaldo plasmid with wild type POT sequence), 0.4 μM forward primer, 0.4 μM reverse primer, 200 μM (each) dNTPs, 1.5 mM $MgSO_4$, 1 μl HotStart KOD polymerase and H_2O to 50 μl. A PCR reaction with 17 elongation cycles was typically run (table 2-1), followed by removal of methylated DNA with 2 μl DpnI restriction enzyme per reaction (1h at 37°C). 5 μl of each DpnI-treated reaction was used to transform *E. coli* OmniMAX cells, which were plated on kanamycin (Kan) plates as described. In the case of

double amino acid mutations, both pairs of primers were added to the template as part of the same reaction.

Step	Temperature [°C]	Time [min]	No. of cycles
Denaturation	94	2	1
Denaturation	94	0.25	17
Hybridisation	65	1	17
Elongation	68	8	17
Final elongation	68	7	1
End	4	∞	1

Table 2-1: Typical polymerase chain reaction (PCR) conditions for amplifying the pWaldo-GFPd plasmid in site-directed mutagenesis.

Plasmids containing the mutated genes were purified by DNA extraction from the *E. coli* Omnimax cells, using the Fermentas miniprep kit and protocol. Cells were grown at 37°C overnight in 5 ml cultures (LB), centrifuged and resuspended in 250 μl resuspension buffer. Cells were lysed and the DNA denatured by adding 250 μl alkaline lysis buffer, then plasmid DNA was renatured with 350 μl neutralisation buffer. Denatured protein and chromosomal DNA were separated by centrifugation at 8000 g for 10 min, and the plasmid DNA was bound to a spin column. Two 500 μl washing steps were performed with 1 min centrifugation steps in between. Plasmid DNA was eluted in 3-40 μl ddH_2O to concentrations typically between 100 and 200 ng/μl. All constructs were verified by gene sequencing (SourceBioscience).

2.1.3. Cloning the wild type GkPOT construct

The pET28b plasmid containing the wild type GkPOT DNA was a gift from Dr Ryuichiro Ishitani (University of Tokyo). The GkPOT gene was re-cloned into the pWaldo-GFPd plasmid. To amplify the gene sequence, primers were designed containing restriction sites for NdeI (forward) and BamHI (reverse), to match the MCS of the pWaldo vector (fig. 2-1). These primers were used in a PCR reaction, set up as described in 2.1.2 (but using 30 elongation cycles and a lower hybridisation temperature, 55°C). The PCR product (amplified GkPOT gene) was run for 50 min at 110 V on a 1% agarose gel made with 1x TAE buffer (40 mM Tris base, 20 μM acetic acid and 1 mM EDTA) and stained with 1x SYBR® Safe DNA gel stain (fig. 2-2, A). The DNA was gel-extracted using the Fermentas gel extraction kit (gel band weighed and dissolved in 1:1 w/v binding buffer for 10 min at 55°C, then bound to a spin column, washed and eluted), to remove any traces of the original plasmid.

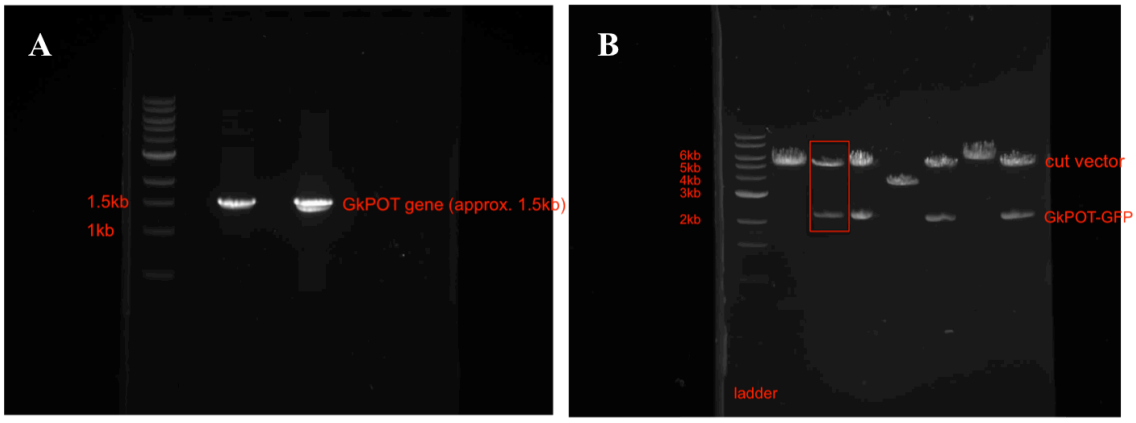


Fig. 2-2: Cloning the wild type GkPOT gene into pWaldo-GFPd.
(A) Agarose gel of the initial PCR product showing the amplified GkPOT gene before and after the PCR cleanup.
(B) Agarose gel of double-digested plasmids purified from seven *E. coli* colonies (NdeI and HindIII digest). The red square shows the restriction pattern confirming successful insertion of the GkPOT-GFP construct.

The gel-extracted PCR product and the empty pWaldo template were double-digested using NdeI and BamHI from New England Biolabs (NEB) (1x reaction buffer, 1 μl of each

enzyme, plus ddH_2O and PCR product/template to 50 µl), to create dsDNA fragments with compatible sticky ends. After 1-2h incubation at 37°C, the double-digested products were cleaned up using the Fermentas PCR cleanup kit. DNA concentration and purity were estimating using a NanoDrop system, by measuring the absorption maximum at 260 nm (A_{260}) (equal to 1 for a pure 50 ng/µl solution), as well as the A_{260}/A_{280} ratio (the latter indicating the presence of protein) (pure DNA typically shows a ratio of 1.8-2) (Glasel, 1995).

The cleaned up products were ligated together using the QuickLigase kit from NEB (1x reaction buffer, 1 µl QuickLigase plus the insert and vector in a molar ratio of at least 3:1). After the ligation reaction (15 min at RT), 5 µl were transformed into *E. coli* Omnimax cells. A control ligation reaction was performed in parallel without the insert. Colonies from the ligation plate were used to extract the pWaldo-GkPOT-GFP plasmid, which was double-digested with BamHI and HindIII to check the correct insertion of the GkPOT-GFP gene construct (fig. 2-2, B).

2.2. Expression, purification and crystallisation of POT family transporters

2.2.1. Expression trials of POT homologues

Expression trials were carried out with pWaldo-GFP constructs of POT proteins in Terrific Broth (TB) using IPTG induction, as well as in two types of minimal autoinduction media:

PASM-5052 and ZYM-5052 (Studier *et al*, 2005). PASM-5052 was made up as follows: 20 mM $MgSO_4$, 0.2x trace metal mixture, 1x 5052 solution (0.5% glycerol, 0.05% D-glucose and 0.2% α-lactose), 1x "P" solution (50 mM Na_2HPO_4, 50 mM KH_2PO_4 and 25 mM $(NH_4)_2SO_4$, pH 6.7), 0.2 mg/ml (each) amino acid mixture (no cysteine/methionine) and 20 mg/ml methionine. ZYM-5052 contained: 1 mM trisodium citrate, 0.2x trace metal mixture, 2 mM $MgSO_4$, 1x 5052 solution, 1x "M" solution (25 mM Na_2HPO_4, 25 mM KH_2PO_4, 50 mM NH_4Cl and 1 mM Na_2SO_4, pH 6.4) and (per 200 ml) 2 g tryptone and 1 g yeast extract.

For the starter cultures, one colony of *E. coli* C43(DE3) cells freshly transformed with the pWaldo-POT-GFP construct was inoculated in MDG minimal medium (1 mM trisodium citrate, 0.2x trace metal mixture, 2 mM $MgSO_4$, 0.5% D-glucose, 0.25% aspartate and 1x "M" solution), with 50 μg/ml kanamycin. After growth at 37°C overnight, the cultures were back-diluted 1:100 into 10ml LB, PASM-5052 and ZYM-5052 respectively (+ Kan) and grown at 37°C in a shaker-incubator (280 rpm) to an OD_{600} of 0.6-0.8, when expression was initiated with 0.4 mM IPTG (TB samples only), by inducing the T7 promoter which in turn induces expression of the T7 polymerase used to express the recombinant gene. For the expression trials using cells with the *pLemo* plasmid (Wagner *et al*, 2008), expression was induced using a gradient of L-rhamnose concentrations from 0.1 to 2 mM. In all cases, the temperature was then dropped to 25°C after induction and maintained overnight. The OD_{600} of the cultures was measured and the GFP fluorescence of 1 ml culture (normalised to OD=8) was measured using a fluorescence plate reader. Relative fluorescence units (RFU) at 488 nm excitation and 512 nm emission were converted to expression levels in mg/ml according to a previously generated standard curve plotting fluorescence against protein amount with purified GFP protein, as described (Drew *et al*, 2006).

2.2.2. Large-scale expression of GFP-fusion POT constructs and membrane isolation

GFP-tagged constructs of POT proteins were typically expressed in 2-6 L batchesin a shaker-incubator (280 rpm), using 2.5 L TunAir aerated culture flasks, each flask carrying 1 L culture in TB medium. Overnight cultures were set up in LB. For the SeMet incorporation, cells expressing $PepT_{St}$ were grown in PASM-5052 minimal medium as described in 2.2.1, but replacing the methionine component with a mixture of 10 mg/ml methionine and 125 mg/ml selenomethionine. The ^{15}N-labelled proteins for NMR experiments were expressed in a variation of the 5052 minimal medium enriched with ^{15}N isotope; more specifically, 1 L mixtures of 50 mM Na_2HPO_4, 50 mM KH_2PO_4, 5 mM Na_2SO_4, 2 mM $MgSO_4$ and 0.75% glycerol were prepared and autoclaved before adding 50 mM $^{15}NH_4Cl$, 0.2x trace metals, 0.05% D-glucose and 0.01% α-lactose to make up the final mixture. The SeMet and ^{15}N cultures were grown in the same conditions as the TB cultures (shaking at 280 rpm and lowering the temperature from 37 to 25°C after reaching OD=0.6-0.8), and IPTG induction was found to favour expression yields even when using autoinduction media.

After overnight expression at 25°C, cultures typically reached an OD_{600} of 6-10. The cells were spun down at 5,000 g (20 min at 4°C) and the pellets resuspended in 1x PBS (8.1 mM Na_2HPO_4, 1.9 mM KH_2PO_4, 137 mM NaCl and 2.7 mM KCl), using 40-45 ml buffer per pellet (from 1 L culture), with the help of a glass homogeniser. A spatula's tip of DNase I (bovine) was added to the homogenised mixture. The cells were then lysed at high pressure (30 kpsi) and 5°C in a cell disruptor (one pass). The lysed cell debris was spun down at 25,000 g (20 min at 4°C) and discarded. The supernatants were distributed to 70 ml ultracentrifuge tubes (Beckman) and spun down at 41,000 rpm in a TI-45 rotor (~130,000 g)

for 2h at 4°C. The pellets (isolated membranes) were resuspended in 1x PBS (up to ~20 ml for 1 L) using the glass homogeniser (or a syringe for small volumes), and GFP fluorescence was measured as described in 2.2.1. Typical RFU counts per 100 μl membrane suspension were between 20,000 and 100,000. Membranes were flash-frozen in liquid N_2 and stored indefinitely at -80°C.

2.2.3. Purification of GFP-fusion POT constructs

The purification protocol was based on the one previously described (Drew *et al*, 2006). All steps were carried out at 4°C unless specified. Membranes (typically obtained from 6 L culture, containing an estimated ~50 g fusion protein at the time of harvesting) were thawed and detergent-extracted in 360 ml solubilisation buffer (1x PBS, 150 mM NaCl, 10% glycerol and 1% DDM), stirring for 1h in the cold room. For each purification, the volume of the solubilisation reaction (and implicitly the amount of DDM) was adjusted to the amount of expressed protein (for example, for a yield of 25 g the solubilisation volume would be 180 ml). Non-solubilised membranes were isolated by centrifugation at ~130,000 g (45 min), and GFP fluorescence of the supernatant (solubilised membranes) was compared to the estimate before solubilisation; typical yields in DDM were over 90%. During the centrifugation step, an affinity chromatography column was set up using Ni-NTA or Ni-sepharose resin (~1 ml resin per 5 mg estimated protein) added to a drip column and pre-equilibrated with wash buffer (1x PBS, 150 mM NaCl, 10 mM imidazole and 0.1% DDM). Adding glycerol to the wash buffer did not influence protein stability, and thus was discontinued after initial trials.

10 mM imidazole was also added to the solubilised membrane fraction after centrifugation (the imidazole is needed to prevent unspecific binding of contaminants to the resin).

For the affinity step, the solubilised membrane fraction was incubated with the equilibrated resin (stirring vigourously for 2h); the relatively long incubation time is necessary as binding of the His_8 tag protein to the resin can be hindered by the detergent micelles. After incubation, the mixture was added to the drip column, the flowthrough was collected and fluorescence was measured. A considerable amount of protein was collected in the flowthrough (50-60% of the total protein for $PepT_{So}$, 20-30% for $PepT_{St}$) and did not bind to the resin even after a second incubation. The amount of protein lost in the flowthrough at this step varied with each protein, but remained constant between different purifications of the same protein (table 2-2). After collecting the flowthrough, the resin was washed with 8 column volumes (CV) wash buffer at 30 mM imidazole. Some protein was typically lost in the washing step, and if the amount did not exceed 5% a second washing step at 40 mM imidazole was carried out; nevertheless, the 30 mM wash was generally sufficient to remove most contaminants. The bound protein was eluted in wash buffer at 250 mM imidazole (typically in 30-40 ml), and the GFP fluorescence was measured.

$PepT_{St}$	$PepT_{So}$	$PepT_{Xc}$	$PepT_{Bc}$	$PepT_{Bs}$	$PepT_{Vc}$	YdgR	GkPOT
40-50%	70-80%	60-65%	55-60%	60-65%	55-60%	65-70%	45-50%

Table 2-2: Binding efficiency of POT constructs at 10 mM imidazole, calculated from GFP fluorescence counts.

To cleave the GFP tag, TEV protease (in 20 mM Tris HCl pH 7.5, 300 mM NaCl, 20% glycerol and 5 mM DTT) was added to the eluted protein in a ratio of approximately 1:1 (w/w) for complete cleavage; DDM was added to 0.1% to make up for the added TEV volume. The mixture was transferred inside a Spectra/Por dialysis membrane (3.5 kDa cutoff) and dialysed overnight against 20 mM Tris HCl (pH 7.5), 150 mM NaCl and 0.03% DDM. The dialysis

step was used to remove the excess imidazole (the volume of dialysis buffer was chosen such that the imidazole concentration would drop to or below 10 mM overnight), as well as to gradually lower the detergent concentration to only 3 times the critical micelle concentration (CMC).

On the second day, the dialysate was filtered using 200 nm Millex®GP filter units and supplemented with imidazole to 10 mM if necessary. For the reverse-purification step, the sample was loaded onto a 5 ml HisTrap™ column (GE Healthcare) (equilibrated with crystallisation buffer – the same as the dialysis buffer) at 2 ml/min, in order to bind the His-tagged products of the TEV protease cleavage (GFP and the protease itself). The flowthrough (containing untagged protein) was collected and concentrated to ~500 µl using a 30 kDa cutoff Amicon centrifugal filter unit (concentrator), while the GFP and TEV were eluted from the HisTrap column with 500 mM imidazole.

For the final step, the concentrated protein was run at 0.4 mg/ml on an Superdex™200 (S200) gel filtration column (GE Healthcare) pre-equilibrated with either crystallisation buffer (for crystallisation experiments) or reconstitution buffer (50 mM KPi pH 7, 150 mM KCl and 0.3% DM) (for liposome reconstitution with the rapid dilution method). The protein concentration (here and elsewhere in this thesis) was estimated before the gel filtration step by measuring A_{280} on a NanoDrop machine (UV light absorption by aromatic residues) and using the protein-specific extinction coefficient (ε) estimated from the amino acid sequence with the ProtParam tool on the ExPASy website (http://web.expasy.org/protparam/). No more than 5-6 mg protein were loaded on the gel filtration column at any one time to avoid signal saturation. For crystallisation trials, different detergents were added to the crystallisation buffer during gel filtration, as detailed below (table 2-3).

Detergent	CMC [%]	Buffer concentration [%]
n-dodecyl β-D-maltoside (DDM)	0.0076	0.03
n-decyl β-D-maltoside (DM)	0.08	0.3
n-nonyl β-D-maltoside (NM)	0.28	0.8
n-octyl β-D-glucopyranoside(OG)	0.45	1.2
lauryldimethylamine-oxide (LDAO)	0.023	0.07
octaethylene glycol monododecyl ether ($C_{12}E_8$)	0.0048	0.015
β-D-glucopyranoside,5-cyclohexylpentyl 4-O-α-D-glucopyranoside (Cymal-5)	0.12	0.4
3α-hydroxy-7α,12α-di-((O-ß-D-maltosyl)-2-hydroxyethoxy)-cholane (Façade EM)	0.02	0.06

Table 2-3: List of detergents used in the purification of POT family transporters. The typical detergent concentration in the buffers was approx. 3 times the critical micelle concentration (CMC). The CMC values are reproduced from the Avanti Polar Lipids website.

After gel filtration, the protein fractions were pooled and concentrated again (50 kDa cutoff to limit excessive detergent concentration), typically to ~10 mg/ml (for crystallisation) and ~2 mg/ml (for liposome reconstitution). Samples were flash-frozen in liquid N_2 and stored at -80C. Samples were taken after each purification step and checked by SDS-PAGE with Tris-Glycine gels. Samples were mixed with 1x Tris-Gly sample buffer (Novex®) and run for 1.5 h at 125 V and 35 mA alongside a prestained protein marker (PageRuler™). The running buffer contained 125 mM Tris, 1.25 M glycine and 0.5% SDS, and the components of the gel are shown below (table 2-4). Gels were stained with Instant Blue dye (Expedeon) and destained in water before visualising.

Gel	Component	Volume
Separation (10%)	Acrylamide (30%)	2 ml
	Tris (pH 8.8)	1.5 ml
	H_2O	2.44 ml
	APS	60 µl
	TEMED	4 µl
Stacking (5%)	Acrylamide (30%)	333 µl
	Tris (pH 6.8)	500 µl
	H_2O	1.17 ml
	APS	20 µl
	TEMED	3 µl

Table 2-4: List of components for one gel mixture.

2.2.4. Right angle light scattering and refractive index profiles

RALS measures the average light scattering intensity caused by the protein-detergent micelle particles in solution at a fixed (90°) angle and calculates the average molecular weight of the particles by comparing the measurements to those of a standard sample of known molecular weight. In addition, the refractive index profile is used to determine the size of the uncomplexed detergent micelles (separated by the size exclusion column), based on standard values for the given detergent found in the literature (Strop and Brunger, 2005), and consequently the amount of detergent contained in the protein-complexed micelles. Together, the data from the two profiles is used to determine average molecular weights for the micelle as well as the protein:detergent ratios therein (Slotboom *et al*, 2008).

For the experiment, an S200 size exclusion column was equilibrated with standard crystallisation buffer (20 mM Tris HCl pH 7.5, 150 mM NaCl and 0.03% DDM) and a RALS/RI/UV profile was collected for a standard BSA sample of known molecular weight

(200 µl at 2 mg/ml), using a Viscotek TDAmax tetra-detector from Malvern. The $PepT_{So}$ and $PepT_{St}$ samples (200 µl at 2 mg/ml) were then each run on the column and similar profiles were collected. The data was analysed by Matthew Jennions at the Membrane Protein Laboratory at the Diamond Light Source.

2.2.5. Purification of the His-tagged $PepT_{So}$ construct for SPA

The wild type $PepT_{So}$ gene was cloned into a pET11 plasmid (with a C-terminal His_6 tag) by Dr Joanne Parker. Expression and membrane isolation were carried out as described in 2.2.2. The isolated membranes (from 6 L culture) were added to a seven-step sucrose density gradient (25 to 55% in 20 mM Tris, pH 8), and the inner membrane fraction was separated through ultracentrifugation at 113,000 g (18h at 4°C). This was followed by another centrifugation step (2h at 130,000 g) to remove the glucose, then the inner membranes were resuspended in 20 mM Tris buffer, flash-frozen and stored at -80°C.

The membranes were detergent-extracted with solubilisation buffer (25 mM Tris pH 8, 300 mM NaCl, 10% glycerol and 1% DDM) and incubated with 10 ml Ni-sepharose resin. The affinity step was carried out as described, with two washes (200 ml and 100 ml) at 20 and 40 mM imidazole respectively (wash buffer the same as the solubilisation buffer with 0.1% DDM and no glycerol), before eluting the protein with 250 mM imidazole. The protein was concentrated and run on a size exclusion column. During this step, half of the sample was exchanged into 20 mM Tris HCl (pH 7.5), 150 mM NaCl and 0.03% DDM, and the other half in liposome reconstitution buffer (50 mM KPi pH 7, 150 mM KCl and 0.3% DM).

2.2.6. Crystallisation of POT family transporters

In optimal conditions, a supersaturated protein solution can give rise to stable molecular aggregates (nucleation centres), to which free molecules can gradually bind to form crystals. If the crystal lattice is sufficiently well ordered, the X-ray diffraction pattern of the crystal can lend itself to high resolution structure determination. To achieve this, crystallisation screening is used to identify favourable conditions for the protein solution to reach supersaturation and for nucleation to occur. On average, this depends on the protein concentration as well as the concentration of the precipitant in the crystallisation solution (fig. 2-3). Moreover, the types of salt and precipitant used, the pH of the solution and the temperature can all influence the process, such that identifying crystallisation conditions is mostly an empirical process. However, commercially available crystallisation screens have been compiled from previously successful conditions and provide a useful starting point for further optimisation.

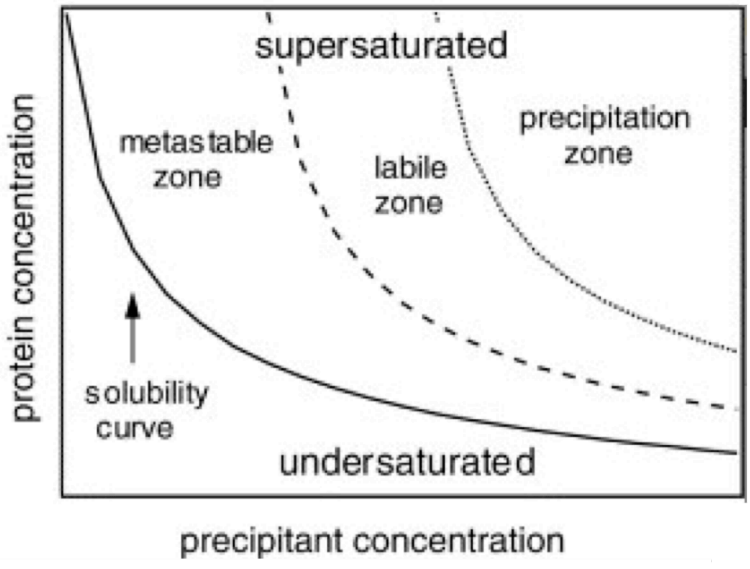


Fig. 2-3: Phase diagram for protein crystallisation. Depending on the protein and precipitant concentrations, the protein solution can be undersaturated (not sufficient to allow formation of nucleation sites) or supersaturated (favourable for crystal growth). The supersaturated region consists of the metastable zone (low saturation, crystals can form but may be unstable) and the labile zone (high saturation, favouring spontaneous nucleation). Beyond the labile zone, high protein and/or precipitant concentrations can cause precipitation (Asherie, 2004).

Initial crystallisation trials with POT proteins were carried out in different detergents as summarised in the results chapter, using the MemGold (Newstead *et al*, 2008) and MemGold 2 (Parker and Newstead, 2012) screens from Molecular Dimensions. All trials were set up in 96-well SwissSci plates filled up manually with the crystallisation conditions in the screen (70 µl per well), using a Mosquito pipetting robot to set up sitting drops (fig. 2-4, A) containing 200 nl protein solution (typically at 7-10 mg/ml) and 200 nl reservoir solution (the crystallisation condition). For each screen, plates were incubated at 4 and 19°C.

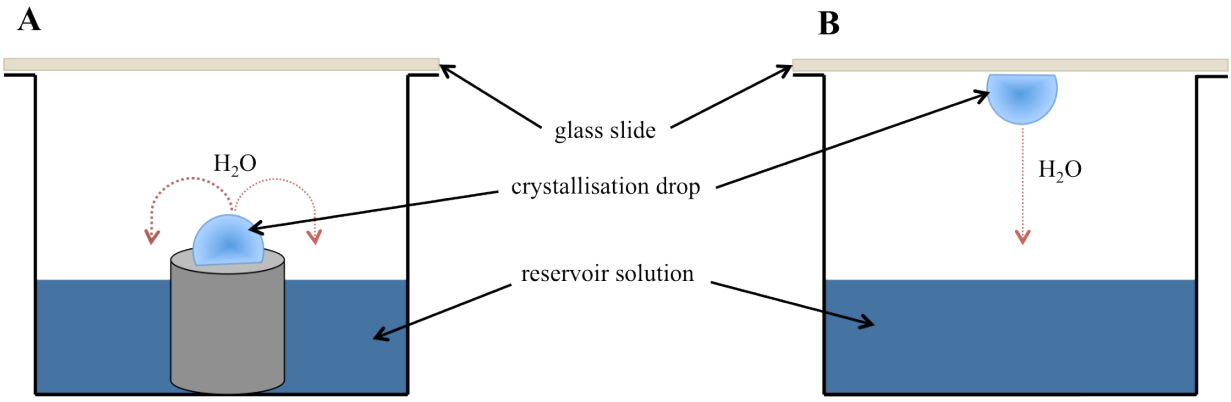


Fig. 2-4: Vapour diffusion crystallisation using the sitting drop (A) and hanging drop (B) methods. The crystallisation drop contains the protein solution mixed with the reservoir solution. As the precipitant and salt concentrations in the drop are lower than those in the reservoir, water will gradually diffuse out of the drop until the concentrations are equilibrated. This process decreases the volume of the drop and progressively increases the concentrations of its components, allowing crystals to form.

Following initial hits, crystallisation conditions were optimised using the hanging drop method (fig. 2-4, B) in 24-well crystallisation plates (500 µl reservoir solution per well), using 1 µl protein solution and 1 µl reservoir solution to set up the drops. To optimise, the concentration of each component in the initial crystallisation screen condition (precipitant, salt) was modified. For example, $PepT_{So}$ formed crystals in the Façade EM detergent in 0.1 M NaCl, 0.1 M $CdCl_2$, 0.1 M Tris (pH 8) and 33% PEG400. Starting from this condition, an optimisation plate was set up varying the PEG400 concentration (26 to 35% in 6 steps), while a second plate featured additional varying concentrations of salt (from 0.01 to 0.2 M), and a third used the series of PEG concentrations in combination with different buffers (MES at pH

6.5, HEPES/NaOH at pH 7 and 8, and sodium acetate at pH 4). Similar optimisation screens were set up for the other proteins starting from different crystallisation conditions.

Crystals typically grew within 2-3 days and were collected for diffraction screening a few days later. For cryoprotection, crystals were transferred to a solution containing the same components as the reservoir solution but supplemented with 0.03% DDM and the precipitant (PEG) concentration increased to 35%, prior to flash-freezing and storing in liquid N_2. Crystal screening and X-ray diffraction data collection for both native and SeMet/mercury-derivatised crystals were carried out at Diamond Light Source and the European Synchrotron Radiation Facility, on both regular and microfocus beamlines.

To aid with the phasing of $PepT_{St}$ crystals, cysteine mutations were introduced at several positions in the protein, as described in the results chapter. The mutant variants were purified with the standard protocol and were derivatised before the final gel filtration step by incubation with 2.5mM methylmercury chloride (MMC) for 2h at RT. Protein precipitated during the reaction was separated by centrifugation, and the sample was run on the gel filtration column to remove the excess label before setting up crystallisation drops in the standard $PepT_{St}$ condition (26% PEG400, 30 mM $MgCl_2$, 1 mM $CdCl_2$ and 0.1 M MES/NaOH pH 6.5). Alternatively, non-derivatised crystals were soaked in 5mM mercury acetate overnight (1 μl drops), followed by a fast soak in 10mM of the same before flash-freezing. In addition to mercury, other large anomalous scatterers were added to the $PepT_{St}$ crystals in the form of tantalum bromide crystals (a small amount added with a hypodermic needle) and I3C (magic triangle) (1 μl per crystallisation drop).

Several other methods were tested that would improve the quality of $PepT_{St}$ crystals. The first of these was lysine methylation, first introduced as a method for improving the chance of crystal contact formation at the surface of proteins (Rayment, 1997). The method

involves creating a methylated species by having the ε-amino group of lysine side chains react with a ketone or aldehyde in an aqueous environment, followed by further reaction with formaldehyde to yield a dimethylated product (Walter *et al*, 2006), thereby decreasing protein solubility and potentially favouring crystallisation. For this, purified $PepT_{St}$ (in 50 mM HEPES buffer pH 7.5, 250 mM NaCl and 0.03% DDM) was diluted to ~1 mg/ml and incubated with 20 mM borane-dimethylamine complex (ABC) and 40 mM formaldehyde for 2h at 4°C. The concentrations of ABC and formaldehyde were then doubled and the mixture was incubated for another 2h. For the third step, the ABC concentration was further increased to 50 mM for an overnight incubation. The methylation reaction caused no visible precipitation, and the protein was exchanged back into crystallisation buffer on the gel filtration column the following day before setting up crystal trials in the standard condition.

Crystal dehydration is a frequently used strategy for improving data quality, sometimes with spectacular results (Heras *et al*, 2003). Its previous successful application to membrane protein crystals (Efremov *et al*, 2010) prompted a similar attempt using crystals of $PepT_{St}$ (wild type and M96C/V256C variant). To initiate gradual dehydration, $PepT_{St}$ crystals grown in hanging drops were transferred from the standard condition (at 26% PEG400) to a new drop with cryoprotectant solution (35% PEG400), and the new drop was placed above a new reservoir with 35% PEG400. The following day, the cover slide was directly transferred above another reservoir with 40% PEG400, and the process was repeated every day until the precipitant concentration reached 50%. Crystals were then looped out of the drops and flash-frozen. Alternatively, crystals were transferred directly from 26% to 35% PEG400 on the first day, with the subsequent steps unchanged. A third batch of crystals was transferred directly from 26 to 50% drops.

Crystallisation trials were also set up with $PepT_{St}$ and peptide substrates dialanine (AlaAla) and diphenylalanine (PhePhe) in the standard condition. For this, the peptides were dissolved in crystallisation buffer to 10 mM (5 mM for PhePhe due to limited solubility), and incubated with the protein (at 10 mg/ml) overnight at 4°C prior to setting up the crystallisation drops.

The influence of lipids on membrane protein crystallisation is not well understood, but phospholipid head groups (particularly negatively charged) are known to interact with positive charges on the protein surface in several published crystal structures (Palsdottir and Hunte, 2004). Furthermore, co-purification of MFS transporter GlpT with phospholipids and glycerol favours crystallisation and sample monodispersity in different detergents (Lemieux *et al*, 2003), while gradual re-addition of lipids to purified LacY improves crystal resolution and creates crystals with different crystallographic symmetry (Guan *et al*, 2006). To investigate the effect of re-lipidation on purified $PepT_{St}$ in DDM, three lipid stocks from Avanti Polar Lipids in chloroform solution (50 mg each) were dried under an argon stream overnight with a vacuum pump: *E. coli* polar lipids (EPL), egg phosphatidylcholine (PC) and a 3:1 (w/w) mixture of the two (EPL:PC). The lipids were dissolved in crystallisation buffer (0.03% DDM) to 5 mg/ml and incubated for 1h at RT to facilitate detergent solubilisation. Purified $PepT_{St}$ in the same buffer concentrated to 30 mg/ml was then mixed with each lipid solution in a 1:0.02 protein:lipid ratio (w/w), to a final protein concentration of 13.6 mg/ml. The mixes were incubated overnight at 4°C on a rotating wheel. The following day, the samples were centrifuged to remove any precipitated protein, and crystallisation drops were set up in the standard condition at 19°C. The protocol was repeated using the same lipids with $PepT_{St}$ protein purified in NM, OG, Cymal-5 and $C_{12}E_8$.

2.3. Transport assays

Radioactively (tritium) labelled peptides and drugs were used as reporter substrates in the transport assays. They are listed below, along with the concentrations and specific activities of each stock (table 2-5).

Compound (^{3}H-labelled)	Conc. [µg/ml]	Specific activity [Ci/mmol]
L-alanyl-L-alanine	400.6	0.4
Glycylsarcosine	5.2	28.3
D-phenylalanyl-L-glutamine	5600	25
Cefadroxil	476.8	0.8
Valacyclovir	95.5	3.4

Table 2-5: List of ^{3}H-labelled compounds used as reporters for transport assays. All compounds were purchased from Moravek Biochemicals and American Radiolabeled Chemicals. All stocks were at 1 mCi/ml concentration.

2.3.1. Whole cell assays

$PepT_{So}$ and $PepT_{St}$ were overexpresed in *E. coli* BL21(DE3) cells as pWaldo-GFP-His_8 constructs (as typically used for expression), as well as pTTQ18-His_6 constructs (without the GFP tag). The cells were grown in 10 ml cultures (TB) at 37°C to an OD_{600} of 0.5-0.7 and expression was induced with 0.4 mM IPTG followed by growth at 25°C overnight. Approximately $5 \cdot 10^9$ cells (equivalent to 1 ml of cells at OD=5) were harvested 20h after induction and centrifuged at 2,500 g for 5 min. The cells were resuspended in 1 ml transport

buffer (25 mM HEPES/NaOH pH 7.5, 150 mM NaCl, 5 mM glucose, sterilised and filtered). For the assay, 375 µl resuspended cells were further diluted in transport buffer to a total volume of 500 µl and incubated at 37°C in a heat block. To start the reaction, radiolabelled peptide/drug was added to the desired concentration (typically 0.4 µM per assay for both ^{3}H-GlySar and ^{3}H-D-Phe-L-Gln, but 2 µM for ^{3}H-cefadroxil). At each time point, 100 µl samples were removed and the reaction was stopped by diluting the samples in 1.5 ml ice-cold transport buffer. The cells were then isolated on a 0.4 µm nitrocellulose filter. The filters were washed with 2x2 ml transport buffer and added to superpolyethylene scintillation vials (PerkinElmer) containing 7.5 ml Ultima Gold scintillation fluid.

The radioactivity contained by the filters was estimated with a scintillation counter, which measures the luminescence exhibited by the scintillation fluid in response to the ionising radiation. The signal in counts per minute (CPM), averaged from 300 measurements, was then converted to molar amounts of labelled substrate according to a standard curve (fig. 2-5). Further conversions were applied to correct for the presence of cold (unlabelled) substrate if necessary. All the data from whole cell and liposome-based transport assays was analysed in Prism 5 (GraphPad).

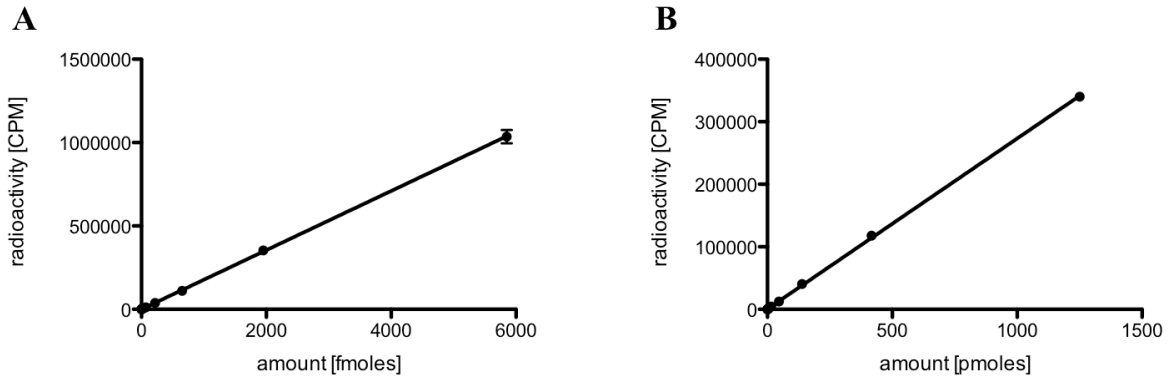


Fig. 2-5: Standard curves for ^{3}H-GlySar (A) and ^{3}H-AlaAla (B) showing the relationship between scintillation counts per minute (CPM) and the amount of radiolabelled peptide in the stocks used for transport assays of POT proteins.

An additional assay was performed using fluorophore-labelled peptide β-Ala-Lys-AMCA at 50 μM concentration. In this case, the cells were centrifuged after 15 min incubation and washed three times with ice-cold transport buffer to remove the excess fluorophore, and the AMCA fluorescence emission was measured in a plate reader (excitation at 340 nm, emission at 460 nm).

Finally, for the whole cell drug competition assay, YdgR was expressed as a His-tagged construct in a pBAD vector (Guzman *et al*, 1995). Expression was induced with 0.2% arabinose, and the assay was carried out as described above, but using 500 μM ^{3}H-GlySar, the uptake of which was measured after 10 min incubation in the presence of 5 mM competitor.

After preliminary transport data from whole cell assays, it was decided that developing an *in vitro* assay would offer a more robust, cleaner system in which to test POT-mediated uptake activity. Other peptide transporters had been assayed by reconstitution from a detergent solution state to a lipid bilayer vesicle (liposome). This system allows for specific modulation of buffer conditions inside and outside the liposome, creating for instance a membrane potential or pH gradient across the liposome membrane to provide energy for transport.

2.3.2. Adsorption onto biobeads and permeability check

A common reconstitution method previously applied to peptide transporters uses biobeads to forcibly remove detergent from the protein solution during incubation with preformed, detergent-destabilised liposomes (fig. 2-6). The biobead method was initially described for *L. lactis* transporter DtpT (Fang *et al*, 1999). In a preliminary study, Triton X-100 and DM were

identified as optimal detergents for destabilising the liposomes by partial or complete solubilisation prior to incubation with the protein. Meanwhile, the efficiency of reconstitution itself (amount of protein incorporated in the liposomes) is generally influenced by the critical micelle concentration (CMC) of the detergent in the protein solution (the higher the CMC, the easier to remove).

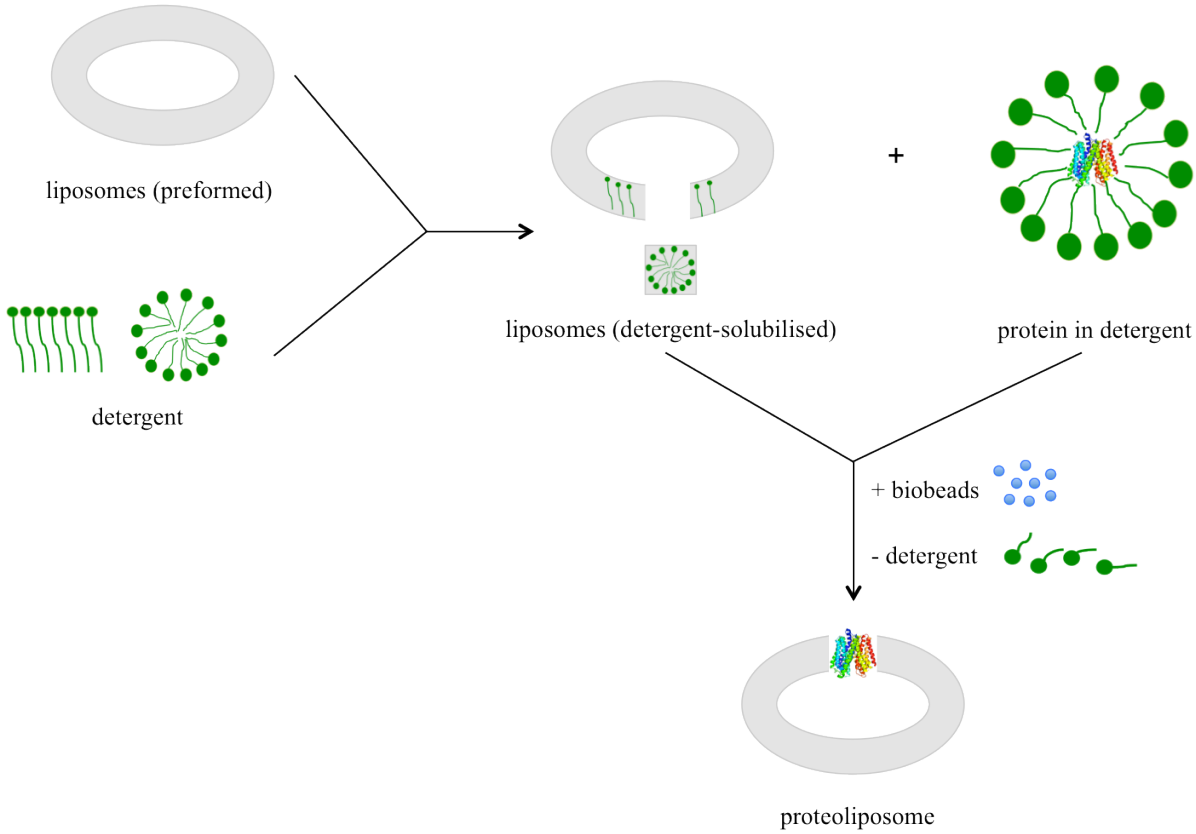


Fig. 2-6: Reconstitution of a peptide transporter (biobead method). Preformed liposomes are destabilised by detergent and incubated with protein in detergent solution. Biobeads are added to remove the protein from the detergent micelles and allow the formation of the proteoliposome.

$PepT_{St}$ was first reconstituted with the biobead method, using a variation on the method above and a standard protocol described for ABC transporters in DDM (Geertsma *et al*, 2008). *E. coli* polar lipids (Avanti) were dried and resuspended in 150mM NaCl and 20mM Tris buffer (at pH 6, 6.5, 7 and 9) under a N_2 stream, then sonicated to obtain small unilamellar vesicles (SUVs), followed by 10 freeze-thaw cycles to form large multilamellar vesicles (LMVs). To make the liposomes (large unilamellar vesicles or LUVs), the LMVs were extruded through 400 nm Nuclepore polycarbonate filters (Whatman). Liposomes were

destabilised by gradual addition of Triton X-100 until the absorbance of the solution at 540nm reached either the onset of solubilisation (R_{sat}) or full solubilisation (R_{sol}, clear solution). Liposomes were then incubated with pure $PepT_{St}$ (same buffer plus 0.03% DDM) in a 1:100 (w/w) protein:lipid ratio, with gradual addition of biobeads to strip away the detergent (40 mg biobeads per 40 μg protein, repeated three times with 1h incubation steps in between and followed by an overnight incubation). The following day, biobeads were removed, and proteoliposomes were spun down at 100,000 g, resuspended in 50 mM KPi at pH 7 and stored at -80°C.

Reconstitution efficiency was checked by silver staining and found to be approx. 100% regardless of pH (fig. 2-7). For this, equal volumes of the resuspended proteoliposomes and the supernatant from the final centrifugation step were run on a NuPage gel, which was then fixed in 50% ethanol, 10% acetic acid and 10% fixative enhancer (Bio-Rad) for 20 min at RT on a shaker and rinsed 5 times with water. The gel was stained and developed using the Bio-Rad kit; when the bands became visible, the reaction was stopped with 5% acetic acid.

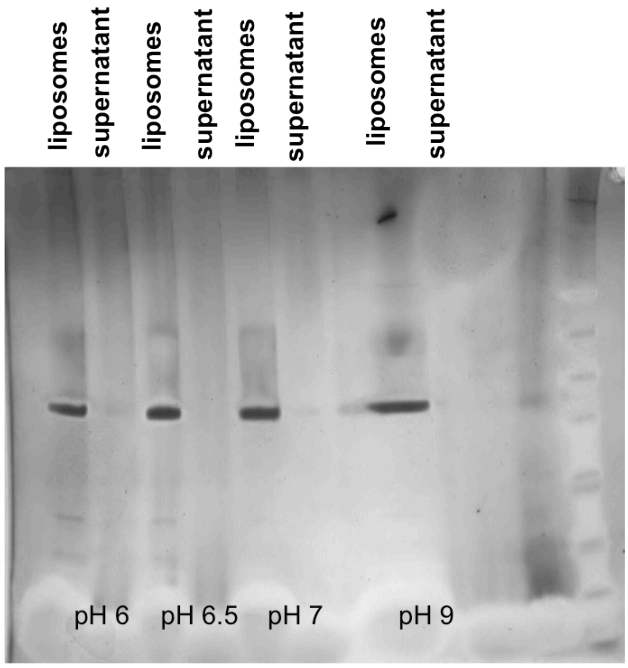


Fig. 2-7: $PepT_{St}$ liposome reconstitution (biobead method) at different pH values (silver stain). The protein is found only in centrifuged liposome fractions, not in the supernatant.

Despite efficient reconstitution, the method produced non-functional liposomes, as indicated by a lack of activity in the experimental setup previously used for DtpT and the *E. coli* transporters using radiolabelled transport substrates like GlySar and AlaAla; no activity was observed for protein purified from either the GFP or the His-tagged constructs. To investigate, the degree of proton permeability across the liposome bilayer was determined using a fluorescence assay as previously described (Fang *et al*, 1999; Rigaud and Lévy, 2003). Proteoliposomes were reconstituted with 100 μM 8-hydroxy-1,3,6-pyrenetrisulfonate (pyranine) (inside buffer at pH 7) and the fluorescence emission was measured after resuspension in buffer at pH 4 and pH 7, and compared to a detergent-disrupted liposome sample. The fluorescent signal was instantaneously quenched at pH 4, suggesting the liposomes were leaky (permeable to protons in the absence of protonophores). Therefore, no transmembrane ion gradient could be generated, which is central to performing transport assays in this system.

2.3.3. Lipid purification and rapid dilution

A second method was developed based on the rapid dilution approach, in which detergent is removed via rapid dilution of the protein-lipid mix into a large volume of buffer below the CMC of the detergent (fig. 2-8). Although not previously applied to peptide transporters, the rapid dilution method has frequently been used to reconstitute lactose permeases LacY and LacS, using OG for both protein purification and liposome destabilisation (Newman and Wilson, 1980; Foucaud and Poolman, 1992). Similarly, TRAP transporter SiaPQM was reconstituted by rapid dilution, using protein purified in DDM and exchanged into DM prior

to reconstitution (Mulligan *et al*, 2009). Given the stability of $PepT_{St}$ and $PepT_{So}$ in DM, the method was performed with both proteins after DM exchange on the gel filtration column.

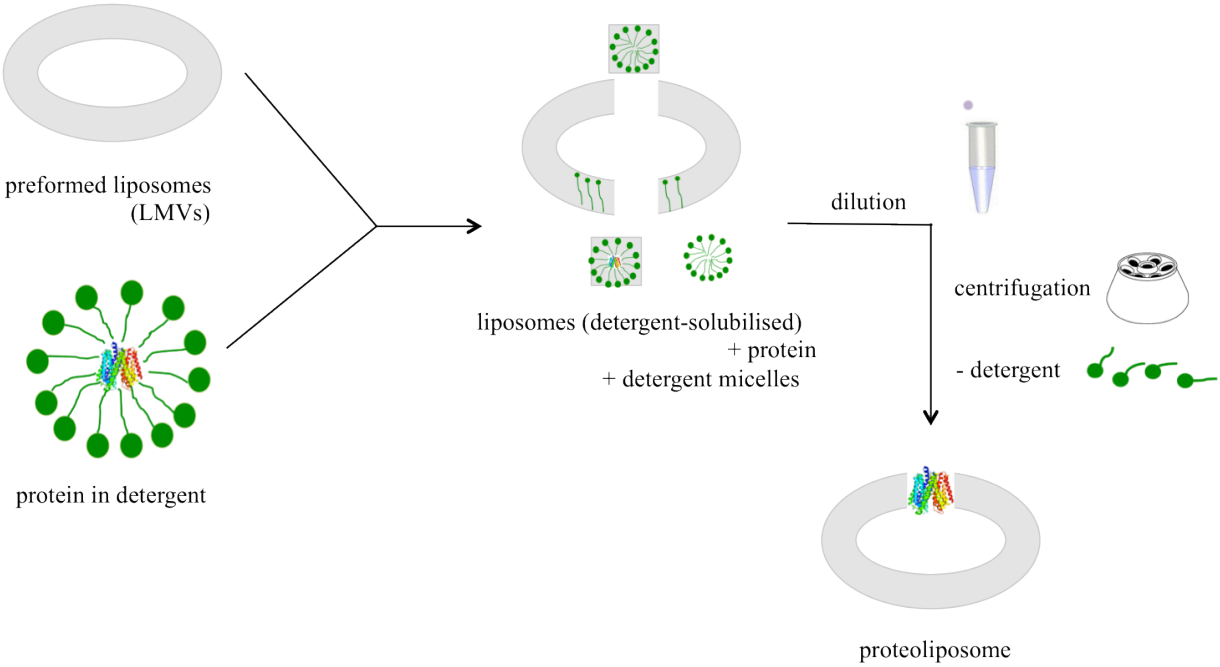


Fig. 2-8: Reconstitution of a peptide transporter (rapid dilution method). Preformed liposomes are destabilised by incubation with the protein in detergent solution. The mixture is diluted below the CMC of the detergent and spun down in an ultracentrifuge to separate detergent micelles from the proteoliposomes.

In order to create a physiological lipid environment for the transporter, liposomes were prepared with *E. coli* total lipid extract (Avanti). To remove contaminants, the extract was purified according to previous indications (In't Veld *et al*, 1992; Geertsma *et al*, 2008). Specifically, 1 g of total lipid extract in chloroform (20 mg/ml) was dried in a round-bottom flask using a rotary evaporator (30 min), then dissolved in ~5 ml chloroform (200 mg/ml). To precipitate the lipids, the mixture was added drop-wise to a conical flask containing 150 ml acetone with 1 mM DTT to prevent oxidation, while stirring vigourously at 4°C. After an overnight incubation, the precipitated lipids were spun down in Falcon tubes (10 min at 3,000 g) and the supernatants were discarded. The lipid pellets were exposed to a N_2 stream until completely dry (approx. 2x15 min), resuspended in 150 ml diethyl ether and incubated in a

conical flask while stirring for 1h at RT. Non-dissolved contaminants were spun down (10 min at 3,000 g), and the supernatant (pure lipids) was transferred to a pre-weighed round-bottom flask and dried on the rotary evaporator (30 min). The flask was re-weighed to calculate the yield (typically 80-90%). Chloroform was added to 100 mg/ml and the lipids (containing phosphatidylethanolamine, phosphatidylglycerol and cardiolipin) were mixed with egg phosphatidylcholine (Avanti) in a 3:1 ratio (w/w), as previously described (Knol *et al*, 1996). The drying step was repeated twice, with ethanol added in between steps to remove any chloroform traces. The dried lipid mixture was resuspended in liposome storage buffer (50 mM KPi at pH 7). To form SUVs, the mixture was sonicated on ice with a QSonica 700 sonicator at 4 μm intensity, in four cycles of 15 s each with 45 s between cycles. Finally, SUVs were subjected to three freeze-thaw cycles to fuse them into LMVs.

$PepT_{St}$/$PepT_{So}$ in reconstitution buffer (150mM KCl, 50mM KPi pH 7 and 0.3% DM) was incubated for 10 min on ice with the LMVs at a 1:40 (w/w) protein:lipid ratio (typically 250 μg protein and 10 mg LMVs). This was followed by rapid dilution into liposome storage buffer (diluting the detergent 1:150) and ultracentrifugation at 130,000 g to pellet the proteoliposomes. The pellets were resuspended in storage buffer to 0.4 mg/ml protein concentration, followed by five freeze-thaw cycles to scramble the protein orientation in the membrane, as previously described (Geertsma *et al*, 2008). Proteoliposomes were stored at -80°C for up to 2 years. Reconstitution efficiency was checked by running proteliposomes on a Tris-Gly gel and comparing the protein bands to standards of known amount, as shown here for a number of $PepT_{So}$ variants (fig. 2-9, A). Between 70 and 90% of all proteins were generally reconstituted with this method, as estimated by quantifying band intensities with the GelEval software (fig. 2-9, B and C); subsequent transport assay measurements were corrected for the reconstitution efficiency.

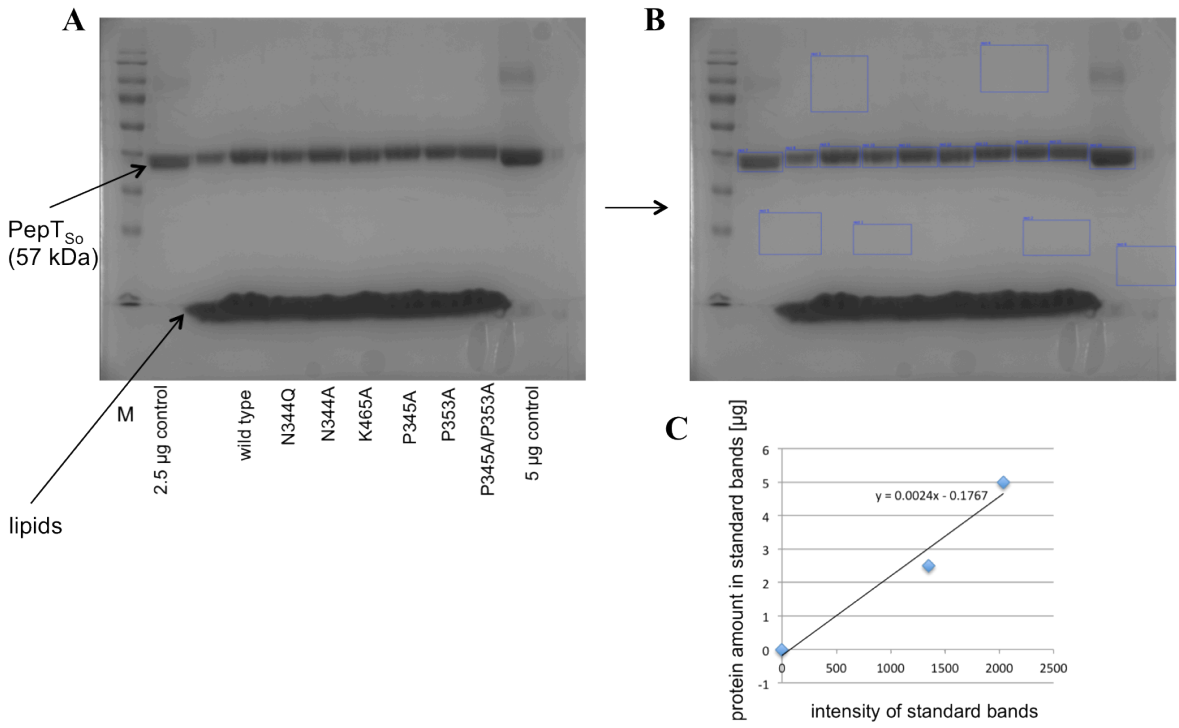


Fig. 2-9: Checking reconstitution efficiency (dilution method).
(A) Equal amounts (3.6 µg) of reconstituted $PepT_{So}$ (wild type and variants) are run on a TrisGly gel together with standards (pure protein in detergent).
(B) Using GelEval, band intensities are calculated against a selected background (no protein).
(C) Standard curve plotting protein amount against average intensities of standard bands (the amount of protein in purified standards was estimated by measuring absorbance at 280 nm).

2.3.4. Designing transport assays: ΔΨ-driven and counterflow

Liposome-based transport assays were successfully set up to test the activity of $PepT_{St}$ and $PepT_{So}$ reconstituted by rapid dilution. Typically, proteoliposomes (LMVs) were thawed and extruded through 400 nm polycarbonate filters (Avanti) on the day of the assay to create large, unilamellar vesicles (LUVs) of uniform size. Prior to extrusion, samples were centrifuged (50,000 rpm in a Beckman TLA100,2 rotor, equivalent to ~90,000 g, for 30 min at 4°C) and resuspended in the desired buffer (0.5-1 ml), as the extrusion process breaks open the liposomes, allowing for replacement of the storage buffer with the internal buffer required for the assay. After extrusion using a mini-extruder set (Avanti) and 1 ml Hamilton syringes (11 passes), the centrifugation step was repeated (at 15°C to prevent membrane fusion), and the

liposomes were resuspended in a small volume of internal buffer, ready for the assay. Two types of assays were performed: ΔΨ-driven (in which an artificial, + out/ – in membrane potential is generated by addition of an ionophore) and counterflow (driven by the substrate gradient) (fig. 2-10, A and B).

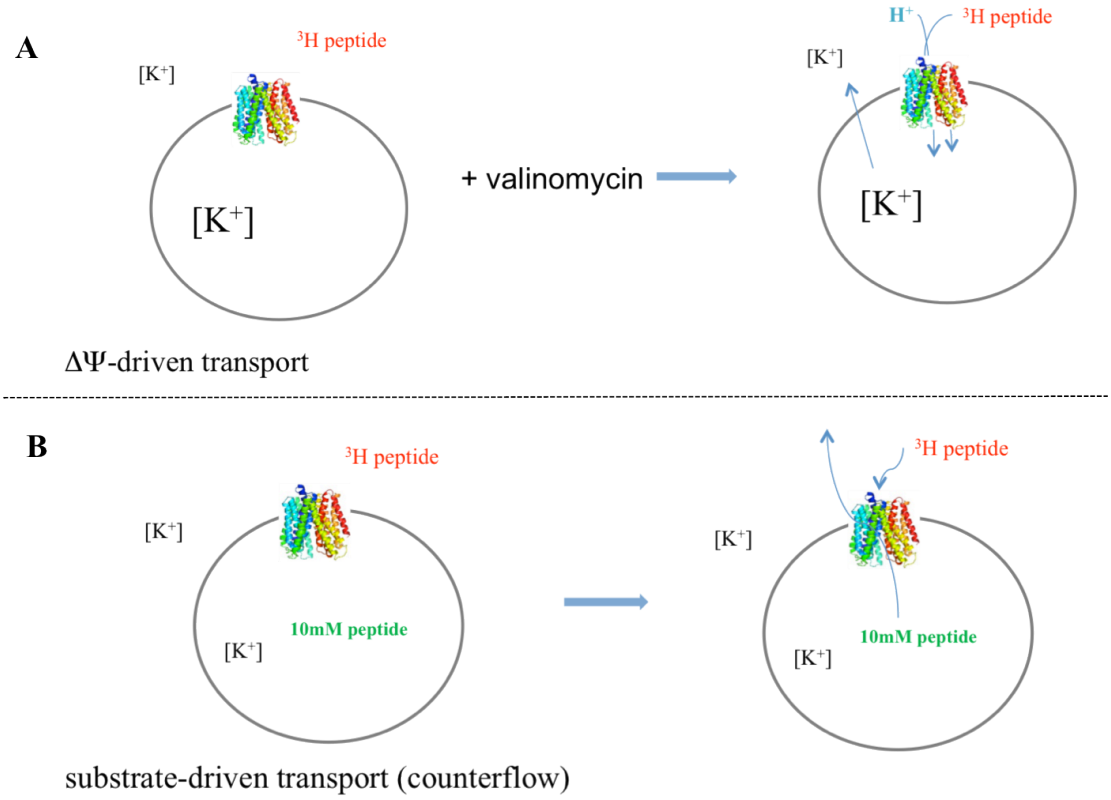


Fig. 2-10: Transport assays for peptide transporters.
(A) ΔΨ-driven: valinomycin causes efflux of K^+ ions, building up an external positive charge which in turn activates the transporter.
(B) Counterflow: in the absence of an ion gradient, the transporter exchanges unlabelled for labelled peptide in response to its concentration gradient.

For the ΔΨ-driven assay, proteoliposomes containing an excess of K^+ ions were diluted 1:50 into an external buffer without K^+, followed by addition of K^+ ionophore valinomycin (fig. 2-11, A); valinomycin makes the membrane permeable to K^+ ions, thereby creating an outward K^+ flux to establish a transmembrane electrical potential (ΔΨ), according to the equation previously described (Viitanen *et al*, 1986):

$$\Delta\Psi \text{ (mV)} = -59 \log([K^+]_{in}/[K^+]_{out})$$

, where $[K^+]_{in}$ and $[K^+]_{out}$ are the concentrations of K^+ ions inside and outside the liposomes (after dilution). In this particular case, liposomes containing 120 mM K^+ ions (internal buffer) were diluted 1:50, resulting in a $\Delta\Psi$ of –100mV. The transporter is protonated due to the positive charge build-up in the external buffer, which facilitates uptake of the radiolabelled peptide substrate (fig. 2-10, A). In other words, creating the $\Delta\Psi$ initiates the transport mechanism. Accumulation of the label in the liposome lumen is then quantified by scintillation reading and converted to molar units according to a standard curve (fig. 2-5). Finally, $\Delta\Psi$-driven transport can be abolished by addition of proton ionophore CCCP (carbonyl cyanide m-chlorophenylhydrazone) (fig. 2-11, B), which allows proton influx to effectively cancel the membrane potential.

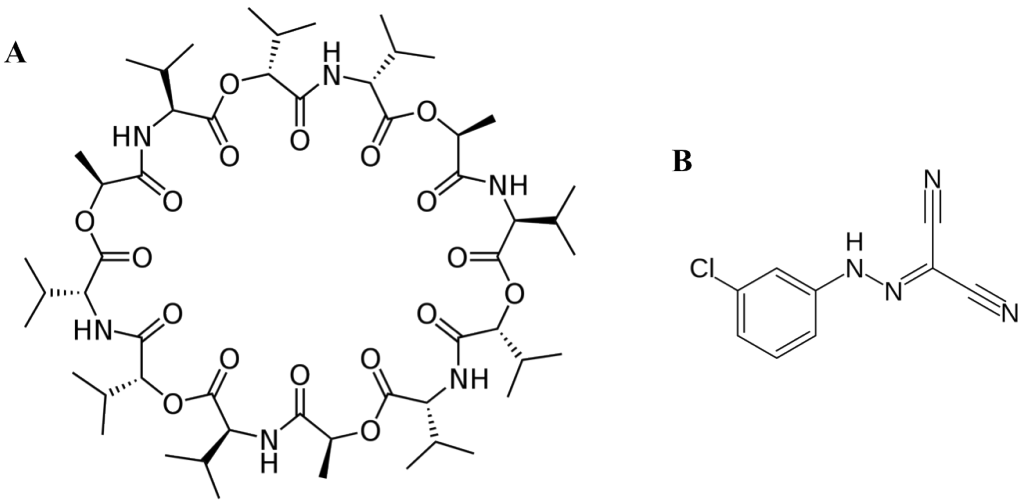


Fig. 2-11: Chemical structures of valinomycin and CCCP.
(A) Valinomycin, a cyclic depsipeptide, is a potassium-specific ionophore. It initiates $\Delta\Psi$-driven transport by creating an efflux of K^+ from the internal to the external buffer.
(B) CCCP (carbonyl cyanide m-chlorophenylhydrazone), a protonophore, makes the liposome membrane permeable to protons and neutralises the membrane potential, effectively stopping $\Delta\Psi$-driven transport.

In the counterflow assay, no potassium gradient is generated, and all ion concentrations are the same inside and out. Instead, liposomes are made with an excess of peptide substrate in the internal buffer. Like LacY (Viitanen *et al*, 1986) and FucP (Dang *et al*,

2010), POTs transport the unlabelled substrate down its concentration gradient and exchange it for labelled molecules at the other side of the membrane, resulting in accumulation of radioactive signal in the liposome lumen (fig. 2-10, B). As has been proposed for LacY (Kaback *et al*, 2001), counterflow occurs without the net transfer of protons across the membrane. Instead, the transporter, in its protonated state, binds the substrate at the luminal side and transports into the external buffer, where it binds the labelled substrate and reorients back to the initial state; the cycle is repeated until the peptide concentrations on both sides of the membrane are equilibrated. Due to the substrate being present at saturating conditions, deprotonation does not occur at the external side before the binding of the labelled substrate (Garcia *et al*, 1983), making the process solely depend on the substrate gradient. Both $PepT_{St}$ and $PepT_{So}$ perform counterflow, making this the first instance of observed bidirectional, substrate-dependent transport in POT proteins.

In conclusion, activity in the counterflow assay speaks to the transporter's ability to perform uncoupled transport and undergo the full cycle of conformational changes associated with reorientation between inward- and outward-open states, while the ΔΨ-driven assay establishes whether the transporter is able to couple peptide transport to the membrane potential.

2.3.5. Assay parameters

Internal and external buffers for both ΔΨ-driven and counterflow assays were prepared at the same pH (typically 6.5) as previously described (Fang *et al*, 2000; Dang *et al*, 2010). The buffer compositions typically used in the POT transport assays are presented below (table 2-6). A minimum of 10 μg protein was used per assay, as 10 μg was typically the amount of

protein sampled per time point. If a protein sample were to be used for more than an endpoint measurement, the amount of proteoliposomes thawed and extruded on the day of the experiment would be 10 μg multiplied by the number of intended measurements; this was done as the proteoliposomes could not be frozen and reused once extruded. All experiments in the results chapters were performed in duplicate, unless otherwise specified.

	ΔΨ-driven assay	**Counterflow**
Internal buffer	20 mM KPi, 100 mM KAc, 2 mM $MgSO_4$	50 mM KPi, 2 mM $MgSO_4$, 10 mM cold substrate (e.g. AlaAla)
External buffer	120 mM NaPIPES, 2 mM $MgSO_4$, 10 μM valinomycin	50 mM KPi, 2 mM $MgSO_4$

Table 2-6: Internal and external buffer compositions for the ΔΨ-driven and counterflow assays. Some experiments described in the thesis were performed at lower valinomycin concentrations like 2 and 5 μM (*valinomycin was typically added before the start of each assay, from a concentrated stock in DMSO).

For a typical ΔΨ-driven time course experiment (5 time points), 50 μg proteoliposomes were extruded and resuspended in internal ΔΨ buffer as described in 2.3.4, then diluted 1:50 in external ΔΨ buffer (total reaction volume 500 μl), containing 10 μM valinomycin and 15 μM ^{3}H-AlaAla (the latter added as a mixture of labelled and unlabelled peptide), at 25°C and pH 6.5. The timer was started upon addition of the liposomes, and 100 μl samples were taken at each time point and diluted in 1.5 ml $LiCl_2$ to stop the reaction. Liposomes were isolated on 0.22 μm MF-Millipore^{TM} mixed cellulose filters using a vacuum manifold and pump. The filters were washed with 2x2 ml $LiCl_2$, added to scintillation vials and measured as described in 2.3.1.

For the competition assays, ΔΨ-driven uptake of ^{3}H-AlaAla (15 μM) or ^{3}H-GlySar (50 μM) by 1 μM protein (10 μg protein per assay in 175 μl reaction volume) was measured as above in the presence of 2 mM cold substrate. A single measurement was taken after 2 min

($PepT_{St}$, GkPOT, YdgR) and 10 min ($PepT_{So}$). For the IC_{50} experiments, the proteins were assayed in the same conditions in the presence of different competitor concentrations, ranging from 0 to 10 mM.

For the kinetic analysis of ΔΨ-driven ^{3}H-GlySar uptake, time course experiments were performed as above, with GlySar concentrations ranging from 41 μM to 30 mM, in a total volume of 500 μl. Protein concentration in each experiment was 417 nM, and experiments were performed in triplicate. Transport rates were calculated at each GlySar concentration from the linear part of the curves, and the V_{max} and apparent K_m of transport were generated using the standard Michaelis-Menten model for enzyme kinetics.

To measure pH-dependent uptake of ^{3}H-AlaAla under ΔΨ-driven conditions, a different buffer system was used in order to perform the assay across a wide range of pH values (equimolar KPi and NaPi buffers inside and outside the liposomes). Specifically, proteoliposomes preloaded with 50 mM KPi at varying pH values were diluted 1:50 into 50 mM NaPi of the same pH (ranging from 5 to 8), to a total volume of 150 μl, and ^{3}H-AlaAla uptake (at 15 μM) was measured after 1 min incubation at 25°C.

For the peptide-driven counterflow assays, proteoliposomes containing internal counterflow buffer (typically 30 μg protein per assay, sufficient for 3 time points) were diluted 1:50 in external counterflow buffer (table 2-5) containing 5 μM ^{3}H-AlaAla (the total external peptide concentration was 77 μM made up from 72 μM additional cold peptide) at 25°C and pH 6.5, to a total reaction volume of 600 μl. The timer was started upon addition of the liposomes, and samples were taken at 2, 5 and 10 min ($PepT_{St}$, GkPOT) or 3, 10 and 20 min ($PepT_{So}$), as described for the ΔΨ-driven experiment above. For clarity, only single measurements were shown in the results (5 min for $PepT_{St}$ and GkPOT, 10 min for $PepT_{So}$).

2.3.6. Transport assays with the $PepT_{So}$ DEER samples

To investigate the conformational states adopted by POT family transporters, double electron-electron resonance (DEER) measurements were performed on spin-labelled $PepT_{So}$ proteins. To create sites for labelling, double cysteine mutations were introduced at the cytoplasmic and periplasmic sides of the protein, equivalent to the positions in a similar study of LacY (Smirnova *et al*, 2007); the DEER experiments were used to measure the distance distributions between the labelled residue pairs, as the protein adopts different conformational states in detergent solution and lipid bilayers (Mchaourab *et al*, 2011).

The mutant proteins were expressed as GFP-fusion constructs, purified as described in 2.2.3 and concentrated to 10 mg/ml. The concentrated sample was incubated with 10 mM DTT for 1h at RT, then run on a 5 ml HiTrap™ desalting column to remove the DTT. The reduced proteins were labelled by incubation with 1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl (MTSL) at 10:1 molar excess per cysteine, for 3h at RT. The HiTrap step was repeated to remove the unbound label, and the eluted protein was concentrated to 10 mg/ml. The labelled protein was then run on a Superdex S200 gel filtration column pre-equilibrated with reconstitution buffer (50 mM KPi, 150 mM KCl, 0.3% DM at pH 7). Collected fractions were pooled and concentrated to 10 mg/ml.

The proteins were reconstituted with the rapid dilution method, as described in 2.3.3. A 1:1000 (molar) protein:lipid ratio was used (in contrast to the 1:3000 ratio typically used for reconstitution), in order to optimise the DEER sample for small volume and high protein concentration. Proteoliposomes were prepared as described in 2.3.5. Before using the samples

in the DEER experiments, ΔΨ-driven and counterflow uptake of ^{3}H-AlaAla were assayed in highly concentrated proteoliposome samples (250 μM protein compared to 1 μM in a typical transport assay) using the same substrate concentrations as in the regular assays but very low (~15 μM) sample volumes. The purpose of these assays was to replicate the generation of the membrane potential (ΔΨ, via valinomycin) and test the effect of the peptide gradient (counterflow) in "mock DEER samples" – as a typical DEER sample used for this study is ~30 μl at a protein concentration of 1 mM (protein concentration for the mock samples could not be increased above 250 μM for the desired reaction volume, due to lipids taking up a large part of the sample volume).

The assay conditions were replicated in the DEER samples; shortly before transfer to the DEER sample tubes, 10 μl external buffer (for ΔΨ and counterflow) containing AlaAla at regular assay concentrations (unlabelled) was added to the proteliposome pellet and mixed with it to create a suspension, initiating transport. Control samples were prepared in parallel containing no peptide in the external buffer (ΔΨ-driven) and the internal buffer respectively (counterflow). The sample was immediately transferred to a 3 mm high precision quartz tube, flash-frozen and stored in liquid N_2. The DEER experiments were performed by Patricia Dijkman on a Bruker Elexsys 680 spectrometer.

2.4. Binding experiments

2.4.1. Isothermal titration calorimetry

ITC measures the binding of ligands to large macromolecules by recording the changes in the electrical current needed to maintain constant temperature between a sample cell (protein and ligand in buffer solution) and a reference cell (buffer) caused by the increase or decrease in temperature upon ligand binding (Freire *et al*, 1990).

To measure the binding of AlaAla and AlaTyr to $PepT_{So}$, an ITC microcalorimeter was used. At 25°C, the sample cell was equilibrated in 1x PBS and 0.03% DDM (pH 7.4) (3 x 250 µl washes), then 250 µl $PepT_{So}$ in the same buffer at 30 µM concentration was added to the cell. The peptide was then titrated into the sample cell in 16 x 2.5 µl increments from an 80 µl stock solution of varying concentration (30 µM, 300 µM and 1 mM), with a spacing of 180 s between injections. The data was analysed using the Origin™ software.

2.4.2. Scintillation proximity assay

SPA is a highly sensitive method for measuring the binding of radiolabelled ligands, involving attaching the protein to scintillating beads via a histidine tag; the beads emit light in the proximity of the ligand (when it is bound to the protein), but do not detect the unbound ligand free in solution, resulting in highly specific detection of protein-ligand interactions (Udenfriend *et al*, 1985).

For the assay, $PepT_{So}$ was expressed with a His_6 tag and purified as described in 2.2.5. Before the assay, different amounts of YSi-copper SPA scintillating beads (PerkinElmer) were centrifuged for 1 min at 7,000 g and washed three times with 1 ml gel filtration/crystallisation buffer (20 mM Tris HCl pH 7.5, 150 mM NaCl and 0.03% DDM), then resuspended in the same buffer to varying working concentrations (2, 8 and 16 mg/ml). The beads were distributed to a 96-well SPA plate (Isoplate-96, PerkinElmer) and the protein was added to varying concentrations (200 nM, 2 μM) and incubated with the beads (150 μl total volume per well) for 1 h at RT. Increasing concentrations of ^{3}H-AlaAla were then added to the wells, ranging from 15 nM to 5 mM (total concentrations made up using radiolabelled as well as cold peptide). Control rows on the plate included the same reaction mixtures with no protein. The plate was sealed with aluminium foil and incubated overnight at 4°C. The following day, the plate was centrifuged at 4,000 g to spin down condensation water, and the signal was measured in a microplate scintillation counter. For this, a protocol was set up to record the average of 10 scintillation count measurements once every 30 min at RT, over 24 h. A parallel experiment was set up in which the plate was measured immediately after setting up the reactions.

Alternatively, $PepT_{So}$ was reconstituted into liposomes with the rapid dilution method, and used to set up a similar assay. Varying amounts of reconstituted protein were immobilised on the SPA beads as described above, and the ΔΨ-driven assay conditions were replicated in the wells of the SPA plate (beads were resuspended in external ΔΨ buffer and 10 μl valinomycin was added to the wells to initiate the reaction upon adding a 1:50 dilution of proteoliposomes in internal ΔΨ buffer). The signal was measured approx. 10 min after the start of the reaction.

To estimate the efficiency of $PepT_{So}$ binding to the SPA beads (both in liposomes and detergent solution), the protein and beads were incubated together for 1h at RT. The entire pellets (containing the beads and immobilised protein) were centrifuged and loaded on a Tris-Gly gel, alongside 10 µl of the 200 µl supernatant and protein standards of known amount (fig. S8).

2.4.3. NMR spectroscopy

Despite issues like large particle size and insufficient stability, membrane proteins have recently been the objects of NMR spectroscopy experiments (Nietlispach and Gautier, 2011), with a recent example from the transporter field (Morrison and Henzler-Wildman, 2011). $PepT_{So}$, $PepT_{St}$ and GkPOT were expressed in ^{15}N isotope-enriched minimal media and purified with the standard protocol for GFP-fusion proteins, as described in 2.2.2 and 2.2.3. At the final gel filtration step, ^{15}N-labelled proteins were exchanged into NMR buffer (50 mM NaPi, 20 mM NaCl and 0.03% DDM at pH 6.5) and concentrated to make up ~350 µl samples at concentrations ranging from 80 to 150 µM. As a calibration standard, the samples were supplemented with 2 mM 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) in D_2O solution before the experiments.

Alternatively, ^{15}N-$PepT_{So}$ was incorporated into bicelles via liposome reconstitution, as previously described (Morrison and Henzler-Wildman, 2011). In the process, the protein is first reconstituted into liposomes made up of long-chain phospholipids, then the bicelle is formed by adding short-chain phospholipid that disrupts the bilayer and forms a "capping" structure around it (fig. 2-12).

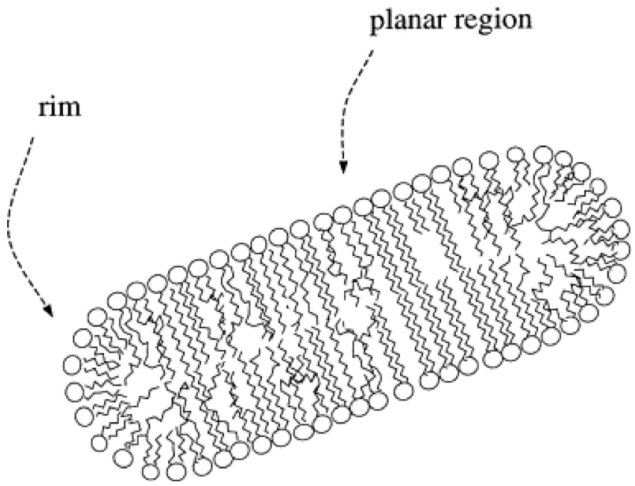


Fig. 2-12: Schematic representation of a bicelle. The planar region contains long chain lipids, while the short chain lipids form the rim region "capping" the bicelle (Whiles *et al*, 2002).

Long-chain lipids 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) and 1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine (DPPC) (containing C_{14} and C_{16} chains respectively) were dissolved in 50 mM KPi (pH 7) and used to make large multilamellar liposome vesicles through sonication and freeze-thawing as described in 2.3.3. ^{15}N-PepT_{So} was reconstituted via rapid dilution using a protein:lipid molar ratio of 1:320 (for example 11.5 mg DMPC per 3 mg protein). Proteliposomes were isolated by ultracentrifugation (130,000 g for 3h at 4°C), and bicelles were formed by resuspending the liposomes in 50 mM KPi containing the short-chain lipid 1,2-diheptanoyl-*sn*-glycero-3-phosphocholine (DHPC) in a 1:3 molar ratio (DMPC:DHPC), equivalent to a q-value of 0.3, which has been previously found to best reflect the properties of the cellular membrane (Glover *et al*, 1990). A 350 µl NMR sample of bicelle-reconstituted PepT_{So} would thus contain 150 µM protein, 50 mM long-chain lipid and 150 mM DHPC. Alternatively, DM was used in one experiment instead of DHPC.

For both detergent and bicelle samples, 2D heteronuclear multiple quantum correlation (HMQC) spectra were collected at 650 and 950 MHz on Omega spectrometers, using the SoFast protocol optimised for rapid data collection (Schanda and Brutscher, 2005). PepT_{So} spectra (in the apo state and in the presence of 80 mM concentrations of dipeptides

AlaTyr and AlaPhe) were collected at 40°C, while GkPOT spectra were collected at 47°C from 256 scans. To measure binding of AlaPhe to GkPOT, increasing amounts of the dipeptide were added to the protein to concentrations ranging from 50 μM to 20 mM, and spectra were collected in the same conditions after each titration point. The HMQC protocol was designed and the experiments were performed by Dr Jason Schnell and Dr Jolyon Claridge. The spectra were visualised and processed in NMRView.

Chapter 3: Screening bacterial POT family transporters for structural studies

Membrane proteins represent between 20 and 30% of the proteome of most organisms (Krogh *et al*, 2001) and up to 60% of identified molecular targets for all classes of therapeutic drugs in humans (Overington *et al*, 2006). Despite their importance, less than 1% of the structures in the PDB are of membrane proteins, currently barely over 400 individual protein structures from a total of over 80,000 solved (http://blanco.biomol.uci.edu/mpstruc/). Several challenges account for membrane proteins being such difficult targets for structural studies. Overexpression commonly produces low yields, due to toxic effects on the host cell, incorrect membrane insertion or non-optimal lipid composition; this is all the more true for eukaryotic proteins. Furthermore, the choice of detergent is crucial for membrane extraction and stability during purification, as well as successful crystallisation. Lastly, poor diffraction and anisotropy are frequent issues with membrane protein crystals, creating a number of difficult bottlenecks (Bill *et al*, 2011). In addition, thermal instability and the large molecular weights of detergent-protein micelles often render membrane proteins intractable to other methods such as NMR spectroscopy.

Various strategies have been proposed and implemented to overcome these challenges. The so-called "Walker strains" of *E. coli* (C41 and C43), obtained empirically from mutant BL21(DE3) cells, proved beneficial for incorporating toxic proteins into the membrane as they delay the onset of transcription and prevent overcrowding the cells with overexpressed

protein (Miroux and Walker, 1996). Other studies sought to improve the balance between overexpression and proper membrane insertion under conditions of rapid cell growth and division by introducing new strains (Wagner *et al*, 2008). Fusing membrane protein constructs to a C-terminal GFP molecule was initially used as a reporter of soluble protein localisation (Waldo *et al*, 1999; Feilmeier *et al*, 2000), and later became a common strategy for controlling proper insertion and folding of membrane proteins (as GFP fluoresces in the cytoplasm but not the periplasm), as well as reporting on expression yields (Drew *et al*, 2001; 2002). Expression and purification protocols for membrane proteins were elaborated based on the GFP fusion approach, in both bacterial and yeast expression systems (Drew *et al*, 2006; Newstead *et al*, 2007). Building on this, fluorescence size exclusion chromatography (FSEC) was established as a means to ascertain the size and monodispersity of expressed GFP-fusion constructs (Kawate and Gouaux, 2006).

A great deal of effort has gone into optimising membrane protein stability. While alanine scanning revealed some thermostabilising mutations in G-protein coupled receptors (Serrano-Vega *et al*, 2008; 2009) and thermostability assays were developed to guide the purification of proteins from different families (Alexandrov *et al*, 2008), most studies have been focused on choosing the right detergents and buffers for membrane solubilisation, purification and crystallisation. Most available X-ray structures of α-helical membrane proteins were obtained from crystallisation with maltopyranoside and glucopyranoside detergents of various chain lengths. In particular, the long-chain (C_{12}) detergent DDM (n-dodecyl-β-D-maltopyranoside) is the most popular choice for purifying and crystallising transporters, while shorter-chain detergent OG (n-octyl-β-D-glucopyranoside) has proven the most successful for crystallising channels (Newstead *et al*, 2008).

Detergents forming large micelles – for example, a DDM micelle is typically 70-75 kDa on average (Strop and Brunger, 2005) – provide stability but can have a negative impact on the diffraction quality as they provide limited hydrophilic space for the formation of crystal contacts (Bill *et al*, 2011). However, in the case of transporters (both H^+- and Na^+-coupled), good protein stability (measured by its half-life in a certain detergent) directly correlates to high resolution of the X-ray diffraction data (Sonoda *et al*, 2011), so detergents that increase protein stability are likely to favour the formation of well diffracting crystals. Consequently, DDM is currently the best candidate for both transporter purification and high resolution structure solving, while transporter stability in other detergents can be used as an indicator of the ultimate crystal quality. Several membrane protein crystallisation screens were developed using the available protein structures to analyse the relative success of detergents and precipitants, salts and additives (Newstead *et al*, 2008; Parker and Newstead, 2012). These screens include a variety of conditions likely to facilitate crystallisation of α-helical membrane proteins, such as detergents, precipitants, salts and buffers, offering a starting point for preliminary crystallisation trials.

Other strategies for making membrane proteins more amenable to structural studies include the use of branched-chain detergents that can assist with crystal contact formation (Hong *et al*, 2010) or detergents with added polar side groupd to increase their amphiphilic character (Zhang *et al*, 2007; 2011). Adding phospholipids to the protein-detergent solution resulted in better diffraction of LacY crystals (Guan *et al*, 2006), while LeuT was trapped in a different conformation by co-crystallisation with a monoclonal antibody (Krishnamurthy and Gouaux, 2012). Fab fragments have been used to crystallise other transporters as well, as they can not only assist with crystal contacts, but also provide phasing information as they introduce a higher degree of order in the crystals and therefore contribute substantially to the overall scattering (Rasmussen *et al*, 2007). More frequently however, the phasing problem is

overcome introducing anomalous scatterers through the use of heavy metal atoms such as mercury, platinum, lead and gold (Sonoda *et al*, 2010; Parker and Newstead, 2013). Data collection on microfocus beamlines can compensate for small crystal size and radiation damage (Carpenter *et al*, 2008), while data from poorly diffracting crystals can sometimes benefit from using low resolution molecular replacement in the initial stages of phasing (Pedersen *et al*, 2010). Additionally, the *in meso* method (commonly referred to as lipid cubic phase crystallisation or more generally crystallisation in lipid environments) has been an importance source of high resolution membrane protein structures in recent years (Caffrey *et al*, 2012). Despite these strategies, membrane proteins remain difficult targets for structural studies. Arguably the most efficient approach starts with the screening multiple homologous proteins for expression, followed by detergent screening and purification, narrowing down the targets to a few that can then enter crystallisation screening and optimisation.

3.1. Expression, purification and crystallisation of POT homologues

Several prokaryotic peptide transporters were chosen for crystal screening, based on their sequence similarity to the human PepT1 and PepT2 (as well as $PepT_{So}$ and $PepT_{St}$, for which crystallisation had been optimised and native data had been collected to 3.6 and 3.3Å resolution respectively prior to the start of this DPhil project). Sequence similarity between the bacterial POTs chosen for screening and their human homologues were generally between

30 and 45% (table 3-1), with $PepT_{So}$ itself showing the highest sequence conservation. The purpose of the screening was to establish which other transporters could be overexpressed in *E. coli* and under what conditions, to carry out their purification in different detergents, and potentially discover optimal crystallisation conditions and determine crystal structures in different conformations that would provide further insight into the structural aspects of the transport cycle.

Protein	Organism	Sequence similarity to human proteins	
		PepT1	PepT2
$PepT_{So}$	*Shewanella oneidensis*	43.3%	46.6%
$PepT_{St}$	*Streptococcus thermophilus*	36.9%	35.8%
$PepT_{Xc}$	*Xanthomonas campestris*	41.9%	41.9%
$PepT_{Bc}$	*Bacillus cereus*	40.5%	39.5%
$PepT_{Bs}$	*Bacillus subtilis*	37.6%	36.8%
$PepT_{Vc}$	*Vibrio cholerae*	33.8%	32.7%
YdgR	*Escherichia coli*	35.3%	34.1%

Table 3-1: Prokaryotic POT family transporters used for screening and their sequence similarity to the human transporters.

The genes encoding the POT family transporters were cloned into pWaldo plasmids as C-terminal GFP-fusions. All proteins were successfully expressed. *E. coli* C43(DE3) cells were typically grown in Terrific Broth medium, but a small-scale expression trial with selected targets resulted in slightly higher yields when using a variant of the PASM-5052 autoinduction medium (Sreenath *et al*, 2005), when compared to TB (fig. 3-1). Expression

trials with $PepT_{So}$ in the ZYM-5052 autoinduction medium (Studier, 2005) did not improve on the expression in TB. Additional expression trials were performed in the *pLemo* expression system (Wagner *et al*, 2008), which uses an L-rhamnose inducible promoter (Muiry *et al*, 1993) to regulate the activity of the T7 polymerase, in turn preventing overcrowding of the Sec translocon with overexpressed protein. However, all peptide transporters tested gave poor expression in the *pLemo* system when using a gradient of 0.1 to 2mM L-rhamnose for induction, compared to the expression yields in TB.

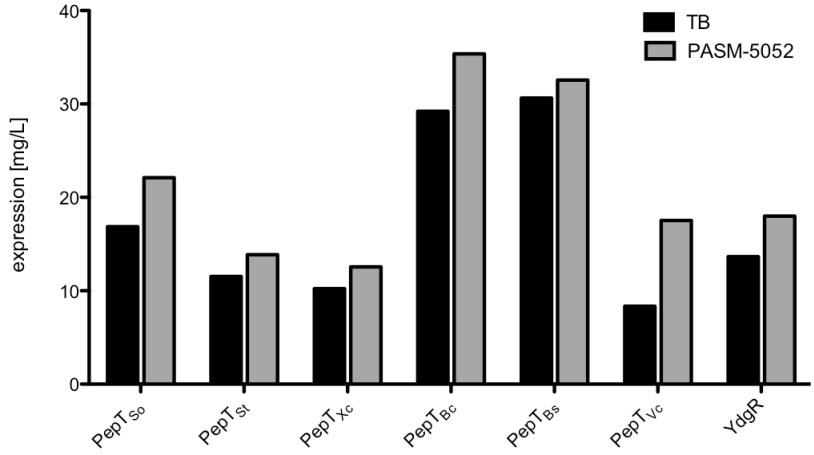


Fig. 3-1: Expression levels of bacterial peptide transporters in TB (IPTG-induced) and PASM-5052 autoinduction medium. pWaldo-GFP constructs were grown at 37°C to OD_{600} = 0.6-0.8 and induced (TB samples only). Cultures were grown overnight at 25°C. Expression levels were converted from relative GFP fluorescence units (RFU) per 1ml of cell culture (normalised to OD=8) to mg/L according to a standard curve.

As expression levels were high in the initial trials, all proteins were expressed in large-scale cultures (4 to 6 L). Membranes were isolated, the membrane protein fractions extracted, and the proteins were purified as described in the methods chapter. All proteins were stable in DDM throughout purification, were over 90% pure on TrisGlycine gels and showed monodisperse peaks on the gel filtration trace, as shown here for the typical example of $PepT_{So}$ (fig. 3-2).

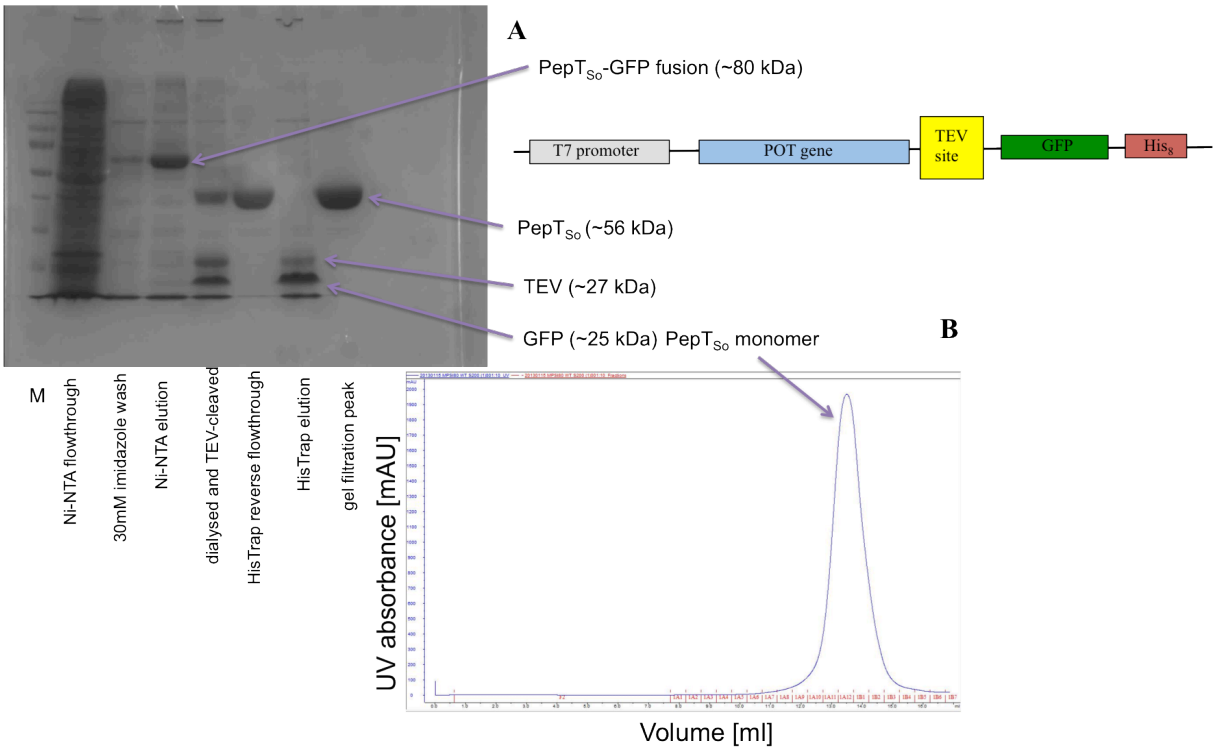


Fig. 3-2: Purification of $PepT_{So}$ from *Shewanella oneidensis*.
(A) TrisGlycine gel showing the purification steps (schematic of the construct on the right). The detergent-extracted membrane protein fraction is run on a Ni-NTA resin and the column is washed with 8CV PBS/DDM buffer at 30mM imidazole; bound GFP-fusion protein is eluted with 250mM imidazole and the GFP tag is cleaved with TEV protease overnight while dialysing into Tris/NaCl/DDM buffer; the cleaved protein is run on a HisTrap column and collected, while the bound GFP (His-tagged) is eluted with 500mM imidazole; the cleaved protein is concentrated and run on a gel filtration column.
(B) Gel filtration profile. Typical of most peptide transporters tested, $PepT_{So}$ shows a monodisperse peak corresponding to a monomer.

For $PepT_{St}$, other detergents were tested, by exchanging the protein from DDM solution into the desired detergent at the final step on the gel filtration column. The detergents used were: the maltosides DM and NM and the glucoside OG, all three of which had been successful in crystallising secondary active transporters (Sonoda *et al*, 2010), as well as the amphiphilic detergent Façade EM (Zhang *et al*, 2007) and the neopentyl glycol detergent Cymal-5. In all these cases, the protein remained stable during the exchange and displayed gel filtration traces similar to that in DDM (fig. 3-3). In addition, $PepT_{St}$ was detergent-extracted and purified in DM, NM and OG in parallel to the DDM purification. The DM-purified protein precipitated heavily after GFP cleavage and remained bound to the reverse affinity column. The protein was more stable in NM, but the cleaved GFP tag separated poorly from the cleaved protein during the reverse step. The OG purification fared better, and $PepT_{St}$

formed crystals in the same conditions as when exchanged into OG during the final gel filtration step; the crystals diffracted to similar resolution in both cases (8-9 Å), so there is no evidence that carrying out the entire purification in a given detergent yields better results than detergent exchange after purification in DDM.

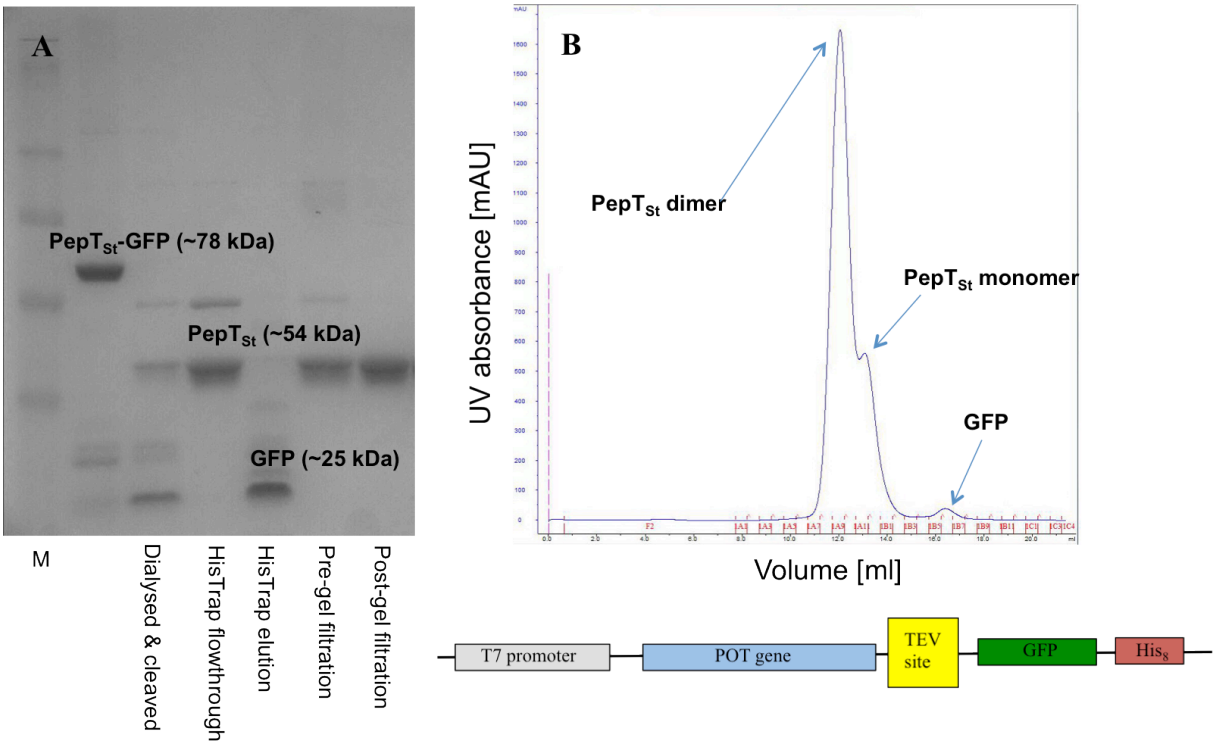


Fig. 3-3: Purification of $PepT_{St}$ from *Streptococcus thermophilus*.
(A) TrisGlycine gel showing the purification steps. The GFP-fusion protein is eluted from the Ni-NTA column, dialysed and cleaved with TEV protease overnight; the GFP tag binds to the HisTrap affinity column and the cleaved protein is in the flowthrough; the protein is concentrated and run on a gel filtration column.
(B) Gel filtration profile, depicted here for a sample exchanged from DDM into DM. In contrast to most peptide transporters tested, $PepT_{St}$ elutes as a mixture of two species.

Like $PepT_{So}$ (fig. 3-2), most proteins in this study elute as monodisperse species from the size exclusion column. $PepT_{St}$ is the sole exception as it forms two peaks, suggesting the existence of two oligomeric states in the detergent solution (fig. 3-3). To establish the molecular weights of these species and the relative preponderance of protein and detergent in each peak, refractive index and right angle scattering (RALS) profiles were determined for $PepT_{So}$ and $PepT_{St}$ in DDM solution, as described in the methods chapter. In both profiles (fig.

3-4), the main peaks (one for $PepT_{So}$ at 15ml, two for $PepT_{St}$ at 13.5ml and 15ml) absorb UV light at 280nm (due to aromatic amino acids), indicating the presence of protein; a smaller peak to the right of the main ones is identified as the uncomplexed detergent micelles, as it causes a shift in the refractive index but does not contain protein.

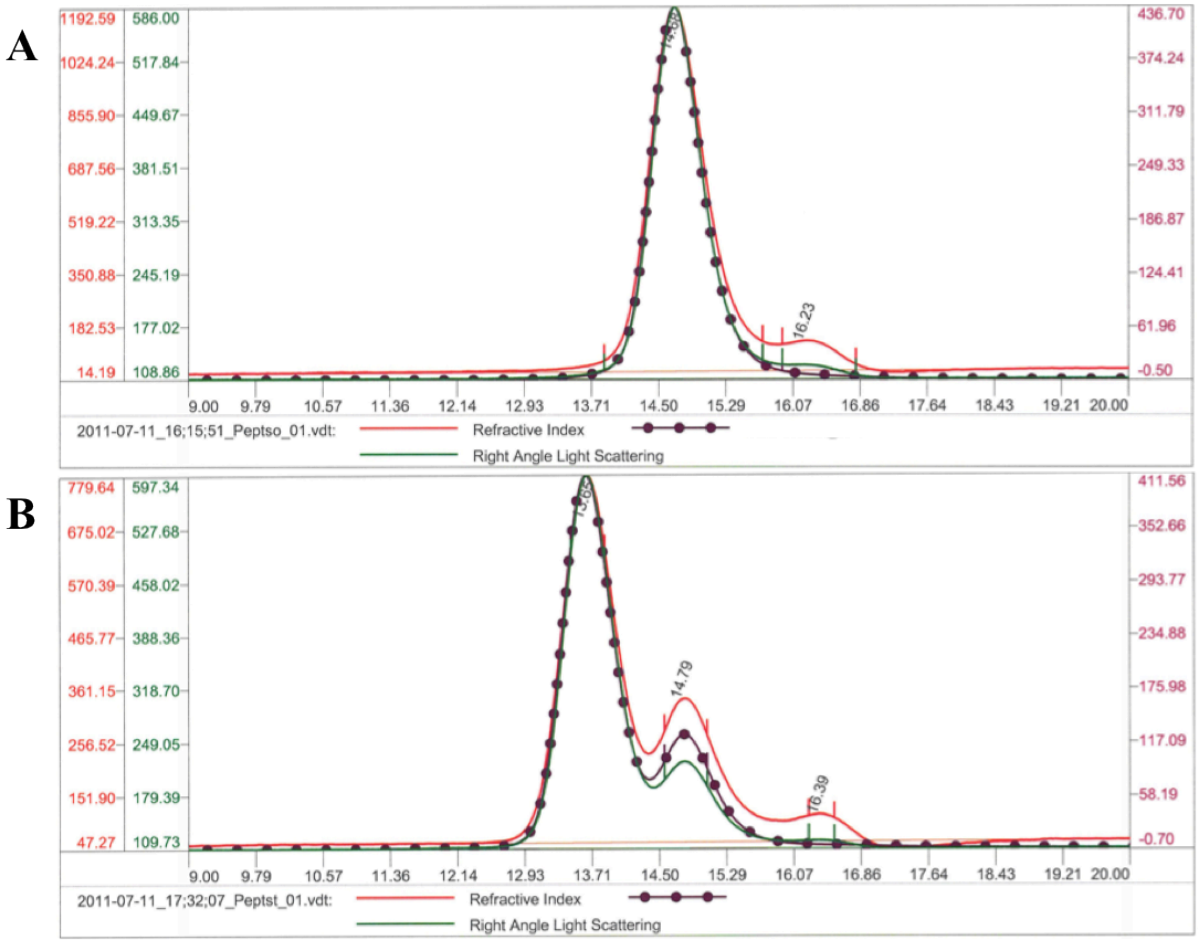


Fig. 3-4: UV, refractive index and right angle light scattering profiles of $PepT_{So}$ (A) and $PepT_{St}$ (B). Individually analysed peaks are marked with vertical lines on each trace. Proteins were run on a 25ml gel filtration column as 2mg/ml samples in 20mM Tris HCl pH 7.5, 150mM NaCl and 0.03% DDM. Refractive index is shown in red, RALS in green, UV in purple.

The average molecular weight (MW) was estimated for each peak to less than 1% error, and the fraction of the total MW represented by the protein (MW_p) was calculated as described in the methods chapter. Thus, the total MW corresponds to the average weight of the protein-detergent micelle, while the MW_p is the molecular weight of the protein species (table 3-2). The results indicate that $PepT_{So}$ exists as a monomer in DDM (15 ml peak), whereas the $PepT_{St}$ sample contains a mixture of dimer (13.5 ml) and monomer (15 ml), the

former being the predominant species (the estimated values agree with the actual MW of the two proteins). Moreover, the presence of the detergent increases the size of the monomeric particle approx. 3-fold and that of the dimer approx. 2.5-fold. The $PepT_{St}$ dimer could be indicative of a functional role for dimerisation in the membrane, however the presence of all other tested transporters as monomers would seem to suggest it is more likely a consequence of non-specific interactions between $PepT_{St}$ monomers in solution.

Protein	1st peak (13.5 ml)		2nd peak (15 ml)		Actual MW (monomer) [kDa]
	MW [kDa]	MW_p [kDa]	MW [kDa]	MW_p [kDa]	
$PepT_{So}$	-	-	154	53	57
$PepT_{St}$	267	113	164	52	53

Table 3-2: Estimated molecular weights of gel filtration species for $PepT_{So}$ and $PepT_{St}$. Molecular weights of the protein-detergent micelle (MW) and the protein alone (MW_p) were calculated based on the refractive index and static light scattering profile of each protein.

After crystallisation trials and optimisation, $PepT_{So}$ formed crystals in Façade EM under different buffer conditions (sodium acetate at pH 4 to Tris at pH 8), largely deviating from the standard crystallisation condition in DDM (MES at pH 6.5); however, none of these diffracted farther than 4.5Å (table 3-3). $PepT_{St}$ crystals grew in Cymal-5, $C_{12}E_8$, NM and OG, but none of them improved on the diffraction of the DDM crystals. Of the other proteins, $PepT_{Xc}$, $PepT_{Bs}$ and $PepT_{Vc}$ produced crystals, but they were either not reproducible or poorly diffracting. Most crystals grew in PEG400 (concentrations between 26 and 33%), with the exception of the $PepT_{St}$/OG crystals in higher molecular weight PEG4000.

Protein	Detergent and condition	Crystals	Best diffraction
$PepT_{So}$	Façade EM (33% PEG400, 0.1 M NaCl, 0.1 M $CdCl_2$, 0.1 M trisodium acetate pH 4) Façade EM (30% PEG400, 0.1 M NaCl, 0.01 M $CdCl_2$, 0.1 M Tris pH 8)	yes	5.5Å (pH 4 crystals) 4.5Å (pH 8 crystals)
$PepT_{St}$	DDM (26% PEG400, 0.03 M $MgCl_2$, 0.001 M $CdCl_2$, 0.1 M MES pH 6.5)	yes	3.3Å
	DM	no (unstable)	-
	NM (33% PEG400, 0.1 M NaCl, 4% v/v ethylene glycol, 0.1 M MES pH 6.5)	yes (not reproducible)	-
	OG (12% PEG4000, 0.2 M $MgCl_2$, 0.1 M Tris pH 8.5) OG (11% PEG8000, 0.1 M Tris pH 8.5)	yes	weak (8-9Å)
	LDAO	no (unstable)	-
	$C_{12}E_8$ (26% PEG400, 0.1 M NaCl, 0.001 M $CdCl_2$, 0.1 M Tris pH 8)	yes	weak (9Å)
	Cymal-5 (27% PEG400, 0.1 M NaCl, 0.1 M MES pH 6.5)	yes	weak (13Å)
	Façade EM	no	-
$PepT_{Xc}$	DDM (33% PEG200, 0.02 M NaCl, 0.02 NaAc pH 4)	yes (not reproducible)	-
$PepT_{Bc}$	DDM	no	-
$PepT_{Bs}$	DDM (27% PEG400, 0.1 M KCl, 0.2 M $MgCl_2$, 0.025 M MES pH 6.5)	yes	Weak (9-10Å)
$PepT_{Vc}$	DDM (32% PEG400, 0.05 M NaCl, 0.04 M $MgCl_2$, 0.1 M HEPES pH 7.5)	yes	none
YdgR	DDM	no	-

Table 3-3: Detergent screening of peptide transporters for crystallisation. DDM was used throughout the purification, and proteins were exchanged into different detergents during the final gel filtration step. All crystals grew at 19°C in hanging drops. Diffraction data was collected at Diamond Light Source and the ESRF.

Varying crystallisation conditions for $PepT_{St}$ revealed that a cadmium (Cd^{2+}) salt was required for the crystal to form. Some of the other POT crystals also formed in conditions with either cadmium, zinc or magnesium salts present, suggesting a role for divalent metal ions in the formation and/or stabilisation of crystal contacts.

3.2. $PepT_{St}$ crystallisation

An early part of the bacterial POT project dealt with optimising crystallisation of wild type $PepT_{St}$ for structure determination, and producing heavy atom data for phasing. Phasing of the $PepT_{St}$ datasets was performed by Dr Simon Newstead using MIRAS (multiple isomorphous replacement with anomalous scattering), with the aid of heavy atom data collected from selenomethionine (SeMet)-labelled and mercury-derivatised crystals. For the former, wild type protein was expressed in SeMet-labeled minimal autoinduction medium (Studier *et al*, 2005). After purification in DDM, $PepT_{St}$-SeMet crystals grew in the standard condition at 19°C. Interestingly, the SeMet protein eluted as a monomer from the gel filtration column, contrasting with the dimer/monomer mixture of the unlabelled wild type – a possible indication of methionines being involved in forming the $PepT_{St}$ dimer in solution. $PepT_{St}$ (not labelled) was also crystallised with selenium-labelled DDM after exchange into Se-DDM buffer on the gel filtration column; the protein precipitated in the standard crystallisation condition and no other crystals grew in any condition in the MemGold screen.

Several Cys mutants of $PepT_{St}$ were produced for mercury derivatisation, as the native protein has no cysteines (table S1). Cysteine mutations were first introduced at or near the positions where native cysteines are found in $PepT_{So}$ (Met96, Val256), followed by several mutations at additional positions. Derivatisation with methylmercury chloride (MMC) yielded several crystals and data was collected for these. The alternative method (soaking crystals in mercury acetate) did not prove successful, as it reduced diffraction quality compared to the MMC-derivatised crystals; the same was true of adding large anomalous scatterers like magic triangle (I3C) and tantalum bromide to the crystals. Of the mutants purified, M96C/V256C, S225C and S130C/M424C crystallised after derivatisation. In contrast, the A307C, S353C and S449C mutations destabilised the protein, causing heavy precipitation during purification (possibly due to increased disorder in the transporter's more flexible C-terminal domain), and did not form crystals.

In addition to the heavy atom crystallisation, a number of methods were used to improve crystallisation conditions and/or the diffraction quality of existing crystals of wild type $PepT_{St}$. Lysine methylation, crystal dehydration, re-lipidation and under oil crystallisation are explained in the methods chapter. Generally a harsh method, lysine methylation of purified $PepT_{St}$ did not destabilise the protein, however it failed to produce crystals in either the standard condition or the crystallisation screen. The second method, crystal dehydration, was applied to crystals of $PepT_{St}$ wild type and the M96C/V256C variant. Gradual dehydration by increasing the precipitant concentration in the crystallisation well was generally well taken, but did not improve on resolution. Similarly, drying the crystals under a gas stream followed by quick flash-freezing seemed to have no significant effect, crystals diffracting to only 4.5Å. Harsher, more drastic dehydration resulted in crystal lattice disorder and worsened resolution.

To determine the influence of lipids on crystallisation, purified $PepT_{St}$ in DDM was re-lipidated by incubation with three lipid types, as described in the methods chapter: *E. coli* polar lipid extract (EPL) (containing PE, PG and CL), eukaryotic membrane lipid phosphatidylcholine (PC) and a 3:1 (w/w) mixture of the two, followed by crystallisation screening (MemGold screen). The lipids were chosen due to their use in the functional reconstitution of $PepT_{St}$ (as described in the next chapter). Crystals of $PepT_{St}$ in EPL and EPL-PC grew in the standard crystallisation condition at 19°C and were typically rod-shaped albeit of smaller size than usual. Selected crystals in both conditions gave good diffraction, comparable with or improving over typical native $PepT_{St}$ datasets, suggesting phospholipid addition is beneficial to $PepT_{St}$ crystal quality, despite not producing new crystal forms. In addition, the EPL and EPL-PC lipid mixes were added in the same ratios to $PepT_{St}$ in other detergents (NM, OG, Cymal-5 and $C_{12}E_8$), but none of these samples crystallised.

$PepT_{St}$ was also co-crystallised with dipeptide substrate dialanine (AlaAla) in the standard condition. The crystals diffracted to 3.6Å, but no density for the peptide was identified in the binding site. Finally, under-oil crystallisation trials of $PepT_{St}$ using Al's oil, as well as a lipidic cubic phase crystallisation trial with a monoolein-prefilled plate were set up, but neither method yielded crystals.

3.3. Conclusions

The road to understanding the structural biology of membrane proteins is a difficult one, marked by numerous challenges at every step. Toxic effects on the expression system, insufficient yields, the harsh process of detergent extraction, instability due to hydrophobicity, and large molecular weights rendering proteins incompatible with certain experimental methods – these are as many potential bottlenecks associated with the study of membrane proteins. These challenges are perhaps best overcome using a high throughput approach, in which several homologous protein targets are screened for expression, detergent solubilisation, purification and crystallisation, using the screening process to select valuable candidates and learn how to circumvent the pitfalls along the way. The approach is exemplified by this chapter, which details the process of identifying potential candidates for structural studies of POT proteins, as well as optimising $PepT_{St}$ crystallisation and obtaining new structural information for crystallographic analysis.

Expression as C-terminal GFP-fusion constructs was successful for all peptide transporters, producing consistently high yields and stable, easily quantifiable protein in both rich and minimal media. DDM was found to be the best detergent for membrane extraction and purification, but also crystallisation, in agreement with previous observations on other transporters (Newstead *et al*, 2008; Sonoda *et al*, 2010). However, both $PepT_{St}$ and $PepT_{So}$ maintain stability in other detergents such as glucosides or neopentylglycols after exchange on the gel filtration column; some of them are compatible with crystallisation, although crystals of neither transporter diffract to higher resolution than the DDM crystals.

Addition of bacterial polar lipids to $PepT_{St}$ resulted in slightly improved resolution, reinforcing the role of phospholipids in creating a stabilising environment for membrane

proteins. Again, the presence of DDM favoured crystallisation in the lipid environment. Systematic controlled dehydration of $PepT_{St}$ crystals was performed to varying results; rapid dehydration was detrimental to crystal quality, while the gradual method had no noticeable effects on diffraction. Finally, SeMet incorporation and MMC derivatisation were successfully used to improve phasing of the native $PepT_{St}$ dataset.

Chapter 4: Structural analysis of $PepT_{St}$

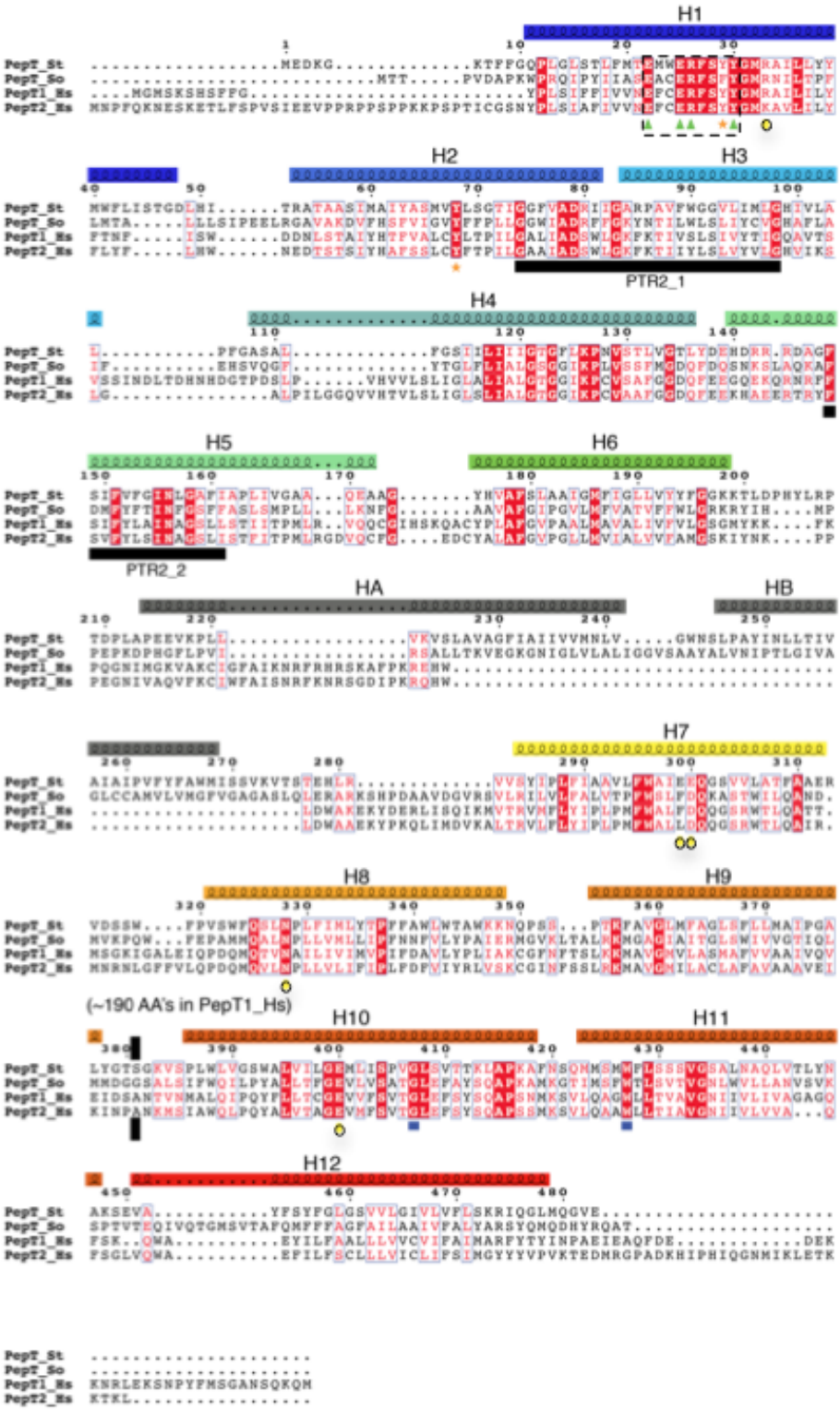


Fig. 4-1: Identity-based sequence alignment of bacterial and human POT members: $PepT_{St}$ (Uniprot:Q5M4H8), $PepT_{So}$ (Q8EKT7), hPepT1 (B2CQT6) and hPepT2 (Q16348), generated using MAFFT and JalView. Identical residues in red, transmembrane helices marked as coils above the sequence and coloured blue to red. The ExxERFxYY motif is marked with a discontinuous line, and the PTR2 motifs with black bars. The extracellular segment in PepT1 and PepT2 between TMs 9 and 10 was omitted for clarity (Solcan *et al*, 2012).

$PepT_{St}$ is an oligopeptide transporter from Gram-positive bacterium *Streptococcus thermophilus*. Its amino acid sequence is 50% identical to that of $PepT_{So}$, and it is approximately 30% identical to the sequences of both hPepT1 and hPepT2. In all four proteins, clusters of particularly well conserved residues can be identified, such as the POT family signature ExxERFxYY motif on H1 and the FYxxINxG motif on H5 (albeit with the tyrosine replaced by a valine in $PepT_{St}$). The PTR2_1 motif between H2 and H3 (GxxxADxxxGKxxTI) is also seen in all four proteins, but the last five residues are not conserved in $PepT_{St}$ (fig. 4-1). Compared to other POT homologues screened for crystallisation, $PepT_{St}$ yielded the best results (reproducible, well diffracting crystals). Its structure was solved and revealed a novel snapshot of the transport mechanism in the POT family.

4.1. Data collection and crystallography

Wild type $PepT_{St}$ crystals grew in the standard condition, at 0.1 M MES/NaOH pH 6.5, 0.03 M $MgCl_2$, 1 mM $CdCl_2$ and 26% PEG400 (Solcan *et al*, 2012). Crystals were generally rod-shaped or polyhedral, around 200 μm in length (fig. 4-2, A). They diffracted on average to 3.6-3.5Å (in one direction, as diffraction was anisotropic), with a previous crystal (the main source for the 3D model) reaching 3.3Å (table 2-1). $PepT_{St}$ forms an inverted dimer in the asymmetric unit, and the space group was identified as orthorhombic $P2_12_12_1$.

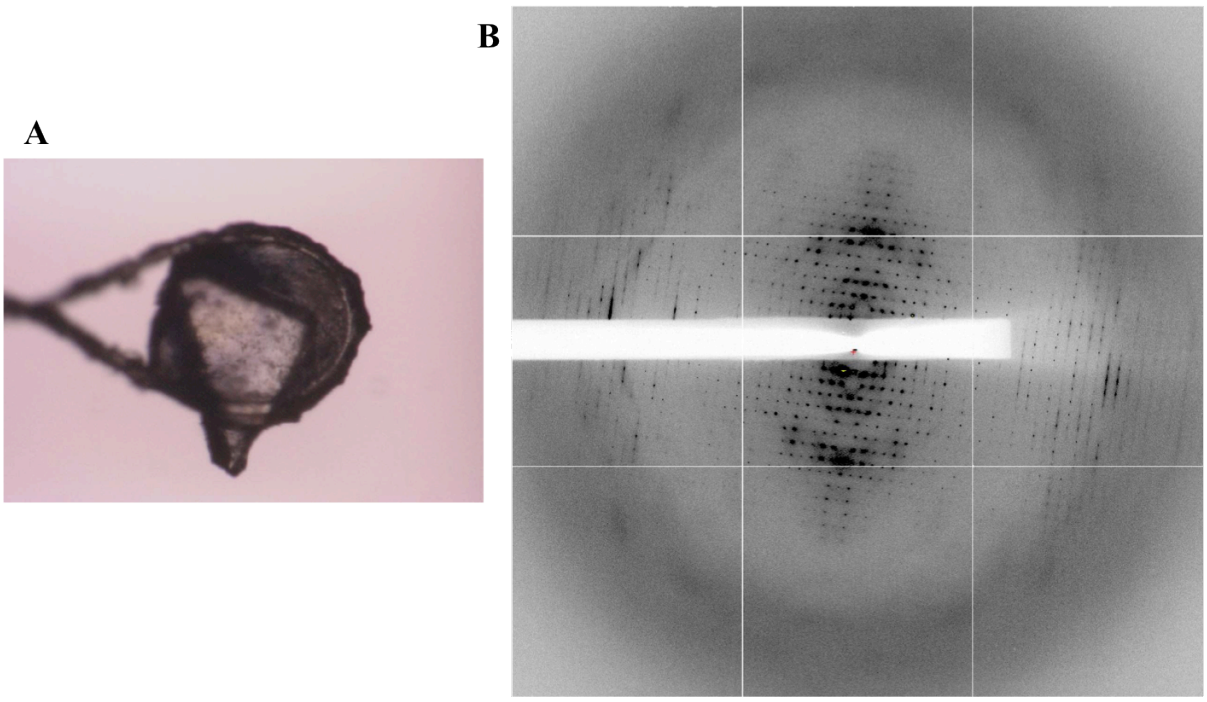


Fig. 4-2:
(A) Typical $PepT_{St}$ crystal in the crystallisation loop, approx. 200µm.
(B) Diffraction pattern of wild type $PepT_{St}$. Diffraction spots are distributed anisotropically, the resolution reaching approx. 3.5Å in the horizontal direction and 5.5Å in the vertical. Data collected at Diamond Light Source.

Diffraction patterns of the cysteine variants resembled those of the wild type $PepT_{St}$ (fig. 4-2, B). Anomalous data was collected from both SeMet and Hg-derivatised crystals to resolutions comparable to the wild type. A further dataset was collected at 7.5KeV (1.8Å wavelength) to detect any anomalous signal from cadmium; a cadmium ion was found to be involved in the crystal contact between two monomers in adjacent asymmetric units (fig. 4-3).

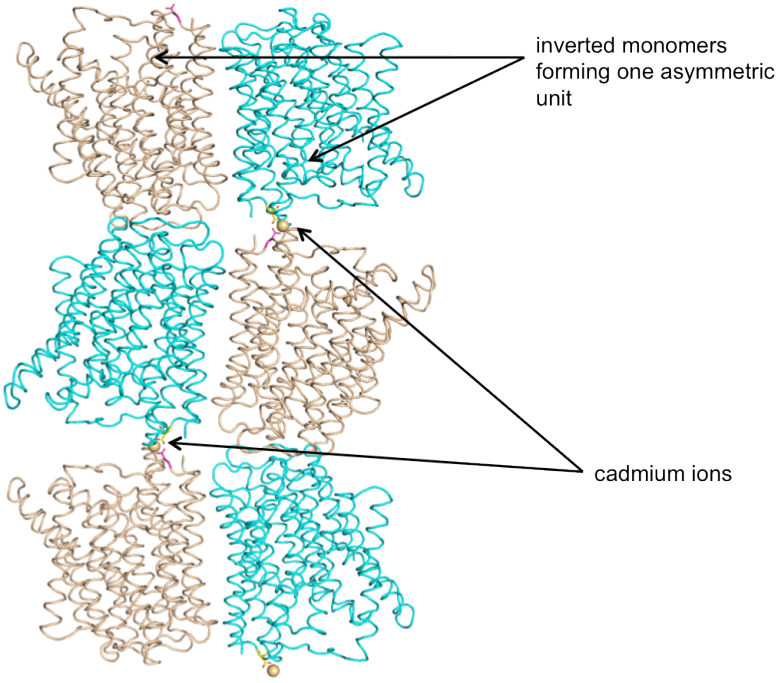


Fig. 4-3: Cadmium ions forming crystal contacts between $PepT_{St}$ monomers in adjacent asymmetric units. Anomalous peaks identified after data collection at the cadmium edge (7.5KeV).

	Native (high)	Native (low)	MMC-1[a]	MMC-2	Se
Data collection					
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions					
a, *b*, *c* (Å)	89.6, 113.1, 215.3	90.4, 113.6, 223.4	90.7, 113.8, 222.5	90.7, 113.7, 224.9	90.6, 113.6, 222.61
α, β, γ (deg)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	107–3.3 (3.4–3.3)	74–4.0 (4.1–4.0)	70–4.5 (4.6–4.5)	70–5.9 (6.0–5.9)	90–4.5 (4.6–4.5)
R_{merge}[b]	10.1 (111.2)	8.1 (85.0)	9.4 (82.2)	6.4 (60.3)	6.6 (47.9)
$I/\sigma I$	5.1 (1.1)	9.6 (1.9)	10.9 (2.7)	16.8 (2.0)	9.0 (2.7)
Completeness (%)	99.3 (99.9)	94.0 (97.4)	97.2 (96.2)	98.5 (93.6)	96.8 (90.9)
Redundancy	3.5 (3.5)	3.0 (3.0)	6.7 (6.5)	6.8 (6.9)	3.5 (3.5)
R_{cullis} (%)			90.3/95.3	109.2/93.5	94.6/90.8
Isomorphous/anomalous					
Phasing power[c]			0.583/0.459	0.345/0.709	0.658/0.646
Isomorphous/anomalous					
Refinement					
Resolution (Å)	29–3.3				
No. of reflections	27 943				
R_{work}/R_{free}	27.3/28.9				
No. of atoms	13 898				
Protein					
r.m.s.d.					
Bond lengths (Å)	0.01				
Bond angles (deg)	0.98				

Table 2-1: Data collection, phasing and refinement statistics for $PepT_{St}$ (adapted from Solcan *et al*, 2012).

4.2. The $PepT_{St}$ structure

The $PepT_{St}$ model was built and refined by Simon Newstead, to a final R/R_f of 27/29 (table 4-1). The structure is typical of a membrane protein, with a hydrophobic surface making up

most of the central, membrane-embedded region as well as solvent-exposed positive charges concentrated at the cytoplasmic side, in compliance with the "positive-inside" rule (de Heijne, 1986) (fig. 4-4, A). As seen in $PepT_{So}$, the residues of the ExxERFxYY motif on H1 are situated in the middle of the membrane, where they contribute to the putative peptide binding site, contradicting earlier topology predictions for hPepT1 that indicated it might be found in an extracellular loop (Daniel, 2004). Similarly, the FYxxINxG motif on H5 is membrane-embedded, while the PTR 2_2 motif itself (GGFVADRIIGARPAV in $PepT_{St}$) forms an exposed surface at the H1/H2 hairpin, the latter suggestive of a membrane-anchoring role for the loosely conserved sequence (fig. 4-4, B).

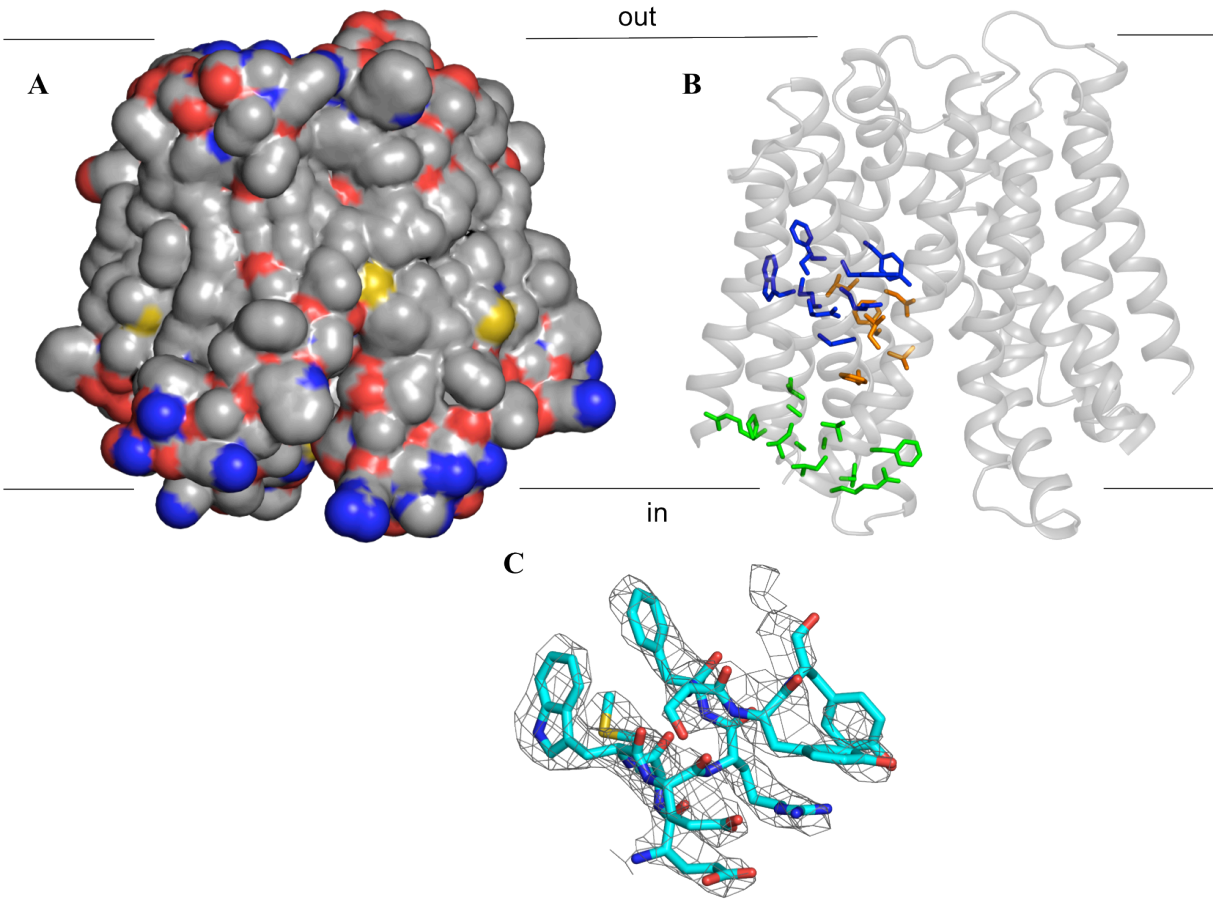


Fig. 4-4: Features of the $PepT_{St}$ structure.
(A) Electrostatic surface of $PepT_{St}$ viewed in the plane of the membrane. The central, membrane-embedded region is mostly hydrophobic, while positive charges are mainly concentrated at the cytoplasmic side.
(B) Localisation of POT family signature motifs in the inward-open $PepT_{St}$ structure (PDB:2APS): ExxERFxYY (H1, blue), GxxxADxxxGAxxAV (H2/H3, green) and FVxxINxG (H5, orange).
(C) 2Fo-Fc electron density map for the residues of the ExxERFxYY motif, contoured at 3.0 sigma. Residues in the final model are shown as sticks.

The $PepT_{St}$ model reveals a similar topology to $PepT_{So}$, with two 6-helical bundles (H1-H6 and H7-H12) forming the core MFS fold, completed by connecting helices HA and HB (fig. 4-5, A). The latter two pack to one side of the transporter (fig. 4-5, C) and are believed to move independently from the 12-helix bundle based on atomistic simulations (Solcan *et al*, 2012). The "intracellular gate" seen in the $PepT_{So}$ structure is fully open in $PepT_{St}$, giving rise to an inward-open conformation, with the central cavity accesible to the intracellular side (fig. 4-5, B).

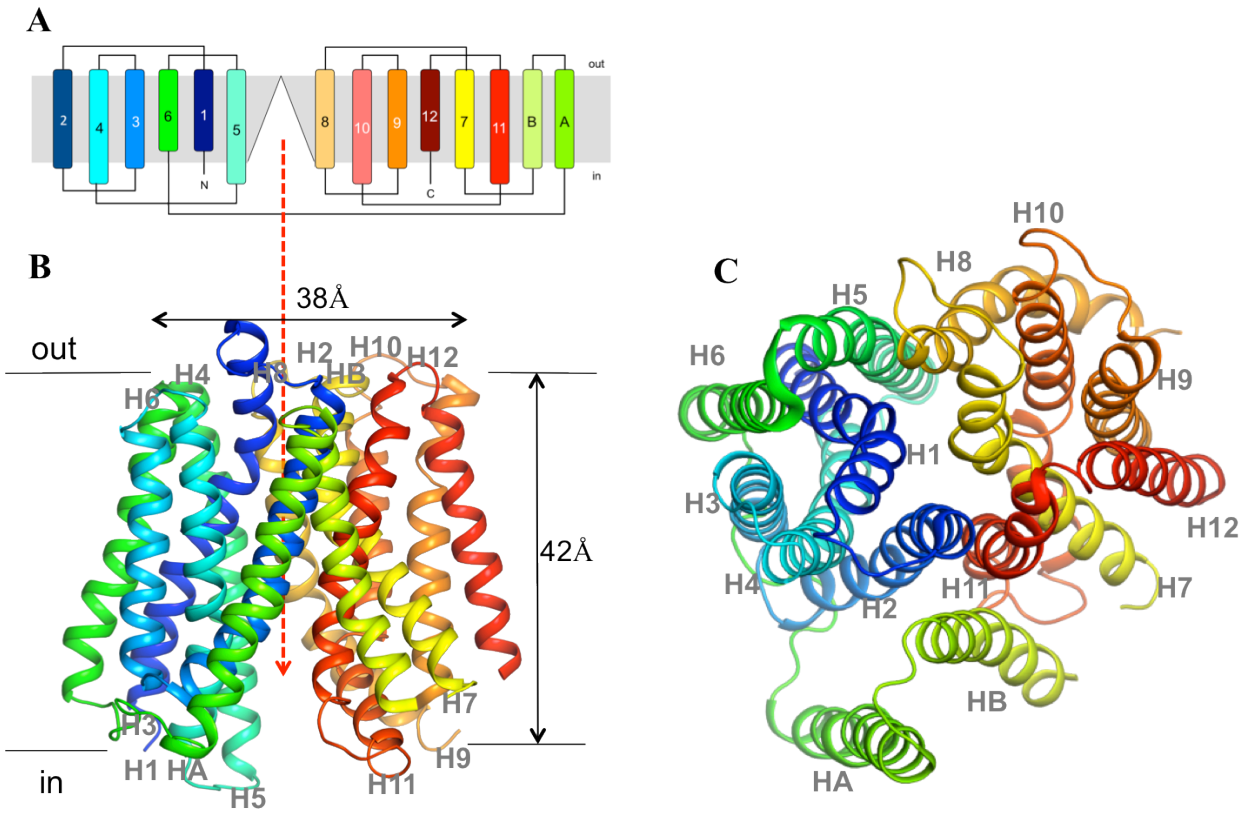


Fig. 4-5: Structure of $PepT_{St}$.
(A)Topology diagram of $PepT_{St}$. showing the central cavity and 14 transmembrane helices.
(B) Crystal structure of $PepT_{St}$ (PDB:2APS), viewed in the plane of the membrane. Transmembrane helices coloured blue to red from the N to the C terminus. The red arrow indicates the inward-open central cavity.
(C) View of the $PepT_{St}$ structure from the extracellular side of the membrane.

While overall similar to $PepT_{So}$, the most notable difference seen in $PepT_{St}$ is the different positions of helices H7, 10 and 11, which move away from H4 and H5 to open the intracellular gate. At the extracellular side, H1/2 and H7/8 seal the transporter by packing against each other, forming an "extracellular gate". As will be discussed in the next chapters, pairs of oppositely charged residues on the N- and C-terminal sides of the central cavity stabilise the extracellular gate and potentially orchestrate the conformational changes between inward-open and occluded states (and from occluded to the proposed outward-open state), reinforcing the alternating access model for MFS transport.

4.3. Comparison to the $PepT_{So}$ structure

The structures of $PepT_{So}$ and $PepT_{St}$ are the earliest example of MFS transporters from the same family being crystallised in different conformations (occluded and inward-open, respectively); the structure of sugar transporter XylE was later solved in three different conformations (Sun *et al*, 2012; Quistgaard *et al*, 2013). $PepT_{So}$ and $PepT_{St}$ have the same fold and high sequence similarity, and therefore a direct structural comparison between the two is possible.

The two structures can be superimposed with an overall RMSD of 1.68Å, calculated for backbone atoms (fig. 4-6, A). The main structural rearrangement going from the $PepT_{So}$ to the $PepT_{St}$ structure is indeed the opening of the intracellular gate, through the movement of helices H10 and H11 away from H4 and H5, as previously proposed (Newstead *et al*, 2011).

However, the main rearrangement occurs in the cytoplasmic halves of H10 and H11, their movement coupled with the straightening of H7 (which is kinked in the $PepT_{So}$ structure). This is reflected in the RMSD calculations, which show that the N-terminal domains (H1-H6) superimpose better than the C-terminal domains (H7-H12) when comparing the two structures (1.39Å in contrast to 1.86Å) (fig. 4-6, B and C).

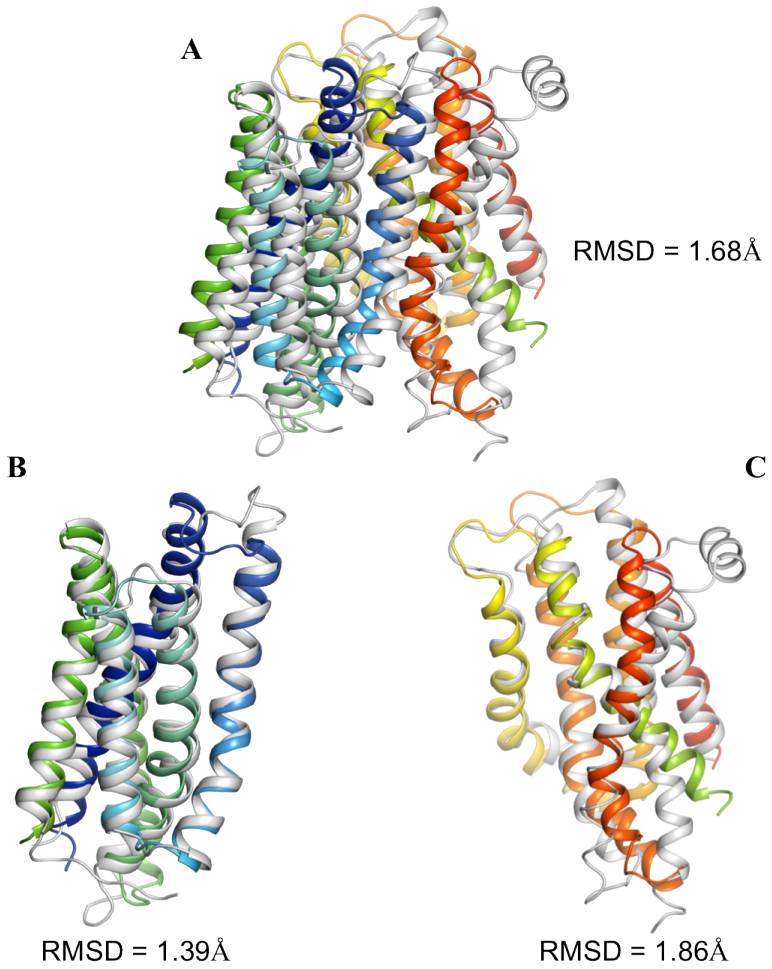


Fig. 4-6: Structural comparison between $PepT_{St}$ and $PepT_{So}$.
(A) Superposition of the inward-open $PepT_{St}$ structure (PDB:4APS, coloured blue to red) onto the occluded $PepT_{So}$ structure (PDB:2XUT, grey). Helices HA and HB are omitted for clarity.
(B) and (C) Superposition of the N-terminal and C-terminal six-helix bundles, respectively. RMSD values were calculated for backbone atoms.

4.4. Conclusions

The structure of $PepT_{St}$ in a different conformation from $PepT_{So}$ is a significant addition to the database of MFS structures, as it allows for a direct structural comparison between members of the same family of MFS transporters. $PepT_{St}$ exhibits the typical 12TM fold and shows the same topology as its homologue. Moreover, the structure confirms the localisation of the peptide binding site and the POT family signature motifs. As $PepT_{St}$ crystallises in an inward-open conformation, the main structural changes from the $PepT_{So}$ structure (inward occluded) are the opening of the intracellular cavity and the closing of the extracellular one, through what appears to be the coordinated opening and closing of mobile gates.

Both N- and C-terminal helices participate in forming the intra- and extracellular gates that are likely to regulate the structural transitions associated with the transport cycle. However, in the case of the intracellular gate (the most prominent conformational change between the structures), its opening is primarily enabled by rearrangements in the cytoplasmic regions of C-terminal helices; in contrast, the N-terminal helix bundle appears to be less mobile. This observation may be related to the overall better sequence conservation in the N-terminal domain of the two proteins, which may indicate that the two domains are involved in different aspects of the transport cycle, some of them conserved (N-terminal) and others more subject to variability (C-terminal). At any rate, the two crystal structures provide a starting point for a detailed functional investigation of the peptide transport mechanism in the POT family, which will be detailed in the following chapters.

Chapter 5: Understanding the binding site of proton-coupled oligopeptide transporters

The structure of $PepT_{So}$ in the inward-occluded conformation brought to light the peptide binding site, consisting of generally well conserved residues in the central hydrophilic cavity of the transporter. Its architecture was typical of MFS binding sites, consisting of polar residues contributed by both N- and C-terminal helices, and several aromatic side chains reaching into the centre of the cavity. Although captured in a different conformation, the $PepT_{St}$ binding site proved to be largely similar, most residues conserved (with a few notable differences), pointing towards similar interactions and a potentially similar transport mechanism (fig. 5-1).

Both binding sites are asymmetric, with negatively charged side chains predominantly on the C-terminal side (H7 and H11) counterbalanced by positively charged ones on the N-terminal side (H1 and H4). This arrangement of opposite charges may function as a dipole, making it possible to orientate asymmetric substrates of varying length and charge. With the exception of Glu299 and Glu300 in $PepT_{St}$ (Phe315 and Asp316 respectively in $PepT_{So}$ and variable in other transporters), these residues are conserved in both proteins as well as all other POT family members. The first arginine on H1 is part of the universally conserved ExxERFxYY signature motif, as are the two glutamates on H1 (Glu22 and Glu25 in $PepT_{St}$), the latter two positioned just left of the binding site. The two tyrosines on H1 (the first of these substituted for phenylalanine in $PepT_{So}$) are part of the same motif, and, together with a

third aromatic residue (Tyr68), protrude into the centre of the binding site between the polar/charged side chains. Many of these residues form polar interactions in the crystal structures, and some of these interactions appear to form or break depending on the conformation of the transporter (inward-open for $PepT_{St}$, inward-occluded for $PepT_{So}$).

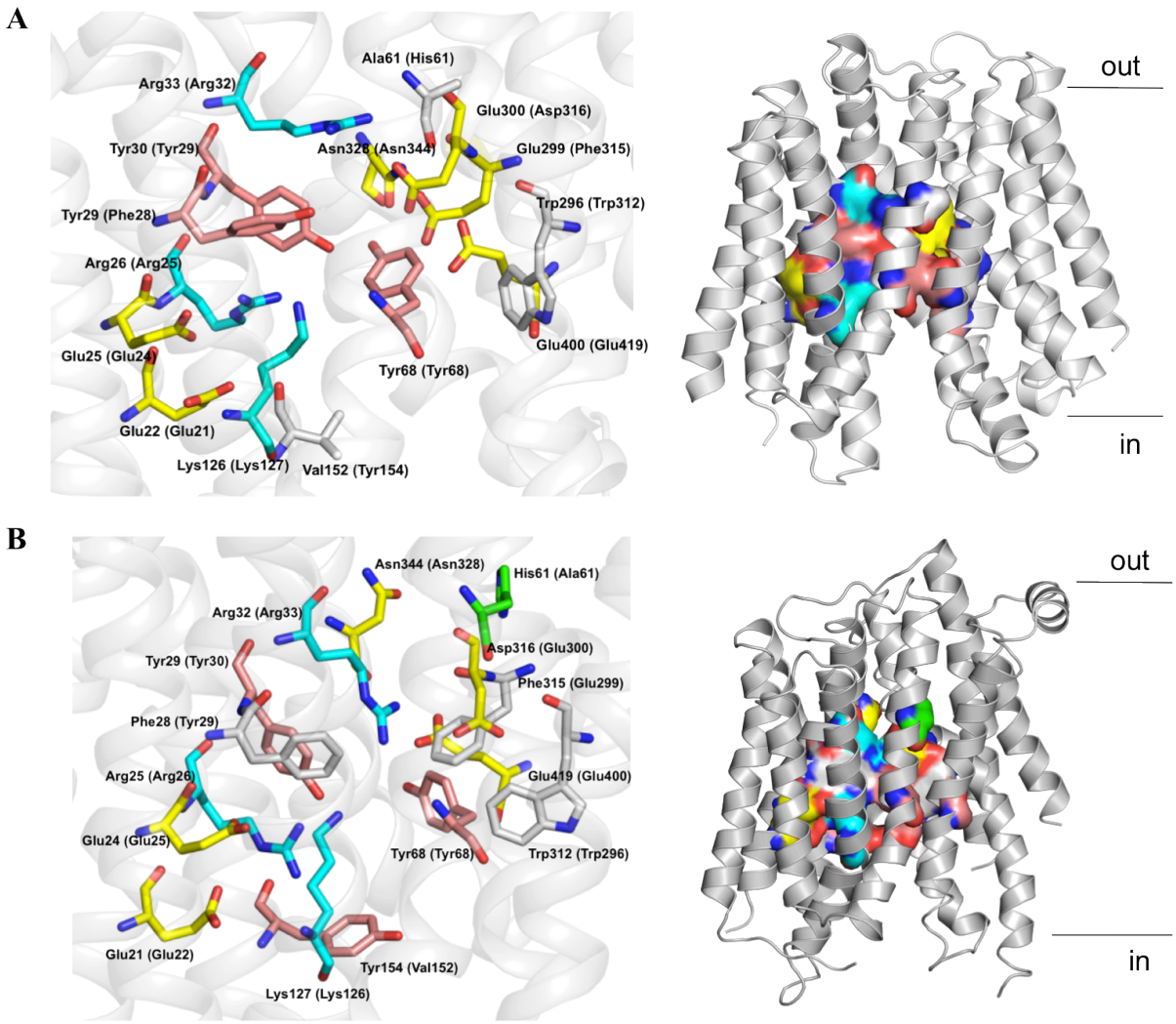


Fig. 5-1: The peptide binding sites of $PepT_{St}$ (A) and $PepT_{So}$ (B). The localisation of the binding sites in the protein structure is shown in surface representation (right). Equivalent residues from the other protein are listed between brackets. View along the plane of the membrane. PDB codes: 2APS ($PepT_{St}$) and 2XUT ($PepT_{So}$).

Apart from the Glu299/Phe315 substitution, two more striking differences exist between the $PepT_{St}$ and $PepT_{So}$ binding sites. Firstly, the PTR2_2 motif (FYxxINxG) contains Tyr154 ($PepT_{So}$ and also a tyrosine in most POTs), whereas $PepT_{St}$ has a rare substitution for

a valine at this position. Secondly, His61 is situated at the top of the $PepT_{So}$ central cavity, its imidazole side chain pointing into the extracellular cavity, but in $PepT_{St}$ it is replaced by an alanine. In light of these differences, $PepT_{So}$ resembles hPepT1 and hPepT2 more closely, as both of them, like $PepT_{So}$, have a tyrosine and a histidine at these two positions respectively.

The architecture of the binding site, with its asymmetric arrangement of charges and clusters of highly conserved residues, is likely the key to understanding how substrate specificity is regulated and how transport is coupled to the proton translocation mechanism. These aspects have frequently been addressed in the literature pertaining to the eukaryotic peptide transporters (PepT1 and PepT2) through transport assays. However, in the absence of a purification protocol, functional studies of PepT1/2 have been hitherto restricted to *in vivo* experiments, either using cell lines that natively express the proteins or other cell types transfected with PepT-encoding DNA. Fortunately, the use of prokaryotic transporters as model systems to understand PepT function has brought with it more ways to perform peptide and drug transport assays.

Whole cell assays (in *E. coli* cells overexpressing protein) are a quick way of assaying transport by bacterial POTs. However, the whole cell approach limits the choice of reporter substrates to non-hydrolysable peptide analogues or non-peptide molecules, as physiological peptides would face enzymatic degradation in the cytoplasm, making quantification of transport levels problematic. Substrates typically used when assaying POTs in *E. coli* cells include fluorophore-attached peptide β-Ala-Lys-AMCA and radiolabeled (^{3}H or ^{14}C) peptide analogue glycylsarcosine (GlySar). The two were used as reporter substrates for YdgR (DtpA), YhiP (DtpB) and YjdL (DtpC), all of which were assayed in whole cells to investigate transport kinetics and determine substrate specificity by calculating IC_{50} values for different non-labelled peptides and drugs. Using whole cell assays, it was established that cephalosporin antibiotics compete against GlySar transport via YdgR (but not YhiP) (Harder

et al, 2008), and that YjdL displays much lower affinities for tripeptides compared to dipeptides (high mM compared to low mM) (Ernst *et al*, 2009).

In vitro POT transport assay protocols are based on earlier work on other H^+-coupled MFS transporters, specifically LacY and other lactose permeases. The *E. coli* lactose transport system was the first to be reconstituted into liposomes (Newman and Wilson, 1980), and transport assays were designed using radiolabeled substrate and artificially imposed electrical and chemical gradients across the liposome membrane (Viitanen *et al*, 1984; 1986; Lolkema *et al*, 1991). These formed the basis for transport assays with reconstituted DtpT (peptide transporter from *L. lactis*) (Hagting *et al*, 1997). Using dipeptide ^{14}C-ProAla as a reporter, it was determined that DtpT substrates are exclusively di- and tripeptides, and that transport via DtpT, like LacY, is H^+-coupled and dependent on either the membrane potential ($\Delta\Psi$-driven) or a concentration gradient of the substrate itself (counterflow) (Fang *et al*, 1999; 2000). More recently, similar experiments confirmed H^+ coupling and the ability to perform counterflow for other MFS transporters such as GlpT (Fann *et al*, 2003) and FucP (Dang *et al*, 2010).

While other *in vitro* methods have been used to assess POT function, the data from these remains scarce. For instance, measurements of transmembrane electrical currents in liposomes were performed with YdgR in the presence of substrate to confirm the rheogenic nature of transport (Weitz *et al*, 2007), but the small size of the liposomes makes the method impractical for accurate measurement of proton/peptide stoichiometry (unlike the patch clamp method in oocytes); instead, recording proton fluxes directly in liposomes tends to yield better results when assaying transporters with rapid uptake rates, such as cation/anion exchangers (Walden *et al*, 2007). Therefore, measuring uptake of labelled substrates in liposomes remains the most feasible *in vitro* method for assaying POT function.

Transport assays of DtpT and the *E. coli* POTs in whole cells and liposomes have provided some preliminary data on substrate preference and the general dynamics of transport. Many issues however remain to be investigated, such as drug transport *in vitro*, the pH dependence of transport, and distinguishing between substrates and inhibitors. The recent crystal structures of $PepT_{So}$ and $PepT_{St}$ and their similarity to the human PepT1 and PepT2 make them exciting new targets for a more in-depth functional study, not only to establish bacterial POTs as model systems, but also to understand the structural implications of biochemical assay data.

5.1. Whole cell transport assays of wild type $PepT_{So}$ and $PepT_{St}$

A cell-based assay was set up to analyse transport of radiolabelled compounds in *E. coli* cells overexpressing peptide transporters, as described in the methods chapter. GFP-fusion and His-tagged constructs of $PepT_{So}$ and $PepT_{St}$ were tested, and experiments used ^{3}H-GlySar as the substrate. Tritium-labelled non-physiological dipeptide D-Phe-L-Gln (a substrate of rabbit PepT1) (Meredith *et al*, 2000) and fluorescent peptide β-Ala-Lys-AMCA were also tested for uptake, but neither was transported in a quantifiable amount.

$PepT_{St}$ and $PepT_{So}$ were assayed 20h after induction, using pWaldo-GFP-His_8 constructs (as typically used for expression), as well as pTTQ18-His_6 constructs (without the GFP tag). ^{3}H-GlySar uptake by the His-tagged constructs was approximately linear ($PepT_{So}$)

and approached saturation ($PepT_{St}$) over 30 minutes, while the GFP-fusion protein only transported the peptide at the level of the control cells carrying empty pTTQ18 plasmid (fig. 5-2). The similar activity profiles of the two proteins prove that peptide transporters fused to GFP are not active in cells, perhaps due to physical hindrance by the C-terminal GFP molecule in the cytoplasm. Nevertheless, while constitutive transport is observed in the control cells (endogenous *E. coli* transporters), the His-tagged proteins are active to a level comparable to that seen in a previous study with a similar construct of YdgR (Harder *et al*, 2008).

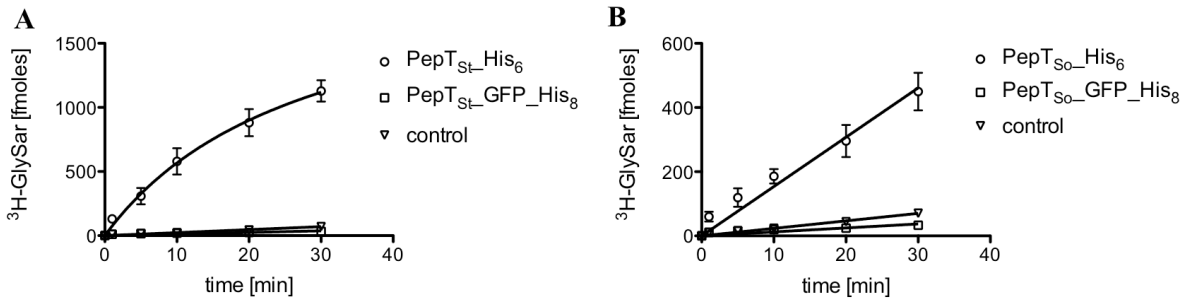


Fig. 5-2: Uptake of ^{3}H-GlySar in *E. coli* BL1(DE3) cells expressing GFP- and His-tagged constructs of $PepT_{St}$ (A) and $PepT_{So}$ (B). $5 \cdot 10^9$ cells were harvested 20h post-induction and 0.4 μM ^{3}H-GlySar was used per assay. Cells overexpressing empty pTTQ18 plasmid were used as control.

Alternatively, the assay was performed 3h after induction; however, results proved difficult to reproduce, with readings varying greatly between experiments, and control cells showing increased activity. Furthermore, using His-tagged constructs limits precise quantification of protein expression levels (in contrast to the GFP system), particularly as issues with leaky expression are common in BL21 cells. To address these concerns, four peptide transporter deletion strains of *E. coli* were tested for GlySar uptake under the same conditions in order to identify a strain with reduced background levels for further experiments. Each strain had a deletion of one of the four genes encoding for POTs in the *E.* coli genome (*tppB*, *yhiP*, *yjdL* and *ybgH*). Surprisingly, this resulted in increased activity compared to the normal BL21 control cells, suggesting that knocking out one transporter causes upregulation

of the remaining three (fig. 5-3). This compensating effect may be linked to the association of peptide transporter expression with stress response, as all the knockout strains grew exceptionally slowly in the TB cultures and reached stationary phase at a lower OD_{600} than usual.

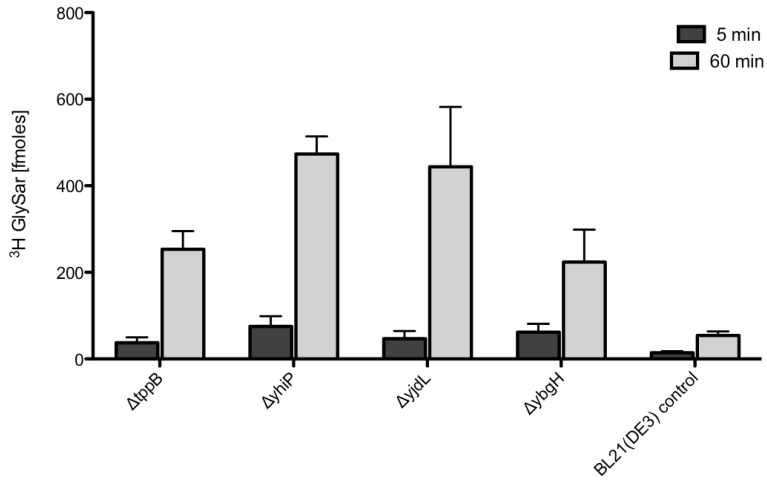


Fig. 5-3: Uptake of 0.4 μM ^{3}H-GlySar by *E. coli* Keio peptide transporter knockout strains.

To overcome the limitations of whole cell assays, a decision was taken to develop a more robust, *in vitro* transport assay for $PepT_{St}$ and $PepT_{So}$. To this end, a reconstitution protocol was developed and proteoliposome-based transport assays (ΔΨ-driven and counterflow) were designed as described in the methods chapter.

5.2. Substrate specificity of wild type $PepT_{St}$ and $PepT_{So}$ *in vitro*

The uptake of several radioactively labelled molecules by liposome-reconstituted transporters was tested at different temperatures, pH values, concentrations of both liposomes and substrate, and across different time spans. Assay conditions were optimised for direct uptake (ΔΨ-driven), counterflow (substrate-driven), competition experiments and kinetics, in an effort to generate a general profile for peptide transport by wild type $PepT_{St/So}$, as well as explain the implications of substrate selectivity on the amino acid composition of the binding sites.

5.2.1. Time-dependent uptake of radiolabelled substrates

The uptake of ^{3}H-labelled AlaAla, GlySar and D-Phe-L-Gln by $PepT_{St}$ was recorded under ΔΨ-driven conditions over a time span of 5 min (fig. 5-4, A). AlaAla and GlySar are substrates of hPepT1 and hPepT2, while D-Phe-L-Gln is transported by the rabbit PepT1. Here, $PepT_{St}$ only takes up AlaAla and GlySar; AlaAla transport is fast, the liposomes quickly saturated with peptide, whereas GlySar is taken up slowly and linearly at this concentration. While AlaAla is the preferred substrate, GlySar uptake is visible when comparing measurements taken at the 5 min time point (fig. 5-4, B). The different transport rates and the lack of D-Phe-L-Gln uptake echo the whole cell assay results discussed previously.

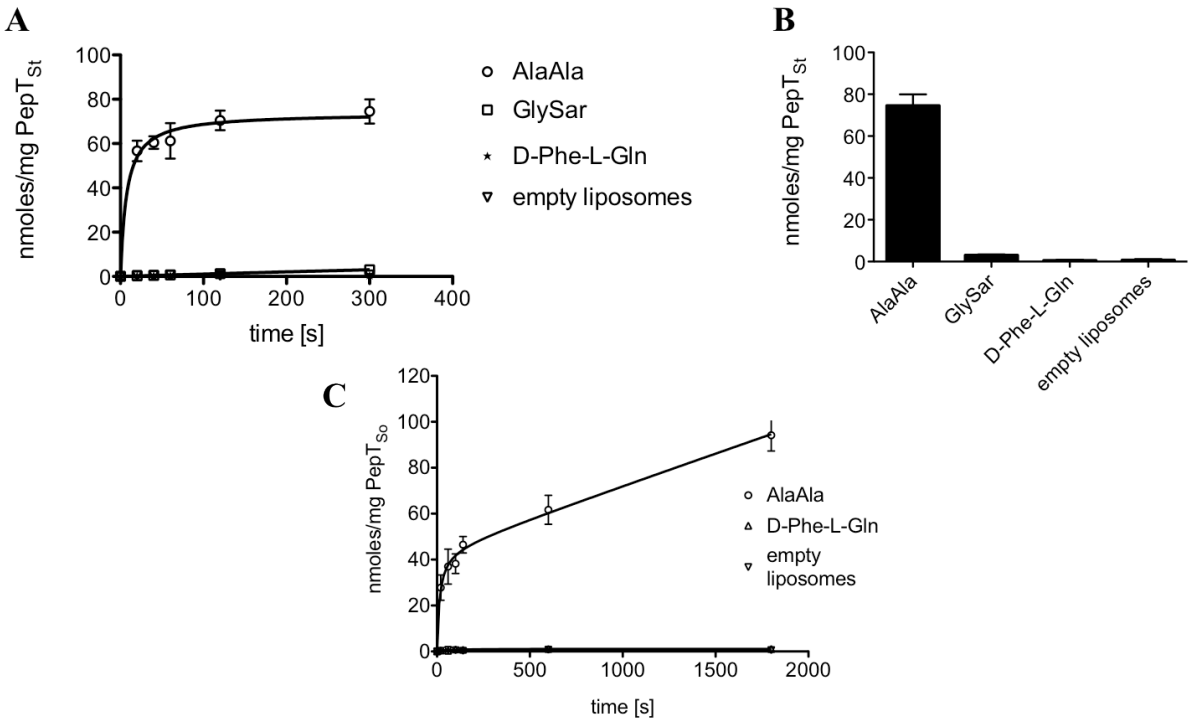


Fig. 5-4: Comparative uptake of three ^{3}H-labelled substrates by liposome-reconstituted POTs at pH 6.5.
(A) AlaAla, GlySar and D-Phe-L-Gln transport by $PepT_{St}$. Proteoliposomes containing internal buffer (20 mM KPi, 100 mM KAc and 2 mM $MgSO_4$) were diluted 1:50 in external buffer (120 mM NaPIPES and 2 mM $MgSO_4$) containing 10 μM substrate and 2 μM valinomycin.
(B) Differences in transport at the 5 min time point ($PepT_{St}$).
(C) AlaAla and D-Phe-L-Gln transport by $PepT_{So}$.

In identical assay conditions, $PepT_{So}$ also transports AlaAla but not its D-isomer; however, AlaAla is taken up at a slower rate, and saturation takes longer than 30 min to reach (fig. 5-4, C). While all reconstituted and assayed transporters were purified from cleavable GFP-fusion constructs as described in the methods chapter (pWaldo vector), $PepT_{So}$ was additionally purified with a His_6 tag (pTTQ18 construct), but no differences were observed in terms of either reconstitution efficiency or uptake profile. Together, the data show that the binding sites of both $PepT_{St}$ and $PepT_{So}$ are similar to those of PepT1 and PepT2 inasmuch as they distinguish between L- and D-isomers of peptides, as well as between high and low affinity substrates.

A similar time course experiment was performed to test AlaAla counterflow via both $PepT_{St}$ and $PepT_{So}$. The exchange of unlabelled peptide from the liposome lumen with ^{3}H-AlaAla from the external buffer was monitored over 5 min (fig. 5-5). The amount of labelled peptide accumulated in the liposomes is similar to that seen in the ΔΨ-driven assay, suggesting that, although the membrane potential can drive transport, the presence of a substrate gradient can compensate for its absence. Furthermore, substrate affinity seemingly does not change with the membrane potential.

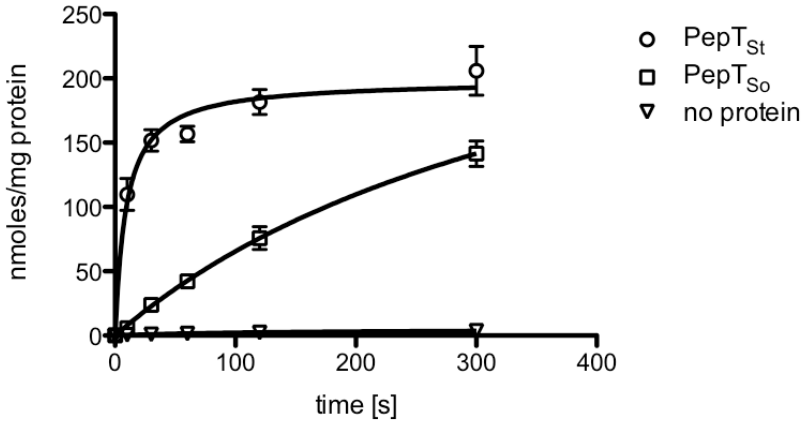


Fig. 5-5: Substrate-driven uptake (counterflow) of AlaAla by $PepT_{St}$ and $PepT_{So}$. Proteoliposomes containing 50 mM KPi (pH 6.5), 2 mM $MgSO_4$ and 10 mM AlaAla were diluted 1:50 in 50 mM KPi (pH 6.5) and 2 mM $MgSO_4$ containing 5 μM ^{3}H-AlaAla.

5.2.2. Peptide size preference

To establish that $PepT_{St}$ is a POT family transporter (di- and tripeptide specific), a ΔΨ-driven competition experiment was set up using ^{3}H-AlaAla as a reporter and alanine peptides of various lengths (including alanine itself) as the competitors. Uptake of the reporter peptide was measured after 2 min (fig. 5-6). Indeed, only AlaAla and AlaAlaAla display competition, the dipeptide preferred over the tripeptide. Although whole cell assays on YdgR indicated a

higher affinity for the tripeptide (Weitz *et al*, 2007), this finding is in agreement with existing liposome data on DtpT (Fang *et al*, 2000). D-Ala-D-Ala is also not transported, as seen previously for hPepT2 (Lin and King, 2007), proving $PepT_{St}$ is selective towards L-isomers. Some competition by alanine and tetraalanine is noticeable, possibly due to non-specific crowding of the transporter by the excess peptide and/or hydrophobic interactions. Finally, CCCP (proton ionophore) stops the reaction, confirming the transport mechanism is proton-coupled.

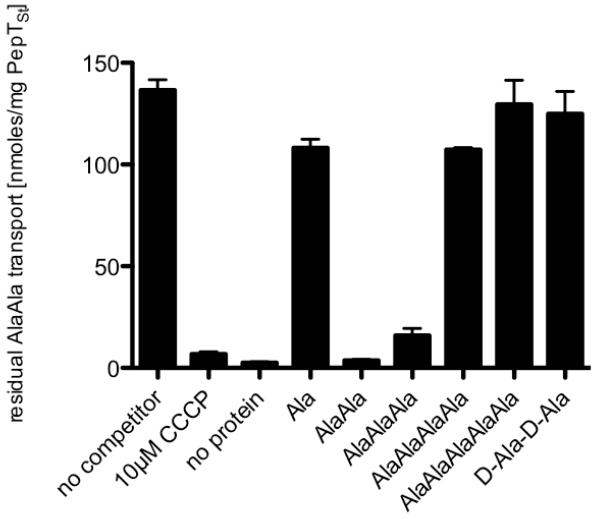


Fig. 5-6: Competition assay with alanine peptides. ΔΨ-driven uptake of 15 µM ^{3}H-AlaAla by $PepT_{St}$ at 2 min (pH 6.5, 25°C) was quantified in the presence of competing peptides. Competitor concentration was 2 mM unless specified.

5.2.3. pH dependence of AlaAla uptake

A single-measurement ΔΨ-driven assay was performed with $PepT_{St}$ at different pH values, showing that pH 6.5 is optimal for AlaAla transport (fig. 5-7). Comparable amounts of peptide are taken up between pH 5.5 and 7.5, but readings decrease significantly below and

above that interval. The preference for a neutral to weakly acidic pH matches that observed in eukaryotic peptide transporters like rabbit PepT1 (Steel *et al*, 1997) and mouse PepT1 (Jappar *et al*, 2010), and agrees with the preferred pH range for transport of neutral peptides and drugs by peptide transporters in general (Wenzel *et al*, 1996; Steel *et al*, 1997).

Additionally, the destabilising effect of salt on peptide transport was shown by recording AlaAla uptake with increasing concentrations of NaCl in the external buffer, from 0 to 600 mM (fig. S1).

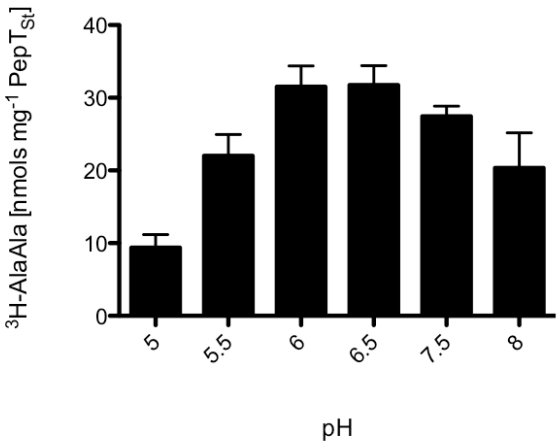


Fig. 5-7: pH dependence of AlaAla uptake by $PepT_{St}$. Proteoliposomes containing 50 mM KPi at each pH were diluted 1:50 in 50 mM NaPi at the same pH containing 10 μM valinomycin. Uptake of 15 μM ^{3}H-AlaAla was measured after 1 min at 25°C.

5.2.4. Kinetics of GlySar uptake

As illustrated in figure 5-4 (A and B), AlaAla is rapidly transported by $PepT_{St}$ even at low μM concentration, leading to quick saturation of the liposomes with the peptide. Thus, the region of the curve where transport is linear is very steep. This is likely to become more pronounced at higher peptide concentrations (Cornish-Bowden, 1979), making AlaAla unsuitable for kinetic studies, as transport rates would be difficult to calculate. Consequently, $PepT_{St}$

kinetics were analysed using GlySar, as it is transported by the human proteins and is better suited for monitoring linear transport across a range of concentrations, albeit being a lower affinity substrate than AlaAla.

For the kinetic measurements, individual ΔΨ-driven transport assays were carried out at different GlySar concentrations, and transport rates were calculated from each (fig. S2). The rates were then plotted against substrate concentration to generate a Michaelis-Menten profile (fig. 5-8). The apparent K_m of transport (K_m^{app}) is 9.3 ± 3 mM with a V_{max} of 631 ± 87 nmoles/min x mg protein. The K_m value agrees with a recent *in vivo* study on mouse peptide transporters (Jappar *et al*, 2010) and is in the same range as reported values for rabbit PepT1 in oocytes (Steel *et al*, 1997) and *E. coli* transporters YdgR and YhiP in liposomes (Harder *et al*, 2008). These findings further suggest that $PepT_{St}$ is a valid model system for the study of eukaryotic peptide transport.

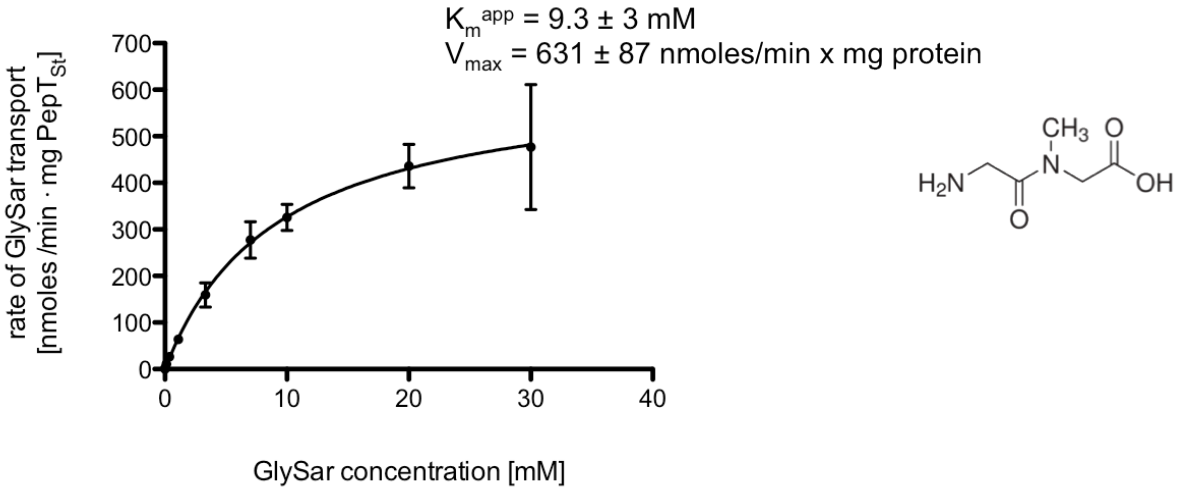


Fig. 5-8: Kinetic analysis of ΔΨ-driven ^{3}H-GlySar uptake (GlySar chemical structure shown on the right). $PepT_{St}$ proteoliposomes in internal buffer (20 mM KPi, 100 mM KAc and 2 mM $MgSO_4$) were diluted 1:50 in external buffer (120 mM NaPIPES and 2 mM $MgSO_4$) containing 10 μM valinomycin and GlySar concentrations ranging from 41 μM to 30 mM. The Y-axis shows transport rates at each substrate concentration, calculated from triplicate measurements.

5.2.5. Di- and tripeptide specificity

While the use of ^{3}H-labelled AlaAla and GlySar as representative POT substrates (high and low affinity) is useful for kinetic experiments as well as studying time and pH dependence of transport, identifying other substrates depends primarily on competition experiments. Consequently, competition assays with a small library of di- and tripeptides were carried out using ^{3}H-AlaAla as a reporter substrate, in order to create a peptide specificity profile for $PepT_{St}$ and $PepT_{So}$. Peptides of various side chain lengths and charges were selected as the competitors. For ease of comparison, the amounts of transported ^{3}H-AlaAla were converted to percentages of the amount transported in the absence of competitor (residual uptake).

The results of the competition assay reveal and overall agreement regarding substrate specificity between $PepT_{St}$ and $PepT_{So}$ (fig. 5-9). As expected considering the prominent presence of aromatic groups in the centre of the binding site, both proteins show a preference for dipeptides with bulky and/or hydrophobic side chains, with PhePhe, AlaPhe, AlaTyr and LeuLeu completely inhibiting AlaAla uptake. This trait is shared with human PepT1 and PepT2, as shown by several other studies (Eddy *et al*, 1995, Kottra *et al*, 2002, Daniel *et al*, 2004, Biegel *et al*, 2006). Smaller dipeptides like the Gly-containing CysGly and GlyGln are better competitors for $PepT_{St}$ than $PepT_{So}$, whereas the situation is reversed for acidic dipeptide GluGlu. Remarkably, neither protein is inhibited by LysLys.

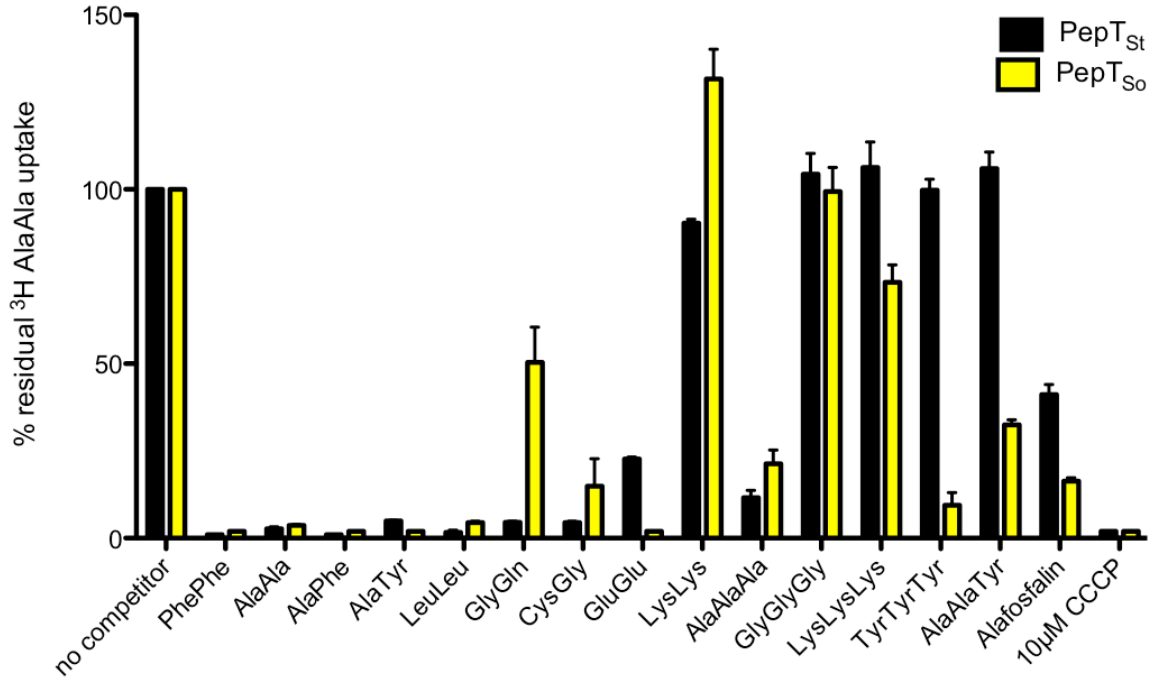


Fig. 5-9: Comparative study of peptide selectivity in PepT$_{St}$ (black) and PepT$_{So}$ (yellow). Uptake of 15 μM ^{3}H-AlaAla by 1 μM protein after a 2 min incubation was measured in the presence of 2 mM competitor unless specified. Data shown as percentages. CCCP refers to the proton ionophore carbonyl cyanide *m*-chlorophenyl hydrazine.

As shown previously for PepT$_{St}$ (5.2.2), AlaAlaAla appears to be a weaker substrate than AlaAla, but is nevertheless a good competitor for both proteins. Other tripeptides (GlyGlyGly and LysLysLys) do not inhibit transport at all, whereas hydrophobic tripeptides like TyrTyrTyr and AlaAlaTyr only compete against PepT$_{So}$-mediated uptake, suggesting that PepT$_{So}$ may be an overall better tripeptide transporter than PepT$_{St}$. A similar distinction has been observed between bacterial POT homologues in *E. coli*, where YjdL was shown to be a near-exclusive dipeptide transporter, compared to its canonical di/tripeptide transporter homologues YdgR and YhiP (Jensen *et al*, 2012a; 2012b). Finally, alafosfalin, a dipeptide analogue typically used in growth assays to determine functional overexpression of POTs (Harder *et al*, 2008; Malle *et al*, 2011), acts as a competitor for both proteins, to different extents.

5.2.6. IC_{50} experiments ($PepT_{St}$)

While measuring competition at a fixed concentration in an endpoint assay offers an indication of the peptide preference, it is not an accurate quantification of affinity. In order to confirm the validity of that result, a more in-depth competition assay was carried out with $PepT_{St}$ and selected peptide substrates at different concentrations (fig. S3). IC_{50} values (competitor concentrations at which transport of the reporter substrate is reduced by 50%) were calculated as expressions of affinity relative to dialanine (table 5-1).

Competing peptide	IC_{50} (µM)
PhePhe	6 ± 1
AlaPhe	22 ± 3
AlaAla	36 ± 2
GlyGln	37 ± 3
AlaTyr	70 ± 8
CysGly	342 ± 54
AlaAlaAla	404 ± 67
GluGlu	598 ± 75
LysLys	> 10000

Table 5-1: IC_{50} values for di- and tripeptides relative to $PepT_{St}$-mediated uptake of 15µM ^{3}H-AlaAla.

The IC_{50} values are generally consistent with the endpoint measurements of $PepT_{St}$, but offer a clearer distinction between high affinity competitors: PhePhe is the preferred peptide, and comparably low values are recorded for AlaPhe, AlaTyr, GlyGln and AlaAla

itself. GluGlu and CysGly register lower affinities, similar to AlaAlaAla, the latter being approximately 10-fold weaker than the affinity for AlaAla. An IC_{50} value could not be calculated for LysLys, indicating it may be either a low affinity substrate or not transprted at all. In summary, the $PepT_{St}$ binding site can distinguish between hydrophobic and charged dipeptides (of which it prefers the former), as well as between negative and positive charges, in the same way the human transporters do (Eddy *et al*, 1995). In the absence of the apparent K_m of transport for dialanine, the IC_{50} is not equivalent to an absolute K_i value, and as such is merely defined in relation to the affinity for dialanine. However, as dialanine competes against its own uptake (at 15 μM) with an IC_{50} of 36 μM, it can be assumed that the K_m^{app} itself is in the same concentration range, and therefore the K_i values for the other peptides are not far removed from the IC_{50} values calculated here.

5.2.7. Competition vs. transport: counterflow assays

When assaying peptide transporter function through competition assays, it is important to distinguish between *de facto* substrates and inhibitors or non-specific binders. One way to make this distinction is to measure direct uptake of a substrate, either by *in vivo* electrophysiology measurements or using labelled substrates (Brandsch, 2009). Although the former is not applicable to liposome-reconstituted proteins and the latter would require a large number of radiolabelled molecules, the counterflow assay setup provides a second option. As both $PepT_{St}$ and $PepT_{So}$ were shown to transport AlaAla bidirectionally in the absence of an ion gradient (exchanging excess unlabelled peptide for its tritiated form), the same should occur when liposomes are loaded with other unlabelled peptide substrates (taking up ^{3}H-

AlaAla in exchange). In contrast, if the peptide is not a substrate, the transport cycle should not be initiated, and therefore no ^{3}H-AlaAla uptake should be detectable.

Several di- and tripeptides were tested for counterflow via $PepT_{St}$ and $PepT_{So}$, alongside empty liposomes as negative controls; ^{3}H-AlaAla uptake levels higher than background indicate the peptide is transported (fig. 5-10). The data shows that all peptides that inhibited reporter uptake in the competition assay (including peptide analogue alafosfalin) are indeed transported substrates. Also, as implied by the competition data, all tested tripeptides are $PepT_{So}$ substrates, whereas $PepT_{St}$ transports trialanine but no uptake is detected for larger tripeptides like TyrTyrTyr and AlaAlaTyr.

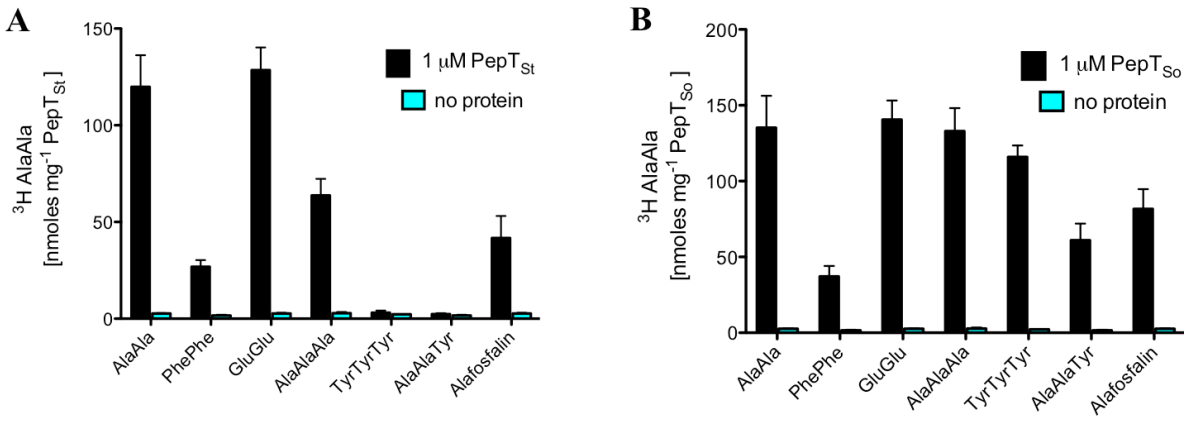


Fig. 5-10: Counterflow of di- and tripeptides by $PepT_{St}$ (A) and $PepT_{So}$ (B). Proteoliposomes were loaded with 10 mM peptide (except 2 mM for PhePhe), and peptide-driven uptake of 5 μM ^{3}H-AlaAla by 1 μM protein was measured after 5 min ($PepT_{St}$) and 10 min ($PepT_{So}$) at pH 6.5 and 25°C.

5.3. Transport profile of $PepT_{St}$ and $PepT_{So}$ binding site mutants

The wild type transport profiles of $PepT_{St}$ and $PepT_{So}$ demonstrate the general similarities between the two proteins in terms of their peptide specificity as well as their dependence on the driving forces for transport like the electrochemical gradient and substrate concentration. However, the data also shows the proteins to have different affinities for certain peptides, with tripeptides in particular proving to be better transported by $PepT_{So}$ than $PepT_{St}$. To an extent, these results can be traced back to the structure of the binding sites (fig. 5-1) to explain how specificity is regulated (for instance, aromatic side chains may serve to accommodate hydrophobic substrates), but without knowing the function of individual residues the explanations remain conjectural. In order to investigate the significance of these differences and the role of the binding site residues in the transport mechanism, site-directed mutagenesis was used in combination with transport assays, in an attempt to distinguish between residues involved in peptide recognition and proton coupling.

For this, the first mutagenesis-based *in vitro* functional study of peptide transporters, all binding site residues were exchanged to alanine in both proteins, using site-directed point mutations. After expression and purification, the proteins were reconstituted and tested in transport assays alongside the wild type variants, in order to assess the involvement of each mutated residue in coupling transport to the membrane potential ($\Delta\Psi$-driven) and the ability of the protein to undergo conformational changes (counterflow).

In addition to alanine, some of the charged amino acids were also mutated to those of opposite charge (such as Lys to Glu) or similar charge but shorter or longer side chains (for

example, Glu to Asp). In particular, the role of Lys126/127, strictly conserved among POT members, was interrogated by mutating it to alanine, arginine, histidine and glutamine (table 5-2). Similarly, the three tyrosines in $PepT_{St}$ were mutated to both alanine and phenylalanine to examine the eventual significance of both side chain length and the hydroxyl group, while the equivalent Phe28 in $PepT_{So}$ was in turn mutated to a tyrosine.

Certain mutations did not express, such as all of the tested Glu299 substitutions in $PepT_{St}$ (alanine, glutamine and even aspartic acid), suggesting the glutamate side chain at that position is crucial to protein stability. Moreover, alanine mutations of two conserved residues, Asn328 and ExxERFxYY motif amino acid Arg26 (both in $PepT_{St}$), did not express and thus had to be replaced by more conservative substitutions (Asp and Lys, respectively), although interestingly the equivalent alanine substitutions in $PepT_{So}$ expressed well and yielded stable proteins (table 5-2).

All proteins were assayed as described in the methods chapter, using ^{3}H-AlaAla as a reporter substrate. As $\Delta\Psi$-driven AlaAla uptake via $PepT_{St}$ is fast (fig. 5-4, A), a single measurement was taken after 2 min; to match the slower uptake rate, $PepT_{So}$ variants were measured at the 10 min time point. For the counterflow measurements, three samples were taken at different time points to check the gradual signal increase indicative of uptake, but only one measurement was plotted for clarity (5 min for $PepT_{St}$, 10 min for $PepT_{So}$). $PepT_{St}$ variants were assayed first, to complete the functional characterisation of the protein (Solcan *et al*, 2012), and the comparative $PepT_{So}$ study followed later. However, for the purpose of this thesis, $PepT_{St}$ and $PepT_{So}$ results are presented and discussed side by side, to allow for a direct comparison of their similarities and differences.

$PepT_{St}$ mutant	(equivalent) $PepT_{So}$ mutant	Helix	Expression ($PepT_{St}$)	Expression ($PepT_{So}$)
E22A	E21A	H1	yes	yes
E22Q	-	H1	yes	-
E25A	E24A	H1	yes	yes
E25Q	-	H1	yes	-
R26A	R25A	H1	no	yes
R26K	-	H1	yes	-
Y29A	F28A	H1	yes	yes
Y29F	F28Y	H1	yes	yes
Y30A	Y29A	H1	yes	yes
Y30F	Y29F	H1	yes	yes
R33A	R32A	H1	yes	yes
R33K	-	H1	no	-
-	H61A	H2	-	yes
-	H61K	H2	-	yes
-	H61D	H2	-	yes
Y68A	Y68A	H2	yes	yes
Y68F	Y68F	H2	yes	yes
K126A	K127A	H4	yes	yes
K126R	-	H4	yes	-
K126H	-	H4	yes	-
K126E	-	H4	yes	-
-	Y154A	H5	-	yes
-	Y154F	H5	-	yes
E299A	-	H7	no	-
E299D	-	H7	no	-
E299Q	-	H7	no	-
E300A	D316A	H7	yes	yes
E300D	-	H7	yes	-
E300Q	-	H7	yes	-
N328A	N344A	H8	-	yes
N328D	-	H8	yes	-
N328Q	N344Q	H8	yes	yes
E400A	E419A	H10	yes	yes
E400D	-	H10	yes	-
E400Q	-	H10	yes	-

Table 5-2: Summary of expression results for $PepT_{St}$ and $PepT_{So}$ binding site mutations.

5.3.1. Polar residues involved in proton coupling

A salt bridge network is observed in the $PepT_{St}$ structure, involving the side chains of Glu22, Glu25 and Arg26 on H1 (fig. 5-11, A). Likewise, the equivalent Glu21, Glu24 and Arg25 in $PepT_{So}$ also have closely adjacent side chains (fig. 5-11, B). The positions of these side chains are similar in both proteins, despite overall different conformations, suggesting they do not take part in any conformational rearrangements transitioning from the occluded state to the inward-open.

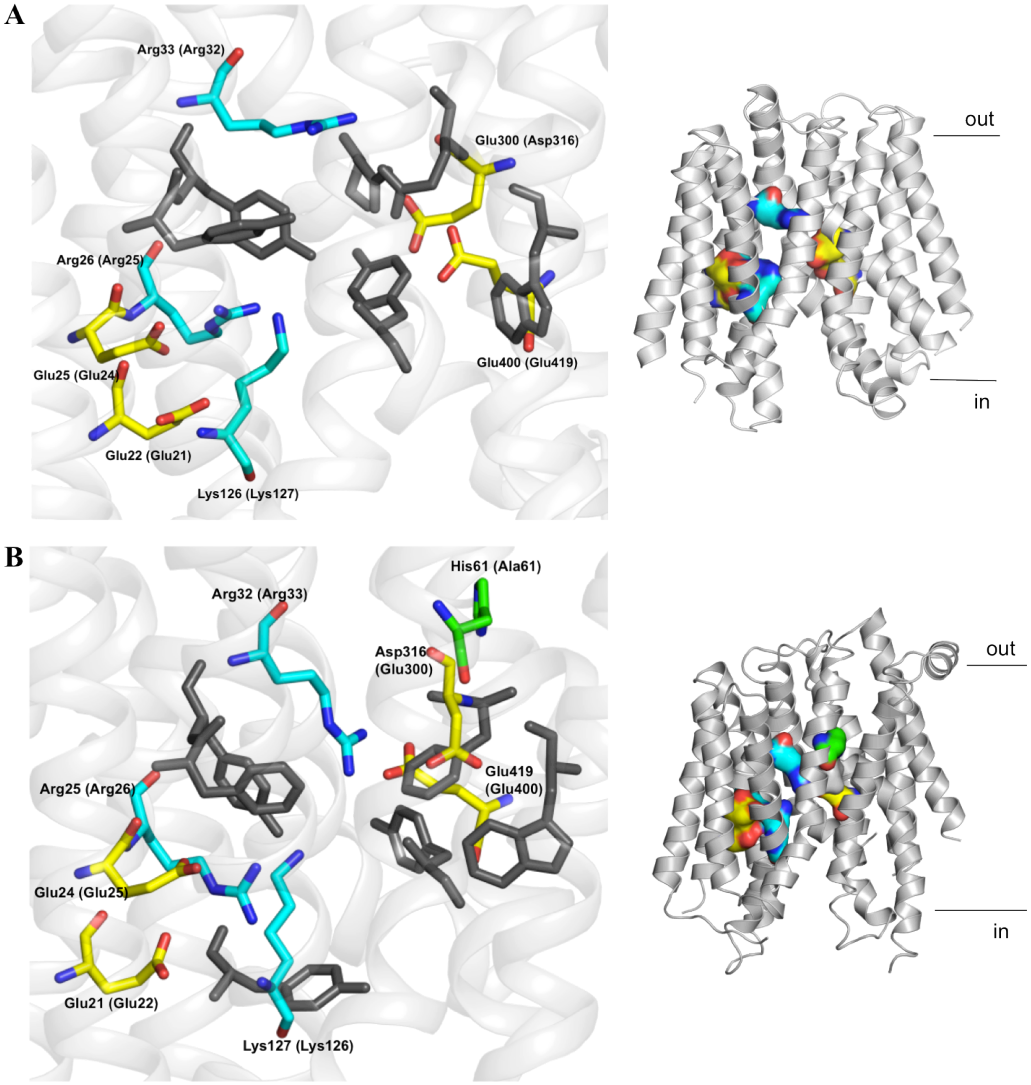


Fig. 5-11: The peptide binding sites of $PepT_{St}$ (A) and $PepT_{So}$ (B), with the polar residues mutated in this study outlined in colour (the other binding site residues in grey). The localisation of the residues in the protein structure is shown in surface representation (right). Equivalent residues from the other protein are listed between brackets. View from the plane of the membrane. PDB codes: 2APS ($PepT_{St}$) and 2XUT ($PepT_{So}$).

To test this, the residues were mutated to alanine and proteins were assayed for ΔΨ-driven transport and counterflow. In $PepT_{St}$, Arg26Ala did not express, so a lysine mutant was assayed instead. All mutations (including additional glutamine mutations for the glutamates in $PepT_{St}$ preserving side chain length) decoupled the transporter from the membrane potential (no ΔΨ-driven activity), while resulting in reduced counterflow activity (peptide-driven uptake) (fig. 5-12). This suggests a connection between the universally conserved ExxERFxYY motif (of which all three residues are part) and the proton coupling mechanism during peptide transport.

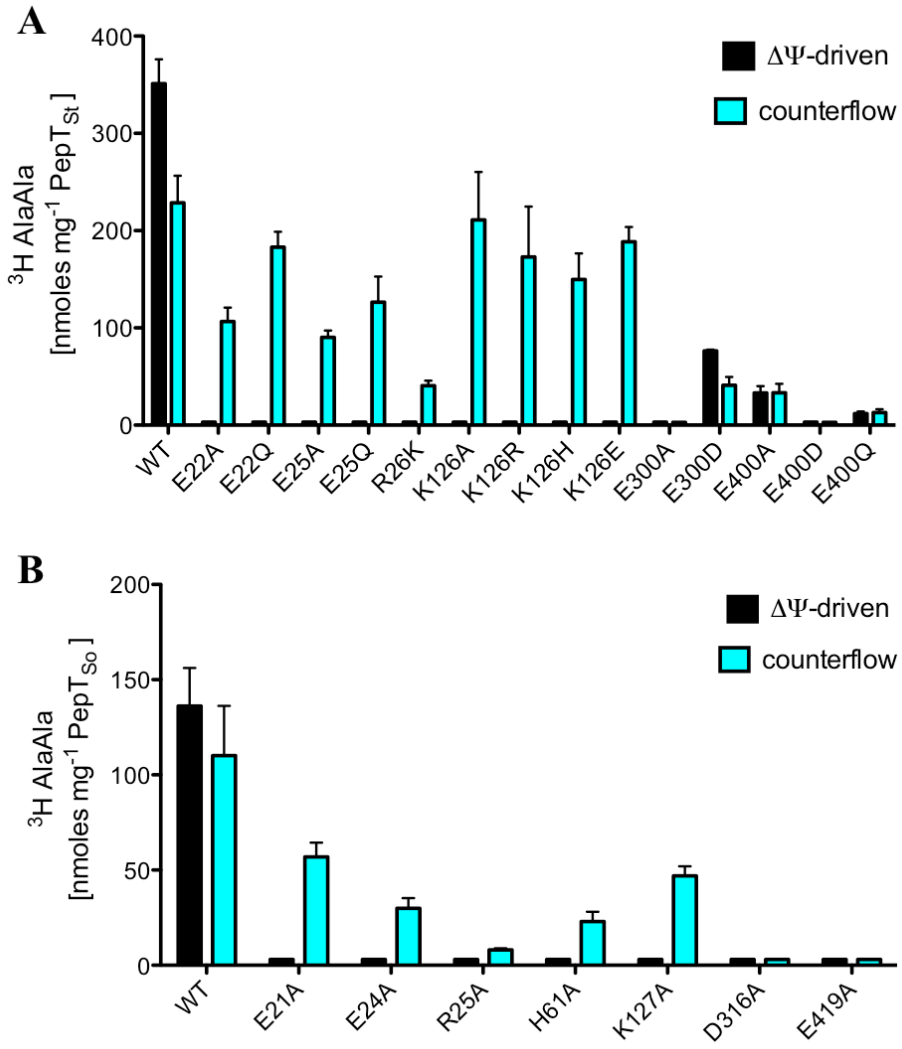


Fig. 5-12: Functional analysis of the role of polar amino acids outlining the central cavity.
(A) Effect of $PepT_{St}$ mutations on ΔΨ-driven (black) and peptide driven counterflow uptake (blue) is shown. Uptake of 15 µM ^{3}H-AlaAla by 1 µM protein was measured at pH 6.5 and 25°C after 2 min (ΔΨ-driven) and 5 min (counterflow).
(B) Effect of mutations on the equivalent residues in $PepT_{So}$ (where applicable). Measurements for both assays were taken after 10 min.

Protruding into the binding site is also the side chain of Lys126/127 (H4), universally conserved in the POT family. Its striking position in both structures, close to the charged residues on H1, prompted further investigation. When assayed, the alanine mutations behave similarly to those of the glutamates and arginine on H1, disrupting proton coupling without hindering counterflow (fig. 5-12, A and B). Interestingly, other mutations, some of them more conservative (arginine, histidine) have the same effect, with no detectable activity in the ΔΨ-driven assay (fig. 5-12, A). This points towards a central, highly specific role for the lysine in the coupling mechanism. Whether the residue is itself protonated or simply serves to position an acidic side chain during the transport cycle to facilitate its deprotonation remains open to discussion, as it is difficult to predict the pK_a of ionisable side chains. Nevertheless, Lys126/127 appears to be essential to proton coupling. In addition, the lysyl side chain may accommodate the carboxylate end of peptide substrates, as proposed for *E. coli* peptide transporter YjdL (Jensen *et al*, 2012b).

On the opposite side of the $PepT_{St}$ binding site, three glutamate side chains form an acidic cluster (fig. 5-11, A). Of these, the middle glutamate (Glu299) is not conserved (for example, in $PepT_{So}$ it is replaced by Phe315), but it seems to play a crucial structural/stabilising role, as none of the Glu299 mutants expressed (table 5-2). Despite great variability, side chains at this position may serve to accommodate peptide substrates (Jensen *et al*, 2012a), as will be discussed in more detail in the next subchapter.

The other two residues in the cluster (Glu300 and Glu400) are conserved as acidic residues in $PepT_{So}$ (Asp316 and Glu419 respectively). Alanine mutations of these amino acids severely reduce transport in both ΔΨ-driven and counterflow assays, with the Glu400Asp substitution completely abolishing it (fig. 5-12, A and B). In agreement with this, the analogous mutation in hPepT1 was shown to abolish transport (Bolger *et al*, 1998).

Furthermore, previous studies have suggested this glutamate recognises the amino group of the N-terminal residue of peptide substrates in rabbit PepT1 (Meredith *et al*, 2000), hPepT1 (Xu *et al*, 2009) and YjdL (Jensen *et al*, 2012a). As glutamates have been identified as protonation sites in other MFS transporters such as FucP (Dang *et al*, 2010) and LacY (Kaback *et al*, 2011), it seems likely that they may also play a similar role in peptide transporters. Either way, their presence is crucial to peptide uptake regardless of its driving source, indicating a more complex role in not just proton coupling but also the structural integrity of the transporter.

In the $PepT_{St}$ structure, Glu400 is in close proximity of another conserved residue on H8, Asn328 (Asn344 in $PepT_{So}$). Another example of a residue with a complex role, this asparagine will be discussed in the next chapter, along with Arg33/Arg32 on H1. The Glu300/Asp316 data will also be reprised in that context, as all three residues are involved in regulating conformational changes during transport.

5.3.2. The role of His61 in $PepT_{So}$ and PepT1-like transporters

His61 is found on H2, at the bottom of the extracellular cavity in the $PepT_{So}$ structure, close to Asp316 on H7 (fig. 5-11, B). The histidine is not universally conserved among POT family members, but it is found in both hPepT1 and hPepT2, where it has been identified as the main protonation site, as well as a few other bacterial transporters. Mutating it in the human proteins abolishes transport activity in HEK cells (Uchiyama *et al*, 2003) and *Xenopus* oocytes (Fei *et al*, 1998). Similarly, using histidine-specific inhibitors like diethypyrocarbonate and zinc on hPepT1 in oocytes and Caco-2 cells completely abolishes

uptake of both peptide and drug substrates (Terada *et al*, 1996; Okamura *et al*, 2003). Moreover, electrophysiological measurements on rabbit PepT1 variants have identified it as a critical H^+-binding site (Chen *et al*, 2000). Similarly, the His61Cys mutation in $PepT_{So}$ has the same effect in a whole cell assay (Newstead *et al*, 2010). This was investigated in more detail in the liposome-reconstituted system, as the "ΔΨ vs. counterflow" setup would make it possible to correctly identify the residue's function. Indeed, the His61Ala variant does not transport dialanine under ΔΨ-driven conditions, but still performs counterflow (fig. 5-12, B), suggesting the residue may indeed be a crucial site for protonation. Moreover, the lysine and aspartate variants expressed poorly, and then became unstable and precipitated during purification, which points towards a critical structural role as well as a functional one.

In comparison, $PepT_{St}$ has an alanine at this position, while in other bacterial transporters like GkPOT and YdgR this residue is a serine. In the latter's case, a whole cell study has ruled out the involvement of the serine in the transport mechanism (Malle *et al*, 2011). This may be an indication of divergence in the evolution of peptide transporters, to allow for a slightly different proton coupling mechanism in an otherwise conservative family of proteins. In the case of $PepT_{So}$ and the eukaryotic proteins, the presence of a titratable histidine side chain at the periplasmic/extracellular side of the central cavity could be seen as a refinement of the mechanism in order to allow for more successful adaptation to changes in environmental pH.

5.3.3. Aromatic side chains and substrate recognition

Three aromatic side chains protrude into the central cavity of $PepT_{St}$ and $PepT_{So}$: Tyr29 (Phe28), Tyr30 (Tyr29) and Tyr68 (Tyr68), contributed by helices 1 and 2 (fig. 5-13). The latter two are universally conserved in the POT family, while the first is always an aromatic residue. These residues have been suggested to play a role in determining substrate specificity for rabbit PepT1 in oocytes (Pieri *et al*, 2009), and a similar role was proposed for the equivalent tyrosines in a recent study on bacterial transporter $PepT_{So2}$ (Guettou *et al*, 2013).

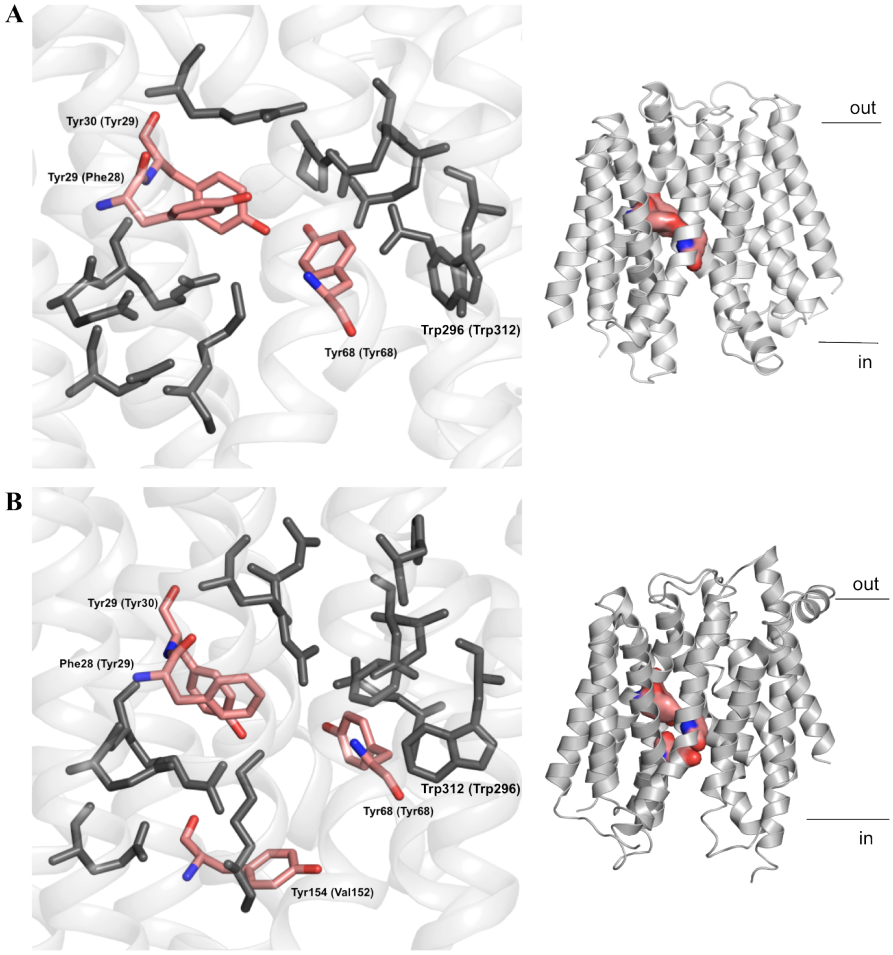


Fig. 5-13: The peptide binding sites of $PepT_{St}$ (A) and $PepT_{So}$ (B), with the aromatic residues mutated in this study outlined in colour (the other binding site residues in grey). The localisation of the residues in the protein structure is shown in surface representation (right). Equivalent residues from the other protein are listed between brackets. View from the plane of the membrane. PDB codes: 2APS ($PepT_{St}$) and 2XUT ($PepT_{So}$).

In $PepT_{St}$, phenylalanine mutations of Tyr29 and Tyr68 do not influence transport (either ΔΨ-driven or counterflow), whereas alanine substitutions result in decreased uptake compared to the wild type (fig. 5-14, A). Therefore, the hydroxyl groups of these residues are not involved in transport, but shortening the side chain is detrimental, suggesting a role for these tyrosines in the recognition/stabilisation of the substrate within the binding site. In $PepT_{So}$, the opposite Phe28Tyr mutation also does not influence uptake; similarly, the bulky side chain is required at Tyr68 (the alanine mutation yields decreased uptake levels), but not so the hydroxyl group (the tyrosine mutation has no noticeable effect) (fig. 5-14, B). This observation is in line with the competition data (fig. 5-9), the overall substrate preference of PepT1 and PepT2, known to have high affinities for hydrophobic substrates (Eddy *et al*, 1995; Biegel *et al*, 2006), as well as the close connection between binding site aromatic groups and hydrophobic substrates in the MFS in general, as discussed in the introduction.

The only one of these three amino acids where the OH group is required for proton coupling is Tyr30 in $PepT_{St}$ (Tyr29 in $PepT_{So}$), its side chain close to the guanidinium group of Arg26 (Arg25, respectively) (fig. 5-1). In $PepT_{St}$, the change to a phenylalanine terminates ΔΨ-driven uptake, whereas counterflow activity is still detectable; meanwhile, the alanine substitution is completely inactive in both assays (fig. 5-14, A). In contrast, the equivalent $PepT_{So}$ phenylalanine mutation loses both ΔΨ-driven and counterflow activity, but its mutation to alanine partially rescues activity in the first assay (fig. 5-14, B). Concerning the partial rescue of coupling by the alanine substitution in $PepT_{So}$, it could be speculated that an alternative polar residue could compensate for the absence of the OH group, albeit only when the benzyl side chain is removed. It is unclear what this other residue may be or if indeed a replacement is necessary in $PepT_{So}$, but one potential candidate is the universally conserved Asn158 on H5, the equivalent of which is speculated to enter the binding site of peptide transporter GkPOT during the transport cycle, and is essential to both types of transport (Doki

et al, 2013). Notwithstanding, the data suggests a more subtle significance for Tyr29/Tyr30, wherein the hydroxyl group is used for proton coupling while the aromatic group itself serves a different structural role in each protein.

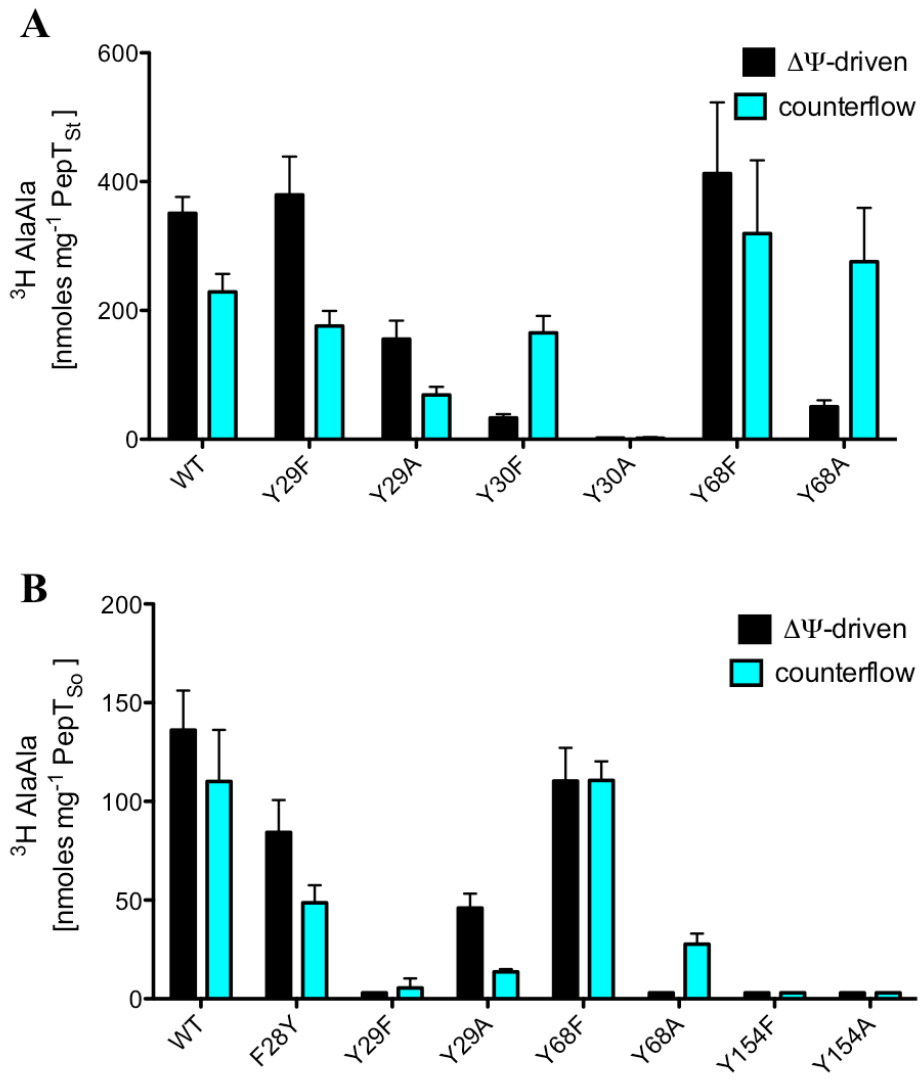


Fig. 5-14: Aromatic side chains mediate substrate recognition in the binding site.
(A) Effect of mutating aromatic residues in $PepT_{St}$ on ΔΨ-driven (black) and peptide driven counterflow uptake (blue).
(B) Effect of mutating the equivalent residues in $PepT_{So}$ (where applicable).

In addition to these aromatic groups, the binding site also contains the side chain of Trp296/Trp312 (H7). When compared to the inward-open structure of $PepT_{St}$, in $PepT_{So}$ (occluded) the indole ring is rotated to face the central cavity, due to the repositioning of the

cytoplasmic half of H7 that closes the intracellular gate and brings the transporter into the occluded state (fig. 5-13). The function of this conserved tryptophan has been investigated in *E. coli* peptide transporter YjdL, where it is proposed to interact with the conserved tyrosines in the binding site (Jensen *et al*, 2012a). In GkPOT, the equivalent Trp306 loses proton coupling, but not counterflow function, when exchanged to alanine (Doki *et al*, 2013). Similarly, leucine and serine mutations of the equivalent Phe289 in YdgR abolish uptake of fluorescent peptide in whole cells (Malle *et al*, 2011).

Lastly, another tyrosine, Tyr154, sits at the bottom of the central cavity in $PepT_{So}$, where helices 5 and 10 pack against one another to occlude the $PepT_{So}$ binding site from the cytoplasm (fig. 5-13, B). Tyr154 is structurally essential; compared to the other conserved tyrosines in the binding site, it appears to be more important, since not only the alanine, but also the phenylalanine mutation results in abolished uptake (fig. 5-14, B). In the structure, Tyr154 reaches across to the C-terminal domain, while stacking against Phe150 below it in a π-π interaction (not shown). Phe150, like Tyr154, is part of the PTR2_2 signature motif, and their interaction may serve to stabilise the occluded state of the transporter by sealing the binding site at the cytoplasmic side. This hypothesis is reinforced by the inward-open structure of another *S. oneidensis* transporter, $PepT_{So2}$, in which the phenylalanine side chain rotates away from the tyrosine, opening the intracellular cavity (Guettou *et al*, 2013).

In addition to the structural role, the complete lack of activity in both transport assays by the Tyr154Phe variant points to the essential nature of the OH group. This hypothesis is in agreement with findings on the role of Tyr167 in human PepT1 (Yeung *et al*, 1998), rabbit PepT1 (Pieri *et al*, 2009) and the equivalent Tyr156 in YdgR (Malle *et al*, 2011), which identify the residue as involved in either substrate binding or translocation. Although highly conserved among POT proteins, the tyrosine is replaced by a valine in $PepT_{St}$, whose role in the transport mechanism remains to be examined.

5.4 Sequence variability in the binding site dictates peptide specificity

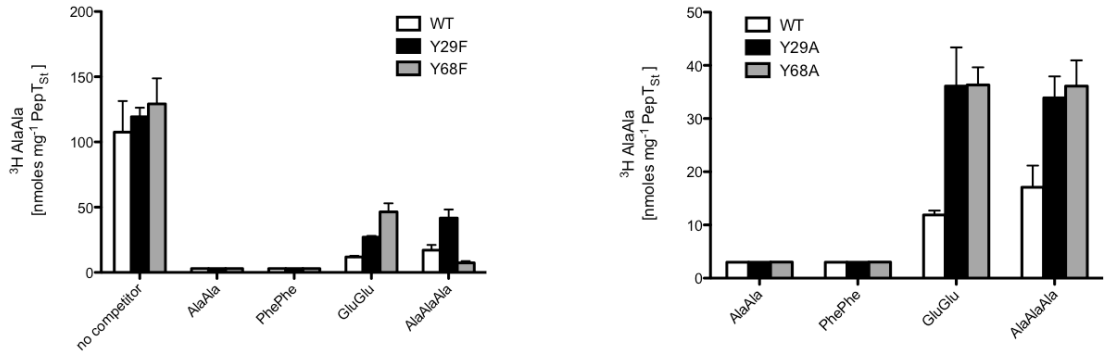
The site-directed mutagenesis approach served to identify key amino acids in the peptide binding site that are involved in aspects of substrate recognition and coupling. While substrate recognition was shown to depend mainly on the presence of aromatic side chains, the data does not rule out the involvement of other residues. The use of a small, neutral dipeptide (AlaAla) as the sole substrate for the mutant transport assays clearly limits the interpretation of the results, but fortunately the di/tripeptide competition data (fig. 5-9) provides more information on how other substrates might be accommodated and which other residues might be involved. Together, the mutagenesis and competition data can be traced back to the crystal structures to identify the features of POT binding sites that may account for how these differences arise.

For a first example, as Tyr29 and Tyr68 ($PepT_{St}$) are thought to be involved in substrate recognition, individual competition experiments (competition at 2 mM plus IC_{50} measurements) were performed to test the effect of the respective phenylalanine and alanine substitutions on the affinities for four peptide substrates (fig. 5-15).

In the ΔΨ-driven competition assay (fig. 5-15, A), the phenylalanine mutants show changes in their affinity for GluGlu and AlaAlaAla compared to the wild type transporter, while the alanine mutants completely lose their affinity for both substrates but continue to transport AlaAla and PhePhe. The data shows that Tyr29 and Tyr68 are directly involved in regulating peptide specificity. To quantify their individual contributions, IC_{50} values were

calculated (fig. 5-15, B). These show more than a 3-fold reduction in the affinity for AlaAlaAla in the case of the Tyr29Phe mutant (1.4 mM compared to 0.4 mM for the wild type protein), while its affinity for GluGlu is unchanged. As this residue is natively a phenylalanine in $PepT_{So}$, the absence of a hydroxyl group at this position may explain the slightly lower affinity of $PepT_{So}$ for AlaAlaAla, as shown in the competition data (fig. 5-9). In contrast, the Tyr68Phe mutant has decreased affinity for GluGlu (1.63 mM compared to 0.56 mM), but transports AlaAlaAla as well as the wild type.

A



B

IC_{50} (mM)	WT	Y29F	Y68F
GluGlu	0.6 ± 0.06	0.5 ± 0.05	1.6 ± 0.1
PhePhe	0.005 ± 0.002	0.006 ± 0.004	0.02 ± 0.002
AlaAla	0.04 ± 0.002	0.05 ± 0.006	0.05 ± 0.005
AlaAlaAla	0.4 ± 0.06	1.4 ± 0.02	0.35 ± 0.003

Fig. 5-15: Binding site tyrosines and peptide affinities.
(A) Effect of phenylalanine and alanine substitutions of Tyr29 and Tyr68 on the ability of four oligopeptides to compete against AlaAla uptake in a ΔΨ-driven assay. Competing peptides at 2 mM, residual ^{3}H-AlaAla uptake measured after 2 min.
(B) Change in IC_{50} values after substitution of Tyr29 and Tyr68 to phenylalanine.

While the tyrosines on H1 and H2 regulate peptide specificity, a look at the other binding site residues identifies more potential contributors. For example, the difference in the affinity for GluGlu (inhibits AlaAla uptake completely in $PepT_{So}$, but only decreases it to 23% of normal activity in $PepT_{St}$) (fig. 5-9) may be explained by the presence of the benzyl group of Phe315 just underneath the acidic side chain at Asp316 (fig. 5-1, B). Although the

phenylalanine itself is not conserved in the POT family, the side chain found at this position is thought to interact with the N-terminal side chain of dipeptides (Jensen *et al*, 2012a). In $PepT_{St}$, this residue is a glutamate (Glu299) (fig. 5-1, A), which may repel the acidic side chain in GluGlu, thus reducing the transporter's affinity for it.

This glutamate could also explain the strong preference shown by $PepT_{St}$ for another dipeptide, GlyGln (in contrast, uptake by $PepT_{So}$ is only barely affected by GlyGln). The glutamine side chain of the peptide would provide Glu299 with a positive charge to interact with, thus stabilising the peptide-bound intermediate state of the transporter. On the other hand, hPepT1, which $PepT_{So}$ most closely resembles, also has a Phe at this position, which is critical to *in vivo* uptake of dipeptide analogue glycylsarcosine (Kulkarni *et al*, 2003). Its backbone is also thought to form hydrogen bonds to the C-terminal carboxylate of dipeptides in docking experiments using a homology model of hPepT1 (Pedretti *et al*, 2008). In short, side chain variability at this position could be either the cause or the evolutionary consequence of different substrate requirements in different peptide transporters.

Another position of variable sequence among POT members is Glu300/Asp316 ($PepT_{St}$ and $PepT_{So}$ numbers). For example, mutating the glutamate at this position to a glutamine is crucial to crystallising bacterial peptide transporter GkPOT with modified dipeptide alafosfalin (Doki *et al*, 2013). More recently, $PepT_{So2}$, a transporter in which this residue is natively an asparagine, was also co-crystallised with alafosfalin, reinforcing the side chain's significance. Taking the example of GkPOT, the other four residues that coordinate the alafosfalin molecule in its binding site (an arginine, two tyrosines and an asparagine) are all conserved in $PepT_{St}$ and $PepT_{So}$ (fig. 5-16), with just Gln310 being different. Therefore, the difference in relative affinities for alafosfalin between $PepT_{St}$ and $PepT_{So}$ (fig. 5-9) might be connected to the different side chains at this position. Having an aspartate rather than a glutamate here could improve the strength of the ligand coordination in $PepT_{So}$ by providing

more space, thus explaining the increased affinity for alafosfalin. Indeed, like the adjacent Phe315, the residue at position 316 is subject to great variability among peptide transporters (it is also an aspartate in hPepT1 and hPepT2, but changes to alanine, serine, glutamine and phenylalanine in other bacterial homologues). The data so far suggest this variability may be responsible for protein-specific changes in substrate preference.

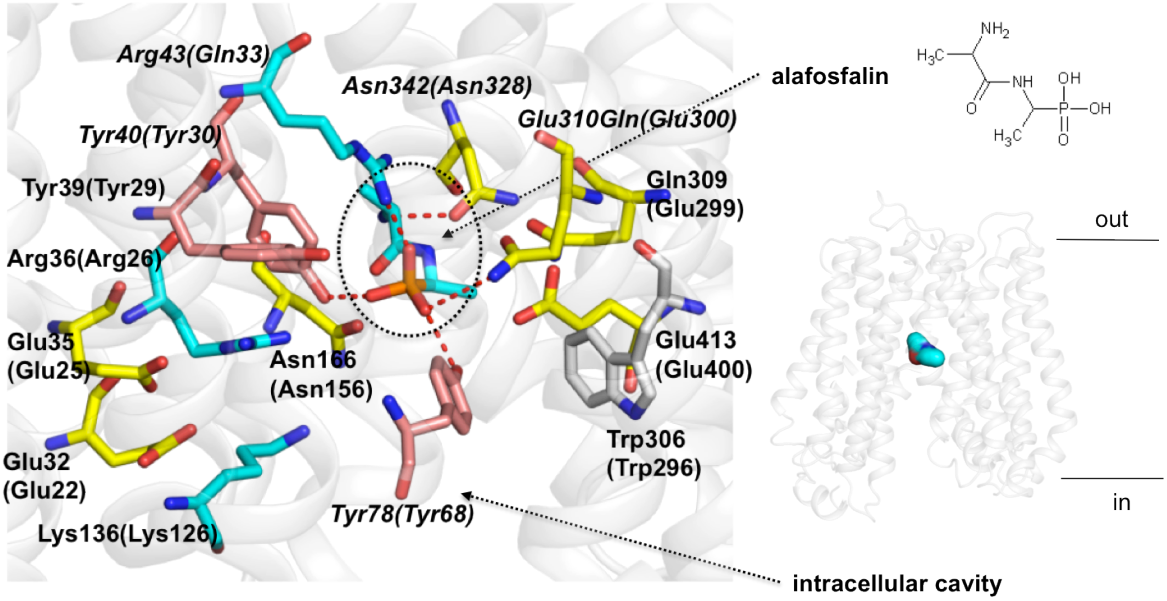


Fig. 5-16: Binding site of GkPOT, a bacterial peptide transporter in an inward-open conformation, with bound peptide analogue alafosfalin (chemical structure on the right). The GkPOT structure (PDB:4IKZ) is shown on the right, with alafosfalin represented as spheres. The central cavity is open to the intracellular side of the membrane, and the alafosfalin molecule is coordinated by five conserved residues (in italics). Equivalent $PepT_{St}$ residue numbers are listed in brackets.

However, whether the N- or C-terminal side chain of a dipeptide substrate is preferentially accommodated at this position remains an open question. For instance, an interaction for both the N- and C-terminal side chains has been proposed with the more conserved Ser423 on H10. The equivalent Ser626 in hPepT2 is thought to accommodate the nitro group of high affinity dipeptide derivative inhibitor LysZNO_2-Ala (Pedretti *et al*, 2011). This serine is conserved in most POT's, with the exception of the "non-canonical" *E. coli* transporters YjdL and YbgH, where it is replaced by an aspartate. This substitution is believed to account for the reported preference shown by these two transporters for a lysyl group at the

C-terminus of dipeptides (Casagrande *et al*, 2009, Ernst *et al*, 2009). Indeed, neither $PepT_{So}$ nor $PepT_{St}$ transport LysLys, while rabbit PepT1 and PepT2 display modest affinity for GlyLys (Steel *et al*, 1997, Döring *et al*, 2002) – but more data on a variety of peptides would be needed to establish whether the serine is involved in regulating substrate specificity.

Finally, some notable differences concern the affinity of the two proteins for tripeptides. Non-aliphatic tripeptides (GlyGlyGly, LysLysLys, TyrTyrTyr) have no detectable effect on dialanine uptake by $PepT_{st}$. However, of these, TyrTyrTyr abolishes uptake via $PepT_{So}$ completely, while the others do not compete (fig. 5-9). Triglycine may be too small to form the necessary contacts for tripeptide binding and translocation, although previous evidence exists of its uptake in both intestinal epithelia and basolateral kidney cells (Thwaites *et al*, 1994, Terada *et al*, 2000). Positively charged peptides such as di- and trilysine are likely transported in their neutral form (Daniel, 2004), so the pH of the assay may be too low for these to show any inhibitory effects.

The difference in TyrTyrTyr affinity between $PepT_{So}$ and $PepT_{St}$ suggests the notion of a more accommodating binding site in $PepT_{So}$, in which a distinct configuration is assembled for holding the bulky tripeptide in the occluded, substrate-bound conformation during transport. In keeping with this model, a recent paper from our group (Lyons *et al*, in preparation) reports the crystal structures of $PepT_{St}$ with a dipeptide (AlaPhe) and a tripeptide (AlaAlaAla), both transported substrates. In the structure with AlaPhe, the dipeptide is oriented horizontally, while in the other structure the tripeptide sits vertically in the binding site and forms fewer contacts, the two being coordinated by different residues (fig. 5-17).

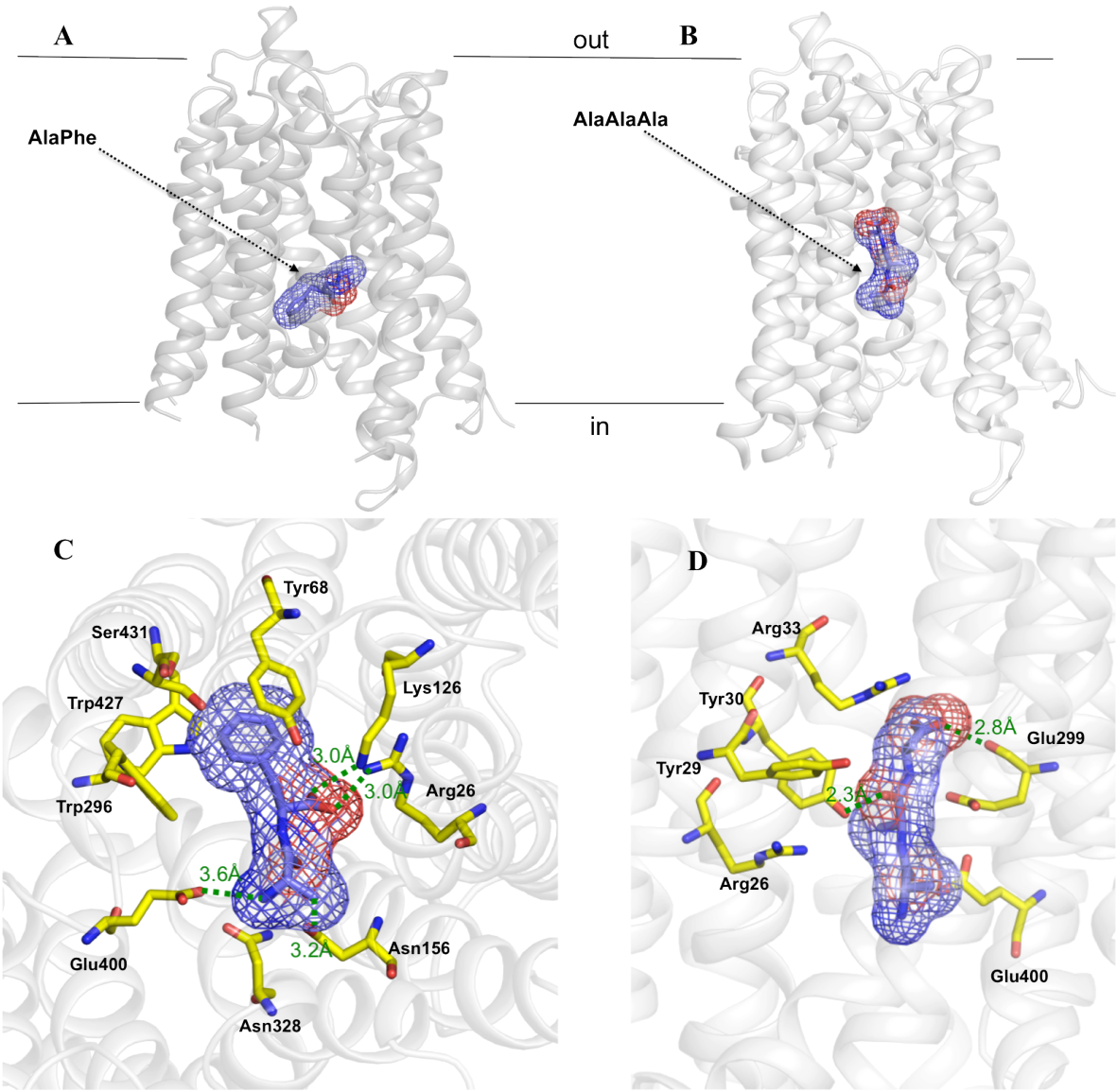


Fig. 5-17: $PepT_{St}$ structures with AlaPhe and AlaAlaAla.
(A) $PepT_{St}$ viewed in the plane of the membrane with dipeptide AlaPhe bound in horizontal orientation.
(B) $PepT_{St}$ with tripeptide AlaAlaAla bound vertically in the central cavity.
(C) Top view of $PepT_{St}$ showing AlaPhe (purple) in horizontal orientation, coordinated within the binding site. Hydrogen bonds shown as dashed green lines, with distances in Ångströms.
(D) View of $PepT_{St}$ in the plane of the membrane, showing AlaAlaAla (purple) positioned vertically within the binding site, coordinated by different residues.

The two structures are the first to show physiological peptides bound to an oligopeptide transporter. Both of them confirm that binding site residues are organised to form a dipole across the central cavity, serving to orient the peptide. In contrast to the GkPOT/alafosfalin structure (alafosfalin is not transported by GkPOT, or transported with

low affinity, as will be discussed in chapter 8), the $PepT_{St}$/AlaPhe also confirms the predicted interactions of Glu400 and Lys126 with the amino group and the carboxyl group of the dipeptide, respectively. Of significance is the involvement of Tyr68 in coordinating AlaPhe (as seen above, the Tyr68Ala mutation changes the affinity of the transporter for another dipeptide, GluGlu); in contrast, AlaAlaAla is in close proximity of another tyrosine, Tyr29 (which was found to regulate the affinity for AlaAlaAla). Also interestingly, in the AlaAlaAla structure, the carboxyl group of the peptide contacts the carbonyl of Glu299, pointing back to the role that sequence variability at this position plays in determining how different substrates are accommodated.

The different orientations of the two peptides are likely to illustrate different binding modes for di- and tripeptides within the same binding site, with a high affinity binding configuration for dipeptides and a low affinity one for tripeptides. This is particularly relevant to understanding drug transport, as many of the PepT1/PepT2 drug substrates like cephalosporin β-lactam antibiotics match the stereochemistry of the L-isomers of tripeptides, and thus are believed to be oriented similarly (Bailey *et al*, 2000).

5.5. Conclusions

Building on the two POT crystal structures, this parallel study on $PepT_{So}$ and $PepT_{St}$ was undertaken for two main reasons: to characterise and compare the transporters in terms of their peptide preference and to investigate the roles of binding site residues in regulating the two aspects of secondary transport (substrate specificity and coupling). Peptide preference

was studied through direct uptake assays (in whole cells and proteoliposomes), kinetics and competition assays, calculating and comparing relative affinities and identifying transported substrates through counterflow assays with different peptides. The coupling question was addressed by testing the activity of protein variants obtained by site-directed mutagenesis of the binding site, in both ΔΨ-driven and counterflow experiments, thereby distinguishing between residues involved in substrate recognition, coupling or both.

Although it showed direct uptake of labelled substrates, the *in vivo* assay had a number of drawbacks: the reliance on His-tagged constructs instead of GFP-tagged ones (making precise quantification of expression levels difficult), occasionally high backgrounds caused by endogenous transport, and the requirement for non-hydrolysable substrates. As a result, a reconstitution protocol was established and assays were designed to test ΔΨ- and substrate-driven transport of radioactive substrates (a third assay was set up in which a pH gradient would drive transport, but no uptake was detected).

Like other MFS transporters, $PepT_{St}$ and $PepT_{So}$ were functional under both ΔΨ-dependent and counterflow conditions, and they were shown to transport AlaAla and GlySar at different rates, echoing the results of earlier studies on DtpT and the *E. coli* POTs. AlaAla transport was found to be optimal at pH values between 5.5 and 7.5, and a Michaelis-Menten kinetic profile was calculated for GlySar indicating low affinity for the substrate, in agreement to similar studies on eukaryotic PepT proteins in whole cells. Furthermore, competition experiments confirmed the exclusive preference for L-isomers of di- and tripeptides and the proton-coupled nature of uptake, as well as a detailed profile of di- and tripeptide specificity for both proteins. Together, the data on the wild type $PepT_{St}$ and $PepT_{So}$ establish them as valid working models for studying eukaryotic peptide transport.

The binding site mutagenesis study reveals many similarities between the proteins with regard to the organisation of the binding site and the conservation of residues therein. Residues of the ExxERFxYY motif in particular are involved in the coupling aspect of transport, as are several other charged residues on both N-terminal and C-terminal halves of the binding site. Meanwhile, the tyrosines in the centre of the cavity are mainly responsible for recognising substrates and regulating affinity. While the distinction between a structural and a proton coupling role is not yet clear for some of these residues, this study shows that several of them are most likely involved in the proton translocation mechanism, either by taking direct part or by mediating conformational changes that are required to enable it.

Sequence variation at certain equivalent positions in $PepT_{St}$ and $PepT_{So}$ (such as the presence of His61 in the latter) hint at a possibly different coupling mechanism for $PepT_{So}$, one that may involve a titratable side chain and a slightly different hydrogen bonding network between polar residues in the central cavity. It is possible that $PepT_{So}$ and $PepT_{St}$ have evolved to accommodate different ways to couple peptide transport to the proton gradient and/or the membrane potential. The presence of His61 at the border between the central binding site and the extracellular cavity may well be representative of this interpretation. However, most conserved aspects are found in both proteins, such as the interaction between the glutamates and arginine on H1, the role of the aromatic side chains and the lysine at the bottom of the cavity, as well as three inter-domain salt bridges that will be discussed in the following chapter.

More data is needed to fully understand how peptide specificity is regulated. It appears that, despite an overall similarity in substrate preference between POT family members (no doubt a consequence of conserved interactions within the binding site that are needed to recognise the amino and carboxyl termini of the peptides), there can be variation, such as the increased and decreased affinities for GluGlu/alafosfalin and GlyGln, respectively, in the case

of $PepT_{So}$. As in the case of the coupling mechanism, it is apparent that sequence variability at certain positions can give rise to significant differences in substrate specificity from one transporter to another. For instance, the identity of the residues at positions Tyr29/Phe28, Tyr68 and Glu299/Phe315 is likely to influence the way different peptides are accommodated. This is supported by comparison with the recently solved substrate-bound structures of $PepT_{St}$, which also provide valuable clues to explain the striking differences in tripeptide affinity between $PepT_{So}$ and $PepT_{St}$, and serve to extrapolate the findings of this study to understand how peptide and drug specificity is regulated in the mammalian peptide transporters.

Chapter 6: Conformational changes and the transport cycle

As discussed in the introduction, the alternating access model for secondary active transport has evolved in recent years, from the initial "rocker-switch" theory (coordinated rigid body motion of the N- and C-terminal domains during transport) to a more detailed model involving a series of helical movements and salt bridge configurations, both between and within the two domains of the transporters. Specifically in the case of the Major Facilitator Superfamily, crystal structures of transporters from several subfamilies were solved (table 1-2), revealing similarities as well as differences. Structural and functional data were ultimately assembled to advance individual models for transport and coupling.

Generating a unified model for transport is made difficult by the fact that crystal structures are restricted to a singular conformation, and trapping the same transporter in different states has proven challenging. Even though structures of MFS transporters in several conformational states are currently available (fig. 1-10), they are not directly comparable, as the details of the transport cycle are likely to differ in each case. Therefore, studying the conformational changes involved in the transport mechanism typically relies on a combination of transport assays and other biochemical methods (such as FRET and EPR), homology modeling and molecular dynamics simulations, as exemplified by LacY in the introduction. In particular, the discovery of inverted repeat topology has provided a useful

tool for generating structural models of transporters in alternate conformations based on available crystal structures.

The structures of $PepT_{So}$ and $PepT_{St}$ (Newstead *et al*, 2011; Solcan *et al*, 2012) are the earliest example of MFS transporters from the same family being crystallised in different conformations (inward-occluded and inward-open, respectively). As discussed in the $PepT_{St}$ structural analysis, the two structures can be superimposed with an overall RMSD of 1.68Å, but the positions of the TM helices in the N-terminal domain (H1-H6) remain relatively unchanged in the two structures, whereas the C-terminal domain (H7-H12) undergoes a few notable helical rearrangements (fig. 4-6). Using the crystal structure comparison as the starting point, an inquiry was made into the conformational aspects of the peptide transport cycle, using site-directed mutagenesis, transport assays and data from molecular modeling, MD simulations and double electron-electron resonance (DEER) experiments.

6.1. Alternating access to the central cavity

Despite protein-specific differences, a unifying feature of the transport mechanisms among the MFS seems to be the role of salt bridges in regulating the alternating access of substrates to the binding site, by mediating the opening and closing of "gates" (Forrest *et al*, 2011). Substrate binding in the outward-open state of LacY was proposed to disrupt a salt bridge between a glutamate and an arginine on opposite sides of the central cavity, initiating the transition to the inward-open state (Abramson *et al*, 2003). Other interhelical salt bridges were identified as crucial to transport through modeling of the outward-open state (Guan and Kaback, 2006), and were later confirmed by the inverted repeat swap model of LacY

(Radestock and Forrest, 2011). MD simulations based on the crystal structure of GlpT also indicated that salt bridges between basic and acidic residues stabilise individual conformations in the transport cycle (Law *et al*, 2008a). More recently, a comparison between three conformations of XylE identified three pairs of residues forming inter-domain salt bridges to facilitate substrate binding at the periplasmic side, the shift to the occluded intermediate state, and substrate release at the cytoplasmic side (Quistgaard *et al*, 2013).

6.1.1. Inter-domain salt bridges facilitate alternating access

In the inward-open structure of $PepT_{St}$, two such inter-domain salt bridges are observed between highly conserved residues bridging the N- and C-terminal helical bundles and stabilising the extracellular gate (H1/H2 and H7/H8) in the closed position (Solcan *et al*, 2012). One of them, between Arg53 (H1) and Glu312 (H7), seals the transporter at the extracellular end, while the other one links Arg33 (H1) and Glu300 (H7) at the top of the binding site, locking the central cavity in the inward-open position (fig. 6-1, A, red). A third pair of conserved, oppositely charged residues is found at the lower end of the binding site: Lys126 (H4) and Glu400 (H10); the latter do not interact in the inward-open structure, but a similar interaction was observed at the same position stabilising the outward-open structure of fucose transporter FucP (Dang *et al*, 2010). In comparison, of the equivalent residues in $PepT_{So}$, only Arg32 and Asp316 (the central pair) interact in the inward occluded structure (fig. 6-1, B, grey).

To generate a structural representation of the peptide transport cycle, our collaborators Dr Lucy Forrest and Dr Sebastian Radestock at the Max Planck Institute of Biophysics in Frankfurt have modeled $PepT_{St}$ and $PepT_{So}$ in alternate conformational states. The occluded state of $PepT_{St}$ was modeled using the $PepT_{So}$ structure (homology modeling), and an outward-open model was generated through swapping inverted repeat elements in the inward-open structure (H1-H3 with H4-H6 and H7-H9 with H10-H12). While both the upper and central salt bridges are observed in the structure, the occluded state only features the central contact, with the upper residues farther apart as H1 and H7 begin to separate and open the extracellular gate (fig. 6-1, A, grey). The outward-open model (fig. 6-1, A, green) sees the extracellular gate fully open, the upper and central residues disassociated and the proposed salt bridge between Lys126 and Glu400 holding the protein in the outward-open state at the intracellular end.

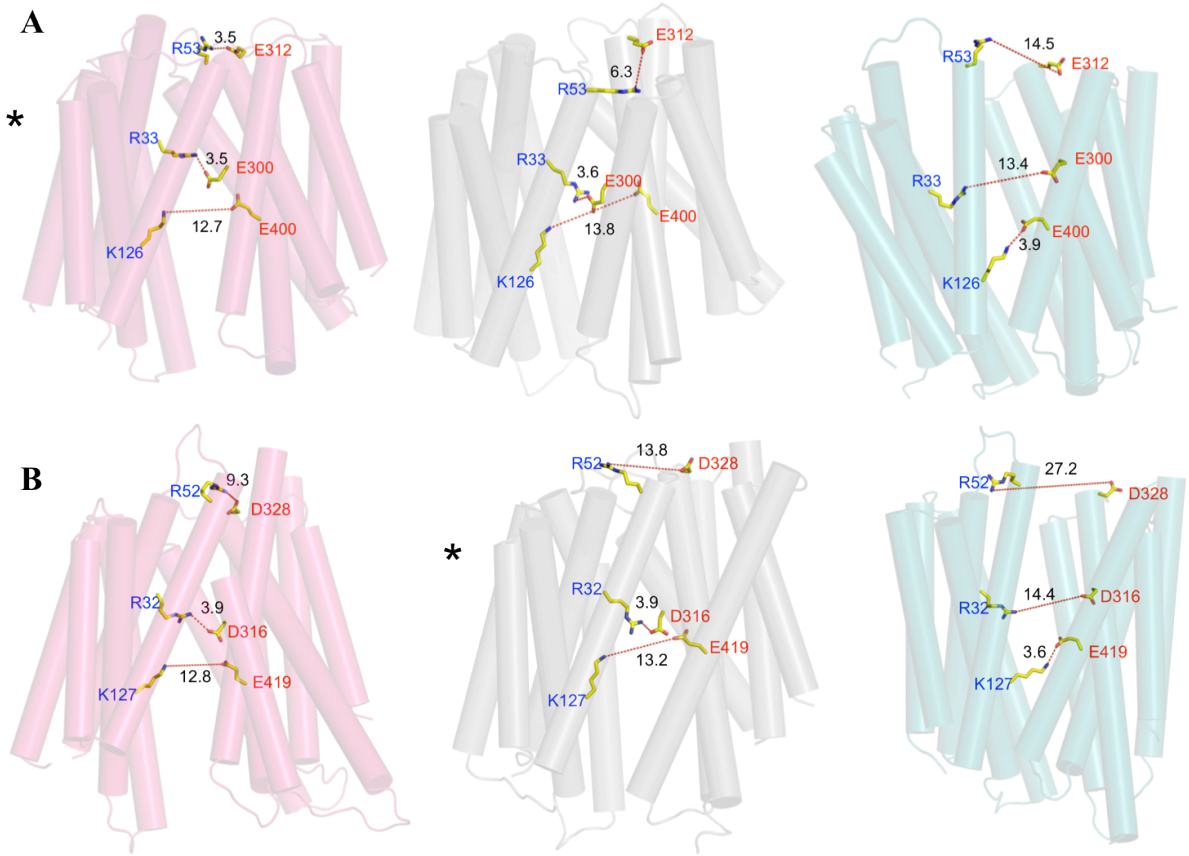


Fig. 6-1: Mapping the polar contacts between pairs of charged residues bridging the gap between the N- and C-terminal domains of $PepT_{St}$ and $PepT_{So}$. All distances in Ångströms. Crystal structures are marked with an asterisk.
(A) $PepT_{St}$: inward-open structure (red), occluded model (grey) and outward-open model (green).
(B) $PepT_{So}$: inward-open model (red), occluded structure (grey) and outward-open model (green).

Similarly, outward- and inward-facing conformations of $PepT_{So}$ were modeled based on the occluded crystal structure of $PepT_{So}$ and the inward-open structure of $PepT_{St}$, respectively. As proposed, Lys127 and Glu419 (the lower pair) come together within hydrogen-bonding distance in the outward-open model of $PepT_{So}$ (fig. 6-1, B, green) when compared to the occluded structure (grey) and the inward-open model (red). The central salt bridge seen in the $PepT_{So}$ structure (Arg32-Asp316), also observed at the same position in $PepT_{St}$, is preserved in the inward-open state, but breaks when the transporter faces outward. Finally, the top pair of residues, Arg52 and Glu328, although never forming a clear polar contact, is significantly closer in the inward-open model.

The amino acids involved in forming the salt bridges were mutated and their function was tested in ΔΨ-driven and counterflow assays. In $PepT_{St}$, alanine mutations of Arg53 and Glu312 reduce uptake by up to 70%, while a charge-swapped variant (R53E/E312R) partially rescues activity to about 50% of wild type levels (fig. 6-2, A). This result suggests a supportive role for the upper salt bridge, and implies that the distribution of charges at these positions is not essential; indeed, this salt bridge is conserved in other POTs but the charges are often reversed (acidic residue on H1, basic residue on H7), such as in GkPOT (Glu63-Lys322) and $PepT_{So2}$ (Asp46-Arg304). Correspondingly, the upper salt bridge in $PepT_{So}$ appears to play a non-crucial role in the transport mechanism, with alanine mutations of both Arg52 and Asp328 resulting in reduced, but not abolished ΔΨ-dependent uptake, while having no effect on counterflow (fig. 6-2, B).

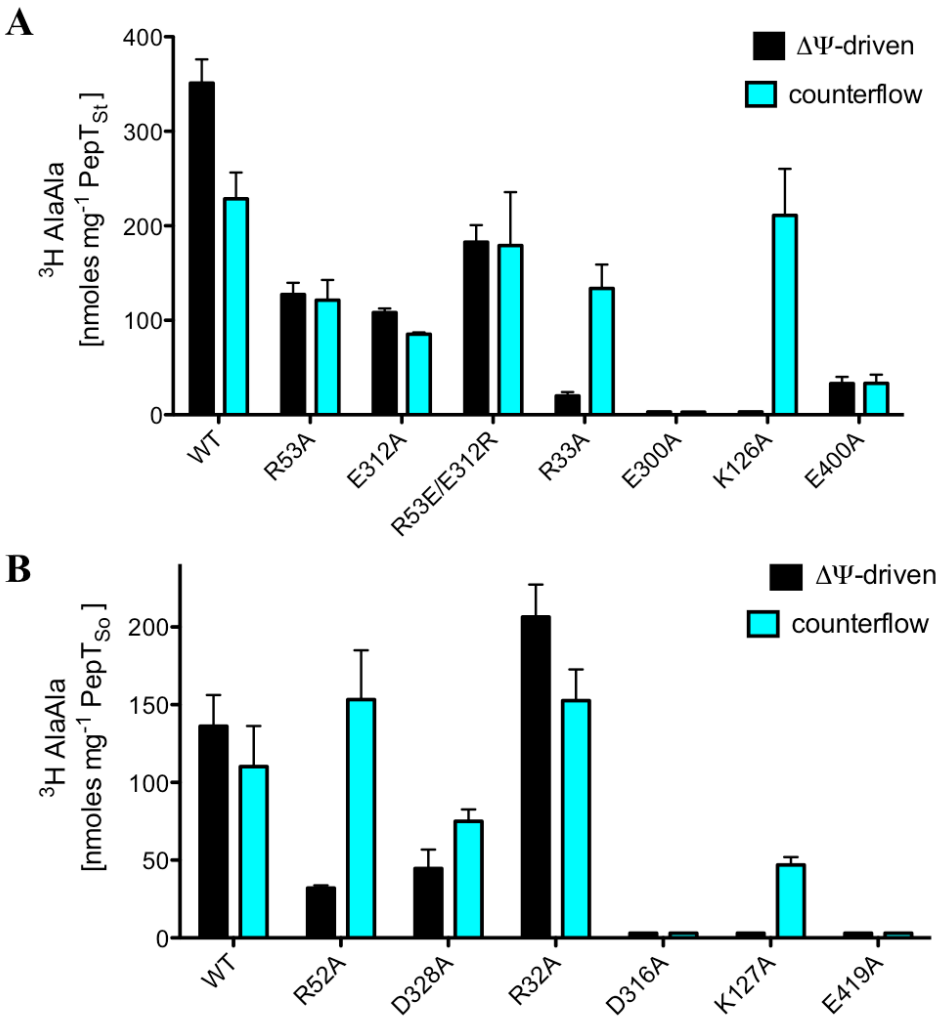


Fig. 6-2: Functional role of inter-domain salt bridges in $PepT_{St}$ and $PepT_{So}$.
(A) Effect of mutating the three charged residue pairs on ΔΨ-driven (black) and peptide driven uptake or counterflow (blue). Uptake of 15 μM ^{3}H-AlaAla by 1 μM protein was measured at pH 6.5 and 25°C after 2 min (ΔΨ-driven) and 5 min (counterflow).
(B) Effect of mutations on the equivalent residues in $PepT_{So}$. Measurements for both assays were taken after 10 min.

In contrast, the integrity of the salt bridge at the bottom of the central cavity is essential to transport. As shown in the previous chapter, the universally conserved Lys126/Lys127 (H4) plays a role similar to that of the glutamates and arginine on H1: the alanine mutation disrupts coupling without significantly hindering counterflow. The same lysine is predicted to form a salt bridge to the first glutamate on H1 in the occluded conformation of GkPOT (Doki *et al*, 2013). Lys126/127 may therefore be the link between the charged residues on H1 and those of opposite charge on the C-terminal side of the cavity,

being involved in both proton translocation and the structural changes that lead to the outward-open conformation. Its partner on H10, Glu410/Glu419, is important to both ΔΨ- and substrate-driven uptake, with alanine substitutions severely reducing ($PepT_{St}$) or abolishing transport ($PepT_{So}$). In addition to forming the salt bridge closing the intracellular gate in the outward-open state, the potential significance of the lysine and glutamate in substrate recognition and proton coupling was discussed previously (5.3.1). Both residues are therefore likely to play a complex role in the transport mechanism. For example, it is possible that the lysine provides a positively charged environment that serves to modify the pK_a of the protonated glutamate through electrostatic interactions, facilitating the release of the proton into the cytoplasm and subsequently binding to it to close the gate. Furthermore, the two side chains may accommodate the functional groups on the peptide substrates, by either hydrogen bonds or electrostatics (the C-terminal carboxylate by the lysine, the N-terminal amino group by the glutamate), as seen in the $PepT_{St}$-AlaPhe structure (fig. 5-17, A).

Lastly, the central salt bridge, formed between Arg33 and Glu300 in $PepT_{St}$, is required for ΔΨ-dependent uptake (fig. 6-2, A), suggesting a role in coupling. As in the case of the charged pair at the bottom of the cavity, a possible explanation would be that the arginine side chain modifies the pK_a of Glu300 such that it may form a salt bridge to it, thereby stabilising the binding site in the sealed conformation seen in the inward-open and occluded states. It is possible that Glu300 acts as a protonation site, as the equivalent glutamate was found to be protonated in the ligand-bound structure of GkPOT, its protonated state seemingly required for binding (Doki *et al*, 2013). Indeed, a similar role (site of the initial protonation event that increases ligand affinity at the extracellular side) has been proposed for acidic residues at similar positions in FucP (Dang *et al*, 2010) and PipT (Pedersen *et al*, 2013). Therefore, it seems likely that Glu300 is protonated before substrate

binding, and is involved in coordinating the substrate in the initial outward-facing substrate-bound state (before the shift to the occluded conformation) (fig. 6-3).

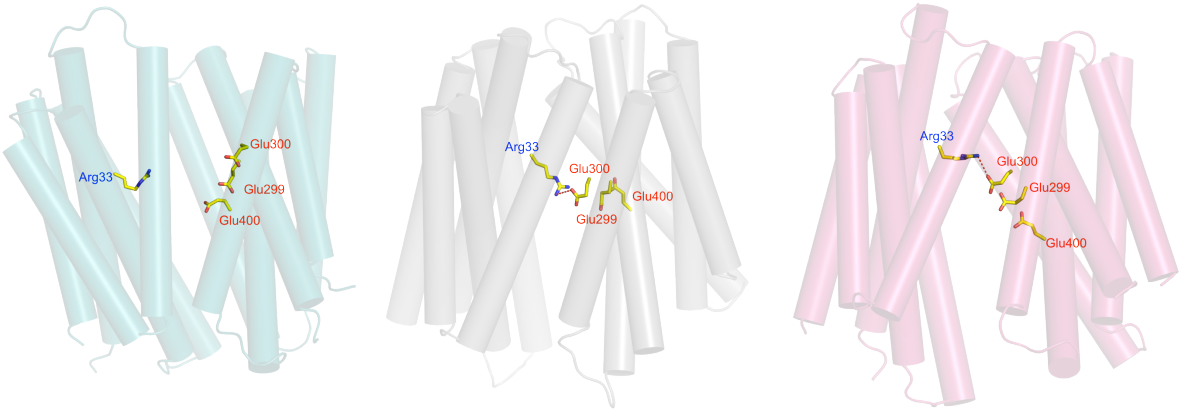


Fig. 6-3: Top views of the $PepT_{St}$ outward-open model (green), occluded model (grey) and inward-open structure (red), showing the relative positions of the central inter-domain salt bridge and the C-terminal glutamate cluster potentially involved in proton coupling.

As the protonated glutamate cannot form the salt bridge to the Arg33 in the occluded state, it needs to pass the proton to another negatively charged side chain. This reveals a new possible role for Glu299, the side chain of which is indeed in close proximity of both Glu300 and Glu400 (fig. 6-3). The cluster of three glutamates may therefore form a proton relay network (potentially along with structural water molecules and/or with the involvement of the substrate), with the Glu400 side chain the last to be protonated before releasing the proton into the cytoplasmic space.

In $PepT_{So}$, the situation appears slightly different. Similarly to Glu300, the equivalent Asp316Ala mutation in $PepT_{So}$ abolishes both types of transport (fig. 6-2, B). However, whereas the equivalent Arg33Ala mutation decouples $PepT_{St}$ from the charge gradient, the Arg32Ala substitution ($PepT_{So}$) does not seem to hinder the transporter's activity (fig. 6-2, B). How then does $PepT_{So}$ revert back to the outward-open state for a new transport cycle even without the arginine side chain in position to form the salt bridge to Asp316? A possible interpretation of that result lies in the adjacency of Asp316 to His61. As discussed in the

previous chapter (5.3.2), the histidine is believed to be the main protonation site in the mammalian oligopeptide transporters and appears to play a similar role in $PepT_{So}$. Comparably, a histidine acts as a protonation site in GlpT, where it interacts with an acidic side chain to coordinate an interhelical salt bridge (Law *et al*, 2008b). As mentioned in the introduction (1.3.5), a similar role is played by His322 in LacY, acting as a hydrogen bonding partner to a glutamate involved in a salt bridge between N- and C-terminal domains (Radestock and Forrest, 2011). In $PepT_{So}$, His61 is similarly found in close proximity to Asp316 in both the occluded structure and the inward-open model, but they are separated in the outward-open model, in which the histidine faces the open extracellular cavity (fig. 6-4).

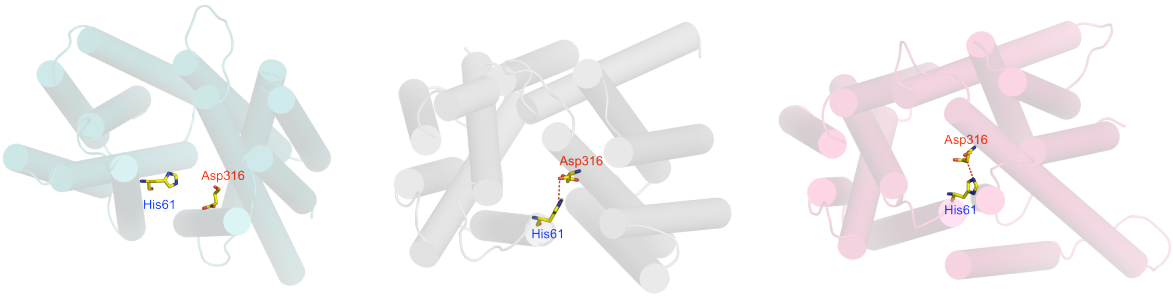


Fig. 6-4: Top views of the $PepT_{So}$ outward-open model (green), occluded structure (grey) and inward-open model (red), showing the relative positions of His61 and Asp316.

Therefore, His61 may contact Asp316 in the inward-open state, initiating the transition to the occluded state, and bridging the N- and C-terminal domains in the process. This new, flexible gate would then break easily when the transporter adopts the outward-open state (as Asp316 moves with H7 to open the cavity), but at the same time would be strong enough to stabilise the occluded and inward-open conformations, by keeping H7 packed against H1 and H2 even in the absence of the positive side chain at Arg32. Once in the outward-open conformation, His61 may bind the extracellular proton, then protonate Asp316 to allow substrate binding and the shift to the occluded state, their synchronised motion accompanying the transition back into the inward-open conformation. In keeping with this

model, mutating the arginine would not abolish transport, as the central salt bridge would no longer be required. In contrast, $PepT_{St}$ has an alanine in the place of His61, therefore making the positive side chain at Arg33 a necessity, as only it can contact Glu300.

6.1.2. Cytoplasmic salt bridges stabilise the outward-open conformation

In addition to modeling, structural aspects of the transport mechanism were also investigated through molecular dynamics simulations. Performed by our collaborator Dr Philip W. Fowler in the SBCB Unit at the University of Oxford, the MD work was focused on $PepT_{So}$ and aimed to validate the models (as well as DEER measurements discussed later in this chapter) and identify new interactions and structural changes accompanying the transport cycle. One MD experiment in particular examined the propensity of several pairs of charged residues at the extra- and intracellular side of the transporter to form salt bridges in the outward-open conformation compared to the occluded crystal structure (fig. 6-5, A). The study identified two pairs of residues at the cytoplasmic side not interacting in the structure but predicted to do so in the outward-open state (fig. 6-5, B): Asp79 and Lys84 (on opposite sides of the H2-H3 hairpin), and Asp136 and Lys439 (connecting H4 to H11).

As the repeat swapped outward-open model and the MD data agreed on these interactions, the residues were mutated to alanine, which completely abolished both ΔΨ- and peptide-driven uptake in most cases (fig. 6-5, C and D). The sole exception, the Lys84Ala substitution nevertheless resulted in 50 to 75% reduction of uptake levels. Like the Lys127-Glu419 inter-domain salt bridge discussed previously, Asp79-Lys84 and Asp136-Lys439 also

interact to stabilise the outward-open transporter, with the latter pair in particular having the crucial function of bridging H4 and H11 in this conformation.

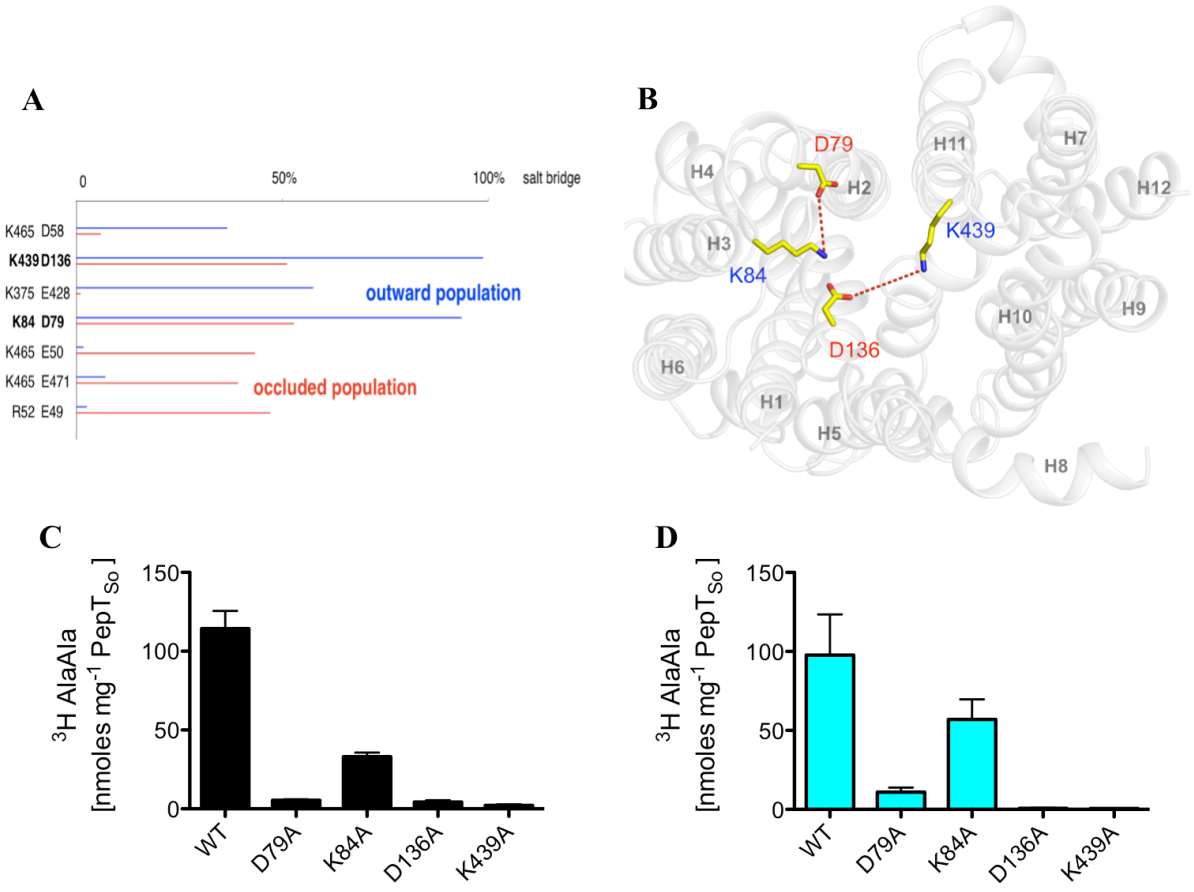


Fig. 6-5: Salt bridges at the cytoplasmic end of PepT$_{So}$ TM helices stabilise the outward-open conformation.
(A) Predicted likelihood of salt bridge formation by pairs of N- and C-terminal residues at the cytoplasmic and periplasmic side of the protein, based on comparing the occluded PepT$_{So}$ structure to MD simulations of the outward-open state (figure provided by Philip W. Fowler).
(B) Cytoplasmic view of the occluded PepT$_{So}$ structure showing two pairs of charged residues predicted to form salt bridges in the outward-open state. Helices HA and HB were ommited for clarity.
(C) Effect of alanine mutations on ΔΨ-driven dialanine uptake.
(D) Effect of alanine mutations on peptide-driven uptake (counterflow).

6.2. Structural movements in the C-terminal domain

Superimposing the $PepT_{So}$ and $PepT_{St}$ structures shows the C-terminal helix bundle to be more mobile than its N-terminal counterpart in the transition from the occluded to the inward-open states (fig. 4-6). While salt bridge formation and dissociation between the two domains explains the closing and opening of gates that allow alternating access to the central cavity, the movement of C-terminal TMs themselves (mainly in their cytoplasmic halves, as explained) is made possible by their ability to kink at conserved positions, resulting in a "swinging" sideways rotation with respect to a "hinge" region.

6.2.1. Rearrangement of H10 and H11: opening the intracellular gate

Sealed in the $PepT_{So}$ structure, the intracellular gate is opened in $PepT_{St}$ by the coordinated movement of H7 and H10/H11 away from H4/H5. More specifically, H7 straightens compared to its kinked conformation in the occluded structure, providing sufficient space for the cytoplasmic halves of H10 and H11 to rotate sideways. The bending of H10 and H11 occurs at two membrane-embedded residues: Gly407 and Trp427 (Gly426 and Trp446 in $PepT_{So}$) (fig. 6-6, A). The glycine and tryptophan side chains are situated at the same position with respect to the membrane plane and are highly conserved in the POT family. Moreover, glycines are found at equivalent positions in LacY (Abramson *et al*, 2003), GlpT (Lemieux *et*

al, 2004) and FucP (Dang *et al*, 2010), suggesting that flexibility in this region is a conserved feature of the MFS.

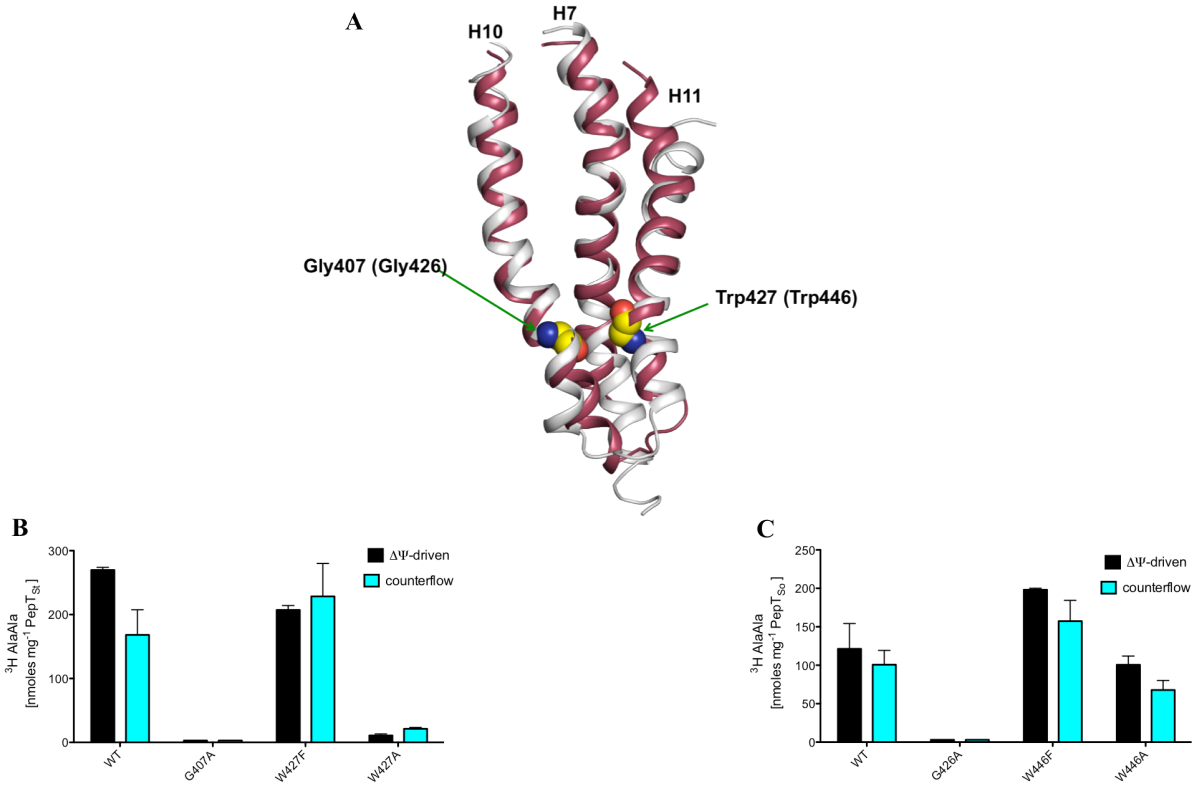


Fig. 6-6: Conserved glycine and tryptophan residues on H10 and H11 enable the opening of the intracellular gate.
(A) Superposition of H7, H10 and H11 from the crystal structures of $PepT_{So}$ (grey) and $PepT_{St}$ (red). The two residues forming the hinge region in $PepT_{St}$ are shown as spheres ($PepT_{So}$ numbers in brackets).
(B) Effect of mutating the hinge residues in $PepT_{St}$ on ΔΨ-driven uptake (black) and counterflow (blue).
(C) Effect of the equivalent mutations in $PepT_{So}$.

The importance of these residues was confirmed by mutagenesis and transport assays. In both proteins, glycine-to-alanine substitutions abolish ΔΨ-coupled uptake and counterflow, indicating a disruption of the transporter's ability to undergo the conformational changes required for transport (fig. 6-6, B and C). Alanine mutations of the tryptophan residues have comparable, albeit less drastic effects, while replacing the tryptophans with phenylalanines has no effect, as it fulfills the requirement for a bulky, hydrophobic side chain. Thus,

Glu407/426 and Trp427/446 both appear to be essential to the opening of the intracellular gate.

6.2.2. Rearrangement of H8: proline residues and coupling

The models of $PepT_{So}$ and $PepT_{St}$ in different conformational states are based on both the crystal structures and the inverted repeat symmetry typical of MFS transporters. The models confirm that the largest structural reconfigurations underlying the transition from outward to inward-open states occur in the C-terminal domain. The outward-open models in particular introduce some pronounced changes in the orientation of several helices. For instance, H10 and H11 move outward from the central cavity, while H7, found in a kinked orientation in the occluded structure, moves away from H1/H2 to allow for the opening of the extracellular gate. In addition, these movements trigger the reordering of the salt bridges that connect the N- and C-terminal bundles, as discussed. However, the most prominent rearrangement seen in the outward-open models is that of helix H8, which is featured packing tightly against H3 and H11. To achieve this, the upper and lower parts of H8 separately undergo twisting motions that result in the straightening of the helix compared to its bent conformation in the occluded state (fig. 6-7, A). Exemplified here for $PepT_{So}$, H8 reorients itself similarly in the outward-open model of $PepT_{St}$.

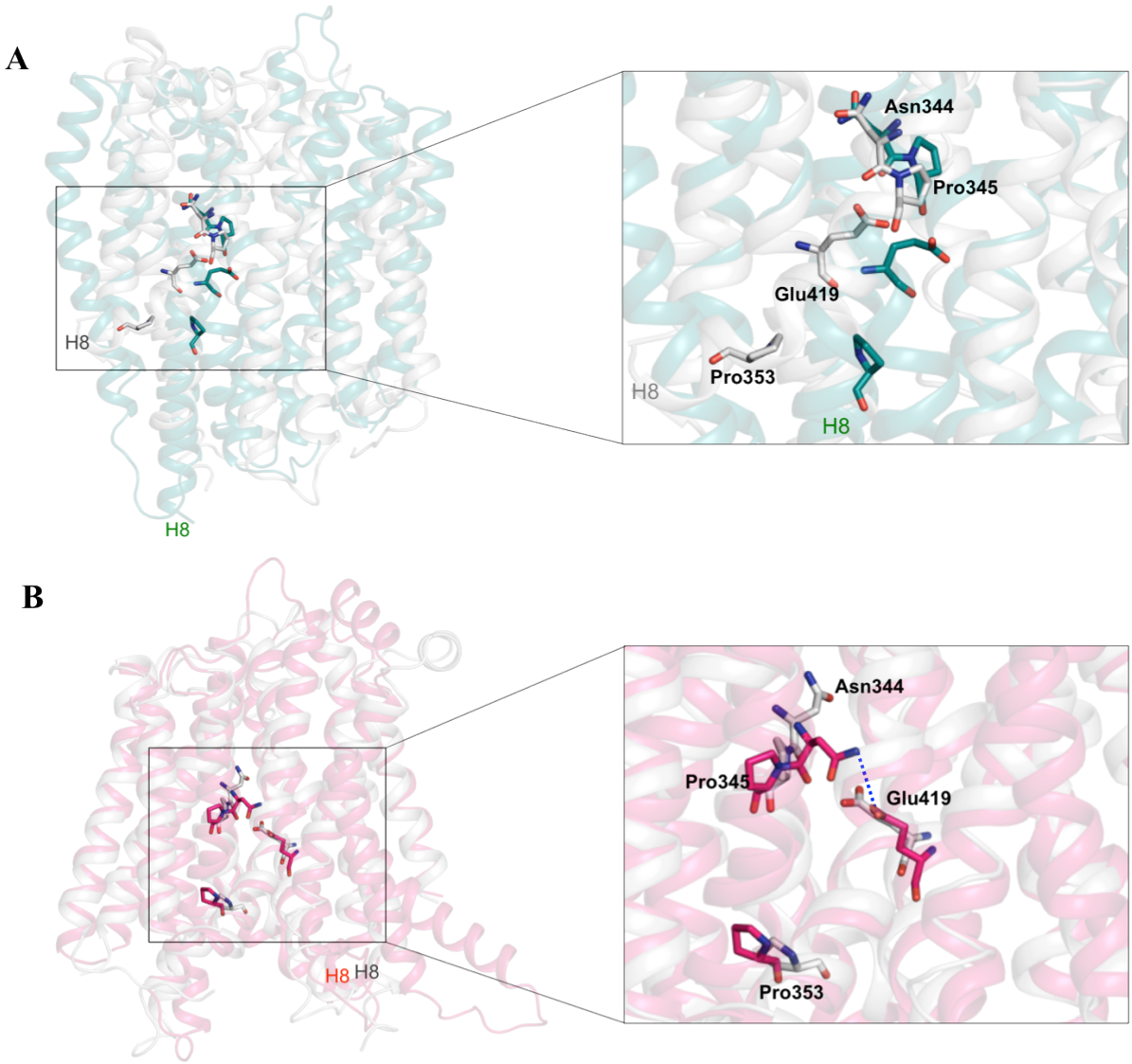


Fig. 6-7: Conserved proline and asparagine residues and the movement of H8 in $PepT_{So}$.
(A) Comparison between the occluded structure (grey) and outward-open model (green) (back view). The Pro353 side chain reorients itself to facilitate the movement of H8. Grey and green labels indicate the positions of the helix in the two states.
(B) Comparison between the occluded structure (grey) and inward-open model (pink) (front view). Reorientation of the Pro345 side chain is accompanied by that of the neighbouring Asn344, facing back into the central cavity in the inward-facing state.

Not surprisingly, these motions occur at two conserved proline residues on H8, Pro345 and Pro353 ($PepT_{So}$ numbers). When either of these is mutated to alanine, the transporter largely loses its $\Delta\Psi$-coupling capacity (fig. 6-8, A). The more drastic effect is seen for Pro353 in the lower half of the helix, which according to the model describes the wider movement of the two. The equivalent Pro329 and Pro337 mutations in $PepT_{St}$ yield the same effect, albeit with a less prominent reduction in uptake levels (fig. 6-8, B). Interestingly,

peptide-driven transport in both proteins is not hindered by any of these mutations, suggesting the movement of H8 is not essential to the cycle of conformational changes that accompany transport, but rather that it is needed to couple the transporter to the proton gradient. A lack of helical flexibility due to the alanine substitution may also affect the mobility of the neighbouring H7 and H10, both of which contain residues critical to the coupling mechanism, which is hence likely to depend on their coordinated movement during the transport cycle.

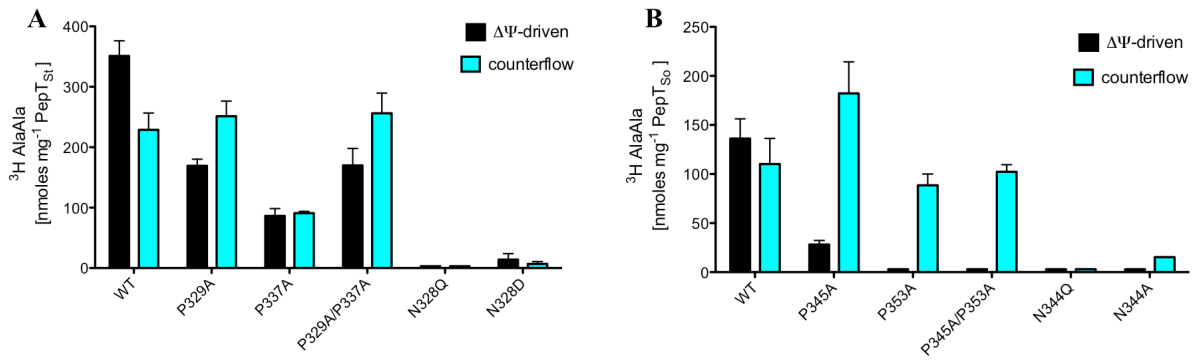


Fig. 6-8: Conserved proline and asparagine residues on H8 mediate conformational changes linked to the proton coupling mechanism.
(A) Effect of mutations on PepT$_{So}$-mediated ΔΨ-driven uptake (black) and counterflow (blue).
(B) Effect of the equivalent mutations in PepT$_{St}$.

The transition from the occluded to the inward-open state appears to involve less drastic helical movements. However, the ring in Pro345 reorients, causing the side chain of the adjacent Asn344 to face down from the extracellular cavity towards the now open intracellular gate, where it comes within hydrogen bonding distance of Glu419 (now released from the salt bridge to Lys127 seen in the outward-facing state) (fig. 6-7, B). The asparagine (Asn 328 in PepT$_{St}$) is absolutely conserved in the POT family, and is similarly positioned to contact the Glu400 side chain in the PepT$_{St}$ structure (fig. 5-1, A). A similar interaction in LacY is considered essential for substrate binding (He *et al*, 1997) and, indeed, Asn344/328 is also crucial to both ΔΨ- and ligand-dependent transport in PepT$_{So}$ and PepT$_{St}$, neither of the

variants tested (alanine, glutamine and aspartate substitutions) showing any activity (fig. 6-8). Furthermore, the Asn328Ala variant of $PepT_{St}$ failed to express (table 5-2).

Although the glutamate does not change its position considerably in any of the conformations like the asparagine does, they are both essential for transport, and as such it is likely that their interaction is key to transporter function. As this interaction is only seen in the inward-open model of $PepT_{So}$ and the inward-open structure of $PepT_{St}$, its formation must depend on the movement of the upper half of H8, and thus on the first of the two prolines. Since $PepT_{So}$ continues to function with an immobile H8 only in the absence of the membrane potential (proline-to-alanine mutations affect ΔΨ-driven uptake but not counterflow), it follows that the Asn344-Glu419 salt bridge, although dependent on the movement of H8, is critical not to peptide transport, but to coupling. The rearrangement of H8 therefore has a double function: it opens the extracellular gate in the outward-open state and it mediates coupling to the electrochemical gradient by facilitating the contact between Asn344 and Glu419 in the inward-open state.

6.3. Double electron-electron resonance (DEER) experiments

DEER is an electron paramagnetic resonance (EPR) method that measures dipolar coupling distances between electrons in a range of 1.5 to 8 nm (Reginsson and Schiemann, 2011). In order to be amenable to DEER studies, proteins require labeling with a molecule containing

an unpaired electron, which may act as a paramagnetic centre during the measurement. Typically, a nitroxide label like MTSL (methanethiosulfonate spin label) is used to label cysteine side chains, as it reacts with the –SH group to form a disulfide bond. Having two such labels in a region of the protein that is expected to undergo conformational changes makes it possible to study these changes through comparing the measured distance distributions for the labeled pair. The method is highly sensitive to distance, as the intensity of the signal depends on the cube of the distance between the two spin centres (Elsaesser *et al*, 2005).

In recent years, DEER has been used to study the conformations adopted by membrane proteins in detergent and lipid bilayer environments, in the presence of electrochemical gradients or upon ligand binding, and thereby provide insight into their mechanism (Mchaourab *et al*, 2011). The effect of ATP binding and hydrolysis on the conformational dynamics of ABC transporters (Borbat *et al*, 2007), the influence of lipid composition on liposome-reconstituted ion channels (Vamvouka *et al*, 2008) and the activity of intramembrane enzymes in detergent solution and bicelles (Upadhyay *et al*, 2008) are a few examples of aspects investigated by DEER. In secondary active transporters, DEER measurements have been used to study the alternating access mechanism. For example, Na^+ binding was found to induce the outward-open state in LeuT (Claxton *et al*, 2010), while inhibitors of Glt_{Ph} were shown to favour different conformations compared to the native substrate in a manner dependent on the membrane environment (Georgieva *et al*, 2013; Hänelt *et al*, 2013). Two comprehensive DEER studies were carried out with LacY, measuring distances at both the periplasmic and cytoplasmic surfaces of the protein to show that all three conformations of the transport cycle are detectable in detergent solution, with an inward-open bias in the apo state and a shift towards the outward-open conformation upon glucose binding (Smirnova *et al*, 2007; Madej *et al*, 2012).

6.3.1. Transport activity of $PepT_{So}$ DEER mutants

Following the successful DEER work on LacY (the first such study on an MFS transporter), a collaboration was set up for a DEER study on $PepT_{So}$, with Dr Marcella Orwick and Patricia Dijkman in Prof. Anthony Watts' group at the University of Oxford performing the experiments. After a series of measurements on selected periplasmic and cytoplasmic spin-labeled double cysteine mutants (chosen based on the equivalent positions studied in LacY), the proteins were reconstituted into liposomes and their activity assayed to confirm that the cysteines (introduced through mutagenesis) or the spin labels had not affected function. All the DEER mutants were active in the assay, with the anticipated exception of the variant carrying an additional double alanine mutation at Pro345/Pro353 (fig. 6-9) – the latter shown to be conformationally "immobile" under ΔΨ-driven conditions in the previous chapter.

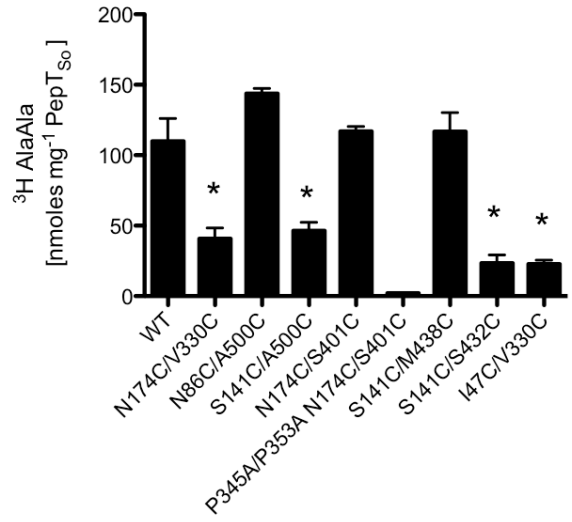


Fig. 6-9: ΔΨ-driven uptake of 15 μM ^{3}H-AlaAla at 10 min by 1 μM double cysteine mutant variants of $PepT_{So}$ used for DEER experiments. Proteins that dimerised during purification are marked with asterisks.

During the buffer exchange before reconstitution, several mutants eluted as a mixture of ~75% dimer and 25% monomer, likely due to either the forming of disulfide bonds between unlabeled surface cysteines on adjacent molecules, or unspecific interactions induced by the label. The dimer could be only partially removed by reducing with 5 mM DTT, and unsurprisingly, the respective proteins (marked with asterisks) were active to only ~25% of wild type levels in the assay (fig. 6-9). However, it is unlikely that dimerisation would have affected the DEER distance distributions, as the resulting additional distances inside the dimer would have been below the detectable range for DEER.

6.3.2. The effect of peptides and the electrochemical gradient on DEER distance distributions

After measuring distance distributions in DDM solution, several mutant $PepT_{So}$ samples were reconstituted into liposomes in order to study conformational dynamics in a lipid bilayer, and test the effect of the electrochemical gradient across the liposome membrane and the presence of a peptide substrate on the adopted conformations.

Proteins were expressed as GFP-fusion constructs, purified and reconstituted as described in the methods chapter. Before the DEER measurements, ΔΨ-driven and counterflow uptake of ^{3}H-AlaAla were assayed in highly concentrated proteoliposome samples using the same substrate concentrations as in the regular transport assays but very low sample volumes. The purpose of these assays was to replicate the generation of the membrane potential (ΔΨ, via valinomycin) as well as test the effect of the peptide gradient

(counterflow) in "mock DEER samples" (i.e. high protein concentration). The data shows that wild type $PepT_{So}$ and variants transport the labelled substrate to similar extents in both assays, with the exception of the cytoplasmic mutant S141C/S432C which only shows limited counterflow activity (fig. 6-10).

Despite the sample volume not permitting to generate an optimal (–100mV) membrane potential, ΔΨ-driven uptake is detectable and depends on the presence of valinomycin. The proteins also perform counterflow in the typical conditions (10 mM peptide inside the liposomes), and both graphs indicate that proteins are performing active transport 10 min after time zero. Therefore, reconstituted proteins performing uptake under these conditions that are prepared for DEER measurements at the 10 min time point should yield distance distributions characteristic of the structural dynamics involved in both ΔΨ-driven transport and counterflow.

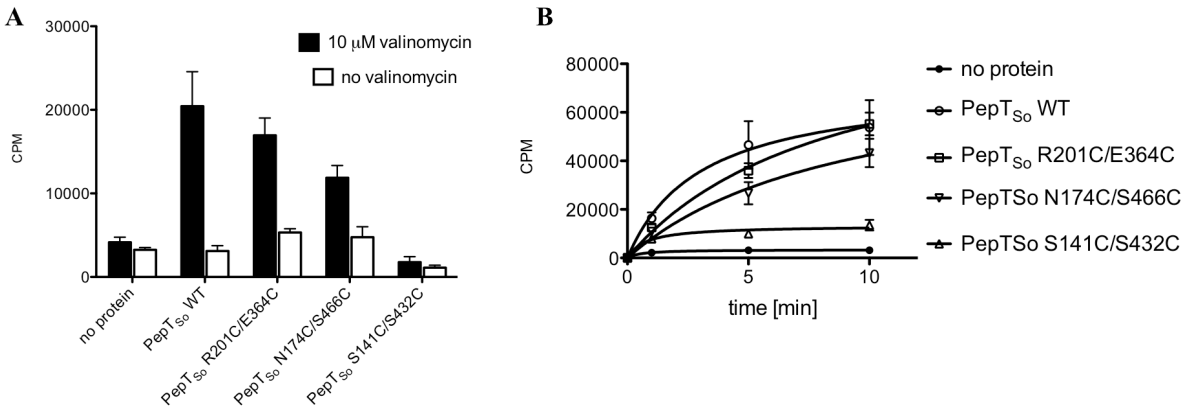


Fig. 6-10: The effect of ΔΨ and peptide on transport in concentrated DEER samples of $PepT_{So}$.
(A) ΔΨ-driven assay. 5 µl proteoliposomes containing internal buffer (20 mM KPi, 100 mM KAc and 2 mM $MgSO_4$) were resuspended in 10 µl external buffer (120 mM NaPIPES and 2 mM $MgSO_4$) containing 15 µM ^{3}H-AlaAla in the presence and absence of valinomycin, to a final protein concentration of 250 µM. Uptake was measured at pH 6.5 and 30°C after 10 min. The Y-axis shows uptake of the ^{3}H label in counts per minute (CPM) of β radiation emission.
(B) Counterflow assay. 5 µl proteoliposomes containing 50 mM KPi (pH 6.5), 2 mM $MgSO_4$ and 10 mM AlaAla were resuspended in 10 µl of 50 mM KPi (pH 6.5) and 2 mM $MgSO_4$ containing 5 µM ^{3}H-AlaAla, to a final protein concentration of 250 µM. Peptide-driven uptake was measured over time at 30°C.

Under this assumption, liposomes containing $PepT_{So}$ R201C/E364C (cytoplasmic pair) were assayed for uptake in four conditions (ΔΨ-driven with and without valinomycin and counterflow with and without peptide) using non-radioactive peptide as the substrate, then prepared for the DEER experiments within 10 min from time zero (as described in the methods chapter), flash-frozen and measured. In addition, $PepT_{So}$ S174C/S446C (periplasmic) and S141C/S432C (cytoplasmic) were only measured in the counterflow setup (with and without peptide). The DEER distributions for the three $PepT_{So}$ variants in these conditions are presented below (fig. 6-11).

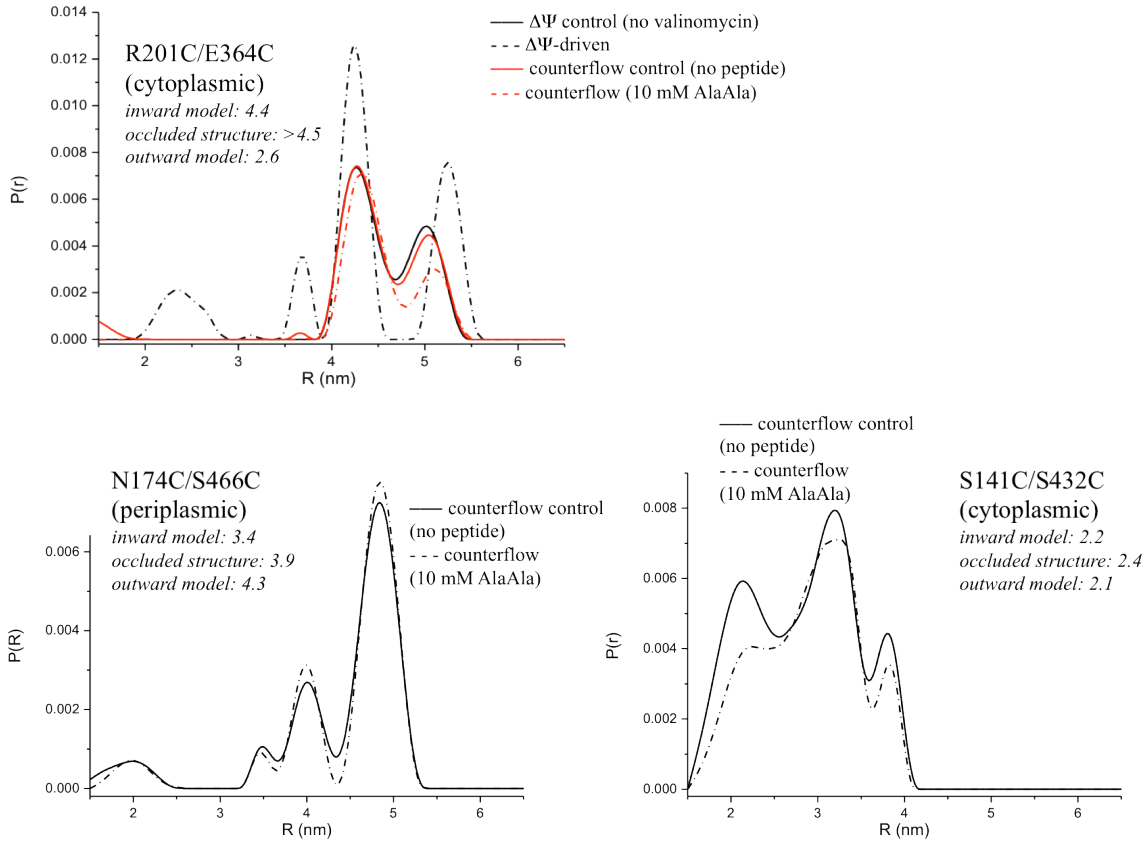


Fig. 6-11: DEER distance distributions for three liposome-reconstituted $PepT_{So}$ spin-labelled variants, in the presence and absence of peptide and membrane potential. For comparison, the equivalent distances in the crystal structures/models of the inward-open, occluded and outward-open states are listed in italics.

As seen above, the presence or absence or peptide in the liposome lumen does not affect the distance distributions of the counterflow samples. The S141/S466 profile (fig. 6-11, C) is difficult to interpret due to poor resolution of the peaks; however, the three peaks correspond to short distances (below 4 nm), consistent with those seen in the crystal structure and models. The three periplasmic distances (N174/S446) (fig. 6-11, B) roughly match the distances in the models, suggesting that the protein adopts all three conformational states in the liposome membrane, regardless of the presence of peptide. Interestingly, the R201/E364 pair only shows two peaks, matching the inward-open and occluded states but no peak for the outward-open state (fig. 6-11, A, red). The most interesting observation however is that the R201/E364 profile changes when the membrane potential ($\Delta\Psi$) is generated upon the addition of valinomycin (fig. 6-11, A, black), with the greater distance (potentially indicating an occluded state) increasing from 5 to 5.2 nm, and two new distances registering at 2.4 and 3.7 nm (corresponding to a shift towards the outward-open conformation).

While there is not enough data to make a definite statement at this point, the distance distributions above indicate that $PepT_{So}$ adopts multiple conformations in a liposome membrane, and that these are not affected by the presence of a peptide gradient, but by the presence of a transmembrane gradient. The data on R201C/E364C in particular seems to suggest that a protonation event, rather than a peptide binding event, may be the initial trigger for the conformational changes associated with transport.

6.4. Conclusions

When combined with crystallographic data on POT proteins, mutagenesis and *in vitro* transport assays are powerful tools for identifying individual amino acids involved in substrate recognition and proton coupling. However, in order to study the structural aspects of a complex process like the peptide transport cycle, functional characterisation needs to be combined with data from other biophysical and computational methods, such as spectroscopy, simulations and modelling. This approach was used with $PepT_{So}$ and $PepT_{St}$ to generate models of alternate conformational states and identify the structural elements that trigger them, with one method often being used to confirm the data obtained through another in a combined effort to create a model for peptide transport.

Homology and inverted repeat topology modelling provided possible snapshots of the three main conformations required for alternating access. Based on these models and the crystal structures, key conserved salt bridges between the N- and C-terminal domains were identified and confirmed to have either essential or supportive roles in inducing the outward-open, inward-open and occluded conformations. Links were proposed between the formation of these salt bridges and the proton coupling and peptide recognition mechanisms within the binding site, with essential differences between the two proteins pointing towards a different proton coupling mechanism in $PepT_{So}$, centred around the role of His61 and its interaction with Asp316, a mechanism possibly shared by other POTs like PepT1 and PepT2, where the equivalent histidine is present and thought to have a similar function. Additional salt bridges matching the inverted repeat swap model were identified through molecular dynamics simulations of the outward-open state of $PepT_{So}$, and were confirmed in transport assays.

Helical rearrangements in the C-terminal helix bundle were also found to play a crucial role in opening the intra- and extracellular gates. Highly conserved residues were identified as potential key players in these rearrangements, based on superimposing the available crystal structures and comparing them with the alternate state models. Transport assays then established a structural role for the glycine and tryptophan residues in the "hinge" region of H10 and H11 in opening the intracellular gate, and a more complex role for the proline residues on H8 (involved in opening the extracellular gate), with an adjacent asparagine making a critical link with the binding site and the proton coupling mechanism.

Finally, EPR spectroscopy in the form of DEER measurements was used to create distance distribution profiles characteristic of the conformational states adopted by $PepT_{So}$ in detergent and lipid bilayer environments. The data was compared and contrasted with the results of the simulations and modelling, and transport assays were once again used to optimise conditions for testing the effect of proton and peptide gradients (the main driving forces of peptide transport) on the sampled conformations. Still a work in progress, the project aims to establish a link between the biochemical, biophysical and structural aspects of the peptide transport mechanism.

Chapter 7: Prokaryotic POT proteins and drug transport

The pharmacological importance of PepT1 and PepT2 has been discussed in literature reviews (Daniel and Kottra, 2004; Brandsch, 2009) and the introduction to this thesis (1.4.4). A range of β-lactam antibiotics, antiviral compounds, peptide-based prodrugs and others have been established as either substrates or inhibitors of the mammalian peptide transporters, by assembling the data from direct uptake and competition experiments as well as electrophysiology in cells containing either native or overexpressed protein.

Despite high sequence conservation making the bacterial POTs potential targets for studying drug uptake *in vitro*, data so far has been scarce. Before the publication of the first POT crystal structures, the only indication of drug transport by bacterial POTs consisted of competition experiments in whole cells overexpressing *E. coli* transporters YdgR (DtpA) and YhiP (DtpB). In the first of these, YdgR-mediated uptake of β-Ala-Lys-AMCA (fluorescent peptide) was inhibited by ~70-90% by a large excess (200-fold) of cephalosporin β-lactams cefadroxil, cefalexin and cephradine (Weitz *et al*, 2007). The same study found no proof of competition by ampicillin or amoxicillin, but ACE inhibitors enalapril and captopril were identified as possible competitors, causing between 20 and 30% reduction in uptake levels. The second study used peptide analogue GlySar as a reporter (Harder *et al*, 2008). This time using a mere 10-fold excess of competitor over reporter substrate, the competitive effect of

the three cephalosporins was replicated for YdgR, whereas uptake by YhiP was affected to a much lesser extent.

7.1. Drug transport by YdgR, $PepT_{St}$ and $PepT_{So}$

The whole cell assays showing cephalosporin competition for YdgR were performed using non-physiological, low affinity substrates as reporters of uptake. As seen in a previous chapter (5-1), altering the expression patterns of endogenous peptide transporters in *E. coli* cells can cause significant changes in background uptake levels. After the assays on $PepT_{So}$ and $PepT_{St}$ were performed, a similar study was undertaken with YdgR and YhiP, with the purpose of replicating the results of the competition experiment in the previous paper as well as testing direct uptake of radiolabeled drugs.

7.1.1. Whole cell experiments

The competition assay was carried out using *E. coli* BL21(DE3) cells transformed with His-tagged constructs of YdgR and YhiP in pBAD plasmids, as described in the methods chapter. A selection of confirmed PepT1 substrates (fig. 7-1) was tested for competition against ^{3}H-GlySar.

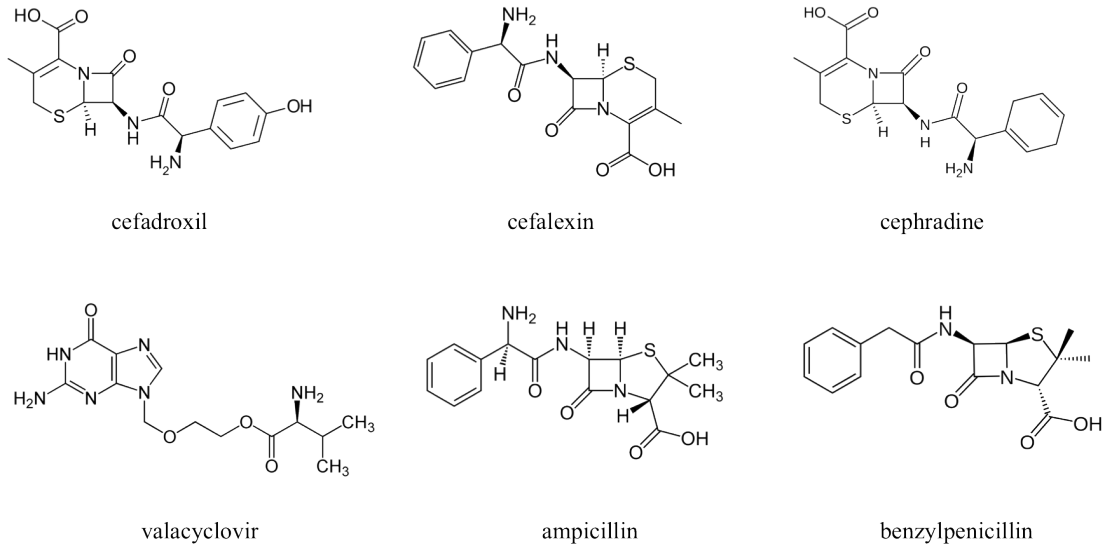


Fig. 7-1: Chemical structures of the drug compounds used in the competition assays.

As seen for $PepT_{St}$ and $PepT_{So}$ previously, uptake of the reporter is proton-coupled, as it is abolished by protonophore CCCP (fig. 7-2). The results are similar to the previous, with YdgR-mediated uptake being reduced by ~60-70% on average in the presence of cefadroxil, cefalexin and cephradine, as well as previously untested antiviral compound valacyclovir (fig. 7-2, A). In contrast, the drugs' effect on YhiP function is limited (a maximum of 25% reduction in the case of cefadroxil) (fig. 7-2, B). As in the previous study (Harder *et al*, 2008), ampicillin has no discernible effect, while benzylpenicillin, an anionic β-lactam, seems to act as a weak competitor in the case of both proteins.

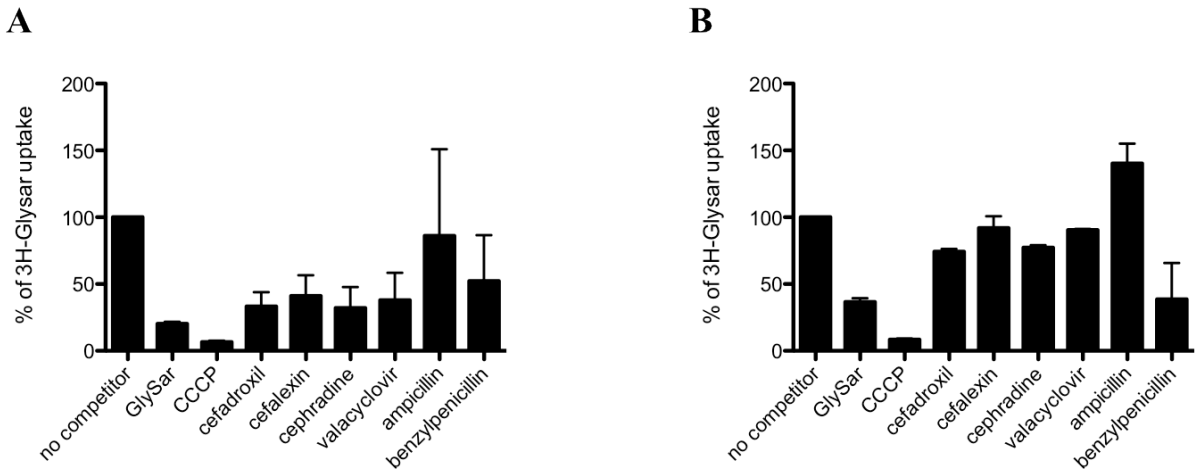


Fig. 7-2: Drug competition assay in whole cells with YdgR (A) and YhiP (B). $5 \cdot 10^9$ *E. coli* BL21(DE3) cells overexpressing the transporters as His-tagged constructs were harvested 20h after induction and resuspended in transport buffer (25 mM Tris pH 7.5, 150 mM NaCl and 5 mM glucose). Uptake of 500 µM ^{3}H-GlySar was measured after 10 min at 37°C in the presence of 5 mM competitor unless specified. Data shown as percentages. CCCP refers to the proton ionophore carbonyl cyanide *m*-chlorophenyl hydrazine. The experiments were performed by NS and Daphne Amevenu.

Having replicated the YdgR competition results, direct uptake of ^{3}H-labeled drugs was tested next. In agreement with the competition data, cells overexpressing YdgR transport ^{3}H-cefadroxil in a time-dependent manner, whereas YhiP overexpression results in uptake levels similar to the control cells (fig. 7-3). Like the competition assay, the time course uptake experiment with cefadroxil was performed in collaboration with Daphne Amevenu. This is the first indication of direct uptake of a drug substrate by a bacterial member of the POTs, providing further confirmation that they are valid working models for studying drug transport.

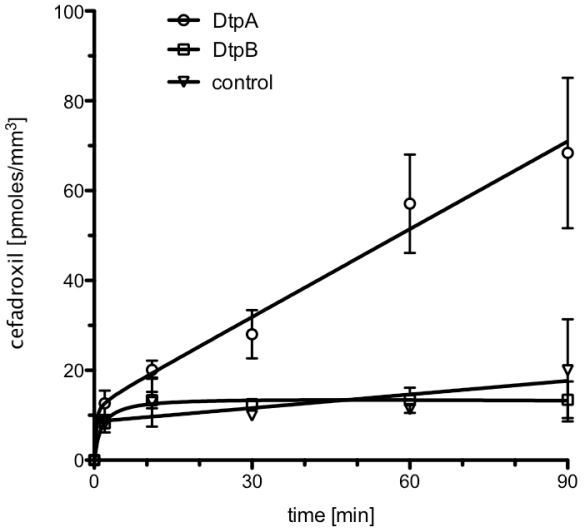


Fig. 7-3: Time-dependent uptake of ^{3}H-cefadroxil in *E. coli* BL1(DE3) cells expressing His-tagged constructs of YdgR and YhiP. $5 \cdot 10^{9}$ cells were harvested 20h post-induction and resuspended in transport buffer at 37°C. 2 μM ^{3}H-cefadroxil was used per assay. Cells overexpressing empty pBAD plasmid were used as control. The experiment was performed by NS and Daphne Amevenu.

Interestingly however, neither the direct uptake nor the competition data could be replicated for $PepT_{So}$ or $PepT_{St}$. Moreover, uptake of ^{3}H-labeled valacyclovir by either YdgR or YhiP could not be measured in similar conditions, despite the competition assay indicating it as a potential substrate. Consequently, as whole cell assay data using overexpressed native protein are difficult to interpret, YdgR was purified from a GFP-fusion construct and reconstituted into liposomes to examine the effect of the drugs in an unbiased, *in vitro* system.

7.1.2. *In vitro* experiments

As direct uptake of ^{3}H-cefadroxil or ^{3}H-valacyclovir was not detected *in vitro*, a competition experiment was carried out instead. Competition was measured as described in the methods chapter, with ^{3}H-AlaAla as the reporter substrate. Surprisingly, the result is markedly different from that of the whole cell assay, with weak effects on uptake levels seen for all the previously tested drugs, as well as for two other PepT1 cephalosporin substrates, cefaclor and ceftibuten (fig. 7-4, A). However, cefadroxil causes the greatest reduction in AlaAla uptake, by ~30-35%, making it the strongest competitor, as in the whole cell experiment. Despite the greatly increased excess of competitor over reporter (133-fold compared to 10-fold in the whole cell assay), the limited effect could be due to the relatively lower affinity of the transporter for GlySar compared to AlaAla: the apparent K_m of transport for the former in whole cells is 1 mM (Harder *et al*, 2008), whereas AlaAla is transported with an affinity below 100 μM by homologue $PepT_{St}$ (as discussed in 5.2.6), an affinity in the same range being likely for YdgR as well, based on a time course experiment showing direct uptake of AlaAla at a rate similar to $PepT_{St}$ (fig. S4). This suggests that drugs may be indeed transported, but with low affinity compared to physiological substrates like dipeptides.

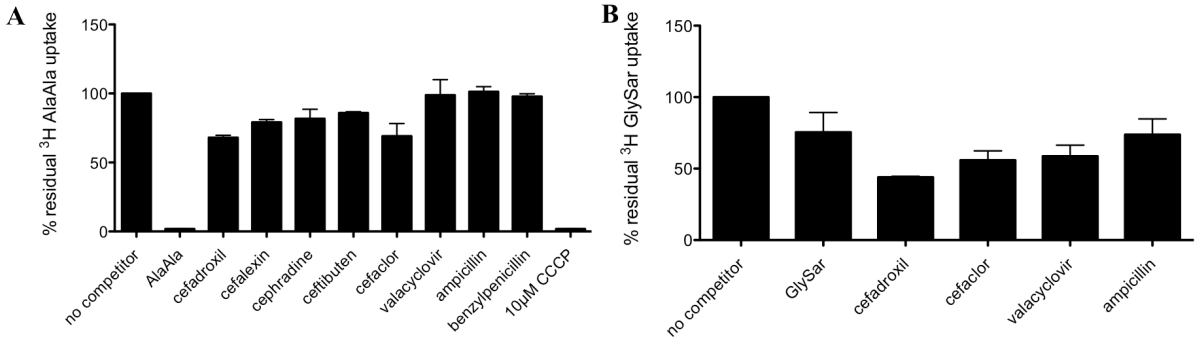


Fig. 7-4: Drug competition assays in liposomes reconstituted with YdgR.
(A) ΔΨ-driven uptake of 15 μM ^{3}H-AlaAla by 1 μM protein after 2 min incubation was measured at pH 6.5 and 25°C, in the presence of 2 mM competitor unless specified. Data shown as percentages.
(B) Uptake of 50 μM ^{3}H-GlySar in the same conditions.

To test this, the competition experiment was repeated in liposomes using GlySar as the reporter instead of AlaAla (fig. 7-4, B). As expected, the competitive effect of the drugs is stronger, with cefadroxil, cefaclor and valacyclovir all reducing uptake by ~50-60%. Disconcertingly however, ampicillin itself appears to compete against ^{3}H-GlySar uptake, to a higher extent than GlySar itself. This result is difficult to interpret, as the 40:1 excess of unlabelled over labelled GlySar was expected to result in a stronger inhibition than the comparably lower 10:1 excess used in the whole cell assay (fig. 7-2); instead, GlySar inhibits its own transport by ~80% in whole cells, but only ~25% in liposomes. This points towards a decrease in GlySar affinity *in vitro* when compared to the whole cell measurements, possibly due to additional GlySar transport by endogenously expressed transporters in the *E. coli* cells. The latter might also be the cause for the easily detected uptake of cefadroxil in whole cells despite the liposome assay suggesting that the protein's affinity for cefadroxil is very low. An additional experiment showed, however, that GlySar does inhibit its own uptake in liposomes if used in sufficient excess (fig. S5), so GlySar affinity is still present after reconstitution.

To establish whether the competitive effect was indicative of *de facto* transport, counterflow assays were performed in liposomes loaded with known transported substrate AlaAla (showing 100% competitive inhibition in the assay), POT reporter peptide alafosfalin (inhibitor of AlaAla uptake by 80%), and cefadroxil itself (weak competitor in the assay, direct uptake not detectable). As both AlaAla and alafosfalin are YdgR substrates, any counterflow activity in their case should be possible to compare to that seen for the drug, thereby establishing the latter as a substrate. Surprinsingly, no counterflow activity was detected for either cefadroxil or alafosfalin, but only for AlaAla (fig. 7-5). The background level readings for alafosfalin suggest the affinity for the molecule is very low, which would also explain the null readings for cefadroxil (affinity for the latter being even lower according to the competition data). The data seems to indicate that the relatively strong competition by

alafosfalin on YdgR may be in part due to an inhibitory effect, as the transport affinity itself appears to be much lower than suggested by the competition assay. At any rate, the data shows that cefadroxil is likely to act as either a low affinity substrate or a weak inhibitor of YdgR function *in vitro*.

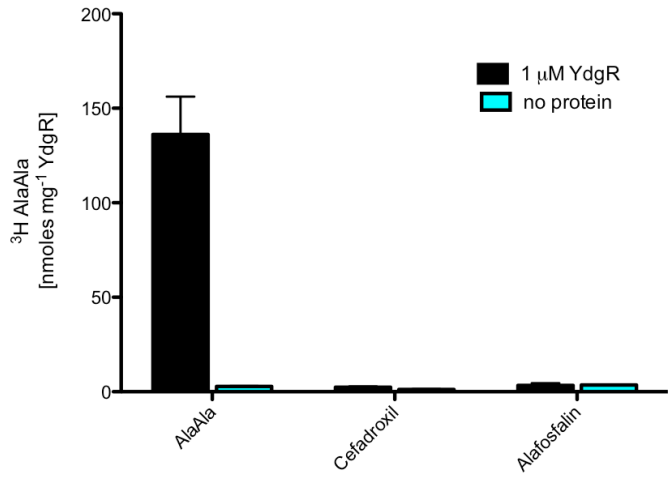


Fig. 7-5: Counterflow of three molecules by liposome-reconstituted YdgR. Proteoliposomes were loaded with 10 mM unlabelled peptide/drug, and substrate-driven uptake of 5 μM ^{3}H-AlaAla was measured after 5 min at pH 6.5 and 25°C.

Finally, to investigate how the drugs affect uptake by different POTs *in vitro*, $PepT_{St}$ and $PepT_{So}$ were used in similar competition experiments. Despite not transporting the radiolabelled drugs *in vivo*, $PepT_{St}$ shows an *in vitro* competition profile similar to YdgR, with cefadroxil inhibiting AlaAla uptake by ~25-30% and yet more pronounced effects from cefalexin, cephradine and benzylpenicillin (fig. 7-6, A). Interestingly, when switching to GlySar as a reporter, the drugs show very little competition (fig. 7-6, B). In contrast, $PepT_{So}$ is not inhibited by cefadroxil but shows increased affinity for valacyclovir (fig. 7-6, C).

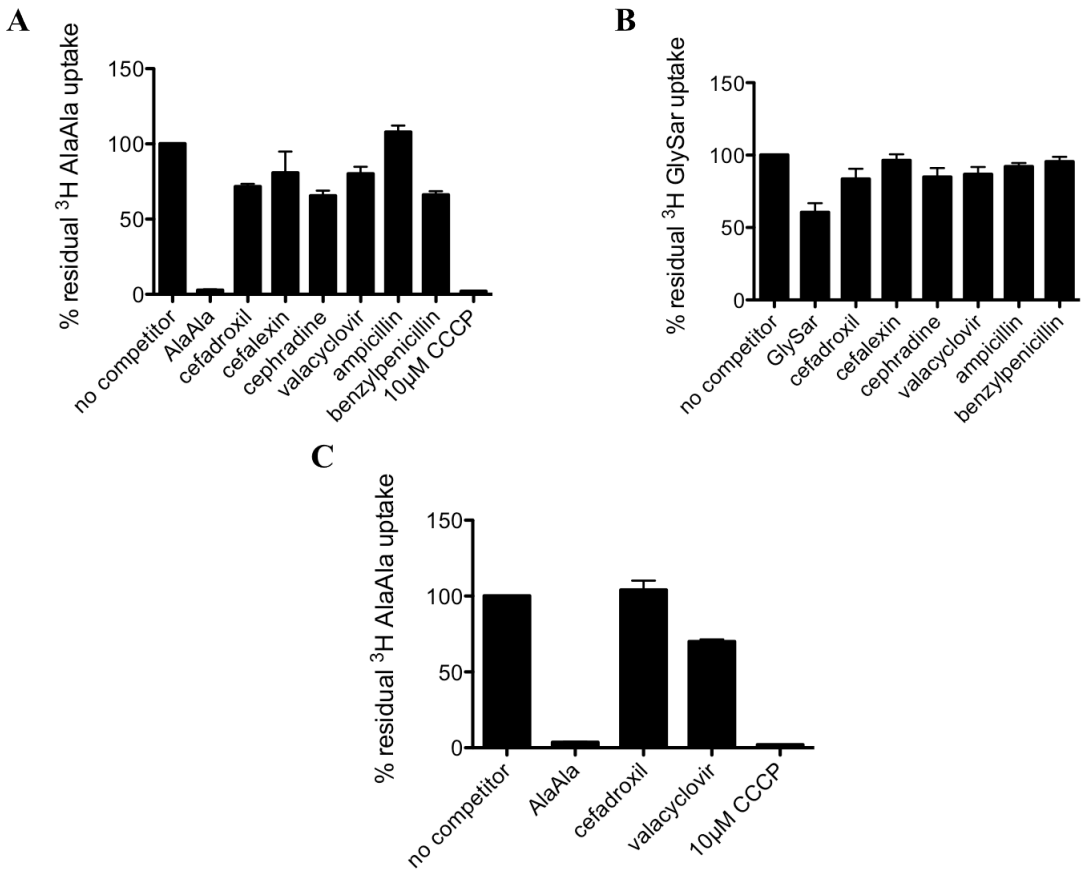


Fig. 7-6: Drug competition assays in $PepT_{St}$ and $PepT_{So}$ proteoliposomes.
(A) Competition against ΔΨ-driven uptake of 15 µM ^{3}H-AlaAla by 1 µM $PepT_{St}$ after 2 min (data as percentages).
(B) Competition using 50 µM ^{3}H-GlySar as a reporter in the same conditions.
(C) Competition against ΔΨ-driven uptake of 15 µM ^{3}H-AlaAla by $PepT_{So}$ in the same conditions. Measurement taken at the 10 min time point.

As high affinity peptide substrates cannot be used as reporters in whole cells, it is currently difficult to establish a definitive correlation between the *in vivo* and *in vitro* data. However, although direct drug uptake could not be measured *in vitro* for any of the proteins, the liposomes competition assay clearly shows that cephalosporin antibiotics compete against uptake of high affinity substrate AlaAla via both YdgR and $PepT_{St}$, as does valacyclovir in the case of $PepT_{So}$, making it likely that they are transported or at least bind to the proteins in a specific manner *in vitro*, albeit with low affinity. In fact, this notion agrees with previously determined affinity constants for cefadroxil by PepT1, like the K_m^{app} of 1.1 mM in oocytes

(Boll *et al*, 1994) or the K_i of 3.4 mM in SK-ChA-1 cells measured relative to GlySar (Neumann and Brandsch, 2003). Similarly, PepT1 displays low affinity (in the low mM range) for all cephalosporin substrates such as cefalexin, ceftibuten and cefaclor, as well as for valacyclovir (Brandsch, 2009). A great contrast to these values is seen in the literature for PepT2, which displays ~1000-fold greater affinities (compared to PepT1) for cefalexin (Ganapathy *et al*, 1995), cephradine (Terada *et al*, 1997) and cefadroxil (Knütter *et al*, 2007). There is therefore a significant difference in drug affinities for the two human transporters, of which PepT1 may resemble the prokaryotic homologues more closely. While three different bacterial POTs have been tested for drug uptake to date, with data indicating low affinity competition as well as confirmed direct uptake in the case of YdgR, finding a good bacterial model for PepT2-mediated drug transport is a pursuable goal.

7.2. Conclusions

The medical relevance of the mammalian oligopeptide transporters is primarily linked to their ability to transport a wide variety of drug compounds. Extensive studies on both PepT1 and PepT2 have identified several substrates, of which many are β-lactam antibiotics, zwitterionic ones like the cephalosporins being generally preferred to penicillins (Wenzel *et al*, 1996). However, different mechanisms have been proposed for the transport of cephalosporin antibiotics via PepT1 and PepT2, with the former preferring substrates that match the stereochemistry of tripeptides (and presumably transporting them by the same mechanism) (Bailey *et al*, 2000), and the latter recognising those with an α-amino group (Luckner and

Brandsch, 2005). Different recognition and transport mechanisms are therefore believed to account for the significantly increased affinity for cephalosporins displayed by PepT2, which serves as the main re-entry point for these drugs after the initial filtering process in the kidneys, improving their overall retention (Takahashi *et al*, 1998).

The drug competition data on liposome-reconstituted YdgR and $PepT_{St}$ in particular indicates an affinity profile similar to that of PepT1, with limited but discernible competitive inhibition by all cephalosporin antibiotics tested. The choice of reporter substrate seems significant when evaluating the competitive effect, as high affinity reporters may mask the effect, while using low affinity reporters can produce unreliable data. Additionally, although $PepT_{St}$ has a similar profile to that of YdgR *in vitro*, it does not transport radioactively labelled cefadroxil in whole cells, whereas YdgR, the native *E. coli* transporter, does. In general, the YdgR data suggests higher affinities *in vivo* for both the drugs and GlySar when compared to the reconstituted protein. This points towards the potential side effects of overexpressing an endogenous POT in cells that constitutively express three other homologues, and warns of the possibility of over-interpreting whole cell assay results.

No POT family transporters have yet been co-crystallised with a drug substrate. However, recent structural data from $PepT_{St}$ with di- and tripeptides (discussed in chapter 5.4) identifies some of the specifics of substrate recognition within the binding site that may allow us to distinguish between high and low affinity pathways of transport (Lyons *et al*, in preparation). Of these, the low affinity path might be shared by tripeptides and large drug molecules like cephalosporins in the "low affinity, high capacity" transporters like PepT1. If, as the data suggests, the bacterial POTs are more similar to PepT1 with respect to drug affinity, then the next step in explaining drug recognition and transport in the POT family will be identifying a PepT2-like bacterial member for drug-focused structural and functional studies.

Chapter 8: Functional studies of other POT family homologues

Although this research project was focused mainly on studying the transport mechanism in $PepT_{St}$ and $PepT_{So}$, additional experiments were carried out with other bacterial POT homologues. Firstly, *E. coli* transporter YdgR was used as a model for drug transport, as discussed in chapter 7, and that study prompted further inquiries into the transport profile of wild type YdgR to establish correlations between the observed drug uptake and its general substrate selectivity in comparison with its homologues. Secondly, the recently solved structure of homologue GkPOT from *G. kaustophilus* formed the basis of a collaborative project with Dr. Ryuichiro Ishitani and Prof. Osamu Nureki at the University of Tokyo focused on characterising the transport mechanism in GkPOT (Doki *et al*, 2013). The data on YdgR and GkPOT are presented in this chapter.

8.1. Substrate specificity in YdgR

In the wake of the intriguing results of the drug competition experiments, a small library of oligopeptides was tested for competition against ^{3}H-AlaAla uptake by liposome-reconstituted

YdgR (fig. 8-1). As with $PepT_{St}$, transport is proton-coupled (inhibited by protonophore CCCP), and no competition is observed with either tetra- and pentaalanine, the single amino acid alanine, or the D-isomer of the dipeptide. These results are also consistent with a previous report from whole cell assays (Harder *et al*, 2008). Instead, L-isomers of di- and tripeptides have competitive effects, with a handful of notable exceptions such as triglycine and the lysine di/tripeptides, all of them matching the pattern seen for $PepT_{St}$ and $PepT_{So}$ (fig. 5-9). Other similarities include the general preference for dipeptides with hydrophobic side chains.

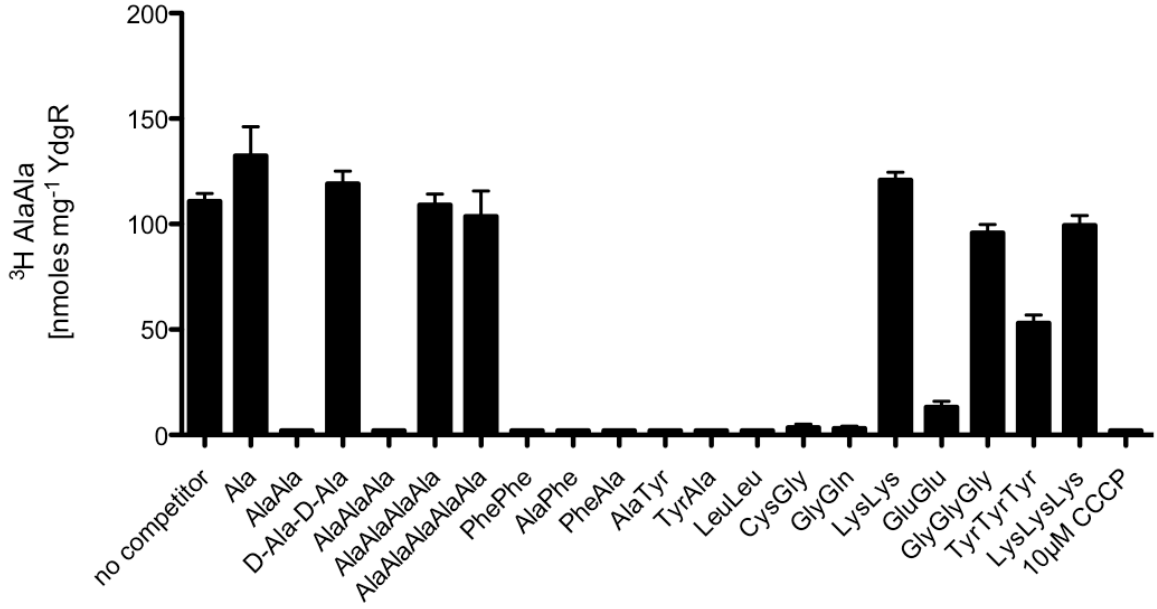


Fig. 8-1: Peptide competition assays in liposomes reconstituted with YdgR. ΔΨ-driven uptake of 15 μM ^{3}H-AlaAla by 1 μM protein after 2 min incubation was measured at pH 6.5 and 25°C, in the presence of 2 mM competitor unless specified.

The strong competitive effect of dipeptides CysGly and GlyGln (abolishing AlaAla uptake) as well as the less prominent effect of GluGlu (~80% reduction) resemble $PepT_{St}$ more closely than $PepT_{So}$, which may indicate a conserved mechanism for dipeptide transport in $PepT_{St}$ and YdgR. In contrast, TyrTyrTyr (a substrate of $PepT_{So}$ but not $PepT_{St}$, as shown previously) inhibits YdgR uptake by ~50%. Together with a seemingly high affinity for

AlaAlaAla (which completely outcompetes the reporter dipeptide), these results identify YdgR as a potentially better tripeptide transporter, more similar in this respect to $PepT_{So}$ than $PepT_{St}$. Indeed, of the residues involved in the "low affinity" tripeptide binding configuration of the $PepT_{St}$ central cavity seen in the co-crystal structure with AlaAlaAla (fig. 5-17, B), the equivalent of Glu299 is a tyrosine in YdgR – resembling the phenylalanine residue found at this position in $PepT_{So}$. It seems possible then that YdgR transports tripeptides using a "$PepT_{So}$-like" mechanism, whereas dipeptides follow a route more similar to that of $PepT_{St}$, although currently the available data is not sufficient to validate this hypothesis, or establish a connection between these observations and drug transport.

8.2. GkPOT: deciphering the mechanism

Recently, the structure of bacterial oligopeptide transporter GkPOT was solved from lipidic cubic phase crystal data to unprecedentedly high resolution for POT family members. GkPOT is an interesting target for structural and functional studies, as its origin is a thermophilic bacterium (*Geobacillus kaustophilus*) and the protein itself is unusually stable. More specifically, the structure of the apo protein and a co-complex structure with a sulfate ion in the binding site were solved to 1.9 and 2Å respectively, followed by an additional structure (of the Glu310Gln) with an alafosfalin molecule in the binding site. All three structures feature the central cavity in an inward-open conformation (Doki *et al*, 2013). GkPOT shares 48% sequence identity with $PepT_{So}$ and 69% with $PepT_{St}$, as well as 52% with $PepT_{So2}$, another homologue that was recently co-crystallised with alafosfalin bound in the central cavity, also in the inward-open state (Guettou *et al*, 2013).

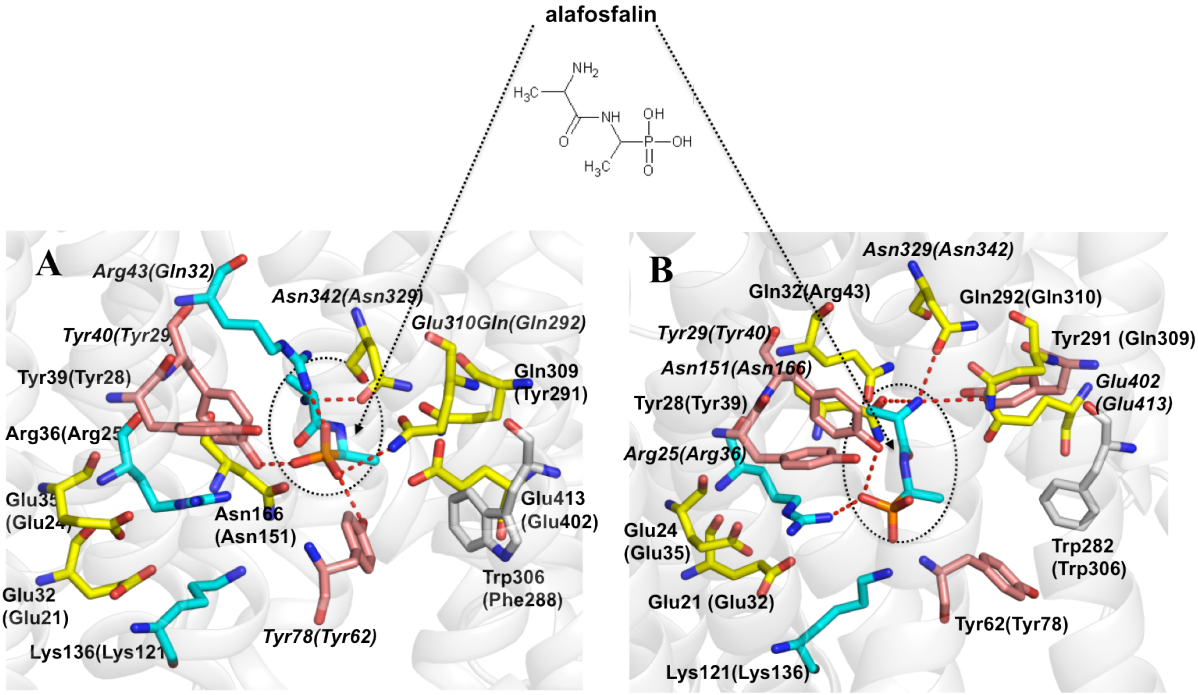


Fig. 8-2: Binding sites of GkPOT, PDB:4IKZ (A) and $PepT_{So2}$, pDB:4LEP (B) in the inward-open conformation, with bound peptide analogue alafosfalin (chemical structure in the centre). The alafosfalin molecule is coordinated by five different residues in each protein (in italics); hydrogen contacts in red. Equivalent residue numbers in the other protein are listed in brackets.

In the structure of the complex, the alafosfalin molecule contacts five conserved binding site residues (fig. 8-2, A), with the Asn342 side chain contacting the N-terminal amino group and the other four coordinating the phosphonate group (in the sulfate-bound structure, the ion is coordinated by the same residues that contact the phosphonate group). Alafosfalin-interacting residues in GkPOT are conserved in $PepT_{St}$, as previously discussed (5.4). In $PepT_{So2}$ however, the molecule is oriented and coordinated differently (for example the amino group is coordinated by a total of three side chains), with some important substitutions among the binding site residues when compared to GkPOT (fig. 8-2, B). The difference between the two structures once more stresses the diversity of binding modes and transport mechanisms in the POT family. The $PepT_{So2}$ structure was published during the write-up for this thesis and there is currently little biochemical data relating to it, so it is

difficult to extract functional implications from the two structures. Nevertheless, a functional analysis was earlier performed on GkPOT, and the results are presented below.

8.2.1. Peptide specificity in wild type GkPOT

For all functional assays, GkPOT was purified from a GFP-fusion construct and reconstituted into liposomes as described. ΔΨ-driven uptake of ^{3}H-AlaAla was first assayed (fig. 8-3, A). The peptide was transported in a time- and temperature-dependent manner, with a rate similar to that recorded for $PepT_{St}$ (fig. 5-4, A).

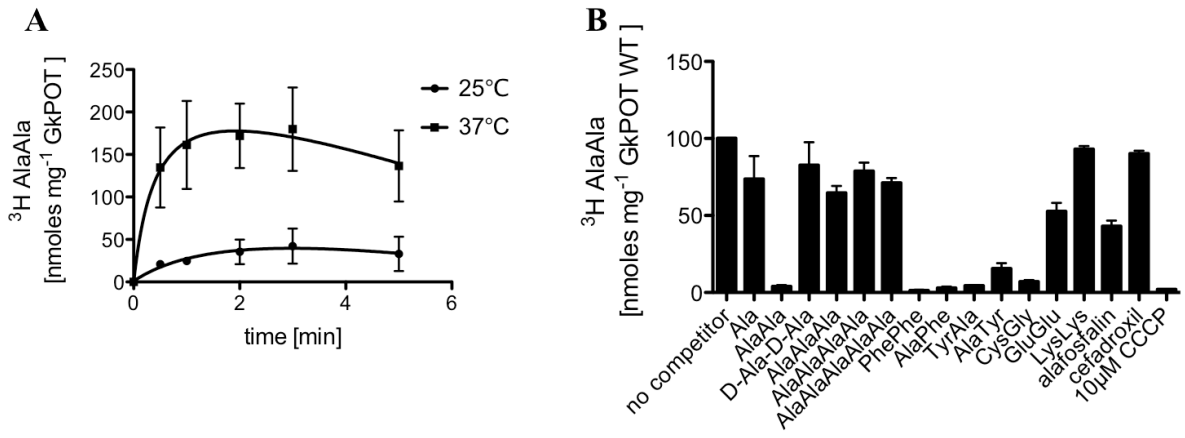


Fig. 8-3: Transport profile of wild type GkPOT.
(A) Time-dependent uptake of ^{3}H-AlaAla by liposome-reconstituted GkPOT at pH 6.5. Proteoliposomes containing internal buffer (20 mM KPi, 100 mM KAc and 2 mM $MgSO_4$) were diluted 1:50 in external buffer (120 mM NaPIPES and 2 mM $MgSO_4$) containing 10 μM ^{3}H-AlaAla and 10 μM valinomycin, to a final protein concentration of 0.7 μM.
(B) Peptide competition assays. Under the same conditions, uptake of 15 μM ^{3}H-AlaAla by 1 μM protein was measured at 25°C after 2 min, in the presence of 2 mM competitor unless specified.

Peptides were then tested for competition against AlaAla uptake, establishing a profile similar to that of $PepT_{St}$, $PepT_{So}$ and YdgR (fig. 8-3, B). In contrast to $PepT_{St}$ and YdgR however, no competition by cefadroxil was detected. Also intriguingly, affinity for the tripeptide AlaAlaAla appears very low (only a reduction of ~30% in uptake levels, compared

to 80-100% typically seen with the other proteins). Counterflow assays with cefadroxil and AlaAlaAla also did not show any evidence of their transport. It is therefore possible that GkPOT does not transport any tripeptides with high affinity (perhaps connected to its high sequence identity to $PepT_{St}$, which also prefers dipeptides), but more data would be needed to prove that assumption.

Alafosfalin itself has a weak competitive effect on uptake, similar to that seen for GluGlu. A counterflow assay with alafosfalin also did not reveal any detectable substrate-driven uptake, suggesting the affinity for it is very low (likely in the milimolar range). There is no data to show transport of alafosfalin by $PepT_{So2}$, but it has been shown that adding alafosfalin to $PepT_{So2}$ increases its stability in a thermal shift assay (Guettou *et al*, 2013). If this is suggestive of a higher affinity for alafosfalin in the case of $PepT_{So2}$, then it is conceivable that the different orientations of the molecule in the two structures represent low affinity (GkPOT) and high affinity ($PepT_{So2}$) substrate binding modes.

Lastly, IC_{50} values were calculated for dipeptides AlaAla (31 μM), AlaPhe (5.4 μM) and GluGlu (4.5 mM) (fig. 8-4). As with $PepT_{St}$, AlaPhe and AlaAla are high affinity substrates, whereas the affinity for GluGlu is significantly lower compared to $PepT_{St}$ (10-fold), in keeping with the weak competitive effect seen in the endpoint assay.

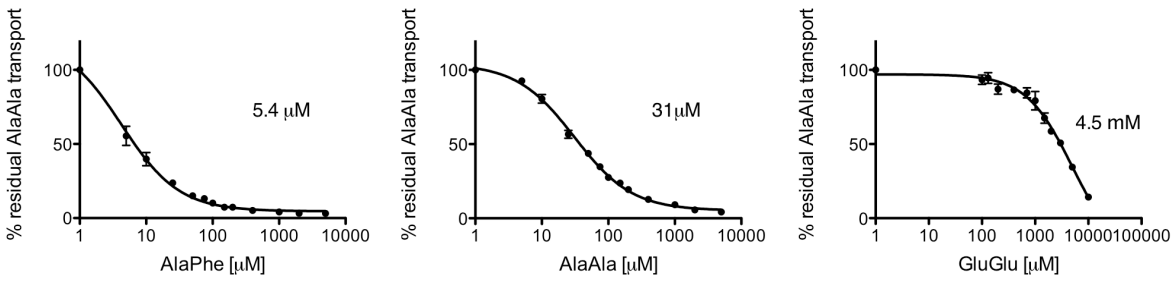


Fig. 8-4: IC_{50} curves measuring competition by dipeptides against ^{3}H-AlaAla uptake by GkPOT. Residual uptake of the reporter (Y-axis) is shown in percentages (relative to the uptake level without competing peptides). Competitor concentrations (X-axis) are shown in logarithmic scale. Measurements were taken after 2 min incubation in ΔΨ-driven conditions. IC_{50} values are listed beside each graph.

8.2.2. Binding site mutagenesis

The main aim of the GkPOT study was to elucidate the roles of key binding site residues in the transport and coupling mechanism. To this end, several mutant variants were purified, reconstituted and assayed for ΔΨ-driven transport and counterflow of AlaAla (table 8-1).

GkPOT mutation	$PepT_{St}$ equivalent residue	$PepT_{So}$ equivalent residue
E32Q	E22	E21
E35Q	E25	E24
R36A	R26	R25
Y40A	Y30	Y29
R43Q	R33	R32
Y78F, Y78A	Y68	Y68
K136A	K126	K127
N166A	N156	N158
W306A	W296	W312
E310Q	E300	D316
N342Q, N342A	N328	N344
E413Q	E400	E419

Table 8-1: Binding site mutagenesis in GkPOT, with equivalent amino acids from $PepT_{St}$ and $PepT_{So}$.

All the amino acids mutated in GkPOT are conserved in $PepT_{St}$ and $PepT_{So}$, and their roles appear to be mostly conserved as well, with some exceptions (fig. 8-5). As in $PepT_{St/So}$, mutating the two glutamates on H1 (here, Glu32 and 35) to glutamine has a severe effect on both ΔΨ-driven uptake and counterflow, suggesting a role in coupling. Other residues of the

ExxERFxYY motif, Arg36 and Tyr40, have a critical role as well, as alanine substitutions completely inactivate the transporter. Since the tyrosine is involved in binding alafosfalin in both GkPOT and $PepT_{So2}$ (fig. 8-2), the interaction between substrate and its OH group is likely to be essential (the same tyrosine is essential for transport in $PepT_{St}$ and binds the tripeptide in the recent co-complex structure – fig. 5-17, B). Similarly, the arginine, which is stabilised by salt bridge contacts to the second glutamate on H1 in all inward-open structures, contacts the dipeptide in the $PepT_{St}$-AlaPhe complex (fig. 4-8, A) and alafosfalin in the $PepT_{So2}$ structure (fig. 8-2, B). All four residues are thus likely to play roles in proton coupling and substrate interactions, either directly or indirectly through their relative positioning in the binding site.

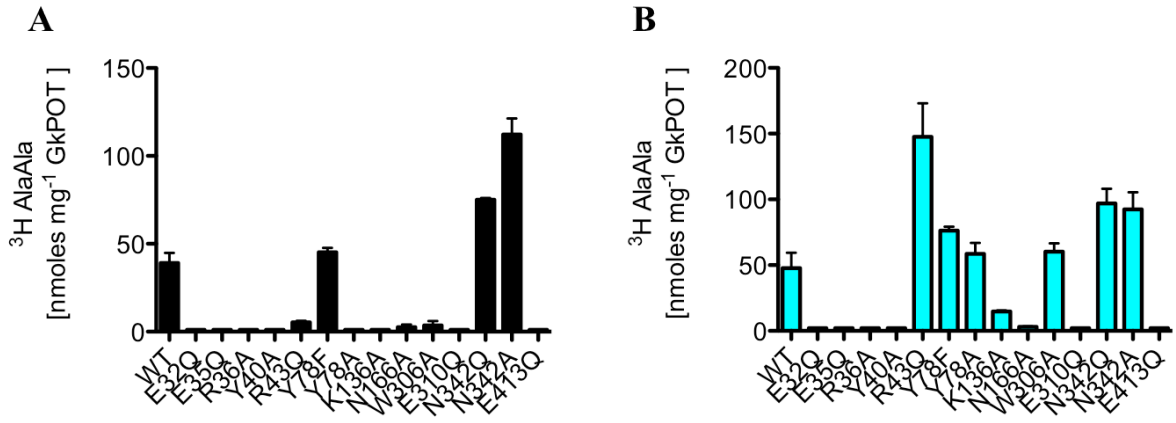


Fig. 8-5: Functional role of binding site residues in GkPOT.
(A) Effect of mutations on ΔΨ-driven transport. Uptake of 15 μM ^{3}H-AlaAla by 1 μM protein was measured at pH 6.5 and 25°C after 2 min.
(B) Effect of mutations on peptide-driven transport (counterflow). Uptake of 5 μM ^{3}H-AlaAla by liposomes loaded with 10 mM AlaAla containing 1 μM protein was measured at pH 6.5 and 25°C after 10 min.

Tyr78 (H2) plays the same structural role in GkPOT as its conserved equivalents in $PepT_{St/So}$ (the alanine mutation is detrimental, but the phenylalanine variant is indistinguishable from the wild type), a role reinforced by the position of this tyrosine with respect to the alafosfalin molecule in both structures. Also similar to their $PepT_{St/So}$ counterparts are the roles of Lys136 and Glu413, the pair of residues forming the salt bridge at the bottom of the central cavity to close the intracellular gate in the outward-open conformation. As seen before, mutating both amino acids in GkPOT decouples the transporter

from the membrane potential and only the Lys136Ala variant still performs counterflow, reinforcing the importance of the inter-domain salt bridge.

Another universally conserved residue, Asn166, is essential to both types of transport. The residue does not interact with the peptide in the GkPOT structure, but is predicted to enter the binding site in the occluded state (Doki *et al*, 2013) and indeed it is seen contacting the phosphonate group in the $PepT_{So2}$ structure (fig. 8-2, B), as well as the N-terminus of the dipeptide in the $PepT_{St}$ structure with AlaPhe (fig. 5-17, A). Interestingly, mutations of the other strictly conserved asparagine, Asn342, do not hinder GkPOT function. The asparagine is essential in $PepT_{St/So}$, where it is involved in the structural rearrangements that open and close the extracellular gate and facilitate coupling, as discussed (6.2.2). Additionally, in GkPOT and $PepT_{So2}$, it forms a hydrogen bond to the amino group of alafosfalin (fig. 8-2, A and B). However, the recent structures of $PepT_{St}$ with bound physiological peptide substrates seem to suggest this contact is not essential (neither the N-terminal amino group of the dipeptide nor that of the tripeptide interact with any residues in the binding site, although Asn328 is in a similar position with respect to the AlaPhe molecule as that seen in GkPOT with alafosfalin) (fig. 5-17, A and B). Indeed, repeating the GkPOT competition experiment with the Asn342Gln and Asn342Ala proteins results in a profile identical to that of the wild type (fig. S6), so the asparagine does not influence peptide selectivity. Therefore, the critical role of the residue in $PepT_{St}$ and $PepT_{So}$ might be due to its involvement in coupling (as it contacts Glu400/Glu419 in the inward-open state), whereas in GkPOT it might serve to contact the amino terminus of peptides, which is not essential for transport. In addition, the equivalent glutamate in GkPOT contacts another carboxamide side chain in the inward-open state (Gln309, which is a glutamate/phenylalanine in $PepT_{St/So}$); the role of this is unclear, but it may substitute for the Asn342 interaction.

Trp306 (H7) is conserved in $PepT_{St/So}$, and its mutation to alanine abolishes GkPOT-mediated transport. As discussed (5.3.3), the tryptophan facilitates the transition from the inward-open state to the occluded, opening the intracellular gate by allowing H7 to kink at this position. This movement also brings another residue on H7, Glu310 (Glu300 in $PepT_{St}$, Asp316 in $PepT_{So}$) into the position where it can form the central salt bridge to Arg43, as seen in the $PepT_{So}$ and $PepT_{St}$ structures as well as in MD simulations of the occluded GkPOT (Doki *et al*, 2013).

In the sulfate-bound structure of GkPOT, the negatively charged ion is 2.7Å from Glu310, which suggests the glutamate is protonated (Doki *et al*, 2013). The phosphonate group of alafosfalin exhibits a tetrahedral symmetry similar to the sulfate ion and is coordinated in a similar way within the binding site. Although the Glu310Gln variant binds alafosfalin in the crystal, possibly by mimicking the H^+-bound state of the glutamate (fig. 8-2, A), it is inactive in both ΔΨ-driven and counterflow assays (fig. 8-5), as it cannot form the salt bridge to Arg43 and induce the occluded conformation during transport.

On the opposite side of the cavity, the mutation of Arg43 to glutamine terminates ΔΨ coupling but allows counterflow. As in $PepT_{St}$, the guanidinium group of the arginine likely regulates the pK_a of the Glu310 side chain, facilitating its protonation (fig. 8-6, A and B).

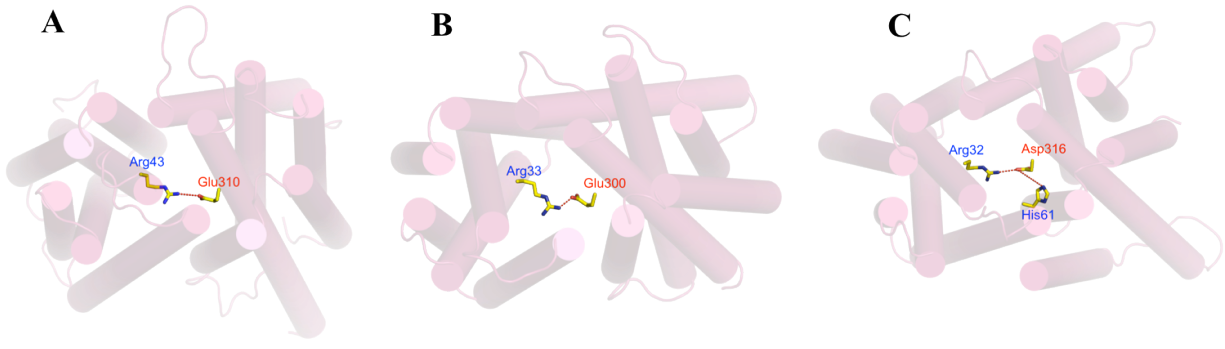


Fig. 8-6: Top views of the GkPOT (A) and the $PepT_{St}$ (B) structures, both in the inward-open conformation, and the inward-open model of $PepT_{So}$ (C). In all three proteins, the inward-open state features the central inter-domain salt bridge between an arginine and a glutamate/aspartate, with an additional contact between the aspartate and His61 in $PepT_{So}$.

When the arginine is mutated in GkPOT (as in $PepT_{St}$), the glutamate remains strongly protonated and can no longer form the salt bridge to the arginine and thus the transporter is blocked (fig. 8-5). In $PepT_{So}$ however, the arginine is not required for transport, likely due to the additional contact that Asp316 makes with His61 (fig. 8-6, C), which stabilises the inward-open conformation by keeping H7 packed against H1/H2 (as discussed in 6.1.1).

To test the role of the arginine in GkPOT, the Arg43Gln variant was assayed over a range of pH values. Indeed, increasing the pH of the transport assay gradually rescues the activity of the Arg43Gln variant, as it increases the likelihood of the glutamate being deprotonated, thereby allowing the conformational change back to the occluded state (fig. 8-7). However this does not result in a complete rescue, as the external pH cannot fully compensate for the absence of the arginine side chain, particularly as the arginine may also be involved in binding the substrate. This reinforces the role of Arg43 in regulating the pK_a of Glu413, and the importance of the central inter-domain salt bridge.

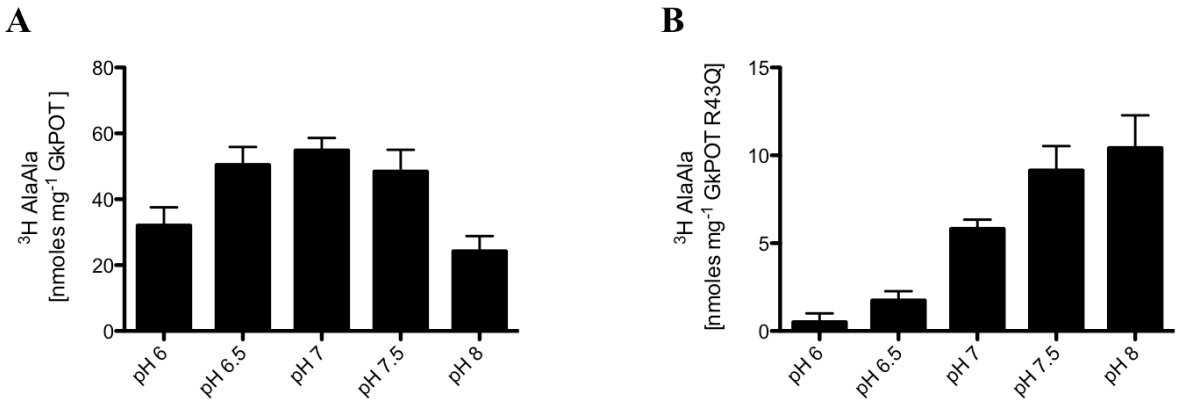


Fig. 8-7: pH dependence of ^{3}H-AlaAla uptake by GkPOT wild type (A) and R43Q (B). Proteoliposomes containing 50 mM KPi at each pH were diluted 1:50 in 50 mM NaPi at the same pH containing 10 μM valinomycin. Uptake of 15 μM ^{3}H-AlaAla was measured after 2 min at 25°C.

8.3. Conclusions

GkPOT is the first oligopeptide transporter to be co-crystallised with a peptidomimetic molecule in the binding site. The structure revealed a number of potentially significant interactions within a central cavity completely conserved with respect to $PepT_{St}$. The functional characterisation of GkPOT indeed revealed a substrate specificity not dissimilar to that of $PepT_{St}$, but with the important distinction that trialanine acted as a particularly weak substrate for GkPOT. In contrast, YdgR (the only bacterial POT to date to show direct uptake of a drug substrate) has a higher affinity for tripeptides, perhaps linked to its affinity for cephalosporin antibiotics. Generally however, both YdgR and GkPOT prefer hydrophobic dipeptides, and particularly those with aromatic side chains, a trait that can be linked to the structural role of the tyrosine residues in the binding site, as identified in GkPOT through site-directed mutagenesis and subsequent transport assays.

Overall, the data on GkPOT suggests a transport mechanism akin to that of $PepT_{St}$, with the conserved salt bridge network on H1 and conserved interhelical salt bridges regulating conformational changes as well as coupling. However, a potentially different, non-essential role for Asn342 in substrate binding is likely, and new key residues are identified in Asn166 and Trp306. The roles of Arg43 and Glu310 are central to the mechanism, as the glutamate is likely a protonation site and the arginine stabilises its charge to enable structural changes. Furthermore, both residues are involved in substrate interactions, with the glutamate in particular contacting negatively charged functional groups when in protonated form (although this may not be the case for other transporters like $PepT_{So2}$ and/or may depend on the substrate).

Chapter 9: On the path to binding: NMR spectroscopy of oligopeptide transporters

As outlined in the previous chapters, transport assays in whole cells and liposomes have represented the main approach to measuring the activity of secondary active transporters to date. In addition to these, several other methods have been used to measure direct binding of transported substrates *in vitro*. This can be problematic, as transporters (and membrane proteins in general) normally require a lipid bilayer for function, and as such the detergent can make it impractical to measure binding *in vitro*. A telling example, surface plasmon resonance (SPR) is perhaps the most popular technique for binding studies of soluble proteins, but has rarely yielded successful results for membrane proteins (Maynard *et al*, 2009). Furthermore, transporter proteins present an additional problem in the transient nature of "binding" in the case of transported substrates compared to conventional ligand-receptor interactions, making it difficult to measure affinities using traditional binding models.

Despite the limitations, some methods have proved successful. Of these, isothermal titration calorimetry (ITC) was used to calculate dissociation constants for high affinity (low nM) protein-ligand interactions like the one between LeuT and leucine (Piscitelli *et al*, 2010) or between EmrE and tetraphenylphosphonium (TPP^+) (Morrison *et al*, 2012). ITC was also used in conjunction with site-directed mutagenesis to identify binding site amino acids involved in regulating the affinity for arginine in APC superfamily transporter AdiC (Kowalczyk *et al*, 2011; Tsai *et al*, 2012). More recently, K_D values in the high micromolar

range were calculated for nitrate/nitrite exchanger NarU in DDM (Yan *et al*, 2013). Nevertheless, measuring ligand binding by ITC can be hindered by the presence of the detergent micelles, as indicated by a study on MFS transporter FucP, for which fucose binding was detectable in a neopentyl glycol detergent (Cymal-7) but not in nonylglucoside (NG) (Dang *et al*, 2010). Alternatively, binding of lactose and TDG to LacY was measured in lipid vesicles by quantifying the efficiency of labelling alkylated residues upon addition of substrates (Guan and Kaback, 2004).

Apart from the low affinity, binding measurements of transport substrates can be affected by the low sensitivity of the method employed (especially when the transporter cannot be purified in high amounts). One way to overcome this is using a radioactivity-based method such as a scintillation proximity assay (SPA). Designed to measure short distance interactions, SPA involves attaching proteins (via a histidine tag) to scintillating beads, which emit light in the presence of radioactively-labelled ligands in their proximity (bound to the protein) but do not detect the free label present in solution (unbound) (Udenfriend *et al*, 1985). In recent years, the technique has been applied to measure substrate binding by several transporters, such as the binding of tyrosine to NSS transporter Tyt1 in detergent-solubilised membranes (Quick and Javitch, 2007). Leucine binding to DDM-purified LeuT was assayed by SPA and ITC in parallel, offering similar results (Piscitelli *et al*, 2011). More recently, K_D values for ATP and AMP-PNP were calculated for an ABC exporter using SPA with ^{3}H-labelled nucleotides (Hohl *et al*, 2012).

Another method traditionally used for structure determination as well as binding studies is NMR spectroscopy, and like the methods discussed above its use with membrane proteins is limited by several challenges. In particular, NMR experiments typically require large amounts of protein as well as good stability, as spectra are often collected at high temperatures to increase the tumbling rate of particles in solution and, with it, the resolution

(Kim *et al*, 2009). In addition, protein size can be a severe limitation, as slow tumbling rates for molecules above 30-40 kDa cause increased signal linewidths and consequent deterioration of the NMR spectra. NMR studies of detergent-solubilised membrane proteins are therefore difficult, but nevertheless structure determination was possible in the case of some smaller proteins such as intramembrane enzyme DsbB (Zhou *et al*, 2008) and 7-helical GPCR-like receptor pSRII (Gautier *et al*, 2010). Several strategies have been employed in recent years to improve conditions for alpha-helical membrane protein NMR, such as performing the experiments in bicelles, which despite being larger than detergent micelles can confer a more native environment contributing to the protein's integrity as well as increase the isotropy of the sample to compensate for slow tumbling rates (Nietlispach and Gautier, 2011). An alternative approach is the use of nanolipoprotein particles (nanodiscs), which are smaller in size and have been found to increase protein stability and functionality in several cases (Raschle *et al*, 2010). Moreover, solid state NMR is becoming more frequently used due to its compatibility with large particles (micelles, bicelles or liposomes), as the signal quality is independent of the molecular size (McDermott, 2009). Recently, a protocol for NMR data collection was recently developed for multidrug transporter EmrE reconstituted into bicelles (Morrison and Henzler-Wildman, 2011), and a later study showed the protein was functional and yielded higher resolution spectra in bicelles compared to DDM micelles (Morrison *et al*, 2012). The quality of the EmrE data is encouraging, as it indicates that NMR can be successfully used for structural and functional studies of transporter proteins.

Outside of transport assays, functional data on POT family transporters has remained scarce. Peptide binding by ITC was measured for *E. coli* transporter YjdL (Jensen *et al*, 2012b), but the data is conflicting: AlaLys binds with high affinity (K_D 14 μM) while no binding is detected for AlaGln, despite a previous whole cell transport assay showing them to be substrates with similar affinities (Jensen *et al*, 2012a). More recently, ATR (attenuated

total reflectance) was used in conjunction with FTIR (Fourier transform infrared spectroscopy) to estimate the binding of AlaAla to GkPOT by comparing IR spectra at different peptide concentrations, generating a tentative K_D of 3.5 mM (Doki *et al*, 2013). The difference in affinity between the two examples is significant, and the K_D for AlaAla is very different from the estimated transport affinity (Solcan *et al*, 2012), making the interpretation of these results difficult. Considering the literature summarised above and the scarcity of data on POT substrate binding, ITC, SPA and NMR spectroscopy were attempted with $PepT_{So}$, $PepT_{St}$ and GkPOT and some preliminary data was generated, as detailed below.

9.1. Preliminary binding studies (ITC and SPA)

The first attempt to measure peptide binding, an ITC experiment was carried out with $PepT_{So}$ in detergent solution (DDM), purified from a GFP-fusion construct. Two peptides were used for the measurements: AlaAla and AlaTyr, both known to be high affinity substrates in the liposome-based assay. Different concentrations and protein:peptide ratios were tested as described in the methods chapter, but no binding was detected.

For the scintillation proximity assay, $PepT_{So}$ was purified as a His-tagged construct as described in the methods chapter. The yield was significantly lower compared to the usual yields obtained from the GFP-tagged construct (~1mg pure protein from 6L culture in contrast to 10-15mg), but purity was high and the gel filtration profile was characteristic of $PepT_{So}$ (fig. S7).

Before the assay, the protein samples (purified protein in detergent solution and protein reconstituted into liposomes) were incubated with copper beads using different concentrations and protein:bead ratios, and the efficiency of binding was checked by running samples on a Tris-Gly gel. The protein in detergent solution bound poorly to the SPA beads, as most of it was present in the supernatant (fig. S8, A). In contrast, the liposome-reconstituted protein bound the beads with high efficiency, independent of the concentration of beads (fig. S8, B). However, a similar test showed that reconstituted $PepT_{So}$ without a His tag also bound the SPA beads, suggesting unspecific interactions between the liposomes and the hydrophobic bead surface.

An SPA protocol was set up as described, and measurements were taken over the course of 7 days at different time intervals using $PepT_{So}$ both in detergent and liposomes and increasing concentrations of ^{3}H-AlaAla. Additionally, transport assay conditions were re-created in the SPA plates to generate a $\Delta\Psi$ across the liposome membrane. Peptide binding was not detected in any of the conditions.

9.2. NMR experiments

Solution NMR is a potentially powerful tool for investigating structural changes that occur upon substrate binding. A well-resolved 2D NMR spectrum can provide highly specific information on the residues that are affected by the binding process, as well as the overall dynamics of binding; this is achieved through measuring the chemical shifts corresponding to isotope-labelled residues in the protein during titration experiments with substrate molecules.

A project was initiated in collaboration with Dr Jason Schnell and Dr Jolyon Claridge at the University of Oxford to explore new avenues for studying peptide and drug binding to POT family transporters using solution state NMR, as well as to optimise conditions for NMR studies of large membrane proteins in general. The initial target of this study was $PepT_{So}$, and after preliminary experiments the focus shifted to GkPOT due to stability issues, as detailed below.

9.2.1. 2D amide spectra of $PepT_{So}$

$PepT_{So}$ was expressed in ^{15}N-labelled minimal media and purified as described in the methods chapter. Typically, the yield was ~20mg pure protein from 6L growth culture (fig. S9, A).

A heteronuclear multiple quantum correlation spectrum (HMQC) of the native protein in DDM solution was acquired at 40°C (fig. 9-1, A). The spectrum shows the correlation between each nitrogen-hydrogen pair in the protein's amide groups, and thus each peak corresponds to an individual residue in the protein (all residues except prolines). As expected for a protein with high α-helical content, the centre of the spectrum contains a cluster of poorly resolved peaks. Elsewhere however the peak resolution is relatively high, particularly when considering the large size of the protein-DDM micelles (56.7 kDa protein and ~100 kDa detergent). Resonances are visible for the glycine backbones and asparagine/glutamine side chains (upper part of the spectrum), as well as for the tryptophan indole amines (lower left). Most peaks in the spectrum are homogeneous in width, which indicates that $PepT_{So}$ is overall rigid in detergent solution. A second spectrum was collected after addition of 80 mM AlaTyr (a known transported substrate), but the dipeptide did not cause any changes in the spectrum (fig. 9-1, B). Furthermore, a spectrum of $PepT_{St}$ was acquired in similar conditions but the

resolution was extremely poor (fig. S10), likely due to the protein being predominantly a dimer in solution, as shown previously (fig. 3-3).

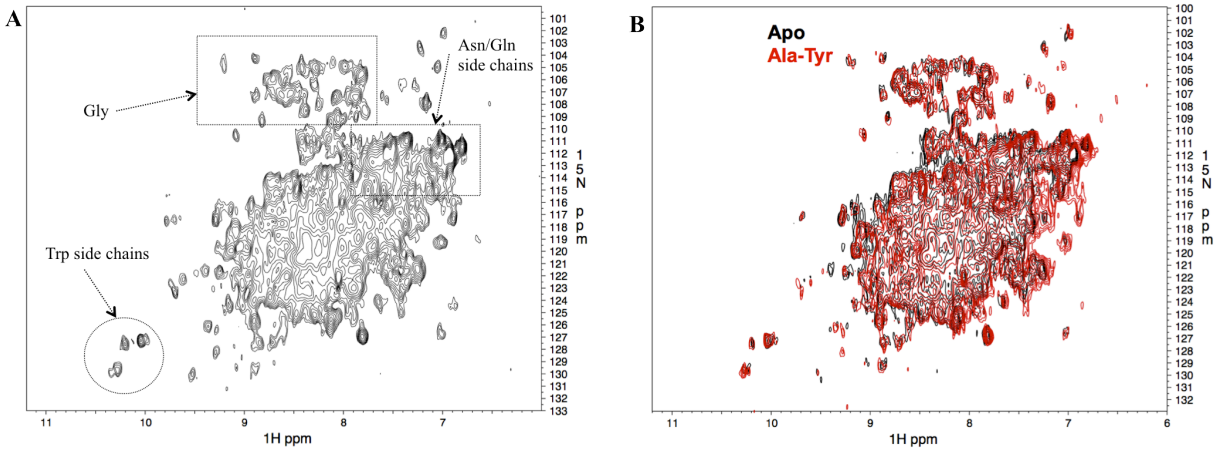


Fig. 9-1: 2D SoFast-HMQC spectra of ^{15}H-PepT$_{So}$ in DDM micelles.
(A) Apo protein spectrum collected at 950 MHz, 40°C, protein concentration 80 µM.
(B) Overlay of the apo spectrum (black) and the spectrum acquired after addition of 80 mM AlaTyr peptide (red).

PepT$_{So}$ in DDM solution was unstable at 40°C beyond the 6h experiment time needed to collect the two HMQC spectra, making it difficult to use the sample for further peptide binding studies. To study the influence of lipid environments on protein stability and the quality of the spectra, isotope-labelled PepT$_{So}$ was incorporated into DHPC:DMPC bicelles, based on the promising results obtained with EmrE in similar conditions (Morrison and Henzler-Wildman, 2011; Morrison *et al*, 2012). Bicelle reconstitution was over 75% successful and did not affect the integrity of the protein (fig. S9, B). Surprisingly however, PepT$_{So}$ precipitated heavily in the DHPC bicelles, and the HMQC spectrum was poorly resolved, with no visible glycine peaks and strong peaks in the middle of the spectrum suggesting partial unfolding (fig. S11, A). An additional spectrum was collected after addition of 80 mM AlaPhe; no chemical shifts were detected, and the protein itself was mostly degraded after the first 3h acquisition time (fig. S11, B).

$PepT_{So}$ stability was greatly increased when DM detergent was added to the DMPC proteoliposomes instead of DHPC to form the bicelles. However, the spectrum resolution did not improve, as DM is less likely to form isotropic bicelles than DHPC due to its long-chain lipid tail (Whiles *et al*, 2002) (fig. 9-2, A). Although the increased stability allowed for longer acquisition times, a spectrum collected in the presence of AlaPhe did not reveal any changes in the positions of the peaks, suggesting the peptide does not bind to the protein under these conditions (fig. 9-2, B).

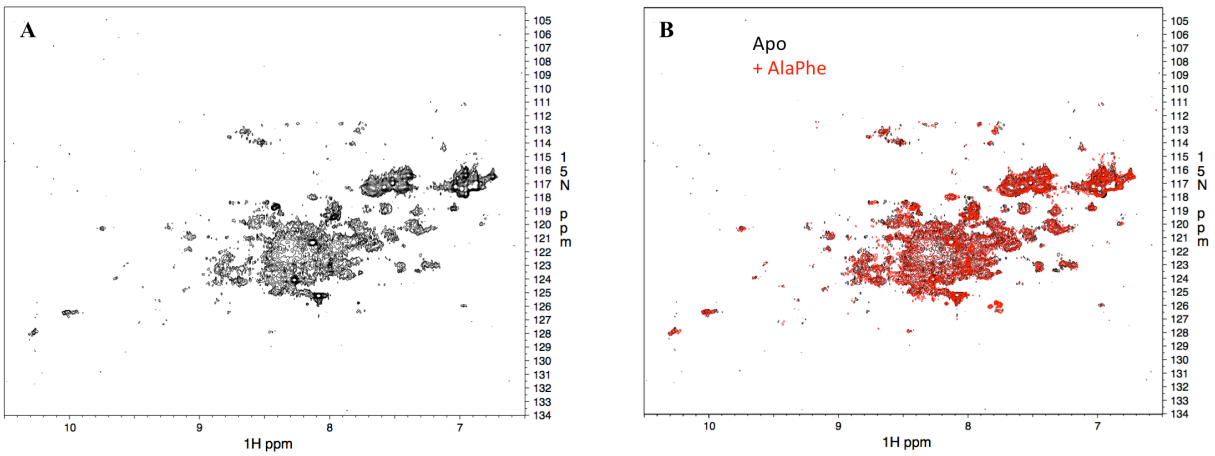


Fig. 9-2: SoFast-HMQC spectra of ^{15}H-$PepT_{So}$ in DMPC:DM bicelles (950 MHz, 40°C, protein concentration 100 µM).
(A) Apo protein spectrum. $PepT_{So}$ was reconstituted into DMPC liposomes (molar protein:lipid ratio 1:130) and DM was added to a 1:3 ratio (DMPC:DM).
(B) Overlay of the apo spectrum (black) and the spectrum acquired after addition of 80 mM AlaPhe (red).

9.2.2. GkPOT as NMR candidate: the search for thermostability

Despite relatively well resolved 2D spectra, $PepT_{So}$ proved a difficult target for NMR studies due to its instability at high temperature, while $PepT_{St}$ was not amenable due to the large size of the dimer. An alternative target was revealed in the form of GkPOT when a thermostability experiment performed at 50°C showed it to be completely stable in DDM solution over a 24h

time course, while $PepT_{So}$ precipitated out of solution shortly after the first measurement (fig. 9-3).

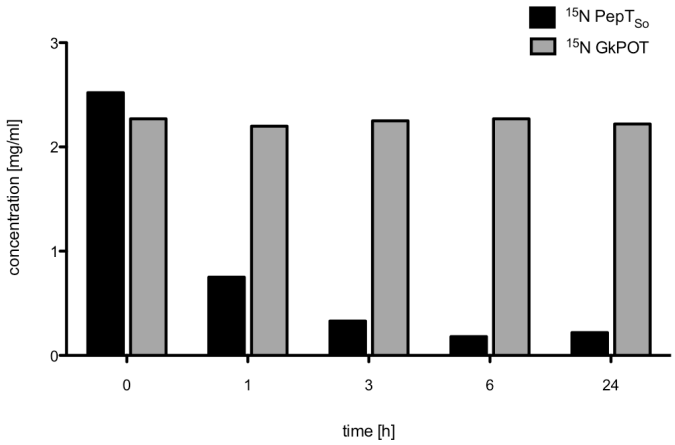


Fig. 9-3: Comparative stability of ^{15}N-labelled $PepT_{So}$ and GkPOT in detergent solution. 20 µl samples at approx. 2.5 mg/ml starting concentration (in 20 mM Tris pH 7.5, 150 mM NaCl and 0.03% DDM) were incubated at 50°C. At the given time points, the samples were centrifuged, the precipitated pellets were discarded and the protein concentration of the supernatant solution measured by A_{280}.

^{15}N-labelled GkPOT was expressed and purified using the same protocol as for $PepT_{So}$ and yielded ~24mg pure protein from 6L (fig. 9-4).

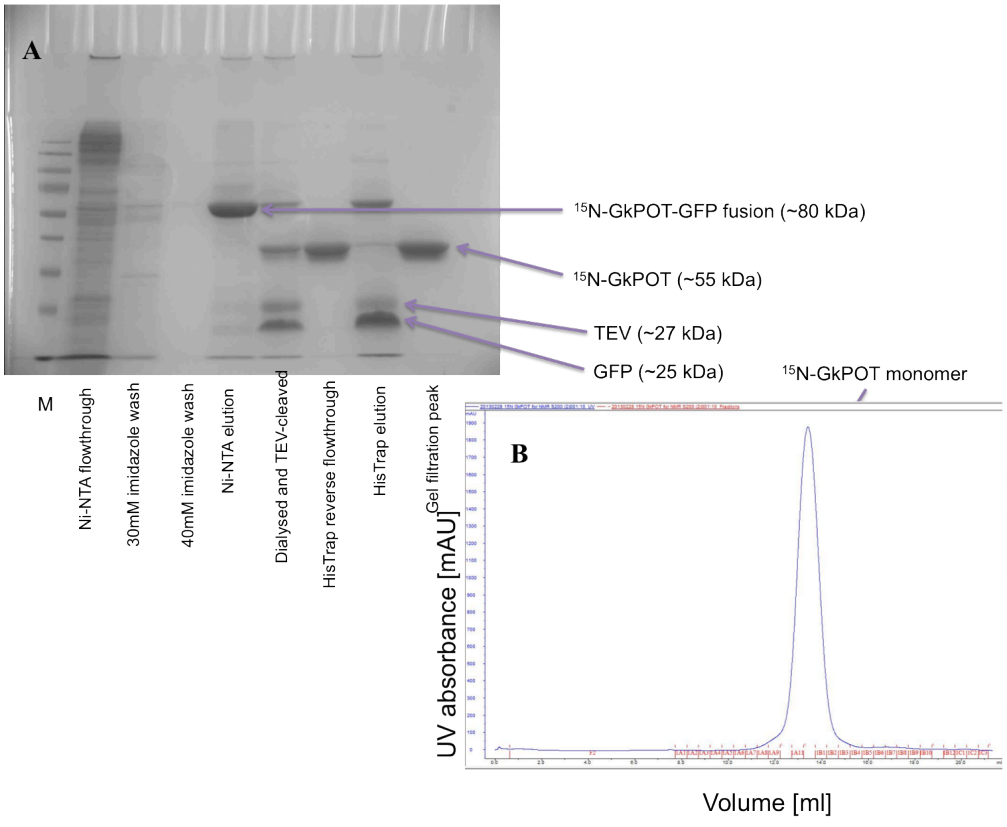


Fig. 9-4: Purification of ^{15}N-GkPOT for NMR experiments.
(A) TrisGlycine gel showing the purification steps.
(B) Gel filtration profile.

The SoFast-HMQC spectrum of GkPOT in DDM micelles collected at 47°C is similar to the $PepT_{So}$ spectrum, with a central cluster of amide peaks in the centre and several clearly distinguishable regions such as the glycine and tryptophan peaks (fig. 9-5, A). Overall, the resolution is lower compared to $PepT_{So}$; however, the protein sample was highly stable over time, and a titration experiment was carried out with dipeptide substrate AlaPhe. As shown previously (8.2.1), AlaPhe competes with high affinity against GkPOT-mediated AlaAla uptake in the liposome-based assay (IC_{50} = 5.4 µM). Spectra were collected at increasing concentrations of peptide, from 50 µM to 20 mM and superimposed onto the apo spectrum.

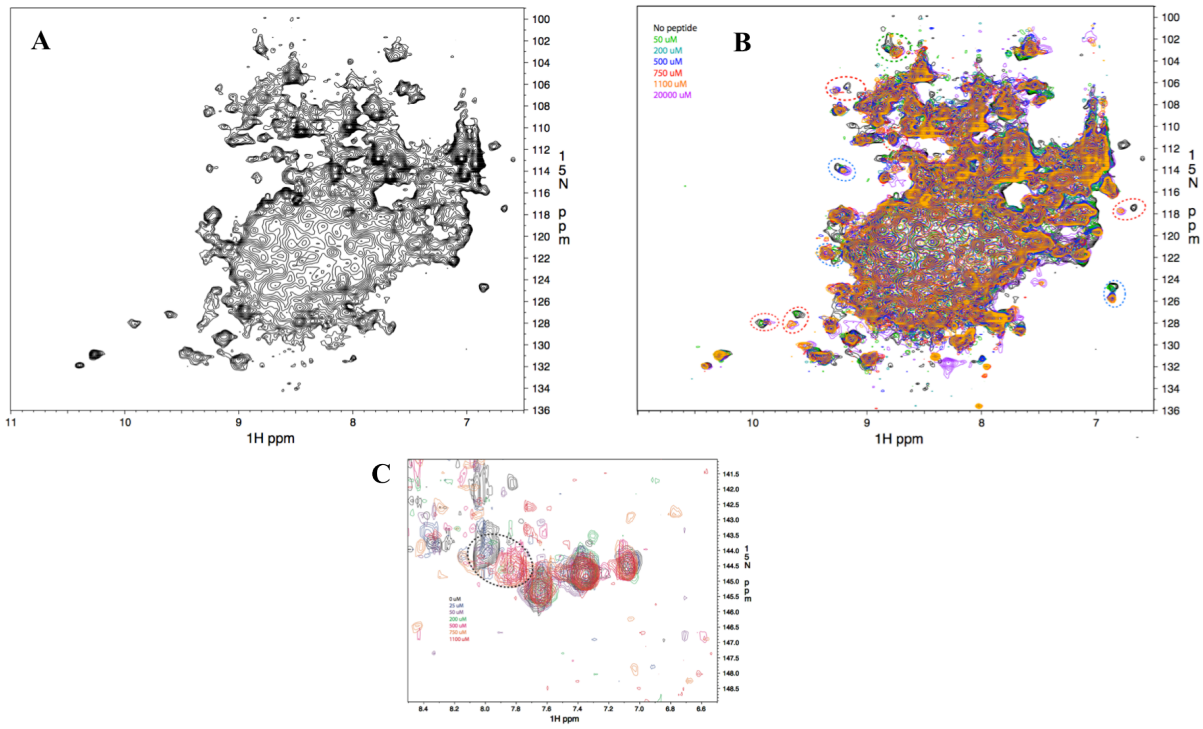


Fig. 9-5: NMR titration experiment showing AlaPhe binding to ^{15}N-GkPOT in DDM solution.
(A) SoFast-HMQC spectrum of the apo protein collected at 950 MHz and 47°C, protein concentration 150 µM.
(B) Overlay of the apo spectrum (black) and the titration spectra (green to purple). AlaPhe concentrations in matching colours, notable chemical shifts encircled (red for slow exchange, blue for fast exchange).
(C) Overlay of the arginine amide spectra from the same experiment. One of the four peaks (circle) shifts during the titration.

The overlay (fig. 9-5, B) shows chemical shifts in several peaks occuring between the 200 µM and 1 mM titration points, indicative of binding. It can be estimated that saturation occurs between 500 µM and 1 mM. Some peaks (red circles in the figure) disappear from the

spectrum and reappear in a different position at a certain titration point; this exchange regime is indicative of "slow" exchange. Other peaks (blue circles) show the opposite behaviour: they shift gradually upon ligand addition and the radius of their shift is smaller, indicating "fast" exchange. An intrinsic spectroscopic property, the exchange regime provides information on the k_{ex} (the rate constant that describes the overall rapidity of exchange), and through it, on the binding kinetics. More specifically, based on the visible chemical shifts, the GkPOT spectrum shows an overall "intermediate" exchange regime, for which the k_{ex} can be determined as:

$k_{ex} = 2\pi \cdot |\delta_f - \delta_b|$

, where δ_f and δ_b are the chemical shifts (in Hz) of the free and bound states respectively (Fejzo *et al*, 1999). In GkPOT, one of the arginine side chain amide peaks (found in a different region of the spectrum, between 140 and 150 ppm in the nitrogen dimension) is the most representative example of intermediate exchange (fig. 9-5, C). Therefore, the k_{ex} was calculated as the area of the circle described by the chemical shift of the arginine peak in the hydrogen dimension (with a radius of 20 Hz): $k_{ex} = 2\pi \cdot 20$ Hz $= 1250$ s^{-1}. As k_{ex} is also equal to the sum of the dissociation and association rates ($k_{on} + k_{off}$), it follows that the lifetimes of both the unbound and bound states are in the order of 1 ms.

The dissociation constant (K_D) can also be approximated by plotting a characteristic chemical shift like that of the arginine amide against the peptide concentration. Albeit a rough estimate, the K_D is likely in the hundred μM range, suggesting that the binding affinity for AlaPhe in detergent solution is over 10-fold lower than the estimated transport affinity *in vitro*, and represents an intermediate value compared to previously determined K_D's for AlaLys/YjdL (Jensen *et al*, 2012b) and AlaAla/GkPOT (Doki *et al*, 2013).

9.3. Conclusions

Understanding the structural implications of a binding event can be a challenging task, particularly in the case of transporter proteins, in which binding sites are likely designed for transient accommodation of substrates. Assigning specific amino acid-substrate interactions is especially difficult for transporters like the POTs, which recognise and transport a wide range of peptide and non-peptide substrates. In addition, very few POT inhibitors have been identified to date (Theis *et al*, 2002b; Ekins *et al*, 2005), and no single structural requirement has been deemed essential for the recognition of substrate molecules (Daniel and Kottra, 2004). Under these conditions, identifying which residues interact with a bound substrate depends on indirect deductions from transport assays and crystallographic data. While the functional work on $PepT_{St}$ and $PepT_{So}$ presented here has provided considerable information on how substrate affinity is regulated within the binding site, more data from alternative biochemical methods is needed for an accurate mapping of the binding events.

The NMR spectroscopy approach in particular has proven successful, as it is the first to show direct proof that peptide binding in solution affects the transporters on a structural level, in a concentration-dependent manner. Although the data resolution is insufficient to assign the peaks to specific residues, the chemical shifts seen upon AlaPhe addition in GkPOT indicate that binding induces discrete, localised structural changes rather than long-range conformational changes. The data also indicates that at least one arginine side chain reorients itself upon binding, and it can be speculated that this is one of the two conserved binding site arginines on H1.

Originating from a thermophilic organism, GkPOT has proven a good target for NMR experiments, as it is stable for long time periods at high temperatures in the NMR magnet,

making it possible to collect more data and perform titration experiments using a single sample. Currently a work in progress, the GkPOT NMR project aims to characterise peptide binding in detergent as well as in lipid bilayer environments (as attempted earlier with $PepT_{So}$), and identify the binding site residues involved in binding. This can be achieved by performing titration experiments on samples with specifically labelled residues (such as tyrosines) as well as on mutant versions of GkPOT (for example, mutating the H1 arginines and comparing the arginine chemical shifts in the mutant protein spectrum to the wild type might serve to determine their role in substrate binding).

Chapter 10: Final discussion & perspectives

In 2010 when this research project began, structural studies of oligopeptide transporters were in the early stages. The $PepT_{So}$ crystal structure had been solved and would be published shortly, and diffraction data was being collected from crystals of a second transporter, $PepT_{St}$. At this stage, the project branched out in two main directions. The first of these involved optimising $PepT_{St}$ crystallisation (as well as collecting additional data for phasing the native dataset), seeking to improve crystal resolution and identifying new targets for POT structural studies – as summarised in chapters 3 and 4. The latter aspect was focused on screening POT homologues for expression, purification and crystallisation, based on the information gathered from previous studies on transporters and other membrane proteins with regard to expression systems, detergents, and crystallisation conditions.

The second direction was the functional characterisation of $PepT_{St}$ and $PepT_{So}$, building on the two crystal structures viewed in the context of MFS transporters, as well as on over two decades of functional work on the mammalian oligopeptide transporters. The results of the functional study are presented in chapters 5 and 6, with more recent contributions (including studies of other POT homologues) discussed separately in chapters 7-9.

The structures of $PepT_{So}$ and $PepT_{St}$ proved that POT proteins belong to the Major Facilitator Superfamily, confirming earlier predictions in spite of their relative sequence divergence compared to most other MFS families. The structures were also the first examples in the literature of two MFS transporters from the same family captured in different conformations, making it possible to draw a direct comparison between the structural aspects

of their transport mechanisms. The conformational changes from the occluded ($PepT_{So}$) to the inward-open structure ($PepT_{St}$) appeared to occur at conserved positions, and regions in both proteins were identified as putative "gates" that serve to provide alternating access to the central cavity from the intracellular and the extracellular/periplasmic sides of the membrane. The structures also revealed the peptide binding site, located in the central cavity, containing conserved residues in an arrangement that would potentially accommodate a variety of peptide substrates. More recently, other structures of oligopeptide transporters were solved (GkPOT, $PepT_{So2}$ and $PepT_{St}$), some of them with bound peptides. The new structures were used to refine the interpretation of the biochemical, biophysical and computational data, as they suggest different binding modes for different substrates (high affinity for dipeptides, low affinity for tripeptides) and highlight some novel interactions within the binding site.

The $PepT_{St/So}$ functional study required a robust experimental system for assaying peptide transport. Preliminary assays in whole cells confirmed that the proteins transported typical PepT1/PepT2 substrates with comparable rates, and validated the use of radiolabelled substrates for direct uptake measurements. However, the *in vivo* system was not ideal for accurate measurements, and thus a protocol was developed for functional reconstitution into liposomes. The *in vitro* setup allowed for accurate protein quantification and made it possible to alter buffer conditions to examine the effect of ion gradients, pH, salt and substrates on the uptake of reporter peptides and drugs.

Like their eukaryotic counterparts, all proteins tested ($PepT_{So}$, $PepT_{St}$, GkPOT and YdgR) take up a wide range of di- and tripeptides under the influence of an artificially generated membrane potential, but they also transport peptide substrates down their concentration gradient in the absence of ion gradients (although proton translocation may be involved). The proteins have a distinct preference for hydrophobic peptides (mainly due to the presence of aromatic side chains in the binding site), and general display greater affinity for

dipeptides than for tripeptides. Overall, the peptide specificity profiles of the wild type proteins establish bacterial POT members as models for studying coupled peptide transport in eukaryotes.

Binding site residues are generally highly conserved in bacterial as well as eukaryotic POT members, and most have important contributions to the transport mechanism. In particular, the residues of the ExxERFxYY motif on H1 are involved in coupling, as are the other conserved polar residues lining the C-terminal side of the central cavity, as mutations of these residues tend to abolish ΔΨ-dependent transport. Additionally, many of these polar residues are seen interacting with the peptide main chain or side chain in the peptide-bound structures. In contrast, two of the aromatic residues at the centre of the binding site play a purely structural role in all proteins, serving to accommodate hydrophobic substrates. An important distinction is the presence of His61 at the contact point between the extracellular cavity and the binding site in $PepT_{So}$, only conserved in some bacterial members but found in both PepT1 and PepT2. Its intermediate pK_a makes the histidine side chain easily accept or donate a proton, and indeed it has been confirmed as the main protonation site in the human transporters, an observation sustained by mutagenesis data in $PepT_{So}$. The histidine forms a crucial contact with Asp316 on H7 in the inward-open and occluded states of the transport cycle, closing the extracellular gate. This interaction is unique to $PepT_{So}$ and suggests a slightly altered proton coupling mechanism compared to transporters that do not have a titratable residue at this position.

The role of the aspartate on H7 (glutamate in $PepT_{St}$ and GkPOT) is otherwise conserved, as it forms the central inter-domain salt bridge to Arg32 on H1, one of three such salt bridges stabilising the three conformational states in the alternating access model (inward-open, occluded and outward-open). In $PepT_{So}$, the arginine side chain is not required for transport, as the presence of the histidine may compensate for its absence. In the other two

proteins, the arginine-glutamate salt bridge is critical to both coupling and the conformational transitions between states, echoing a similar interaction at the centre of the LacY transport mechanism. Both the arginine and the glutamate (in its protonated state) are also likely to bind the peptide.

A similar function can be attributed to the lower inter-domain salt bridge (Lys126-Glu410 in $PepT_{St}$), the mutation of which affects ΔΨ-coupled transport and counterflow in all proteins. Both residues interact with dipeptide AlaPhe in the substrate-bound structure of $PepT_{St}$ (unlike the H7 glutamate/aspartate, this glutamate binds the N-terminal amino grup of AlaPhe, suggesting it is not protonated), but neither of them contacts alafosfalin in the GkPOT or $PepT_{So2}$ structures. However, transport assays show alafosfalin to be a weak substrate (and there is no evidence of its transport by GkPOT), so the contacts made by the molecule in the two structures may not be an accurate representation of how transported dipeptides are normally accommodated.

Concerning the structural changes in the C-terminal domain, results are less open to interpretation. The occluded-inward-open transition is accompanied by the rotation of the cytoplasmic halves of H10 and H11 around a conserved hinge region (a glycine and a tryptophan), and their movement is coordinated with the straightening of H7 (around another conserved, essential tryptophan). Similarly, the rearrangement of H8 occurs at conserved proline residues, allowing an adjacent asparagine to rotate downwards from the extracellular cavity into the binding site and contact Glu400 on H10 as the intracellular cavity opens. The role of the asparagine itself is unclear, as it is essential in $PepT_{St/So}$ but not required in GkPOT; it appears to be involved in positioning the substrate in all co-crystal structures, albeit without contacting it through hydrogen bonds or electrostatics.

Overall, the structures and the functional data combined show that the two main processes of the coupled transport mechanism (substrate translocation and coupling) are interlinked, with many of the binding site residues involved in both, while higher order conformational rearrangements facilitate alternating access as well as coupling. To achieve this complex level of coordination, conserved residues are found at strategic positions throughout the proteins, making the alternating access mechanism more than a series of rigid body motions described by the N- and C-terminal helix bundles. Instead, the transport cycle appears to be a complex process dependent on intra- and interhelical salt bridge networks, helix rotations and a tightly regulated gating mechanism. A tentative summary of the transport mechanism (dipeptide/H^+ symport) is presented below for $PepT_{St}$ (fig. 10-1).

- In the outward-open state (A), the transporter is easily protonated due to the high proton concentration in the extracellular environment. The conformation, stabilised by the Lys126-Glu400 salt bridge that closes the intracellular cavity, is energetically favoured.
- The C-terminal carboxylate of the dipeptide binds with high affinity to the protonated side chain of Glu300 (B), causing helices H7 and H8 to kink at pivot positions (Trp296 and Pro337 respectively).
- As the extracellular gate closes (H7 and H8 packing against H1/H2 on the opposite side), the peptide is bound by the protonated Glu300 and Arg33; Asn328 and Tyr30 may also bind the peptide at this stage (C).
- With the extracellular gate fully closed, the transporter is in the high-energy occluded conformation (D). The central cavity is sealed at the top by the Arg33-Glu300 salt bridge, while the proton is passed on to the side chain of Glu299, which binds the carboxyl end of the peptide, with additional contacts from the binding site tyrosines. Meanwhile, the cytoplasmic half of H11 starts rotating sideways at Trp427 to open the intracellular gate (as does H10 at Gly407, not pictured); the H10/H11 movement is coordinated with the straightening of H7 (as Trp296 rotates away from the central cavity), and H8 also rotates (at Pro329).

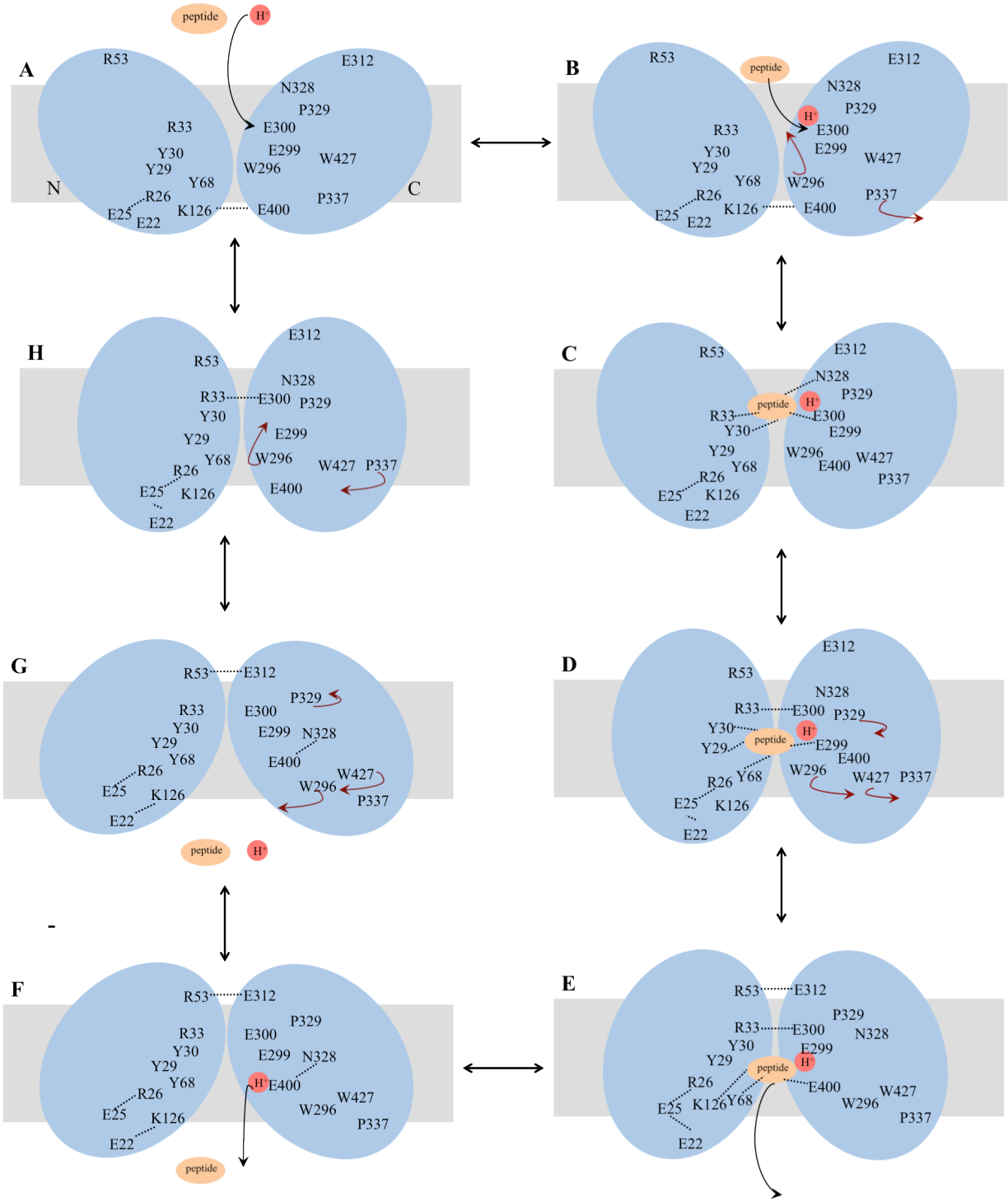


Fig. 10-1: Proposed mechanism for $PepT_{St}$-mediated dipeptide/H^+ symport. Key residues are labelled, and polar contacts are depicted as dashed lines. The movements of C-terminal helices are marked by red arrows, originating at the residues that facilitate the movements (e.g. an arrow at W296 indicates the rotation of H7).

- As the intracellular gate opens and the Arg53-Glu312 salt bridge forms at the extracellular side, the transporter is in a substrate-bound, inward-occluded state (E). Meanwhile, the peptide is reoriented: the negatively charged Glu400 side chain binds the N-terminal amino group, while Lys126 and Tyr68 bind the C-terminal carboxylate and side chain respectively.
- With the intracellular gate fully open, the peptide is released into the cytoplasm, Glu400 is protonated, and the transporter is in the low-energy inward-open state (F).
- Glu400 is easily deprotonated in the intracellular environment (G), and the empty transporter cycles back through an occluded state (H) to the initial outward-open conformation (A).

Some aspects of the transport mechanism are not yet understood. For example, the disruption of the central salt bridge (Arg33-Glu300) is proposed in the GkPOT study (Doki *et al*, 2013), which suggests that intracellular deprotonation occurs at the Glu300 side chain before the transporter transits back into the apo occluded state. Indeed, the protonated Glu300 may bind the carboxylate of the substrate after the initial protonation event at the extracellular side, but both the $PepT_{St}$-AlaPhe and $PepT_{So2}$-alafosfalin structures seem to suggest that a negatively charged Glu400 binds the N-terminal amino group, perhaps at a later stage in the mechanism, as depicted here. Furthermore, forming the central salt bridge in the occluded and inward-open states, as seen in the structures, would require that Glu300 be deprotonated before opening the intracellular gate, thus making it unlikely that it would have retained the proton after the release of the peptide. It is therefore likely that Glu299 and Glu400 help to translocate the proton into the cytoplasm after initial protonation at Glu300. This hypothesis is supported by the very similar mutagenesis data on Glu300 and Glu400 (the involvement of Glu299 in proton translocation is only speculative, as none of the Glu299 mutants expressed and thus no transport data is available). The model remains to be confirmed through other substrate-bound structures and binding data on specific residues.

While the mechanism is proposed here for $PepT_{St}$, it is possible that other transporters have different ways of coordinating dipeptide substrates. Moreover, tripeptides and drugs are likely to form different interactions compared to dipeptides. As discussed in 6.1.1, His61 is

most likely involved in proton binding in $PepT_{So}$, and a potentially different translocation path was suggested in the GkPOT discussion (8.2.2). Lastly, the model assumes that one proton is translocated with the peptide, but the stoichiometry of the transport process may change with the protein and/or the protonation state of the substrate. The role of the ExxERFxYY motif on H1 in proton coupling is difficult to assess, but it is conceivable that an additional proton (or more) is translocated through the hydrogen bond network on H1 (involving Arg26, Glu25 and Glu22), resulting in a stoichiometry of 2:1 (proton:peptide) or higher. This would also control the position of Lys126 (and the formation of the Lys126-Glu400 salt bridge), as suggested by the Lys126-Glu22 interaction observed in the inward-open structures. More data is needed to elucidate how the H1 residues are involved in translocation and to determine the stoichiometry of the transport cycle.

From an evolutionary standpoint, there may be a causal relationship between the complexity of the mechanism and the diversity of transported substrates. In this view, the sequence variability at certain positions in the binding site (like the presence of the histidine and the different side chains at Glu299/Phe315) may have come about as an adaptation/survival strategy in a novel or changing environment, characterised by the abundance or scarcity of particular substrates. An example appears to be the limited competitive effect of drug molecules on bacterial POT transport, despite them being transported by the human proteins. In this case, bacterial members are likely to have adapted to resist the detrimental intracellular accumulation of, for instance, β-lactam antibiotics. In some transporters such as $PepT_{St}$, the low affinity for the drugs may be connected to the lack of uptake of large tripeptides, as tripeptides and cephalosporin antibiotics are likely recognised and transported via a similar pathway.

The results presented here outline the importance of a combined approach to the study of mechanism, using structural, biochemical, biophysical and computational tools to gain a better understanding of the transport cycle. Aside from the main assay work, valuable input was received from modelling, molecular dynamics simulations and DEER measurements, while NMR spectroscopy provided a novel approach to peptide binding studies.

Far from exhausting the addressed topics, the results of this study lay the foundation for future work on the POT family transporters. On the structural front, there is still much to be done in terms of screening for homologues and capturing proteins in different conformations. For example, in light of the findings in chapter 6, cysteine mutations could be used to introduce disulfide crosslinks between residues predicted to form salt bridges in the outward-open state, thereby trapping the transporter in that conformation, as previously achieved for the inward-facing structure of Glt_{Ph} (Reyes *et al*, 2009). Otherwise, the outward-open state could be stabilised by co-crystallisation with a native substrate, as seen in the case of the first XylE structure (Sun *et al*, 2012).

Alternatively, the use of lipid environments for cystallisation may have the potential to favour alternative conformational states, as indicated by a recent study on LacY which shows the transporter to preferentially adopt the outward-open conformation in membranes (Jiang *et al*, 2012). In particular, *in meso* crystallisation (lipidic cubic phase) has provided many examples of high resolution membrane protein structures in recent years (Caffrey *et al*, 2012), and it has recently started to be successful in crystallising transporters as well, GkPOT and the recently solved co-crystal structures of $PepT_{St}$ with peptides being some of the more recent examples. Another lipid-based approach is bicelle crystallisation, which has produced a novel structure for bacteriorhodopsin (Faham and Bowie, 2002) and a substrate-bound structure of LeuT (Wang *et al*, 2012). Thus, screening POT homologues for LCP and bicelle crystallisation could identify new targets for functional studies and provide more information on binding and the structural dynamics of the transport cycle.

The preliminary NMR data on POT proteins presented in chapter 9 points out some of the limitations of the method (protein and micelle size, instability at high temperatures), but also identifies GkPOT as a good target for solution binding studies due to its atypical thermostability. The next step in the GkPOT NMR project is to collect 2D spectra of ^{13}C-labelled samples, which would allow to scan for specific functional groups in the protein like the aromatic rings, thus potentially singling out chemical shifts contributed by the binding site tyrosines upon peptide binding; specific substrate-interacting residues could then be identified through mutating potential candidates to non-aromatic amino acids and collecting new spectra.

The arginine peak seen changing its position in the ^{1}H-^{15}N spectra could be identified in a similar way. Furthermore, better resolved spectra could be collected from protein grown in minimal media supplemented with labelled amino acids (McIntosh and Dahlquist, 1990); this would allow investigating the effect of binding on specific side chains, further narrowing down the residues potentially involved in binding. As GkPOT is stable and can be purified in high amounts, the approach is feasible, albeit time consuming. In addition, the peptide titration experiment can be repeated with different peptides (for example, a control experiment using a low affinity substrate like GluGlu is needed to confirm that AlaPhe-induced chemical shifts are caused by a specific binding event). Alternatively, experiments with bicelle-reconstituted protein may show better resolution (Morrison *et al*, 2012) and would make it possible to examine the influence of the lipid bilayer on substrate binding.

Finally, binding studies (as well as co-crystallisation) with pharmaceutically important substrates would provide valuable information for clinical applications. The drug binding study (chapter 7) is still a work in progress, and would benefit from a bacterial model that transports PepT1/PepT2 drug substrates with high affinity. Mapping specific transporter-drug interactions onto the structures of oligopeptide transporters and explaining them in the context of the transport mechanism would be the first major step towards designing other drugs and inhibitors targeted to POT proteins, a process that, in the absence of structures, has been restricted to mostly empirical deductions based on the available functional data on the mammalian transporters (Ekins *et al*, 2005). Naturally, obtaining crystal structures of eukaryotic POT members would advance the field considerably, and setting up an *in vitro* system to characterise their function would be a useful addition to the existing *in vivo* protocols. Thus, an effort needs to be made to extend the knowledge gained from studying bacterial POT proteins to the structural and functional study of PepT1 and PepT2.

APPENDIX

$PepT_{St}$ mutation	Crystals	Diffraction
M96C/V256C	yes	3.6-3.5Å
S225C		3.6-3.5Å
S130C/M424C		Weak (8-9Å)
A307C	no	-
S353C		
S449C		

Table S1: Cysteine mutagenesis of $PepT_{St}$ for mercury derivatisation. Proteins were purified in DDM and derivatised by incubaton with 2.5mM methylmercury chloride, before crystallisation at 19°C in the standard condition (0.1M MES pH 6.5, 0.03M $MgCl_2$, 1mM $CdCl_2$, 26% PEG400). Diffraction data collected at Diamond Light Source and the ESRF.

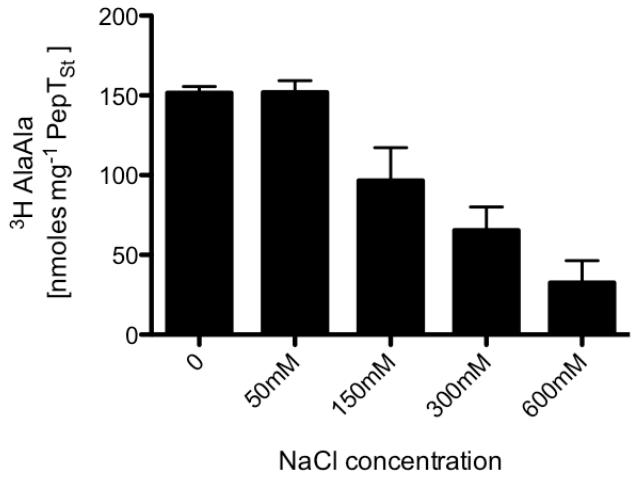


Fig. S1: Effect of NaCl concentration on $PepT_{St}$-mediated AlaAla uptake. Proteoliposomes in internal buffer (20 mM KPi, 100 mM KAc and 2 mM $MgSO_4$) were diluted 1:50 in external buffer (120 mM NaPIPES and 2 mM $MgSO_4$) containing 10 μM valinomycin, 15 μM ^{3}H-AlaAla and increasing concentrations of NaCl. Uptake was recorded after 1 min at pH 6.5 and 25°C.

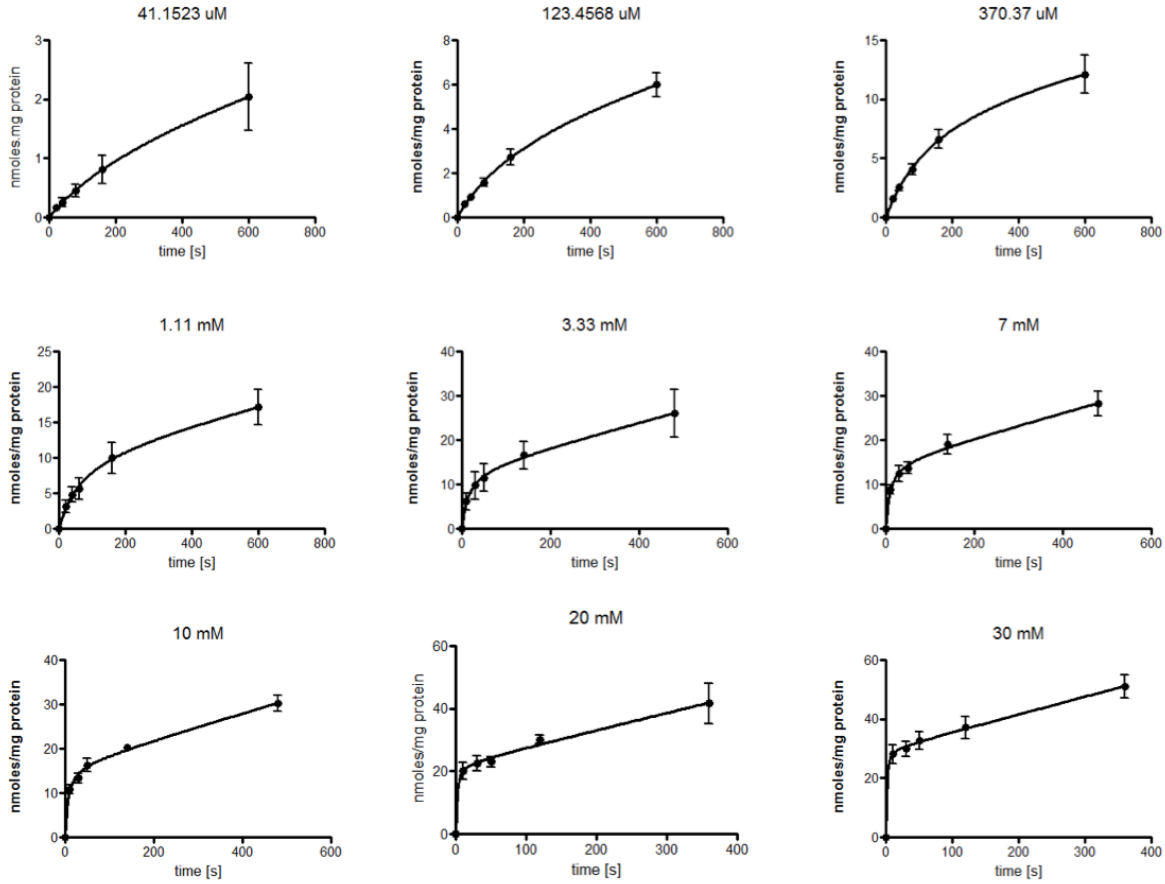


Fig. S2: Kinetic analysis of ΔΨ-driven ^{3}H-GlySar uptake by $PepT_{St}$, raw data. Y-axis values were converted from CPM measurements on a scintillation counter, according to a ^{3}H-GlySar standard curve. Uptake of substrate at different concentrations was monitored over time, using a fixed amount of ^{3}H-labelled GlySar supplemented with unlabelled GlySar to the desired concentration.

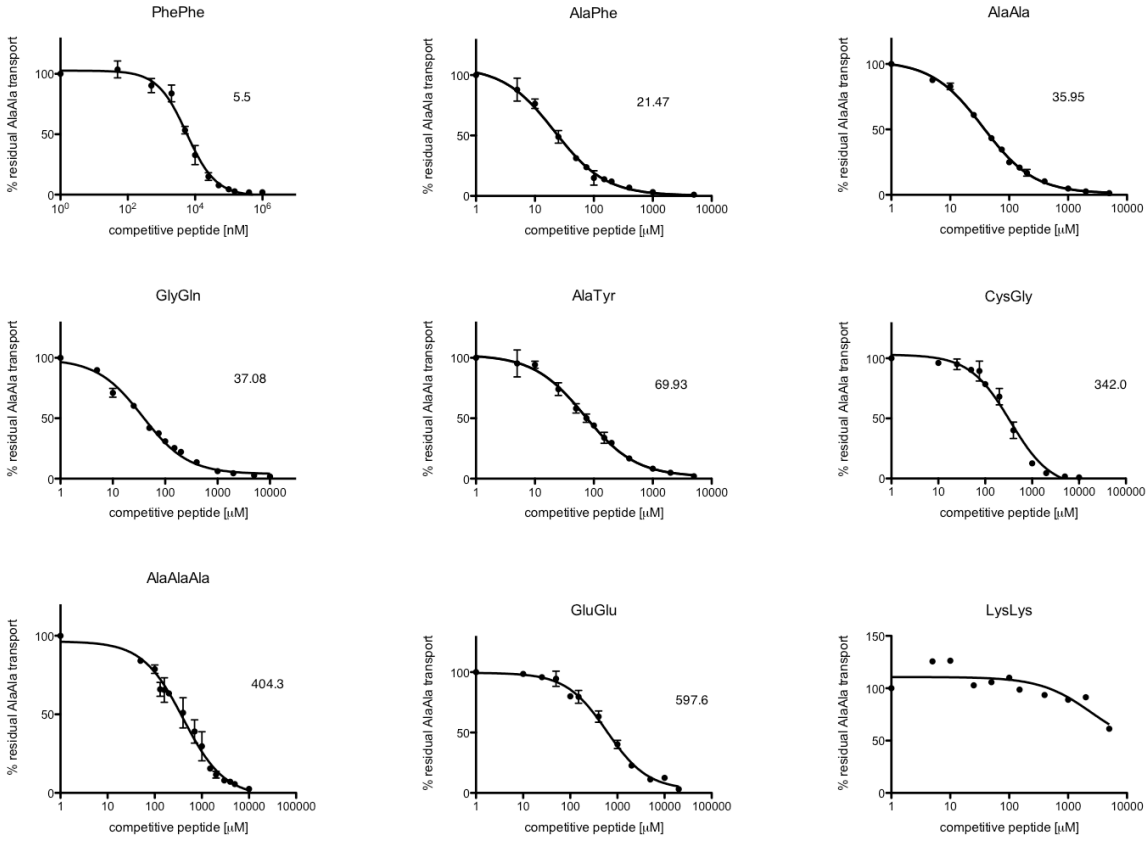


Fig. S3: IC_{50} curves measuring competition by di- and tripeptides against ^{3}H-AlaAla uptake by $PepT_{St}$, raw data. Residual uptake of the reporter (Y-axis) is shown in percentages (relative to the uptake level without competing peptides). Competitor concentrations (X-axis) are shown in logarithmic scale. Measurements were taken after 2 min incubation in ΔΨ-driven conditions. IC_{50} values are listed beside each graph in μM.

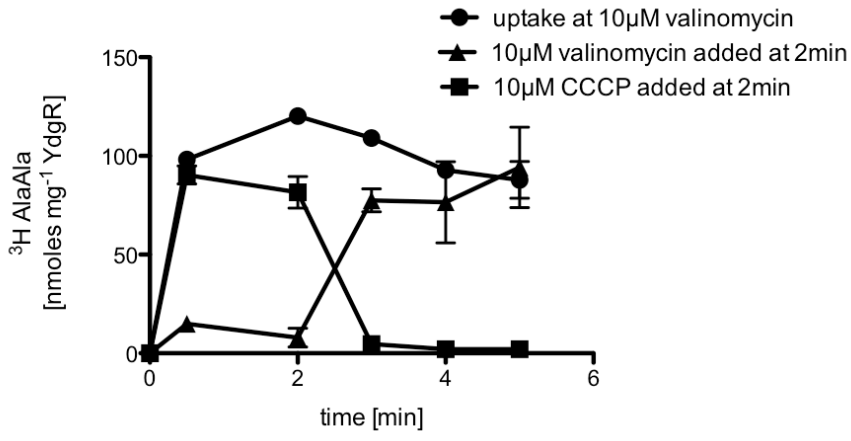


Fig. S4: Time-dependent uptake of ^{3}H-AlaAla by liposome-reconstituted YdgR at pH 6.5 and 25°C. Proteoliposomes containing internal buffer (20 mM KPi, 100 mM KAc and 2 mM $MgSO_4$) were diluted 1:50 in external buffer (120 mM NaPIPES and 2 mM $MgSO_4$) containing 15 μM substrate and 10 μM valinomycin (circle). A second experiment was performed in the same conditions with the addition of proton ionophore CCCP after 2 min (square). A third experiment was started without valinomycin in the external buffer, causing a "slow start" after its addition at the 2 min time point, showing that transport is dependent on the ΔΨ.

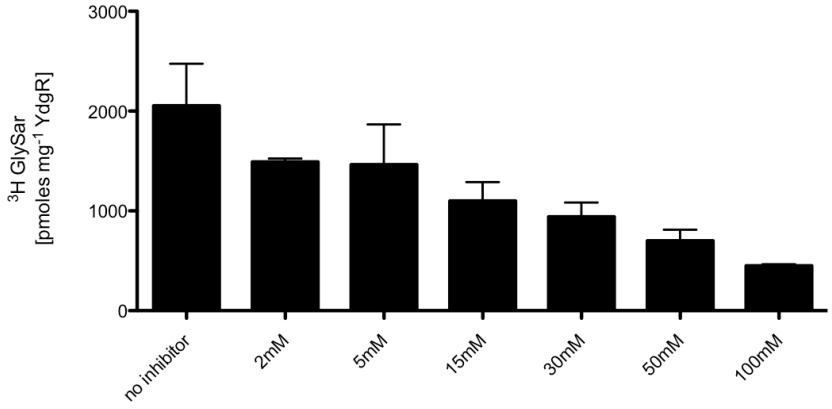


Fig. S5: GlySar competing against its own uptake in liposomes reconstituted with YdgR. ΔΨ-driven uptake of 50 μM ^{3}H-AlaAla by 1 μM protein after 2 min incubation was measured at pH 6.5 and 25°C, in the presence of increasing concentrations of unlabelled GlySar, from 2 to 100 mM.

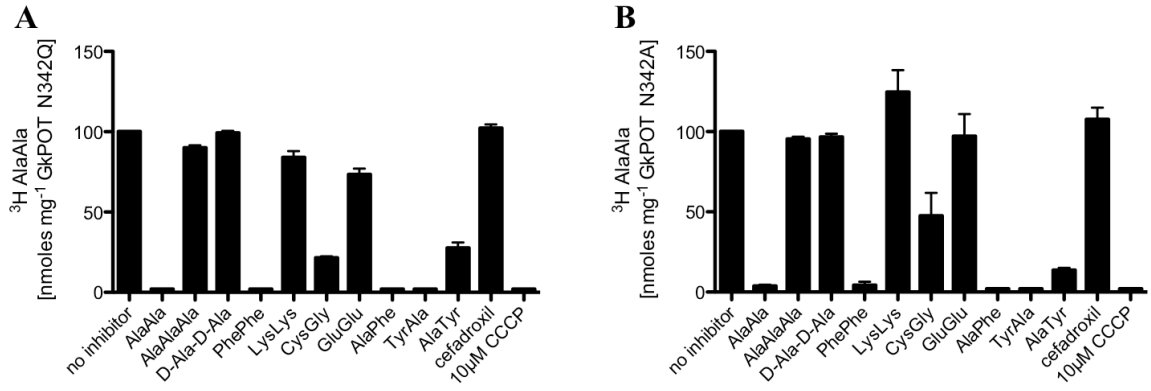


Fig. S6: Competition assays in liposomes reconstituted with GkPOT variants N342Q (A) and N342A (B). ΔΨ-driven uptake of 15 μM ^{3}H-AlaAla by 1 μM protein after 2 min incubation was measured at pH 6.5 and 25°C, in the presence of 2 mM competitor unless specified.

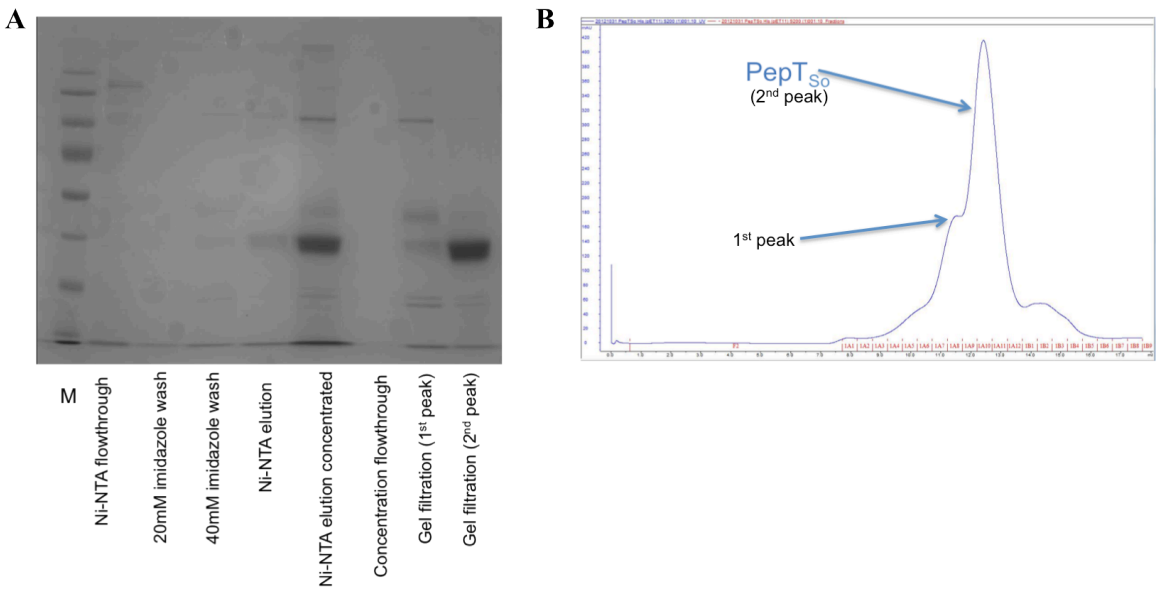


Fig. S7: Purification of the $PepT_{So}$ His-tagged construct for the scintillation proximity assay.
(A) TrisGlycine gel showing the purification steps.
(B) Gel filtration profile. As seen on the gel, the smaller peak on the left is made up of contaminants, while the main peak is pure $PepT_{So}$.

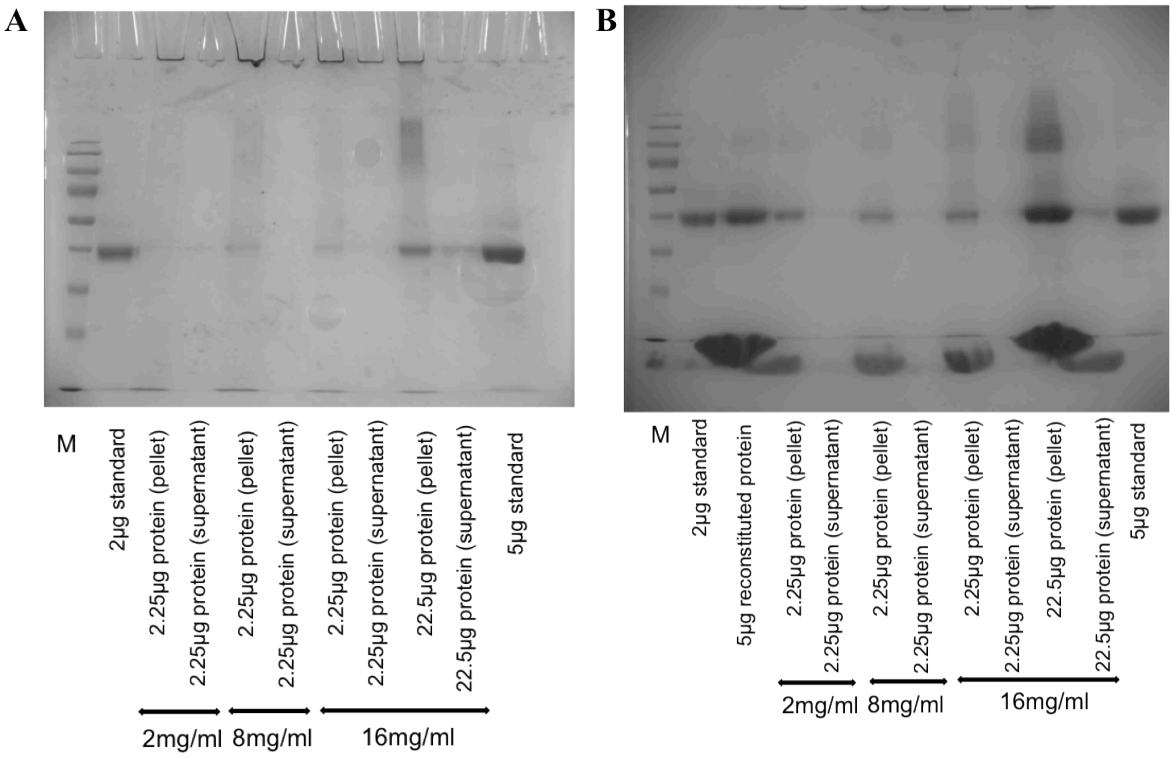


Fig. S8: Checking the efficiency of $PepT_{So}$-His binding to YSi-copper SPA beads in detergent solution (A) and after liposome reconstitution (B). Different protein amounts (2.25 or 22.5 µg) were incubated with different concentrations of beads (2, 8 and 16 mg/ml) for 1h at RT. Pellets containing the beads and immobilised protein were spun down and loaded on a TrisGly gel, along with 10 µl of the 200 µl supernatant and protein standards of known amount.

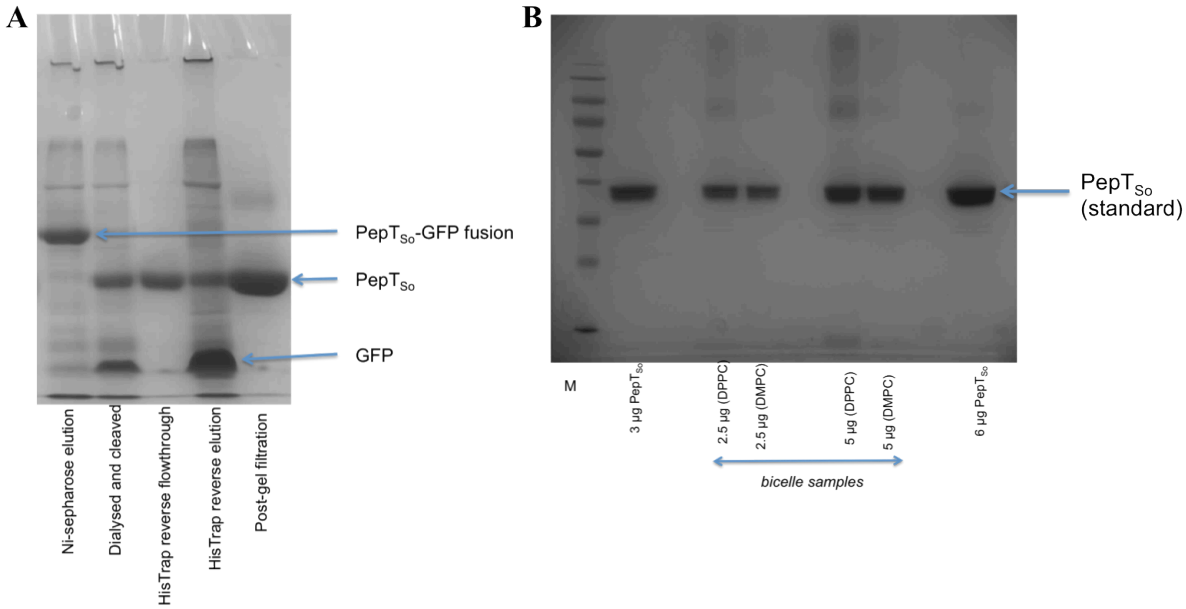


Fig. S9: Purification and bicelle reconstitution of ^{15}N-PepT$_{So}$.
(A) TrisGly gel showing the purification steps.
(B) TrisGly gel showing the efficiency of reconstitution into DPPC and DMPC liposomes, prior to the addition of DHPC to form isotropic bicelles. Different amounts of reconstituted protein (2.5 and 5 µg assuming 100% efficiency) were run on the gel alongside standards of known amount.

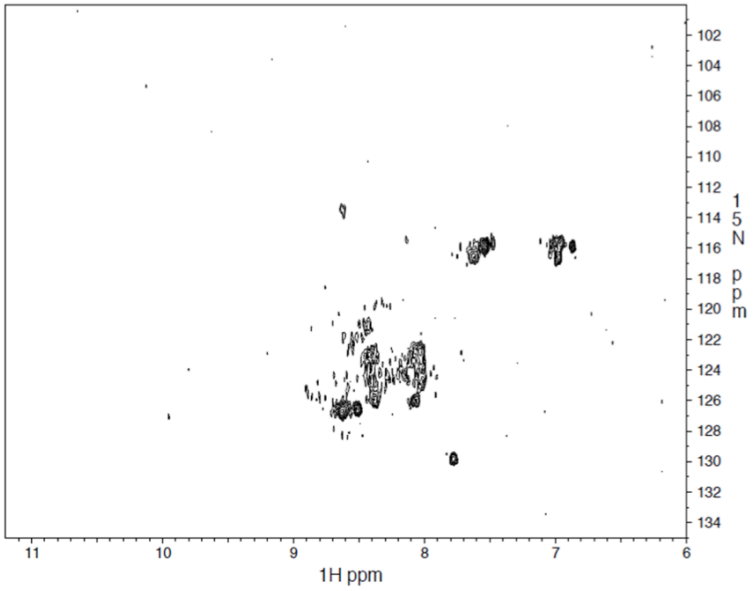


Fig. S10: SoFast-HMQC spectrum of ^{15}H-PepT$_{St}$ (950 MHz high field, 40°C, protein concentration 100 µM).

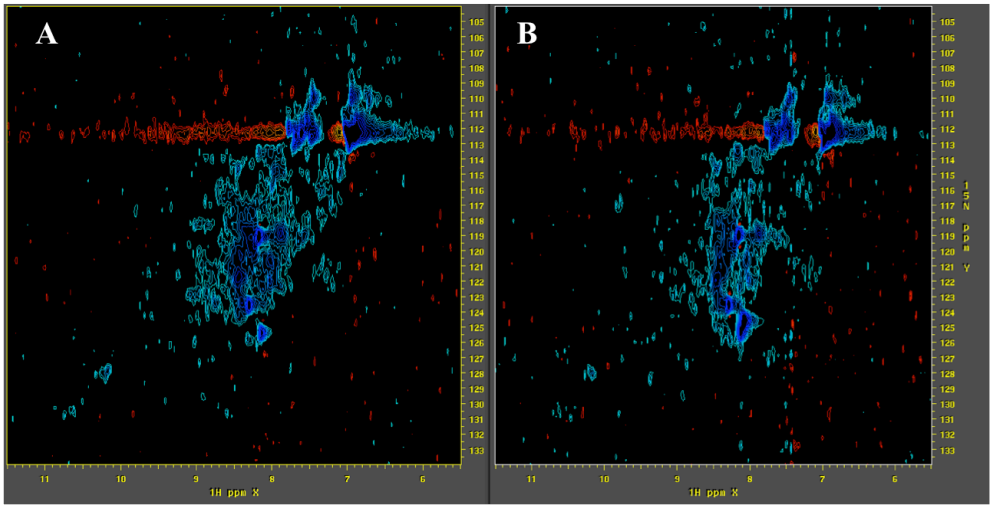


Fig. S11: SoFast-HMQC spectra of ^{15}H-PepT$_{So}$ reconstituted into DMPC/DHPC bicelles (protein:lipid ratio 1:130 molar, Q value 0.3).
(A) Apo protein spectrum collected at 650 MHz, 40°C, protein concentration 100 µM.
(B) Spectrum collected in the same conditions after addition of 80 mM AlaPhe (protein partially degraded).

PUBLICATIONS

- Lyons J, Solcan N, Parker JL, Li D, Brinth A, Caffrey M and Newstead S, „A multi site peptide-binding pocket explains the molecular basis of peptide recognition in the bacterial homologues of PepT1 and PepT2" in preparation;

- Doki S, Kato HE, Solcan N, Iwaki M, Koyama M, Hattori M, Iwase N, Tsukazaki T, Sugita Y, Kandori H, Newstead S, Ishitani R, Nureki O, „Structural basis for dynamic mechanism of proton-coupled symport by the peptide transporter POT" **PNAS,** 2013 Jul 9; 110(28): 11343-8;

- Solcan N, Kwok J, Fowler PW, Cameron AD, Drew D, Iwata So and Newstead S, „Alternating access mechanism in the POT family of oligopeptide transporters" **EMBO Journal,** 2012 Aug 15;31(16): 3411-21.

BIBLIOGRAPHY

1. Abramson, J., Iwata, S. & Kaback, H. R. (2004) Lactose permease as a paradigm for membrane transport proteins (Review). *Mol. Membr. Biol.* **21,** 227–236.

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

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

4. Adibi, S. (1977) Clearance of dipeptides from plasma: role of kidneys and intestine. *Ciba Found Symp.* **50,** 265–285.

5. Alexandrov, A. I., Mileni, M., Chien, E. Y. T., Hanson, M. A. & Stevens, R. C. (2008) Microscale fluorescent thermal stability assay for membrane proteins. *Structure* **16,** 351–359.

6. Aller, S. G., Yu, J., Ward, A., Weng, Y., Chittaboina, S., Zhuo, R., Harrell, P. M., Trinh, Y. T., Zhang, Q., Urbatsch, I. L. & Chang, G. (2009) Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science* **323,** 1718–1722.

7. Amunts, A., Toporik, H., Borovikova, A. & Nelson, N. (2010) Structure determination and improved model of plant photosystem I. *J. Biol. Chem.* **285,** 3478–3486.

8. Anderson, C. M. H. & Thwaites, D. T. (2010) Hijacking solute carriers for proton-coupled drug transport. *Physiology Bethesda* **25,** 364–377.

9. Asatoor, A. M., Bandoh, J. K., Lant, A. F., Milne, M. D. & Navab, F. (1970) Intestinal absorption of carnosine and its constituent amino acids in man. *Gut* **11,** 250–254.

10. Asherie, N. (2004) Protein crystallization and phase diagrams. *Methods* **34,** 266–272.

11. Bailey, P. D., Boyd, C. A., Bronk, J. R., Collier, I. D., Meredith, D., Morgan, K. M. & Temple, C. S. (2000) How to make drugs orally active: a substrate template for peptide transporter PepT1. *Angew. Chem. Int. Ed. Engl.* **39**, 505-508.

12. Baradaran, R., Berrisford, J. M., Minhas, G. S. & Sazanov, L. A. (2013) Crystal structure of the entire respiratory complex I. *Nature* **494,** 443–448.

13. Berrisford, J. M. & Sazanov, L. A. (2009) Structural basis for the mechanism of respiratory complex I. *J. Biol. Chem.* **284,** 29773–29783.

14. Biegel, A., Knütter, I., Hartrodt, B., Gebauer, S., Theis, S., Luckner, P., Kottra, G., Rastetter, M., Zebisch, K., Thondorf, I., Daniel, H., Neubert, K. & Brandsch, M. (2006) The renal type H^+/peptide symporter PEPT2: structure-affinity relationships. *Amino acids* **31,** 137–156.

15. Bill, R. M., Henderson, P. J. F., Iwata, S., Kunji, E. R. S., Michel, H., Neutze, R., Newstead, S., Poolman, B., Tate, C. G. & Vogel, H. (2011) Overcoming barriers to membrane protein structure determination. *Nat. Biotechnol.* **29,** 335–340.

16. Blattner, F. R. (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* **277,** 1453–1462.

17. Bogdanov, M., Heacock, P. N. & Dowhan, W. (2002) A polytopic membrane protein displays a reversible topology dependent on membrane lipid composition. *EMBO J*. **21,** 2107–2116.

18. Bolger, M. B., Haworth, I. S., Yeung, A. K., Ann, D., von Grafenstein, H., Hamm-Alvarez, S., Okamoto, C. T., Kim, K.-J., Basu, S. K., Wu, S. & Lee, V. H. L. (1998) Structure, Function, and Molecular Modeling Approaches to the Study of the Intestinal Dipeptide Transporter PepT1. *J. Pharm. Sci*. **87,** 8–13.

19. Boll, M., Markovich, D., Weber, W.-M., Korte, H., Daniel, H. & Murer, H. (1994) Expression cloning of a cDNA from rabbit small intestine related to proton-coupled transport of peptides, beta-lactam antibiotics and ACE-inhibitors. *Pflugers Arch.: Eur. J. Physiol*. **429,** 146–149.

20. Borbat, P. P., Surendhran, K., Bortolus, M., Zou, P., Freed, J. H. & Mchaourab, H. S. (2007) Conformational motion of the ABC transporter MsbA induced by ATP hydrolysis. *PLoS Biol*. **5,** 2211-2219.

21. Brandsch, M. (2009) Transport of drugs by proton-coupled peptide transporters: pearls and pitfalls. *Expert Opin. Drug Metab. Toxicol*. **5,** 887–905.

22. Bretscher, M.S. (1972) Asymmetrical lipid bilayer structure for biological membranes. *Nat. New Biol*. **236**, 11-12.

23. Bretschneider, B., Brandsch, M. & Neubert, R. (1999) Intestinal transport of β-lactam antibiotics: Analysis of the affinity at the H^+/peptide symporter (PepT1), the uptake into Caco-2 cell monolayers and the transepithelial flux. *Pharm. Res*. **16**, 55-61.

24. Caffrey, M., Li, D. & Dukkipati, A. (2012) Membrane protein structure determination using crystallography and lipidic mesophases: recent advances and successes. *Biochemistry* **51,** 6266–6288.

25. Carpenter, E. P., Beis, K., Cameron, A. D. & Iwata, S. (2008) Overcoming the challenges of membrane protein crystallography. *Curr. Opinion Struct. Biol*. **18,** 581–586.

26. Casagrande, F., Harder, D., Schenk, A., Meury, M., Ucurum, Z., Engel, A., Weitz, D., Daniel, H. & Fotiadis, D. (2009) Projection structure of DtpD (YbgH), a prokaryotic member of the peptide transporter family. *J. Mol. Biol*. **394,** 708–717.

27. Chen, C. C. & Wilson, T. H. (1984) The phospholipid requirement for activity of the lactose carrier of Escherichia coli. *J. Biol. Chem*. **259,** 10150–10158.

28. Chen, X. Z., Steel, A. & Hediger, M. A. (2000) Functional roles of histidine and tyrosine residues in the H^+-peptide transporter PepT1. *Biochem. Biophys. Res. Commun*. **272,** 726–730.

29. Chen, X. Z., Zhu, T., Smith, D. E. & Hediger, M. A. (1999) Stoichiometry and kinetics of the high-affinity H^+-coupled peptide transporter PepT2. *J. Biol. Chem*. **274,** 2773–2779.

30. Cho, Y. & Salton, M. R. J. Y. (1966) Fatty acid composition of bacterial membrane and wall lipids. *Biochim. Biophys. Acta* **116,** 73–79.

31. Claxton, D.P. (2010) Ion/substrate-dependent conformational dynamics of a bacterial homologue of neurotransmitter:sodium symporters. *Nat. Struct. Mol. Biol* **17,** 822–829.

32. Cornish-Bowden, A. (1979) Fundamentals of Enzyme Kinetics. Butterworths, London.

33. Crisman, T. J., Qu, S., Kanner, B. I. & Forrest, L. R. (2009) Inward-facing conformation of glutamate transporters as revealed by their inverted-topology structural repeats. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 20752–20757.

34. Cronan, J. E. (2003) Bacterial membrane lipids: where do we stand? *Annu. Rev. Microbiol.* **57,** 203–224.

35. Curnow, P., Lorch, M., Charalambous, K. & Booth, P. J. (2004) The reconstitution and activity of the small multidrug transporter EmrE is modulated by non-bilayer lipid composition. *J. Mol. Biol.* **343,** 213–222.

36. Dang, S., Sun, L., Huang, Y., Lu, F., Liu, Y., Gong, H., Wang, J. & Yan, N. (2010) Structure of a fucose transporter in an outward-open conformation. *Nature* **467,** 734–738.

37. Daniel, H., Morse, E. L. & Adibi, S. A. (1992) Determinants of substrate affinity for the oligopeptide/H^+ symporter in the renal brush border membrane. *J. Biol. Chem.* **267,** 9565–9573.

38. Daniel, H. (2004) Molecular and integrative physiology of intestinal peptide transport. *Annu. Rev. Physiol.* **66,** 361–384.

39. Daniel, H. & Kottra, G. (2004) The proton oligopeptide cotransporter family SLC15 in physiology and pharmacology. *Pflugers Arch.: Eur. J. Physiol.* **447,** 610–618.

40. Daniel, H., Spanier, B., Kottra, G. & Weitz, D. (2006) From bacteria to man: archaic proton-dependent peptide transporters at work. *Physiology (Bethesda)* **21,** 93–102.

41. Dantzig, A. H. (1997) Oral absorption of b-lactams by intestinal peptide transport proteins. *Adv. Drug Delivery Rev.* **23,** 63–76.

42. Davidson, A. L., Dassa, E., Orelle, C. & Chen, J. (2008) Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol. Mol. Biol. Rev.* **72,** 317–364.

43. Deisenhofer, J., Epp, O., Miki, K., Huber, R. & Michel, H. (1985) Structure of the protein subunits in the photosynthetic reaction centre of Rhodopseudomonas viridis at 3Å resolution. *Nature* **318,** 618-624.

44. Detmers, F. J., Kunji, E. R., Lanfermeijer, F. C., Poolman, B. & Konings, W. N. (1998) Kinetics and specificity of peptide uptake by the oligopeptide transport system of Lactococcus lactis. *Biochemistry* **37,** 16671–16679.

45. Dieck, S. T., Heuer, H., Ehrchen, J., Otto, C. & Bauer, K. (1999) The peptide transporter PepT2 is expressed in rat brain and mediates the accumulation of the fluorescent dipeptide derivative beta-Ala-Lys-Nepsilon-AMCA in astrocytes. *Glia* **25,** 10–20.

46. Doki, S., Kato, H.E., Solcan, N., Iwaki, M., Koyama, M., Hattori, M., Iwase, N., Tsukazaki, T., Sugita, Y., Kandori, H., Newstead, S., Ishitani, R., Nureki, O. (2013) Structural basis for dynamic mechanism of proton-coupled symport by the peptide transporter POT. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 11343-11348.

47. Dougherty, D. A. (2007) Cation-π interactions involving aromatic amino acids. *J. Nutr.* **137**, 1504S-1508S.

48. Döring, F., Dorn, D., Bachfischer, U., Amasheh, S., Herget, M. & Daniel, H. (1996) Functional analysis of a chimeric mammalian peptide transporter derived from the intestinal and renal isoforms. *J. Physiol.* **497,** 773–779.

49. Döring, F., Martini, C., Walter, J. & Daniel, H. (2002) Importance of a small N-terminal region in mammalian peptide transporters for substrate affinity and function. *J. Membr. Biol.* **186,** 55–62.

50. Döring, F., Walter, J., Will, J., Föcking, M., Boll, M., Amasheh, S., Clauss, W. & Daniel, H. (1998) Delta-aminolevulinic acid transport by intestinal and renal peptide transporters and its physiological and clinical implications. *J. Clin. Invest.* **101,** 2761–2767.

51. Drew, D. E., von Heijne, G., Nordlund, P. & de Gier, J. W. (2001) Green fluorescent protein as an indicator to monitor membrane protein overexpression in Escherichia coli. *FEBS Lett.* **507,** 220–224.

52. Drew, D., Lerch, M., Kunji, E., Slotboom, D.-J. & de Gier, J.-W. (2006) Optimization of membrane protein overexpression and purification using GFP fusions. *Nat. Methods* **3,** 303–313.

53. Drew, D., Sjöstrand, D., Nilsson, J., Urbig, T., Chin, C., de Gier, J.-W. & von Heijne, G. (2002) Rapid topology mapping of Escherichia coli inner-membrane proteins by prediction and PhoA/GFP fusion analysis. *Proc. Natl. Acad. Sci. U.S.A.* **99,** 2690–2695.

54. Dringen, R., Hamprecht, B. & Bröer, S. (1998) The peptide transporter PepT2 mediates the uptake of the glutathione precursor CysGly in astroglia-rich primary cultures. *J. Neurochem.* **71,** 388–393.

55. Eddy, E. P., Wood, C., Miller, J., Wilson, G. & Hidalgo, I. J. (1995) A comparison of the affinities of dipeptides and antibiotics for the di- / tripeptide transporter in Caco-2 cells. *Int. J. Pharm.* **115,** 79–86.

56. Edeling, M. A., Byriel, K. A., Jones, A., Raina, S. & Martin, J. L. DsbG (2003) Crystal Diffraction from Low to High Resolution Notes from the Bench. *Bioscience* **11,** 139–145.

57. Efremov, R. G., Baradaran, R. & Sazanov, L. A. (2010) The architecture of respiratory complex I. *Nature* **465,** 441–445.

58. Efremov, R. G. & Sazanov, L. A. (2011) Structure of the membrane domain of respiratory complex I. *Nature* **476,** 414–420.

59. Ekins, S., Johnston, J. S., Bahadduri, P., D'Souza, V. M., Ray, A., Chang, C. & Swaan, P. W. (2005) In vitro and pharmacophore-based discovery of novel hPEPT1 inhibitors. *Pharm. Res.* **22,** 512–517.

60. Elsaesser, C., Brecht, M. & Bittl, R. (2005) Treatment of spin-coupled metal-centres in pulsed electron-electron double-resonance experiments. *Biochem. Soc. Trans.* **33,** 15–19.

61. Ernst, H. A., Pham, A., Hald, H., Kastrup, J. S., Rahman, M. & Mirza, O. (2009) Ligand binding analyses of the putative peptide transporter YjdL from E. coli display a significant selectivity towards dipeptides. *Biochem. Biophys. Res. Commun.* **389,** 112–116.

62. Faham, S. & Bowie, J. U. (2002) Bicelle crystallization: a new method for crystallizing membrane proteins yields a monomeric bacteriorhodopsin structure. *J. Mol. Biol.* **316,** 1–6.

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

64. Fang, G., Friesen, R., Lanfermeijer, F., Hagting, A., Poolman, B. & Konings W. N. (1999) Manipulation of activity and orientation of membrane-reconstituted di-tripeptide transport protein DtpT of Lactococcus lactis. *Mol. Membr. Biol.* **16,** 297–304.

65. Fang, G., Konings, W. N. & Poolman, B. (2000) Kinetics and Substrate Specificity of Membrane-Reconstituted Peptide Transporter DtpT of Lactococcus lactis. (2000). *J. Bacteriol.* **182**, 2530-2535.

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

67. Fann, M., Busch, A. & Maloney, P. C. (2003) Functional Characterization of Cysteine Residues in GlpT, the Glycerol 3-Phosphate Transporter of Escherichia coli. *J. Bacteriol.* **185**, 3863-3870.

68. Faria, T. N., Timoszyk, J. K., Stouch, T. R., Vig, B. S., Landowski, C. P., Amidon, G. L., Weaver, C. D., Wall, D. A. & Smith, R. L. (2004) A novel high-throughput pepT1 transporter assay differentiates between substrates and antagonists. *Mol. Pharm.* **1,** 67–76.

69. Farrell, J. & Rose, A. (1967) Temperature effects on microorganisms. *Annu. Rev. Microbiol.* **21,** 101–120.

70. Fei, Y.-J., Kanai, Y., Nussberger, S., Ganapathy, V., Leibach, F. H., Romero, M. F., Singh, S. K., Boron, W. F. & Hediger, M. A. (1994) Expression cloning of a mammalian proton-coupled oligopeptide transporter. *Nature* **368,** 563-566.

71. Feilmeier, B. J., Iseminger, G., Schroeder, D., Webber, H. & Phillips, G. J. (2000) Green Fluorescent Protein Functions as a Reporter for Protein Localization in *Escherichia coli*. *J. Bacteriol.* **182**, 4068-4076.

72. Fejzo, J., Lepre, C. A., Peng, J. W., Bemis, G. W., Murcko, M. A. & Moore, J. M. (1999) The SHAPES strategy: an NMR-based approach for lead generation in drug discovery. *Chem. Biol.* **6,** 755–769.

73. Finean, J.B. (1969) Biophysical contributions to membrane structure. *Quart. Rev. Biophys.* **2**, 1-23.

74. Forrest, L. R., Krämer, R. & Ziegler, C. (2011) The structural basis of secondary active transport mechanisms. *Biochim. Biophys. Acta* **1807,** 167–188.

75. Forrest, L. R. & Rudnick, G. (2009) The rocking bundle: a mechanism for ion-coupled solute flux by symmetrical transporters. *Physiology* **24,** 377–386.

76. Forrest, L. R., Zhang, Y.-W., Jacobs, M. T., Gesmonde, J., Xie, L., Honig, B. H. & Rudnick, G. (2008) Mechanism for alternating access in neurotransmitter transporters. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 10338–10343.

77. Foucaud, C. & Poolman, B. (1992) Lactose transport system of Streptococcus thermophilus. Functional reconstitution of the protein and characterization of the kinetic mechanism of transport. *J. Biol. Chem.* **267,** 22087–22094.

78. Freire, E., Mayorga, O. L. & Straume, M. (1990) Isothermal titration calorimetry. *Anal. Chem.* Vol. **62,** 950-959.

79. Frillingos, S. & Kaback, H. R. (1996) Cysteine-scanning mutagenesis of helix VI and the flanking hydrophilic domains in the lactose permease of *Escherichia coli*. *Biochemistry* **35**, 5333-5338.

80. Gadsby, D. C. (2010) Ion channels versus ion pumps: the principal difference, in principle. *Nat. Rev. Mol. Cell. Biol.* **10,** 344–352.

81. Ganapathy, M. E., Brandsch, M., Prasad, P. D., Ganapathy, V. & Leibach, F. H. (1995) Differential recognition of beta-lactam antibiotics by intestinal and renal peptide transporters, PEPT 1 and PEPT 2. *J. Biol. Chem.* **270,** 25672–25677.

82. Ganapathy, V. & Leibach, F. H. (1986) Carrier-mediated reabsorption of small peptides in renal proximal tubule. *Am. J. Phys.* 251, 945-953.

83. Ganapathy, V., Mendicino, J. F. & Leibach, F. H. (1981) Transport of glycyl-L-proline into intestinal and renal brush border vesicles from rabbit. *J. Biol. Chem.* **256,** 118–124.

84. Ganapathy & Leibach, F. H. (1985) Is intestinal peptide transport energized by a proton gradient? *Am. J. Phys.* **249,** G153–G160.

85. Gao, X., Zhou, L., Jiao, X., Lu, F., Yan, C., Zeng, X., Wang, J. & Shi, Y. (2010) Mechanism of substrate recognition and transport by an amino acid antiporter. *Nature* **463,** 828–832.

86. Garcia, M. L., Viitanen, P., Foster, D. L. & Kaback, H. R. (1983) Mechanism of lactose translocation in proteoliposomes reconstituted with lac carrier protein purified from *Escherichia coli*. Effect of pH and imposed membrane potential on efflux, exchange, and counterflow. *Biochemistry* **22,** 2524–2531.

87. Gautier, A., Mott, H. R., Bostock, M. J. & Kirkpatrick, J. P. (2010) Structure determination of the seven-helical transmembrane receptor sensory rhodopsin II by solution NMR spectroscopy. *Nat. Struct. Mol. Biol.* **17,** 768–774.

88. Gbaguidi, B., Hakizimana, P., Vandenbussche, G. & Ruysschaert, J.-M. (2007) Conformational changes in a bacterial multidrug transporter are phosphatidylethanolamine-dependent. *Cell. Mol. Life Sci.* **64,** 1571–1582.

89. Geertsma, E. R., Nik Mahmood, N. a B., Schuurman-Wolters, G. K. & Poolman, B. (2008) Membrane reconstitution of ABC transporters and assays of translocator function. *Nat. Protoc.* **3,** 256–266.

90. Georgieva, E. R., Borbat, P. P., Ginter, C., Freed, J. H. & Boudker, O. (2013) Conformational ensemble of the sodium-coupled aspartate transporter. *Nat. Struct. Mol. Biol.* **20**, 215-221.

91. Giacomini, K. M., Huang, S.-M., Tweedie, D. J., Benet, L. Z., Brouwer, K. L. R., Chu, X., Dahlin, A., Evers, R., Fischer, V., Hillgren, K. M., Hoffmaster, K. A., Ishikawa, T., Keppler, D., Kim, R. B., Lee, C. A., Niemi, M., Polli, J. W., Sugiyama, Y., Swaan, P. W., Ware, J. A., Wright, S. H., Yee, S. W., Zamek-Gliszczynski, M. J. & Zhang, L. (2010) Membrane transporters in drug development. *Nat. Rev. Drug Discov.* **9,** 215–236.

92. Glasel, J. A. (1995) Validity of nucleic acid purities monitored by 260/280 absorbance ratios. *Biotechniques* **18**, 62-63.

93. Glover, K. J., Whiles, J. A., Guohua, W., Yu, N., Deems, R., Struppe, J. O., Stark, R. E., Komives, E. A. & Vold, R. R. (2001) Structural evaluation of phospholipid bicelles for solution-state studies of membrane-associated biomolecules. *Biophys. J.* **81,** 2163–2171.

94. Goh, E., Siino, D. F. & Igo, M. M. (2004) The *Escherichia coli* tppB (ydgR) Gene Represents a New Class of OmpR-Regulated Genes. *J. Bacteriol.* **186,** 4019–4024.

95. Goldfine, H. (1984) Bacterial membranes and lipid packing theory. *J. Lipid Res.* **25,** 1501–1507.

96. Goldkorn, T., Rimon, G. & Kaback, H. R. (1983) Topology of the lac carrier protein in the membrane of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **80,** 3322–3326.

97. Gouaux, E. (2009) Review. The molecular logic of sodium-coupled neurotransmitter transporters. *Philos. Trans. R. Soc. London Ser. B* **364,** 149–154.

98. Groeneveld, M. & Slotboom, D.-J. (2007) Rigidity of the subunit interfaces of the trimeric glutamate transporter GltT during translocation. *J. Mol. Biol.* **372,** 565–570.

99. Groneberg, D. A., Döring, F., Theis, S., Nickolaus, M., Fischer, A. & Daniel, H. (2002) Peptide transport in the mammary gland: expression and distribution of PEPT2 mRNA and protein. *Am. J. Phys.* **282,** E1172–E1179.

100. Groneberg, D. A., Nickolaus, M., Springer, J., Döring, F., Daniel, H. & Fischer, A. (2001) Localization of the Peptide Transporter PEPT2 in the lung. *Am. J. Pathol.* **158,** 707–714.

101. Guan, L., Hu, Y. & Kaback, H. R. (2003) Aromatic Stacking in the Sugar Binding Site of the Lactose Permease. *Biochemistry* **126,** 1377–1382.

102. Guan, L. & Kaback, H. R. (2004) Binding affinity of lactose permease is not altered by the H^+ electrochemical gradient. *Proc. Natl. Acad. Sci. U.S.A.* **101,** 12148–12152.

103. Guan, L. & Kaback, H. R. (2006) Lessons from lactose permease. *Annu. Rev. Biophys. Biomol. Struct.* **35,** 67–91.

104. Guan, L., Mirza, O., Verner, G., Iwata, S. & Kaback, H. R. (2007) Structural determination of wild-type lactose permease. *Proc. Natl. Acad. Sci. U.S.A.* **104,** 15294–15298.

105. Guan, L., Smirnova, I. N., Verner, G., Nagamori, S. & Kaback, H. R. (2006) Manipulating phospholipids for crystallization of a membrane transport protein. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 1723–1726.

106. Guettou, F., Quistgaard, E. M., Trésaugues, Moberg, P., Jegerschöld, C., Zhu, L., Jong, A. J. O., Nordlund, P. & Löw, C. Structural insights into substrate recognition in proton-dependent oligopeptide transporters. *EMBO Reports* **14**, 804-810.

107. Guskov, A., Kern, J., Gabdulkhakov, A., Broser, M., Zouni, A. & Saenger, W. (2009) Cyanobacterial photosystem II at 2.9-A resolution and the role of quinones, lipids, channels and chloride. *Nat. Struct. Mol. Biol.* **16,** 334–342.

108. Guzman, L.-M., Belin, D., Carson, M. J. & Beckwith, J. (1995) Tight regulation , modulation , and high-level expression by vectors containing the arabinose PBAD promoter. *J. Bacteriol.* **177**, 4121-4130.

109. Hagting, A., Knol, J., Streutker, M. R., Fang, G., Poolman, B. & Konings, W. N. (1997) Amplified expression, purification and functional reconstitution of the dipeptide and tripeptide transport protein of Lactococcus lactis. *Eur. J. Biochem.* **247,** 581–587.

110. Hanahan, D. & Harbor, C. S. (1983) Studies on Transformation of *Escherichia coli* with Plasmids. *J. Mol. Biol.* **166,** 557–580.

111. Hänelt, I., Wunnicke, D., Bordignon, E., Steinhoff, H.-J. & Slotboom, D. J. (2013) Conformational heterogeneity of the aspartate transporter Glt(Ph). *Nat. Struct. Mol. Biol.* **20**, 210-214.

112. Harder, D., Stolz, J., Casagrande, F., Obrdlik, P., Weitz, D., Fotiadis, D. & Daniel, H. (2008) DtpB (YhiP) and DtpA (TppB, YdgR) are prototypical proton-dependent peptide transporters of *Escherichia coli*. *FEBS J.* **275,** 3290–3298.

113. He, M. M. & Kaback, H. R. (1997) Interaction between residues Glu269 (helix VIII) and His322 (helix X) of the lactose permease of *Escherichia coli* is essential for substrate binding. *Biochemistry* **36,** 13688–13692.

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

115. Hellier, M. D., Holdsworth, C. D., McColl, I. & Perrett, D. (1972) Dipeptide absorption in man. *Gut* **13,** 965–969.

116. Higgins, C. F. & Linton, K. J. (2004) The ATP switch model for ABC transporters. *Nat. Struct. Mol. Biol.* **11,** 918–926.

117. Hirai, T., Heymann, J. A. W., Shi, D., Sarker, R., Maloney, P. & Subramaniam, S. (2002) Three-dimensional structure of a bacterial oxalate transporter. *Nat. Struct. Biol.* **9,** 597–600.

118. Hohl, M., Briand, C., Grütter, M. G. & Seeger, M. A. (2012) Crystal structure of a heterodimeric ABC transporter in its inward-facing conformation. *Nat. Struct. Mol. Biol.* **19,** 395–402.

119. Hong, W.-X., Baker, K. A., Ma, X., Stevens, R. C., Yeager, M. & Zhang, Q. (2010) Design, synthesis, and properties of branch-chained maltoside detergents for stabilization and crystallization of integral membrane proteins: human connexin 26. *Langmuir* **26,** 8690–8696.

120. Huang, Y., Lemieux, M. J., Song, J., Auer, M. & Wang, D.-N. (2003) Structure and mechanism of the glycerol-3-phosphate transporter from *Escherichia coli*. *Science* **301,** 616–660.

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

122. In't Veld, G., Elferink, M. G., Driessen, A. J. & Konings, W. N. (1992) Reconstitution of the leucine transport system of *Lactococcus lactis* into liposomes composed of membrane-spanning lipids from *Sulfolobus acidocaldarius*. *Biochemistry* **31,** 12493–12499.

123. Jamieson, D. J. & Higgins, C. F. (1984) Anaerobic and leucine-dependent expression of a peptide transport gene in Salmonella typhimurium. *J. Bacteriol.* **160,** 131–136.

124. Jappar, D., Wu, S.-P., Hu, Y. & Smith, D. E. (2010) Significance and regional dependency of peptide transporter (PEPT) 1 in the intestinal permeability of glycylsarcosine: in situ single-pass perfusion studies in wild-type and Pept1 knockout mice. *Drug Metab. Dispos.* **38,** 1740–1746.

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

126. Jensen, J. M., Simonsen, F. C., Mastali, A., Hald, H., Lillebro, I., Diness, F., Olsen, L. & Mirza, O. (2012) Biophysical characterization of the proton-coupled oligopeptide transporter YjdL. *Peptides* **38,** 89–93.

127. Jensen, J. M., Ismat, F., Szakonyi, G., Rahman, M. & Mirza, O. (2012) Probing the Putative Active Site of YjdL: An Unusual Proton-Coupled Oligopeptide Transporter from E. coli. *PLoS One* **7,** e47780 (2012).

128. Jessen-Marshall, A. E., Paul, N. J., Brooker, R. J. (1995) The conserved motif, GXXX(D/E)(R/K)XG[X](R/K)(R/K), in hydrophilic loop 2/3 of the lactose permease. *J. Biol. Chem.* **270**, 16351-16257.

129. Jiang, X., Guan, L., Zhou, Y., Hong, W.-X., Zhang, Q. & Kaback, H. R. (2012) Evidence for an intermediate conformational state of LacY. *Proc. Natl. Acad. Sci. U.S.A.* **109,** E698–704.

130. Jordan, K. N., Oxford, L. & O'Byrne, C. P. (1999) Survival of low-pH stress by Escherichia coli O157:H7: correlation between alterations in the cell envelope and increased acid tolerance. *Appl. Environ. Microbiol.* **65,** 3048–3055.

131. Kaback, H. R., Sahin-Tóth, M. & Weinglass, A. B. (2001) The kamikaze approach to membrane transport. *Nat. Rev. Mol. Cell. Biol.* **2,** 610–620.

132. Kaback, H. R., Dunten, R., Frillingos, S., Venkatesan, P., Kwaw, I., Zhang, W. & Ermolova, N. (2007) Site-directed alkylation and the alternating access model for LacY. *Proc. Natl. Acad. Sci. U.S.A.* **104,** 491–494.

133. Kaback, H. R., Smirnova, I., Kasho, V., Nie, Y. & Zhou, Y. (2011) The alternating access transport mechanism in LacY. *J. Membr. Biol.* **239,** 85–93.

134. Kato, M., Maegawa, H., Okano, T., Inui, K. & Hori, R. (1989) Effect of various chemical modifiers on H^+ coupled transport of cephradine via dipeptide carriers in rabbit intestinal brush-border membranes: role of histidine residues. *J. Pharmacol. Exp. Ther.* **251,** 745–749.

135. Kawate, T. & Gouaux, E. (2006) Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins. *Structure* **14,** 673–681.

136. Kennedy, D. J., Leibach, F. H., Ganapathy, V. & Thwaites, D. T. (2002) Optimal absorptive transport of the dipeptide glycylsarcosine is dependent on functional Na^+/H^+ exchange activity. *Pflugers Arch.: Eur. J. Phys.* **445,** 139–146.

137. Kim, H. J., Howell, S. C., van Horn, W. D., Jeon, Y. H. & Sanders, C. R. (2011) Recent advances in the application of solution NMR spectroscopy to multi-span integral membrane proteins. *Prog. Nucl. Magn. Reson. Spectrosc.* **55**, 335-360.

138. Knol, J., Veenhoff, L., Liang, W.-J., Henderson, P. J. F., Leblanc, G. & Poolman, B. (1996) Unidirectional reconstitution into detergent-destabilized liposomes of the purified lactose transport system of Streptococcus thermophilus. *J. Biol. Chem.* **271,** 15358–15366.

139. Knütter, I., Hartrodt, B., Toth, G., Keresztes, A., Kottra, G., Mrestani-Klaus, C., Born, I., Daniel, H., Neubert, K. & Brandsch, M. (2007) Synthesis and characterization of a new and radiolabeled high-affinity substrate for H+/peptide cotransporters. *FEBS J.* **274,** 5905–5914.

140. Kottra, G. & Daniel, H. (2001) Bidirectional electrogenic transport of peptides by the proton-coupled carrier PEPT1 in Xenopus laevis oocytes: its asymmetry and symmetry. *J. Phys.* **536,** 495–503.

141. Kottra, G., Stamfort, A. & Daniel, H. (2002) PEPT1 as a paradigm for membrane carriers that mediate electrogenic bidirectional transport of anionic, cationic, and neutral substrates. *J. Biol. Chem.* **277,** 32683–32691.

142. 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., Vazquez-Ibar, J. L., Palacin, M. (2011) Molecular basis of substrate-induced permeation by an amino acid antiporter. *Proc. Natl. Acad. Sci. U.S.A.* **108,** 3935–3940.

143. Krishnamurthy, H. & Gouaux, E. (2012) X-ray structures of LeuT in substrate-free outward-open and apo inward-open states. *Nature* **481,** 469–474.

144. Krishnamurthy, H., Piscitelli, C. L. & Gouaux, E. (2009) Unlocking the molecular secrets of sodium-coupled transporters. *Nature* **459,** 347–355.

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

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

147. Kulkarni, A. A., Haworth, I. S., Uchiyama, T. & Lee, V. H. L. (2003) Analysis of transmembrane segment 7 of the dipeptide transporter hPepT1 by cysteine-scanning mutagenesis. *J. Biol. Chem.* **278,** 51833–51840.

148. Kunji, E. R. S., Smid, E. J., Plapp, R., Poolman, B. & Konings, W. I. (1993) Di-Tripeptides and Oligopeptides Are Taken Up via Distinct Transport Mechanisms in *Lactococcus lactis*. *J. Bacteriol.* **175,** 2052-2059.

149. Law, C. J., Almqvist, J., Bernstein, A., Goetz, R. M., Huang, Y., Soudant, C., Laaksonen, A., Hovmöller, S. & Wang, D.-N. (2008) Salt-bridge dynamics control substrate-induced conformational change in the membrane transporter GlpT. *J. Mol. Biol.* **378,** 828–839.

150. Law, C. J., Maloney, P. C. & Wang, D.-N. (2008) Ins and outs of major facilitator superfamily antiporters. *Annu. Rev. Microbiol.* **62,** 289–305.

151. Lemieux, M. J., Huang, Y. & Wang, D.-N. (2004) Glycerol-3-phosphate transporter of *Escherichia coli*: structure, function and regulation. *Res. Microbiol.* **155,** 623–629.

152. Lemieux, M. J., Song, J., Kim, M. J., Huang, Y., Villa, A., Auer, M., Li, X.-D. & Wang, D.-N. (2003) Three-dimensional crystallization of the Escherichia coli glycerol-3-phosphate transporter : A member of the major facilitator superfamily. *Protein Sci.* **12**, 2748–2756.

153. Liang, R., Fei, Y.-J., Prasad, P. D., Ramamoorthy, S., Han, H., Yang-Feng, T. L., Hediger, M. A., Ganapathy, V. & Leibach, F. H. (1995) Human intestinal H^+/peptide cotransporter: cloning, functional expression, and chromosomal localization. *J. Biol. Chem.* **270**, 6456-6463.

154. Lin, H. & King, N. (2007) Demonstration of functional dipeptide transport with expression of PEPT2 in guinea pig cardiomyocytes. *Pflugers Arch.: Eur. J. Physiol.* **453,** 915–922.

155. Liu, W., Liang, R., Ramamoorthy, S., Fei, Y.-J., Ganapathy, M. E., Hediger, M. A., Ganapathy, V. & Leibach, F. H. (1995) Molecular cloning of PepT2, a new member of the H^+/peptide cotransporter family, from human kidney *Biochim. Biophys. Acta* **1235,** 461–466.

156. Lolkema, J. S., Carrasco, N. & Kaback, H. R. (1991) Kinetic analysis of lactose exchange in proteoliposomes reconstituted with purified lac permease. *Biochemistry* **30,** 1284–1290.

157. Lubkowitz, M. (2011) The Oligopeptide Transporters: A Small Gene Family with a Diverse Group of Substrates and Functions? *Mol. Plant* 1–9.

158. Luckner, P. & Brandsch, M. (2005) Interaction of 31 beta-lactam antibiotics with the H+/peptide symporter PEPT2: analysis of affinity constants and comparison with PEPT1. *Eur. J. Pharm. Biopharm.* **59,** 17–24.

159. Luli, G. W. & Strohl, W. R. (1990) Comparison of Growth , Acetate Production , and Acetate Inhibition of *Escherichia coli* Strains in Batch and Fed-Batch Fermentations. *Appl. Environ. Microbiol.* **56**, 1004-1011.

160. Lyons, J., Solcan, N., Parker, J.L., Li, D., Brinth, A., Caffrey, M. and Newstead, S. (2014) A multi site peptide-binding pocket explains the molecular basis of peptide recognition in the bacterial homologues of PepT1 and PepT2 (*in preparation*).

161. Mackenzie, B., Loo, D. D. F., Fei, Y.-J., Liu, W., Ganapathy, V., Leibach, F. H. & Wright, E. M. (1996) Mechanisms of the Human Intestinal H^+-coupled Oligopeptide Transporter hPEPT1. *J. Biol. Chem.* **271,** 5430–5437.

162. Madej, M. G., Soro, S. N. & Kaback, H. R. (2012) Apo-intermediate in the transport cycle of lactose permease. **109**, E2970–E2978.

163. Majumdar, D. S., Smirnova, I., Kasho, V., Nir, E., Kong, X., Weiss, S. & Kaback, H. R. (2007) Single-molecule FRET reveals sugar-induced conformational dynamics in LacY. *Proc. Natl. Acad. Sci. U.S.A.* **104,** 12640–12645.

164. Malle, E., Zhou, H., Neuhold, J., Spitzenberger, B., Klepsch, F., Pollak, T., Bergner, O., Ecker, G. F. & Stolt-Bergner, P. C. (2011) Random mutagenesis of the prokaryotic peptide transporter YDGR identifies potential periplasmic gating residues. *J. Biol. Chem.* **286**, 23121-23131.

165. Maynard, J. A., Linquist, N. C., Sutherland, J. N., Lesuffleur, A., Warrington, A. E., Rodriguez, M. & Oh, S.-H. (2010) Next generation of SPR technology of membrane-bound proteins for ligand screening and biomarker discovery. *Biotechnol. J.* **4**, 1542–1558.

166. McDermott, A. (2009) Structure and dynamics of membrane proteins by magic angle spinning solid-state NMR. *Annu. Rev. Biophys.* **38,** 385–403.

167. McHaourab, H. S., Steed, P. R. & Kazmier, K. (2011) Toward the fourth dimension of membrane protein structure: insight into dynamics from spin-labeling EPR spectroscopy. *Structure* **19,** 1549–1561.

168. McIntosh, L. P. & Dahlquist, F. W. (1990) Biosynthetic incorporation of ^{15}N and ^{13}C for assignment and interpretation of nuclear magnetic resonance spectra of proteins. *Q. Rev. Biophys.* **23,** 1–38.

169. Meredith, D. & Price, R. A. (2006) Molecular modeling of PepT1--towards a structure. *J. Membr. Biol.* **213,** 79–88.

170. Meredith, D., Temple, C. S., Guha, N., Sword, C. J., Boyd, C. A. R., Collier, I. D., Morgan, K. M. & Bailey, P. D. (2000) Modified amino acids and peptides as substrates for the intestinal peptide transporter PepT1. *Eur. J. Biochem. FEBS* **267,** 3723–3728.

171. Miller, C. (2006) ClC chloride channels viewed through a transporter lens. *Nature* **440,** 484–489.

172. Miller, S., Edwards, M. D., Ozdemir, C. & Booth, I. R. (2003) The closed structure of the MscS mechanosensitive channel. Cross-linking of single cysteine mutants. *J. Biol. Chem.* **278,** 32246–32250.

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

174. Mitchell, P. (1961) Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature* **191**, 144-148.

175. Mitsuoka, K., Miyoshi, S., Kato, Y., Murakami, Y., Utsumi, R., Kubo, Y., Noda, A., Nakamura, Y., Nishimura, S. & Tsuji, A. (2008) Cancer detection using a PET tracer, 11C-glycylsarcosine, targeted to H+/peptide transporter. *J. Nucl. Med.* **49,** 615–622.

176. Morrison, E. A., DeKoster, G. T., Dutta, S., Vafabakhsh, R., Clarkson, M. W., Bahl, A., Kern, D., Ha, T. & Henzler-Wildman, K. A. (2012) Antiparallel EmrE exports drugs by exchanging between asymmetric structures. *Nature* **481,** 45–50.

177. Morrison, E. A. & Henzler-Wildman, K. A. (2012) Reconstitution of integral membrane proteins into isotropic bicelles with improved sample stability and expanded lipid composition profile. *Biochim. Biophys. Acta* **1818,** 814–820.

178. Morth, J. P., Pedersen, B. P., Buch-Pedersen, M. J., Andersen, J. P., Vilsen, B., Palmgren, M. G. & Nissen, P. (2011) A structural overview of the plasma membrane Na+,K+-ATPase and H+-ATPase ion pumps. *Nat. Rev. Mol. Cell. Biol.* **12,** 60–70.

179. Mueckler, M. & Makepeace, C. (2010) Model of the exofacial substrate-binding site and helical folding of the human Glut1 glucose transporter based on scanning mutagenesis. *Biochemistry* **48,** 5934–5942.

180. Muiry, J. A. R., Gunn, T. C., McDonald, T. P., Bradley, S. A., Tate, C. G. & Henderson, P. J. F. (1993) Proton-linked L-rhamnose transport, and its comparison with L-fucose transport in Enteriobacteriaceae. *Biochem. J.* **290**, 833-842.

181. Mulligan, C., Geertsma, E. R., Severi, E., Kelly, D. J., Poolman, B. & Thomas, G. H. (2009) The substrate-binding protein imposes directionality on an electrochemical sodium gradient-driven TRAP transporter. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 1778–1783.

182. Nakanishi, T., Tamai, I., Takaki, A. & Tsuji, A. (2000) Cancer cell-targeted drug delivery utilizing oligopeptide transport activity. *Int. J. Cancer* **88,** 274–280.

183. Nässl, A.-M., Rubio-Aliaga, I., Fenselau, H., Marth, M. K., Kottra, G. & Daniel, H. (2011) Amino acid absorption and homeostasis in mice lacking the intestinal peptide transporter PEPT1. *Am. J. Physiol.* **301,** G128–137.

184. Neumann, J. & Brandsch, M. (2003) δ-Aminolevulinic Acid Transport in Cancer Cells of the Human Extrahepatic Biliary Duct. *Pharmacology* **305,** 219–224 (2003).

185. Newman, M. J., Fostera, D. L., Wilson, T. H. & Kaback, H. R. (1981) Purification and Reconstitution of functional lactose carrier from *Escherichia coli*. *J. Biol. Chem.* **256,** 11804–11808.

186. Newman, M. J. & Wilson, T. H. (1980) Solubilization and reconstitution of the Lactose Transport System from Escherichia coli. *J. Biol. Chem.* **255,** 10583-10586.

187. Newstead, S., Ferrandon, B. & Iwata, S. (2008) Rationalizing α-helical membrane protein crystallization. *Protein Sci.* **17**, 466–472.

188. Newstead, S., Drew, D., Cameron, A.D., Postis, V.L.G., Xia, X. Fowler, P.W., Ingram, J.C., Carpenter, E.P., Sansom, M.S.P., McPherson, M.J., Baldwin, S.A. & Iwata, S. (2011) Crystal structure of a prokaryotic homologue of the mammalian oligopeptide-proton symporters, PepT1 and PepT2. *EMBO J.* **30**, 417-426.

189. Newstead, S., Kim, H., Heijne, G. Von, Iwata, S. & Drew, D. (2007) High-throughput fluorescence-based optimization of eukaryotic membrane protein overexpression and purification in Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 13936-13941.

190. Nicholls, D.G., Ferguson, S.J. (2002) Bioenergetics, 3rd Edition. Academic Press Inc., London.

191. Nietlispach, D. & Gautier, A. (2011) Solution NMR studies of polytopic α-helical membrane proteins. *Curr. Opin. Struct. Biol.* **21,** 497–508.

192. Nissen, P. (2012) A POTluck of peptide transporters. *EMBO J.* **31,** 3382–3383.

193. Nour-Eldin, H. H., Andersen, T. G., Burow, M., Madsen, S. R., J ørgensen, M.E., Olsen, C. E., Dreyer, I., Hedrich, R., Geiger, D. & Halkier, B. A. (2012) NRT/PTR transporters are essential for translocation of glucosinolate defence compounds to seeds. *Nature* **488,** 531–534.

194. Okamura, M., Terada, T., Katsura, T., Saito, H. & Inui, K. (2003) Inhibitory effect of zinc on PEPT1-mediated transport of glycylsarcosine and beta-lactam antibiotics in human intestinal cell line Caco-2. *Pharm. Res.* **20,** 1389–1393.

195. Overington, J. P., Al-Lazikani, B. & Hopkins, A. L. (2006) How many drug targets are there? *Nat. Rev. Drug Discov.* **5,** 993–996.

196. Palsdottir, H. & Hunte, C. (2004) Lipids in membrane protein structures. *Biochim. Biophys. Acta* **1666,** 2–18.

197. Pao, S. S., Paulsen, I. A. N. T. & Saier, M. H. (1998) Major Facilitator Superfamily. *Microbiology* **62,** 1–34.

198. Parker, J. L. & Newstead, S. (2012) Current trends in α-helical membrane protein crystallization: an update. *Protein Sci*. **21,** 1358–1365.

199. Parker, J. L. & Newstead, S. (2013) Phasing statistics for alpha helical membrane protein structures. *Protein Sci*. doi: 10.1002/pro.2341.

200. Pebay-Peyroula, E., Dahout-Gonzalez, C., Kahn, R., Trezeguet, V., Lauquin, G. J.-M., Brandolin, G. (2003) Structure of mitochondrial ADP/ATP carrier in complex with carboxyatractyloside. *Nature* **426,** 39–44.

201. Pedersen, B. P., Morth, J. P. & Nissen, P. (2010) Structure determination using poorly diffracting membrane-protein crystals: the H^+-ATPase and Na^+, K^+-ATPase case history. *Acta Crystallogr*. **66,** 309–313.

202. Pedersen, B. P., Kumar, H., Waight, A. B., Risenmay, A. J., Roe-Zurz, Z., Chau, B. H., Schlessinger, A., Bonomi, M., Harries, W., Sali, A., Johri, A. K. & Stroud, R. M. (2013) Crystal structure of a eukaryotic phosphate transporter. *Nature* **496**, 533-536.

203. Pedretti, A., De luca, L., Marconi, C., Negrisoli, G., Aldini, G. & Vistoli, G. (2008) Modeling of the intestinal peptide transporter hPepT1 and analysis of its transport capacities by docking and pharmacophore mapping. *ChemMedChem* **3,** 1913–1921.

204. Pedretti, A., De Luca, L., Marconi, C., Regazzoni, L., Aldini, G. & Vistoli, G. (2011) Fragmental modeling of hPepT2 and analysis of its binding features by docking studies and pharmacophore mapping. *Bioorg*. *Med*. *Chem*. **19,** 4544–4551.

205. Perry, J. L., Dembla-Rajpal, N., Hall, L. A. & Pritchard, J. B. (2006) A three-dimensional model of human organic anion transporter 1: aromatic amino acids required for substrate transport. *J*. *Biol*. *Chem*. **281,** 38071–38079.

206. Pieri, M., Gan, C., Bailey, P. & Meredith, D. (2009) The transmembrane tyrosines Y56, Y91 and Y167 play important roles in determining the affinity and transport rate of the rabbit proton-coupled peptide transporter PepT1. *Int*. *J*. *Biochem*. *Cell*. *Biol*. **41,** 2204–2213.

207. Piscitelli, C.L., Krishnamurthy, H. & Gouaux, E. (2010) Neurotransmitter/sodium symporter orthologue LeuT has a single high-affinity substrate site. *Nature* **468,** 1129–1132.

208. Quick, M. & Javitch, J. A. (2007) Monitoring the function of membrane transport proteins in detergent-solubilized form. *Proc*. *Natl*. *Acad*. *Sci*. *U.S.A*. **104,** 3603–3608.

209. Quistgaard, E. M., Löw, C., Moberg, P., Trésaugues, L. & Nordlund, P. (2013) Structural basis for substrate transport in the GLUT-homology family of monosaccharide transporters. *Nat*. *Struct*. *Mol*. *Biol*. **20,** 766–769.

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

211. Raschle, T., Hiller, S., Etzkorn, M. & Wagner, G. (2011) Non-micellar systems for solution NMR spectroscopy of membrane proteins. *Curr*. *Opin*. *Struct*. *Biol*. **20,** 471–479.

212. Rasmussen, S. G. F., Choi, H.-J., Rosenbaum, D. M., Kobilka, T. S., Thian, F. S., Edwards, P. C., Burghammer, M., Ratnala, V. R. P., Sanishvili, R., Fischetti, R. F., Schertler, G. F. X., Weis, W. I. & Kobilka, B. K. (2007) Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature* **450,** 383–387.

213. Rayment, I. (1997) Reductive alkylation of lysine residues to alter crystallisation properties of proteins. *Methods Enzymol*. **276,** 171-179.

214. Reginsson, G. W. & Schiemann, O. (2011) Pulsed electron-electron double resonance: beyond nanometre distance measurements on biomacromolecules. *Biochem. J.* **434,** 353–363.

215. Ren, Q. & Paulsen, I. T. (2007) Large-scale comparative genomic analyses of cytoplasmic membrane transport systems in prokaryotes. *J. Mol. Microbiol. Biotechnol.* **12,** 165–179.

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

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

218. Rigaud, J.-L. & Lévy, D. (2003) Reconstitution of membrane proteins into liposomes. *Methods Enzymol.* **372,** 65–86.

219. Roepe, P. D. & Kaback, H. R. (1989) Site-directed mutagenesis of tyrosine residues in the lac permease of *Escherichia coli*. *Biochemistry* **28,** 6127–6132.

220. Roszak, A. W., Howard, T. D., Southall, J., Gardiner, A. T., Law, C. J., Isaacs, N. W. & Cogdell, R. J. (2003) Crystal structure of the RC-LH1 core complex from Rhodopseudomonas palustris. *Science* **302,** 1969–1972.

221. Rottenberg, H. (1979) The measurement of membrane potential and ΔpH in cells, organelles and vesicles. *Methods Enzymol.* **LV,** 547-569.

222. Rubio-Aliaga, I., Frey, I., Boll, M., Groneberg, D. A., Eichinger, H. M., Balling, R. & Daniel, H. (2003) Targeted Disruption of the Peptide Transporter Pept2 Gene in Mice Defines Its Physiological Role in the Kidney. *Mol. Cell. Biol.* **23,** 3247–3252.

223. Sahin-Tóth, M., Karlin, A. & Kaback, H. R. (2000) Unraveling the mechanism of the lactose permease of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **97,** 10729–10732.

224. Saier, M. H. (2000) A functional-phylogenetic classification system for transmembrane solute transporters. *Microbiol. Mol. Biol. Rev.* **64,** 354–411.

225. Saier, M. H. (2003) Tracing pathways of transport protein evolution. *Mol. Microbiol.* **48,** 1145–1156.

226. Saito, H., Okuda, M., Terada, T., Sasaki, S. & Inui, K.-I. (1995) Cloning and Characterization of a Rat H^+/Peptide cotransporter mediating absorption of beta-lactam antibiotics the Intestine and kidney. *J. Pharmacol. Exp. Ther.* **275**, 1631-1637.

227. Sawada, K., Terada, T., Saito, H., Hashimoto, Y. & Inui, K.-I. (1999) Effects of glibenclamide on glycylsarcosine transport by the rat peptide transporters PEPT1 and PEPT2. *Br. J. Pharmacol.* **128**, 1159-1164.

228. Schanda, P. & Brutscher, B. (2005) Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds. *J. Am. Chem. Soc.* **127,** 8014–8015.

229. Schmidt, D., Cross, S. R. & MacKinnon, R. (2009) A gating model for the archeal voltage-dependent K^+ channel KvAP in DPhPC and POPE:POPG decane lipid bilayers. *J. Mol. Biol.* **390,** 902–912.

230. Serrano-Vega, M. J., Magnani, F., Shibata, Y. & Tate, C. G. (2008) Conformational thermostabilization of the β1-adrenergic receptor in a detergent-resistant form. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 877–882.

231. Serrano-Vega, M. J. & Tate, C. J. (2009) Transferability of thermostabilizing mutations between β-adrenergic receptors. *Mol. Membr. Biol.* **26**, 385–396.

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

233. Shen, H., Smith, D. E., Yang, T., Huang, Y. G., Schnermann, J. B. & Brosius III, F. C. (1999) Localization of PEPT1 and PEPT2 proton-coupled oligopeptide transporter mRNA and protein in rat kidney Localization of PEPT1 and PEPT2 proton-coupled oligopeptide transporter mRNA and protein in rat kidney. *Am. J. Physiol.* **276**, F658-F665.

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

235. Shimamura, T., Weyand, S., Beckstein, O., Rutherford, N. G., Hadden, J. M., Sharples, D., Sansom, M. S. P., Iwata, S., Henderson, P. J. F. & Cameron, A. D. (2010) Molecular basis of alternating access membrane transport by the sodium-hydantoin transporter Mhp1. *Science* **328,** 470–473.

236. Shu, C., Shen, H., Teuscher, N. S., Lorenzi, P. J., Keep, R. F. & Smith, D. E. (2002) Role of PEPT2 in peptide/mimetic trafficking at the blood-cerebrospinal fluid barrier: studies in rat choroid plexus epithelial cells in primary culture. *J. Pharmacol. Exp. Ther.* **301,** 820–829.

237. Slotboom, D. J., Duurkens, R. H., Olieman, K. & Erkens, G. B. (2008) Static light scattering to characterize membrane proteins in detergent solution. *Methods* **46,** 73–82.

238. Smid, E. J., Driessen, A. J. & Konings, W. N. (1989) Mechanism and energetics of dipeptide transport in membrane vesicles of Lactococcus lactis. *J. Bacteriol.* **171,** 292–298.

239. Smirnova, I., Kasho, V., Choe, J.-Y., Altenbach, C., Hubbell, W. L. & Kaback, H. R. (2007) Sugar binding induces an outward facing conformation of LacY. *Proc. Natl. Acad. Sci. U.S.A.* **104,** 16504–16509.

240. Smirnova, I., Kasho, V., Sugihara, J. & Kaback, H. R. (2009) Probing of the rates of alternating access in LacY with Trp fluorescence. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 21561–21566.

241. Shu, C., Shen, H., Hopfer, U. & Smith, D. E. (2001) Mechanism of intestinal absorption and renal reabsorption of an orally active ACE inhibitor: Uptake and transport of fosinopril in cell cultures. *Drug Metab. Dispos.* **29,** 1307–1315.

242. Solcan, N., Kwok, J., Fowler, P.W., Cameron, A.D., Drew, D., Iwata, S. & Newstead, S. (2012) Alternating access mechanism in the POT family of oligopeptide transporters. *EMBO J.* **31,** 3411–3421.

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

244. 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. & Drew, D. (2011) Benchmarking Membrane Protein Detergent Stability for Improving Throughput of High-Resolution X-ray Structures. *Structure* **19,** 17–25.

245. Steel, A., Nussberger, S., Romero, M. F., Boron, W. F., Boyd, C. A. R. & Hediger, M. (1997) Stoichiometry and pH dependence of the rabbit proton-dependent oligopeptide transporter PepT1. *J. Physiol.* **498,** 563–569.

246. Steiner, H.-Y., Song, W., Zhang, L., Naider, F., Becker, J. M. & Stacey, G. (1994) An Arabidopsis peptide transporter is a member of a new class of membrane transport proteins. *Plant Cell* **6,** 1289–1299.

247. Steiner, H.-Y, Naider, F. & Becker, J. M. (1995) MicroReview The PTR family : a new group of peptide transporters. *Science* **16,** 825–834.

248. Stolz, J., Stadler, R., Opekarová, M. & Sauer, N. (1994) Functional reconstitution of the solubilized Arabidopsis thaliana STP1 monosaccharide-H^+ symporter in lipid vesicles and purification of the histidine tagged protein from transgenic *Saccharomyces cerevisiae*. *Plant J*. **6,** 225–233.

249. Strop, P. & Brunger, A. T. (2005) Refractive index-based determination of detergent concentration and its application to the study of membrane proteins. *Protein Sci*. **14**, 2207–2211.

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

251. Sun, L., Zeng, X., Yan, C., Sun, X., Gong, X., Rao, Y. & Yan, N. (2012) Crystal structure of a bacterial homologue of glucose transporters GLUT1-4. *Nature* **490,** 361–366.

252. Takahashi, K., Nakamura, N., Terada, T., Okano, T., Futami, T., Saito, H. & Inui, K.-I. (1998) Interaction of beta-lactam antibiotics with H^+/peptide cotransporters in rat renal brush-border membranes. *J*. *Pharmacol*. *Exp*. *Ther*. **286,** 1037–1042.

253. Tamai, I., Nakanishi, T., Nakahara, H., Sai, Y., Ganapathy, V., Leibach, F. H. & Tsuji, A. (1998) Improvement of L-dopa absorption by dipeptidyl derivation, utilizing peptide transporter PepT1. *J*. *Pharm*. *Sci*. **87,** 1542–1546.

254. Tan, B. K., Bogdanov, M., Zhao, J., Dowhan, W., Raetz, C. R. H. & Guan, Z. (2012) Discovery of a cardiolipin synthase utilizing phosphatidylethanolamine and phosphatidylglycerol as substrates. *Proc*. *Natl*. *Acad*. *Sci*. *U*.*S*.*A*. **109,** 16504–16509.

255. Teather, R. M., Müller-Hill, B., Abrutsch, U., Aichele, G. & Overath, P. (1978) Amplification of the lactose carrier protein in Escherichia coli using a plasmid vector. *Mol*. *Gen*. *Genet*. **159,** 239–248.

256. Terada, T., Saito, H., Mukai, M. & Inui, K. I. (1996) Identification of the histidine residues involved in substrate recognition by a rat H^+/peptide cotransporter, PEPT1. *FEBS Lett*. **394,** 196–200.

257. Terada, T., Sawada, K., Ito, T., Saito, H., Hashimoto, Y. & Inui, K.-I. (2000) Functional expression of novel peptide transporter in renal basolateral membranes Functional expression of novel peptide transporter in renal basolateral membranes. *Am*. *J*. *Physiol*. **279**, F851–F857.

258. Terada, T., Irie, M., Okuda, M. & Inui, K. (2004) Genetic variant Arg57His in human H^+/peptide cotransporter 2 causes a complete loss of transport function. *Biochem*. *Biophys*. *Res*. *Commun*. **316,** 416–420.

259. Terada, T., Saito, H., Mukai, M. & Inui, K. (1997) Recognition of β-lactam antibiotics by rat peptide transporters, PEPT1 and PEPT2, in LLC-PK 1 cells. *Am*. *J*. *Physiol*. **273**, F706-F711.

260. Thamotharan, M., Bawani, S. Z., Zhou, X. & Adibi, S. A. (1999) Functional and Molecular Expression of Intestinal Oligopeptide Transporter (Pept-1) After a Brief Fast. *Metabolism* **48,** 681–684.

261. Theis, S., Hartrodt, B., Kottra, G., Neubert, K. & Daniel, H. (2002) Defining minimal structural features in substrates of the H^+/peptide cotransporter PEPT2 using novel amino acid and dipeptide derivatives. *Mol*. *Pharmacol*. **61,** 214–221.

262. Theis, S., Knütter, I., Hartrodt, B., Brandsch, M., Kottra, G., Neubert, K. & Daniel, H. (2002) Synthesis and characterization of high affinity inhibitors of the H^+/peptide transporter PEPT2. *J*. *Biol*. *Chem*. **277,** 7287–7292.

263. Thwaites, D. T., Hirst, B. H. & Simmons, N. L. (1994) Substrate specificity of the di/tripeptide transporter in human intestinal epithelia (Caco-2): identification of substrates that undergo H^+-coupled absorption. *Br. J. Pharmacol.* **113,** 1050–1056.

264. Tomita, Y., Katsura, T., Okano, T., Inui, K. & Hori, R. (1990) Transport mechanisms of bestatin in rabbit intestinal brush-border membranes: role of H^+/dipeptide cotransport system. *J. Pharmacol. Exp. Ther.* **252,** 859–862.

265. Tsai, M.-F., Fang, Y. & Miller, C. (2012) Sided functions of an arginine-agmatine antiporter oriented in liposomes. *Biochemistry* **51,** 1577–1585.

266. Uchiyama, T., Kulkarni, A. A., Davies, D. L. & Lee, V. H. L. (2003) Biophysical evidence for His57 as a proton-binding site in the mammalian intestinal transporter hPepT1. *Pharm. Res.* **20,** 1911–1916.

267. Udenfriend, S., Gerber, L. D., Brink, L. & Spector, S. (1985) Scintillation proximity radioimmunoassay utilizing 125I-labeled ligands. *Proc. Natl. Acad. Sci. U.S.A.* **82,** 8672–8676.

268. Upadhyay, A. K., Borbat, P. P., Wang, J., Freed, J. H. & Edmondson, D. E. (2008) Determination of the oligomeric states of human and rat monoamine oxidases in the outer mitochondrial membrane and octyl beta-D-glucopyranoside micelles using pulsed dipolar electron spin resonance spectroscopy. *Biochemistry* **47,** 1554–1566.

269. Vamvouka, M., Cieslak, J., Eps, N. V., Hubbell, W. & Gross, A. (2008) The structure of the lipid-embedded potassium channel voltage sensor determined by double-electron – electron resonance spectroscopy. *Protein Sci.* **17**, 506–517.

270. Vance, D.E., Vance, J.E. (2002) Biochemistry of Lipids, Lipoproteins and membranes, 4th Edition. Elsevier Science, Amsterdam.

271. Viitanen, B. P., Newman, M. J., Foster, D. L., Wilson, T. H. & Kaback, H. R. (1986) Purification, reconstitution and characterization of the lac permease of *Escherichia coli*. *Methods Enzymol.* **125**, 429-452.

272. Viitanen, P., Garcia, M. L. & Kaback, H. R. (1984) Purified reconstituted lac carrier protein from *Escherichia coli* is fully functional. *Proc. Natl. Acad. Sci. U.S.A.* **81,** 1629–1633.

273. Wagner, S., Klepsch, M. M., Schlegel, S., Appel, A., Draheim, R., Tarry, M., Hogbom, M., van Wijk, K. L., Slotboom, D. J., Persson, J. O. & de Gier, J.-W. (2008) Tuning Escherichia coli for membrane protein overexpression. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 14371–14376.

274. Walden, M., Accardi, A., Wu, F., Xu, C. Williams, C. & Miller, C. (2007) Uncoupling and turnover in a Cl^-/H^+ exchange transporter. *J. Gen. Physiol.* **129,** 317–329.

275. Waldo, G. S., Standish, B. M., Berendzen, J. & Terwilliger, T. C. (1999) Rapid protein-folding assay using green fluorescent protein. *Nat. Biotechnol.* **17,** 691–695.

276. Walter, T. S., Meier, C., Assenberg, R., Au, K.-F., Ren, J., Verma, A., Nettleship, J. E., Owens, R. J., Stuart, D. I. & Gimes, J. M. (2006) Lysine methylation as a routine rescue strategy for protein crystallization. *Structure* **14,** 1617–1622.

277. Wang, H., Elferich, J. & Gouaux, E. (2012) Structures of LeuT in bicelles define conformation and substrate binding in a membrane-like context. *Nat. Struct. Mol. Biol.* **19,** 212–219.

278. Wang, X., Bogdanov, M. & Dowhan, W. (2002) Topology of polytopic membrane protein subdomains is dictated by membrane phospholipid composition. *EMBO J.* **21,** 5673–5681.

279. Webb, K. E., Matthews, J. C. & DiRienzo, D. B. (1992) Peptide absorption: a review of current concepts and future perspectives. *J. Anim. Sci.* **70,** 3248–3257.

280. Weitz, D., Harder, D., Casagrande, F., Fotiadis, D., Obrdlik, P., Kelety, B. & Daniel, H. (2007) Functional and structural characterization of a prokaryotic peptide transporter with features similar to mammalian PEPT1. *J. Biol. Chem.* **282,** 2832–2839.

281. Wenzel, U., Gebert, I., Weintraut, H., Weber, W.-M., Clauss, W. & Daniel, H. (1996) Transport characteristics of differently charged cephalosporin antibiotics in oocytes expressing the cloned intestinal peptide transporter PepT1 and in human intestinal Caco-2 cells. *J. Pharmacol. Exp. Ther.* **277,** 831–839.

282. 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. & Cameron, A. D. (2008) Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter. *Science* **322,** 709–713.

283. Whiles, J. A., Deems, R., Vold, R. R. & Dennis, E. A. (2002) Bicelles in structure–function studies of membrane-associated proteins. *Bioorg. Chem.* **30,** 431–442.

284. Widdas, W. F. (1952) Inability of diffusion to account for placental glucose transfer in the sheep and consideration of the kinetics of a possible carrier transfer. *J. Physiol.* **118,** 23–39.

285. Xie, Y., Chou, L., Cutler, A. & Weimer, B. (2004) DNA Macroarray Profiling of Lactococcus lactis subsp. lactis IL1403 Gene Expression during Environmental Stresses. *Appl. Environ. Microbiol.* **70,** 6738–6747.

286. Xu, L., Haworth, I. S., Kulkarni, A. A., Bolger, M. B. & Daryl, L. (2009) Mutagenesis and cysteine scanning of transmembrane domain 10 of the human dipeptide transporter. *Pharm. Res.* **26,** 2358–2366.

287. Yamashita, A., Singh, S. K., Kawate, T., Jin, Y. & Gouaux, E. (2005) Crystal structure of a bacterial homologue of Na^+/Cl^--dependent neurotransmitter transporters. *Nature* **437,** 215–223.

288. Yamashita, E., Zhang, H. & Cramer, W. A. (2007) Structure of the cytochrome b6f complex: quinone analogue inhibitors as ligands of heme cn. *J. Mol. Biol.* **370,** 39–52.

289. Yan, H., Huang, W. Yan, C., Gong, X., Jiang, S., Zhao, Y., Wang, J. & Shi, Y. (2013) Structure and mechanism of a nitrate transporter. *Cell Rep.* **3,** 716–723.

290. Yang, Q., Wang, X., Ye, L., Mentrikoski, M., Mohammadi, E., Kim, Y.-M. & Maloney, P. C. (2005) Experimental tests of a homology model for OxlT, the oxalate transporter of *Oxalobacter formigenes*. *Proc. Natl. Acad. Sci. U.S.A.* **102,** 8513–8518.

291. Yeung, A. K., Basu, S. K., Wu, S. K., Chu, C., Okamoto, C. T., Hamm-Alvarez, S. F., von Grafenstein, H., Shen, W.-C., Kim, K.-J., Bolger, M. B., Haworth, I. S., Ann, D. K. & Lee, V. H. (1998) Molecular identification of a role for tyrosine 167 in the function of the human intestinal proton- coupled dipeptide transporter (hPepT1). *Biochem. Biophys. Res. Commun.* **250,** 103–107.

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

293. Zhang, Q., Ma, X., Ward, A., Hong, W.-X., Jaakola, V.-P., Stevens, R. C., Finn, M. G. & Chang, G. (2007) Designing facial amphiphiles for the stabilization of integral membrane proteins. *Angew. Chem.* **46**, 7023 –7025.

294. Zhang, Q., Tao, H. & Hong, W.-X. (2011) New amphiphiles for membrane protein structural biology. *Methods* **55,** 318–323.

295. Zhang, Y.-M. & Rock, C. O. (2008) Membrane lipid homeostasis in bacteria. *Nat. Rev. Microbiol.* **6,** 222–233.

296. Zhou, Y., Cierpicki, T. Jimenez, R. H. F., Lukasik, S. M., Ellena, J. F., Cafiso, D. S., Kadokura, H., Beckwith, J. & Bushweller, J. H. (2009) NMR solution structure of the integral membrane enzyme DsbB: functional insights into DsbB-catalyzed disulfide bond formation. *Mol. Cell* **31,** 896–908.

297. Zhu, T., Chen, X., Steel, A., Hediger, M. A. & Smith, D. E. (2000) Differential recognition of ACE inhibitors in Xenopus Laevis oocytes expressing rat PepT1 and PepT2. *Pharm. Res.* **17,** 526–532.

298. Ziegler, T. R., Fernández-Estívariz, C., Gu, L. H., Bazargan, N., Umeakunne, K., Wallace, T. M., Diaz, E. E., Rosado, K. E., Pascal, R. R., Galloway, J. R., Wilcox, J. N. & Leader, L. M. (2002) Distribution of the H^+/peptide transporter PepT1 in human intestine: up-regulated expression in the colonic mucosa of patients with short-bowel syndrome. *Am. J. Clin. Nutr.* **75,** 922–930.