Structural and functional analyses of the *Escherichia coli* peptide transporter DtpA

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Yonca Ural-Blimke

Aus Istanbul, Türkei

Tübingen

2018

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	27.11.2018
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Samuel Wagner
2. Berichterstatter:	Prof. Dr. Thilo Stehle

The work presented in this dissertation was performed in the group of
Dr Christian Löw at the European Molecular Biology Laboratories (EMBL) Hamburg
under his supervision.

Abstract

The proton-dependent oligopeptide transporters (POTs) are ubiquitous in all kingdoms of life and the only peptide transport system in humans. Coupled to co-transport of protons, they take up di- and tripeptides into the cell. Two human homologues are crucial for the nutrient (re)absorption and drug delivery. Human PepT1 (hPepT1, SLC15A1) is highly expressed in the small intestines, whereas hPepT2 (SLC15A2) is mainly found in the kidney. Indeed, antibiotics such as ceftibuten, medication against high blood pressure such as enalapril and antiviral drugs such as valacyclovir and valganciclovir are taken up by hPepT1 and hPepT2. Although these transporters have been studied biochemically in the past, structural information on any of the eukaryotic transporters is missing. High-resolution structures from bacterial homologues contributed to our understanding of ligand binding and selectivity but these transporters do not transport drugs. Exceptionally, a homologue from *Escherichia coli*, DtpA, shows a striking similarity to hPepT1 in terms of ligand selectivity and transports several different drug molecules.

To characterize the ligand selectivity of DtpA further, I used a number of *in vitro* and *in vivo* methods: A ligand library consisting of di- and tripeptides and selected drugs was screened by differential scanning fluorimetry and microscale thermophoresis. Further, the screening was complemented by the established *in vivo* transport and competition assay. Both *in vivo* and *in vitro* data were in good agreement with the following conclusions: (1) DtpA has a significantly higher preference for tripeptides than dipeptides, and (2) prefers di- and tripeptides with hydrophobic and aromatic residues compared to charged residues. (3) Glycine containing ligands are less likely to bind and be transported by DtpA. From the drug library, valacyclovir and valganciclovir stood out with binding affinities similar to the ones the measured for di- and tripeptides.

The structure of DtpA was determined using a nanobody as a crystallization chaperone in a ligand-free state at 3.3 Å and a valganciclovir-bound state at 2.7 Å. Valganciclovir, L-valyl ester of ganciclovir, is an antiviral prodrug used in the treatment of cytomegalovirus. The valine moiety increases the solubility and absorption rate of the compound. In fact, ganciclovir without the N-terminal valine

cannot be transported by hPepT1 or hPepT2. Because the N-terminal residue of a di-/tripeptide plays an important role for the ligand coordination in previous ligand-bound bacterial structures, the valine moiety was suggested to be coordinated in the position of the N-terminal residue of a di-/tripeptide. However, in the DtpA-valganciclovir structure, the guanine ring mimics the N-terminal residue and the valine moiety is coordinated in a previously uncharacterized pocket. This pocket is larger in DtpA compared to other studied POTs due to a structural difference in transmembrane helix 10 caused by an intrahelical loop. Transmembrane helix 11 is therefore shifted away from the ligand-binding site, which increases the pocket size and allows the transport of various drug molecules.

In summary, I present the functional characterization of DtpA by combining *in vivo* and *in vitro* studies with a focus on ligand selectivity and bring new insights to this transporter. Furthermore, the DtpA-valganciclovir structure reveals an unexpected binding mode of the drug and an uncharacterized pocket within the ligand-binding site. Finally, we provide a homology model of hPepT1 laying the groundwork for further drug development.

Zusammenfassung

Protonen-abhängige Oligopeptidtransporter (POT) kommen in nahezu allen Organismen vor und sind in Eukaryota das einzige Transportsystem für kurze Peptide. Die Aufnahme von Di- und Tripeptide in die Zelle ist an den Cotransport von Protonen gekoppelt. Die zwei humanen Transportproteine, hPepT1 und hPepT2, sind unentbehrlich für die Aufnahme von Eiweißen und Medikamenten. hPepT1 wird vor allem im Dünndarm und hPepT2 in den Nieren exprimiert. Das Antibiotikum Ceftibuten, das Bluthochdruck-Medikament Enalapril, die antiviralen Präparate Valacyclovir und Valganciclovir sowie weitere Medikamente werden von diesen Transportern in den Körper aufgenommen.

Obwohl zahlreiche biochemische Studien mit hPepT1 und hPepT2 durchgeführt wurden, fehlen strukturelle Kenntnisse über diese pharmakologisch wichtigen Transportproteine. Die hochauflösende Strukturanalyse von bakteriellen Homologen bereicherten das Verständnis über die Ligandenbindung und –selektivität, aber diese Proteine transportieren keine Medikamente. Die Ausnahme bildet ein Homolog aus Escherichia coli, DtpA. DtpA nimmt wie hPepT1 Medikamente auf und zeigt darüberhinaus Ähnlichkeiten in der Ligandenselektivität.

Um das Verständnis der Ligandenselektivität von DtpA zu vertiefen, habe ich *in vitro* und *in vivo* Methoden verwendet. Mittels *Differential-Scanning*-Fluorometrie (DSF) und *Microscale Thermophoresis* (MST) wurde die Ligandenbindung von DtpA anhand einer Bibliothek von Di- und Tripeptiden sowie Medikamenten untersucht. Diese Studie wurde mit einem etablierten *in vivo* Konkurrenztransportassay komplementiert. Die Ergebnisse waren übereinstimmend und sind in folgenden drei Punkten aufgelistet. (1) DtpA bevorzugt Tripeptide gegenüber Dipeptiden. (2) Di- und Tripeptide mit hydrophoben und aromatischen Seitenketten werden bevorzugt im Gegensatz zu Di- und Tripeptide mit geladenen Seitenketten. (3) Di- und Tripeptide, die Glycin beinhalten, werden nicht bevorzugt. Von den Medikamenten stachen nur Valacyclovir und Valganciclovir hervor.

Die Struktur von DtpA wurde mit Hilfe eines Nanobodys als Kristallisationschaperon jeweils ohne einen Liganden und gebunden an Valganciclovir mit 3.3 Å und 2.7 Å Auflösung bestimmt. Valganciclovir ist der L-Valinylester von Ganciclovir, welches

in der Therapie von Cytomegalie eingesetzt wird. Der Valinrest erhöht die Löslichkeit und die Absorptionsrate von Valganciclovir. hPepT1 und hPepT2 transportieren nur Valganciclovir aber nicht Ganciclovir. Valacyclovir und Valganciclovir ähneln sich sehr stark in der chemischen Struktur. In vorherigen Strukturanalysen wurde eine wichtige Rolle der N-terminalen Aminosäure des Liganden zugeschrieben, darauffolgend wurde eine ähnliche Rolle für den N-terminalen Valinrestes für die Koordination von Valacyclovir in der Bindefalte von hPepT1 vorausgesagt. Entgegen der Voraussage nimmt der Guaninrest von Valganciclovir die Rolle der N-terminalen Aminosäure in der Bindefalte von DtpA ein, während der Valinrest in einer bisher unbekannten Bindetasche koordiniert wird. Diese Bindetasche in DtpA ist größer im Vergleich zu den anderen bekannten POT Strukturen. Dies wird durch eine Strukturänderung an den Transmembranhelices 10 und 11 hervorgerufen. Transmembranhelix 10 wird durch einen Loop unterbrochen, wodurch Transmembranhelix 11 von der Bindefalte weggeschoben wird. Dadurch entsteht eine größere Bindetasche, die Valganciclovir und höchstwahrscheinlich auch Tripeptide besser koordinieren kann.

In dieser Arbeit wurde mittels *in vivo* und *in vitro* Methoden weitere Kenntnisse über die Ligandenbindung von DtpA gewonnen. Die Strukturanalyse von DtpA in Komplex mit Valganciclovir hat eine neue Bindetasche in den Fokus gebracht, die in der Bindung von Medikamenten und Tripeptide eine Rolle spielt. Zusätzlich wurde eine Homologiemodell von hPepT1 mit Valganciclovir auf Basis der neuen strukturellen Erkenntnisse des prototypischen DtpA berechnet, welches für die Medikamentenentwicklung verwendet werden kann.

Table of Contents

List of Tables

List of Figures

Abbreviations

7.8 MAG	1-(7Z-pentadecenoyl)-rac-glycerol
DTT	1,4-Dithiothreitol
A_{280nm}	Absorbance at 280 nm
AK-AMCA	β-AlaLys-(7-amino-4-methylcoumarin-3-yl) acetic acid
CV	Column volume
DMSO	Dimethyl sulfoxide
K_d	Dissociation constant
E. coli	*Escherichia coli*
EtBr	Ethidium bromide
pBpF	H-p-Bz-Phe-OH
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LCP	Lipidic cubic phase
LB	Lysogeny broth
MST	Microscale thermophoresis
DDM	n-Dodecyl-β-D-maltopyranoside
nanoDSF	Nano differential scanning fluorimetry
N00	Nanobody 00
OD_{600nm}	Optical density at 600 nm
PEG 400	Polyethylene glycol 400
PEG 600	Polyethylene glycol 600
PCR	Polymerase chain reaction
PDB	Protein data bank
rmsd	Root mean square deviation
NaCl	Sodium chloride
NaP	Sodium phosphate
SN	Supernatant
TEV	*Tobacco etch virus*
TB	Terrific Broth media
TCEP	Tris(2-carboxyethyl)phosphine

Table 1: List of amino acids in three-letter- and one-letter-code

Amino acid	Three-letter-code	One-letter-code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

1. Introduction

1.1. The membrane and membrane proteins

Membranes play a crucial role in cell biology: they isolate the cellular compartments from the environment and they subdivide the cell into independent and specialized compartments. This is vital for keeping molecules within the cell and cellular organelles, and to maintain chemical gradients inside and outside of the cell or the organelles. Chemically, the membrane is a lipid bilayer with a hydrophobic core and a hydrophilic surface, where proteins are anchored or embedded. Thus, it is a dynamic and semi-permeable barrier to molecules, where many biochemical reactions occur.

The membrane is a complex environment consisting of various types of lipids. It differs in lipid composition and lipid ratios across species (Hannich, Umebayashi, and Riezman 2011), between the different organelle membranes or even as an adaptation to the environmental growth conditions (O'Brien and Frerman 1980). Furthermore, the lipid composition in the two leaflets of the bilayer is asymmetric due to the high-energy barrier for a lipid to swap to the opposite side. The lipids found in biomembranes can be grouped into three categories: (i) glycerophospholipids, (ii) sphingolipids and (iii) sterols (Figure 1).

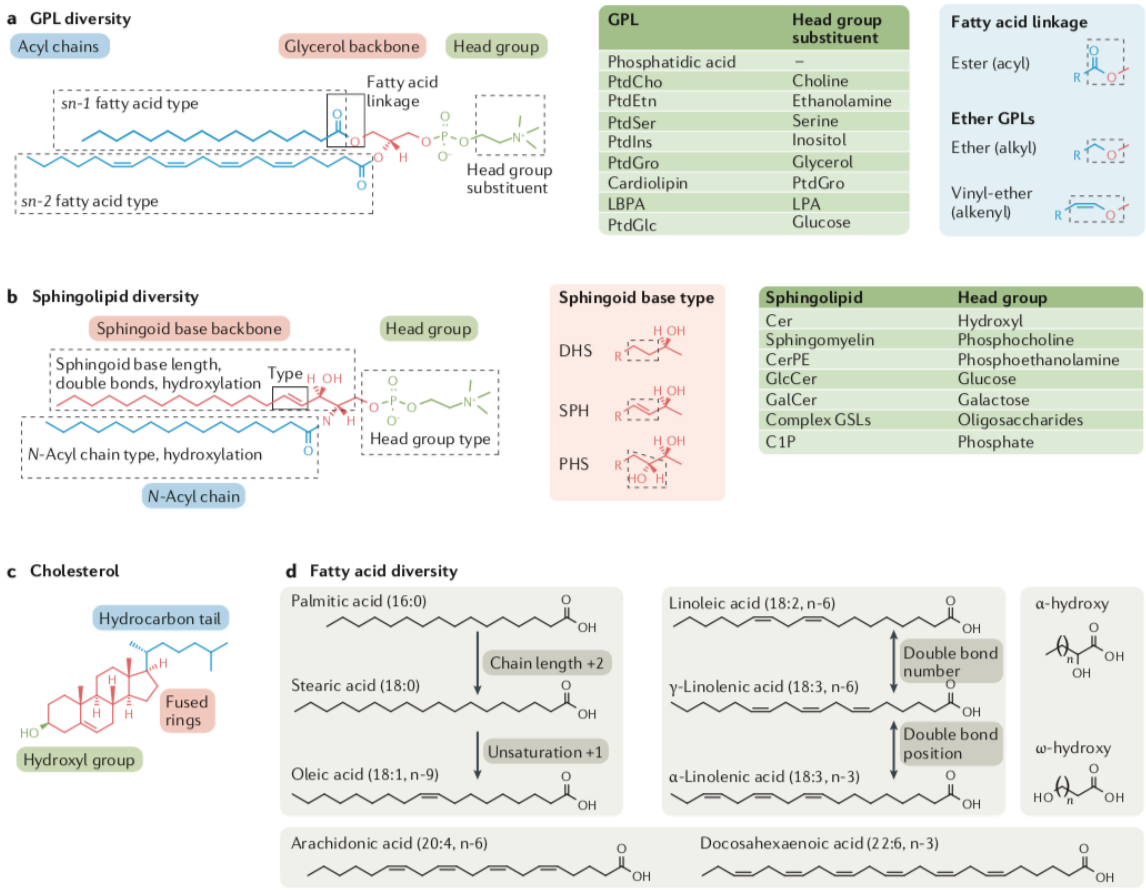


Figure 1: Diversity of membrane lipids.
(a) Glycerophospholipids (GPL) are illustrated with the glycerol backbone attached to two fatty acids and a head group. On the right possible head group substituents are listed (green box) as well as the chemical linkage between the glycerol backbone and the fatty acids (blue box). (b) Sphingolipids are presented with the sphingoid backbone linked to a N-acyl chain and a head group with the different sphingoid base types (pink box) and the head groups (green box). (c) Cholesterol molecule. (d) Common fatty acids in the lipid membrane. Figure modified from (Harayama and Riezman 2018).

Glycerophospholipids are built on a glycerol-3-phosphate backbone, which is esterified to fatty acids at two positions and the phosphate is esterified with an alcohol head group, e.g. serine or inositol. In case of sphingolipids, instead of glycerol they are built on a sphingosine backbone linked to one fatty acid via a peptide bond and one alcohol esterified with the phosphate group. The fatty acids are typically 10 to 24 carbon long acyl chains, which are called saturated if no double bond is present and unsaturated if at least one double bond is present within the carbon tail. Additionally, the phospholipids or sphingolipids can accommodate an oligosaccharide head-group. Finally, sterols are tetracyclic compounds formed of several isoprenyl groups.

Cholesterol is the most common sterol in the membrane of animals, ergosterol in fungi and stigmasterol and sitosterol in plants. Prokaryotes don't have sterols in the membrane in the strict sense. Instead, hopanoids were identified in prokaryotic membranes, which are pentacyclic triterpenoids and functional sterol surrogates (Dufourc 2008; Sáenz *et al*. 2015). Functionally, lipids play additional roles in signaling, energy storage and as cofactors, precursors and pigments.

Overall, the membrane is composed of several different lipids and their relative proportions strongly affect dynamics of the membrane. Although the transverse movement of the lipids (movement across the leaflets) is energetically unfavorable, they move laterally (within a leaflet) according to Brownian motion but due to their interaction with proteins, they are not randomly distributed across the bilayer. The membrane is crowded with proteins affecting the permeability. Only small neutral molecules can diffuse through a membrane but due to the crowdedness of the membrane by integral membrane proteins and sugars, the diffusion rates are very low, and even for these molecules transport systems are required (Agre 2004).

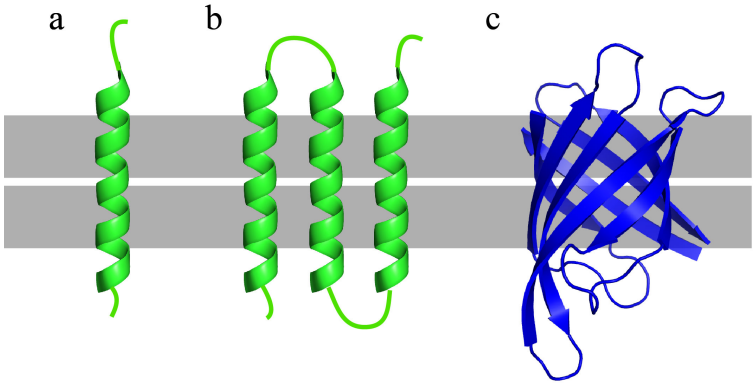


Figure 2: Integral membrane proteins.
There are three types of integral membrane proteins, α-helical proteins with either (a) a single helix or (b) multiple helices spanning through the membrane and (c) β-barrels, which consist of β-sheets. As an example, the structure of the human retinol-binding protein (PDB id: 1RBP) is illustrated.

There are two types of membrane proteins, integral and peripheral. The integral membrane proteins are fully embedded in the membrane, with one or multiple α-helices spanning the hydrophobic core of the membrane (Figure 2). Typically, one transmembrane α–helix consists of 20-25 hydrophobic residues. According to the positive-inside rule positively charged residues (Lys, Arg, His) accumulate on the

cytoplasmic side of the protein and aromatic residues (Trp, Tyr and Phe) are often at the lipid water interface. There are also integral membrane proteins consisting of β-sheets, which form a membrane-spanning β-barrel (Figure 2). Every second residue of the β-sheets is a hydrophobic residue, orienting its side chain towards the "outside" of the barrel. The peripheral membrane proteins do not intrude through the hydrophobic core of the membrane. Their α-helices or loops interact with the hydrophilic lipid surface or other integral membrane proteins via electrostatic interactions or they are anchored covalently to lipids. Proteins on the extracellular peripheral site are often involved in the extracellular matrix, whereas on the intracellular side they have variable functions from interactions with the cytoskeletal proteins to soluble enzymes, which act on lipid molecules.

Integral membrane proteins make up 20-30 % of most genomes (Krogh *et al*. 2001) and perform a variety of roles. For example, the electron transport chain consists of integral membrane proteins and is crucial for ATP synthesis, the energy unit of the cell. As in the case of the electron transport chain, the membrane proteins often move the substrate from one protein to the next like an assembly line with the membrane serving as the platform. The cells communicate with their environment and with each other via membrane receptors and adhesion proteins. There are enzymes, which act within the membrane, such as proteases. Transporters and channels carry the task of moving molecules across the membrane. The difference between transporters and channels is that channels have an open pore, which has a specific size for its substrate and remains open during the conducting phase to both sides of the membrane, whereas transporters open only to one side of the membrane at a given time point and need to go through conformation changes for the transport activity. In the next chapter the role of transporters will be discussed in more detail.

1.2. Transporters

The semi-permeable membrane allows a very limited number of molecules to pass through the membrane. Only small and neutral molecules fall into this category. The high number of proteins within the membrane further limits the simple diffusion. The membrane creates an electrochemical gradient, which is built by the different concentrations of molecules and of charges on both sides of the membrane. The

electrochemical gradient is an important factor for all transport activities across the membrane. The different types of transporter proteins are classified by the stoichiometry of solutes and/or the energy dependency of the transport mechanism. Depending on the stoichiometry, a transporter is called uniporter if only one molecule is transported during a transport cycle. If multiple molecules are transported, depending on the direction of the molecules relative to each other they are called symporters, for same direction, or antiporters, for opposite directions (Figure 3).

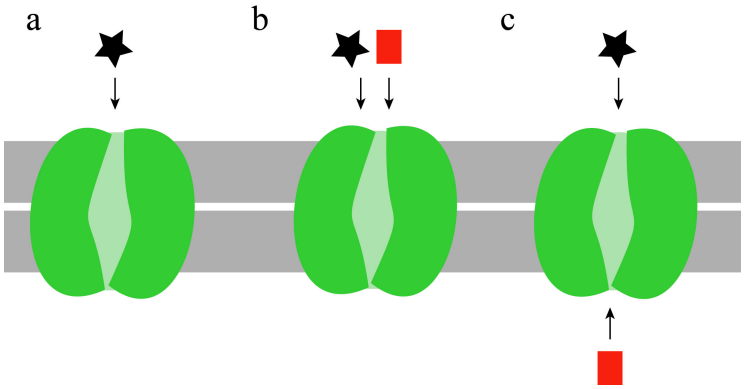


Figure 3: Types of transport based on the stoichiometry
Transporters, illustrated in green, are named (a) uniporters, if only one ligand is moved across the membrane. In case of multiple ligands, they are called (b) symporters, when the ligands are moved in the same direction or (c) antiporters, if in opposite directions.

All processes of transport require energy. The electrochemical gradient is the driving force for the facilitated diffusion, where a molecule is transported from the high concentration side of the membrane to the low concentration side. Often a molecule needs to be transported against its concentration gradient. There are primary active transporters, which use the energy gained from ATP hydrolysis for the transport. ATP-binding cassette (ABC) transporters build the largest superfamily of primary active transporters. The secondary active transporters couple the transport of the molecule of interest to the co-transport of a proton, ion or solute along its electrochemical gradient. The major facilitator superfamily (MFS) is the largest and best studied superfamily of secondary active transporters. The ion channels transport ions along their concentration gradient. They are, as mentioned before, different from transporters in their mechanism and are open to both sides of the membrane. Another type of transport is via ionophore-mediated ion transport, in which a molecule binds

to the ions shielding its charge and moves it across the membrane to release it on the other side. Figure 4 shows the different types of transporters that are described above. In this thesis, the secondary active transporters and more specifically the major facilitator superfamily is in the focus.

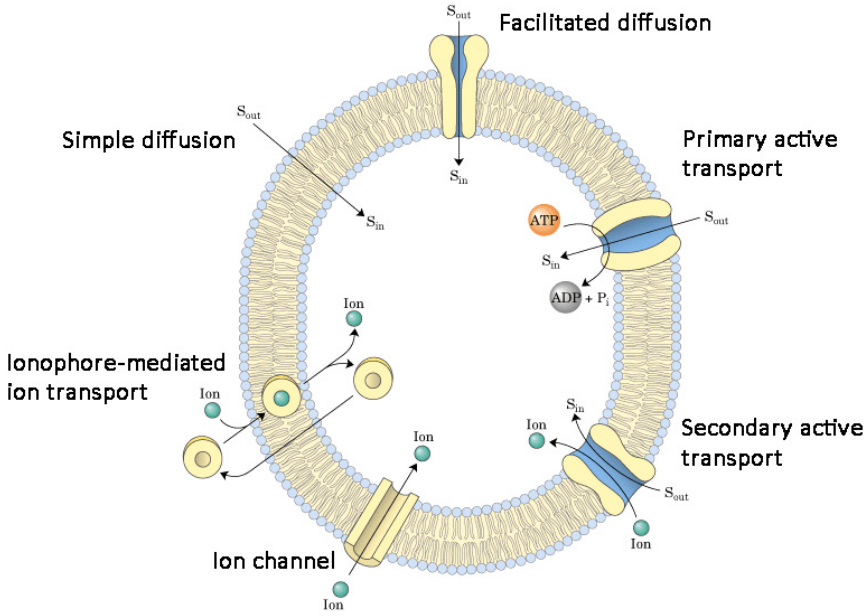


Figure 4: Types of transporters across the membrane
Only small neutral molecules can trespass the membrane via simple diffusion without the need of a transporter. Other molecules need transporter proteins. Along their concentration gradient, they pass through the transporter protein via facilitated diffusion. Often molecules need to move against their concentration gradient, which requires energy. Primary active transporters use hydrolyses of ATP, whereas the secondary active transporters move an additional ion or solute along its concentration gradient to obtain the necessary energy for the transport. Other transport mechanisms are via ion channels and ionophore mediated ion transport. Figure modified from (Nelson, Lehninger, and M 2000)

There are two commonly used classification systems for transporters. The Transport Classification Database (TCDB) is accepted by the International Union of Biochemistry and Molecular Biology. In this system, all transporters are categorized based on their transport mechanism, phylogenetic relations and substrate. They are given an identifier (ID), which is composed of five digits: V.W.X.Y.Z. The first digit, V is a number and classifies the transport mechanism, such as channel or primary active transporter. W is a letter and groups transporters to subclasses, followed by X, which is a number and defines the superfamily/family of transporters. The superfamily/family organization is not well defined in this classification system. Y, a

number, defines the subfamily/family of transporters and Z, a number, corresponding to a specific transporter. The other classification system is the solute carrier system (SLC) proposed by the HUGO Gene Nomenclature Committee (HGNC), which concentrates only on facilitators and secondary active transporters found in the human genome. The transporters are given an identifier as SLCnXm. n is a number and represents the protein family, X, a letter corresponding to the subfamily and m is a number given to a specific transporter. For example, the proton-dependent oligopeptide transporter family is 2.A.17 in the TCDB system and SLC15A according to the SLC classification.

1.3. Major facilitator superfamily

The major facilitator superfamily (MFS) of transporters is one of the oldest superfamilies of membrane proteins in evolution, being present in archea, prokaryotes and eukaryotes. It facilitates the transport of solutes through the membrane, either along their concentration gradient or against the concentration gradient coupled to co-transport of other solutes or ions. MFS is responsible for the transport of a wide variety of molecules such as organic and inorganic ions, nucleotides, sugars, lipids and short peptides. Thus, they play an important role in the transport of nutrients into the cell and in the import and export of toxins and drugs. Therefore, they are highly relevant for drug delivery and pharmacokinetics in humans and drug resistance in bacteria.

All transporters belonging to the MFS have a common structural fold consisting of 12 transmembrane α-helices (TM) building the transporter core. The core is organized into two bundles, each 6 TMs long, called the N-terminal and C-terminal bundle. The arrangement of the two bundles is in 2-fold pseudo-symmetry with the symmetry axis going through the center of the molecule perpendicular to the membrane plane. Each bundle can be further dissected into two groups of 3 TMs each (Yan 2013). The first TM of each 3-TM group, namely TM1, -4, -7 and 10 are in the center of the protein and build the transport path. The second TM of each 3-TM group (TM2, -5, -8, -11) is on the periphery of the center and mediates the interactions between the N- and C-terminal bundles. The third TM of each 3-TM group (TM3, -6, -9 and -12) is on the

outer shell of the transporter. There are additional soluble or transmembrane domains in some protein families.

The transport mechanism of MFS transporters was postulated over 50 years ago before any structural information was available (Jardetzky 1966). The three described criteria were the following: (i) The transporter should have a cavity in the center of the protein where the molecule can fit in. (ii) The transporter should have two conformations, which are open to opposite sides of the membrane. (iii) The binding affinity of the molecule should vary to the different conformations of the transporter. With the structural information the first two points were confirmed and further extended with the occluded conformation. During a transport cycle, the transporter is open to one side of the membrane, where it binds the ligand. The "switch" of the conformation occurs by the coordinated movement of the two bundles like a rigid body motion. Through an occluded conformation, the transporter opens to the other side of the membrane. After the release of the ligand, the transporter switches back to the initial conformation. The transporter is at no time point open to both sides during the transport cycle. Several structures of MFS transporters have been determined from different families capturing different states of this process (Figure 5). Quistgaard *et al.* further refined the model by the observation of the crystal structures of partially-occluded transporter structures, in which one TM end from each bundle comes closer as an intermediary step, called "clamping", prior to the switching to the other conformation (Quistgaard *et al*. 2016). For a long time, only one conformation of the transport cycle was captured by structural analysis of a given protein (Figure 5) making it difficult to study and understand the mechanism of conformational changes. XylE and GLUT3 were the first cases with structures of the same protein captured in multiple conformations (Quistgaard *et al*. 2013; Deng *et al*. 2015).

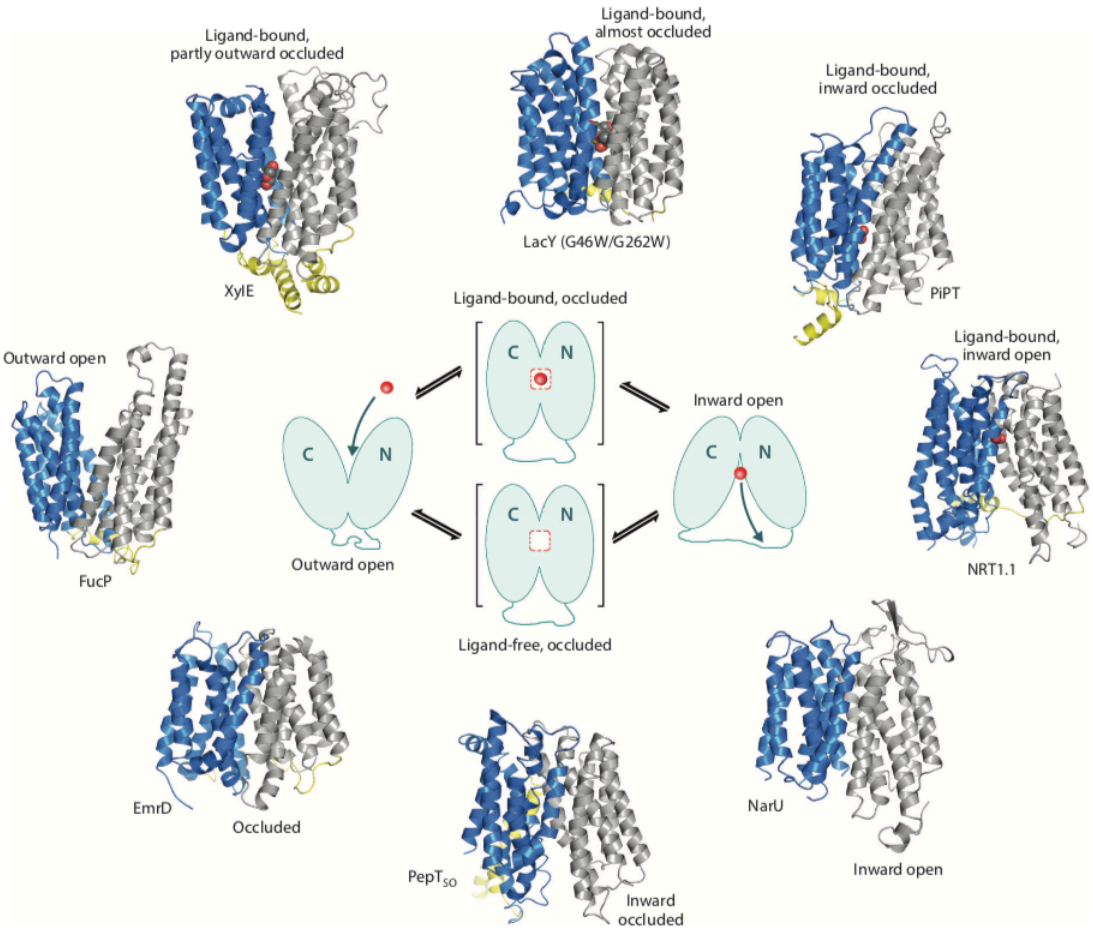


Figure 5: Alternate access mechanism of MFS transporters
The different conformations of MFS transporters are illustrated in the center starting in the outward open conformation, where the ligand, in red, can bind to the ligand-binding site in the center of the molecule. Next, through an occluded state the transporter moves to the inward open state, where the release of the ligand takes place. Through a ligand-free occluded state, the starting state of outward open conformation is reached, and the cycle is complete. On the outer circle, selected MFS transporter structures are presented in blue (N-terminal bundle) and in grey (C-terminal bundle). The protein family specific additional structural moieties are shown in yellow. Figure modified from (Yan 2015)

The next point of this dissertation focuses on the description of a family of transporters classified within the MFS superfamily: the proton-dependent oligopeptide transporters.

1.4. Proton-dependent oligopeptide transporters

Prokaryotes have several transport systems to move short peptides across the membrane. Dipeptide permease (Dpp) and oligopeptide permease (Opp) systems

belong to the ATP-binding cassette (ABC) transporter superfamily, which are primary active transporters. Proton-dependent oligopeptide transporters (POTs) were initially named tripeptide permease (Tpp) until it was discovered that these transporters, unlike the Dpp and Opp families, belong to secondary active transporters and the MFS superfamily (Steiner, Naider, and Becker 1995). POTs are present in prokaryotes and eukaryotes, but no members have been found in archea. In eukaryotes, they are the sole protein family that can deliver short peptides into the cells and play a crucial role in nutrient uptake. With the inwardly directed proton electrochemical gradient, they transport di- and tripeptides and protons. Some POTs, including human PepT1 and PepT2 as well as DtpA from *Escherichia coli*, also transport peptidomimetic drugs into the cell (Brandsch, Knütter, and Bosse-Doenecke 2008; Prabhala et al. 2017).

POTs have the conserved MFS fold. The N- and C-terminal bundles, each consisting of 6 transmembrane helices (TMs), build the transporter core. The prokaryotic members have two additional TMs, namely HA and HB, which are inserted between the N- and C-terminal bundles as mentioned previously. The function of the HAHB domain remains unknown. Insertions between the N-terminal bundle and HAHB domain showed no change in functional assays but insertion between the HA and HB helices depleted the activity completely (Zhao *et al*. 2014). Some eukaryotic members including mammalian POTs have a soluble extracellular domain, which is connected to the TM9 and TM10 (Figure 6). The crystal structure of the extracellular domain (ECD) of PepT1 from *Mus musculus* and PepT2 from *Rattus norvegicus* was determined and the overall structure is two compactly packed immunoglobin-like folds (Beale *et al*. 2015) (Figure 6). A functional role in recruiting the intestinal protease trypsin to PepT1 was suggested based on the μM-affinity of the PepT1 ECD and trypsin but not chymotrypsin or pepsin but it is expected not to be the only function of this domain (Beale *et al*. 2015).

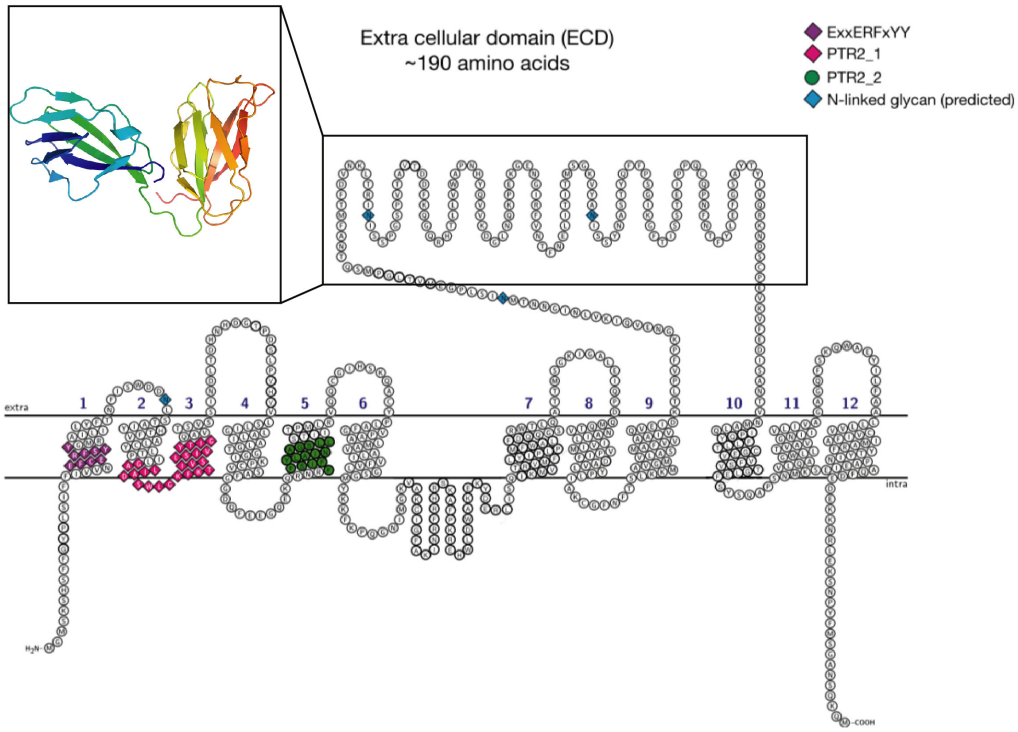


Figure 6: Model of hPepT1 and the crystal structure of PepT1 ECD
The predicted transmembrane helices of hPepT1 are shown with the three signature motives of POTs highlighted in purple, pink and green. The predicted glycosylation sites are in blue. The structure of the extracellular domain of PepT1 from *Mus musculus* (Beale *et al*. 2015) is shown in a box on top left.

There are three conserved sequence motives in POTs, namely the ExxERFxYY motif in TM1, PTR2_1 motif (G***AD***GK**TI***S **Y**G) in the loop between TM2 and TM3 and PTR2_2 motif (*FS*FY*AIN*GSL*) in TM5 (Daniel 2006) (Figure 6). The ExxERFxYY and PTR2_2 motives are involved in ligand binding however the PTR2_1 motif is distant from the ligand-binding site. It is a modified version of the MFS signature motif A and is expected to play a role in the conformational transition between the inward and outward open states (Jiang *et al*. 2013; Zhao *et al*. 2014).

There are to date eight published POT structures, all from bacterial background (Table 2) (Newstead et al. 2011; Solcan et al. 2012; Doki et al. 2013; Guettou et al. 2013; Zhao et al. 2014; Boggavarapu et al. 2015; Parker et al. 2017; Minhas et al. 2018; Quistgaard, Molledo, and Löw 2017). Few of them have been co-crystallized with di-/tripeptides or alafosfalin, a dipeptide mimetic, giving first insights to the ligand promiscuity and selectivity of these peptide transporters. Initially, POTs were

suggested to transport all combinations of amino acids as short peptides, but further studies show that there are differences between POT members in terms of their ligand selectivity. Single amino acids and tetrapeptides are not likely to be ligands and there is stereoselectivity for L-amino acids, although one D-amino acid in the ligand is tolerated in some cases (Daniel 2006). Some of the bacterial POTs have higher affinities towards dipeptides than tripeptides (Martinez Molledo *et al*. 2018), whereas other POTs have higher affinity towards tripeptides as shown in this thesis. The side chains at the N- and C-terminal position of a dipeptide has been shown to play an important role (Weitz *et al*. 2007; Guettou *et al*. 2014). Almost all ligand studies were performed as competition assays and there are scarcely any direct uptake studies, which should be considered in the interpretation of the ligand selectivity in the literature.

Table 2: List of published POT structures

POT/SLC5	Organism	Year	Co-crystallized with
$PepT_{So}$	*Shewanella oneidensis*	2010	Ligand-free
$PepT_{St}$	*Streptococcus thermophilus*	2012	Ligand-free, (AF, AAA)*, (AL, FA, AQ, DE)**, (FAL, FAQ, FAT)**
GkPOT	*Geobacillus kaustophilus*	2013	Ligand-free, Alafosfalin
$PepT_{So2}$	*Shewanella oneidensis*	2013	Ligand-free, Alafosfalin, (AY(Br)A, AY(Br), AAA)***
DtpD	*Escherichia coli*	2014	Ligand-free
YePepT	*Yersinia enterocolitica*	2015	Ligand-free
$PepT_{Xc}$	*Xanthomonas campestris*	2017	Ligand-free
$PepT_{Sh}$	*Staphylococcus hominis*	2018	CG-3M3SH

*published in 2014, ** in 2018, *** published in 2017

All POT structures with dipeptides and alafosfalin share a conserved binding mode. The ligand-binding site is shaped by multiple pockets, which accommodate the side chains of the ligand. Three such pockets have been investigated in further depth for $PepT_{St}$ (Figure 7) (Martinez Molledo *et al*., 2018) and $PepT_{So2}$ (Guettou et al. 2014; Martinez Molledo et al. 2018). The pockets are rich in aromatic residues that can make hydrophilic, hydrophobic or aromatic interactions contributing to the promiscuity of the ligand-binding site. There are several residues involved in the peptide backbone binding such as two Asp (N156, N328 in $PepT_{St}$) and a Glu (E400

in $PepT_{St}$) coordinating the N-terminal end of a dipeptide and an Arg (R26 in $PepT_{St}$) and Tyr (Y30 in $PepT_{St}$) coordinating the peptide bond and the C-terminal carboxy group (Figure 7).

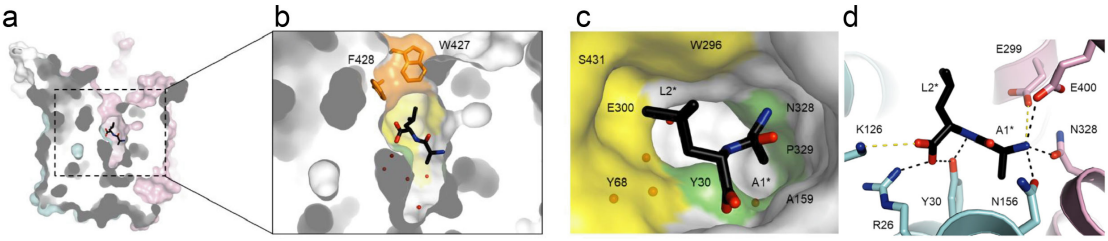


Figure 7: $PepT_{St}$ structure in complex with AlaLeu
(a) Inward open $PepT_{St}$ structure in surface model cut in half perpendicular to the membrane plane with the ligand AlaLeu in sticks model. (b) Zoom into the ligand binding site, the two aromatic residues F428 and W427 which are suggested to have a lid function are highlighted in orange. (c) Surface model of the ligand binding-site with pocket 1, which coordinates the N-terminal residue in green and pocket 2, which coordinates the C-terminal residue in yellow. (d) Residues coordinating the ligand backbone in sticks model and the distances ≥3.2 Å in black and 3.2-4 Å in yellow dashes. Figure from (Martinez Molledo *et al*. 2018).

There are three published studies with tripeptides suggesting two different binding modes. In one of these studies, $PepT_{So2}$ was co-crystallized with AlaTyrAla including a bromide moiety at the tyrosine residue (Guettou *et al*. 2014). The binding mode was identical with a dipeptide, with the N-terminal two residue positions overlapping and the C-terminal residue of AlaTyrAla extending further into the ligand-binding site. In the other study, $PepT_{St}$ was co-crystallized with AlaAlaAla showing a unique vertical binding mode as opposed to the horizontal binding mode of dipeptides (Lyons *et al*. 2014). Nevertheless, further studies in our group with $PepT_{St}$ challenged this model with the observations of a HEPES molecule in an identical position as the AlaAlaAla model during co-crystallization attempts with several ligands (Martinez Molledo *et al*. 2018). $PepT_{St}$ was co-crystallization with three further tripeptides (PheAlaAla, PheAlaGln, PheAlaThr) that agree with the horizontal placement of the tripeptides similar to the dipeptide binding mode (Martinez Molledo, Quistgaard, and Löw 2018). Further functional and structural studies are required for the understanding of di-/tripeptide coordination in POTs and how POTs are promiscuous but retain a level of selectivity towards specific residues. There is also high interest in the investigation of drug binding in POTs for drug development.

1.5. DtpA

Escherichia coli genome has four genes belonging to POT family, named *ydgR*, *ybgH*, *yjdL* and *ydgR*. Correspond proteins were named DtpA to DtpD, Dtp standing for di-/tripeptide permease. DtpA is one of the first and most studied POTs due to its similarity in substrate preference to hPepT1. ydgR gene (DtpA) was shown to be regulated by the EnvZ/OmpR two-component regulatory system, which is involved in the *E. coli* environment stress response (Goh, Siino, and Igo 2004). Unfortunately, no further studies followed the gene regulation or expression profile of DtpA to our knowledge.

Few ligand uptake and competition assays were developed for DtpA to study its ligand selectivity. The most established method is the *in vivo* uptake of AK-AMCA (β-Ala-Lys-N-7-amino-4-methylcoumarin-2-acetic acid) (Weitz *et al*. 2007; Malle *et al*. 2011; Prabhala *et al*. 2014; Prabhala *et al*. 2017). AK-AMCA is a dipeptide mimetic coupled to a fluorescent molecule, which can be also taken up by hPepT1 (Groneberg *et al*. 2001) and hPepT2 (Dieck *et al*. 1999). In this method, cells overexpressing DtpA are incubated with AK-AMCA and an additional substrate for competition assays followed by washing steps. The fluorescence is measured in whole cells, which corresponds to the uptake of AK-AMCA. In case of a good competitor, the AK-AMCA uptake drops to <5%, whereas for poor competitors the AK-AMCA uptake remains close to 100% normalized by the AK-AMCA uptake without a competitor. As examples AlaAla, D-AlaAla, GlyGln, LysAla, AlaAlaAla, cefadroxil and valacyclovir were shown to be good competitors, whereas AspGly, GlyLys, Ala-D-Ala, ampicillin and levadopa are poor competitors (Weitz *et al*. 2007; Prabhala *et al*. 2014). A good competitor is likely to bind and be transported by DtpA, whereas a poor competitor is less likely to bind and be transported by DtpA. From these experiments following conclusions were drawn: (i) DtpA can transport several di- and tripeptides but not all possible combinations. (ii) Combinations of L-amino acids are preferred. A D-amino acid at the N-terminal position of a dipeptide is tolerated (at least in case of D-AlaAla) but a D-amino acid at the C-terminal position of a dipeptide (in case of Ala-D-Ala) is not a good competitor. (iii) A positively charged residue in the N-terminal position of a dipeptide but not in the C-terminal position is an efficient competitor of AK-AMCA. (iv) A negatively charged residue in the C-

terminal position of a dipeptide but not in the N-terminal position is a good competitor of AK-AMCA, although the difference is not as severe as for positively charged residues. Nevertheless, one needs to be careful in the interpretation of the results because it is an indirect method based on competition rate and the direct uptake of a ligand cannot be measured in this method, except for AK-AMCA itself. Furthermore, most conclusions are based on testing very few combinations of residues in the N-terminal and C-terminal positions of a dipeptide.

A new label-free cell based method was developed for DtpA using the pH sensitive dye pyranine (Prabhala *et al*. 2014). Based on the proton-dependent nature of the transport activity, the pH change causes a change in the fluorescence profile of pyranine. Pyranine was added in the buffer and is a cell impermeable chemical. With this method, LysAla, AlaAla and AlaAlaAla were shown to be causing a pH change, which is interpreted as ligand transport. For the dipeptide AlaLys and the amino acid Ala only no change in pH was detected. Additionally, DtpA (and DtpB) showed a pH change with tetraalanine, which was not detected with other methods. The authors suggested that tetraalanine is a low affinity ligand, which cannot compete with AK-AMCA but nevertheless can be transported as observed from the pH change. It remains open, if single amino acids, tetrapeptides and several combinations of di-/tripeptides might be ligands with very low affinity, which were not competing with AK-AMCA.

With a liquid chromatography-mass spectroscopy method (LC-MS) , the direct uptake of six clinical drugs was tested with *E. coli* cells overexpressing DtpA as well as without overexpression (Prabhala *et al*. 2017). From the tested drugs oseltamivir, bestatin, valacyclovir and ampicillin are known to be taken up by hPepT1 but not levodopa (L-DOPA) (Brandsch, Knütter, and Bosse-Doenecke 2008). In the AK-AMCA competition assay only bestatin and valacyclovir were good competitors but all five drugs were uptaken into cells via DtpA according to the LC-MS results and levodopa was not, which supports the similarity between DtpA and hPepT1 in ligand selectivity. Furthermore the direct uptake of radioactively labeled [^{3}H]-Gly-Sar (used also in studies with hPepT1) and [^{3}H]-Ala-Ala were measured with DtpA reconstituted in proteoliposomes (Harder *et al*. 2008; Bippes *et al*. 2013). From these methods, AK-AMCA is the most widely used method due to its relative simplicity, lack of necessity of facilities such as for mass spectroscopy or the need to work with

radioactive compounds. Nevertheless, none of these methods is perfect and there is further need for direct uptake methods and/or methods, which can be scaled up to measure a wider ligand library in parallel.

In a random mutagenesis study, 35 single point mutations were tested with the AK-AMCA uptake assay that led to a full or partial loss-of-function (LOF) (Malle *et al*. 2011). The key findings were (i) Mutation of G78 and G86, two conserved glycines across POTs, led to LOF although they are not in proximity of the ligand-binding site according to the homology model. (ii) Mutation of M295 and F289 led to LOF additional to already characterized residues in proximity of the ligand-binding site. (iii) E56 and R305 single mutations led to LOF, whereas E56R/R305E rescues the LOF showing that E56 and R305 form a salt bridge, which is crucial for the function. Due to their position on a model based on the $PepT_{So}$ structure (only available POT structure at the time of the study), they suggested a periplasmic salt bridge role.

Despite the high interest in DtpA and several functional and mutational studies a high-resolution structure of DtpA has been elusive. The only structural information was the very low resolution negative-stain TEM (transmission electron microscopy) images, which confirmed the monomeric nature of the protein (Weitz *et al*. 2007).

1.6. Valganciclovir and valacyclovir

Herpesviruses (Herpesviridae) are DNA viruses, which are widely spread in humans. There are five species, which infect humans: herpes simplex virus (HSV) 1 and 2, varicella zoster virus, which causes chicken pox and shingles, Epstein-Barr virus and cytomegalovirus (CMV). HSV-1 and -2 cause blisters in the mouth, lips, nose or genitals and the World Health Organization (WHO) estimates that 67 % of people under age of 50 globally have HSV-1 infection (WHO fact sheet on HSV). Although cytomegalovirus infection is often symptomless in healthy people, it can be life threatening for people with immunodeficiency such as HIV patients, organ transplant recipients or newborn infants. A clinical study from USA in 2006 showed that 36 % of children between age 6 and 11 and 90% of all older than 80 years old were infected with CMV (Staras *et al*. 2006). Due to the latent phase of herpesviruses, a complete cure does not exist. Nevertheless, there are a few drugs that inhibit the viral

replication, such as ganciclovir (Fan-Havard, Nahata, and Brady 1989) and acyclovir (Schaeffer *et al*. 1978). Both are acyclic guanosine analogs that are monophosphorylated specifically by viral thymidine kinase to monophosphate derivates and further phosphorylated by host kinases to triphosphate analogues. The triphosphate analogues of guanosine inhibit the viral DNA polymerases, block the DNA chain elongation and thereby limit the virus replication. Acyclovir is highly selective for HSV-1, -2 and varicella zoster virus but is not very effective against CMV. Ganciclovir is effective against HSV-1, -2 and CMV, although due to the higher toxicity than acyclovir it is mainly used against CMV (De Clercq and Field 2006). A common problem with both drugs was the poor water solubility and very low uptake rates via oral administration.

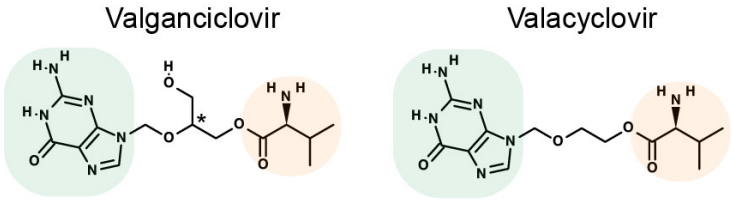


Figure 8: Chemical structures of valganciclovir and valacyclovir.
The guanine ring is highlighted with a green background and the N-terminal valine with an orange background. Both drugs have a chiral center at the Cα of L-valine and valganciclovir a second chiral center marked with an asterisk at the C2-position.

L-valine esters of acyclovir and ganciclovir, namely valacyclovir and valganciclovir, were shown to improve the water solubility and uptake rate remarkably. The uptake rate of valacyclovir improved from 19 % to 60 %, and from 6 % to 60 % in case of valganciclovir (Jung and Dorr 1999). This increase was linked to hPepT1, which is highly expressed in the small intestines and transports valacyclovir and valganciclovir but not acyclovir or ganciclovir into the intestinal cells (Guo *et al*. 1999; Sugawara *et al*. 2000). Valacyclovir and valganciclovir is rapidly converted to acyclovir or ganciclovir and L-valine by cellular esterases. Valganciclovir has in addition to the chiral center at Cα of L-valine a chiral center at the C2 position (marked with an asterisk in Figure 8). The drug is a mixture of both diastereomers, both of which can be transported by hPepT1 and cleaved by the host esterases. The chiral center is lost after the cleavage.

1.7. Aims of the study

The peptide transporter DtpA has been in the focus of the POT field due to its similarity to hPepT1 in ligand selectivity and its ability of drug transport. There is a high interest in hPepT1 because of its physiological role in nutrient uptake and pharmacokinetic role in the transport of a wide variety of peptidomimetic antibiotics, antiviral, anticancer and antiparkinson drugs/prodrugs in the small intestines. Due to several glycosylation sites and the low protein stability, hPepT1 is highly challenging to work with, but bacterial homologue DtpA from *E. coli* turned out to be an excellent model system. There have been a few POT structures published (Table 2), some of which were co-crystallized with di-/tripeptides or alafosfalin in the ligand binding site, but most of these transporters have been shown not to transport any drugs.

To elucidate drug transport in the POT/SLC15 family of peptide transporters and characterize DtpA in further depth, the aims of this thesis were to:

- Characterize the ligand preference of DtpA with a ligand library systematically using *in vitro* and *in vivo* assays
- Determine the structure of DtpA with and without a drug
- Investigate the difference between published POT structure with the aim to understand why DtpA (and hPepT1/2) can transport drugs, whereas other bacterial members cannot

2. Materials and Methods

2.1. Materials

2.1.1. Chemicals and consumables

Table 3: Table of chemicals

Chemicals	Producer	Catalog number
1-(7Z-pentadecenoyl)-rac-glycerol (7.8 MAG)	Avanti Polar Lipids	#850531O
1,4-Dithiothreit (DTT)	Roth	#6908.2
Agarose for DNA electrophoresis	SERVA	#11404.07
Ampicillin sodium salt	Roth	#K029.3
β-AK-AMCA	Biotrend	#BP0352
Calcium chloride	Roth	#5239.2
cOmplete, EDTA-free protease inhibitor cocktail	Roche	#5056489001
D(+)-Glucose	Roth	#X997.2
D(+)-Sucrose	Roth	#4621.1
di-Sodium hydrogen phosphate heptahydrate	Roth	#X987.2
Dimethyl sulfoxide (DMSO)	Sigma Aldrich	#D2650
DNAse I	AppliChem	#A3778
Ethanol	Roth	#9065.3
Ethidium bromide solution, 0.025 % (w/v)	Roth	#HP47.1
Glucose	Roth	#X997
Glycerol	Roth	#3783.1
Glycine, Pufferan	Roth	#3908.2
H-p-Bz-Phe-OH (pBpF)	Bachem	#4017646
HEPES, Pufferan >99.5 %	Roth	#9105.3
Imidazole	Roth	#X998.4
Isopropyl β-D-1-thiogalactopyranoside (IPTG)	Roth	#2316.5
Kanamycin sulphate	Roth	#T832.4
LB Agar (Lennox)	Roth	#X965.2
LB Broth Low Salt Granulated	Melford	#GL1703
Magnesium chloride hexahydrate	Roth	#2189.1
n-Dodecyl-β-D-Maltopyranoside (DDM), anagrade	Anatrace	#D322LA
n-Dodecyl-β-D-Maltopyranoside (DDM), sol-grade	Anatrace	#D310S
Nickel(II) chloride hexahydrate ≥ 98 %, p.a	Roth	#4489
Polyethylen glycol 400 (PEG 400)	Fluka	#91893
Polyethylen glycol 600 (PEG 600)	Fluka	#87333
Sekusept Plus	Ecolab	#104372E
Sodium chloride, >99,5%, p.a., ACS, ISO	Roth	#3957
Sodium dihydrogen phosphate monohydrate	Roth	#K300.2
Sodium hydroxide ≥ 99	Roth	#9356

Sodium phosphate dibasic dodehydrate	Sigma-Aldrich	#71649
Super signal West pico luminol/enhancer solution	Thermo Scientific	#1856136
Super signal West pico luminol/enhancer solution	Thermo Scientific	#1856135
TBE Buffer 10x, Rotiphorese	Roth	#3061.2
TBST	Sigma Aldrich	#91414
Terrific Broth (TB, modified), granular	Melford	#GT1702
Tris Acetate-EDTA (TAE) buffer 10X	Sigma-Aldrich	#T8280
Tris-(hydroxymethyl)-aminomethane (TRIS)	Roth	#5429.3

Table 4: Ligand library

Category	Ligand	One letter code	Producer	Catalog number
Dipeptide	Ala-Ala	AA	Sigma	#A502
Dipeptide	Ala-Asp	AD	Sigma	#A9627
Dipeptide	Ala-Gln	AQ	Sigma	#G8541
Dipeptide	Ala-Glu	AE	Bachem	#4002318
Dipeptide	Ala-Leu	AL	Sigma	#A1878
Dipeptide	Ala-Lys	AK	Bachem	#G-1290
Dipeptide	Ala-Phe	AF	Sigma	#A3128
Dipeptide	Asp-Glu	DE	Sigma	#A1916
Dipeptide	Glu-Glu	EE	Sigma	#G3640
Dipeptide	Gly-His	GH	Sigma	#G1627
Dipeptide	Gly-Ser	GS	Sigma	#G-3127
Dipeptide	His-Ser	HS	Sigma	#H3129
Dipeptide	Leu-Ala	LA	Bachem	#G-2460
Dipeptide	Leu-Leu	LL	Bachem	#M-1535
Dipeptide	Lys-Ala	KA	Bachem	#G-2630
Dipeptide	Met-Ser	MS	Sigma	#M9380
Dipeptide	Phe-Ala	FA	Bachem	#G-2850
Dipeptide	Thr-Gln	TQ	Sigma	#T3275
Tripeptide	Ala-Ala-Ala	AAA	Bachem	#H-1445
Tripeptide	Ala-Ala-Tyr	AAY	Bachem	#G-1445
Tripeptide	Ala-Leu-Ala	ALA	Bachem	#H-5975
Tripeptide	Ala-Phe-Ala	AFA	Bachem	#H-5420
Tripeptide	Ala-Pro-Ala	APA	Bachem	#H-1595
Tripeptide	Gly-Gly-His	GGH	Sigma	#G4541
Tripeptide	Leu-Gly-Gly	LGG	Sigma	#L9750
Tripeptide	Leu-Leu-Ala	LLA	Bachem	#H-3905
Tripeptide	Met-Ala-Ser	MAS	Sigma	#M1004
Tripeptide	Phe-Ala-Ala	FAA	GL Biochem	
Drug	5-Aminolevulinic acid		Sigma	#A3785
Drug	Midodrine		Sigma	#M8277
Drug	Palmidronate		Sigma	#P2371
Drug	Penicillin G sodium salt		Sigma	#13752
Drug	Valacyclovir		Sigma	#PHR1601
Drug	Valganciclovir		Fluka	#PHR 1626

Table 5: Enzymes and commercial kits

Enzymes and commercial kits	Producer	Catalog number
DNase I recombinant	Roche	#4536282001
DpnI restriction enzyme	New England BioLabs	#R0176S
Lysozyme	Roth	#8259.2
Mix & Go! *E. coli* Transformation Kit	Zymo Research	#T3001
Monolith His-Tag Labeling Kit RED-tris-NTA	Nanotemper	#MO-L008
Phusion High-Fidelity DNA Polymerase	New England BioLabs	#M0530S
QIAquick Gel Extraction Kit	Qiagen	#28706
QIAquick Miniprep Kit	Qiagen	#27104
QIAquick PCR Purification Kit	Qiagen	#28106
T4 DNA Ligase	New England BioLabs	#M0202S
T4 DNA Polymerase	New England BioLabs	#M0203S
T4 Polynucleotide Kinase	New England BioLabs	#M0201S
Taq DNA Polymerase	New England BioLabs	#M0267S

Table 6: Table of consumables

Consumables	Producer	Catalog number
CaptureSelect beads for EPEA-tag	ThermoFisher Scientific	#194288010
Monolith NT.115 Capillaries	Nanotemper	#MO-K022
Monolith NT.LabelFree Standard Capillaries	Nanotemper	#MO-Z022
Ni-NTA Agarose	Invitrogen	#R901-15
NuPAGE® LDS Sample Buffer (4X)	ThermoFisher Scientific	#NP008
NuPAGE® Novex® 4-12% gels	ThermoFisher Scientific	#NP0322
Prometheus NT.48 Standard Capillaries	Nanotemper	#PR-C002
Roti®-Mark 10-150 Protein-Marker	Roth	#T850
Spin-X® UF 100 kDa concentrator	Corning	#431491
Spin-X® UF 5 kDa concentrator	Corning	#431487
Spin-X® UF 50 kDa concentrator	Corning	#431490

Table 7: Commercial crystallization screens

Crystallization screens	Method of crystallization	Supplier	Catalog no.
MemGold_HT-96	vapor diffusion	Molecular Dimensions	MD1-41
MemGold2_HT-96	vapor diffusion	Molecular Dimensions	MD1-64
Wizard Cubic LCP Screen	LCP	Rigaku Reagents, Inc	1008650
MemMeso HT-96	LCP	Molecular Dimensions	MD1-87
Additive_screen	additive screen	Hampton Research	HR2-428
MemAdvantage	additive screen	Molecular Dimensions	MD1-70

2.1.2. Devices

Table 8: Table of devices

Device	Producer
AKTA Pure with fraction collector F9C	GE Healthcare
Analytical gel filtration system	Agilent Technologies
Centrifuge Avanti JXN-26	Beckman-Coulter
Electrophoresis chamber for agarose gels	NeoLab
Electrophoresis chamber for SDS-PAGE gels	Invitrogen
EmulsiFlex-C3 cell homogenizer	Avestin
Gel imaging and documentation system	Bio-Rad
HiLoad 16/600 Superdex200 (gel filtration column)	GE Healthcare
HiLoad 16/600 Superdex75 (gel filtration column)	GE Healthcare
JLA 8.1000 rotor	Beckman-Coulter
Monolith NT.115 microscale thermophoresis	Nanotemper
Monolith NT.LabelFree microscale thermophoresis	Nanotemper
Mosquito-LCP	TTP Labtech
New BrunswickTM Innova® 42 small shaking incubator	Eppendorf
New BrunswickTM Innova® 44 shaking incubator	Eppendorf
PCR cycler	Eppendorf
Power supply	Consort
Prometheus NT.48 nano differential scanning fluorimetry	Nanotemper
Rock Imager	Formulatrix
Scorpion dispenser	Art Robbins Instruments
Superdex200 Increase 10/300 GL (gel filtration column)	GE Healthcare
Thermoblock	Eppendorf
Ti 45 rotor	Beckman-Coulter
Ultracentrifuge Optima XE-90	Beckman-Coulter
Ultracentrifuge tubes (polycarbonate, 70 ml tubes)	Beckman-Coulter

2.2. Methods

2.2.1. Molecular Biology methods

2.2.1.1. Transformation

50 µL competent cells were mixed with 2 µL of plasmid and incubated 15 min on ice. The sample was introduced to 42 °C for 45 s followed by addition of 600 µL of LB buffer. The cells were incubated for another 15 min on ice. Cells with ampicillin-resistant plasmids were directly plated out on LB-Agar plates supplemented with 0.1 mg/ml ampicillin. For kanamycin-resistant plasmids, one-hour incubation at 37 °C

shaking at 220 rpm was included before plating on LB-Agar plates supplemented with 0.1 mg/ml kanamycin. Colonies grew on the plates overnight at 37 °C.

2.2.1.2. Colony PCR

Colony PCR reaction was performed with the components listed in Table 9 mixed with a colony that grew on the agar plate with the settings as in Table 10. This method was used at the final stages of a cloning experiment to test the size of a band in gel electrophoresis as described in Methods section 2.2.1.3. Colonies with the right size were sent for sequencing to confirm the sequence of the construct as further explained in Methods section 2.2.1.5.

Table 9: Colony PCR components

	Volume for 1 reaction [μL]
5 x Thermo buffer	2.5
10 mM dNTPs	0.5
10 μM forward primer	0.5
10 μM reverse primer	0.5
Taq polymerase	0.125
fill up with water to	25

Table 10: Settings of the colony PCR

Step	Temperature [°C]	Time	
1	95	1 min	
2	95	30 s	
3	55	30 s	
4	68	1 min	Go to step 2 for 30 cycles
5	68	5 min	
6	4	hold	

2.2.1.3. Agarose gel preparation

1 % Agarose gel was prepared with TBE (TRIS/Borate/EDTA) buffer and a drop of EtBr (ethidium bromide). After polymerization and loading with the DNA samples, the gel was placed in the electrophoresis chamber, filled with 1x TBE buffer and 110 V was applied for 45 min. The gel was imaged with ChemiDoc MP (Bio-Rad).

2.2.1.4. Plasmid preparation

Plasmid preparation was performed with the QIAprep Spin Miniprep Kit (Qiagen) according to the protocol of the manufacturer.

2.2.1.5. Sequencing

Samples were sent to Eurofins (Hamburg, Germany) for sequencing and the T7 forward and T7 reverse primers were used for DtpA wild-type and mutants. For the nanobodies the MP57 and GIII primers (Pardon et al. 2014) were used to sequence the constructs from the N-terminal or C-terminal end.

Table 11: Primers for sequencing

	Primer sequence
T7 forward	5'-TAATACGACTCACTATAGGG-3'
T7 reverse	5'-CTAGTTATTGCTCAGCGGT-3'
MP57 (forward)	5'-TTATGCTTCCGGCTC GTATG-3'
GIII (reverse)	5'-CCACAGACAGCCCTCATAG-3'

2.2.1.6. Blunt-end PCR for site-directed mutagenesis

To introduce site-directed point mutations in DtpA, blunt-end PCR method was used. A set of primers was designed for each mutant, which had the 5'-ends on the aimed site of the point mutation. In one of the primers the modified nucleotides for the mutation were included (highlighted in red in Table 12), the other primer was adjacent to that site. The primers were designed in reverse direction to span the whole vector. The components in the Table 13 were mixed on ice and a polymerase chain reaction (PCR) was performed with the settings as in Table 14.

Table 12: Primers for site-directed mutagenesis of DtpA

Mutant	Direction	Primer sequence
Y38A	forward	5'-GCATAACCAAAACGTTCC-3'
	reverse	5'-GGGCCTACAAGGAATTATG-3'
Y71A	forward	5'-GCAACCAGGGCACTAAAG-3'
	reverse	5'-GGGTCTGGTCGCTATCG-3'
K130A	forward	5'-GCAAACAGGCCGTTACC-3'
	reverse	5'-GGCTAACCCGTCTTCTCTG-3'
N160A	forward	5'-GACGGACATGTAGTACATGGTGAATG-3'
	reverse	5'-GCCATCGGCTCTTTCTTCTCTA-3'
I399A	forward	5'-CATCAGTTCCCCGATG-3'
	reverse	5'-GCGTCTGGTCTGGGT-3'

The changed nucleotides are highlighted in red for introducing the aimed mutation.

Table 13: Blunt-end PCR components

	Volume for 1 reaction [μL]
5 x Phusion buffer	5
10 mM dNTPs	0.5
10 uM fwr primer	1.25
10 uM rev primer	1.25
Phusion polymerase	0.25
DNA template	1
fill up with water to	25

Table 14: Settings of the blunt-end PCR

Step	Temperature [°C]	Time	
1	98	1 min	
2	98	30 s	
3	60	30 s	
4	72	5 min	go to step 2 for 40 cycles
5	98	10 min	
6	4	hold	

The PCR product was inspected via gel electrophoresis (Methods section 2.2.1.3), mixed with 1 μL DpnI and incubated for one hour at 37 °C followed by purification with the QIAquick PCR purification kit according to manufacturer's protocol. Then, 30 μL of PCR product was mixed with 4 μL 10 x Polynucleotide ligase buffer, 1 μL T4 Polynucleotide kinase and filled up to 40 μL with water. After two hours of incubation at 37 °C, it was again purified with a QIAquick PCR purification kit. Next, 30 μL phosphorylated PCR product was mixed with 4 μL of 10x Ligation buffer, 2 μL of T4 DNA ligase and filled up to 40 μL. The ligated vector was transformed into *E. coli* DH5α cells as described in Methods section 2.2.1.1. Several colonies for each construct were used for colony PCR method as described in Methods section 2.2.1.2 to check for the correct length of the vector and sent for sequencing to confirm the sequence (Methods section 2.2.1.5).

2.2.2. Protein expression and purification methods

2.2.2.1. DtpA expression and purification

2.2.2.1.1. Required buffers and stocks for expression

- Plasmid of DtpA in pTH27 vector
- *E. coli* C41 (DE3) competent cells
- Lysogeny broth (LB) medium
- Terrific broth (TB) medium
- Ampicillin stock solution, 100 mg/ml
- Isopropyl β-D-1-thiogalactopyranoside (IPTG) stock solution, 1 M

2.2.2.1.2. Required buffers and stocks for purification

- Lysis buffer: 20 mM NaP pH 7.5, 300 mM NaCl, 5 % glycerol, 15 mM imidazole
- cOmplete, Mini, EDTA-free Protease Inhibitor cocktail tablets
- Lysozyme stock solution, 100 mg/ ml
- TCEP stock solution, 0.5 M
- DTT stock solution, 0.5 M
- DNAse stock solution, 2000 U/ ml
- DDM, sol-grade powder
- DDM stock solution, 3 %
- Wash buffer 1: 20 mM NaP pH 7.5, 300 mM NaCl, 5 % glycerol, 15 mM imidazole, 0.03% DDM, 0.5 mM TCEP
- Wash buffer 2: 20 mM NaP pH 7.5, 300 mM NaCl, 5 % glycerol, 40 mM imidazole, 0.03% DDM, 0.5 mM TCEP
- Elution buffer: 20 mM NaP pH 7.5, 150 mM NaCl, 5 % glycerol, 300 mM imidazole, 0.03% DDM, 0.5 mM TCEP
- Gel filtration buffer: 20 mM HEPES, pH 7.5, 150 mM NaCl, 5 % glycerol, 0.5 mM DTT, 0.03% DDM
- TEV protease stock solution, 1 mg/ml, stored at -80 °C

2.2.2.1.3. Expression and purification protocol

DtpA, wild type (WT) and mutants, were expressed and purified as previously described (Flayhan et al. 2018). The full-length dtpA (Uniprot ID: P77304) gene was amplified from the *E. coli* genome and cloned into the pTH27 vector (Woestenenk et al. 2004). The construct has an N- and C-terminal His-tag and a TEV cleavage site after the N-terminal tag. The C-terminal tag cannot be cleaved. 50 µL of competent *E. coli* C41 (DE3) cells were transformed with 2 µL of DtpA plasmid (Methods section 2.2.1.1) and plated on a LB agar plate supplemented with 0.1 mg/ml ampicillin. 5 mL LB medium supplemented with 0.1 mg/ml ampicillin was inoculated with a colony

from the plate and incubated at 37 °C shaking at 220 rpm. After five to seven hours, the culture was extended to 50 mL and incubated at 37 °C shaking at 220 rpm overnight. 1 L TB supplemented with 0.1 mg/ml ampicillin was inoculated with the overnight culture for a starting OD_{600nm} in 2 L baffled bottles and incubated at 37 °C shaking at 220 rpm until the OD_{600nm} reached 0.7. The bottles were moved to room temperature for 30 min and the shakers cooled to 18 °C. The DtpA expression was induced with 0.2 mM IPTG and the incubation continued at 18 °C shaking at 220 rpm overnight. The cells were harvested by centrifugation at 10,000 xg for 12 min and the pellet stored at -20 °C until purification.

For purification, the cell pellet was resuspended in lysis buffer with 5 units/ml DNase, 1 tablet of proteinase inhibitor/100 mL, 1 mg/ml lysozyme and 0.5 mM TCEP. 3 to 5 mL of lysis buffer was used per 1 g of cell pellet. Cell lysis was performed with an emulsifier (EmulsiFlex-C3, Avestin) in three passages at 10,000-15,000 psi pressure. All following steps were performed at 4 °C. The lysate was centrifuged 12 min at 10,000 xg to remove the unlysed cells. The supernatant (SN) was ultracentrifuged 50 min at 35,000 rpm using the rotor Ti45 (95,000 xg) (Optima XE-90, Beckman Coulter). The pellet containing the membrane fraction was solubilized with a potter resuspending the pellet in equal volume of lysis buffer with 1 tablet of proteinase inhibitor/100 mL and 0.5 mM TCEP. The suspension was incubated with 1 % (m/v) dodecylmaltoside (DDM) (Anatrace, sol-grade) mixing for 1 h to solubilize the membrane proteins. Then, it was ultracentrifuged 50 min at 30,000 rpm using the same rotor. The SN was applied to Ni-IMAC beads (ThermoFischer), which were equilibrated in wash buffer 1. For 1 L cell culture, 1 gravity column filled with 2 mL Ni-IMAC beads were used. SN with beads was incubated for 45 min on a rotating wheel and then was applied on a gravity column. The beads were washed twice with 12 mL wash buffer 1 per column, followed by washing twice with 12 mL wash buffer 2 per column. DtpA was eluted by adding 8 mL elution buffer for each column and combining the elution fractions in one tube. 0.5 mL TEV protease (1mg/ml) was added to the elution and the cleavage was performed in a dialysis bag in 0.5 L gel filtration buffer under light stirring overnight. DtpA was concentrated using a 50 kDa-concentrator (Corning Spin-X UF concentrators) by 12 min centrifugation steps to 5 mL and loaded on a HiLoad 16/600 Superdex 200 pg column (GE Healthcare) in AKTA system, which was previously equilibrated with the gel filtration buffer. The

fractions corresponding to the DtpA peak were collected and concentrated to 11 mg/ml using a 50 kDa cut-off concentrator. Part of the protein sample was used fresh for crystallization and the rest was flash-frozen and stored at -80 °C until further use.

2.2.2.2. Nanobody expression and purification

2.2.2.2.1. Required buffers and stocks for expression

- Plasmid of nanobody in pMESy4 vector
- *E. coli* WK6 competent cells
- Lysogeny broth (LB) medium
- Terrific broth (TB) medium
- Ampicillin, 100 mg/ml stock solution
- Isopropyl β-D-1-thiogalactopyranoside (IPTG), 1 M stock
- Glucose stock solution (filtered), 20 %
- $MgCl_2$ stock solution (filtered), 2 M

2.2.2.2.2. Required buffers and stocks for purification

- TES buffer: 0.2 M TRIS, pH 8, 0.5 mM EDTA, 0.5 M sucrose
- TES/4 buffer: 4-times diluted TES buffer with water
- cOmplete, Mini, EDTA-free Protease Inhibitor cocktail tablets
- Wash buffer: 20 mM NaP, pH 7.5, 20 mM NaCl
- Elution buffer: 20 mM HEPES, pH 7.5, 1.5 M $MgCl_2$
- Gel filtration buffer: 20 mM HEPES, pH 7.5, 150 mM NaCl, 5 % glycerol

2.2.2.2.3. Expression and purification protocol

Nanobodies were generated by Jan Steyaert and Els Pardon at the Vrije University in Brussels with the provided DtpA sample, which was purified as described in Methods section 2.2.2.1. I have received the nanobodies in pMESy4 vector and the construct included a periplasmic signaling peptide and a C-terminal EPEA-tag (GluProGluAla-tag). The expression and purification of nanobodies was performed as described in Pardon *et al*. (Pardon et al. 2014). 50 μL of *E. coli* WK6 cells were transformed with 2 μL of nanobody plasmid and plated on a LB-Agar plate supplemented with 0.1 mg/ml ampicillin. 5 mL LB medium supplemented with 0.1 mg/ml ampicillin was inoculated with a colony from the plate and incubated at 37 °C shaking at 220 rpm. After five to seven hours, the culture was expanded to 50 mL and incubated at 37 °C shaking at 220 rpm overnight. 500 mL TB supplemented with 0.1 mg/ml ampicillin,

0.1 % glucose and 1 mM $MgCl_2$ was inoculated with the overnight culture for a starting OD_{600nm} of 0.05 in a 2 L baffled bottle and was incubated at 37 °C shaking at 220 rpm until the OD_{600nm} reached 0.7. The protein expression was induced with 1 mM IPTG and the incubation continued at 28 °C shaking at 220 rpm overnight. The cells were harvested by centrifugation at 10,000 xg for 12 min and the pellet stored at -20 °C until purification.

The cells were resuspended by stirring the cell pellet with TES buffer supplemented with 1 proteinase inhibitor pellet for 1 h at 4 °C. As a rule of thumb, 3 to 5 mL TES buffer was used per 1 g cell pellet. As the nanobodies were expressed into the periplasmic space, only the outer cell membrane was disrupted by osmotic shock adding two times the volume as in the previous step of TES/4 buffer and stirred for further 45 min. The intact cells were removed by centrifugation for 20 min at 10,000 xg. The SN containing the nanobodies was ultracentrifuged 30 min at 40,000 xg. The SN thereof was applied to a column packed with 10 mL CaptureSelect beads (ThermoFisher Scientific) via sample pump, which were pre-equilibrated with wash buffer in AKTA system using a flow rate of 2.5 ml/min. The column was washed with 5 column volumes (CV) of wash buffer and 5 % elution buffer at a flow rate of 4 ml/min, which was kept constant for all further steps. The nanobody was eluted with 8 CV elution buffer via a gradient from 5 to 100 % elution buffer. The fractions corresponding to the nanobody peak were collected and concentrated using a 5 kDa cut-off to 5 mL, which was injected to HiLoad 16/600 Superdex 75 pg column on AKTA system that was previously equilibrated with gel filtration buffer. The fractions of nanobody were collected and concentrated with a 5 kDa cut-off to 13 mg/ml. Finally, they were aliquoted in 50 or 100 µL batches, flash-frozen and stored at -80 °C until further use.

2.2.2.3. DtpA purification coupled to N00

2.2.2.3.1. Required buffers and stocks for purification

- Gel filtration buffer: 20 mM HEPES, pH 7.5, 150 mM NaCl, 5 % glycerol, 0.5 mM DTT, 0.03 % DDM
- Elution buffer: gel filtration buffer with 2 M $MgCl_2$
- TEV protease stock solution, 1 mg/ml, stored at -80 °C

2.2.2.3.2. Purification protocol

For the purification of DtpA coupled to N00 the same buffers and procedure were used as described in Methods section 2.2.2.1 for DtpA and 2.2.2.2 for N00 except the modified steps as written below.

The DtpA purification was performed until the affinity purification step. Instead of the IMAC purification, the DtpA sample was applied to CaptureSelect beads, which were pre-loaded with N00 and washed with gel filtration buffer for 2 CV. After the application of DtpA, the column was washed with gel filtration buffer for 5 CV and eluted in 2 CV with 75 % elution buffer. The fractions corresponding to the DtpA-N00 complex peak were collected. 0.5 mL of TEV protease (1 mg/ml) was added to the collected fractions and dialysed overnight in gel filtration buffer. Afterwards, the sample was concentrated using a 100 kDa cut-off to 5 mL and injected to HiLoad 16/600 Superdex 200 pg column on AKTA system that was previously equilibrated with the gel filtration buffer. The fractions of the DtpA-N00 complex were collected and concentrated to 10 mg/ml using a 100 kDa concentrator. Part of the protein sample was used directly for crystallization and the rest was flash-frozen and stored at -80 °C until further use.

2.2.3. Biophysical techniques

2.2.3.1. Analytical gel filtration assay

Complex formation of DtpA and the nanobodies were studied with analytical gel filtration assay, which was performed with a Superdex 200 5/150 GL home-packed column (GE Healthcare) connected to the 1260 Infinity liquid chromatography system (Agilent technologies) at 10 °C. The concentration of the DtpA was 0.2 mg/ml and the nanobodies were in 1:1.2 molar ratio to DtpA. Runs of individual proteins as well as DtpA-nanobody mixtures were performed in a row. First, DtpA and one nanobody were mixed and measured and then DtpA was mixed with combinations of two nanobodies. The column was equilibrated with two column volumes of gel filtration buffer followed by the injection of 20 µL of sample. UV absorption and the fluorescence (excitation at 280 nm and emission at 350 nm) were recorded. All

samples were run in duplicates and the data analysis was performed in GraphPad Prism 5.0.

2.2.3.2. Western blot analysis of DtpA-nanobody binding mode

To investigate if nanobodies generated against DtpA are conformational or linear binders, a western blot method was developed. The principle of this approach is shown in the Figure 9. 0.3 and 0.03 μg DtpA, 0.3 and 0.03 μg DtpC (another POT member from *E. coli*) as negative control and 15 μg nanobody 00 (N00) as positive control were loaded on a 4-12 % Bis-TRIS gel (ThermoFisher) and electrophoresis under reducing conditions was performed. The protein was transferred to a polyvinylidene difluoride (PVDF) membrane. 1% bovine serum albumin in TBS-T (TRIS buffer with Tween20, from Sigma) was used for blocking and TBS-T was used as washing buffer. The membrane was incubated 1 h with one of the nanobodies N87, N89, N93 or N00 that contain an EPEA tag at 1 mg/ml concentration, diluted in TBS-T buffer. After washing the membrane three times for five minutes, anti-EPEA antibody coupled to biotin (ThermoFisher) was used as primary antibody, the membrane washed three times for five minutes and streptavidin coupled to horseradish peroxidase (ThermoFisher) applied as secondary antibody. The blot was developed using Super Signal West Pico Substrate (ThermoFisher) and Super Signal West Femto Substrate (ThermoFisher) in 1:10 ratio. The image was acquired with a ChemiDoc MP device (Bio-Rad).

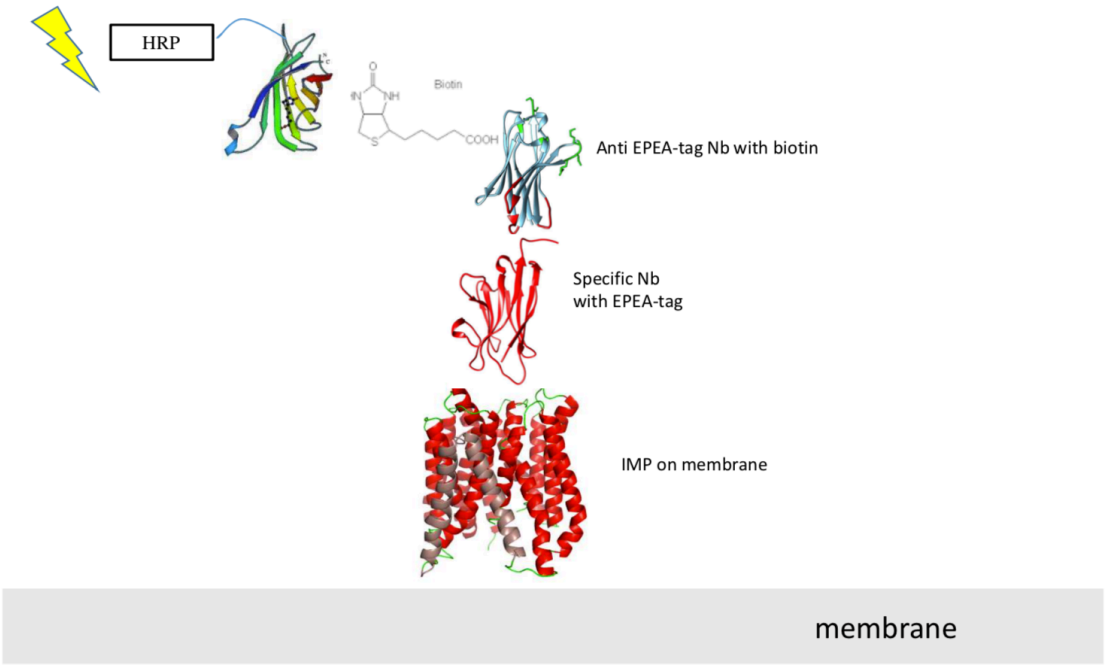


Figure 9: The evaluation of linear or conformational binders via western blot
The integral membrane protein (IMP) was transferred to a PVDF membrane (represented in grey) after SDS-PAGE. First, the membrane was incubated with a nanobody that contains the EPEA-tag. Then, an anti-EPEA-tag antibody coupled to biotin was applied as primary antibody and streptavidin coupled to horseradish peroxidase as secondary antibody. The blot was developed by Super Signal West Pico Substrate and Super Signal West Femto Substrate and imaged with a ChemiDoc MP device.

2.2.3.3. Thermal stability analysis

Nano differential scanning fluorimetry (nanoDSF) was used to assess the thermal stability of proteins and test possible protein-protein and protein-ligand interactions. In this method, the intrinsic signal from the proteins, mostly from tryptophan residues, is followed along a temperature scan. Upon thermal unfolding event, we can detect the change in the fluorescence signal from the tryptophan residues. The fluorescence at 330 nm and 350 nm is recorded over a temperature gradient scan from 20 to 90 °C with 1 °C per minute. From the ratio of 330 nm to 350 nm fluorescence signal plotted against the temperature, the thermal stability is read at the transition point. A shift in the thermal stability in the presence of an additional protein or ligand is interpreted as potential binding. The DtpA concentration was 0.5 mg/ml throughout all experiments, except stated otherwise. For the protein-protein interaction measurements, DtpA was mixed with nanobody in 1:1.2 molar ratio and incubated 1 h prior to measurements. For the protein-ligand interaction measurements, 18 µL of DtpA was mixed with 2 µL

of ligand (final ligand concentration was 5 mM). 10 µL of the solution was filled into a standard capillary (Nanotemper), which is placed into the Prometheus NT.48 device (Nanotemper) for the measurement using 20 % LED and 20 % MST power. The results were analyzed in GraphPad Prsim 5.0.

2.2.3.3.1. RUBIC screen

The thermostability of DtpA and DtpA-N00 complex was analyzed via nanoDSF method (Materials section 2.2.3.3) under different pH, buffer and salt conditions using the commercial RUBIC screen (Molecular Dimensions) (Boivin, Kozak, and Meijers 2013). RUBIC screen was described initially for the thermofluor assay but is also compatible with the nanoDSF method. The conditions included in the RUBIC screen are listed in Appendix Table 17. 21 µL of RUBIC screen was mixed with 2 µL of DtpA (final concentration 0.2 mg/ml) and 2 µL DDM (final concentration 0.03 %), which was important to have enough detergent in the solution for the membrane protein integrity. The samples were loaded to the Prometheus NT.48 device and measured at 30 % LED and 20 % MST power. The results were analyzed in GraphPad Prism 5.0.

2.2.3.4. Binding affinity measurements

2.2.3.4.1. Microscale Thermophoresis (MST) without labeling

The binding affinity of DtpA to the di-/tripeptides was measured by microscale thermophoresis using the Monolith NT.Labelfree device (Nanotemper). This method is based on the detection of intrinsic fluorescence of proteins mostly from the tryptophan residues within the protein. To measure the binding affinity, a concentration series of the ligand is mixed with the protein in glass capillaries. Infrared laser heats a spot within the capillary and the movement of the protein is tracked by fluorescence at 330 and 350 nm. The motion of the protein upon heat is dependent on its size, charge and hydration shell, which changes upon ligand binding. Plotting the ligand concentration against the fluorescence ratio at 330 nm to 350 nm at a given time point gives a ligand-binding curve and from the transition point a value equivalent to dissociation constant (K_d) is withdrawn. The advantage of the technique is that it requires a small volume of the protein and ligand as well as a very low concentration of the protein (in nM range); also, there is no need for labeling.

DtpA was initially diluted to 0.5 μM with assay buffer (100 mM HEPES, pH 7.5, 150 mM NaCl, 0.03 % DDM). A dilution series of the substrate was prepared with the same buffer. 10 μL of DtpA was then mixed with 10 μL of the substrate dilution series and incubated for 10 min at room temperature before loading the samples in standard capillaries to the MST NT.LabelFree device. For the measurements with the DtpA-N00 complex, the concentration of DtpA was kept constant and a 1:1.2 molar ratio of N00 was added to DtpA and incubated for 1 h. As a negative control, the binding affinity of N00 alone, at 2 μM concentration, to selected peptides was measured under same conditions. The settings were 20 % LED and 20 % MST power. All measurements were performed in duplicates and the results were analyzed with GraphPad Prism 5.0.

2.2.3.4.2. Microscale Thermophoresis (MST) with labelling

The limiting factor of the label-free MST method is that some ligands contribute to the signal measured at 330 nm and 350 nm wavelengths, which prevents from detecting the binding effect, as it was the case for the drugs valganciclovir and valacyclovir. In these cases, the label-free method cannot be used and as an alternative, one can label the protein and use another wavelength for the measurements. DtpA has an uncleavable His-tag at the C-terminus, which was used for the labeling with the RED-tris-NTA kit (Nanotemper) according to the protocol of the manufacturer with small changes. The first change was the used buffer; instead of the advised buffer the assay buffer containing the detergent DDM (100 mM HEPES, pH 7.5, 150 mM NaCl, 0.03 % DDM) was used in all steps and second, the concentration of the protein was increased in relation to the dye due to the low binding affinity of the dye to the protein.

DtpA was diluted to 1.6 μM and labeled with RED-tris-NTA dye according to the supplier's protocol but with mentioned changes. 10 μL of the labeled DtpA was mixed with 10 μL of the ligand and incubated 10 min at room temperature. Monolith NT.115 device (Nanotemper) was used with the standard capillaries and 20 % LED and 20 % MST power settings. To compare the results from unlabeled and labeled DtpA, the binding affinity of the tripeptide AFA was measured in both systems. All measurements were performed in duplicates and the results were analyzed with GraphPad Prism 5.0.

2.2.3.5. *In vivo* AK-AMCA uptake and competition assay

The transport activity of DtpA was investigated by the *in vivo* AK-AMCA uptake assay. AK-AMCA stands for β−AlaLys coupled to the fluorescent dye (7-amino-4-methylcoumarin-3-yl) acetic acid by a peptide bond between the side chain amine group of lysine and the carboxyl group of the fluorescent dye (Figure 10).

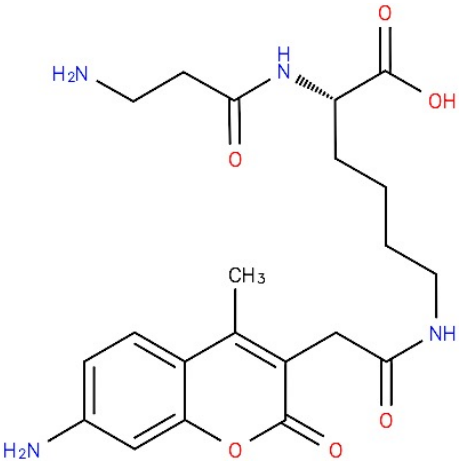


Figure 10: β-Ala-Lys-AMCA chemical structure

DtpA transports AK-AMCA into the cell as well as hPepT1 and hPepT2 and the ligand selectivity of these proteins have been often characterized *in vivo* by the competition assay with this molecule. *E. coli* C41(DE3) cells containing the DtpA plasmid were grown in 5 mL LB medium supplemented with 0.1 mg/ml ampicillin to an OD_{600nm} of 0.7. DtpA expression was induced by the addition of 0.2 mM IPTG, and was incubated for 3 h at 37 °C. The cell culture was then centrifuged for 3 min at 3000 rpm and resuspended to a final OD_{600nm} of 10 in the assay buffer (100 mM HEPES, pH 7.5, 150 mM NaCl, 5 mM glucose). 40 µL of cells, 40 µL of buffer, 10 µL of AK-AMCA (final concentration 50 µM) and 10 µL of the competing ligand (final concentration 5 or 0.5 mM) were mixed in a 96-well plate and incubated for 20 min at 37 °C shaking at 220 rpm. The reaction was stopped by adding 200 µL of ice-cold assay buffer and the cells were washed twice with the assay buffer. Finally, the cells were resuspended in 200 µL assay buffer and the fluorescence was measured with excitation at 350 nm and emission at 450 nm in an M1000 microplate reader (TECAN). All experiments were performed in triplicates. The results were normalized by the fluorescence value of the control, cells overexpressing DtpA incubated with AK-AMCA, and plotted as AK-AMCA uptake percentage. The

analysis was performed in Excel and GraphPad Prism 5.0. The IC50 values were extracted from the log(inhibitor) vs response curves in GraphPad Prism and the K_i value was calculated with the Cheng and Prusoff equation (see below) (Cheng and Prusoff 1973).

$$K_i = \frac{IC50}{1 + \frac{[S]}{K_m}}$$

2.2.3.6. *In vivo* crosslinking of DtpA and N00

The *in vivo* binding of N00 to DtpA was investigated by a crosslinking experiment. The experiment was performed by Kim Bartels from EMBL Hamburg. DtpA was mutated with a UV inducible cross-linker at two positions separately. One position G53 was selected in proximity of the N00 binding site in the periplasmic side of DtpA and the second position F265 was selected on the cytoplasmic side of DtpA further apart from the N00 binding site as a negative control. The crosslinking protocol from Farrell *et al.* was followed (Farrell et al. 2005). The DtpA plasmid carrying the mutation and the amber suppression plasmid pEvol-pBpF, which was kindly provided by Prof. Henning Tidow from University of Hamburg, were co-transformed into *E. coli* C41(DE3) cells. The cells were grown in 100 mL TB buffer until the OD_{600nm} reached 0.7 and the media was supplemented with 1 mM Tyr, 1 mM Trp, 1 mM p-benzoyl-L-phenylalanine (pBpA) and 0.02 % arabinose. After 50 min incubation at 37 °C and shaking at 220 rpm, the protein expression was induced by 0.2 mM IPTG and incubated for further four hours. The cells were harvested by centrifugation at 4000 rpm for 12 min at 4 °C and resuspended in 20 mL lysis buffer (20 mM NaP, pH 7.5, 300 mM NaCl, 5 % glycerol, 15 mM imidazole and 0.5 mM TCEP). Then, the cells were disrupted by sonication two times 1 min, 0.5 s on/off rate and 30 % amplitude. The cells were centrifuged and resuspended again in 20 mL lysis buffer and mixed with 5 µL N00 at 13 mg/ml. Half of the sample was kept in a tube as UV negative control, whereas the other half was poured in a petri dish and shone UV light upon at 365 nm for 60 min. For both fractions, DtpA was purified from the membrane as described in Methods section 2.2.2.1 and western blot analysis was performed as described in Methods section 2.2.3.2, with the exception that a His Probe-HRP antibody was used to detect DtpA.

2.2.4. Structural biology methods

2.2.4.1. Crystallization of DtpA

2.2.4.1.1. Vapor diffusion method

DtpA at 11 mg/ml concentration was mixed with N00 (13 mg/ml) in 1:1.2 molar ratio. The final concentration of DtpA was 8.5 mg/ml in the DtpA-N00 complex. For crystallization of DtpA without nanobodies, the sample at 11 mg/ml concentration was used. For crystallization trials with ligand, the ligand powder was weighed in a fresh tube and DtpA-N00 mix was added on top leading to 30-50 mM ligand concentration. The sample was incubated 1 h on ice and Intelli-plates (Art Robbins Instruments) were filled with the protein sample and reservoir solution in 1:1, 1:2 and 2:1 ratio with a final volume of 300 nL using the liquid dispenser robot mosquito (ttplabtech). The initial crystallization screens are listed in Table 7. The conditions, in which initial crystals grew, were optimized with self-made screens using the scorpion screen builder (Art Robbins Instruments). The crystallization plates were stored in the protein crystallization imager Rock Imager-1000 (Formulatrix) either at 19 or 4 °C and frequent images were taken automatically following the crystallization process. To test the crystals at the beamlines, crystals were fished out of the crystallization drop with a LithoLoop (Molecular Dimensions) of fitting size to the crystal size and stored flash frozen until beam time.

2.2.4.1.2. Lipidic cubic phase method

For LCP crystallization, the protocol from Aherne *et al*. was followed (Aherne, Lyons, and Caffrey 2012). DtpA, at 11 mg/ml, was mixed the lipid 1-(7Z-pentadecenoyl)-rac-glycerol (MAG 7.8) in equal volumes. The lipid was pre-warmed on a thermoblock for 3 min at 45 °C and loaded on a pre-warmed Hamilton gastight syringe for mixing. The DtpA solution was loaded on another syringe, which was then connected to the syringe with the lipid via an LCP syringe coupler (ttplactech). Pushing the content in the syringes from one side to the other 30 to 50 times until the content was transparent and homogenous mixed DtpA and the lipid. The DtpA-lipid mix was moved to one syringe and the other syringe with the coupler was replaced with a mosquito LCP needle (ttplabtech). The syringe was connected to the liquid dispenser robot mosquito (ttplabtech) and 50 nL mesophase and 800 nL precipitant

solution were added on a Laminex UV plastic base with wells of 100 μm depth (Molecular Dimensions). The plate was sealed with a Laminex UV plastic 200-micron film cover (Molecular Dimensions) and stored in the protein crystallization imager Rock Imager-1000 (Formulatrix) at 19 °C. Frequent images were taken automatically following the crystallization process and the plate was also manually inspected under microscope.

2.2.4.2. Data collection and processing

Several hundreds of crystals were screened for DtpA-N00 complex in ligand-free or ligand-bound state throughout this project. Mostly, the beamlines P1318 and P14 operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany) were used but during shutdown times the ID23-1 (Nurizzo et al. 2006) and ID30B (McCarthy et al. 2018) at ESRF (Grenoble, France) were visited or crystals sent to ID30A-1/MASSIF-1 at ESRF (Grenoble, France) for automatic data collection (Bowler et al. 2015). The parameters for data collection were typically 0.1 to 0.2 ° oscillation range, 0.05 to 0.1 s exposure time, 100 % transmission, 1800 to 3600 images covering 180 to 360 ° of the crystal and the detector distance at 2.0 to 2.5 Å resolution in the corners.

The collected diffraction data was indexed and integrated by XDS (Kabsch 2010). Based on the I/σI, $CC_{1/2}$ and completeness values the highest resolution limit was decided. For all reported structures a single crystal was used to determine the structure but scaling three to five datasets from different crystals was tested, which did not improve the statistics or the electron density after molecular replacement and one cycle of refinement. The scaling of the data was performed using XSCALE, followed by the conversion of the file format to mtz format via XDSCONV.

Molecular replacement was performed using the Phaser program from the Phenix suite (Adams et al. 2010; McCoy et al. 2007). Initial molecular replacement tests with at the time available POT structures alone were not successful. Finally, the phases could be estimated by providing two models: DtpD structure (PDB id: 4Q65) modified to a polyalanine chain with loops removed and an unpublished nanobody structure, which I had previously solved in the scope of another project (data not shown). The LLG and TFZ scores were evaluated for choosing a good molecular replacement solution (Oeffner et al. 2013). The refinement was performed in iterative

cycles of the phenix.refine tool in Phenix suite with varying parameters at different stages and manual building of the structure in Coot (Emsley and Cowtan 2004). For the initial refinement strategy, the XYZ coordinates, real space and group B-factors were selected, the final cycles the TLS groups and optimize X-ray/stereochemistry and -/ADP weights were also included. For both diastereomers of valganciclovir and the DDM molecule, pdb and cif files were generated in eLBOW of Phenix suite (Moriarty, Grosse-Kunstleve, and Adams 2009), which were used in the refinement steps. The final models were validated by MolProbity (Chen et al. 2010; Davis et al. 2007) and pdb-redo server (Joosten et al. 2014). The determined structures were provided to the Protein Data Bank (PDB), which will be accessible upon publication:

- DtpA-N00 complex in MES buffer, PDB id: 6GS1
- DtpA-N00 complex in glycine buffer, PDB id: 6GS7
- DtpA-N00 in complex with valganciclovir, PDB id: 6GS4

2.2.4.3. Structural analyses

The structures were inspected and prepared for figures in PyMOL (Delano 2002). The DtpA-N00 and DtpA-valganciclovir interactions were analyzed by LigPlot (Wallace, Laskowski, and Thornton 1995; Laskowski and Swindells 2011) and manual inspection in PyMOL. For the ligand-binding site cavity analysis POVME 2.0 (Durrant, de Oliveira, and McCammon 2011; Durrant et al. 2014) was used. Here, the tested protein structures were first all aligned to DtpA-N00 structure (PDB id: 6GS4) in PyMOL. The position of the Cα atom of the valine residue in valganciclovir was selected due to its central position in the investigated binding pocket. The cavity volume with a radius of 4 and 7 Å was calculated using the software.

3. Results

3.1. Protein expression and purification

3.1.1. Expression and purification of DtpA

For the characterization of DtpA and crystallization experiments, wild type (WT) protein and mutants were expressed and purified as described in the Methods section 2.2.2.1. The results from the DtpA WT purification are shown in Figure 11 and are representative for the purification of the mutants as well. Typically, 20 g cell pellet (wet weight) was harvested from 1 L cell culture of *E. coli* cells and up to 5 mg DtpA was purified. The protein purity was evaluated by SDS-PAGE analysis throughout the purification process. Despite the monodisperse protein profile in the gel filtration chromatogram, the protein was not completely pure; there were additional contaminants of high molecular weight that were not possible to eliminate. Therefore, an optimization of the purification protocol coupled to the nanobody 00 (N00) was developed to improve the overall purity of the sample (see section 3.1.3).

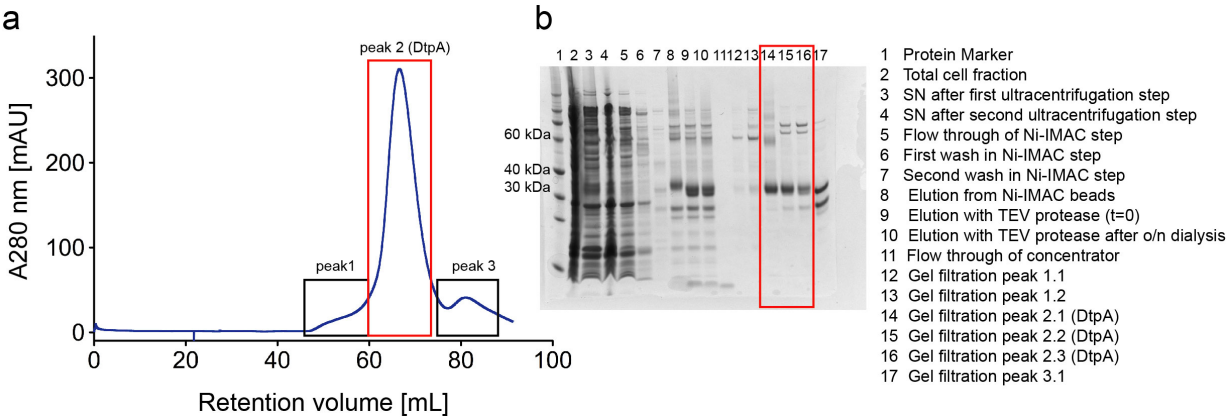


Figure 11: Purification of DtpA
(a) Gel filtration chromatogram. The peak corresponding to DtpA is highlighted with a red box. (b) Coomassie stained gel of all purification steps.

3.1.2. Expression and purification of the nanobodies

Four nanobodies, nanobody 87, -89, -93 and -00, were generated against DtpA as crystallization chaperones. Els Pardon and Jan Steyaert from the Vrije University Brussels performed the nanobody generation with the provided detergent-solubilized DtpA according to the described protocol in Pardon *et al*. (Pardon *et al*. 2014). I have received the plasmids of the nanobodies and expressed and purified the nanobodies as described in Methods section 2.2.2.2. As an example of the purification process, the chromatograms of the nanobody 00 (N00) purification are shown in Figure 12, which are representative for other nanobody purifications as well. The nanobody gene construct contains a C-terminal EPEA-tag (EPEA stands for GluProGluAla), therefore the nanobodies were purified on an affinity column (CaptureSelect™ C-tag affinity matrix) directed against the EPEA-tag.

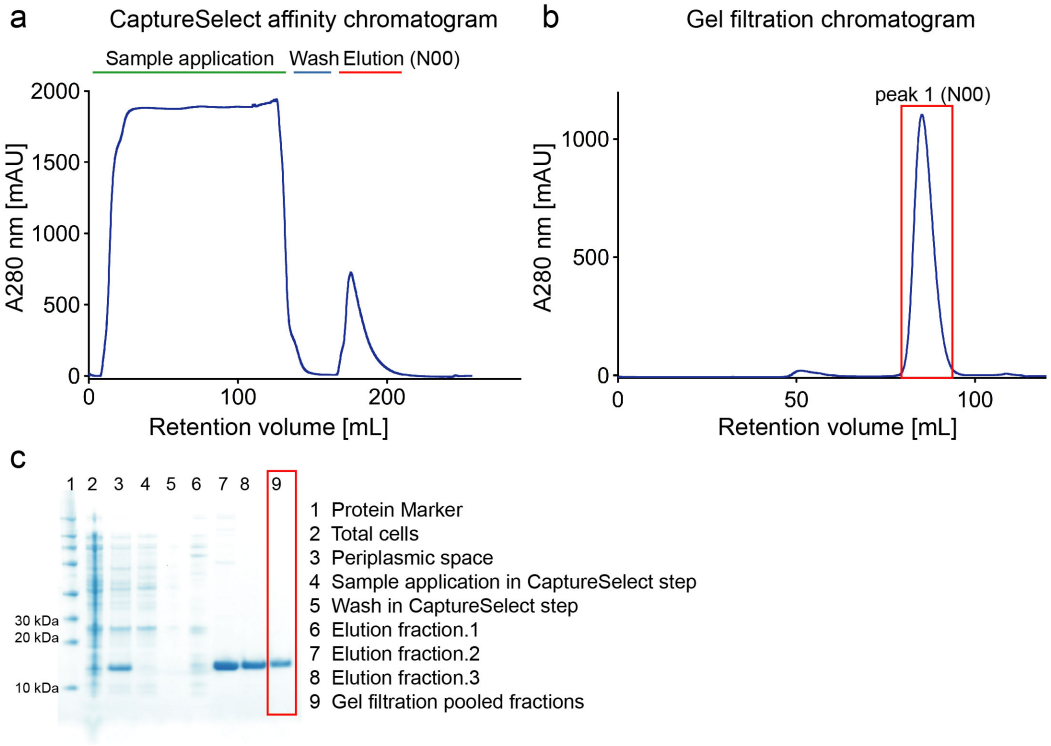


Figure 12: Purification of nanobody 00
(a) CaptureSelect affinity chromatogram. (b) Gel filtration chromatogram with the peak corresponding to N00 marked with a red box. (c) Coomassie stained gel with all purification steps.

The purity of the sample was high. The protein yields were ca. 15 mg from 1 L cell culture starting material. The purified nanobody sample was mixed in 1:1.2 molar

ratio with DtpA for functional characterization and crystallization experiments based on the observation of 1:1 binding mode and to have a slight excess of nanobody to DtpA.

3.1.3. DtpA purification coupled to N00

Combining the purification of DtpA to the purification of N00 a new purification protocol was developed. All steps were followed individually for DtpA and N00 as described until DtpA was applied to the second ultracentrifugation step and N00 was loaded on the CaptureSelect affinity beads (Figure 13). Instead of applying DtpA to Ni-IMAC beads, it was applied to the CaptureSelect affinity beads, which were pre-loaded with N00 and washed with the gel filtration buffer of DtpA containing 0.03 % DDM. After elution of the complex, the TEV cleavage and gel filtration steps were performed as described in Methods section 2.2.2.3.

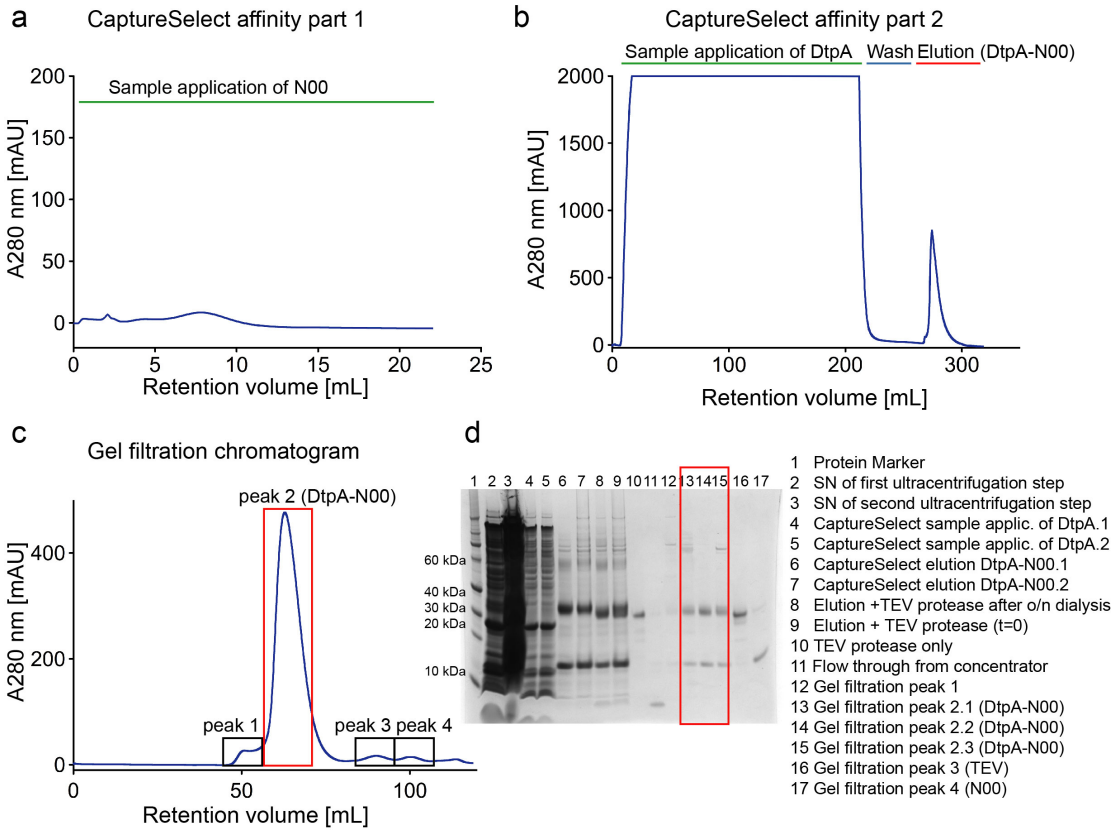


Figure 13: Purification of DtpA coupled to N00
(a) N00 is loaded on the CaptureSelect affinity beads. (b) DtpA is loaded on CaptureSelect beads after the second ultracentrifugation step and eluted together with

N00. (c) Gel filtration chromatogram of DtpA-N00 complex. The peak corresponding to the complex is marked with a red box. (d) Coomassie stained gel from all purification steps.

A highly pure sample of the DtpA-N00 complex was achieved but the yield was with 3 mg from 1 L cell culture starting material lower than the original DtpA purification protocol (5 mg from 1 L cell culture). It is possible that the impurities in the original purification method led to an overestimation of the protein concentrations. Nevertheless, the crystallization attempts were more successful with the original DtpA purification protocol and mixing DtpA and N00 just before setting up crystallization plates; therefore, the DtpA purification coupled to N00 was not optimized further.

3.2. Complex formation of DtpA and nanobodies

3.2.1. DtpA and nanobody interaction monitored by analytical gel filtration assay

The complex formation of DtpA and nanobodies was analyzed by analytical gel filtration chromatography (Methods section 2.2.3.1). Individual protein runs were compared to the potential complexes formed between DtpA and the nanobodies. Furthermore, combinations of two nanobodies were tested with DtpA to investigate a possible cooperative binding of multiple nanobodies. Additionally, a nanobody generated against another membrane protein (N21) was used as a negative control. The evaluation of complex formation chromatograms can be summarized in terms of three major observations: (i) the DtpA peak is shifted to a shorter retention time due to the slightly larger size of the complex relative to DtpA alone, (ii) the absorption signal of the DtpA peak is increased due to the contribution of the nanobody and (iii) the absorption signal of the nanobody peak is strongly reduced, because the nanobody no longer elutes at that retention time but in complex with DtpA. All nanobodies against DtpA form a complex with DtpA as evident on the chromatograms but no complex formation is observed for with N21 (Figure 14). There was no binding observed for any combination of two nanobodies (data shown only for the combination of N00 and N87). Such a binding would suggest that two nanobodies

stabilize the same conformation but bind on distinct sites to DtpA. The absence of the binding could be interpreted in two ways. Either the nanobodies bind to different conformations of DtpA or they bind to the same site on DtpA.

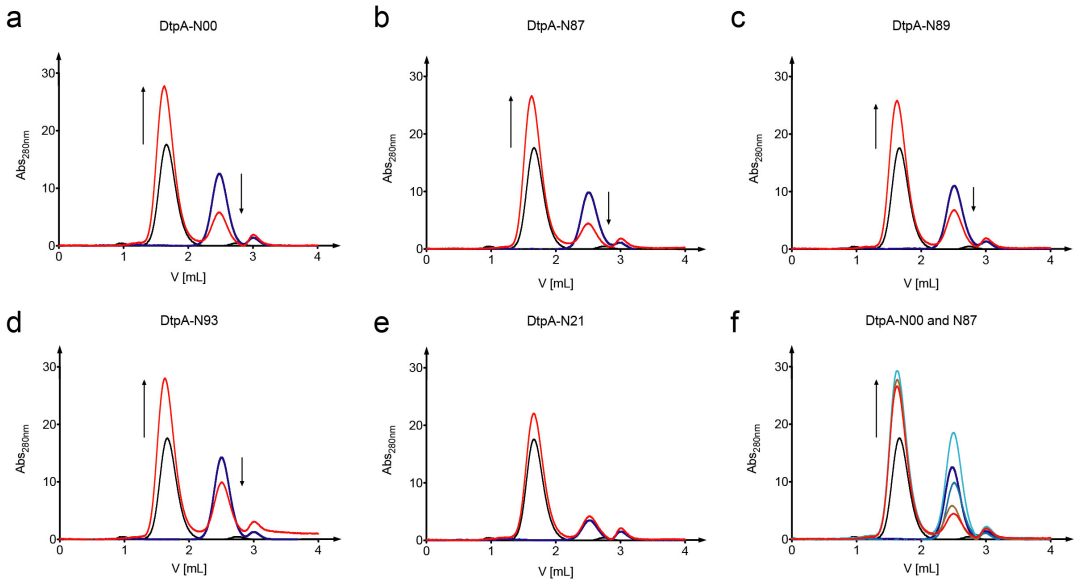


Figure 14: Analytical gel filtration profiles of DtpA and nanobodies
(a-d) The chromatograms of DtpA and the four nanobodies, which were generated against DtpA are plotted with DtpA alone in black, the nanobody alone in blue and the complex in red. Arrows highlight the change in absorbance signal between the DtpA/nanobody alone and the complex. (e) A negative control was included in the experiment using N21, which is a nanobody generated against another membrane protein. (f) An exemplary curve for the testing of two nanobodies (N00 and N87) incubated with DtpA at the same time (in cyan). The curves of DtpA-N00 (red), DtpA-N87 (brown), N00 alone (dark blue), N87 alone (blue) are included for comparison.

3.2.2. Thermal stability of DtpA with nanobodies

The complex formation of DtpA and nanobodies were investigated in addition to the analytical gel filtration assay with the nano differential scanning fluorimetry (nanoDSF) method (Methods section 2.2.3.3). Here, the intrinsic fluorescence of the protein from mostly tryptophan residues are tracked over a temperature gradient from 20 to 85 °C. A value for the thermal stability (also called the melting temperature, T_m) of the protein is obtained from the transition point in the fluorescence ratio 350 nm/330 nm plotted against the temperature.

The thermal stability of DtpA and nanobodies were measured individually and in combination. All nanobodies generated against DtpA led to a significant thermal stabilization. The melting temperature of DtpA was increased by 10 to 18 °C, which was interpreted as complex formation (Figure 15). The negative control, N21, does not thermally stabilize DtpA, which is interpreted as no interaction. Furthermore, higher molar ratios of DtpA to nanobodies, such as 1:2 and 1:3, were tested but these did not show any further stabilization effect (data not shown). Therefore, we expect a 1:1 binding mode.

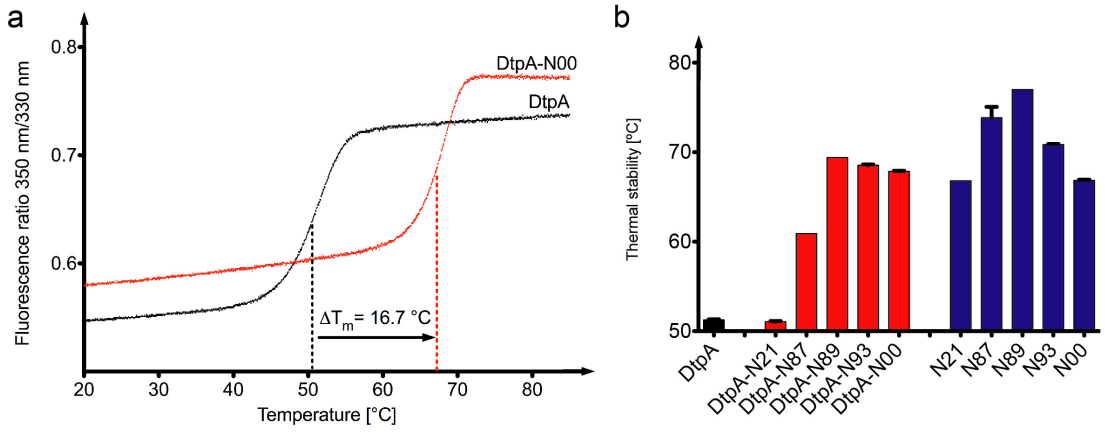


Figure 15: Thermal stabilization effect of nanobodies on DtpA
(a) The fluorescence ratio between 350 nm and 330 nm is plotted for DtpA alone (black) and DtpA-N00 complex (red) across the temperature gradient. (b) Thermal stability of DtpA (black), DtpA-nanobody complexes (red) and nanobody on its own (blue). N21 is a negative control. The error bars indicate the standard deviation calculated from two experiments.

3.2.3. Evaluation of conformational or linear binders

One of the challenges while generating nanobodies against membrane proteins is that often generated nanobodies would recognize linear epitopes, namely parts of the unfolded protein. Most likely because the membrane protein would start to unfold after the immunization of the lama and the selection process for the nanobodies would be then based on the partially unfolded protein. For using nanobodies as crystallization chaperones, it is crucial to have conformational binders, namely nanobodies binding and stabilizing the folded protein. The binding mode of the nanobodies was tested on a western blot analysis as described in the Method section

2.2.3.2. Proteins (DtpA, DtpC which was used as a negative control, and N00, positive control) were run on gel electrophoresis under denaturing conditions (SDS PAGE) and were then transferred to a PVC membrane and incubated with one of the four nanobodies (N87, -89, 93 and 00) for 1 h. After wash steps, the membrane was incubated with a primary antibody against the EPEA-tag of the nanobodies coupled to biotin and a secondary antibody, streptavidin coupled to horseradish peroxidase. On the developed membrane, only a band corresponding to the nanobody could be observed in the positive control lane loaded with N00 (Figure 16). The absence of bands in other lanes is interpreted as all nanobodies being conformational binders and they do not bind to DtpA (or DtpC) in the SDS treated/denatured form. This finding is also supported by the analytical gel filtration and thermal stability experiments, as they are also based on a conformational binding of the nanobody to DtpA.

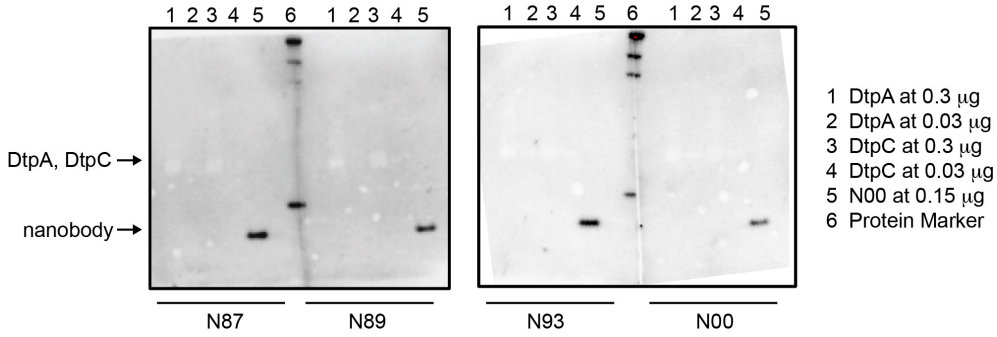


Figure 16: Evaluation of the conformational or linear binding mode
DtpA, DtpC (as negative control) and N00 (as positive control) were loaded on a gel under denaturing conditions and transferred to a PVC membrane. The membrane was incubated with one of the four nanobodies (N87, N89, N93 or N00) and antibodies against nanobodies were used to detect, if the nanobodies could bind to denatured DtpA. No bands were observed in the DtpA lanes but only in the positive control lane loaded with a nanobody (N00).

3.2.4. *In vivo* AK-AMCA uptake in cells co-expressing DtpA and N00

With the previous experiments, the DtpA-nanobody interaction was shown with the purified and detergent-solubilized DtpA, which was necessary for crystallization. The two methods presented here and in the next section were developed to investigate the binding of N00 to DtpA, while DtpA was still in the cell membrane. Thereby, we can

evaluate if the determined DtpA-N00 structure (Results section 3.5) represents a biologically relevant state.

The AK-AMCA uptake assay (Methods section 2.2.3.5) was used for the *in vivo* characterization of DtpA and N00 binding. Cells overexpressing DtpA and co-expressing DtpA and N00 were tested in parallel and the cells co-expressing DtpA and N00 had only 25 % of AK-AMCA uptake relative to the cells overexpressing DtpA only (Figure 17). We concluded from this experiment that N00 binds to DtpA in its biological environment, the native membrane.

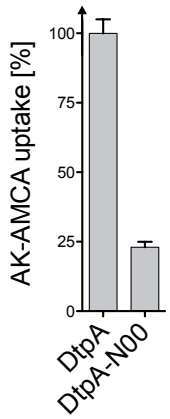


Figure 17: AK-AMCA uptake of DtpA and DtpA-N00 co-expressing cells
The AK-AMCA uptake assay was performed with cells overexpressing DtpA and DtpA co-expressed with N00 and the results were normalized by the uptake of cells overexpressing DtpA only. The error bars indicate the standard deviation calculated from three experiments.

3.2.5. *In vivo* crosslinking of DtpA and N00

Second approach to characterize N00 binding to DtpA *in vivo* was based on crosslinking, which was performed by Kim Bartels (EMBL Hamburg) in collaboration as described in Methods section 2.2.3.6. The UV inducible cross-linker was introduced in two positions within DtpA as an unnatural amino acid. One site was selected in direct proximity to the N00 binding site (G53) and one site in a distant position on the cytoplasmic side of the protein (F265) (Figure 18). DtpA containing the unnatural amino acid was overexpressed in *E. coli* and incubated with N00 after cell lysis with and without UV treatment in parallel. By western blot, we could see a band corresponding to the crosslinked DtpA-N00 complex, which was visible only

with the G53 mutant after UV treatment and not in the control lanes (Figure 18). With this experiment, we could confirm that N00 can bind DtpA, while it is still in the cell membrane.

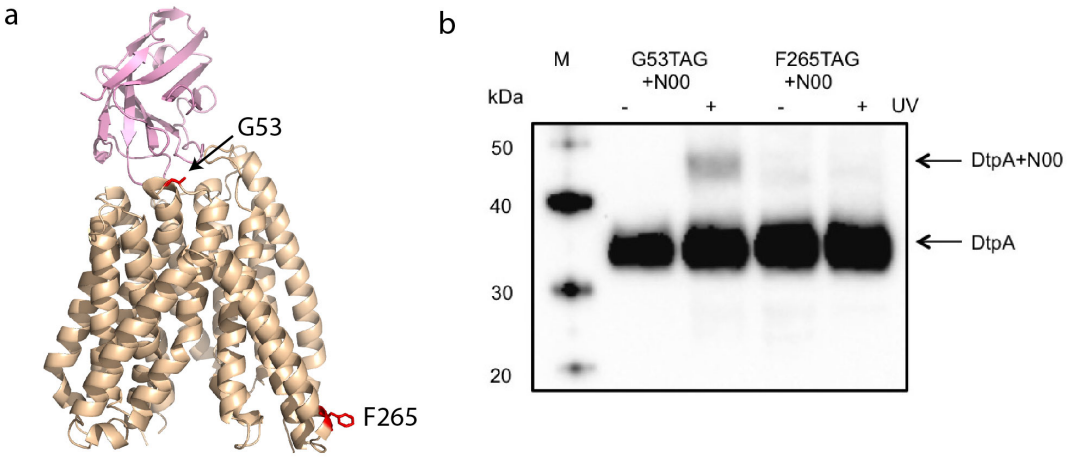


Figure 18: *In vivo* crosslinking of DtpA and N00
(a) The overall structure of DtpA (wheat) and N00 (pink) with the mutation sites, G53 and F265, highlighted in red. (b) *E. coli* cells overexpressing DtpA mutants with the unnatural amino acid were incubated with N00 after cell lysis and treated with UV light. On a western blot analysis, the crosslinking results were analyzed by the occurrence of a band corresponding to the DtpA-N00 complex.

3.3. Protein characterization

3.3.1. Thermal stability of DtpA at different conditions

As a preliminary step towards functional as well as crystallization experiments, the effect of pH, buffer, and salt on DtpA was investigated with the nanoDSF method (Methods section 2.2.3.3) using the commercial RUBIC screen (Methods section 2.2.3.3.1). The thermal stability of DtpA ranged from 44 to 53 °C in the tested conditions (Figure 19).

a

DtpA

	1	2	3	4	5	6	7	8	9	10	11	12
A	51.1°C	48.8°C	51.5°C	52.3°C	53.3°C	51.6°C	51.6°C	52.1°C	52.1°C	49.3°C	50.1°C	52.2°C
B	52.6°C	51.4°C	51.8°C	48.8°C	50.0°C	49.6°C	50.9°C	49.9°C	50.3°C	48.2°C	48.7°C	48.6°C
C	51.2°C	46.8°C	51.7°C	52.2°C	52.1°C	51.7°C	52.0°C	50.2°C	51.1°C	49.2°C	50.1°C	52.4°C
D	50.7°C	49.4°C	51.3°C	48.6°C	47.0°C	47.7°C	49.6°C	48.0°C	47.9°C	46.5°C	49.3°C	46.1°C
E	49.8°C	51.9°C	52.3°C	51.8°C	51.1°C	50.0°C	48.7°C	47.8°C	47.6°C	47.1°C	46.6°C	44.2°C
F	50.0°C	49.4°C	48.9°C	48.6°C	49.0°C	48.5°C	49.1°C	49.9°C	50.8°C	51.0°C	50.8°C	51.4°C
G	48.7°C	48.2°C	47.6°C	47.4°C	49.9°C	47.7°C	49.9°C	49.0°C	48.6°C	48.2°C	48.3°C	49.1°C
H	52.9°C	52.2°C	52.3°C	51.7°C	48.1°C	51.5°C	50.1°C	48.7°C	48.2°C	47.1°C	46.2°C	44.2°C

b

	1	2	3	4	5	6	7	8	9	10	11	12
A	water	0.1 M Citrate pH 4	0.1 M Sodium acetate pH 4.5	0.1 M Citrate pH 5	0.1 M MES pH 6	0.1 M Potassium phosphate pH 6	0.1 M Citrate pH 6	0.1 M Bis-Tris pH 6.5	0.1 M MES pH 6.5	0.1 M Sodium phosphate pH 7	0.1 M Potassium phosphate pH 7	0.1 M HEPES pH 7
B	0.1 M MOPS pH 7	0.1 M Ammonium acetate pH 7.3	0.1 M Tris-HCl pH 7.5	0.1 M Sodium phosphate pH 7.5	0.1 M Imidazole pH 7.5	0.1 M HEPES pH 8	0.1 M Tris-HCl pH 8	0.1 M Tricine pH 8	0.1 M BICINE pH 8	0.1 M BICINE pH 8.5	0.1 M Tris-HCl pH 8.5	0.1 M CHES pH 9
C	0.2 M NaCl	0.2 M NaCl 0.1 M Citrate pH 4	0.2 M NaCl 0.1 M Sodium acetate pH 4.5	0.2 M NaCl 0.1 M Citrate pH 5	0.2 M NaCl 0.1 M MES pH 6	0.2 M NaCl 0.1 M Potassium phosphate pH 6	0.2 M NaCl 0.1 M Citrate pH 6	0.2 M NaCl 0.1 M Bis-Tris pH 6.5	0.2 M NaCl 0.1 M MES pH 6.5	0.2 M NaCl 0.1 M Sodium phosphate pH 7	0.2 M NaCl 0.1 M Potassium phosphate pH 7	0.2 M NaCl 0.1 M CHES pH 7
D	0.2 M NaCl 0.1 M MOPS pH 7	0.2 M NaCl 0.1 M Ammonium acetate pH 7.3	0.2 M NaCl 0.1 M Tris-HCl pH 7.5	0.2 M NaCl 0.1 M Sodium phosphate pH 7.5	0.2 M NaCl 0.1 M Imidazole pH 8	0.2 M NaCl 0.1 M HEPES pH 8	0.2 M NaCl 0.1 M Tris-HCl pH 8	0.2 M NaCl 0.1 M Tricine pH 8	0.2 M NaCl 0.1 M BICINE pH 8	0.2 M NaCl 0.1 M BICINE pH 8.5	0.2 M NaCl 0.1 M Tris-HCl pH 8.5	0.2 M NaCl 0.1 M CHES pH 9
E	0.1 M SPG pH 4	0.1 M SPG pH 4.5	0.1 M SPG pH 5	0.1 M SPG pH 5.5	0.1 M SPG pH 6	0.1 M SPG pH 6.5	0.1 M SPG pH 7	0.1 M SPG pH 7.5	0.1 M SPG pH 8	0.1 M SPG pH 8.5	0.1 M SPG pH 9	0.1 M SPG pH 10
F	0.02 M HEPES pH 7.5	0.05 M HEPES pH 7.5	0.12 M HEPES pH 7.5	0.25 M HEPES pH 7.5	0.02 M Sodium phosphate pH 7.5	0.05 M Sodium phosphate pH 7.5	0.12 M Sodium phosphate pH 7.5	0.25 M Sodium phosphate pH 7.5	0.02 M Tris-HCl pH 8	0.05 M Tris-HCl pH 8	0.12 M Tris-HCl pH 8	0.25 M Tris-HCl pH 8
G	0.05 M NaCl 0.05 M HEPES pH 7.5	0.12 M NaCl 0.05 M HEPES pH 7.5	0.25 M NaCl 0.05 M HEPES pH 7.5	0.5 M NaCl 0.05 M HEPES pH 7.5	0.75 M NaCl 0.05 M HEPES pH 7.5	1 M NaCl 0.05 M HEPES pH 7.5	0.05 M NaCl 0.05 M Tris-HCl pH 8	0.12 M NaCl 0.05 M Tris-HCl pH 8	0.25 M NaCl 0.05 M Tris-HCl pH 8	0.5 M NaCl 0.05 M Tris-HCl pH 8	0.75 M NaCl 0.05 M Tris-HCl pH 8	1 M NaCl 0.05 M Tris-HCl pH 8
H	0.05 M MES/Bis-Tris pH 6	0.05 M MES/Imidazole pH 6.5	0.05 M Bis Tris/PIPES pH 6.5	0.05 M MOPS/Bis-Tris propane pH 7	0.05 M Phosphate/Citrate pH 7.5	0.05 M MOPS/Sodium HEPES pH 7.5	0.05 M BICINE/Tris pH 8.5	0.1 M NaCl 0.05 M Imidazole pH 7.5	0.1 M NaCl 0.12 M Imidazole pH 7.5	0.1 M NaCl 0.25 M Imidazole pH 7.5	0.1 M NaCl 0.35 M Imidazole pH 7.5	0.1 M NaCl 0.5 M Imidazole pH 7.5

Figure 19: Heatmap of RUBIC screen results for DtpA and the conditions
(a) The results from the RUBIC screen for DtpA are presented as a heat map with the lowest thermal stability in °C colored in red and the highest in green. (b) Conditions used in the RUBIC screen. The RUBIC screen consists of buffer and pH screens without NaCl (row A and B, in orange) and with 250 mM NaCl (row C and D, in yellow), followed by a pH range test using the SPG buffer (row E, in lilac) and selected buffers at different concentrations and constant pH (row F, in brown tones). Then, the effect of salt concentration from 50 mM to 1 M is tested for two buffers at constant concentration and pH (row G, in grey tones). Buffer combinations and effect of imidazole concentration is also included (row H, in blue tones).

pH had the strongest effect on DtpA stability, with the highest stability at pH 4.5 to 6.0 and destabilization at higher pH values (Figure 20). Different buffers in this pH range were tested in presence and absence of 0.2 M NaCl. In general, salts are often

included in buffers because of their stabilizing effect on the proteins. Nevertheless, in case of DtpA 0.2 M NaCl had a slightly destabilizing effect in some of the pH and buffer conditions. Sodium phosphate (NaP) buffer at pH 7.5 was used in initial purification experiments. With this screen, it was evident that NaP is a destabilizing buffer. More optimal buffers are TRIS and HEPES (Figure 20); therefore, the purification was continued with HEPES buffer at pH 7.5.

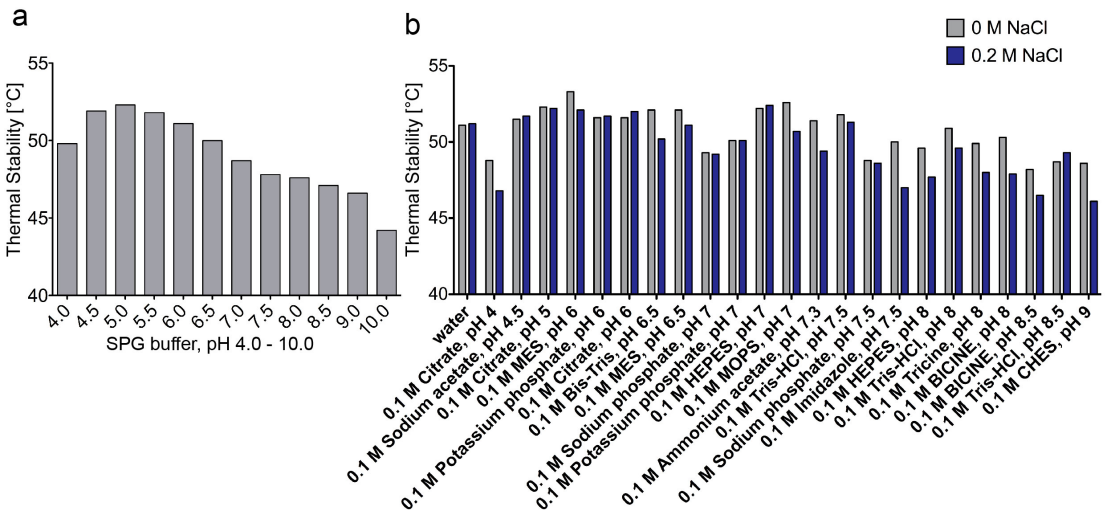


Figure 20: Thermal stability of DtpA in different conditions
(a) The effect of pH from 4.0 to 10.0 was tested as well as (b) a set of different buffers spanning the same pH range in presence and absence of salt (0.2 M NaCl).

3.3.2. Thermal stability of DtpA-N00 complex at different conditions

The thermal stability of DtpA-N00 complex was measured as described for DtpA above. The DtpA-N00 complex is overall more stable than DtpA as mentioned in section 3.2.2 with the melting temperatures ranging from 48 to 67 °C in the RUBIC screen (Figure 21). The thermal stability in position H5 was not possible to determine due to a technical problem, most likely due to air bubbles in the capillary.

DtpA-N00 complex

	1	2	3	4	5	6	7	8	9	10	11	12
A	66.4°C	50.5°C	52.2°C	58.9°C	61.8°C	64.7°C	62.6°C	64.4°C	64.8°C	66.1°C	65.9°C	65.9°C
B	66.0°C	65.7°C	65.1°C	65.5°C	65.0°C	66.2°C	65.8°C	66.9°C	67.2°C	66.9°C	65.4°C	64.6°C
C	57.3°C	48.3°C	56.0°C	55.6°C	59.1°C	62.1°C	62.1°C	62.3°C	61.9°C	63.8°C	62.4°C	62.8°C
D	62.2°C	64.9°C	62.8°C	64.3°C	63.1°C	64.0°C	63.0°C	63.6°C	64.4°C	63.7°C	62.6°C	61.9°C
E	51.6°C	57.5°C	62.1°C	64.9°C	65.6°C	66.2°C	66.3°C	66.1°C	65.9°C	65.3°C	64.8°C	64.2°C
F	67.4°C	67.2°C	66.3°C	65.4°C	66.7°C	66.3°C	64.9°C	64.7°C	66.5°C	66.4°C	65.7°C	64.4°C
G	66.5°C	65.8°C	64.8°C	63.4°C	64.8°C	62.8°C	66.0°C	65.2°C	64.6°C	62.9°C	62.9°C	62.3°C
H	62.3°C	64.3°C	64.0°C	65.8°C	n.d.	66.4°C	66.8°C	64.3°C	63.6°C	61.7°C	59.8°C	57.1°C

Figure 21: Heatmap of RUBIC screen results for DtpA-N00 complex
The results from the RUBIC screen for DtpA-N00 complex are presented as a heat map with the lowest thermal stability in °C colored in red and the highest in green. n.d. stands for not determined due to an error during the measurements.

Interestingly, the DtpA-N00 complex has a different pH dependency than DtpA. The DtpA-N00 complex is stable from pH 5.5 to 10.0 and is only destabilized at lower pH values (Figure 22). This is on the contrary to DtpA, which is stable at lower pH and was destabilized at higher pH values (Figure 20). The effect of salt is more pronounced in the DtpA-N00 complex, possibly due to the destabilizing effect of salt on the complex. Almost for all tested buffer and pH combinations the presence of 0.2 M NaCl lowers the thermal stability by a few °C (Figure 22). The choice of buffer does not make a difference but the change in pH is affecting the thermal stability of the complex. Overall, DtpA-N00 is very stable over a wide range of pH values, which facilitates the crystallization of the complex.

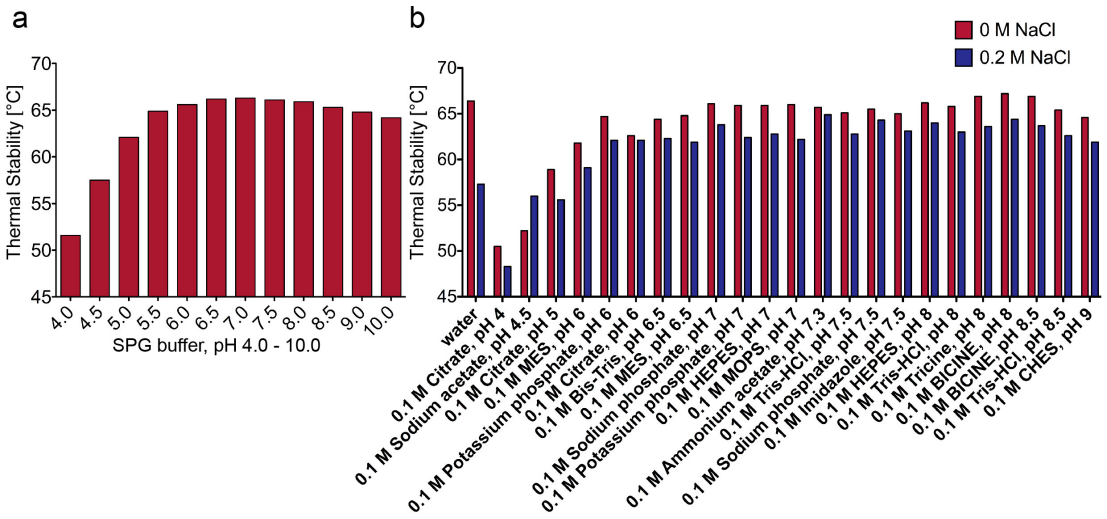


Figure 22: Thermal stability of DtpA-N00 complex in different conditions
(a) The effect of pH from 4.0 to 10.0 was tested as well as (b) a set of different buffers spanning the same pH range in presence and absence of salt (0.2 M NaCl).

3.4. Ligand selectivity

3.4.1. Thermal stability analysis with the ligand library

The ligand selectivity of DtpA was so far tested by *in vivo* uptake and competition assays with a limited number of di-/tripeptides and drugs (Weitz *et al*. 2007; Prabhala *et al*. 2017). In our group, we have established a fast method to screen the binding of ligands *in vitro* consuming only small amounts of purified protein relative to the conventional protein-ligand binding methods, such as the isothermal titration calorimeter. In this assay using the nanoDSF technique, a possible interaction of DtpA with the added ligand can be observed as an increase of the thermal stability.

The ligand library containing eighteen dipeptides, ten tripeptides and seven drugs was tested for DtpA. The selection of the ligands was based on the criteria that they were water soluble and commercially available. The following observations were made from the results shown in Figure 23: (i) All tripeptides, except GlyGlyHis, more than half of the dipeptides and two drugs, valacyclovir and valganciclovir, from the library stabilize DtpA. (ii) The hydrophobic and aromatic residues are more stabilizing than charged residues in the ligand. (iii) Ligands with glycine residues have little to no stabilization effect. A box plot representation was prepared by Jan Strauss from EMBL Hamburg with the thermal stability results, which were subtracted from control (water) and summarized for tripeptides and dipeptides (figure 9b). The median for tripeptides is 7.1 °C and for dipeptides 2.4 °C ($p = 0.035$) concluding that tripeptides have a significantly higher stabilization effect relative to dipeptides on DtpA. The ligands with the highest stabilization effect were further investigated with concentration dependent thermal stability and binding affinity measurements as described in the following sections.

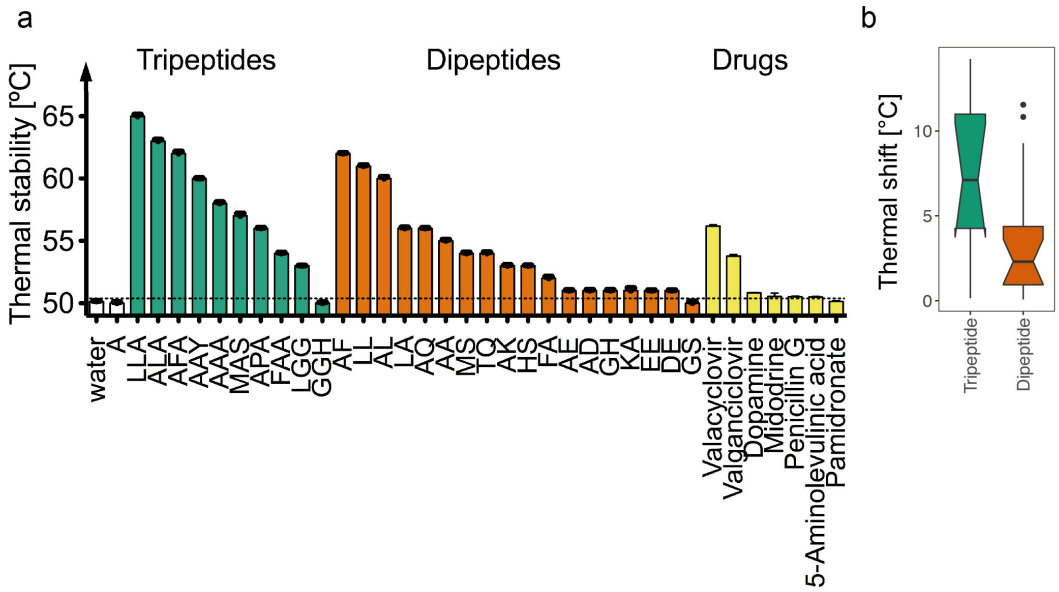


Figure 23: Thermal stability of DtpA with the ligand library
(a) The potential binding of the ligands was tested based on the shift of thermal stability. The ligand concentration was 5 mM. The error bars indicate the standard deviation calculated from three experiments. (b) Box plot analysis of the same results grouped for tripeptides and dipeptides. The 95 % quartile is drawn in green for tripeptides and orange for dipeptides with the lines above and underneath representing the upper and lower extremes. The outliers are shown as dots.

3.4.2. Concentration dependent thermal stability measurements

Following the thermal stability assay with the ligand library, the tripeptide, LeuLeuAla and the dipeptide, AlaLeu, was selected from the ligands with the highest stabilization effect and their stabilization was validated by a concentration dependent assay using the nanoDSF method (Methods section 2.2.3.3). In each case, a concentration dependent stabilization effect was observed (Figure 24). This effect was stronger for DtpA than for DtpA-N00 complex, which has a much higher thermal stability, as expected. It is indeed impressive that DtpA-N00 complex reaches a thermal stability of 74 °C in the presence of 10 mM LeuLeuAla and 70 °C already at 0.3 mM LeuLeuAla. We concluded that not only DtpA but also DtpA-N00 complex binds to LeuLeuAla and AlaLeu.

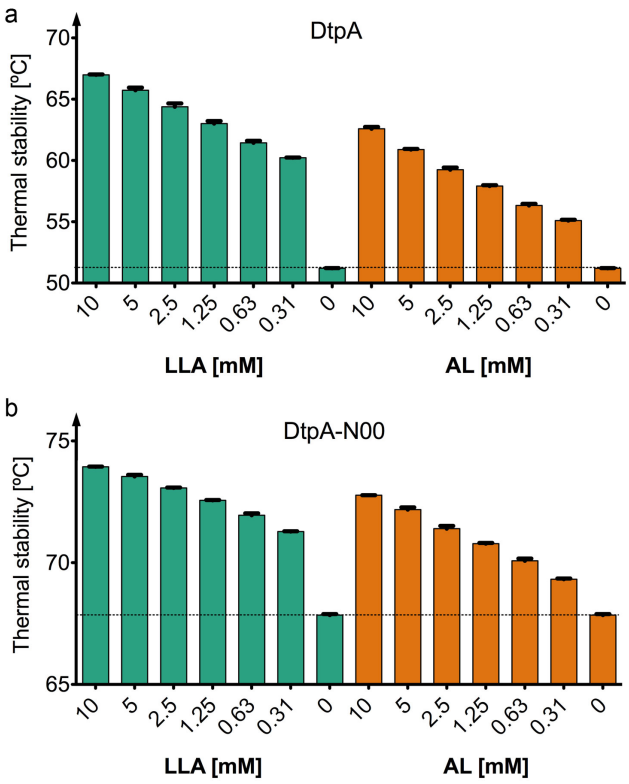


Figure 24: Concentration dependent thermal stabilization effect
The thermal stabilization effect was validated for a tripeptide, LeuLeuAla in green, and a dipeptide, AlaLeu in orange with DtpA and DtpA-N00 complex. The error bars indicate the standard deviation calculated from two experiments.

3.4.3. Acyclovir derivates

Human PepT1 and PepT2 can transport the prodrugs valganciclovir and valaciclovir but not the original drugs ganciclovir or acyclovir (Guo *et al*. 1999; Sugawara *et al*. 2000). After observing the thermal stabilization effect of valganciclovir and valacyclovir on DtpA, we have tested the thermal stabilization effect of ganciclovir, acyclovir and penciclovir, which is a similar drug to the other two. All three drugs are poorly water-soluble and therefore were solubilized in 100% DMSO. As a control, DtpA was mixed with 100% DMSO in the same ratio as with the ligands. The final concentration of DMSO was 5 % in all samples. The addition of the drugs did not change the thermal stability (Figure 25) in contrast to the stabilizing effect of valganciclovir and valacyclovir (Figure 23). Therefore, we conclude that there is no interaction of DtpA with ganciclovir, acyclovir or penciclovir as suggested for hPepT1 as well. Noticeably, the thermal stability of DtpA is slightly lower in presence of 5 % DMSO.

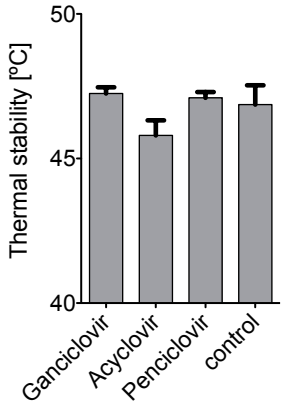


Figure 25: Thermal stability of DtpA in presence of acyclovir derivates
The thermal stabilization effect of the acyclovir, ganciclovir and penciclovir were tested. All samples, including control, are in the presence of 5 % DMSO. The error bars indicate the standard deviation calculated from three experiments.

3.4.4. Binding affinity measurements of peptide and drugs with DtpA and the DtpA-N00 complex

After testing the ligand library based on the change of thermal stability (Figure 23), the ligands with the strongest effect were selected for further investigation. The binding affinity of the selected ligands to DtpA or DtpA-N00 complex was measured with microscale thermophoresis (MST) as described in Methods section 2.2.3.4. The label-free setup of MST is based on the detection of intrinsic fluorescence of proteins, mainly from tryptophan residues. The movement of the protein is followed within a small temperature gradient induced by an infrared laser, and the movement is dependent on the size, charge and hydration shell of the molecule, which changes upon ligand binding. The normalized fluorescence at a selected time point is plotted against the ligand concentration resulting in a ligand-binding isotherm and from the transition point a value equivalent to the dissociation constant (K_d) can be extracted. The advantage of the technique is that it requires only small volumes and a very low concentration of the protein (in nM range), also there is no need for labeling but the limiting factor is possible fluorescence signal from the ligands as observed in the case of valganciclovir and valacyclovir. Therefore, the measurements with the two drugs were performed with the Monolith NT.115 device after labeling the C-terminal His-tag of DtpA with the Monolith NT His-Tag labeling kit RED-tris-NTA (Methods

section 2.2.3.4.2). Here the fluorescence signal from the RED dye was followed as opposed to the intrinsic tryptophan signal. No fluorescence was detected from valganciclovir or valacyclovir at the measured wavelength. The results for all tested ligands are shown in Table 15 and Figure 26.

Table 15: Binding affinity results for di-/tripeptides and the two drugs

Ligands	DtpA K_d [μM]	DtpA-N00 K_d [μM]
valganciclovir	76 ± 16*	113 ± 29*
valacyclovir	60 ± 13*	67 ± 15*
LLA	58 ± 10	59 ± 10
AL	491 ± 112	104 ± 21
ALA	185 ± 52	112 ± 27
AFA	113 ± 24, 33 ± 3*	71 ± 13
AF	106 ± 22	47 ± 13
LL	99 ± 27	45 ± 9
LA	564 ± 250	140 ± 39

*These measurements were performed with labeling the His-tag of DtpA, whereas the others in a label free approach.

The highest affinity was observed for the tripeptide LeuLeuAla, which is also the most thermal stabilizing ligand. Overall, tripeptides had higher binding affinities relative to dipeptides, which is in good agreement with the results from thermal stability experiment. To compare the label-free and labeled method, the tripeptide AlaPheAla was measured using both techniques and indeed there was a 3-fold difference with the labeled method showing a higher binding affinity. Therefore, one needs to be careful while comparing the binding affinity of the drugs with the di-/tripeptides. Overall, they are all in the low to medium μM range. Few ligands, which did not show any thermal stabilization effect, were also measured with this technique but no binding affinity could be determined for any of these ligands (data not shown). Following the observation from concentration dependent thermal stabilization effect of the ligands on DtpA and DtpA-N00 complex, the binding affinity of the same set of ligands was measured with DtpA-N00 complex. In most cases, the binding affinity was similar (2-fold change or less). Nevertheless, in two cases, with AlaLeu and LeuAla, there was a 5-fold difference with the DtpA-N00 complex having a higher binding affinity relative to DtpA. As a control, the binding affinity of N00 alone with

the ligands LeuLeuAla and AlaLeu were measured and no binding could be detected (Figure 26).

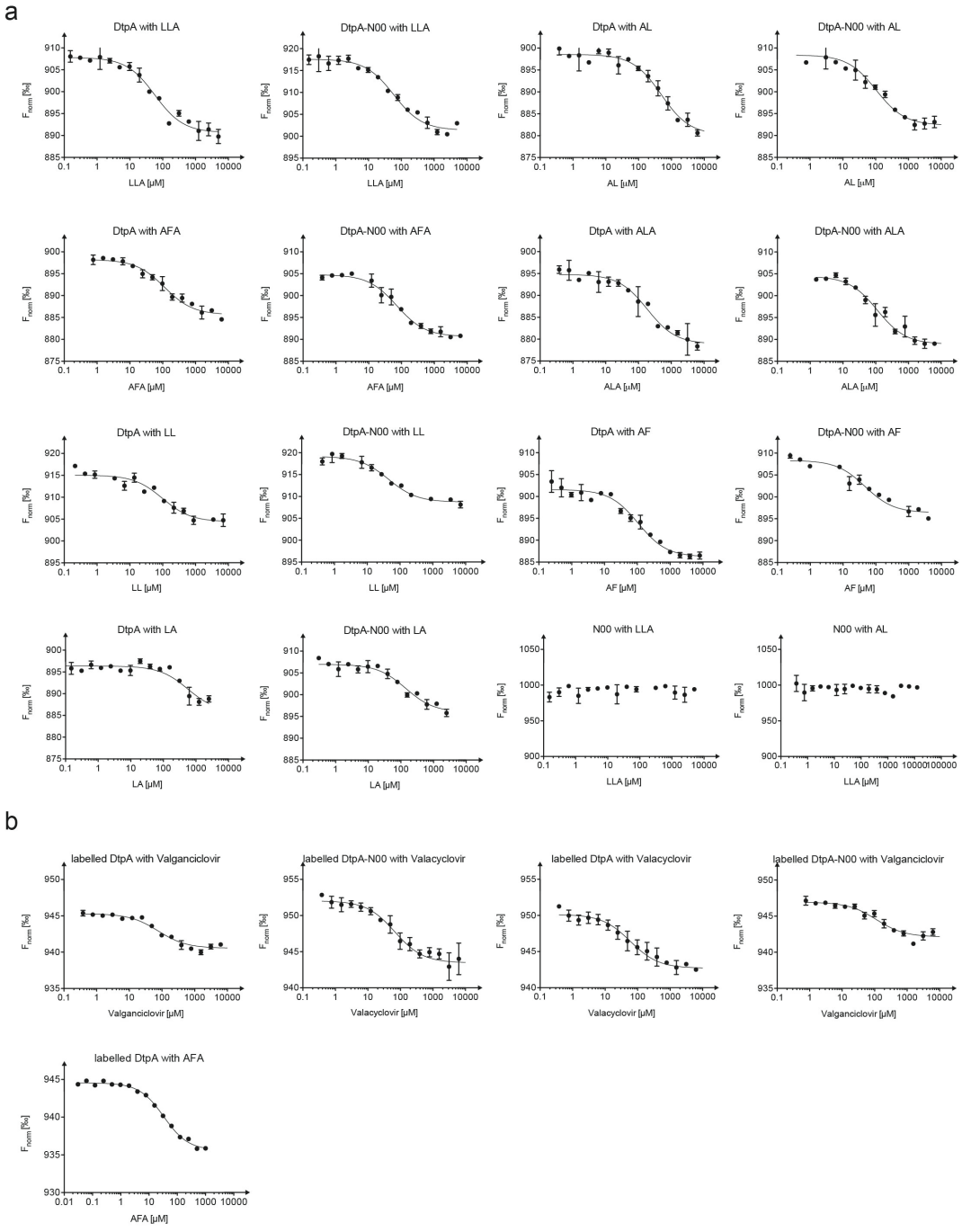


Figure 26: Binding affinity curves for di-/tripeptides and the two drugs
(a) Binding affinity of DtpA and DtpA-N00 complex to the ligands LeuLeuAla, AlaLeu, AlaPheAla, AlaLeuAla, LeuLeu AlaPhe and LeuAla was measured by the labelfree method. N00 was measured only with LeuLeuAla and AlaLeu (b) The binding affinity of valganciclovir and valacyclovir was measured with the same set of ligands after labeling DtpA. The error bars indicate the standard deviation calculated from two measurements.

3.4.5. *In vivo* uptake and competition assay with the ligand library

The ligand selectivity of DtpA was studied previously by other groups using the *in vivo* AK-AMCA (β-Ala-Lys-N-7-amino-4-methylcoumarin-2-acetic acid) uptake and competition assay (Weitz *et al*. 2007). For this method as described in detail in Methods section 2.2.3.5, cells overexpressing DtpA are incubated with the dipeptide mimetic coupled to a fluorescent dye (AK-AMCA), DtpA takes up the fluorescent molecule into the cell and the fluorescence trapped in the cells is measured by a fluorometer. For the competition experiment, a ligand is added along with AK-AMCA, which competes with its uptake. Thereby, the lower the AK-AMCA uptake is, the higher is the competition from the ligand, which is interpreted as binding and possibly transport of the ligand by DtpA. I have established this method in our lab to compare it with the results from the thermal stability and binding affinity measurements that are both *in vitro* assays. At 5 mM ligand concentration all tripeptides, almost all dipeptides and from the drugs valacyclovir and valganciclovir lead to strong competition (Figure 27). The dipeptides, which have a slightly lower competition rate (20-60 % AK-AMCA uptake rate), contain either charged residues in the N-terminal and C-terminal position or a glycine residue. To compare the competition rate of the ligands within the library, the concentration was reduced to 0.5 mM (Figure 27). At this concentration, we observe a similar trend compared to the results from the thermal stability assay (Figure 23). Overall, the tripeptides show a stronger competition relative to the dipeptides and the ligands with hydrophobic residues are better competitors than the ligands with charged residues. From the drug library, valacyclovir and valganciclovir show a similar competition rate to di-/tripeptides.

A box plot analysis was performed by Dr. Jan Strauss from EMBL Hamburg based on the results at 0.5 mM ligand concentration (Figure 27) for comparing the tripeptides and dipeptides. The AK-AMCA competition rate was calculated by subtracting the uptake rate from 100 %. The AK-AMCA competition rate of tripeptides have a median of 68 % whereas the dipeptides only 53 % ($p=0.005$). This is in good agreement with the results from the thermal stability data (Figure 23).

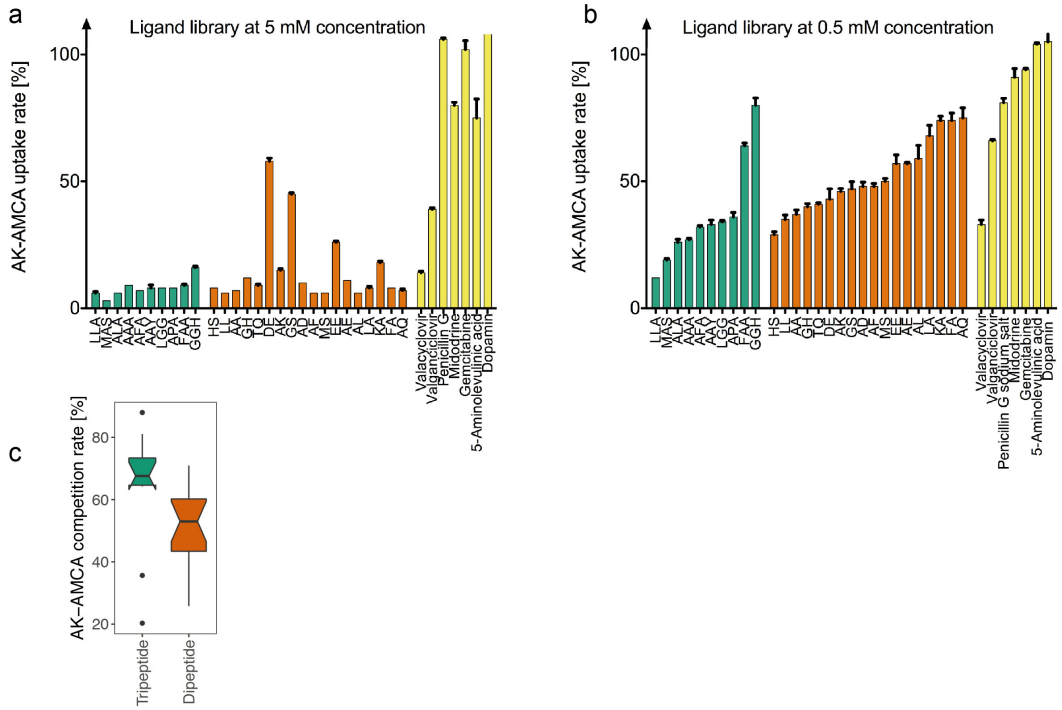


Figure 27: AK-AMCA competition assay with the ligand library
AK-AMCA competition assay with the ligand library at (a) 5 mM and (b) 0.5 mM. The error bars indicate the standard deviation calculated from three experiments. (c) Box plot analysis based on the results at 0.5 mM ligand concentration and subtracted from 100 % for better comparison to the thermal stability results. The 95 % quartile is drawn in green for tripeptides and orange for dipeptides with the lines above and underneath representing the upper and lower extremes. The outliers are shown as dots.

3.4.6. Concentration dependent *in vivo* competition assay

Next, the tripeptide, LeuLeuAla, and the dipeptide, AlaLeu, as well as the two drugs, valacyclovir and valganciclovir, were further investigated by concentration dependent *in vivo* AK-AMCA competition measurements. The IC50 and K_i values were calculated from the curve as described in Methods section 2.2.3.5. All IC50 and K_i values are in low to mid μM range with the lowest value for LeuLeuAla. This is in good agreement with the *in vitro* binding affinity (K_d) values (Table 15, Figure 26) and the described values from the literature (Weitz *et al*. 2007).

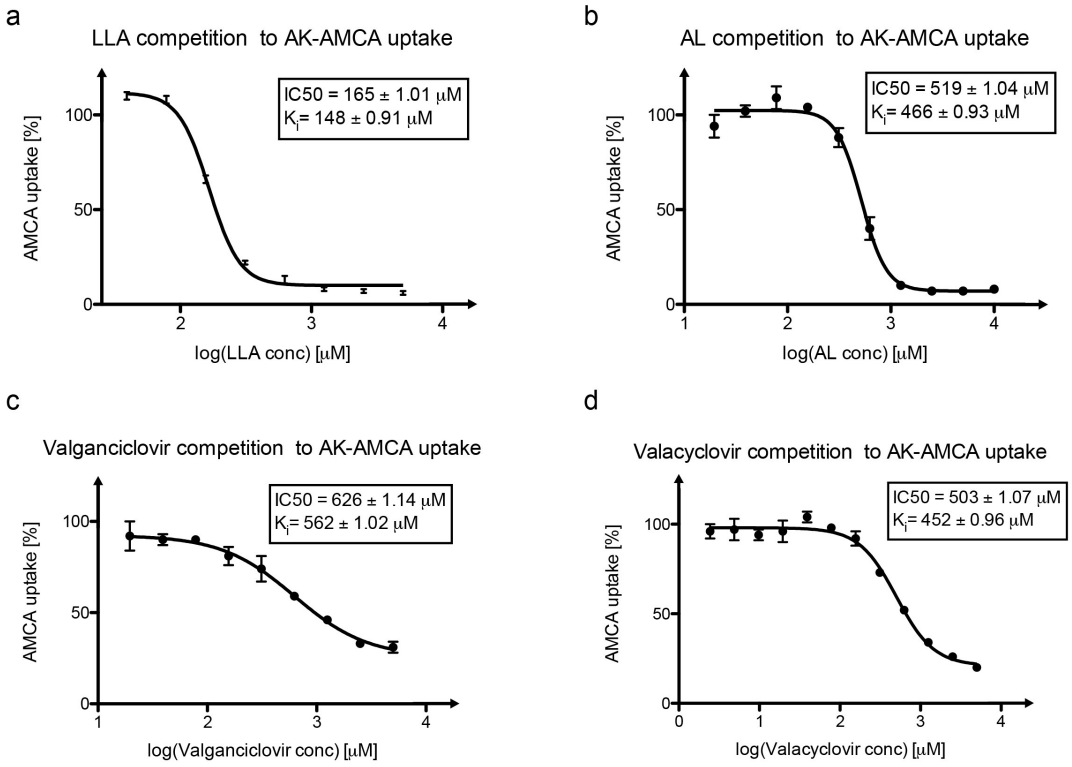


Figure 28: Concentration dependent AK-AMCA competition assay
Concentration dependent AK-AMCA competition assay with (a) LeuLeuAla, (b) AlaLeu, (c) valganciclovir and (d) valacyclovir. The curve is fitted by log(inhibitor) vs response equation in GraphPad Prism v5 and the error bars indicate the standard deviation calculated from three experiments.

3.5. Crystallization and structure determination

Crystallization was performed to determine the structure of DtpA in the road of understanding its ligand selectivity and drug transport capability. We used the vapor diffusion and lipidic cubic phase crystallization methods to solve the structure, although only the former method yielded crystals. Three structures were determined and uploaded in the Protein Data Bank (PDB) as part of this thesis: DtpA-N00 complex in MES buffer (PDB id: 6GS1), DtpA-N00 complex in glycine buffer (PDB id: 6GS7) and DtpA-N00 in complex with valganciclovir (in glycine buffer) (PDB id: 6GS4). The first two are ligand-free structures determined in different crystallization conditions and the third one is co-crystallized with valganciclovir. This is the first structure of a proton-dependent oligopeptide transporter (POT) in complex with a commercial drug.

3.5.1. DtpA-N00 complex in MES buffer

To solve the structure of DtpA, sitting drop vapor diffusion (Methods section 2.2.4.1.1) and lipidic cubic phase (Methods section 2.2.4.1.2) crystallization methods were performed. No initial crystals for DtpA were observed with the used commercial crystallization screens in both methods. Nanobodies were generated against DtpA as crystallization chaperones and all four generated nanobodies could bind and stabilize DtpA as shown in section 3.2.2. Furthermore, initial crystals were observed for DtpA with all four nanobodies with the vapor diffusion method using the same commercial screens but not using the lipidic cubic phase method. Optimization screens were prepared for several conditions including at least one condition for each DtpA-nanobody combination. Although crystals could be obtained in most cases, only the crystals of DtpA-N00 complex significantly improved through the optimization cycles and yielded diffraction in higher resolution ranges. The crystals to determine the first structure of the DtpA-N00 complex (in MES buffer) grew in 0.02 M $MgCl_2$, 0.1 M MES, pH 6.5, 32 % PEG 600 condition. The crystals first appeared on day 14 and grew until day 28 (Figure 29) and were flash frozen without an additional cryoprotectant, because the high concentration of PEG 600 was acting as a natural cryoprotection.

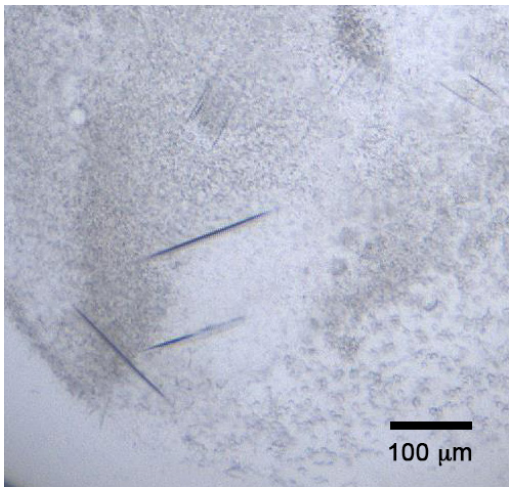


Figure 29: Crystals of DtpA-N00 complex grown in MES buffer condition
Crystals of DtpA-N00 complex grown in the MES buffer condition are either needles or thin plates. Picture was taken on day 28. One of these crystals led to the first structure of DtpA.

The diffraction data were collected at the fully automated MASSIF-1 (ID30A-1) beamline at the European Synchrotron Radiation Facility (ESRF, Grenoble France) under cryo conditions (Bowler *et al*. 2015). XDS and XSCALE packages were used for data processing (Kabsch 2010). The space group was identified as $P2_12_12_1$ and the unit cells were 55.46, 120.72, 63.33 Å and 90° for all three angles (Table 16). The initial phases were obtained by molecular replacement in phaser (McCoy *et al*. 2007) as part of phenix (Adams *et al*. 2010) using the search models of another bacterial POT DtpD (PDB id: 4Q65) and a nanobody. The model was refined in iterative cycles of manual editing of the model in COOT (Emsley and Cowtan 2004) and refinement steps in phenix software. The resolution extended to 3.3 Å and the completeness was over 99 %. DtpA could be modeled starting from residue 17 to 488 (full length 500 residues) without any missing loops and N00 from residue 1 to 125 (full length 132 residues). A single water molecule was included in the model, which is in the ligand-binding site and is coordinated by R34 and N126 of DtpA. Overall, DtpA is in the inward open conformation and the N00 is binding from the periplasmic side of DtpA as shown in Figure 30, which is further discussed in section 3.6.2.

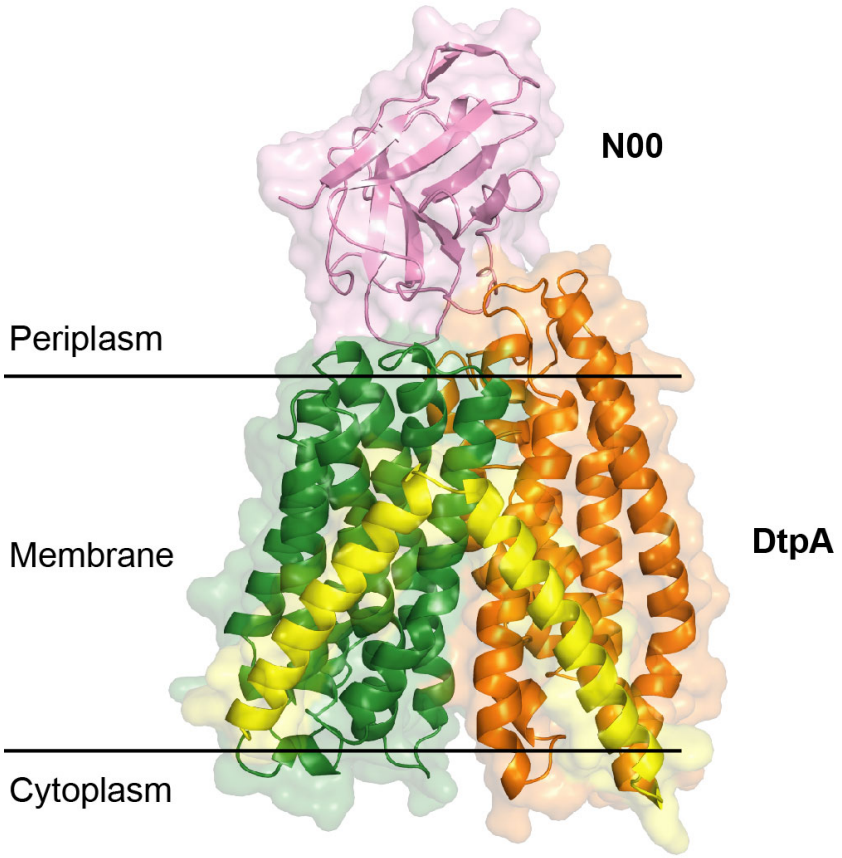


Figure 30: The structure of DtpA-N00 complex in MES buffer
The DtpA-N00 structure is illustrated in surface and cartoon representation. The N-terminal bundle of DtpA is colored green, C-terminal bundle in orange and HAHB domain in yellow and the nanobody N00 in pink. Presumed membrane orientation is represented with black lines.

3.5.2. DtpA-N00 complex in glycine buffer

In parallel to the MES buffer condition, other conditions were further optimized for crystallizing DtpA-N00. The second successful crystallization condition was 0.15 M $CaCl_2$, 0.1 M glycine, pH 9.0 and 35 % PEG 400. Two types of crystals grew under this condition, needle shaped and cubic (Figure 31). Although the cubic crystals were larger in volume, they diffracted only up to 15 Å. The needle shaped crystals were more promising. The crystals leading to the DtpA-N00 complex in the glycine buffer appeared between day 60 and 80 and were flash frozen without an additional cryoprotectant due to the presence of high concentrations of PEG 400.

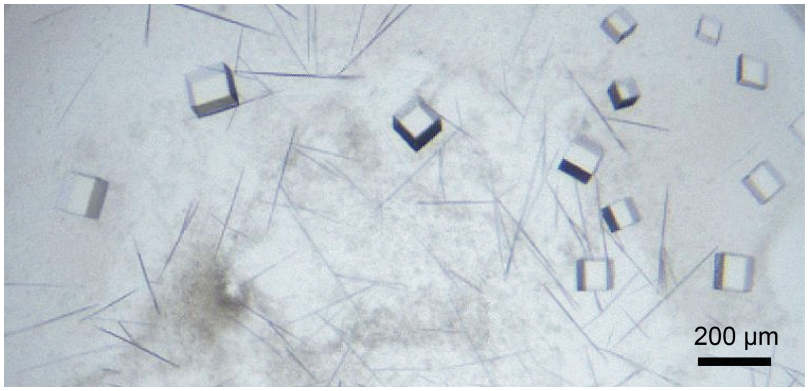


Figure 31: Crystals of DtpA-N00 complex in glycine buffer
In the glycine condition DtpA-N00 complex gave rise to two different crystal forms, a cubic and a needle shape. Only the needle formed crystals yielded diffraction in high resolution range. The picture is taken on day 5.

The diffraction data were collected at P14 beamline operated by EMBL Hamburg at PETRA III (DESY, Hamburg, Germany) under cryo conditions. The same method was used for data processing as for DtpA-N00 complex structure determination (Methods section 2.2.4.2). The space group was the same as for the former structure and the unit cell parameters were almost identical. The resolution extended to 3.3 Å as for the previous structure. The search model for the molecular replacement step was the DtpA-N00 structure in MES buffer (PDB id: 6GS1). The data collection and refinement statistics are listed in Table 16. DtpA could be modeled starting from residue 17 to 487 (full length 500 residues) without any missing loops and N00 from residue 1 to 125 (full length 132 residues). The DtpA-N00 complex in glycine buffer is very similar to the structure in MES buffer (Figure 30) with a rmsd of 0.36 Å over 3985 atoms. The optimization of this condition was crucial due to the difficulties in the reproducibility of the MES condition in contrary to the more robust nature of this condition for co-crystallization with ligands.

3.5.3. DtpA-N00 complex with ligands

Several commercial and self-made optimization crystallization screens based on MES and glycine buffer conditions were tested with the DtpA-N00 mixed with one of the ligands LeuLeuAla, AlaPheAla, valganciclovir and valacyclovir, in low mM concentrations in the crystallization drop. The ligands were selected based on the thermal stabilization and binding affinity results (Figure 23 and Table 15). With each ligand initial crystals were obtained in the glycine condition. The crystals were in all

cases very tiny needles with a high nucleation rate. Only DtpA-N00 with valganciclovir crystals improved after several optimization rounds and diffracted to high resolution.

3.5.3.1. DtpA-N00 complex with Valganciclovir

The crystals of DtpA-N00 in complex with valganciclovir grew in the condition 0.15 M $CaCl_2$, 0.1 M glycine, pH 9.0, 35 % PEG 400 and 0.02 % Anapoe-C12E10. The crystal appeared on day 28 and grew until day 55. They were flash frozen without addition of cryoprotectant, as high concentrations of PEG 400 was present in the drop. The diffraction data were collected at the P13 beamline operated by EMBL Hamburg at PETRA III (DESY, Hamburg, Germany) under cryo conditions. The data processing was performed as described for the structure determination of DtpA-N00 complex in glycine buffer. The resolution extended further to 2.65 Å. The space group was the same and the unit cells were almost identical (Table 16). The residues 17-488 of DtpA (full length 500 residues) and 1-122 of N00 (full length 132 residues) were modeled. The DtpA residues 143-144 and 342-343 from the loops within TM4-TM5 and TM8-TM9 respectively were excluded from the model due to weak electron density. They are both loops on the cytoplasmic side of the protein and thereby on the opposite side of the N00 binding site. Additionally, one valganciclovir, one dodecylmaltoside (DDM) and seven water molecules were built in the electron density. The DDM molecule is on the outer surface of the DtpA in the proximity of the TM9 and TM12 with the sugar moiety close to the expected water lipid interface on the periplasmic side of the protein and the carbon chain aligned along the TM9. The water molecules were not in the ligand-binding site but in several other positions spread across the inner cavity of DtpA. There was a clear electron density to build valganciclovir in the ligand-binding site (Figure 32).

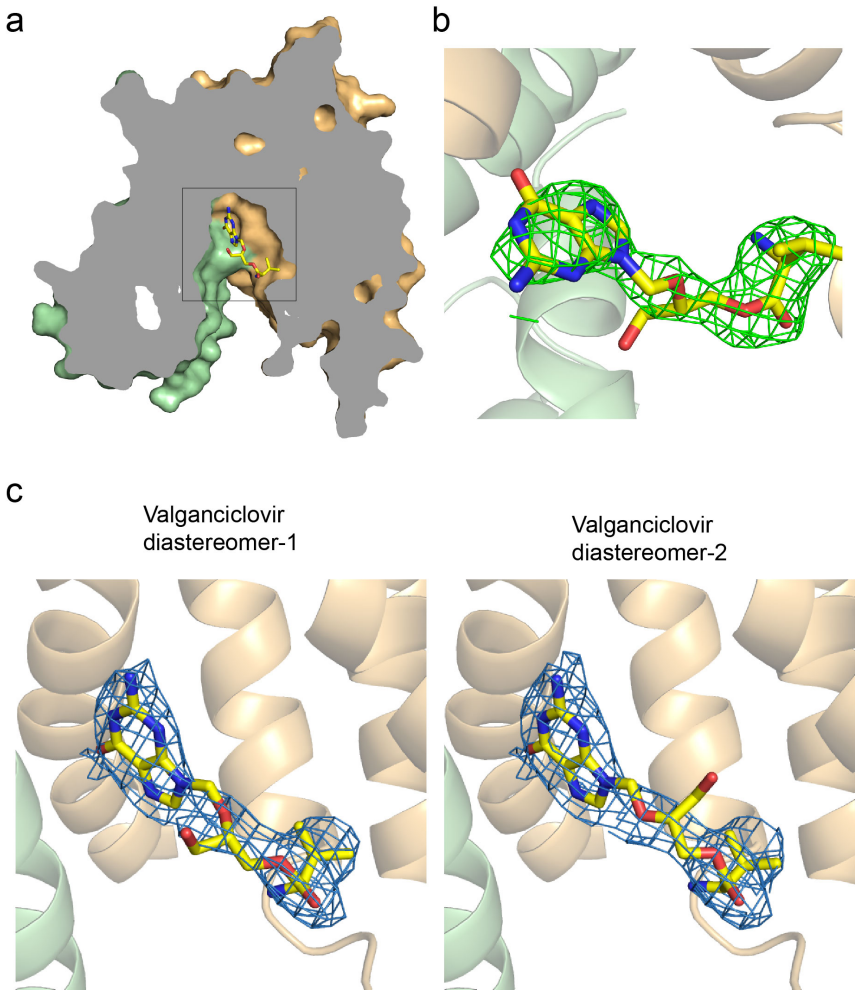


Figure 32: Valganciclovir in the ligand-binding site of DtpA
(a) DtpA in surface representation cut in half with valganciclovir in the ligand-binding site. (b) Omit map of valganciclovir countered at 3 σ level. (c) The $2F_o$-F_c map of DtpA refined with both diastereomers countered at 1 σ level. All further figures are prepared with diastereomer-1 (left).

Valganciclovir is a mix of two diastereomers and both were tested for model building independently. The backbone of valganciclovir fits the electron density in both cases equally well and there is no density corresponding to the -CH_2OH group at the chiral center for either of the diastereomers (Figure 32). This can be explained by a mixture of the two diastereomers in the ligand-binding site of DtpA within the crystal, therefore the electron density of the -CH_2OH group is weaker.

3.6. Analysis of the DtpA-N00 structure

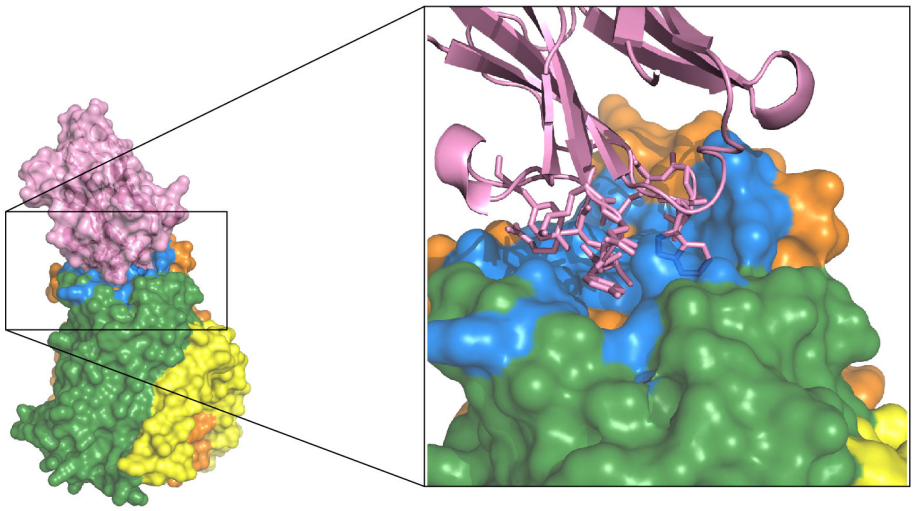
3.6.1. Overall structure

The DtpA structure contains 14 transmembrane α-helices (TM), which are organized in two bundles: TM 1 to 6, the N-terminal bundle, and TM 7 to 12, the C-terminal bundle, forming the transporter core and two additional TMs, HA-HB, placed in between (Figure 30). This is in agreement with the previously published bacterial POT structures (Newstead *et al*. 2011; Martinez Molledo *et al*. 2018; Zhao *et al*. 2014) and the predictions based on the sequence analysis of DtpA. The structure is in inward open conformation as the formerly published structures from the POT family, except $PepT_{So}$ and $PepT_{St}$, which were also captured in partially occluded inward open conformation. DtpA shows high structural similarity in the transporter core with rmsd values between 1.3 to 2.7 Å along the peptide backbone to previous POT structures.

3.6.2. DtpA-N00 interaction site

In the DtpA-N00 complex structure, N00 binds to DtpA on the periplasmic side of the transporter. Nanobodies have three loops, called the complementarity-determining region (CDR) H1, H2 and H3, which are involved in antigen binding. Especially CDR H3 is a longer loop, which can bind into crevices. In the case of the DtpA-N00 interaction, the CDR loops H2 and H3 of the N00 extend into the crevices on the periplasmic side of DtpA, spanning across both the N- and C-terminal bundle (Figure 33).

a



b

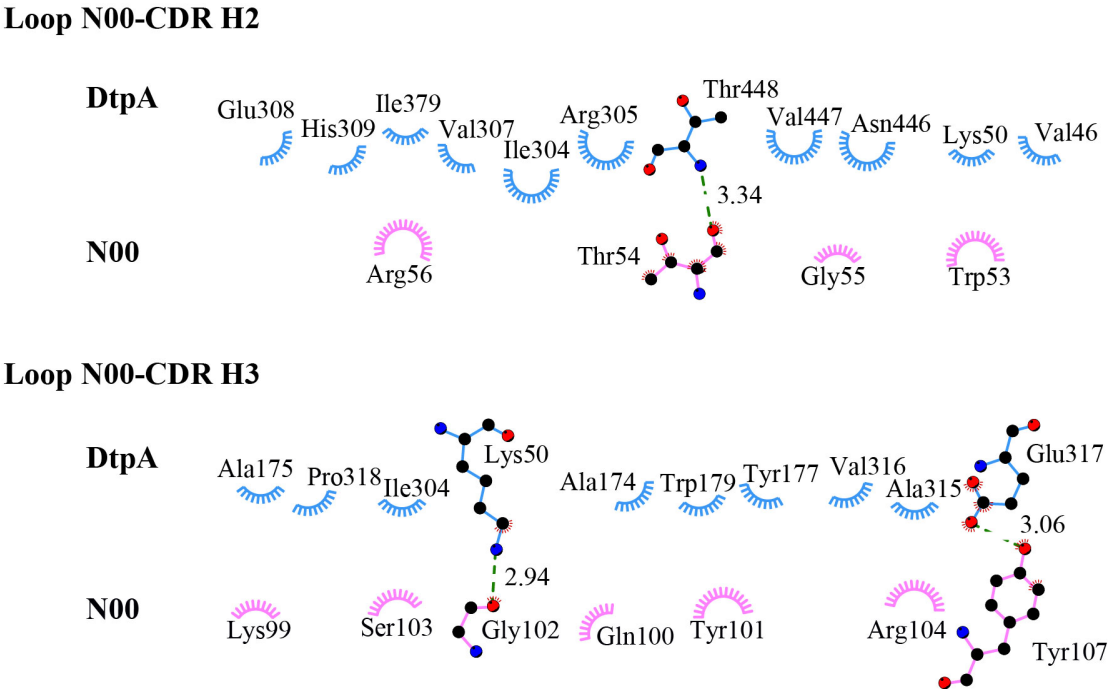


Figure 33: The binding surface of DtpA and N00
(a) Surface representation of DtpA with N-terminal bundle in green and C-terminal bundle in orange, the residues involved in N00 binding are highlighted in blue. The N00 is in cartoon representation and the residues on the two loops, CDR H2 and H3, involved in DtpA binding are shown in sticks model. (b) LigPlot (Laskowski and Swindells 2011) analysis of the DtpA-N00 interaction site.

The interaction site was investigated with LigPlot (Laskowski and Swindells 2011) (Methods section 2.2.4.3). There are hydrophobic interactions between 18 amino acids of DtpA and 8 amino acids of N00 (Figure 33). Additionally, there are three hydrogen bonds contributing to the interaction. Namely, between the main chain

carbonyl group of N00 T54 and the main chain amino group on DtpA T448, between the main chain carbonyl group N00 G102 and side chain primary amine group of DtpA K50 and between the side chains of N00 Y107 and DtpA E317. The protein-protein interaction surface area is approximately 4916 $Å^2$ based on the calculations in PyMOL. With these interactions N00 stabilizes the inward-open conformation of DtpA.

3.6.3. The role of N00 in the crystallization of DtpA

In the crystal structure, the asymmetric unit consists of one DtpA and one N00 molecule and the symmetry is $P2_12_12_1$. Within the symmetry mates, one N00 molecule is in the center of four DtpA molecules forming crystal contacts (Figure 34). One of the four binding sites is the DtpA-N00 binding interface with the N00 CDR H2 and H3 loops intruding into the cavities on the periplasmic side of DtpA (Figure 33). The other three interaction sites are crystal contact points, which do not occur in solution based on our observations that DtpA is a monomer and binds in a 1:1 ratio with N00 using analytical gel filtration analysis. Here, N00 interacts with the loops on the cytoplasmic side of DtpA, different than the main DtpA-N00 interaction, which occurs on the periplasmic side of DtpA. Noticeably, there are no direct interactions between two DtpA molecules but the nanobody plays the central part in crystal contact formation.

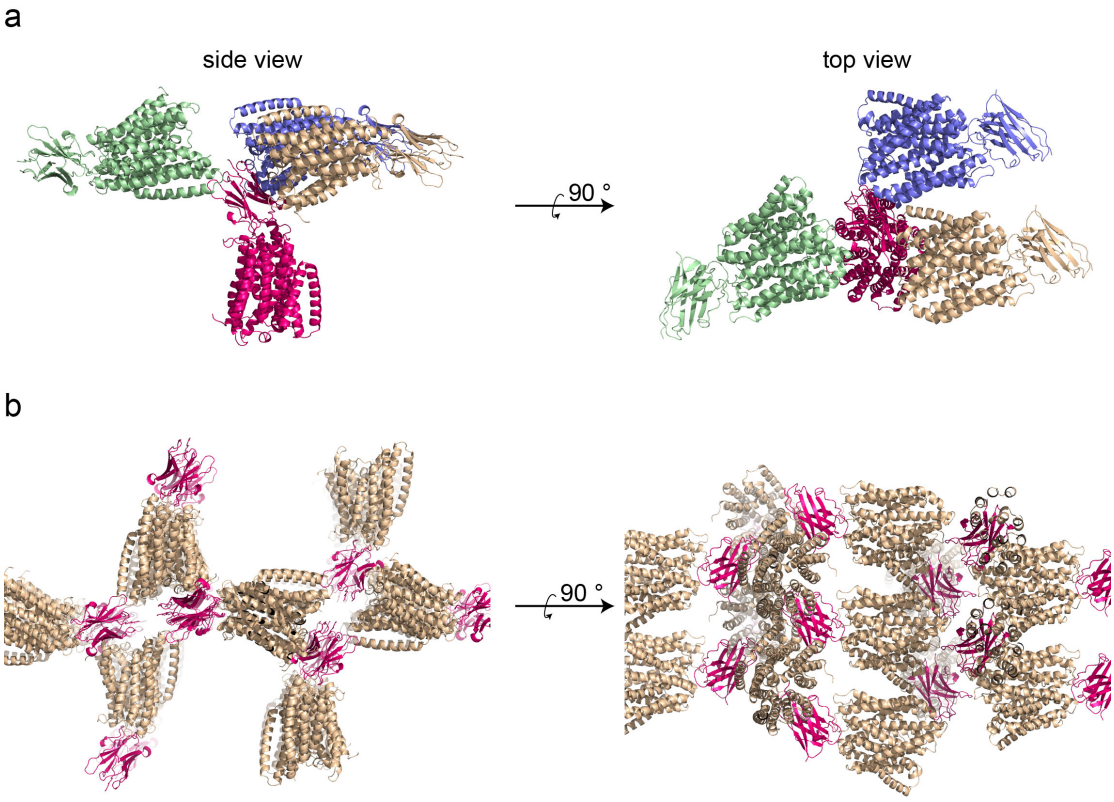


Figure 34: Crystal packing of DtpA-N00
(a) Four DtpA-N00 complexes are shown in different colors: green, blue, wheat and pink. The N00 in pink makes crystal contacts with the loops on the cytoplasmic side of the DtpA in other symmetry mates. (b) DtpA is colored in wheat and N00 in pink for all symmetry mates.

There are large volumes of empty space in the crystal packing, as can be seen in Figure 34, which is typical for type 2 membrane protein crystals (Figure 35). In this crystal type, each membrane protein is covered by the detergent micelle, which is not organized in a crystalline order and therefore, not visible in the electron density map. The type I crystal form is typical for the LCP crystallization method and leads to densely packed crystals with higher signal-to-noise levels and often better resolution; unfortunately no crystals were obtained using the LCP method for DtpA with or without nanobodies.

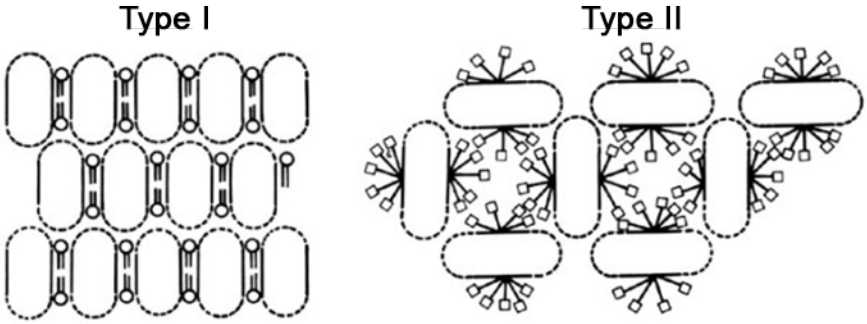


Figure 35: Type I and type II crystal forms of membrane proteins
Membrane protein crystallization is often performed by the lipidic cubic phase or vapour diffusion methods, which lead to different packing features of the membrane protein crystals. In a type I crystal, which is typical for LCP method, the proteins are in a continuous membrane layer and can pack densely, whereas in type 2 crystals, typical for vapor diffusion method, the detergent micelle prevents the proteins from packing densely and there is more space between the proteins. Figure modified from (Gouaux 1998).

3.6.4. Valganciclovir in the ligand-binding site of DtpA

The DtpA-N00 complex was co-crystallized with valganciclovir. We were able to model the valganciclovir with high confidence based on the electron density in the ligand-binding site. Valganciclovir is a L-valyl ester of ganciclovir, which is an antiviral drug, and has a similar size to tripeptides. Based on the previous observations that the N-terminal residue of a di-/tripeptide plays an important role in the coordination of the ligand in the binding site and the N-terminal residue of di- and tripeptides occupies the same position, it was suggested that the N-terminal valine of valganciclovir would also superimpose with the N-terminal residue of di-/tripeptides in the binding site (Figure 36).

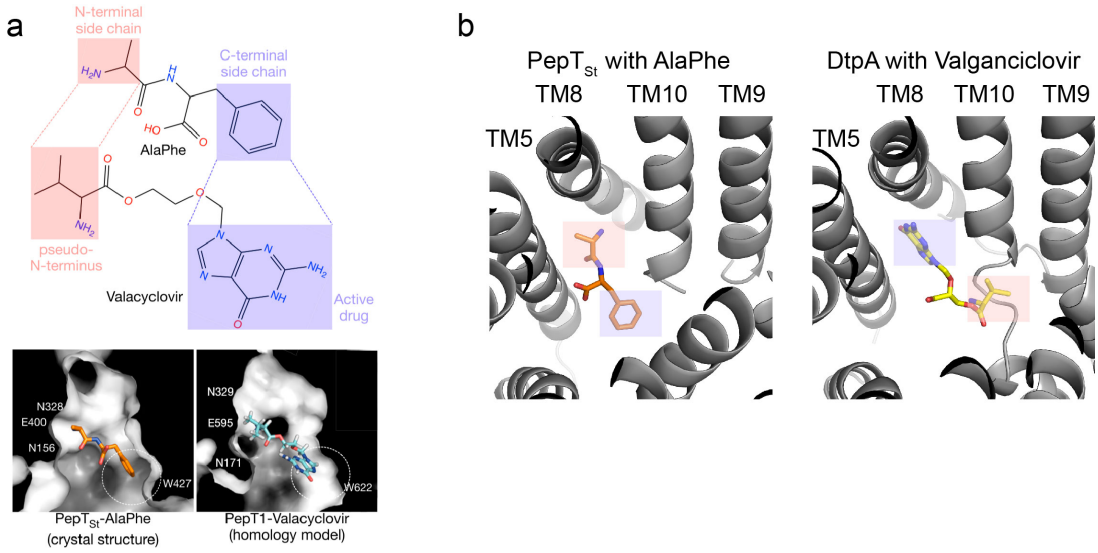


Figure 36: The binding mode of valganciclovir
(a) The expected binding mode in the literature with the N-terminal valine of valacyclovir mimicking the N-terminal residue of a dipeptide, both highlighted in pink box, and the guanine ring mimicking an aromatic side chain, both in blue box. Figure from (Samsudin *et al*. 2016). (b) The comparison of $PepT_{St}$ structure with AlaPhe and DtpA structure with valganciclovir. The guanine ring is mimicking the N-terminal alanine of the dipeptide and the valine of valganciclovir is in a pocket created by the TM10 intrahelical loop that is not present in $PepT_{St}$.

Valganciclovir in this expected binding mode did not fit into the electron density observed in the ligand-binding site but rather in a 180° rotated binding mode. Here, the guanine ring mimics the position of an N-terminal residue, whereas the valine moiety intrudes into a pocket, which is smaller in most other POTs with known structure as discussed in section 3.6.6.

The residues coordinating the guanine ring of valganciclovir is equivalent to the residues coordinating the N-terminal residue of a di-/tripeptide. Specifically, DtpA N160 (TM5), N325 (TM8) and E396 (TM10) make hydrophobic interactions with the guanine ring of valganciclovir and the hydroxyl group of DtpA Y156 is likely to form a hydrogen bond with an amine group within the guanine ring of the valganciclovir (Figure 37). For one of the diastereomers of valganciclovir the hydroxyl group of Y38 is in the distance for a hydrogen bond with the hydroxyl group in the center of the molecule. F289 (TM7) as well as I399 and L402 (TM10), which are part of the intrahelical loop discussed in next section, coordinate the valine residue of valganciclovir. Overall, valganciclovir guanine ring mimics the interactions with a di-/tripeptide but the position of the valine is unique due to this intrahelical loop within TM10.

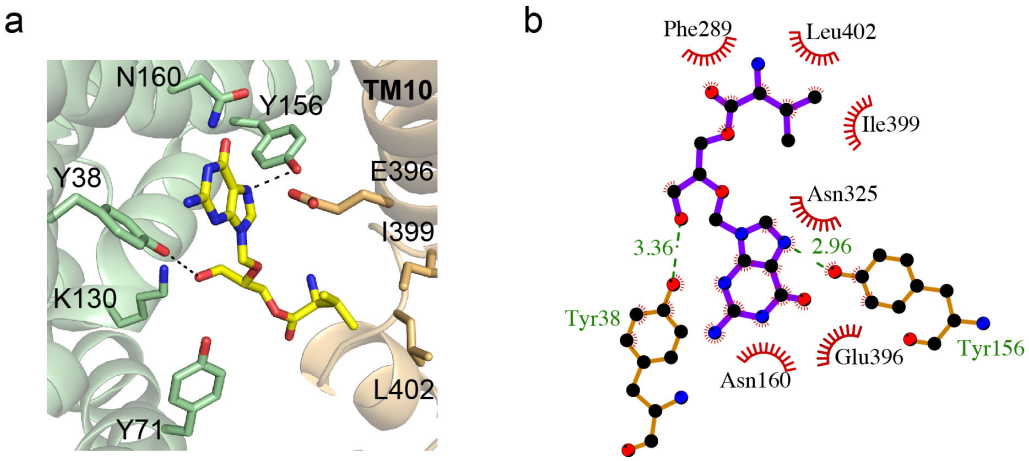


Figure 37: Valganciclovir in the ligand-binding site of DtpA
(a) The residues involved in the ligand coordination are shown in sticks model. The residues F289 and N325 are excluded for clarity and the residues K130 and Y71 are included, because they are typically involved in ligand binding and point towards the ligand (Martinez Molledo *et al*. 2018; Guettou *et al*. 2014), although they are further than 3.5 Å at this conformation. (b) LigPlot (Laskowski and Swindells 2011) analysis of the DtpA-valganciclovir binding.

3.6.5. The difference in TM10-TM11 between DtpA and other POTs

The major structural difference between DtpA and the previously published POT structures is in the TM10 and TM11 region, which is relevant for the coordination of the valganciclovir valine moiety. A discontinuity of the TM10 was observed in $PepT_{So2}$ (Guettou *et al*. 2013) but was not further investigated. In the DtpA structure, this discontinuity is even more pronounced than in $PepT_{So2}$ and is involved in the coordination of valganciclovir. The periplasmic half of TM10-11 in DtpA and $PepT_{So2}$ superimposes well with other POT structures. An intrahelical loop in DtpA (I399-SGLG-L404) and $PepT_{So2}$ at the level of the ligand-binding site causes a shift in the cytoplasmic half of TM10 towards TM11 and pushes TM11 further away creating a larger pocket in the ligand-binding site (Figure 38). To investigate the length of this shift, we have overlaid TM10-11 of DtpA (PDB id: 6GS4) and $PepT_{St}$ (PDB id: 5OXN) and measured the distance between the furthest points of TM10 and TM11 individually. The L404 Cα of TM10 in DtpA and the equivalent L408 Cα atom in $PepT_{St}$, are 6.1 Å apart (represented with an double-headed arrow (d1) in Figure 38) and F424 Cα of TM11 in DtpA and the equivalent F428 Cα in $PepT_{St}$ are 6.8 Å apart

(represented with an double-headed arrow (d2) in Figure 38). The volume of binding pocket near TM10-11 was further investigated as described in the next section.

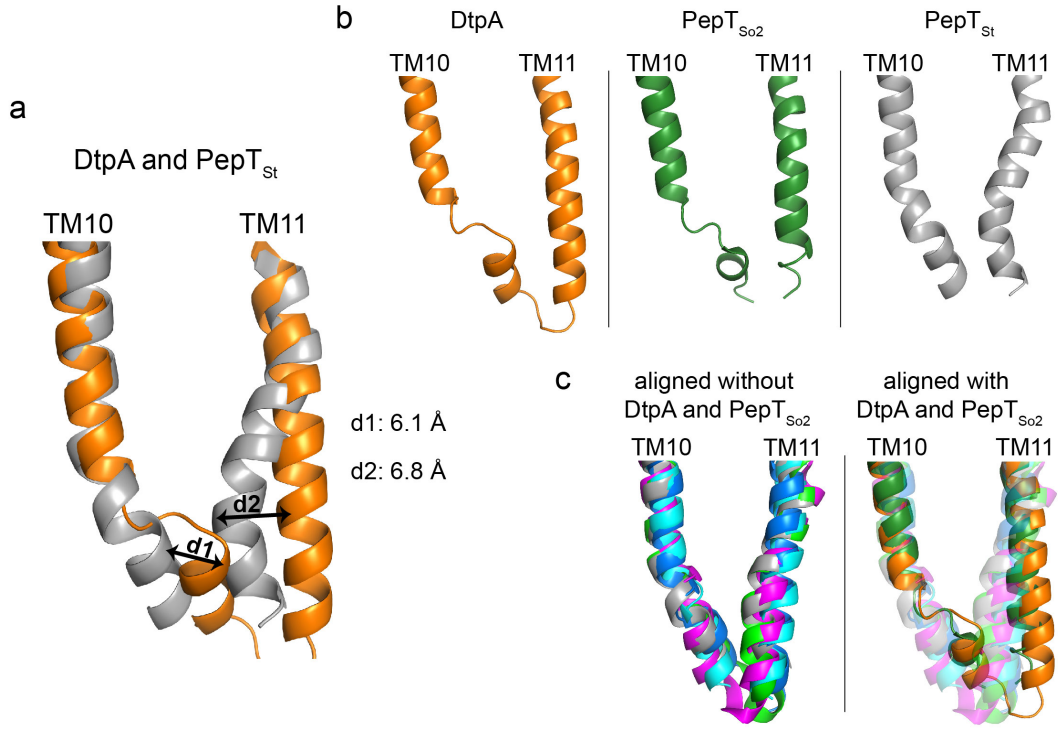


Figure 38: The difference in TM10-TM11 between DtpA and other POT structures
(a) Comparison of DtpA (PDB id: 6GS4 in orange) and PepT$_{St}$ (PDB id: 5OXN in grey). The distance between equivalent residues of DtpA and PepT$_{St}$ at the furthest positions are shown with a double-headed arrow. (b) Side-by-side views of DtpA, PepT$_{So2}$ (PDB id: 4LEP in dark green) and PepT$_{St}$. Notably, DtpA and PepT$_{So2}$ have an intrahelical loop in TM10 and PepT$_{St}$ does not. (c) left: TM10-11 of other POTs than DtpA and PepT$_{So2}$ aligned, namely PepT$_{So}$ (PDB id: 4UVM in magenta), GkPOT (PDB id: 4IKV in light green), DtpD (PDB id: 4Q65 in cyan), and YePepT (PDB id: 4W6V in blue), right: left image is shown in 50% transparency and overlaid with the TM10-11 of DtpA and PepT$_{So2}$ for comparison.

3.6.6. Binding pocket volume analysis with focus on TM10-TM11

The ligand-binding site of POTs is globular with several side pockets, in which different side chains of the di-/tripeptides and (pro)drugs can fit. The binding pocket near TM10-11 was investigated for DtpA and other POT structures by calculating the volume of the ligand binding site using POVME 2.0 (Durrant *et al*. 2014) as described in Methods section 2.2.4.3.

For the volume calculation the Cα atom of the valine residue in valganciclovir was taken as center and the volume at 4 and 7 Å radii was calculated for all aligned structures with default parameters. At both radii, DtpA, $PepT_{So2}$ and DtpD cluster together with a larger pocket near TM10-TM11 and the other POTs cluster together with a smaller pocket (Figure 39). We were surprised to see DtpD in the same category as DtpA and $PepT_{So2}$, because DtpD does not have an intrahelical loop (Figure 38) but TM11 of DtpD is rather similar to DtpA and $PepT_{So2}$ and the slight side chain differences may lead to a bigger pocket compared to other POTs without the TM10 intrahelical loop.

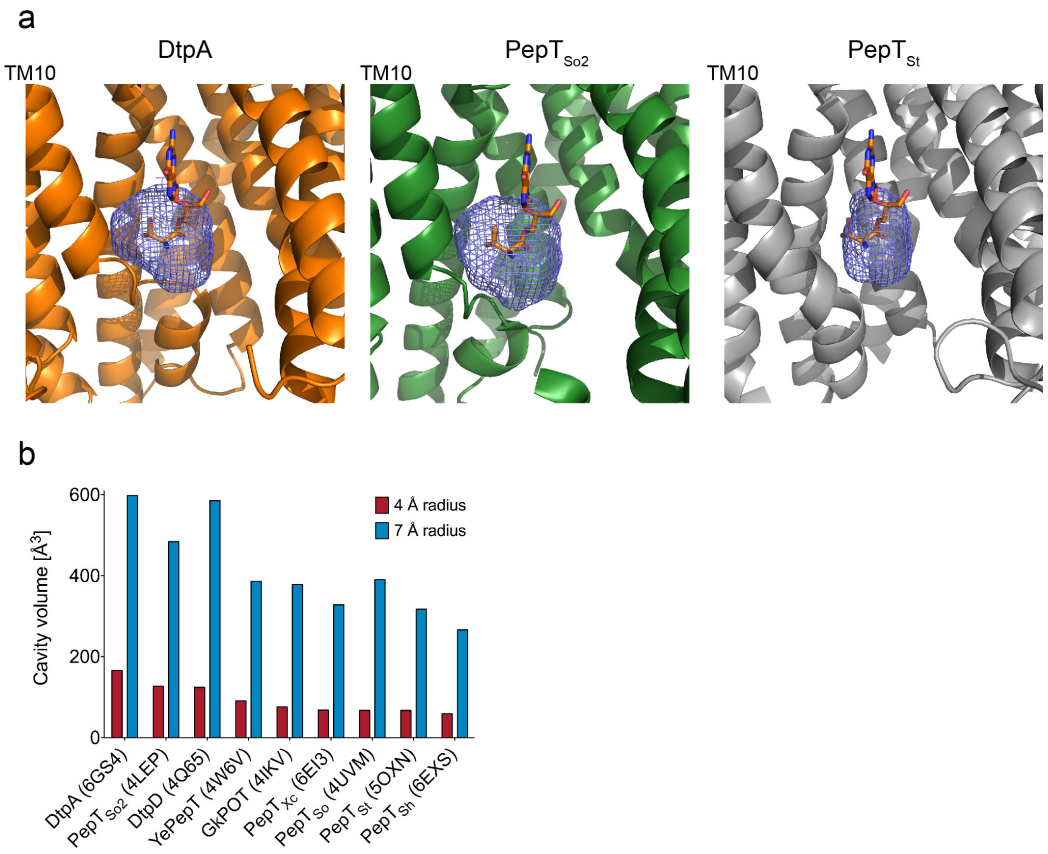


Figure 39: The ligand-binding site volume calculation focusing on TM10-TM11
(a) The calculated volume around the Cα atom of valine in valganciclovir is illustrated for DtpA (PDB id: 6GS4), $PepT_{So2}$ (PDB id: 4LEP) and $PepT_{St}$ (PDB id: 5OXN) in mesh representation at 4 Å radius. Valganciclovir is included in all three pictures for comparison. (b) The cavity volume for all published POT structures at 4 and 7 Å radii around the Cα atom of valine in valganciclovir.

3.7. Mutational characterization of DtpA

To investigate the ligand-binding site of DtpA, five residues were mutated to alanine as described in Methods section 2.2.1.6. The residues Y38, Y71, N160 and I399 were selected due to their interaction with valganciclovir in the DtpA structure (Figure 37). K130 was included because it coordinates di-/tripeptides in previous POT structures (Martinez Molledo *et al*. 2018; Guettou *et al*. 2014), although it was not in proximity of valganciclovir in DtpA structure. The five DtpA mutants were expressed and purified as DtpA WT and characterized by thermal stability assay in presence and absence of N00. The thermal stability of DtpA mutants is very similar to WT and they can all bind to N00 as shown by a strong thermal stabilization effect upon nanobody binding (Figure 40).

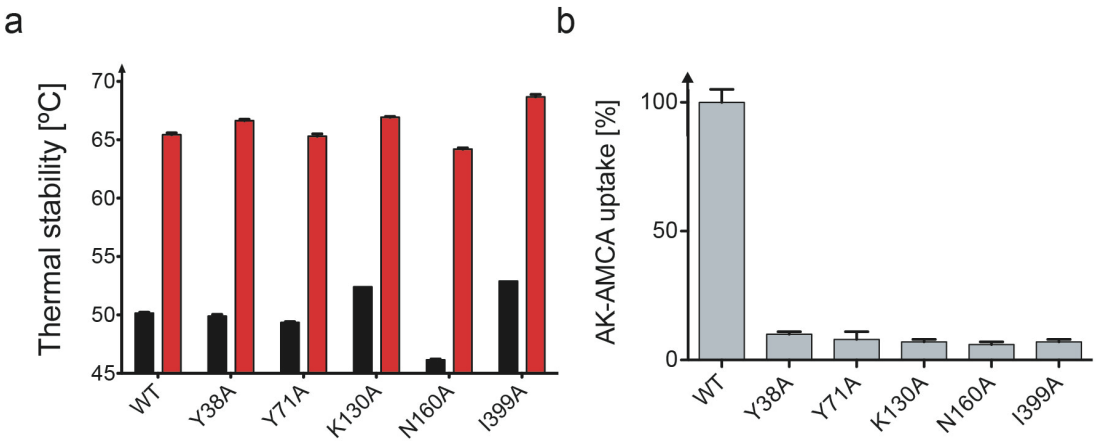


Figure 40: The functional studies with DtpA mutants
(a) Thermal stability analysis of DtpA WT and mutants with (in red) or without N00 (in black). (b) AK-AMCA uptake percentage of DtpA WT and mutants.

Next, the binding affinity of these mutants was tested by the *in vitro* microscale thermophoresis method followed by the *in vivo* AK-AMCA uptake assay. The binding affinity measurements of the mutants to the ligands valganciclovir, valacyclovir, LeuLeuAla and AlaLeu were either not possible to determine or reduced by several orders of magnitude and no AK-AMCA uptake was observed for the mutants (Figure 41 and Figure 42). Surprisingly, the K130A mutation affects the binding affinity to valganciclovir and valacyclovir similar to other mutations, although in the captured DtpA structure it is not in direct proximity of the ligand. It is likely that this residue plays a role in the drug coordination in another conformational

state than in the crystal structure. Overall, the functional characterization of the mutants is in good agreement with the assignment of the tested residues to the ligand-binding site in the crystal structure.

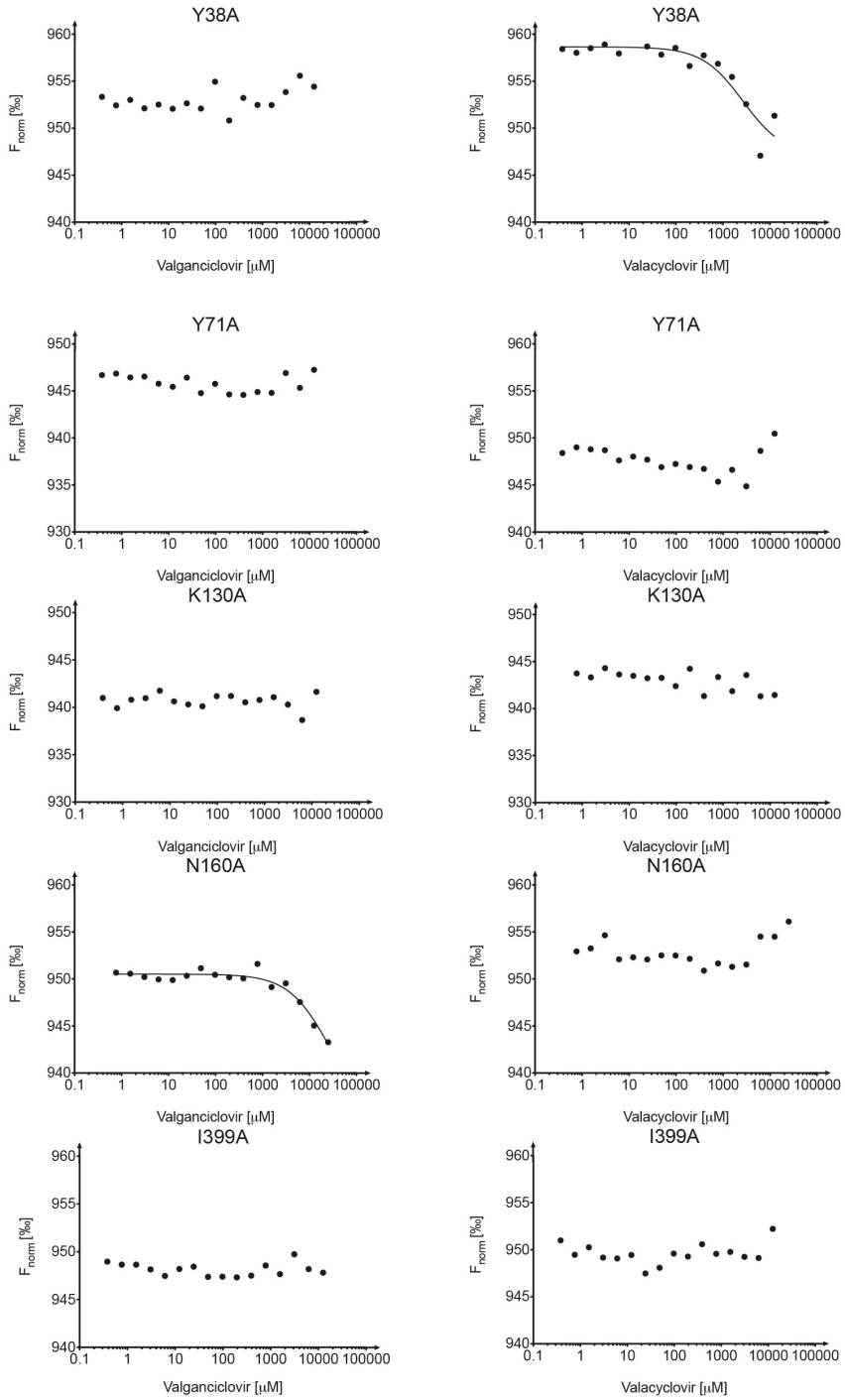


Figure 41: The binding affinity curves of DtpA mutants and drugs
Binding affinity measurements of DtpA mutants with the ligands valganciclovir and valacyclovir.

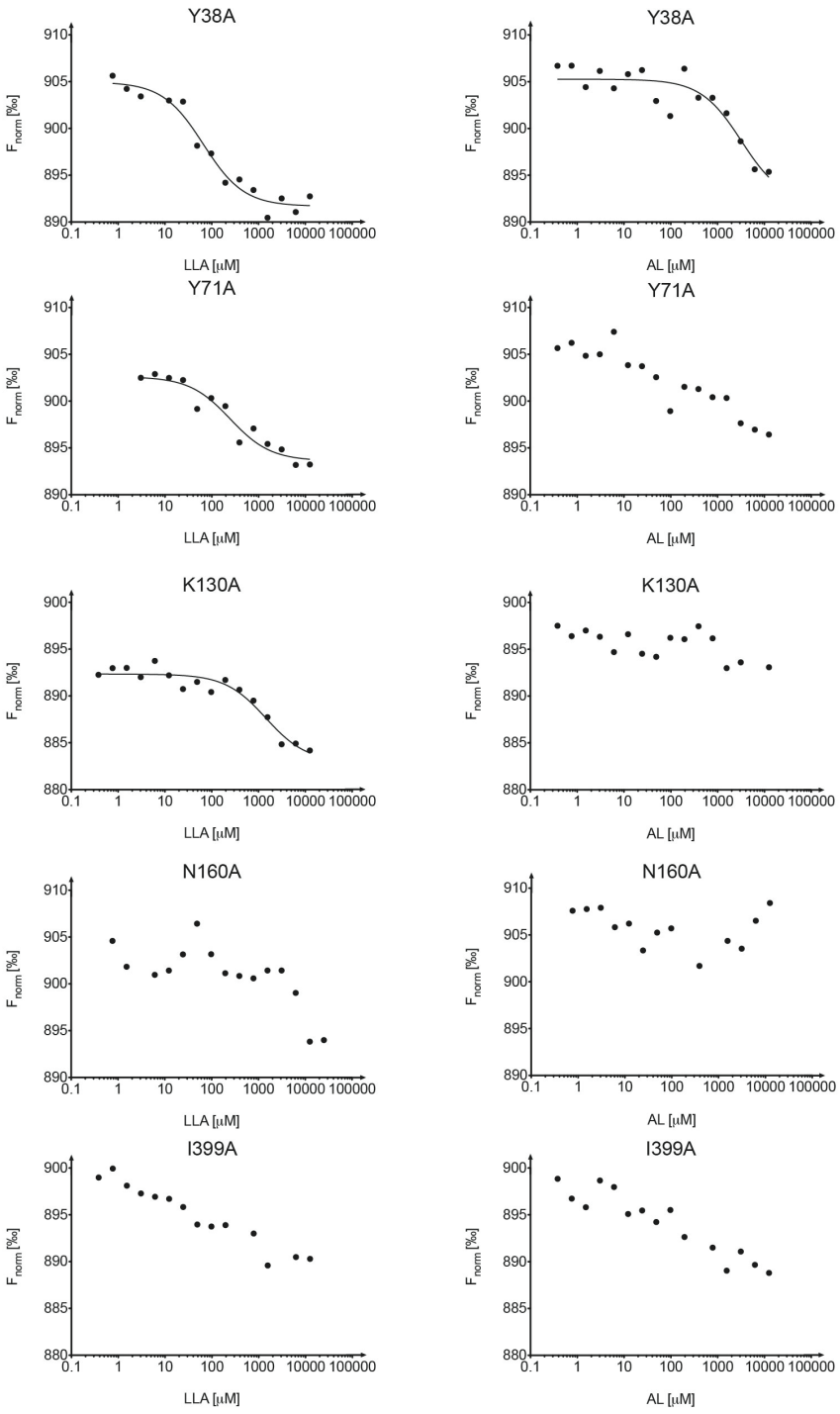
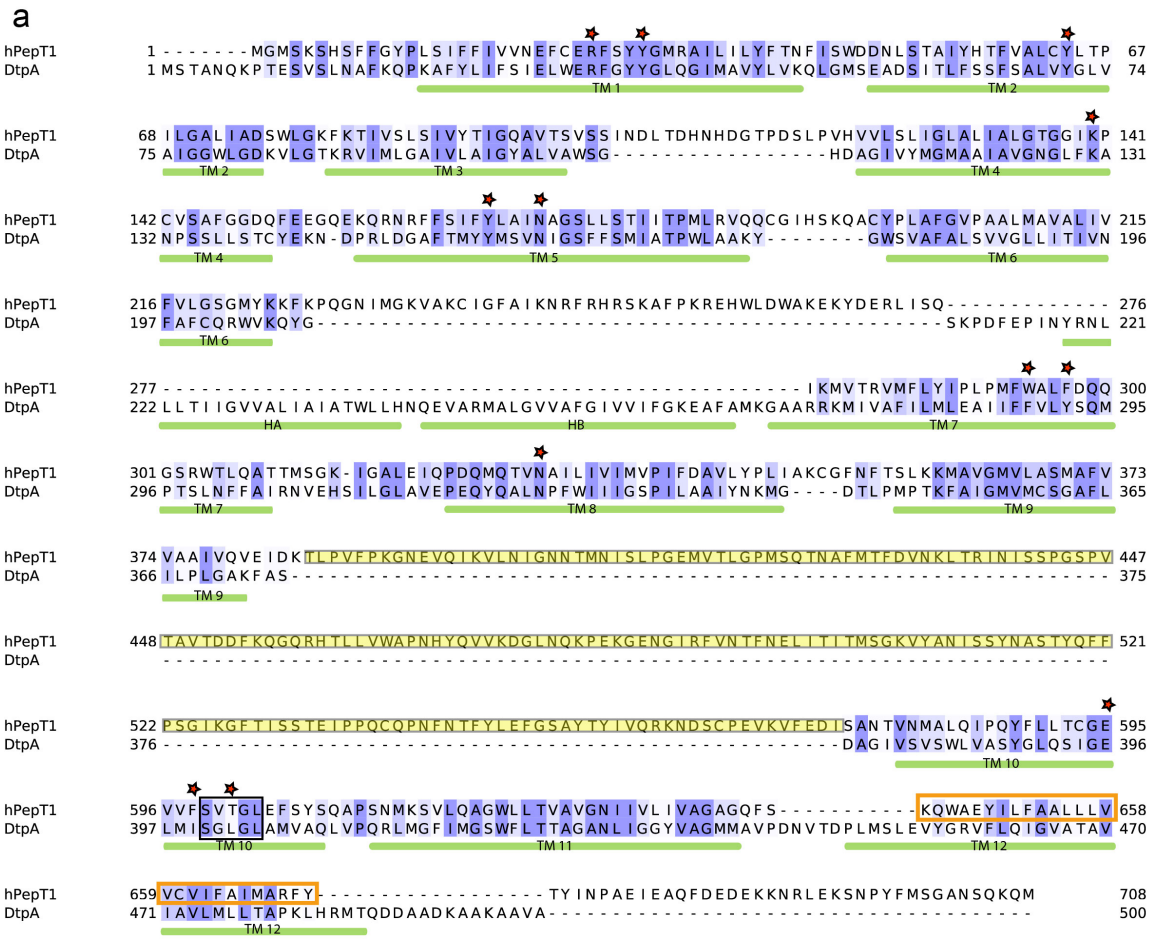


Figure 42: The binding affinity curves of DtpA mutants and di-/tripeptides
Binding affinity measurements of DtpA mutants with the ligands LeuLeuAla and AlaLeu.

3.8. Homology model of the human PepT1 based on DtpA

DtpA has been characterized for its similarity to hPepT1 in its ligand selectivity and transports drugs taken up by hPepT1 (Weitz *et al*. 2007; Prabhala *et al*. 2017). The binding affinity of DtpA to valganciclovir and valacyclovir that we have measured was also similar to the binding affinity recently measured for valacyclovir to hPepT1 (Clémençon, Lüscher, and Hediger 2018) with the same method supporting their conserved ligand binding profile. In the absence of a eukaryotic POT structure, a homology model of the hPepT1 was built by Vasileios Rantos from EMBL Hamburg in collaboration. The sequence alignment used in the homology modeling is shown in Figure 43.

a



b

DtpA	hPepT1
R34	R27
Y38	Y31
Y71	Y64
K130	K140
Y156	Y167
N160	N171
F289	W294
Y292	F297
N325	N329
E396	E595
I399	F598
L402	T601

Figure 43: Sequence alignment of DtpA and hPepT1
(a) The sequence similarity is shown in shades of blue with the darker the color tone representing the higher the similarity between DtpA and hPepT1. The transmembrane helices are marked with green bars underneath the sequence and the ligand-binding site residues with stars above the sequence. The extracellular domain of hPepT1, which is not present in DtpA and not included in the homology model, is highlighted in yellow. The TM12 of hPepT1 model has lower certainty in the model building and therefore shown in an orange box. The black box demonstrates the intrahelical loop within TM10 for both DtpA and hPepT1. (b) The conserved ligand-binding site residues of DtpA and hPepT1 are shown in a separate table.

There are two major structural differences between bacterial and human POTs. The HAHB domain is present only in the bacterial structure; in eukaryotic POTs there is a long linker connecting TM6 and TM7 instead. The extracellular domain is present

only in mammalian POTs and the structure thereof was determined recently (Beale *et al*. 2015), which revealed a globular shape consisting of two tightly packed immunoglobin-like folds. Nevertheless, there is no insight into how this extracellular domain is oriented relative to the transporter core. Therefore, the long linker between TM6-TM7 and the extracellular domain of hPepT1 were omitted and only the transporter core of hPepT1 was modeled (Figure 44).

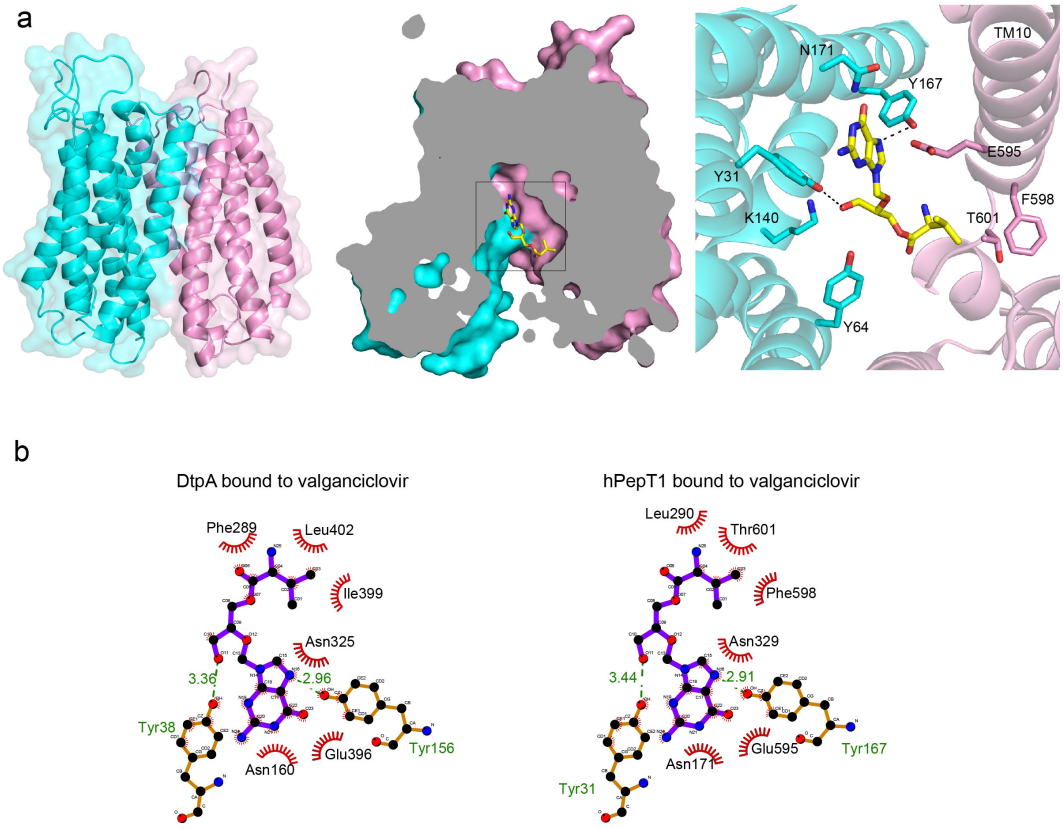


Figure 44: The ligand-binding site analysis of hPepT1 model with valganciclovir (a) Left: Surface and cartoon representation of hPepT1 with the N-terminal bundle colored in cyan and the C-terminal bundle in pink. The extracellular domain and the long loop connecting TM6 and TM7 are excluded from the model. Middle: The surface representation of hPepT1 cut in half to highlight the ligand-binding site in the center of the transporter. Valganciclovir is shown in sticks model (yellow). Right: Same model with the ligand-binding site residues in stick representation. The dashed lines highlight possible hydrogen bonds in less than 3.5 Å distance. The residues N329 and L290 are excluded for clarity and the residues K140 and Y64 are included due to their role in di-/tripeptide coordination as observed in other POT structures (Martinez Molledo *et al*. 2018), although distant from valganciclovir in this case. (b) LigPlot (Laskowski and Swindells 2011) analysis of the ligand-binding site of DtpA and hPepT1 side-by-side.

The ligand-binding site of DtpA and hPepT1 are highly conserved and the ligand-binding site of hPepT1 model resembles the DtpA ligand-binding site (Figure 37). Valganciclovir fits into the ligand-binding site of hPepT1. The side chains coordinating the guanine ring of valganciclovir in hPepT1, namely Y31, Y167, N171, N329 and E595, are the equivalent residues of DtpA, Y38, Y156, N160 and E396 (Figure 44). The hydroxyl group of Y31 is likely to form a hydrogen bond with the hydroxyl group at the center of valganciclovir for one of the diastereomers (Figure 8), because in the other diastereomer the hydroxyl group is pointing towards another direction and is not in proximity. The hydroxyl group of Y167 is likely to form a hydrogen bond with the amine group of the guanine ring. We do not see any interactions with the middle part of valganciclovir, except the aforementioned hydrogen bond. The N-terminal valine is intruding into a pocket and has hydrophobic interactions with L290, F598 and T601. Interestingly, these residues are not as conserved between hPepT1 and DtpA as the residues coordinating the guanine ring but play an equivalent role.

4. Discussion and Outlook

4.1. Homology model of hPepT1

The human PepT1 is highly expressed in the intestinal cells in humans and plays a key role in the nutrient uptake as well as in the absorption of several drugs, such as β-lactam antibiotics, antiviral drugs and anticancer molecules (Brandsch, Knütter, and Bosse-Doenecke 2008). To date, there are no eukaryotic POT structures, but bacterial homologues have proven to be good models. There are a few structures of peptide transporters with di-/tripeptides, their natural ligands, providing insight into the ligand-binding site and the structure-activity relationship. Unfortunately, these bacterial POTs are not able to bind or transport drugs. Therefore, in collaboration an hPepT1 model was built based on the DtpA structure in complex with valganciclovir.

To validate the hPepT1 model, I have examined the mutational studies performed with hPepT1 in the literature. Three transmembrane helices, namely TM5 (Kulkarni, Haworth, and Lee 2003), TM7 (Kulkarni *et al*. 2003) and TM10 (Xu *et al*. 2009), were systematically studied by mutating each residue to cysteine and measuring the uptake of the radioactively labeled [^{3}H]-glycyl-sarcosine (GlySar) in HEK293 cells overexpressing these mutants. Most of the residues, which led to less than 25 % activity, could be mapped in the model in proximity of the ligand-binding site (Figure 45). On TM5, these were Y167, N171 and S174, on TM7 F293 and F297 and on TM10 E595 and F598. Within the same study, the solvent accessibility was investigated and the residues T601-F605 in TM10 were all highly solvent accessible (Xu *et al*. 2009). In a standard continuous α-helix, adjacent residues are not expected to be solvent accessible; however in our model this region forms the intrahelical loop within TM10 (Figure 45), for which one would expect solvent accessibility, supporting the presence of the intrahelical TM10 loop in hPepT1 and our model in general.

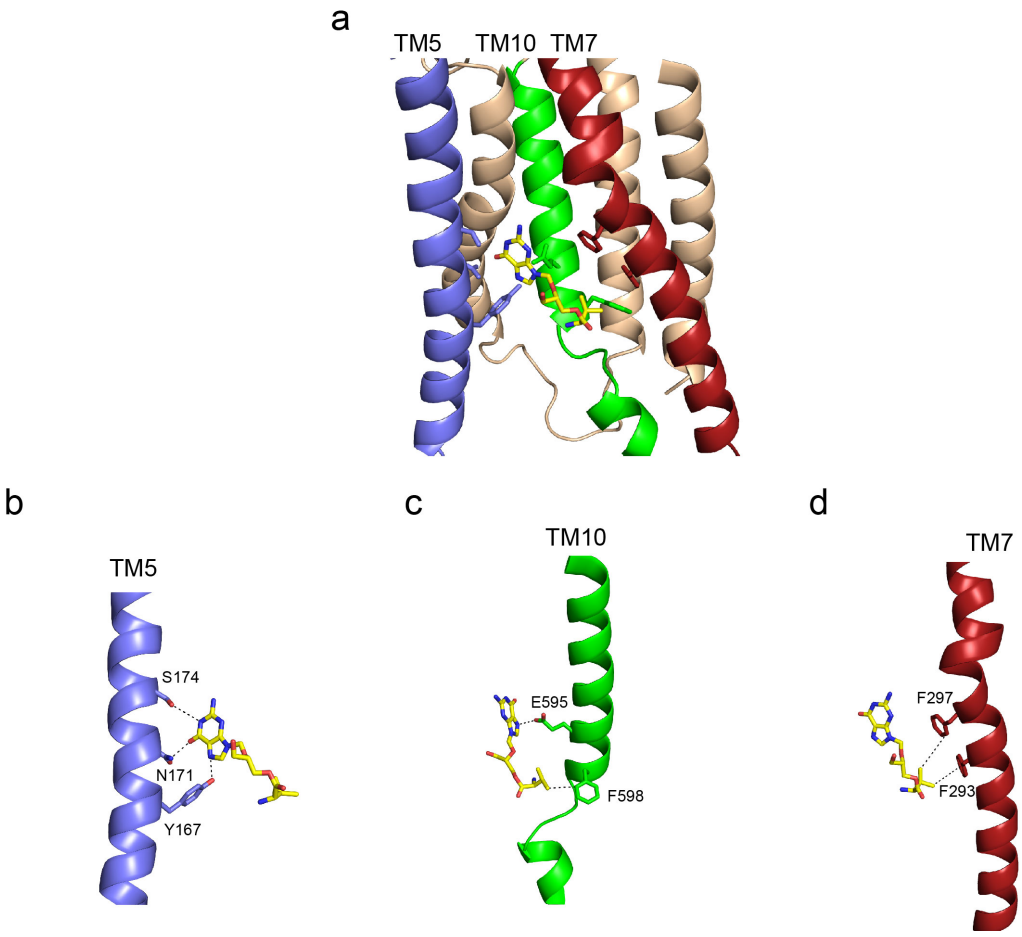


Figure 45: The mutational studies on hPepT1
(a) hPepT1 model in cartoon representation colored in wheat. The transmembrane helices (b) 5, (d) 7 and (c) 10 are colored in blue, red and green respectively. The residues, which were suggested to play a role in ligand uptake as well as valganciclovir (yellow), are shown in sticks model. The dashed lines represent distances up to 5.5 Å.

hPepT1 is one of the most abundant multidrug transporters in the intestines (Drozdzik *et al*. 2014) and of great pharmacological interest. Few homology models have been reported for hPepT1 based on previous bacterial POT structures, which were not capable of transporting drugs and did not have the intrahelical loop in TM10 (Samsudin *et al*. 2016; Beale *et al*. 2015), furthermore they were not uploaded on a server and accessible to the community. Here, we provide a new homology model (accessible in https://www.modelarchive.org upon publication) and expect that it would contribute to further studies on ligand selectivity and drug discovery.

4.2. The binding mode of valganciclovir and the binding pocket near TM10-TM11

The ligand-binding site of POTs is in the center of the protein, is globular and has several side pockets, where different side chains of di- and tripeptides and (pro)drugs can fit in. This organic shape makes the ligand-binding site rather difficult to study and predict which pockets might be involved in the coordination of the different side chains of di- and tripeptides. Furthermore, there are many aromatic residues present in the ligand-binding site. These allow versatile interaction modes contributing to the promiscuity of the transporter. It is even more challenging to predict the binding mode of the peptidomimetic drugs, which mimic the peptide backbone and the side chains. However, the selectivity of the ligand-binding site is defined by the space available for the ligands within these pockets.

The binding mode of valacyclovir in hPepT1 was predicted based on the $PepT_{St}$ structure bound to AlaPhe (Samsudin et al. 2016). Here, the valine moiety of valacyclovir was positioned as the N-terminal residue of the dipeptide AlaPhe and the guanine ring as the aromatic side chain of the C-terminal residue. I have initially tested the same binding mode in the electron density within the ligand-binding site of the DtpA-valganciclovir crystal structure, but this model was not compatible and thus alternative binding modes were tested. The binding mode, where the guanine ring is mimicking the N-terminal residue of a di-/tripeptide and the valine moiety extending into an uncharacterized pocket near TM10-TM11 was the best model based on the electron density map. Indeed, this binding pocket close to TM10-TM11 increases the available space within the ligand-binding site and was the major structural difference between DtpA and other POTs except for $PepT_{So2}$. Interestingly, DtpA and $PepT_{So2}$ have another feature in common, namely both prefer tripeptides to dipeptides, whereas other tested POTs have a preference towards dipeptides. Valganciclovir has a similar backbone length to a tripeptide and we predict that the newly characterized pocket near TM10-TM11 is important not only for the coordination of valganciclovir but also for the coordination of tripeptides.

4.3. Ligand selectivity of DtpA

Ligand selectivity of DtpA was characterized by *in vitro* and *in vivo* assays, using a ligand library consisting of eighteen dipeptides, ten tripeptides and seven drugs. This is the most comprehensive study of DtpA so far and the first one to test ligand binding and binding affinities *in vitro*. The *in vivo* and *in vitro* studies were in good agreement; the former studies were performed in the native environment, the *E. coli* cell membrane, and the latter with the detergent-solubilized protein in solution. The results can be summarized in three points:

First, the tripeptides are preferred over dipeptides. Due to the difficulties of studying the ligand selectivity, there is information only for a handful of POTs. $PepT_{St}$ and $PepT_{So}$ prefer dipeptides over tripeptides in competition assays, and many tripeptides did not compete with the used reporter molecule at all (Martinez Molledo *et al*. 2018; Parker *et al*. 2017). $PepT_{So2}$ is the only other POT similar to DtpA, which has a preference to tripeptides over dipeptides (Guettou *et al*. 2014). To our knowledge, there are no studies on eukaryotic POTs addressing this point so far.

The second observation was that the di-/tripeptides with hydrophobic and aromatic side chains, independent of their position, are preferred over charged residues. This has been indeed observed more than 25 years ago for hPepT1 (Daniel, Morse, and Adibi 1992) and recently for other POTs (Guettou et al. 2014; Parker et al. 2017; Martinez Molledo et al. 2018) and is likely a general feature of POTs. From an evolutionary standpoint, such a preference towards hydrophobic and aromatic residues can be explained by the higher metabolic cost of producing these amino acids *de novo* for bacteria, as well as the essential nature of the amino acids in higher organisms. Nevertheless, it is possible that there are POTs with a different preference, especially in single-cell organisms with multiple POT homologues as suggested for DtpC (Jensen *et al*. 2012). We are still in the early steps of understanding the ligand selectivity of POTs and more systematic studies and new methods are required to characterize it in depth.

The third observation in our studies was that the ligands with glycine residues are less favored. For example, the thermal stabilization effect of MetSer and HisSer was higher than GlySer; similarly, the effect of LeuLeuAla was significantly higher than that of LeuGlyGly (Figure 23). No binding affinity could be determined with a ligand

containing a glycine residue. This is an important point to make, because several *in vivo* competition assays were performed with glycine containing di-/tripeptides, where the effect of the other residue was investigated (Weitz *et al*. 2007). Nevertheless, both GlyGly and GlyGlyGly were shown to be taken up into the cell indirectly in a previous study (Weitz *et al*. 2007). More ligands need to be tested containing glycine residues in a systematic approach to characterize the effect of glycine. Shortly, the ligand selectivity of DtpA was further characterized and the first binding affinity values were determined in low to medium μM range. For ligands with small to no effect on the thermal stabilization of DtpA, the binding affinity could not be determined. The binding affinity of hPepT1 to valacyclovir was recently determined using the same method (MST) as 49 ± 0.3 μM, which is very similar to the binding affinity of DtpA to valacyclovir and valganciclovir with 60 ± 13 μM and 76 ± 16 μM, as shown in this study. These results indicate further similarity between DtpA to hPepT1 and support DtpA as a good bacterial model of the human transporter.

4.4. Limitations in studying the ligand selectivity of POTs

All methods used within this study show binding of the ligand to DtpA, however we are interested in the transport of these ligands. It is a common problem in the POT field and most studies are based on competition assays with a reporter molecule, AK-AMCA (Dieck et al. 1999; Groneberg et al. 2001; Weitz et al. 2007; Prabhala et al. 2017) or a radioactively labeled ligand: [^{3}H]-GlySar (Yeung et al. 1998; Knütter et al. 2001; Newstead et al. 2011) or [^{3}H]-AlaAla (Solcan et al. 2012; Doki et al. 2013; Guettou et al. 2014; Samsudin et al. 2016). The direct transport of the reporter molecule is shown for the POT but the competing ligands can be either really transported or only binding to the transporter without being moved across the membrane. Also the ligands, where no competition is observed, might be nevertheless transported but at a much lower rate than the reporter molecule.

Unfortunately, there are very few methods to study the direct transport of ligands, which are difficult to establish, require specific facilities and not compatible with high throughput screening of many ligands. The reconstitution of the transporter into liposomes with a pH difference inside and out of the liposome to build a proton

gradient is the classical approach but it is necessary to use radioactively labeled ligands. Often only one radioactively labeled ligand is tested in a direct approach and few other ligands are tested in competition. Two major drawbacks from applying this method in a systematic way are the limited access to facilities, which allow working with radioactivity, and the difficulty to acquire a ligand library consisting of di-/tripeptides with different chemical features that are radioactively labeled.

Another approach to measure the direct uptake of the ligands was used for DtpA with mass spectroscopy. Here, the *E. coli* cells were incubated with different drugs and the cell content was then analyzed for the occurrence of the drugs by liquid chromatography-mass spectrometry (LC-MS) (Prabhala et al. 2017). The advantage of this method is that it is *in vivo* and without labeling but it is very challenging to study di-/tripeptides, which occur naturally in the cell, with this method and requires access to a specialized mass spectrometry facility. In summary, although we have contributed to the understanding of the ligand selectivity in DtpA with ligand binding and competition studies, further work is required to understand the transport profile of the ligands.

4.5. Mutational studies on DtpA

Five mutations were introduced to DtpA within this study to test the ligand binding and transport activity. The DtpA mutants Y38A, Y71A, K130A, N160A and I399A lost the AK-AMCA uptake function and had significantly lower binding affinities against the selected ligands (Results section 3.7). This shows that the mutations are critical not only for ligand binding but also for transport.

In a previous study, three of these residues were mutated to different amino acids, namely Y38F, Y71F and K130R (Jensen *et al*. 2012), where the AK-AMCA uptake was strongly reduced. The same study also investigated Y292F/Y292A, S400D/S400N/S400C (Jensen *et al*. 2012). Both residues are part of the binding pocket formed by the TM10 intrahelical loop (Figure 46). In case of Y292F and Y292A, the AK-AMCA uptake increased by 1.5- and 4-fold respectively. On the DtpA structure, this can be explained by an increase of the binding pocket volume upon the mutations. S400 was selected for the mutational study because it is highly

conserved between POTs, except for DtpC and DtpD, which have an aspartate at this position. All mutations of S400 (S400D, S400N and S400C) inhibited the AK-AMCA uptake. Although the side chain of S400 is not pointing towards valganciclovir in the DtpA structure, the backbone carboxyl group is in proximity of the ligand. We cannot explain the effect of the mutation on the uptake activity in the present inward-open conformation of DtpA, but in the occluded or outward-open conformation the side chain may constrict the binding pocket size.

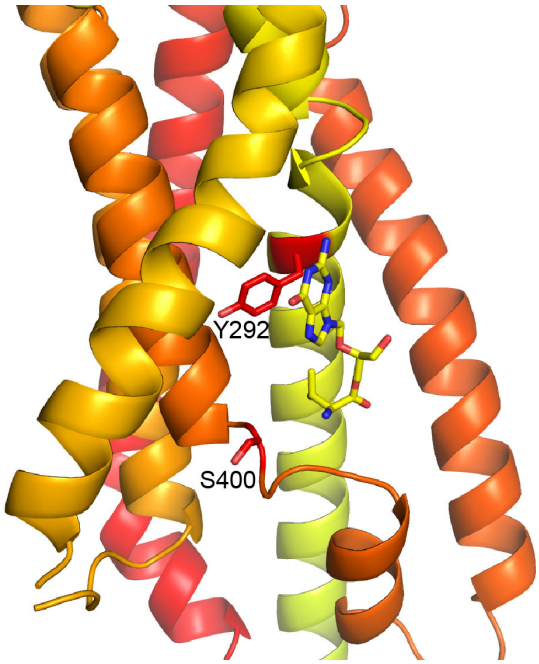


Figure 46: Position of Y292 and S400 in DtpA structure
DtpA is colored form blue (N-terminal end) to red (C-terminal end) and the N-terminal bundle is omitted from the picture for a better view of the ligand-binding site. Valganciclovir is shown in yellow sticks model for illustrating the ligand-binding site.Y292 and S400, which were mutated (Jensen *et al*. 2012), are illustrated on the DtpA structure as sticks model.

DtpA was also subjected to a random mutagenesis study, in which 35 point mutations led to reduced uptake of AK-AMCA and/or alafosfalin (Malle *et al*. 2011) (Figure 47). Most of the mutated residues were likely involved in intramolecular interactions between the transmembrane α-helices. There was only one mutated position within the ligand-binding site, namely F289, which was mutated to leucine and serine, both of which inhibited the AK-AMCA uptake. The most interesting finding of this study was the characterization of the periplasmic gate salt bridge. The single mutations of E56 and R305 inhibited the AK-AMCA uptake, whereas the double-mutation rescued

the uptake activity. In the DtpA structure, E56 and R305 form a hydrogen bond (Figure 47). This periplasmic salt bridge is highly conserved among bacterial POTs, although in the case of $PepT_{So}$, $PepT_{St}$ and YePepT the positions of glutamate and arginine are swapped.

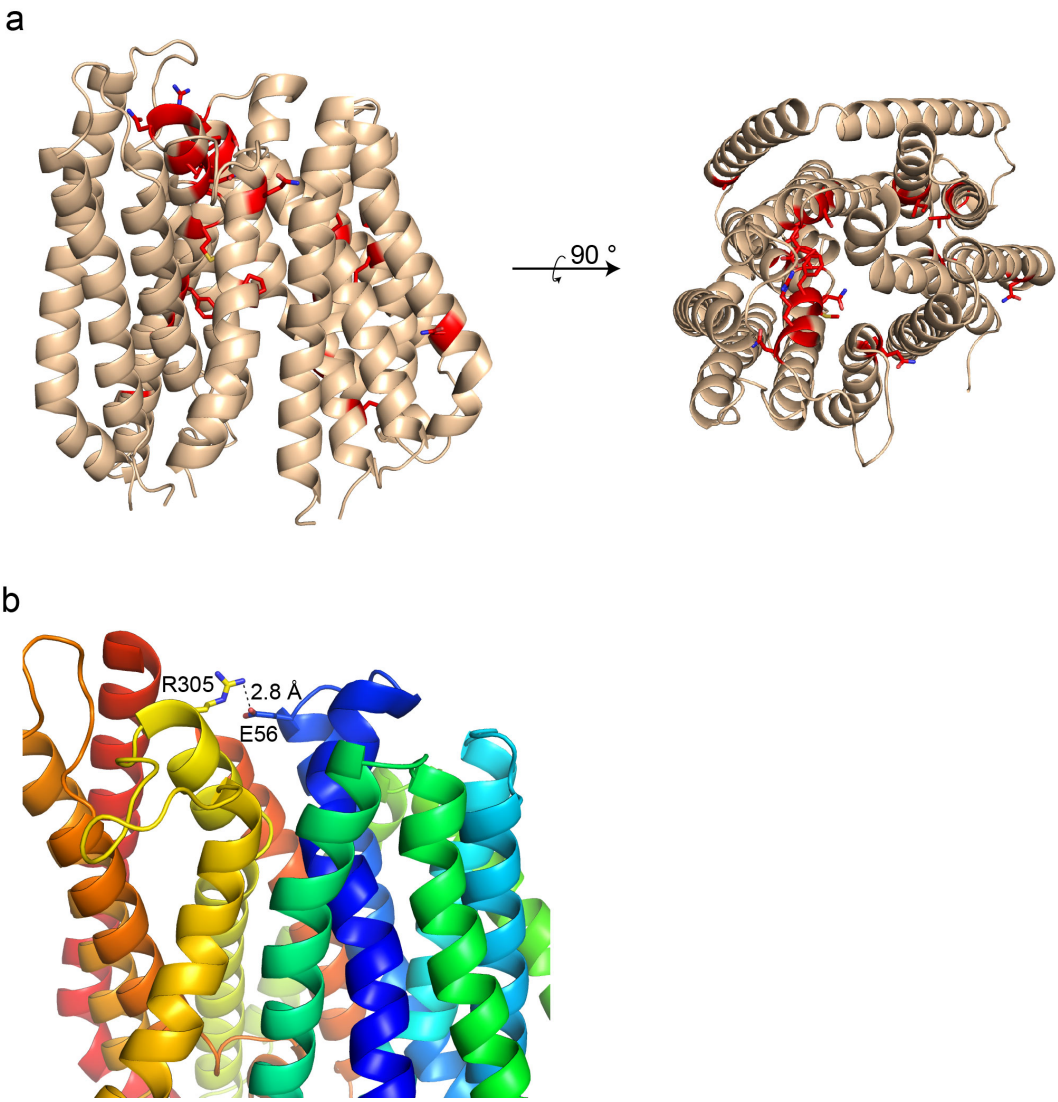


Figure 47: Random mutagenesis results mapped on the DtpA structure.
(a) DtpA in wheat and all mutations (Malle *et al*. 2011), which led to reduced AK-AMCA and/or alafosfalin uptake are colored in red. (b) DtpA is colored from blue (N-terminal end) to red (C-terminal end) and the residues 429-449 (periplasmic loop) are omitted from the structure for clarity. The E56 and R305 are shown in sticks model and the distance between E56 and R305 is shown as a black dashed line.

4.6. The conformational state of DtpA

To support our understanding of ligand binding and transport, the structure of DtpA was determined in both ligand-free and valganciclovir-bound form using a nanobody as crystallization chaperone, which stabilized the inward-open conformation of DtpA. In fact, almost all POT structures are found in the inward-open conformation. Therefore, it is likely that the POTs, at least in the detergent solubilized form, are stabilized in this conformation. Nevertheless, it is challenging to investigate the conformation of a transporter in solution.

An interesting outcome of the *in vitro* assays was that the behavior of DtpA and DtpA-N00 was very similar. The binding affinity measurements are sensitive towards conformational changes, because the method is based on the differences in size, charge and change in hydration shell upon ligand binding; similar changes occur also upon conformational change. For the DtpA-N00 complex, we show that DtpA is in an inward-open conformation. If DtpA without N00 is yielding the same binding affinities towards the same ligands, it is very likely that DtpA is (almost) completely in the inward-open conformation in solution as well. It remains an open question, if DtpA can change its conformation in detergent-solubilized form. Unfortunately, this point was not possible to address within this work and requires further biochemical techniques, such as Förster resonance energy transfer (FRET) measurements.

Noticeably, the ligand binding was studied with the readily formed DtpA-N00 complex, which is locked in the inward-open conformation, namely the release state. For the previous ligand-bound structures of POTs one could assume that the ligand did bind to the outward-open conformation in solution and the transition to inward-open state took place and the crystallization was with the inward open ligand-bound state of the POTs. In case of DtpA, the ligand binds directly to DtpA in inward-open conformation, so it is possible that also the other POTs were in the inward-open conformation in the first place. Why the ligands bind to the transporter in the release state and how they are released is still unclear. Further studies on the transport cycle *in vivo* and *in vitro* will clarify this puzzling observation.

4.7. The role of N00 in the crystallization of DtpA

The first low-resolution structural analysis of a POT was published in 2007 and it was a negative stain transmission electron microscopy (TEM) image of the detergent-solubilized and purified DtpA (Weitz *et al*. 2007). The first high-resolution structure of a POT in 2011 was of a bacterial homologue from *Shewanella oneidensis* (Newstead *et al*. 2011). Considering the early work on DtpA and its strong similarity to human PepT1 in its ligand selectivity, there is a keen interest in the field to determine the high-resolution structure of DtpA. In our hands, no initial crystals were obtained for DtpA using vapor diffusion technique or the lipidic cubic phase method. Only after the addition of nanobodies, initial crystallization of DtpA was observed and could be further optimized leading to the structures of DtpA-N00 complex (Results section 3.5). The nanobody played a crucial part in the crystallization of DtpA by stabilizing DtpA and by forming crystal contact points observed in the determined structure as shown in Figure 34. N00 builds interactions with four DtpA molecules and there are no direct crystal contacts between two DtpA molecules.

The binding of N00 was studied *in vivo* within the natural environment of DtpA by two independent methods that confirmed the occurrence of the DtpA-N00 complex in the native membrane. First, DtpA and N00 were co-expressed in *E. coli* cells and the uptake assay showed a strongly reduced uptake of AK-AMCA in co-expressing cells relative to cells overexpressing DtpA only. The decrease of AK-AMCA uptake (Figure 17) can be explained by the majority of DtpA molecules within the inner membrane of *E. coli* being blocked due to the binding of the N00 from the periplasmic site of the protein. The second approach was *in vivo* crosslinking of N00 to DtpA, which carried a UV-inducible mutation close to the DtpA-N00 binding site. Both results confirm the DtpA-N00 complex formation within the cell membrane and that the structure of DtpA is a biologically relevant one.

4.8. Outlook

DtpA is one of the first characterized bacterial POTs with striking similarity to the human homologue hPept1 (Weitz *et al*. 2007). Within the scope of this work, the structure of DtpA was determined in ligand-free and valganciclovir-bound form and the ligand preference was investigated with *in vitro* and *in vivo* methods. The binding affinity of DtpA was measured for the first time and is in low to medium μM range. A preference for tripeptides over dipeptides was observed in DtpA, which is different than most other studied POTs and may be explained by an intrahelical loop within TM10 and a shift of TM11 that is not observed in the other POT structures. Although a few POT structures with di- and tripeptides are available showing a conserved binding mode of the peptide backbone and side chains intruding into smaller pockets, this is the only structure of a POT with a drug. A homology model of hPepT1 based on this structure was built and will be available in https://www.modelarchive.org upon publication. Drug design efforts have been so far based on the previous POT structures, which cannot transport drugs. We hope that the hPepT1 model based on the DtpA-valganciclovir structure will provide a better platform for further efforts in drug discovery and development until the structure of hPepT1 or other mammalian homologues with similar ligand selectivity are determined.

POTs have been crystallized in partially occluded or fully inward open conformations, including the presented DtpA-N00 complex structures in this work. Although this information is valuable, and the residues of the ligand-binding site determined from the structure is supported by biochemical analyses, the inward open conformation is the release state of the transporter. To improve our understanding of the transport process and to aid in drug design studies, it is crucial to obtain an outward open conformation. To obtain the outward open conformation, we need a better understanding of the conformational changes that take place *in vivo* and *in vitro*. There is a multitude of open questions related to the conformational changes in POTs, such as the trigger for the transition from one conformational state to another, the turnover rates of POTs or even if the outward-open state can be stabilized in the detergent-solubilized form.

With the available structures of DtpA and DtpD (Zhao *et al*. 2014) from *E. coli*, other aspects of POTs can be studied, such as gene regulation and protein expression patterns. *E. coli* is an ideal bacterial background to study this topic as most methods are developed for this organism. *E. coli* has four POTs, named DtpA to -D that genetically cluster into two groups: DtpA and –B in one group and DtpC and –D in the other group. The question why unicellular organisms need multiple POTs in addition to further peptide transport systems such as the Opp and Dpp system, has not been addressed to our knowledge. In *E. coli*, one could study the ligand selectivity of DtpA to -D as well as their expression patterns related to external and internal cues supported by structural information giving us a broader picture of peptide transport in prokaryotic cells.

5. References

Adams, Paul D, Pavel V Afonine, Gábor Bunkóczi, Vincent B Chen, Ian W Davis, Nathaniel Echols, Jeffrey J Headd, et al. 2010. "PHENIX: a Comprehensive Python-Based System for Macromolecular Structure Solution.." *Acta Crystallographica Section D: Biological Crystallography* 66 (Pt 2). International Union of Crystallography: 213–21. doi:10.1107/S0907444909052925.

Agre, Peter. 2004. "Aquaporin Water Channels (Nobel Lecture).." *Angewandte Chemie (International Ed. in English)* 43 (33). Wiley-Blackwell: 4278–90. doi:10.1002/anie.200460804.

Aherne, Margaret, Joseph A Lyons, and Martin Caffrey. 2012. "A Fast, Simple and Robust Protocol for Growing Crystals in the Lipidic Cubic Phase.." *Journal of Applied Crystallography* 45 (Pt 6). International Union of Crystallography: 1330–33. doi:10.1107/S0021889812037880.

Beale, John H, Joanne L Parker, Firdaus Samsudin, Anne L Barrett, Anish Senan, Louise E Bird, David Scott, et al. 2015. "Crystal Structures of the Extracellular Domain From PepT1 and PepT2 Provide Novel Insights Into Mammalian Peptide Transport." *Structure* 23 (10): 1889–99. doi:10.1016/j.str.2015.07.016.

Bippes, Christian A, Lin Ge, Marcel Meury, Daniel Harder, Zöhre Ucurum, Hannelore Daniel, Dimitrios Fotiadis, and Daniel J Müller. 2013. "Peptide Transporter DtpA Has Two Alternate Conformations, One of Which Is Promoted by Inhibitor Binding.." *Proceedings of the National Academy of Sciences of the United States of America* 110 (42). National Acad Sciences: E3978–86. doi:10.1073/pnas.1312959110.

Boggavarapu, Rajendra, Jean-Marc Jeckelmann, Daniel Harder, Zöhre Ucurum, and Dimitrios Fotiadis. 2015. "Role of Electrostatic Interactions for Ligand Recognition and Specificity of Peptide Transporters.." *BMC Biology* 13 (1). BioMed Central Ltd: 58. doi:10.1186/s12915-015-0167-8.

Boivin, Stephane, Sandra Kozak, and Rob Meijers. 2013. "Optimization of Protein Purification and Characterization Using Thermofluor Screens.." *Protein Expression and Purification* 91 (2): 192–206. doi:10.1016/j.pep.2013.08.002.

Bowler, Matthew W, Didier Nurizzo, Ray Barrett, Antonia Beteva, Marjolaine Bodin, Hugo Caserotto, Solange Delagenière, et al. 2015. "MASSIF-1: a Beamline Dedicated to the Fully Automatic Characterization and Data Collection From Crystals of Biological Macromolecules.." *Journal of Synchrotron Radiation* 22 (Pt 6): 1540–47. doi:10.1107/S1600577515016604.

Brandsch, Matthias, Ilka Knütter, and Eva Bosse-Doenecke. 2008. "Pharmaceutical and Pharmacological Importance of Peptide Transporters." *Journal of Pharmacy and Pharmacology* 60 (5). Blackwell Publishing Ltd: 543–85. doi:10.1211/jpp.60.5.0002.

Chen, Vincent B, W Bryan Arendall, Jeffrey J Headd, Daniel A Keedy, Robert M Immormino, Gary J Kapral, Laura W Murray, Jane S Richardson, and David C Richardson. 2010. "MolProbity: All-Atom Structure Validation for Macromolecular Crystallography.." *Acta Crystallographica Section D: Biological Crystallography* 66 (Pt 1). International Union of Crystallography: 12–21. doi:10.1107/S0907444909042073.

Cheng, Y, and W H Prusoff. 1973. "Relationship Between the Inhibition Constant (K1) and the Concentration of Inhibitor Which Causes 50 Per Cent Inhibition

(I50) of an Enzymatic Reaction..” *Biochemical Pharmacology* 22 (23): 3099–3108.

Clémençon, Benjamin, Benjamin P Lüscher, and Matthias A Hediger. 2018. “Establishment of a Novel Microscale Thermophoresis Ligand-Binding Assay for Characterization of SLC Solute Carriers Using Oligopeptide Transporter PepT1 (SLC15 Family) as a Model System..” *Journal of Pharmacological and Toxicological Methods* 92 (March): 67–76. doi:10.1016/j.vascn.2018.03.004.

Daniel, H. 2006. “From Bacteria to Man: Archaic Proton-Dependent Peptide Transporters at Work.” *Physiology* 21 (2): 93–102. doi:10.1152/physiol.00054.2005.

Daniel, H, E L Morse, and S A Adibi. 1992. “Determinants of Substrate Affinity for the Oligopeptide/H+ Symporter in the Renal Brush Border Membrane..” *The Journal of Biological Chemistry* 267 (14). American Society for Biochemistry and Molecular Biology: 9565–73.

Davis, Ian W, Andrew Leaver-Fay, Vincent B Chen, Jeremy N Block, Gary J Kapral, Xueyi Wang, Laura W Murray, et al. 2007. “MolProbity: All-Atom Contacts and Structure Validation for Proteins and Nucleic Acids..” *Nucleic Acids Research* 35 (Web Server issue): W375–83. doi:10.1093/nar/gkm216.

De Clercq, Erik, and Hugh J Field. 2006. “Antiviral Prodrugs - the Development of Successful Prodrug Strategies for Antiviral Chemotherapy..” *British Journal of Pharmacology* 147 (1). Wiley/Blackwell (10.1111): 1–11. doi:10.1038/sj.bjp.0706446.

Delano, W L. 2002. “The PyMOL Molecular Graphics System.” *(2002)*. doi:10.1234/12345678.

Deng, Dong, Pengcheng Sun, Chuangye Yan, Meng Ke, Xin Jiang, Lei Xiong, Wenlin Ren, et al. 2015. “Molecular Basis of Ligand Recognition and Transport by Glucose Transporters..” *Nature* 526 (7573). Nature Publishing Group: 391–96. doi:10.1038/nature14655.

Dieck, Susanne Tom, Heike Heuer, Jan Ehrchen, Christiane Otto, and Karl Bauer. 1999. “The Peptide Transporter PepT2 Is Expressed in Rat Brain and Mediates the Accumulation of the Fluorescent Dipeptide Derivative B-Ala-Lys-Nε-AMCA in Astrocytes.” *Glia* 25 (1). John Wiley & Sons, Inc.: 10–20. doi:10.1002/(SICI)1098-1136(19990101)25:1<10::AID-GLIA2>3.0.CO;2-Y.

Doki, Shintaro, Hideaki E Kato, Nicolae Solcan, Masayo Iwaki, Michio Koyama, Motoyuki Hattori, Norihiko Iwase, et al. 2013. “Structural Basis for Dynamic Mechanism of Proton-Coupled Symport by the Peptide Transporter POT..” *Proceedings of the National Academy of Sciences of the United States of America* 110 (28). National Acad Sciences: 11343–48. doi:10.1073/pnas.1301079110.

Drozdzik, Marek, Christian Gröer, Jette Penski, Joanna Lapczuk, Marek Ostrowski, Yurong Lai, Bhagwat Prasad, Jashvant D Unadkat, Werner Siegmund, and Stefan Oswald. 2014. “Protein Abundance of Clinically Relevant Multidrug Transporters Along the Entire Length of the Human Intestine..” *Molecular Pharmaceutics* 11 (10): 3547–55. doi:10.1021/mp500330y.

Dufourc, Erick J. 2008. “Sterols and Membrane Dynamics..” *Journal of Chemical Biology* 1 (1-4). Springer-Verlag: 63–77. doi:10.1007/s12154-008-0010-6.

Durrant, Jacob D, César Augusto F de Oliveira, and J Andrew McCammon. 2011. “POVME: an Algorithm for Measuring Binding-Pocket Volumes..” *Journal of Molecular Graphics & Modelling* 29 (5): 773–76. doi:10.1016/j.jmgm.2010.10.007.

Durrant, Jacob D, Lane Votapka, Jesper Sørensen, and Rommie E Amaro. 2014.

"POVME 2.0: an Enhanced Tool for Determining Pocket Shape and Volume Characteristics.." *Journal of Chemical Theory and Computation* 10 (11): 5047–56. doi:10.1021/ct500381c.

Emsley, Paul, and Kevin Cowtan. 2004. "Coot: Model-Building Tools for Molecular Graphics.." *Acta Crystallographica Section D: Biological Crystallography* 60 (Pt 12 Pt 1). International Union of Crystallography: 2126–32. doi:10.1107/S0907444904019158.

Fan-Havard, P, M C Nahata, and M T Brady. 1989. "Ganciclovir--a Review of Pharmacology, Therapeutic Efficacy and Potential Use for Treatment of Congenital Cytomegalovirus Infections.." *Journal of Clinical Pharmacy and Therapeutics* 14 (5): 329–40.

Farrell, Ian S, Rebecca Toroney, Jennifer L Hazen, Ryan A Mehl, and Jason W Chin. 2005. "Photo-Cross-Linking Interacting Proteins with a Genetically Encoded Benzophenone.." *Nature Methods* 2 (5). Nature Publishing Group: 377–84. doi:10.1038/nmeth0505-377.

Flayhan, Ali, Haydyn D T Mertens, Yonca Ural-Blimke, Maria Martinez Molledo, Dmitri I Svergun, and Christian Löw. 2018. "Saposin Lipid Nanoparticles: a Highly Versatile and Modular Tool for Membrane Protein Research.." *Structure* 26 (2): 345–45. doi:10.1016/j.str.2018.01.007.

Goh, Ee-Been, Dominic F Siino, and Michele M Igo. 2004. "The Escherichia Coli tppB (ydgR) Gene Represents a New Class of OmpR-Regulated Genes.." *Journal of Bacteriology* 186 (12). American Society for Microbiology: 4019–24. doi:10.1128/JB.186.12.4019-4024.2004.

Gouaux, Eric. 1998. "It's Not Just a Phase: Crystallization and X-Ray Structure Determination of Bacteriorhodopsin in Lipidic Cubic Phases." *Structure* 6 (1). Cell Press: 5–10. doi:10.1016/S0969-2126(98)00002-1.

Groneberg, D A, F Döring, P R Eynott, A Fischer, and H Daniel. 2001. "Intestinal Peptide Transport: Ex Vivo Uptake Studies and Localization of Peptide Carrier PEPT1.." *American Journal of Physiology. Gastrointestinal and Liver Physiology* 281 (3): G697–704.

Guettou, Fatma, Esben M Quistgaard, Lionel Trésaugues, Per Moberg, Caroline Jegerschöld, Lin Zhu, Agnes Jin Oi Jong, Pär Nordlund, and Christian Löw. 2013. "Structural Insights Into Substrate Recognition in Proton-Dependent Oligopeptide Transporters.." *EMBO Reports* 14 (9): 804–10. doi:10.1038/embor.2013.107.

Guettou, Fatma, Esben M Quistgaard, Michael Raba, Per Moberg, Christian Löw, and Pär Nordlund. 2014. "Selectivity Mechanism of a Bacterial Homolog of the Human Drug-Peptide Transporters PepT1 and PepT2." *Nature Structural & Molecular Biology* 21 (8): 728–31. doi:10.1038/nsmb.2860.

Guo, A, P Hu, P V Balimane, F H Leibach, and P J Sinko. 1999. "Interactions of a Nonpeptidic Drug, Valacyclovir, with the Human Intestinal Peptide Transporter (hPEPT1) Expressed in a Mammalian Cell Line.." *The Journal of Pharmacology and Experimental Therapeutics* 289 (1): 448–54.

Hannich, J Thomas, Kyohei Umebayashi, and Howard Riezman. 2011. "Distribution and Functions of Sterols and Sphingolipids.." *Cold Spring Harbor Perspectives in Biology* 3 (5). Cold Spring Harbor Lab: a004762. doi:10.1101/cshperspect.a004762.

Harayama, Takeshi, and Howard Riezman. 2018. "Understanding the Diversity of Membrane Lipid Composition.." *Nature Reviews. Molecular Cell Biology* 19 (5). Nature Publishing Group: 281–96. doi:10.1038/nrm.2017.138.

Harder, Daniel, Jürgen Stolz, Fabio Casagrande, Petr Obrdlik, Dietmar Weitz, Dimitrios Fotiadis, and Hannelore Daniel. 2008. “DtpB (YhiP) and DtpA (TppB, YdgR) Are Prototypical Proton-Dependent Peptide Transporters of Escherichia Coli..” *The FEBS Journal* 275 (13). Blackwell Publishing Ltd: 3290–98. doi:10.1111/j.1742-4658.2008.06477.x.

Jardetzky, Oleg. 1966. “Simple Allosteric Model for Membrane Pumps.” *Nature* 211 (5052): 969–70. doi:10.1038/211969a0.

Jensen, Johanne Mørch, Fouzia Ismat, Gerda Szakonyi, Moazur Rahman, and Osman Mirza. 2012. “Probing the Putative Active Site of YjdL: an Unusual Proton-Coupled Oligopeptide Transporter From E. Coli..” Edited by Andrea Motta. *PloS One* 7 (10). Public Library of Science: e47780. doi:10.1371/journal.pone.0047780.

Jiang, Daohua, Yan Zhao, Xianping Wang, Junping Fan, Jie Heng, Xuehui Liu, Wei Feng, et al. 2013. “Structure of the YajR Transporter Suggests a Transport Mechanism Based on the Conserved Motif a..” *Proceedings of the National Academy of Sciences of the United States of America* 110 (36). National Acad Sciences: 14664–69. doi:10.1073/pnas.1308127110.

Joosten, Robbie P, Fei Long, Garib N Murshudov, and Anastassis Perrakis. 2014. “The PDB_REDO Server for Macromolecular Structure Model Optimization..” *IUCrJ* 1 (Pt 4). International Union of Crystallography: 213–20. doi:10.1107/S2052252514009324.

Jung, Donald, and Albert Dorr. 1999. “Single-Dose Pharmacokinetics of Valganciclovir in HIV- and CMV-Seropositive Subjects.” *The Journal of Clinical Pharmacology* 39 (8). Wiley-Blackwell: 800–804. doi:10.1177/00912709922008452.

Kabsch, Wolfgang. 2010. “Xds..” *Acta Crystallographica Section D: Biological Crystallography* 66 (Pt 2). International Union of Crystallography: 125–32. doi:10.1107/S0907444909047337.

Knütter, I, S Theis, B Hartrodt, I Born, M Brandsch, H Daniel, and K Neubert. 2001. “A Novel Inhibitor of the Mammalian Peptide Transporter PEPT1..” *Biochemistry* 40 (14): 4454–58. doi:10.1021/bi0026371.

Krogh, A, B Larsson, G von Heijne, and E L Sonnhammer. 2001. “Predicting Transmembrane Protein Topology with a Hidden Markov Model: Application to Complete Genomes..” *Journal of Molecular Biology* 305 (3): 567–80. doi:10.1006/jmbi.2000.4315.

Kulkarni, Ashutosh A, Ian S Haworth, and Vincent H L Lee. 2003. “Transmembrane Segment 5 of the Dipeptide Transporter hPepT1 Forms a Part of the Substrate Translocation Pathway..” *Biochemical and Biophysical Research Communications* 306 (1): 177–85.

Kulkarni, Ashutosh A, Ian S Haworth, Tomomi Uchiyama, and Vincent H L Lee. 2003. “Analysis of Transmembrane Segment 7 of the Dipeptide Transporter hPepT1 by Cysteine-Scanning Mutagenesis.” *The Journal of Biological Chemistry* 278 (51). American Society for Biochemistry and Molecular Biology: 51833–40. doi:10.1074/jbc.M308356200.

Laskowski, Roman A, and Mark B Swindells. 2011. “LigPlot+: Multiple Ligand-Protein Interaction Diagrams for Drug Discovery..” *Journal of Chemical Information and Modeling* 51 (10): 2778–86. doi:10.1021/ci200227u.

Lyons, J A, J L Parker, N Solcan, A Brinth, D Li, S T Shah, M Caffrey, and S Newstead. 2014. “Structural Basis for Polyspecificity in the POT Family of Proton-Coupled Oligopeptide Transporters.” *EMBO Reports*, June.

doi:10.15252/embr.201338403.
Malle, Elisabeth, Hongwen Zhou, Jana Neuhold, Bettina Spitzenberger, Freya Klepsch, Thomas Pollak, Oliver Bergner, Gerhard F Ecker, and Peggy C Stolt-Bergner. 2011. "Random Mutagenesis of the Prokaryotic Peptide Transporter YdgR Identifies Potential Periplasmic Gating Residues.." *The Journal of Biological Chemistry* 286 (26): 23121–31. doi:10.1074/jbc.M111.239657.
Martinez Molledo, Maria, Esben M Quistgaard, Ali Flayhan, Joanna Pieprzyk, and Christian Löw. 2018. "Multispecific Substrate Recognition in a Proton-Dependent Oligopeptide Transporter.." *Structure*, February. doi:10.1016/j.str.2018.01.005.
Martinez Molledo, Maria, Esben M Quistgaard, and Christian Löw. 2018. "Tripeptide Binding in a Proton-Dependent Oligopeptide Transporter.." *FEBS Letters*, September. Wiley-Blackwell. doi:10.1002/1873-3468.13246.
McCarthy, Andrew A, Ray Barrett, Antonia Beteva, Hugo Caserotto, Fabien Dobias, Franck Felisaz, Thierry Giraud, et al. 2018. "ID30B - a Versatile Beamline for Macromolecular Crystallography Experiments at the ESRF.." *Journal of Synchrotron Radiation* 25 (Pt 4). International Union of Crystallography: 1249–60. doi:10.1107/S1600577518007166.
McCoy, Airlie J, Ralf W Grosse-Kunstleve, Paul D Adams, Martyn D Winn, Laurent C Storoni, and Randy J Read. 2007. "Phaser Crystallographic Software.." *Journal of Applied Crystallography* 40 (Pt 4). International Union of Crystallography: 658–74. doi:10.1107/S0021889807021206.
Minhas, Gurdeep S, Daniel Bawdon, Reyme Herman, Michelle Rudden, Andrew P Stone, A Gordon James, Gavin H Thomas, and Simon Newstead. 2018. "Structural Basis of Malodour Precursor Transport in the Human Axilla.." *eLife* 7 (July). eLife Sciences Publications Limited: e34995. doi:10.7554/eLife.34995.
Moriarty, Nigel W, Ralf W Grosse-Kunstleve, and Paul D Adams. 2009. "Electronic Ligand Builder and Optimization Workbench (eLBOW): a Tool for Ligand Coordinate and Restraint Generation.." *Acta Crystallographica Section D: Biological Crystallography* 65 (Pt 10). International Union of Crystallography: 1074–80. doi:10.1107/S0907444909029436.
Nelson, David L, Albert L Lehninger, and Cox Michael M. 2000. *Lehninger Principles of Biochemistry*. 3rd ed. Worth.
Newstead, Simon, David Drew, Alexander D Cameron, Vincent L G Postis, Xiaobing Xia, Philip W Fowler, Jean C Ingram, et al. 2011. "Crystal Structure of a Prokaryotic Homologue of the Mammalian Oligopeptide-Proton Symporters, PepT1 and PepT2.." *The EMBO Journal* 30 (2): 417–26. doi:10.1038/emboj.2010.309.
Nurizzo, Didier, Trevor Mairs, Matias Guijarro, Vicente Rey, Jens Meyer, Pablo Fajardo, Joel Chavanne, Jean Claude Biasci, Sean McSweeney, and Edward Mitchell. 2006. "The ID23-1 Structural Biology Beamline at the ESRF.." *Journal of Synchrotron Radiation* 13 (Pt 3). International Union of Crystallography: 227–38. doi:10.1107/S0909049506004341.
O'Brien, W J, and F E Frerman. 1980. "The Effect of Temperature and Membrane Lipid Composition on the Rate of Beta-Oxidation by Escherichia Coli.." *Biochimica Et Biophysica Acta* 617 (1): 20–27.
Oeffner, Robert D, Gábor Bunkóczi, Airlie J McCoy, and Randy J Read. 2013. "Improved Estimates of Coordinate Error for Molecular Replacement.." *Acta Crystallographica Section D: Biological Crystallography* 69 (Pt 11). International Union of Crystallography: 2209–15. doi:10.1107/S0907444913023512.

Pardon, E, T Laeremans, S Triest, and SGF Rasmussen. 2014. “A General Protocol for the Generation of Nanobodies for Structural Biology.” *Nature Protocols*.

Parker, Joanne L, Chenghan Li, Allete Brinth, Zhi Wang, Lutz Vogeley, Nicolae Solcan, Gregory Ledderboge-Vucinic, et al. 2017. “Proton Movement and Coupling in the POT Family of Peptide Transporters..” *Proceedings of the National Academy of Sciences of the United States of America*, November. National Acad Sciences, 201710727. doi:10.1073/pnas.1710727114.

Prabhala, Bala K, N, Nanda G Aduri, a G Aduri, Johanne M Jensen, Heidi A Ernst, Nida Iram, Moazur Rahman, and Osman Mirza. 2014. “New Insights Into the Substrate Specificities of Proton-Coupled Oligopeptide Transporters From E. Coli by a pH Sensitive Assay..” *FEBS Letters* 588 (4): 560–65. doi:10.1016/j.febslet.2014.01.004.

Prabhala, Bala K, Nanda G Aduri, Mazhar Iqbal, Moazur Rahman, Michael Gajhede, Paul R Hansen, and Osman Mirza. 2017. “Several hPepT1-Transported Drugs Are Substrates of the Escherichia Coli Proton-Coupled Oligopeptide Transporter YdgR..” *Research in Microbiology*, February. doi:10.1016/j.resmic.2017.01.005.

Quistgaard, Esben M, Christian Löw, Per Moberg, Lionel Trésaugues, and Pär Nordlund. 2013. “Structural Basis for Substrate Transport in the GLUT-Homology Family of Monosaccharide Transporters..” *Nature Structural & Molecular Biology* 20 (6): 766–68. doi:10.1038/nsmb.2569.

Quistgaard, Esben M, Christian Löw, Fatma Guettou, and Pär Nordlund. 2016. “Understanding Transport by the Major Facilitator Superfamily (MFS): Structures Pave the Way..” *Nature Reviews. Molecular Cell Biology* 17 (2): 123–32. doi:10.1038/nrm.2015.25.

Quistgaard, Esben M, Maria Martinez Molledo, and Christian Löw. 2017. “Structure Determination of a Major Facilitator Peptide Transporter: Inward Facing PepTSt From Streptococcus Thermophilus Crystallized in Space Group P3121.” *PloS One* 12 (3). Public Library of Science: e0173126. doi:10.1371/journal.pone.0173126.

Samsudin, Firdaus, Joanne L Parker, Mark S P Sansom, Simon Newstead, and Philip W Fowler. 2016. “Accurate Prediction of Ligand Affinities for a Proton-Dependent Oligopeptide Transporter.” *Cell Chemical Biology* 23 (2): 299–309. doi:10.1016/j.chembiol.2015.11.015.

Sáenz, James P, Daniel Grosser, Alexander S Bradley, Thibaut J Lagny, Oksana Lavrynenko, Martyna Broda, and Kai Simons. 2015. “Hopanoids as Functional Analogues of Cholesterol in Bacterial Membranes..” *Proceedings of the National Academy of Sciences of the United States of America* 112 (38). National Academy of Sciences: 11971–76. doi:10.1073/pnas.1515607112.

Schaeffer, H J, L Beauchamp, P de Miranda, G B Elion, D J Bauer, and P Collins. 1978. “9-(2-Hydroxyethoxymethyl) Guanine Activity Against Viruses of the Herpes Group..” *Nature* 272 (5654): 583–85.

Solcan, Nicolae, Jane Kwok, Philip W Fowler, Alexander D Cameron, David Drew, So Iwata, and Simon Newstead. 2012. “Alternating Access Mechanism in the POT Family of Oligopeptide Transporters.” *The EMBO Journal* 31 (16). John Wiley & Sons, Ltd: 3411–21. doi:10.1038/emboj.2012.157.

Staras, Stephanie A S, Sheila C Dollard, Kay W Radford, W Dana Flanders, Robert F Pass, and Michael J Cannon. 2006. “Seroprevalence of Cytomegalovirus Infection in the United States, 1988-1994..” *Clinical Infectious Diseases : an Official Publication of the Infectious Diseases Society of America* 43 (9): 1143–51. doi:10.1086/508173.

Steiner, H Y, F Naider, and J M Becker. 1995. “The PTR Family: a New Group of Peptide Transporters..” *Molecular Microbiology* 16 (5): 825–34.
Sugawara, M, W Huang, Y J Fei, F H Leibach, V Ganapathy, and M E Ganapathy. 2000. “Transport of Valganciclovir, a Ganciclovir Prodrug, via Peptide Transporters PEPT1 and PEPT2..” *Journal of Pharmaceutical Sciences* 89 (6): 781–89. doi:10.1002/(SICI)1520-6017(200006)89:6<781::AID-JPS10>3.0.CO;2-7.
Wallace, A C, R A Laskowski, and J M Thornton. 1995. “LIGPLOT: a Program to Generate Schematic Diagrams of Protein-Ligand Interactions..” *Protein Engineering* 8 (2): 127–34.
Weitz, Dietmar, Daniel Harder, Fabio Casagrande, Dimitrios Fotiadis, Petr Obrdlik, Bela Kelety, and Hannelore Daniel. 2007. “Functional and Structural Characterization of a Prokaryotic Peptide Transporter with Features Similar to Mammalian PEPT1.” *The Journal of Biological Chemistry* 282 (5). American Society for Biochemistry and Molecular Biology: 2832–39. doi:10.1074/jbc.M604866200.
Woestenenk, Esmeralda A, Martin Hammarström, Susanne van den Berg, Torleif Härd, and Helena Berglund. 2004. “His Tag Effect on Solubility of Human Proteins Produced in Escherichia Coli: a Comparison Between Four Expression Vectors.” *Journal of Structural and Functional Genomics* 5 (3). Kluwer Academic Publishers: 217–29. doi:10.1023/B:jsfg.0000031965.37625.0e.
Xu, Liya, Ian S Haworth, Ashutosh A Kulkarni, Michael B Bolger, and Daryl L Davies. 2009. “Mutagenesis and Cysteine Scanning of Transmembrane Domain 10 of the Human Dipeptide Transporter..” *Pharmaceutical Research* 26 (10). Springer US: 2358–66. doi:10.1007/s11095-009-9952-9.
Yan, Nieng. 2013. “Structural Advances for the Major Facilitator Superfamily (MFS) Transporters..” *Trends in Biochemical Sciences* 38 (3). Elsevier: 151–59. doi:10.1016/j.tibs.2013.01.003.
Yan, Nieng. 2015. “Structural Biology of the Major Facilitator Superfamily (MFS) Transporters.” *Annual Review of Biophysics* 44. Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, California 94303-0139, USA. doi:10.1146/annurev-biophys-060414-033901.
Yeung, A K, S K Basu, S K Wu, C Chu, C T Okamoto, S F Hamm-Alvarez, H von Grafenstein, et al. 1998. “Molecular Identification of a Role for Tyrosine 167 in the Function of the Human Intestinal Proton- Coupled Dipeptide Transporter (hPepT1)..” *Biochemical and Biophysical Research Communications* 250 (1): 103–7.
Zhao, Yan, Guotao Mao, Min Liu, Laixing Zhang, Xianping Wang, and Xuejun C Zhang. 2014. “Crystal Structure of the E. Coli Peptide Transporter YbgH..” *Structure* 22 (8). Elsevier Ltd: 1152–60. doi:10.1016/j.str.2014.06.008.

6. Appendix

6.1. Crystallographic statistics

Table 16: Crystallographic statistics for the three DtpA-N00 structures

	DtpA-N00 in glycine buffer PDB id: 6GS7	DtpA-N00 in MES buffer PDB id: 6GS1	DtpA-N00-Valganciclovir PDB id: 6GS4
Data collection			
Space group	$P\ 2_12_12_1$	$P\ 2_12_12_1$	$P\ 2_12_12_1$
Cell dimensions			
a, b, c (Å)	55.14 120.53 163.43	55.46 120.72 163.33	54.94 120.19 163.67
α, β, γ (°)	90 90 90	90 90 90	90 90 90
Resolution (Å)	19.94-3.3 (3.42-3.3)	48.16-3.29 (3.41-3.29)	19.68-2.65 (2.74-2.65)
R_{merge}	21.74 (217.7)	26.73 (186.4)	10.74 (207.7)
I / σI	7.04 (0.83)	6.37 (1.06)	10.74 (1.06)
$CC_{1/2}$	0.992 (0.593)	0.993 (0.498)	0.998 (0.402)
Completeness (%)	99.02 (99.70)	99.34 (99.47)	98.98 (97.29)
Redundancy	8.5 (8.8)	6.7 (6.6)	6.4 (6.6)
Refinement			
Resolution (Å)	19.94 3.3 (3.42-3.3)	48.16-3.29 (3.41-3.29)	19.68-2.65 (2.74-2.65)
No. reflections	16919 (1671)	17219 (1679)	32178 (3122)
R_{work} / R_{free}	21.03 / 24.75	25.11 / 26.60	21.57 / 23.96
No. atoms			
Protein	4568	4576	4542
Ligand/ion			60
Water		1	7
B-factors			
Protein	108.69	83.22	89.74
Ligand/ion			133.49
Water		59.18	74.7
R.m.s. deviations			
Bond lengths (Å)	0.004	0.004	0.004
Bond angles (°)	0.92	0.91	0.64

*Values in parentheses are for highest-resolution shell.

6.2. RUBIC screen

Table 17: Conditions in RUBIC screen

Well_No	Conc	Units	Reagent	Conc	Units	Buffer	pH
A1	100	%	Ultrapure water				
A2				0.119	M	Citrate	4
A3				0.119	M	Sodium acetate	4.5
A4				0.119	M	Citrate	5
A5				0.119	M	MES	6
A6				0.119	M	Potassium phosphate	6
A7				0.119	M	Citrate	6
A8				0.119	M	Bis-Tris	6.5
A9				0.119	M	MES	6.5
A10				0.119	M	Sodium phosphate	7
A11				0.119	M	Potassium phosphate	7
A12				0.119	M	HEPES	7
B1				0.119	M	MOPS	7
B2				0.119	M	Ammonium acetate	7.3
B3				0.119	M	Tris-HCl	7.5
B4				0.119	M	Sodium phosphate	7.5
B5				0.119	M	Imidazole	7.5
B6				0.119	M	HEPES	8
B7				0.119	M	Tris-HCl	8
B8				0.119	M	Tricine	8
B9				0.119	M	BICINE	8
B10				0.119	M	BICINE	8.5
B11				0.119	M	Tris-HCl	8.5
B12				0.119	M	CHES	9
C1	0.298	M	Sodium chloride				
C2	0.298	M	Sodium chloride	0.119	M	Citrate	4
C3	0.298	M	Sodium chloride	0.119	M	Sodium acetate	4.5
C4	0.298	M	Sodium chloride	0.119	M	Citrate	5
C5	0.298	M	Sodium chloride	0.119	M	MES	6
C6	0.298	M	Sodium chloride	0.119	M	Potassium phosphate	6
C7	0.298	M	Sodium chloride	0.119	M	Citrate	6
C8	0.298	M	Sodium chloride	0.119	M	Bis-Tris	6.5
C9	0.298	M	Sodium chloride	0.119	M	MES	6.5

C10	0.298	M	Sodium chloride	0.119	M	Sodium phosphate	7
C11	0.298	M	Sodium chloride	0.119	M	Potassium phosphate	7
C12	0.298	M	Sodium chloride	0.119	M	CHES	7
D1	0.298	M	Sodium chloride	0.119	M	MOPS	7
D2	0.298	M	Sodium chloride	0.119	M	Ammonium acetate	7.3
D3	0.298	M	Sodium chloride	0.119	M	Tris-HCl	7.5
D4	0.298	M	Sodium chloride	0.119	M	Sodium phosphate	7.5
D5	0.298	M	Sodium chloride	0.119	M	Imidazole	8
D6	0.298	M	Sodium chloride	0.119	M	HEPES	8
D7	0.298	M	Sodium chloride	0.119	M	Tris-HCl	8
D8	0.298	M	Sodium chloride	0.119	M	Tricine	8
D9	0.298	M	Sodium chloride	0.119	M	BICINE	8
D10	0.298	M	Sodium chloride	0.119	M	BICINE	8.5
D11	0.298	M	Sodium chloride	0.119	M	Tris-HCl	8.5
D12	0.298	M	Sodium chloride	0.119	M	CHES	9
E1				0.119	M	SPG	4
E2				0.119	M	SPG	4.5
E3				0.119	M	SPG	5
E4				0.119	M	SPG	5.5
E5				0.119	M	SPG	6
E6				0.119	M	SPG	6.5
E7				0.119	M	SPG	7
E8				0.119	M	SPG	7.5
E9				0.119	M	SPG	8
E10				0.119	M	SPG	8.5
E11				0.119	M	SPG	9
E12				0.119	M	SPG	10
F1				0.024	M	HEPES	7.5
F2				0.06	M	HEPES	7.5
F3				0.149	M	HEPES	7.5
F4				0.298	M	HEPES	7.5
F5				0.024	M	Sodium phosphate	7.5
F6				0.06	M	Sodium phosphate	7.5

F7				0.149	M	Sodium phosphate	7.5
F8				0.298	M	Sodium phosphate	7.5
F9				0.024	M	Tris-HCl	8
F10				0.06	M	Tris-HCl	8
F11				0.149	M	Tris-HCl	8
F12				0.298	M	Tris-HCl	8
G1	0.06	M	Sodium chloride	0.06	M	HEPES	7.5
G2	0.149	M	Sodium chloride	0.06	M	HEPES	7.5
G3	0.298	M	Sodium chloride	0.06	M	HEPES	7.5
G4	0.595	M	Sodium chloride	0.06	M	HEPES	7.5
G5	0.893	M	Sodium chloride	0.06	M	HEPES	7.5
G6	1.19	M	Sodium chloride	0.06	M	HEPES	7.5
G7	0.06	M	Sodium chloride	0.06	M	Tris-HCl	8
G8	0.149	M	Sodium chloride	0.06	M	Tris-HCl	8
G9	0.298	M	Sodium chloride	0.06	M	Tris-HCl	8
G10	0.595	M	Sodium chloride	0.06	M	Tris-HCl	8
G11	0.893	M	Sodium chloride	0.06	M	Tris-HCl	8
G12	1.19	M	Sodium chloride	0.06	M	Tris-HCl	8
H1				0.06	M	MES/Bis-Tris	6
H2				0.06	M	MES/Imidazole	6.5
H3				0.06	M	Bis-Tris/PIPES	6.5
H4				0.06	M	MOPS/Bis-Tris propane	7
H5				0.06	M	Phosphate/Citrate	7.5
H6				0.06	M	MOPS/Sodium HEPES	7.5
H7				0.06	M	BICINE/Tris	8.5
H8	0.119	M	Sodium chloride	0.06	M	Imidazole	7.5
H9	0.119	M	Sodium chloride	0.149	M	Imidazole	7.5
H10	0.119	M	Sodium chloride	0.298	M	Imidazole	7.5
H11	0.119	M	Sodium chloride	0.417	M	Imidazole	7.5
H12	0.119	M	Sodium chloride	0.595	M	Imidazole	7.5

7. Acknowledgements

First, I would like to thank Dr Christian Löw for trusting me with this exciting project, his scientific enthusiasm, open door and the encouragement to develop new ideas and test them. I am also thankful to the members of the Thesis Advisory Committee: Prof Dr Samuel Wagner, Prof Dr Thilo Stehle, Dr Rob Meijers and Dr John Briggs for their valuable suggestions throughout the project. Next, my thanks go to the Löw group with its present and former members, which felt like a second family: To Dr Maria Martinez Molledo, with whom we started as the first members of this group and learned so much together, to Dr. Ali Flayhan, who helped me to get started with crystallography, to Joanna Pieprzyk, who keeps our lab in great shape, to Kim Bartels who continues to work with the same protein as me and all others for the great time. I would also like to thank Vasileos Rantos and Dr. Jan Kosinski for the nice collaboration and the SPC facility and the beam line scientists for their support. Thanks to EMBL Hamburg admin and IT groups for making my life easier and the EMBL Int. PhD programme and the Christiane Nusslein Volhard Foundation for the funding. Finally, I would like to thank my family for being always there for me.

8. Publications

Yonca Ural-Blimke, Ali Flayhan, Jan Strauss, Vasileios Rantos, Kim Bartels, Rolf Nielsen, Els Pardon, Jan Stayaert, Jan Kosinski, Esben M. Quistgaard, Christian Löw. 2018. (manuscript submitted) 'Structural basis of (pro)drug transport in the SLC15 family of peptide transporters

Ali Flayhan, Haydyn D T Mertens, Yonca Ural-Blimke, Maria Martinez Molledo, Dmitri I Svergun, Christian Löw. 2018. 'Saposin Lipid Nanoparticles: a Highly Versatile and Modular Tool for Membrane Protein Research', Structure, 26.2: 345–45

Esben M Quistgaard, Ulrich Weininger, Yonca Ural-Blimke, Kristofer Modig, Pär Nordlund, Mikael Akke, Christian Löw. 2016. 'Molecular Insights Into Substrate Recognition and Catalytic Mechanism of the Chaperone and FKBP Peptidyl-Prolyl Isomerase SlyD', BMC Biology, 14.1 (BioMed Central): 82

Lisa M Smith, Hernán A Burbano, Xi Wang, Joffrey Fitz, George Wang, Yonca Ural-Blimke, Detlef Weigel. 2015. 'Rapid Divergence and High Diversity of miRNAs and miRNA Targets in the Camelineae', The Plant Journal: for Cell and Molecular Biology, 81.4: 597–610