EMBL European Molecular Biology Laboratory

Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Master Thesis

Faculty of Life Sciences

Department Biotechnology

Pharmaceutical Biotechnology

Characterization and crystallization of proton-dependent oligopeptide transporters

Submitted: 20.04.2015

Svenja Hornung
Matriculation number: 2178863

1. Examiner: **Prof. Dr. Jörg Andrä**
HAW Hamburg

2. Examiner: **Dr. Christian Löw**
EMBL Hamburg

Abstract

POTs (proton-dependent oligopeptide transporters) are integral membrane proteins and essential for maintaining homeostasis in cells by switching between two major conformations during the transport cycle. Di- and tripeptides as well as some small peptide like compounds are recognized and transported across the membrane. This fact leads to the pharmacological interest in these transporters for drug delivery. In humans two transporters, PepT1 and PepT2, occur. So far only structures of five bacterial homologues are available.

Eukaryotic POTs exist of 12 transmembrane helices whereas prokaryotic POTs have two additional transmembrane helices. It was possible to clone different prokaryotic transporter constructs into pET expression vectors as well as the PepT-like transporter (CtPOT) of *Chaetomium thermophilum,* a thermophilic eukaryote, and the human PepT2. Two transporters of *Shewanella oneidensis* ($PepT_{So2}$) modified with the thermostabilized apocytochrome b_{562}RIL as well as the CtPOT and PepT2 were expressed in *Escherichia coli* cells. These transporters except PepT2 were solubilized and purified. Crystals were obtained of all proteins but it was not possible to solve any structure so far due to limited crystal diffraction.

Content

Abbreviations

BRIL	apocytochrome b_{562}RIL
BSA	bovine serum albumin
Cam	Chloramphenicol
CMC	critical micelle concentration
dCTP	deoxycytidine triphosphate
ddH_2O	double-distilled water
DDM	n-Dodecyl-β-D-Maltopyranoside
dGTP	deoxyguanosine triphosphate
DLS	dynamic light scattering
DM	n-Decyl-β-D-Maltopyranoside
DMNG	Decyl Maltose Neopentyl Glycol
dNTP	deoxyribonucleotide triphosphate
E. coli	*Escherichia coli*
GF	gel filtration
HRP	horse-radish peroxidase
IMP	integral membrane protein
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Kan	Kanamycin
LDAO	n-Dodecyl-N,N-Dimethylamine-N-Oxide
LIC	ligation-independent cloning
LT4	T4 lysozyme
Met-Ala-Ser	Methionine-Serine-Alanine
MW	molecular weight
MWCO	molecular weight cut off
Ni-NTA	nickel-nitrilotriacetic acid
OG	n-Octyl-β-D-Glucopyranoside
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
POT	proton-dependent oligopeptide transporter
PVDF	polyvinylidene fluoride
TB	teriffic broth
TBST	Tris buffered saline tween®-20 detergent
TEV	tobacco etch virus
Tyr-Ala	Tyrosine-Alanine

1. Introduction

For maintaining homeostasis in cells, short chain peptides are transported across cell membranes (Guettou *et al.*, 2013; Solcan *et al.*, 2012). Beside peptide transporters like the ABC transporters, the transport occurs via proton-dependent oligopeptide transporters (POTs) which belong to the major facilitator superfamily (Newstead, 2011; Solcan *et al.*, 2012). POT transporters are integral membrane proteins found in all organisms (Newstead, 2011). A wide range of di- and tripeptides are recognized and transported across the membrane as well as compounds with similar stereochemical properties and short peptides. Three different conformational states, inward open, occluded and outward open, are involved in the transport (Solcan *et al.*, 2012) (see figure 1).

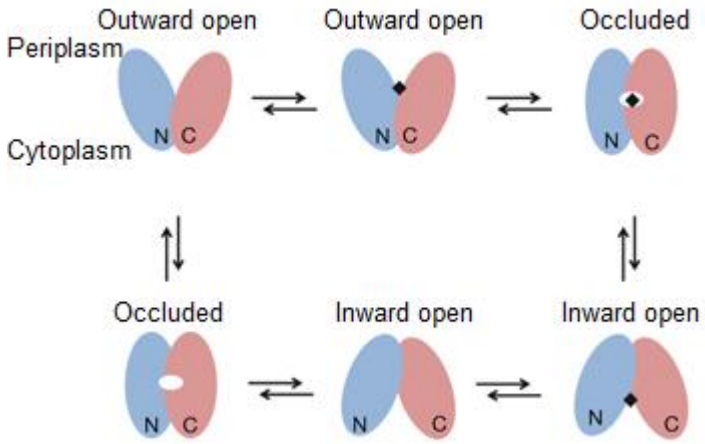


Figure 1: Alternating access transport mechanism of POT transporter. The N-terminal bundle is colored in blue and the C-terminal bundle is colored in red. In the binding pocket the located substrate is shown as a black square. Substrates are transported from periplasm to cytoplasm by conformation changes of the transporter. After substrate binding in the outward open conformation the transporter switches via the occluded state to the inward open form. The substrate is released and the transporter changes its conformation back to the outward open form via the occluded state with no bound substrate (Guettou et al., 2013).

In the outward open conformation the substrate binds in the binding pocket. The transporter changes its conformation to the occluded and inward open form. In this conformation the substrate is released and the transporter changes its conformation back to the outward open form via the occluded stage. For transport and accumulation of substrates above extracellular concentrations the driving force is an electrochemical proton gradient between the periplasm and cytoplasm (Guettou *et al.*, 2013).

In humans two POT transporters, PepT1 and PepT2, are present for the uptake of short peptides. But PepT1 and PepT2 also transport drugs and prodrugs across membranes and therefore are of pharmacological importance. For conversion of pharmacological active

compounds into substrates for human POT transporters the structure and an understanding of the binding mechanisms are needed.

The difference between the prokaryotic and eukaryotic transporter structure is shown schematically in the following figure.

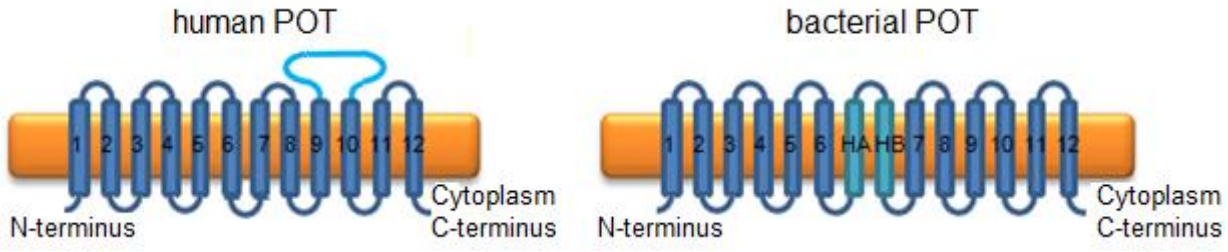


Figure 2: Schematic illustration of the human (left) and bacterial (right) POT transporter. The human POT transporter has 12 predicted transmembrane helices and between helix 9 and 10 a large extracellular loop (Han, Le, Shi, 2003). Bacterial POT transporters instead have 14 transmembrane helices with the two additional helices HA and HB. In human and bacterial transporters both termini are facing the cytoplasm (Guettou et al*., 2014).*

The human PepT1 and PepT2 are predicted to have 12 transmembrane helices with a large loop on the extracellular side between helix 9 and 10 (Han, Le, Shi, 2003). Both termini are facing the cytoplasm side. The bacterial POT transporter consists of 14 transmembrane helices including two additional helices HA and HB which do not occur in human POTs. Up to now there are only structures of five bacterial homologs available: GkPOT (*Geobacillus kaustophilus*) and $PepT_{So2}$ (*Shewanella oneidensis)* co-crystallized with alafosfalin both in the inward-open form, $PepT_{So}$ (*Shewanella oneidensis*) in the occluded structure and $PepT_{St}$ (*Streptococcus thermophilus*) in the inward-open conformation (Guettou *et al.*, 2014) as well as DtpD (*Escherichia coli*) (Zhao *et al*., 2014).

For the co-crystallization of $PepT_{So2}$ with alafosfalin the sitting drop method was used. The crystal beyond a resolution of 4 Å was obtained after several rounds of crystal optimization from the initial crystals of the MemGold screen. After 21 days the best diffracting crystals appeared at 20 °C.

The structure of $PepT_{So2}$ with alafosfalin is shown in the following figure.

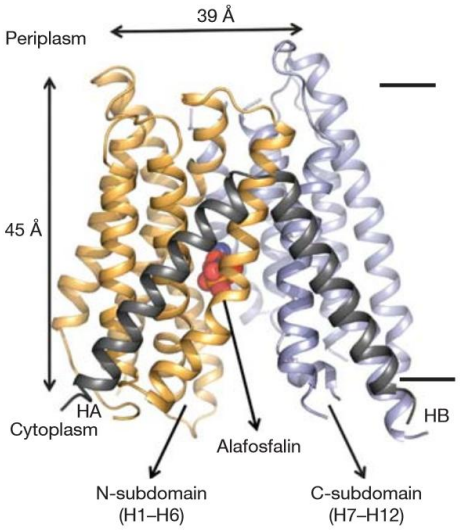


Figure 3: POT transporter $PepT_{So2}$ crystallized in complex with the compound alafosfalin in the inward-open conformation viewed from the plane of the membrane. Yellow and blue colored are the N- and C-terminal subdomains. The helices HA and HB are colored in grey, alafosfalin is shown as red spheres bound in the central cavity between the two subdomains. The approximate dimensions of the transporter are shown as well as the location of the membrane which is illustrated in black bars (Guettou et al., 2013).

The POT transporter from *Shewanella oneidensis* consists of 14 transmembrane helices which form two six-helical bundles and two additional transmembrane helices of unknown function. The two subdomains are connected by the helices HA and HB but their functional role is still unclear. Due to the fact that they are not present in eukaryotic homologues their contribution to the transport mechanism is predicted not to be essential (Guettou *et al.*, 2014).

To obtain higher resolution diffraction data or to trap the transporter in the outward open conformation, $PepT_{So2}$ was fused with the thermostabilized apocytochrome BRIL (b_{562}RIL) protein or LT4 (T4 lysozyme) (see figure 4).

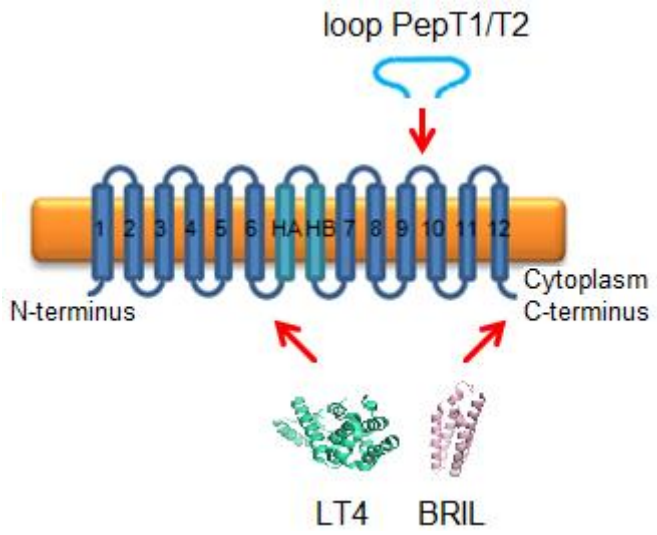


Figure 4: Schematic view of $PepT_{So2}$ constructs. LT4 and BRIL were either inserted in the loop between helix 6 and HA or at the C-terminus. The human loop of PepT1 and PepT2 was inserted into the bacterial loop between helix 9 and 10.

The prokaryotic sequence was replaced by BRIL and LT4 between loop 6 and HA ($PepT_{So2}$-BRIL/LT4-insert) and at the C-terminus ($PepT_{So2}$-BRIL/LT4-Cterm). For the constructs $PepT_{So2}$-PepT1/T2 the loop between helix 9 and 10 was replaced by the human loop of PepT1 and PepT2 to further characterize the folding of the human extracellular loop (see figure 4).

The fungus *Chaetomium thermophilum*, a thermophilic eukaryote, has its optimal growth temperature between 50 °C and 55 °C and therefore its proteins are more heat stable (Bock *et al.*, 2014). In *C. thermophilum* a peptide transporter-like protein (CtPOT) is found (NCBI Reference Sequence: XP_006696555.1) which has not been characterized up to now.

The human PepT1 and PepT2, CtPOT and the above mentioned bacterial constructs were studied in this thesis.

2. Methods

2.1. LIC and PCR

An efficient cloning technique without the use of T4 DNA ligase is LIC (ligation-independent cloning) (Aslanidis, de Jong, 1990). With this method it is possible to insert any PCR product into LIC-compatible vectors (Kato *et al.*, 2013). This technique uses the 3'-5' exonuclease activity of the T4 DNA polymerase (Warren *et al.*, 2013) to generate 5' single-stranded overhangs in the presence of a specific dNTP (deoxyribonucleotide triphosphate) (Kato *et al.*, 2013). The polymerase stops its activity as soon as it hits the dNTP that was added to the reaction (Rupp, 2010). For C-terminal cloning the addition of dCTP and for N-terminal cloning the addition of dGTP to the vector is used. For the genes the complement dNTP is used. The generated sticky ends hybridize with the complementary sticky ends in the vector and form the plasmid (Chen, Janes, 2013).

PCR (Polymerase-Chain-Reaction) is used to amplify a specific region of the DNA or RNA. For the reaction, primers, the DNA template, a thermostable DNA polymerase and nucleotides are needed. Primers are short oligonucleotides which are complementary to the fragment of DNA which should be amplified. The DNA template contains the fragment which is amplified. The DNA between the primers is amplified by the polymerase using the free nucleotides.

The first step in the PCR is the denaturation of the double-stranded DNA. The double-stranded DNA denatures at 94 °C into single-stranded DNA. The second step is the annealing of primers to the complementary sites on the DNA which allows the attachment of the polymerase. This takes place at temperatures between 45 °C and 60 °C. The last step is the elongation which occurs at 72 °C. To repeat these steps thermal cyclers are used (Canene-Adams, 2013).

In this work the vector pNIC-CTHF (Addgene plasmid # 26105) and pNIC28-Bsa4 (Addgene plasmid # 26103) were used for cloning. Primers for LIC were designed for the genes of interest with one part complementary to the genes and the other part complementary to the vector specific sequence.

The vectors are pET expression vectors and are shown in figure 5.

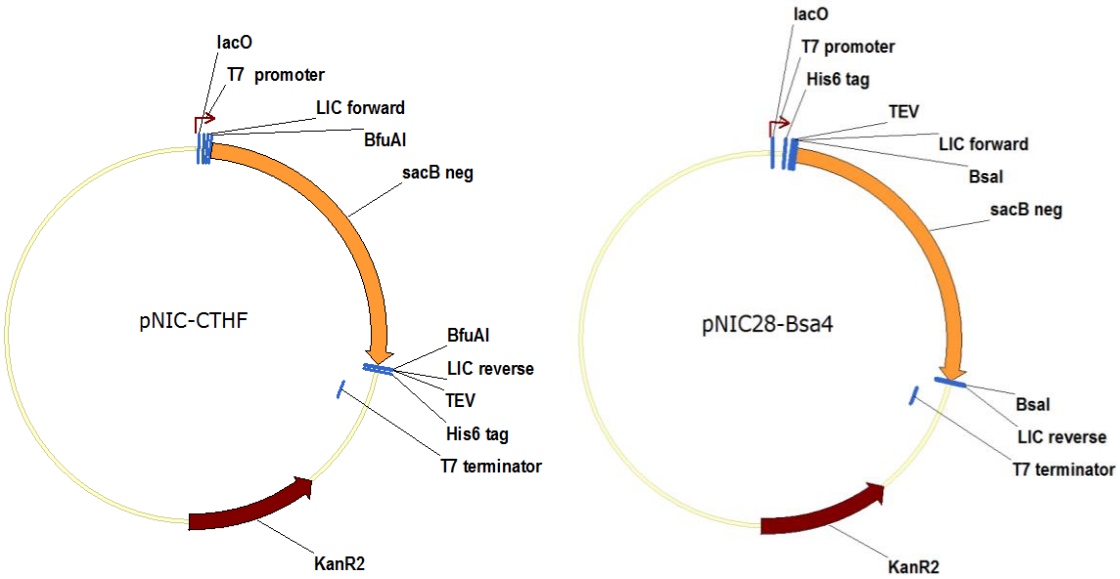


Figure 5: pNIC-CTHF (left) and pNIC28-Bsa4 (right) vectors for C- and N-terminal cloning. The vector pNIC-CTHF is used for cloning with a C-terminal His-tag and pNIC28-Bsa4 for cloning with a N-terminal His-tag. In both vectors a TEV cleavage site is included to cleave the His-tag of the fusion protein off as well as a Kanamycin resistance. During LIC the enzyme cleavage sites for BfuAI and BsaI get destroyed.

The vector pNIC-CTHF has a His_6- and a FLAG-tag at the C-terminus whereas pNIC28-Bsa4 has a His_6-tag at the N-terminal fusion peptide. Both vectors have a TEV protease cleavage site to cleave the His-tag of the fusion protein and a Kan (Kanamycin) resistance. pNIC-CTHF is used for LIC by using the enzyme BfuAI whereas for pNIC28-Bsa4 the enzyme BsaI is used. The *sacB* gene is included in both vectors which encodes the enzyme levansucrase. Hydrolysis of sucrose and synthesis of levans is catalyzed by this enzyme. In the presence of sucrose the expression of the *Bacillus subtilis sacB* gene in *E. coli* is lethal (Pelicic, Reyrat, Gicquel, 1995). Therefore a negative selection on agar plates with 5 % sucrose is possible with these vectors. The pNIC-CTHF vector was used for all genes, for CtPOT, PepT1 and PepT2 additionally the pNIC28-Bsa4 vector was used.

2.1.1. Vector preparation for LIC

Both vectors needed to be prepared for LIC and were cleaved with their specific restriction enzyme. Therefore the components were mixed according to table 1.

Table 1: Cleavage-mix for vector preparation. BfuAI restriction enzyme is used for pNIC-CTHF and BsaI for pNIC28-Bsa4 to linearize the vector and cleave the sacB gene.

pNIC-CTHF		**pNIC28-Bsa4**	
Cleavage mix 1	Volume 1 [µl]	Cleavage mix 2	Volume 2 [µl]
Plasmid prep (~ 5 µg)	65	Plasmid prep (~ 5 µg)	65
10x NEB3 buffer	10	10x NEB3 buffer	10
100x BSA	1	100x BSA	1
BfuAI restriction enzyme	3	BsaI restriction enzyme	3
ddH_2O up to 100 µl	21	ddH_2O up to 100 µl	21

Each mix was incubated at 50 °C for 2-3 hours. The cleaved vectors were purified using the QIAquick PCR Purification Kit (Qiagen) and were eluted in 55 µl elution buffer.

After the cleavage of the vectors the sticky ends were created. Therefore the vectors were treated with the T4 DNA polymerase. The components for the treatment of the vectors were mixed according to table 2.

Table 2: T4 DNA polymerase treatment mixes for the vector preparations. The linearized vector is treated with the T4 DNA polymerase in the present of a vector specific dNTP to create 5' single stranded overhangs.

pNIC-CTHF		**pNIC28-Bsa4**	
T4 DNA polymerase treatment mix 1	Volume 1 [µl]	T4 DNA polymerase treatment mix 2	Volume 2 [µl]
BfuAI cleaved vector	50	BsaI cleaved vector	50
ddH_2O	24	ddH_2O	24
10x NEB2 buffer	10	10x NEB2 buffer	10
dCTP (25mM)	10	dGTP (25mM)	10
100x BSA	1	100x BSA	1
T4 DNA polymerase	5	T4 DNA polymerase	5

Each treatment mix was incubated at 22 °C for 30 minutes and heat inactivated at 70 °C for 20 minutes. 100 µl of ddH_2O (double-distilled water) were added to a final volume of 200 µl.

2.1.2. Amplification of DNA-fragment with appropriate primers

For the amplification of all genes the PCR mix was prepared for each gene according to table 3. For PepT1 the original plasmid was additionally diluted 1:100 and 1:1000 and thereof 0.5 µl were taken for the PCR mix.

Table 3: PCR mix for amplification of DNA-fragments. For the amplification of the genes of interest from the original plasmids the Phusion polymerase was used and the gene specific primers for N- and C-terminal cloning.

PCR mix	Volume [µl]
ddH_2O	33
5x Phusion buffer	10
dNTPs (10 mM)	1
Primer forward	2.5
Primer reverse	2.5
Gene	0.5
Phusion polymerase	0.5

As primers the gene specific primers were used (see Appendix 5.4 Nucleotide sequences and primers of used vectors and POTs). The PCR was run according to the following steps in table 4.

Table 4: PCR parameters for amplifications of transporter genes for LIC. Two cycles were used with different annealing temperatures to amplify the genes.

Step	Temperature [°C]	Time [min]	Repeats
Initial	98	0.5	-
Cycle 1	98	0.16	9
	55	0.5	
	72	2	
Cycle 2	98	0.16	19
	60	0.5	
	72	2	
Final	72	10	-
Hold	4	-	-

After PCR amplification the PCR products of PepT1 with the 1:100 and 1:1000 diluted gene of the original plasmid were treated with 3 µl of the restriction enzyme DpnI for 1 hour at 37 °C because the original plasmid of PepT1 has a Kanamycin resistance, too. This treatment was done to avoid the transformation of the original plasmid.

All PCR products were purified with the QIAquick PCR Purification Kit (Qiagen).

2.1.3. T4 DNA polymerase treatment of PCR products

All amplified PCR products were treated with the T4 DNA polymerase to generate sticky ends. The components of the cleavage-mix are shown in table 6.

Table 5: Components and volumes for the T4 DNA treatment mix for the two vectors. To generate 5' single stranded overhangs the PCR products were treated with T4 DNA polymerase in the present of the complement vector specific dNTP.

PCR product for pNIC-CTHF		**PCR product for pNIC28-Bsa4**	
T4 DNA polymerase treatment mix 3	Volume 3 [µl]	T4 DNA polymerase treatment mix 4	Volume 4 [µl]
ddH2O	2.5	ddH2O	2.5
purified PCR product	5	purified PCR product	5
10x NEB2 buffer	1	10x NEB2 buffer	1
dGTP (25 mM)	1	dCTP (25 mM)	1
100x BSA	0.1	100x BSA	0.1
T4 DNA Polymerase	0.4	T4 DNA Polymerase	0.4

Each reaction mix was incubated for 30 minutes at room temperature followed by incubation for 20 minutes at 70 °C.

2.1.4. Cloning

2.5 µl of the T4 DNA polymerase treated vectors were mixed with 2.5 µl of the prepared PCR fragments and incubated at room temperature for 15 minutes. 5 µl of cloned plasmids were transformed into competent DH5α cells (see 2.2 Transformation of plasmids into competent *E. coli* cells) and were plated on LB-Kan-plates with 5 % sucrose. The plates were incubated overnight at 37 °C.

2.1.5. Screening

For the screening 3 ml TB (teriffic broth) medium with 3 µl Kanamycin were prepared for each colony to be tested. The cultures were inoculated with one colony and grew overnight at 37 °C and 200 rpm. The next day the plasmids were purified with the QIAprep Spin Miniprep Kit (Qiagen).

For the amplification of the three genes a PCR volume of 50 µl was used. The components are shown above in table 3.

Primers, complement to the T7 promoter and T7 terminator were used in the PCR reaction. The PCR was run according to table 6.

Table 6: Parameters of the PCR for the amplification of cloned genes. For PCR the T7 promoter and terminator were used. Therefore the annealing temperature was set to 53 °C and only one cycle was used.

Phase	Temperature [°C]	Time [min]	Repeats
Initial	98	1	-
Cycle	98	0.16	29
	53	0.5	
	72	2	
Final	72	10	-
Hold	4	-	-

The PCR products were purified and analyzed on a 1 % agarose gel.

The following transporter constructs were successfully cloned:

Table 7: Cloned POT transporters with the position of the His-tag, cloned vector and the molecular weight are shown. For PepT2 and CtPOT plasmids with N- and C-terminal His-tag and for the $PepT_{So2}$ constructs only pNIC-CTHF with a C-terminal His-tag was used.

Name	DNA construct	End of His-tag	Vector	MW [kDa] cleaved
His-PepT2	PepT2	N-terminal	pNIC28-Bsa4	81.1
PepT2-His	PepT2	C-terminal	pNIC-CTHF	
His-CtPOT	CtPOT	N-terminal	pNIC28-Bsa4	67.5
CtPOT-His	CtPOT	C-terminal	pNIC-CTHF	
$PepT_{So2}$-BRIL-insert	$PepT_{So2}$-BRIL-insert			67.3
$PepT_{So2}$-BRIL-Cterm	$PepT_{So2}$-BRIL-Cterm			66.4
$PepT_{So2}$-LT4-insert	$PepT_{So2}$-LT4-insert			73.9
$PepT_{So2}$-LT4-Cterm	$PepT_{So2}$-LT4-Cterm			72.7
$PepT_{So2}$-PepT1	$PepT_{So2}$-PepT1			78.7
$PepT_{So2}$-PepT2	$PepT_{So2}$-PepT2			80.0

The above mentioned constructs were successfully cloned into pNIC-CHTF vector, PepT2 and CtPOT were additionally cloned into pNIC-Bsa4 vector.

2.2. Transformation of plasmids into competent *E. coli* cells

The transfer of free DNA molecules into competent bacterial cells is referred as transformation. This process allows the specific integration of DNA molecules like plasmids into bacterial cells and therefore the expression of recombinant proteins (Griffiths *et al.*, 2004).

In this work the competent *E. coli* (*Escherichia coli)* cells DH5α, BL21 (DE3), BL21 Star (DE3), BL21 (DE3) pLysS (Cam (Chloramphenicol) resistance), BL21 (DE3) RIL (Cam resistance), Rosetta 2 (DE3) (Cam resistance) and C41 (DE3) were used.

The DH5α cells were used for plasmid isolation, the other strains for protein expression. BL21 (DE3) cells are used for general purpose expression as well as BL21 Star (DE3) cells but they are modified for reduced mRNA degradation. The strain BL21 (DE3) pLysS is used for high stringency expression, meaning that leaky expression is suppressed. BL21 (DE3) RIL allows the expression of genes encoding tRNAs for rare arginine, isoleucine and leucine codons. Rosetta 2 (DE3) carries 7 tRNAs for rare arginine, isoleucine, leucine, glycine and proline codons. The strain C41 (DE3) was derived from BL21 (DE3) and carries one mutation within the promoter driving expression of T7 RNA polymerase leading to an attenuation of the robust T7 expression of DE3 strains. Therefore this strain is advantageous for the expression of toxic or membrane proteins (Robinson, 2011).

Competent *E. coli* cells were incubated on ice with 5 µl of freshly cloned plasmids or 0.5 µl of sequenced plasmids for 30 minutes. In a water bath the cells were heat-shocked at 42 °C for 45 seconds. 600 µl of TB were added and the cells were incubated for 1 hour at 37 °C. At room temperature the cells were centrifuged at 5,000 rpm for 3 minutes and the supernatant was discarded. The pellet was resuspended in the remaining liquid and plated on a LB-Kan-agar plate and for the cells with Chloramphenicol resistance on a LB-Kan-Cam-agar plate. The plate was incubated over night at 37 °C.

For plasmid isolation two Falcon tubes were each filled with 10 ml TB and 10 µl Kan (stock solution: c = 30 mg/ml). Each tube was inoculated with one colony from the agar plate and incubated in a shaker for about 6 hours at 250 rpm and 37 °C. The isolation of plasmid DNA of DH5α cells was done with the Kit QIAprep Spin Miniprep Kit (QIAGEN). The principal of the isolation is based on alkaline lysis of bacterial cells (Qiagen, 2006). The characteristic of the complete renaturation of plasmid DNA after an alkaline denaturation is used to isolate it from chromosomal DNA (Casali, Preston, 2004).

2.3. Membrane protein over-expression and solubilization

The over-expression of membrane proteins is important to obtain enough material for studies about protein structure and function. The problem of over-expression is the toxicity of these proteins due to accumulation in membranes. Membrane integrity and cell viability are affected by these over-expressed membrane proteins leading to reduced cell growth and hampered division. Therefore membrane protein over-expression often leads to low yields (Wagner *et al.*, 2007).

The expression and insertion of IMPs (integral membrane proteins) in bacteria is shown in figure 6.

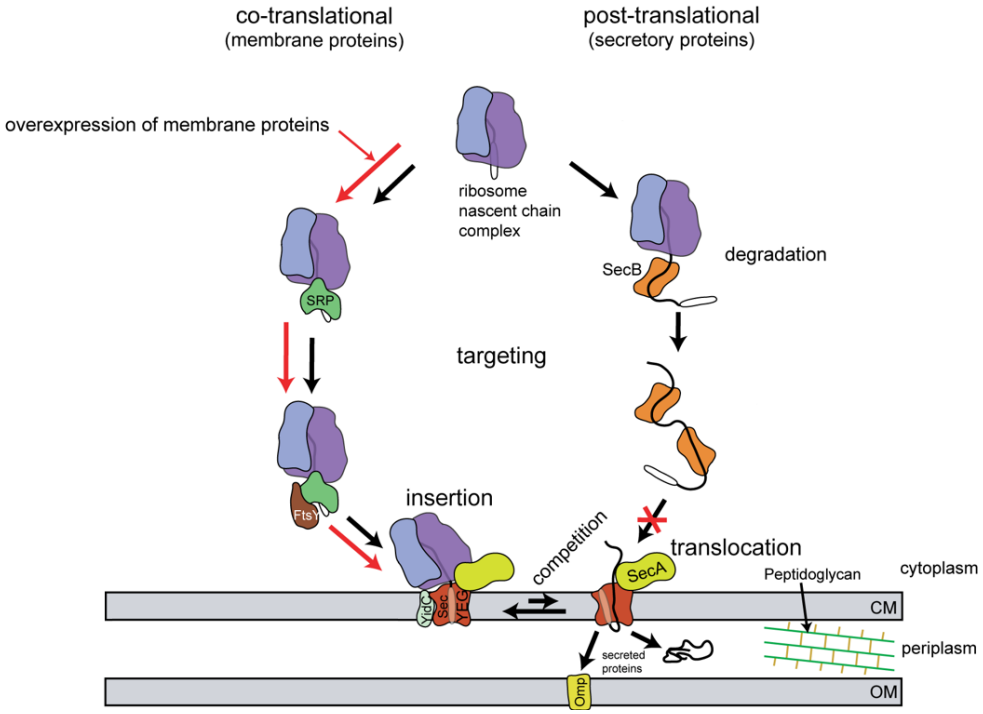


Figure 6: Expression and insertion of integral membrane proteins in bacteria. Via the SRP pathway membrane proteins are targeted to the SecYEG translocon. At the translocon the proteins are inserted into the membrane. Secreted proteins are targeted to SecYEG translocon via the SecB dependent pathway. Both pathways approach the same translocon resulting in a competition between the sorting of these proteins. During over-expression of membrane proteins the saturation of the protein sorting and translocon machineries can be a reason for the hampered cell division and affected membrane integrity (Wagner et al., 2007).

Membrane proteins are guided to the cytoplasmic membrane SecYEG translocon via the signal recognition particle pathway. The ribosome nascent chain complex binds to the SecYEG translocon and the nascent chain is inserted into the membrane. In the SecYEG translocon the transmembrane domains of membrane proteins get kept and move into the lipid bilayer. Secretory proteins are targeted via the SecB-dependent pathway to the cytoplasmic membrane. Due to the fact that both pathways approach at the Sec translocon

there is a competition between the sorting of these proteins. The saturation of the protein sorting and translocon machineries can be a reason for the hampered cell division and affected membrane integrity (Wagner *et al.*, 2007).

IMPs are inserted into the lipid bilayer. Therefore it is necessary to solubilize the membrane proteins with detergents in aqueous solution. In figure 7 the steps of membrane solubilization are shown.

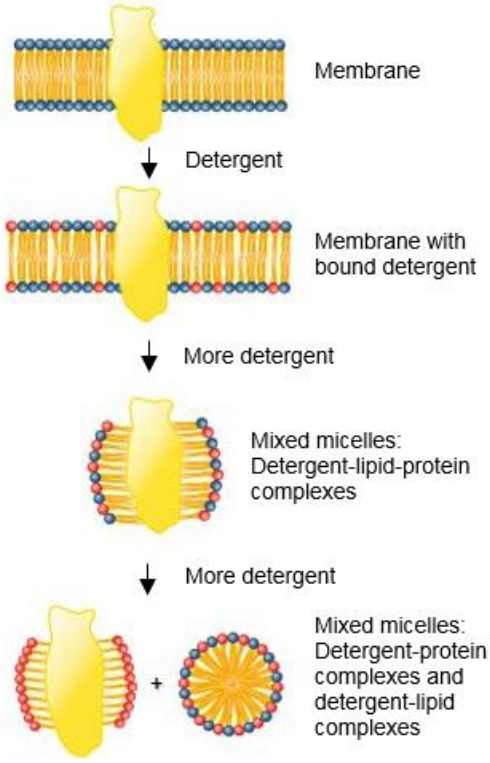


Figure 7: Different steps in membrane solubilization. The lipid bilayer surrounds the integral membrane protein. With addition of detergent these molecules get into the bilayer. The lipid bilayer is disrupted and mixed micelles are formed when adding more detergent. If more detergent is added lipids are displaced from the protein and detergent-protein and detergent-lipid complexes are formed (Luckey, 2014).

The IMP is surrounded by the lipid bilayer. With low amount of detergent some of these molecules insert into the bilayer. With the addition of more detergent the lipid bilayer is disrupted and mixed micelles are formed. When more detergent is added lipids are displaced from the protein resulting in detergent-protein and detergent-lipid complexes (Luckey, 2014).

Non-ionic detergents are used for mild and non-denaturing conditions because they don't break protein-protein interactions. In this way the protein stays biologically active (Seddon, 2004). DDM (n-Dodecyl-β-D-Maltopyranoside), a non-ionic detergent, was therefore used in this work to solubilize the membrane proteins for further purifications.

2.3.1. Small scale protein over-expression of POTs

For the test expression, all plasmids of the cloned POT transporter constructs were transformed (see 2.2 Transformation of plasmids into competent *E. coli* cells) into C41 (DE3) cells. For the CtPOT constructs as well as for the His-PepT2 the cell strains BL21 (DE3), BL21 Star (DE3), BL21 (DE3) pLysS, BL21 (DE3) RIL and Rosetta 2 (DE3) were additionally chosen.

The day before the test expression an overnight culture was set up with 10 ml TB (teriffic broth), 10 µl of the according antibiotics and one colony each.

For test expression a volume of 50 ml and a starting OD_{600} of 0.05 per ml were chosen. The cells grew at 37 °C and 200 rpm until an OD_{600} of 0.6 was reached. The cells were placed on the bench and afterwards at 18 °C and 200 rpm for 30 minutes each. Protein expression was induced with 0.2 mM IPTG (isopropyl β-D-1-thiogalactopyranoside) and proteins were expressed for 20 hours. 40 ml of the cultures were harvested at 4,000 rpm and 4 °C for 10 minutes and the pellets were stored frozen at -20 °C.

2.3.2. Small scale membrane preparation of POTs

All buffer compositions are shown in the appendix (5.5 Buffers). The frozen cell pellets were thawed on ice, resuspended in 15 ml lysis buffer I and incubated under stirring at 4 °C for 45 minutes. Cells were lysed by sonication (SonoPuls HD3200, Bandelin). Therefore the titanium probe Microtip (Sonopuls, Bandelin) was used. Cell lysis was done twice for 1 minute with a pulsation (on/off) of 0.8 seconds and amplitude of 25 %. Unbroken cells and cell debris were removed by centrifugation at 10,000 g and 4 °C for 10 minutes. 8 ml of the supernatant were taken and the membranes were collected by centrifugation at 30,000 rpm (Beckman, Ti45 rotor) and 4 °C for 50 minutes. After centrifugation the supernatant was discarded, the membranes were resuspended in 3 ml solubilization buffer I per 200 OD_{600} units and stored at -20 °C.

2.3.3. Solubilization test of POTs

For the solubilization test of CtPOT, His-PepT2, $PepT_{So2}$-BRIL/LT4-Cterm and $PepT_{So2}$-BRIL/LT4-insert 270 µl of protein solution were mixed with 30 µl of 10 % detergent to obtain a final concentration of 1 % detergent.

In the following table the tested detergents for each protein are shown.

Table 8: Used detergents for the solubilization test of different POT transporter. Tested detergents were used in a final concentration of 1 %.

	DDM	DMNG	$C_{12}E_8$	LDAO	DM	Triton	OG	NG	OGNG
His-CtPOT	X	X	X	X	X	X	X	X	X
CtPOT-His	X	X	X	X	X				
His-PepT2	X	X	X	X	X				
$PepT_{So2}$-BRIL-Cterm	X	X							
$PepT_{So2}$-BRIL-insert	X	X							
$PepT_{So2}$-LT4-Cterm	X	X							
$PepT_{So2}$-LT4-insert	X	X							

The samples and the reference (protein solution without detergent) were incubated for 45 minutes at 4 °C. Afterwards the detergent treated samples, but not the reference, were centrifuged for 30 minutes at 30,000 rpm and 4 °C. The supernatant was analyzed.

2.3.4. Large scale protein over-expression of CtPOT and $PepT_{So2}$-BRIL-Cterm/insert

The day before the expression 50 ml TB medium with 50 µl of antibiotics was inoculated with one colony from the agar plate. For CtPOT-His BL21 (DE3) RIL cells, for His-CtPOT and $PepT_{So2}$-BRIL-Cterm/-insert C41 (DE3) cells were used. The cultures were incubated at 37 °C and 200 rpm overnight. On the next day 500 ml TB medium with 500 µl of antibiotics were inoculated with the overnight culture to a start OD_{600} of 0.05 per ml. The cells were incubated at 37 °C and 200 rpm. At an OD_{600} of 1.0 the cultures were placed on the bench and afterwards in an incubator at 18 °C and 200 rpm each for 30 minutes. The expression was induced with 0.2 mM IPTG. The cells grew overnight and were harvested by centrifugation for 12 minutes at 10,000 g and 4 °C. The supernatant was discarded and the pellets were frozen at -20 °C.

2.4. Purification of CtPOT and $PepT_{So2}$-BRIL-Cterm/insert

2.4.1. Large scale membrane preparation

For the purification of CtPOT-His the cell pellet of a 0.5 l culture and for the His-CtPOT construct the pellet of 4 l culture were used. For the proteins $PepT_{So2}$-BRIL-Cterm and $PepT_{So2}$-BRIL-insert the pellets of 1 l culture were used.

Each frozen cell pellet was thawed on ice. It was resuspended in 5 ml lysis buffer II per 1 g of cell pellet (wet weight) and incubated under stirring for 45 minutes at 4 °C. The cells

were disrupted three times at 10,000-14,000 psi with the high-pressure homogenizer (EmulsiFlex-C3, Avestin) and centrifuged for 12 minutes at 10,000 g and 4 °C.

The supernatant was taken and centrifuged for 1 hour at 35,000 rpm and 4 °C. The supernatant was discarded and the membrane pellet was resuspendend in solubilization buffer II (half volume of lysis buffer). The membrane was solubilized under stirring for 30 minutes at 4 °C and centrifuged for 30 minutes at 30,000 rpm and 4 °C. The supernatant was taken and the membrane pellet was discarded.

2.4.2. Affinity chromatography of CtPOT and $PepT_{So2}$-BRIL-Cterm/insert

Affinity chromatography is used for the purification of specific biomolecules from complex protein mixtures. The principle is based on the specific and reversible absorption of a molecule (adsorbent) to a matrix-bound ligand. This highly affine binding partner is covalently immobilized on a suitable matrix. Due to the specific interactions of the ligand with the target molecule it can be selectively enriched out of the complex mixture. The elution of the target protein can for example take place by competitive replacement or change of the pH value respectively (Burgess, Deutscher, 2009).

In this work Ni-NTA (nickel-nitrilotriacetic acid) agarose matrix (Invitrogen) was used for the purification of the POT transportes with the His-tag because Histidine shows a strong interaction with the immobilized nickel ions by forming coordination bonds with the immobilized transition metal. With the addition of free imidazole to the buffer elution is done (Bornhorst, Falke, 2000).

The following purification and cleavage was done for the proteins at 4 °C.

2 ml of resin were prepared and equilibrated with 5 ml wash buffer I. The beads were incubated with the supernatant for 30 minutes on a rotating wheel and afterwards transferred to a column. The flow through was collected and again passed onto the column. The beads were washed twice with 10 ml of wash buffer I and then twice with wash buffer II. 8 ml of wash buffer were added and the resin was transferred to a Falcon tube. 2 ml of TEV (tobacco etch virus) protease was added and the solution was incubated overnight on a rotating wheel. The next day the resin and the protein solution were transferred to the column. The cleaved protein was eluted and the beads were washed twice with 4 ml wash buffer II. This flow through and the cleaved protein were pooled together. The His-tag was eluted with two times 5 ml of elution buffer.

2.4.3. Ultrafiltration

The ultrafiltration unit Spin-X® Ultrafiltration Concentrators (Corning) were used to concentrate the pooled elution fractions. According to the molecular weight of the proteins a MWCO (molecular weight cut off) of 100 kDa was used. The centrifugation was carried out at 4,000 rpm and 4 °C.

2.4.4. Gel filtration

Gel filtration (GF), also called size exclusion chromatography, allows the separation of proteins due to their size respectively their hydrodynamic radiuses. The stationary phase consists of a porous matrix with defined pore size which proteins with fitting size access. The principle of separation is based on different retention times of the molecules in the matrix pores. Big protein molecules which can't access the internal volume are eluted first while small protein molecules which diffuse into the pores elute later (Burgess, Deutscher, 2009).

A Superose™6 10/300 GL, a Superdex™200 10/300 GL and a Superdex™200 16/600 GL (GE Healthcare) column which were connected to an ÄKTA pure system (GE Healthcare) were used.

The chromatography parameters are shown in table 9.

Table 9: Parameters of the gel filtration runs. For each column the applied POT transporter is shown as well.

Parameter	**Gel filtration**		
Column	Superdex™200 10/300 GL	Superdex™200 16/600 GL	Superose™6 10/300 GL
Protein	CtPOT-His	$PepT_{So2}$-BRIL-insert/Cterm	His-CtPOT
Flow rate	0.5 ml/min	1 ml/min	0.75 ml/min
Column delta pressure limit	1.5 MPa	0.3 MPa	1.5 MPa
Wavelength	280 nm	280 nm	280 nm
Equilibration with	40 ml	130 ml	40 ml
Sample volume	1 ml	5 ml	1 ml
Elution	Isocratic, 1 CV	Isocratic, 1 CV	Isocratic, 1 CV
Fraction size	0.5 ml	1 ml	0.5 ml

The purification was done at 4 °C and the flow through was collected with the fraction collector in 96 deep-well plates.

2.5. Biochemical methods for protein and DNA characterization

2.5.1. DNA sequencing

Cycle sequencing is used to sequence DNA with a PCR-type reaction and it is analyzed by computer. For the PCR the template DNA, a single primer, DNA polymerase, dNTP's and ddNTP's (dideoxy terminator nucleotides), labeled with a fluorescent tag of different color, are needed. The difference to the above mentioned PCR (see 2.1 LIC and PCR) is that during polymerization ddNTP's are incorporated and therefore lead to chain termination (Clark, Pazdernik, 2012). The sample is then injected into a capillary and separated according to the size of the fragments in an electrical field. The fragments pass a detector and due to the fluorescence signal of the ddNTP's the sequence of the DNA can be read. The automated sequencer prints the sequence in a series of peaks, one for each nucleotide position (Brown, 2002).

The DNA samples were amplified by PCR (2.1.2 Amplification of DNA-fragment with appropriate primers), purified with the QIA-KIT (Qiagen) and sent for sequencing to the company Eurofins Genomics.

2.5.2. Determination of protein concentration

For determination of the protein concentration the absorption of the aromatic amino acids tyrosine and tryptophane at 280 nm was used. With the measured absorption, the protein specific molar extinction coefficient and the path length it is possible to calculate the protein concentration according to the L-B-law (Burgess, Deutscher, 2009).

The measurements were done with the ND-1000 spectrophotometer (NanoDrop Technologies) and all used extinction coefficients are listed in table 10.

Table 10: Used extinction coefficients for the different membrane transporter for the determination of protein concentration.

Protein	**Extinction coefficient [$M^{-1}cm^{-1}$]**
CtPOT	116,770
$PepT_{So2}$-BRIL-insert	119,750
$PeoT_{So2}$-BRIL-Cterm	119,750

2.5.3. Gel electrophoresis

Gel electrophoresis is a method to separate DNA and proteins according to their size in an electrical field (Lodish *et al.*, 2000). DNA fragments were mixed with 6X DNA Loading Dye (Thermo Scientific) and the GeneRuler™ 1 kb DNA Ladder (Thermo Scientific) was used as marker. The DNA fragments were separated on a 1 % agarose gel with ethidium bromide (Roth) for 40 minutes at 120 V. Ethidium bromide intercalates between the base pairs of the DNA and therefore increases the intrinsic fluorescence (Lodish *et al.*, 2000). The DNA was visualized under ultraviolet light in the Gel Doc™ 2000 system (BioRad).

The analysis of proteins and protein mixtures was done with SDS-PAGE (polyacrylamide gel electrophoresis) under denaturing conditions. For denaturation of the protein samples NuPAGE® Sample Reducing Agent (life technologies) was mixed with NuPAGE® LDS Sample Buffer (4X) (life technologies). The mixture was added to samples and they were incubated for a few minutes at 37 °C. The samples were loaded on a precast gel and run for 40 minutes at 200 V.

As protein markers the marker Roti®-Mark 10-150 and Roti®-Mark PRESTAINED (Carl Roth) were used. Depending on the chamber size protein samples between 10 µl and 15 µl were loaded. For the PAGE NuPAGE® 4-12 % Bis-Tris Gels (life technologies) were used with the NuPAGE® MES SDS Running Buffer (20X) in 1x concentration (life technologies). For the detection of protein bands the gel was stained with Instant*Blue* (Expedeon) and destained with water for one hour each.

2.5.4. Western Blot

For the chemiluminescent detection of specific proteins the protein mixture is separated first by SDS-PAGE. The proteins are transferred from the gel to a polyvinylidene fluoride membrane via an electrical field. In this work the semi-dry transfer method was used. Semi-dry means that the amount of buffer is limited (Gravel, Golaz, 1996).

The set-up of the semy-dry transfer is shown in the following figure.

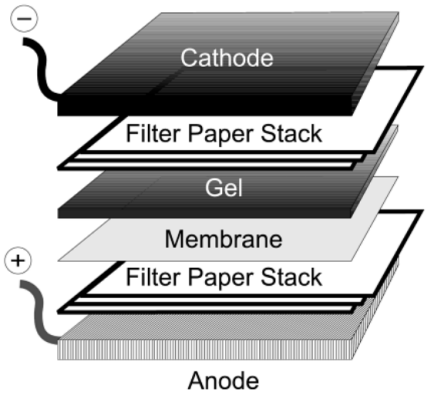


Figure 8: Schematic set-up of the Western Blot. On top of the membrane the gel is placed and both are embedded with filter paper stacks. The membrane is facing the anode and the gel the cathode *(Holtzhauer, 2006; Gravel, Golaz, 1996).*

The buffer-wetted filter papers are in direct contact with the electrodes and in between the filter stacks the gel and blotting membrane are placed (Gravel, Golaz, 1996).

The bound proteins can be detected with specific antibodies. Here a HisProbe-HRP (horse-radish peroxidase) was used which binds to the His-tag of the protein. After addition of substrate the immobilized HRP splits hydroperoxides which results in oxygen. The chemiluminescence of reacting oxygen with luminol can be detected (Holtzhauer, 2006; Thermo Scientific, 2011).

For the protein transfer from the SDS-gel to the membrane the Trans-Blot® Turbo™ Mini PVDF Transfer Pack (BioRad) was used. The transfer was done with the Trans-Blot® Turbo™ Blotting System (BioRad) at a constant electric current of 2.5 A for 10 minutes.

The following block, wash and incubation steps were done on a shaker. The membrane was blocked with 10 ml of BSA/TBST (25 mg/ml BSA (bovine serum albumin) in TBST (Tris buffered saline tween®-20 detergent) for one hour. With 15 ml TBST the membrane was washed three times for 5 minutes. 12.5 ml of HisProbe-HRP working solution were added and the membrane was incubated for one hour. Afterwards the membrane was washed again three times with 15 ml of TBST. 5 ml of SuperSignal West Pico Substrate Working solution was added and the membrane was incubated for 5 minutes. The membrane was exposed in the Gel Doc™ 2000 (BioRad) for 3 minutes and every 30 seconds a picture was taken.

2.6. Biophysical methods for protein characterization

2.6.1. Crystallization

X-ray crystallography is a technique to determine the three dimensional structure of proteins at atomic resolution. Therefore purified protein samples at high concentrations are crystallized (Smyth, Martin, 2000). The hurdle for membrane proteins is to shield the hydrophobic surfaces with detergents to allow extraction, purification and handling of these proteins away from their membrane environment. One major problem is still the growth of well-ordered membrane protein crystals. Often these crystals are tiny, extremely fragile, poorly ordered and very sensitive to radiation damage (Moraes *et al.*, 2014).

Obtained crystals are exposed to an x-ray beam generated from accelerated electrons in a synchrotron storage ring and resulting diffractions are collected by a detector. From the diffraction pattern it is possible to determine the crystal system and the unit cell dimensions. Spot intensities are used to determine the amplitude and the phase resulting in the determination of the structure factors. From the structure factors the position of the atoms in the crystal can be calculated using fast fourier transformation. Into this resulting electron density map the amino acid sequence is built (Smyth, Martin, 2000).

2.6.2. Protein crystallization

Crystallization is a phase transition phenomenon. The requirements for the structural characterization of proteins by x-ray diffraction analysis are crystals. Protein solutions are brought into a super saturation environment from which crystals grow. Crystal growth is achieved by variation of different parameters. Temperature, pH as well as the concentration of precipitant, protein and additives can be varied (see figure 9).

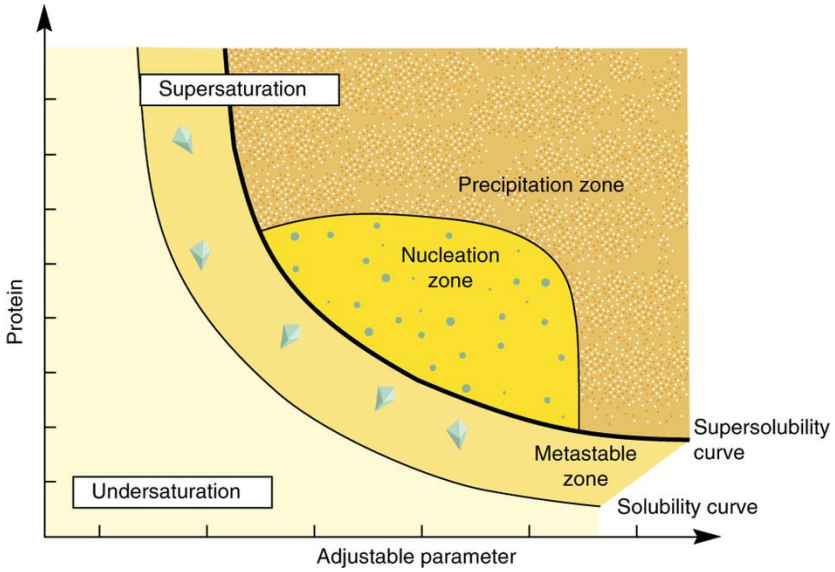


Figure 9: Phase diagram of protein crystallization with the precipitation, nucleation and metastable zone. The protein condition (solution or precipitation) is shown for the according location in the diagram. PH, temperature, etc. can be the adjustable parameter (Chayen, 2004).

Figure 9 shows four areas of different supersaturation degrees. In the area of high supersaturation the protein precipitates and in the moderate zone spontaneous nucleation takes place. In the metastable phase with low supersaturation crystals grow and are stable but no nucleation takes place. The fourth area is the zone of undersaturation in which proteins never crystallize (Chayen, 2004).

2.6.3. Crystallization screens

Crystallization conditions of a protein are not predictable and therefore different screens are tested to find initial conditions (McPherson, 1990).

For the crystallization screens the sitting drop method was used with Intelli-Plate 96 (catalog number: 102-0001-03, ARI). In the following figure the schematic setup is shown.

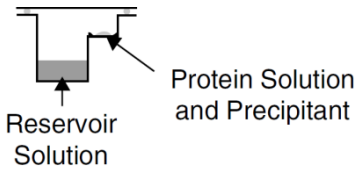


Figure 10: Schematic illustration of sitting drop method. The protein solution mixed with the precipitant are put seated above the reservoir solution on a plateau. A cover slide is used for sealing to create a closed system.

As shown in figure 10 the protein solution and the precipitants are mixed and placed on a plateau above the reservoir solution. The system is sealed with a cover slide which results in a closed system. The concentration gradient between drop and reservoir is slowly

compensated by gas diffusion. Therefore the precipitation concentration in the drop increases and possibly supersaturates (Gad, 2005).

A drop volume of 300 nl and a reservoir volume of 100 µl were used for crystallization.

For the determination of initial crystallization conditions the screens MemGold HT-96 (Molecular Dimensions) and JBScreen Membrane HTS (Jena Bioscience) were used for the proteins as well as the addition of peptides for $PepT_{So2}$-BRIL-insert and -Cterm for co-crystallization. Therefore 54 µl of protein were incubated with 6 µl of each peptide (c = 50 mM) for 30 minutes on ice. As peptides Met-Ser-Ala (Methionine-Serine-Alanine) and Tyr-Ala (Tyrosine-Alanine) were tested. To improve crystallization self-made optimization screens (Appendix: 5.6 Optimization screens for crystallization) were used. The reservoir solutions for the optimization plates were pipetted with the Scorpion (ARI) robot and the crystallization screens were set up with the help of pipetting robot Mosquito-LCP (ttplabtech). The crystallization plates were stored in the imager system RockImager-1000 (Formulatrix) at 19 °C. With the Rockmaker software the plates were imaged for crystal growth up to 80 days and the results were documented.

2.6.4. Dynamic light scattering

DLS (dynamic light scattering) is an optical method to characterize proteins in solution. Monochromatic light is sent through a sample and the fluctuations in the scattered light are measured (Daan *et al.*, 2013). These fluctuations occur due to Brownian motion of macromolecules in solution (Aivaliotis, 2003). An autocorrelation function is used to obtain from these data the hydrodynamic radii of the measured molecules (Narhi, 2013).

Membrane proteins are solubilized in aqueous solutions by detergents. Therefore the detergents replace the membrane lipids and form micelles in different geometries according to the size of protein and the exposed amino acids (Aivaliotis, 2003).

Because it is not clear if DLS can be used to differentiate between detergent bound IMPs and empty detergent micelles some preliminary experiments were done. Therefore different detergents were measured in buffer with and without 5 % glycerol at different concentrations ranging from 0 % to 4 % to follow micelle formation and to determine the micelle size. 20 µl of sample were pipetted and filtered (Ultrafree-MC, GV 0.22 µm, Millipore) for 3 minutes at 5,000 rpm.

To test the differentiation between detergent free and protein bound micelles, measurements were done with $PepT_{So2}$-BRIL-Cterm at a concentration of 2 mg/ml in buffer with 5 % glycerol. DDM was spiked in the sample with different concentrations ranging from 0.03 % to 4 %. Therefore the protein was centrifuged at 30,000 rpm and 4 °C for 30 minutes. The buffer and the detergent solution were filtered (Ultrafree-MC, GV 0.22 µm, Millipore) for 3 minutes at 5,000 rpm. For each DDM concentration 10 µl of sample were prepared in duplicate.

All measurements were done with the DynaPro® NanoStar™ (Wyatt) at 25 °C and per measurement 30 acquisitions were taken.

3. Results and discussion

3.1. Cloning and DNA sequencing of POTs

It was possible to amplify the genes of all constructs. After the T4 DNA polymerase treatment and cloning into the pNIC-CTHF and pNIC28-Bsa4 vector on all LB-Kan-plates with 5 % sucrose grew colonies. From these colonies plasmid preparations were done and the PCR products of the amplification with the T7 terminator and promoter were analyzed on a 1 % agarose gel. The results for the C- and N-terminal cloning of CtPOT and PepT2 are shown in the following figure.

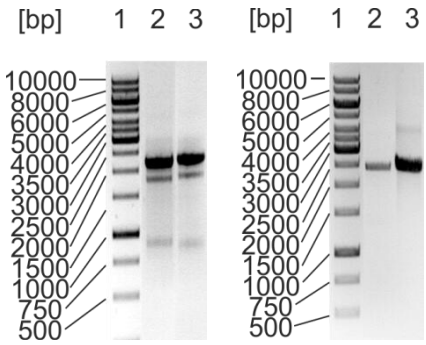


Figure 11: Gel picture of the amplified CtPOT and PepT2 with the T7 promoter and terminator which were cloned with C-terminal (left) and N-terminal (right) His-tag. The samples were purified and sent for sequencing. 1) Marker, 2) CtPOT, 3) PepT2.

The PCR products were purified and sent for sequencing. For DNA sequencing the T7 promoter and T7 terminator were each chosen. The results of DNA sequencing confirmed that all genes except for PepT1 were successfully cloned into the pNIC-CTHF, CtPOT and PepT2 as well into the pNIC28-Bsa4 vector. Either the original plasmid of PepT1 grew on the agar plates or only short fragments of the gene were cloned into the pNIC-CTHF or the pNIC28-Bsa4 vector. Alternatively a different expression vector could be tried with the use of restriction enzymes to create single stranded overhangs. The gene of interest is cut with the same restriction enzyme which is used for the linearization of the vector. The DNA fragments are mixed and attach to each other via base pairing. Ligation is done with the enzyme DNA ligase.

3.2. Small scale over-expression and membrane preparation of POTs

The constructs of CtPOT and PepT2 were expressed in 50 ml TB medium over night at 18 °C in different cell lines. The growth curves for the *E. coli* cells with the His-tagged PepT2 are shown in the following figure.

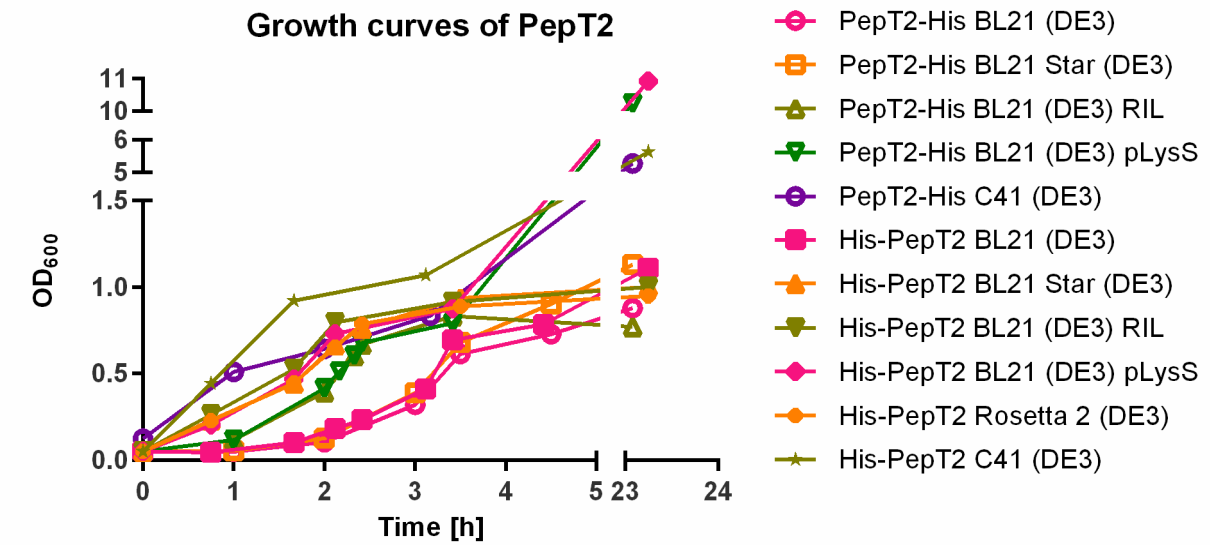


Figure 12: Growth curves of different E. coli cell lines containing either the C- or N-terminal His-tagged PepT2. In the graph the measured OD_{600} is shown against the time. The BL21 (DE3) pLysS cells grew best with an OD_{600} of about 10 to 11. Most cells grew to an end OD_{600} of 0.8 to 1.2.

The cells expressing the PepT2 transporter grew best in BL21 (DE3) pLysS. The lowest cell growth was observed in BL21 (DE3) RIL for the C-terminal His-tag transporter construct with an OD_{600} of about 0.8 and for the N-terminal construct with an OD_{600} of about 1.0 using Rosetta 2 (DE3) cells. This indicates that cell growth stopped after induction.

The growth curves of CtPOT in different *E. coli* strains are shown in the following figure.

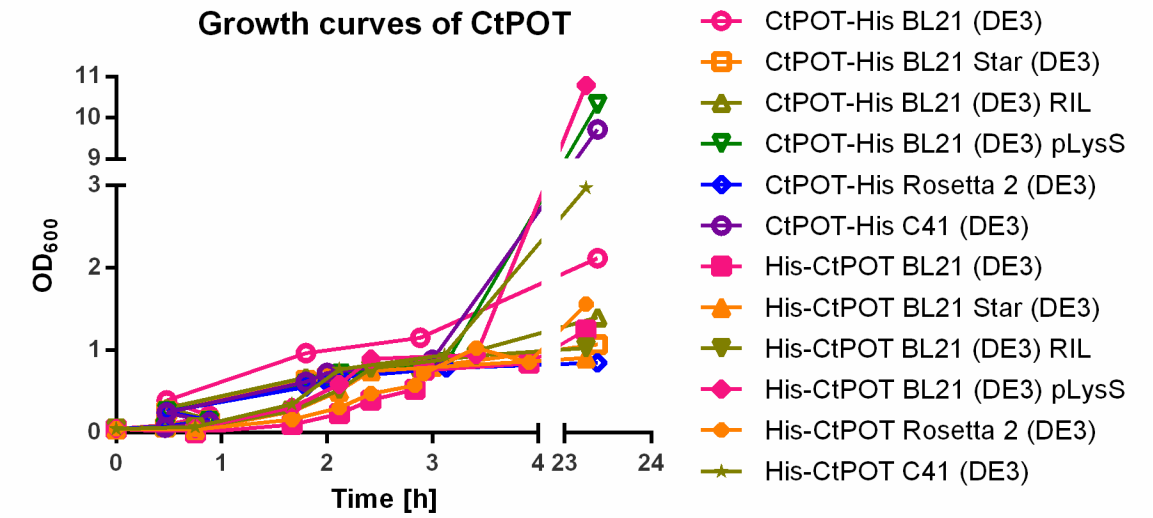


Figure 13: *Growth curves of CtPOT in different E. coli cells.* The measured OD_{600} is shown against time. The transformed plasmids into BL21 (DE3) pLysS cells grew best. Between the C41 (DE3) cells with the C- and N-terminal construct the biggest difference is seen in the end OD_{600} with 9.5 and 2.9.

The BL21 (DE3) pLysS cells with His-CtPOT grew best with a final OD_{600} of 10.8. The Rosetta 2 (DE3) strain with CtPOT-His grew worst. These cells only reached an OD_{600} of about 0.9, meaning that cell growth stopped after induction. The biggest difference in the growth curve is seen in the end OD_{600} of the C41 (DE3) cells with the C- and N-terminal constructs with 9.5 and 2.9.

The membranes of these cells were prepared and normalized according to the final cell density. The samples of the constructs with the C-terminal His-tag are shown on the gel and on the Western Blot in figure 14.

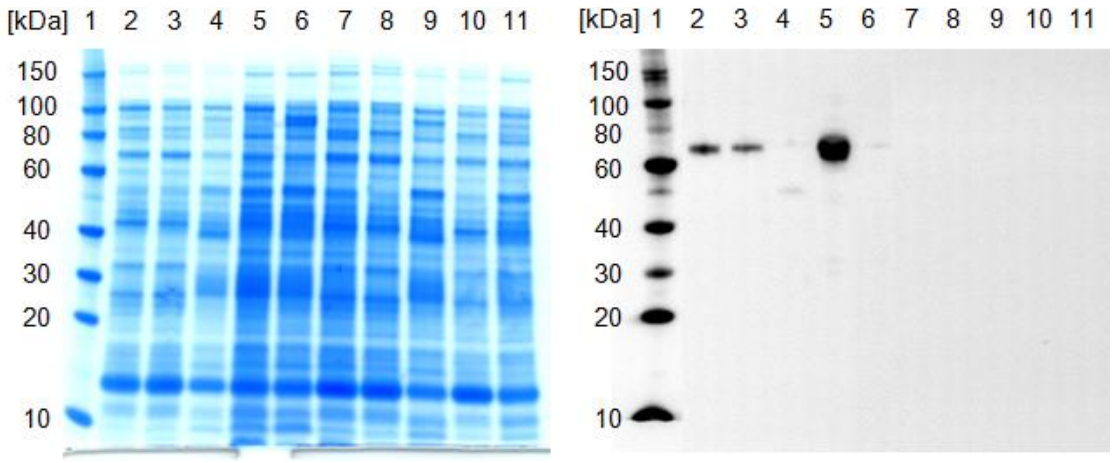


Figure 14: Prepared and normalized membrane fractions of the C-terminal CtPOT and PepT2 of different cell lines are shown. Left: SDS-Gel, right: Western Blot*. 1) Marker, 2) CtPOT in BL21 (DE3), 3) CtPOT in BL21 Star(DE3), 4) CtPOT in BL21 (DE3) pLysS, 5) CtPOT in BL21 (DE3) RIL, 6) CtPOT in C41(DE3), 7) PepT2 in BL21 (DE3), 8) PepT2 in BL21 Star (DE3), 9) PepT2 in BL21 (DE3) pLysS, 10) PepT2 in BL21 (DE3) RIL, 11) PepT2 in C41 (DE3). For PepT2 no expression is visible. CtPOT was expressed in BL21 (DE3), BL21 Star(DE3) and BL21 (DE3) RIL with the best expression.*

For PepT2-His no expression is detectable on the Western Blot (lane 7-11). The transporter was cloned as well into the pNIC28-Bsa4 vector to test the expression with a N-terminal His-tag. These prepared membranes are shown in figure 15.

CtPOT-His was expressed in BL21 (DE3) (lane 2), BL21 Star (DE3) (lane 3) and BL21 (DE3) RIL (lane 5). The best expression was achieved in the BL21 (DE3) RIL cells (lane 5) which is seen in the Western Blot. Therefore these cells were chosen for further cultivations.

The prepared membranes of the N-terminal His-tag constructs were analyzed on a SDS-Gel and Western Blot which are shown in the figure 15.

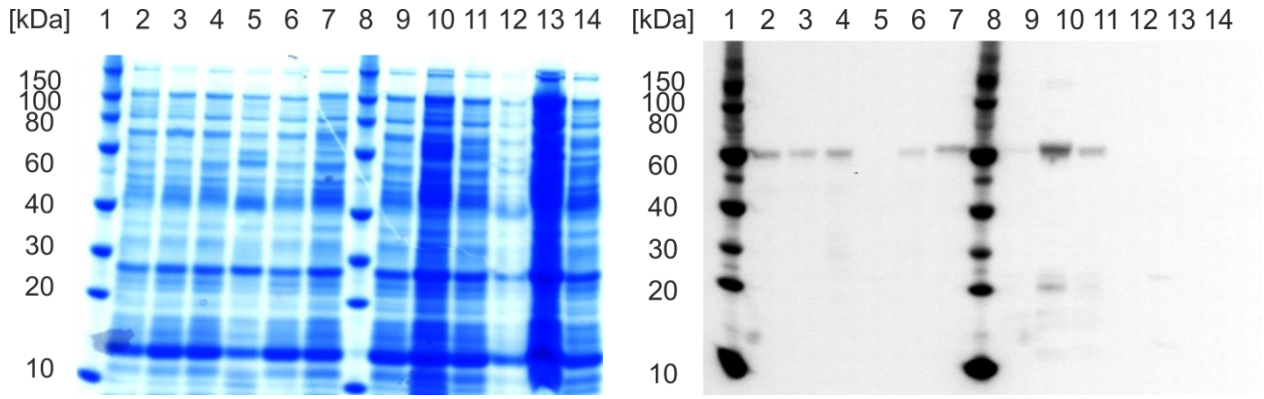


Figure 15: Normalized and prepared membrane fractions of CtPOT and PepT2 with N-terminal His-tag of different E. coli strains are shown. SDS-gel (left) and Western Blot (right). 1) Marker, 2) CtPOT in BL21 (DE3), 3) CtPOT in BL21 Star(DE3), 4) CtPOT in BL21 (DE3) pLysS, 5) CtPOT in BL21 (DE3) RIL, 6) CtPOT in Rosetta 2 (DE3), 7) CtPOT in C41 (DE3), 8) Marker, 9) PepT2 in BL21 (DE3), 10) PepT2 in BL21 Star(DE3), 11) PepT2 in BL21 (DE3) RIL, 12) PepT2 in BL21 (DE3) pLysS, 13) PepT2 in Rosetta (DE3), 14) PepT2 in C41 (DE3). CtPOT was expressed in all cells except in BL21 (DE3) RIL. PepT2 was expressed in BL21 Star (DE3) and BL21 (DE3) RIL.

His-CtPOT was expressed in all cell strains except in BL21 (DE3) pLysS (lane 5) where no band is visible in the Western Blot. Because the biomass obtained from C41 (DE3) cells was the highest and the protein was expressed at levels comparable to the other strains these cells were chosen for the further tests. His-PepT2 was expressed in BL21 (DE3) (lane 9), BL21 Star (DE3) (lane 10) and BL21 (DE3) RIL (lane 11). The lysis of the cells for His-PepT2 didn't work very well which is visible in the different intensities of the bands on the gel (figure 15). Therefore the bands on the Western Blot (figure 15) are not comparable for the over-expression of this protein.

The solubilization test was done for the expressed CtPOT constructs and for PepT2 with the N-terminal His-tag. Therefore different detergents were tried.

The test expression growth curves of the $PepT_{So2}$ constructs are shown in the following figure.

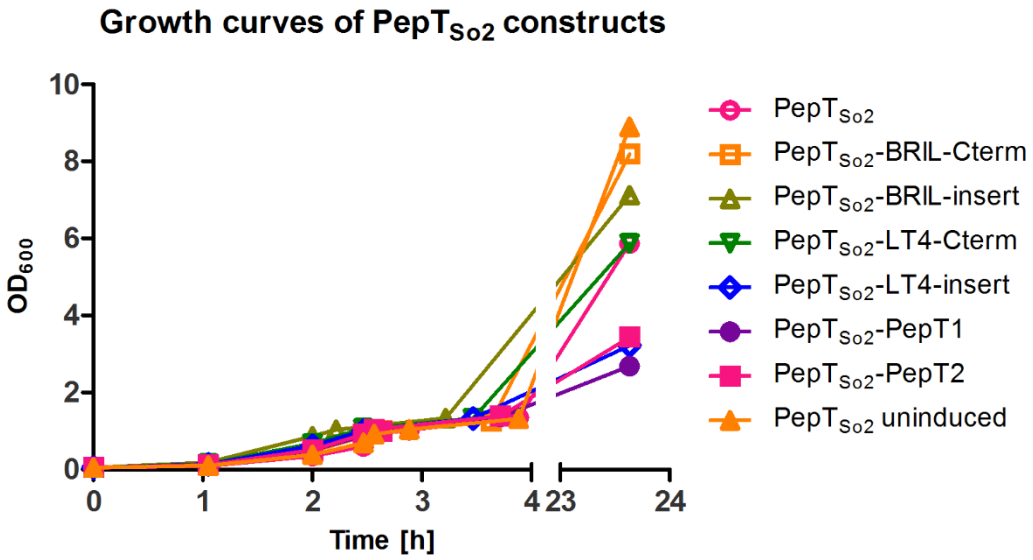


Figure 16: Growth curves of the $PepT_{So2}$ constructs in C41 (DE3) cells. Shown is the measured OD_{600} against the time. The constructs $PepT_{So2}$-BRIL-Cterm and $PepT_{So2}$-BRIL-insert grew best directly after the uninduced cells. The lowest final cell density was observed with $PepT_{So2}$-PepT1.

The C41 (DE3) cells of the not induced $PepT_{So2}$ cells reached the highest OD_{600} and therefore grew best. The reason is the existing causation of cell death due to over-expression of membrane proteins. The lowest cell concentration was reached for the $PepT_{So2}$-LT4-insert construct. For the analysis of the over-expression the membranes of all constructs were prepared and normalized according to the cell density. The results are shown in figure 17.

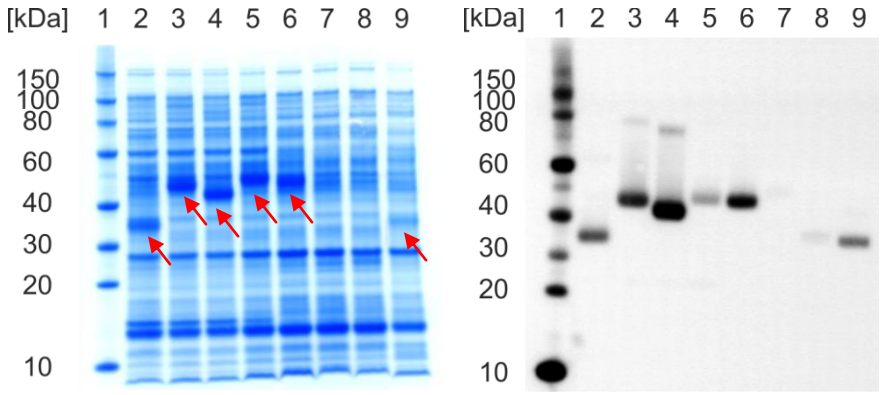


Figure 17: Membrane fractions of different POT transporter normalized according to cell density after test expression. SDS-Page (left) and Western Blot *(right) are shown. 1) Marker, 2) $PepT_{So2}$, 3) $PepT_{So2}$-BRIL-Cterm, 4) $PepT_{So2}$-BRIL-insert, 5) $PepT_{So2}$-LT4-Cterm, 6) $PepT_{So2}$-LT4-insert, 7) $PepT_{So2}$-PepT1, 8) $PepT_{So2}$-PepT2, 9) $PepT_{So2}$ uninduced. For all constructs expression in C41 (DE3) cells was observed except for $PepT_{So2}$-PepT1/PepT2. Best over-expression was achieved with $PepT_{So2}$-BRIL-insert.*

The $PepT_{So2}$ transporter was expressed in C41 (DE3) cells (lane 2) but also in the uninduced cells protein expression can be seen (lane 9). A reason for this could be a leakage in the promoter. Instead of C41 (DE3) cells, BL21 (DE3) pLysS could be used. In this strain the background expression level of the target protein is reduced. Though the cells grew anyway quite well, C41 (DE3) cells were further used in the scale up process. In comparison to $PepT_{So2}$, the proteins $PepT_{So2}$-BRIL-Cterm/-insert (lane 3/4) and $PepT_{So2}$-LT4-Cterm/insert (lane 5/6) are expressed quite well as seen on the SDS gel and Western Blot. All $PepT_{So2}$ constructs run at a lower molecular weight than expected. For example the$PepT_{So2}$-BRIL-insert (lane 4) has a MW (molecular weight) of 68.8 kDa but runs on the SDS-gel at a MW of about 45 kDa. One reason could be the different bound SDS ratio to the protein so that the protein is running faster. The correctness of the sequences was confirmed by DNA sequencing. For the $PepT_{So2}$-BRIL-Cterm, $PepT_{So2}$-BRIL-insert, $PepT_{So2}$-LT4-Cterm and $PepT_{So2}$-LT4-insert a solubilization test was done as well.

3.3. Solubilization test of POTs

To solubilize His-PepT2 different detergents at a final concentration of 1 % were tested. The result of the solubility test for human transporter is shown in the following figure in the Western Blot.

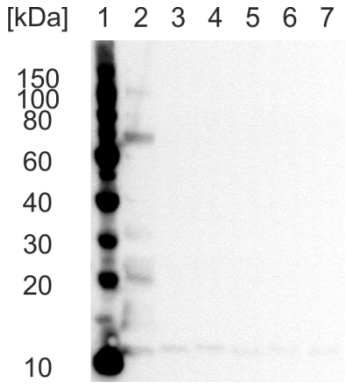


Figure 18: Membrane fractions after solubilization of His-PepT2 analyzed with Western Blot. 1) Marker, 2) His-PepT2 reference, 3) His-PepT2 with 1 % DDM, 4) His-PepT2 with 1 % DM, 5) His-PepT2 with 1 % DMNG, 6) His-PepT2 with 1 % LDAO, 7) His-PepT2 with 1 % $C_{12}E_8$. With the used detergents it was not possible to solubilize PepT2. Only the band of the reference without detergent is visible.

On the Western Blot a band is only visible for the reference (lane 1). Therefore His-PepT2 wasn't solubilized by any of the detergents and PepT2 was skipped for further tests. One reason can be that the used detergents are too mild. Stronger detergents could be tested but as they affect as well the native folding of the protein this was not tried. Another option

is the expression in eukaryotic cells which could be tried. These cells are more qualified for the expression of complex proteins as they can carry out complex post-translational modifications of these proteins which could be needed to stabilize the protein.

The Western Blots of the solubilized CtPOT constructs with the different detergents are shown in figure 19.

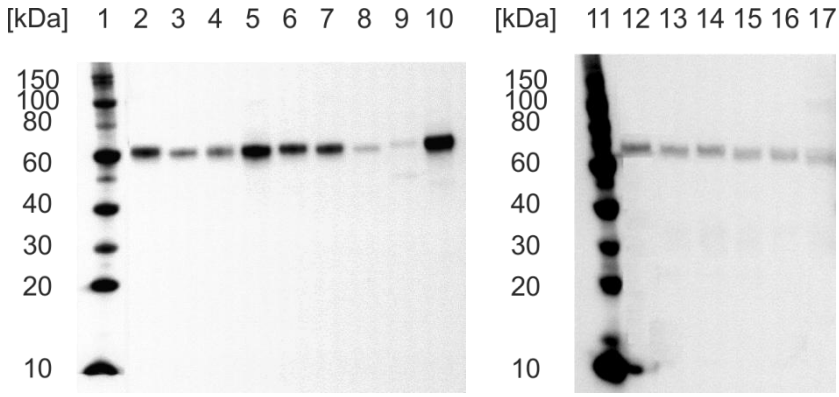


Figure 19: Membrane fractions of CtPOT-His (left) and His-CtPOT (right) solubilized with different detergents. Membrane fractions were analyzed with Western Blot. 1) Marker, 2) CtPOT-His with 1 % DDM, 3) CtPOT-His with 1 % DMNG, 4) CtPOT-His with 1 % $C_{12}E_8$, 5) CtPOT-His with 1 % LDAO, 6) CtPOT-His with 1 % DM, 7) CtPOT-His with 1 % Triton, 8) CtPOT-His with 1 % OG, 9) CtPOT-His with 1 % NG, 10) CtPOT-His reference, 11) Marker, 12) His-CtPOT reference, 13) His-CtPOT with 1 % DDM, 14) His-CtPOT with 1 % DM, 15) His-CtPOT with 1 % DMNG, 16) His-CtPOT with 1 % LDAO, 17) His-CtPOT with 1 % $C_{12}E_8$. For CtPOT it was possible to solubilize it in all tested detergents.

As shown in figure 19 almost every detergent solubilized the transporter. OG (lane 8) and NG (lane 9) solubilized only a little of the CtPOT-His which is seen according to the weak band on the Western Blot. LDAO (lane 5) worked best to solubilize the protein in comparison to the reference (lane 10). But also DDM (lane 2) solubilized the protein. Due to the fact that LDAO is a very harsh detergent DDM was chosen for further purifications. His-CtPOT was well solubilized by all detergents which is seen in the right Western Blot. All bands are as strong as the reference which indicates that almost all protein was solubilized. For His-CtPOT DDM was chosen as well for further purifications.

For the proteins $PepT_{So2}$-BRIL/LT4-Cterm/-insert DDM and DMNG were chosen for the solubilization test and the results are shown in the following figure.

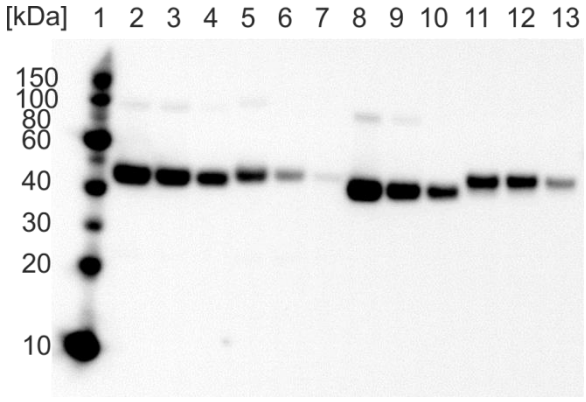


Figure 20: Membrane fractions of $PepT_{So2}$-BRIL/LT4-Cterm/-insert transporter solubilized with DDM and DMNG analyzed with Western Blot. 1) Marker, 2) $PepT_{So2}$-BRIL-Cterm reference, 3) $PepT_{So2}$-BRIL-Cterm with 1 % DDM, 4) $PepT_{So2}$-BRIL-Cterm with 1 % DMNG, 5) $PepT_{So2}$-LT4-Cterm reference, 6) $PepT_{So2}$-LT4-Cterm with 1 % DDM, 7) $PepT_{So2}$-LT4-Cterm with 1 % DMNG, 8) $PepT_{So2}$-BRIL-insert reference, 9) $PepT_{So2}$-BRIL-insert with 1 % DDM, 10) $PepT_{So2}$-BRIL-insert with 1 % DMNG, 11) $PepT_{So2}$-LT4-insert reference, 12) $PepT_{So2}$-LT4-insert with 1 % DDM, 13) $PepT_{So2}$-LT4-insert with 1 % DMNG. All constructs were solubilized in the tested detergents. The solubilization in DDM worked better than in DMNG which can be seen due to the stronger Western Blot signal.

Both detergents solubilized all proteins. But as seen on the gel DDM (lane 3/6/9/12) solubilized the proteins compared to the references (lane 2/5/8/11) best. Therefore 1 % DDM was chosen for the solubilization of these transporters. The proteins $PepT_{So2}$-BRIL-Cterm (lane 2-4) and $PepT_{So2}$-BRIL-insert (lane 8-11) were best expressed and solubilized. Therefore these two proteins were chosen for further purification.

3.4. Purification of CtPOT and PepT$_{So2}$-BRIL-Cterm/insert

The CtPOT was purified with Ni-NTA affinity chromatography. The result of the purification is shown in the following figure.

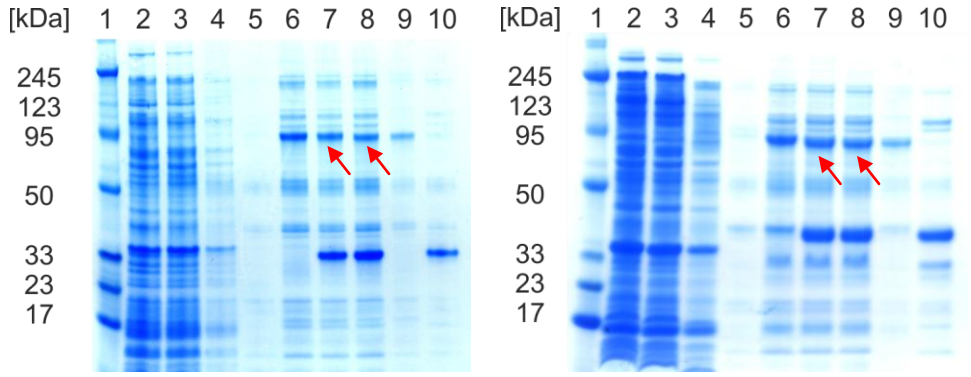


Figure 21: Affinity chromatography of CtPOT-His (left) and His-CtPOT (right). 1) Marker, 2) Supernatant after ultracentrifugation, 3) Flow through, 4) Wash 1, 5) Wash 2, 6) resuspended beads with bound protein solution, 7) resuspended beads with protein solution and TEV, 8) cleaved protein solution with beads and TEV, 9) CtPOT elution, 10) His-tag elution. In both purifications the protein is eluting from the column and the main band in the elution fraction is the protein band.

After the wash steps (lane 4 and 5) the protein solution was incubated over night with the TEV protease to cleave the His-tag. On the gel a slight different of the molecular weight of CtPOT is seen after cleavage (lane 7 and 8) which indicates that the His-tag was cleaved. After tag cleavage, the transporter elutes from the column which is seen in lane 9. In both purifications the main band in the elution fraction refers to the CtPOT. Also the purity of the transporter in these fractions is quite good because only some impurities are visible. The elution fraction was each concentrated to 1 ml and a gel filtration was performed to further purify the protein.

The gel filtration chromatogram of the cleaved C-terminal CtPOT with the SDS-gel is shown in figure 22.

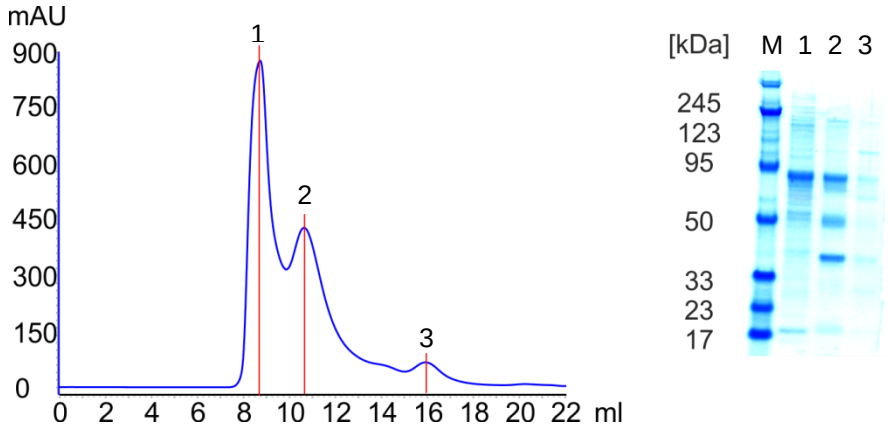


Figure 22: Gel filtration chromatogram (Superdex 200 10/300 GL) of CtPOT with the cleaved C-terminal His-tag and the according SDS-gel with the marked fractions in the chromatogram. M) Marker, 1) sample peak 1, 2) sample peak 2, 3) sample peak 3. In the void peak as well as in the second peak the target protein is eluting from the column. In the second peak some impurities are visible.

The elution profile of the CtPOT with the cleaved C-terminal His-tag shows a huge aggregation peak. The elution amount of the CtPOT-His is too low for further crystallization tests and the elution fraction contains still many impurities. Therefore no further tests were done with this construct.

For the N-terminal cleaved CtPOT the Superose 6 10/300 GL column was used because the resolution of this column for larger proteins is better than the Superdex 200 10/300 GL. The chromatogram and the SDS-gel are shown in the following figure.

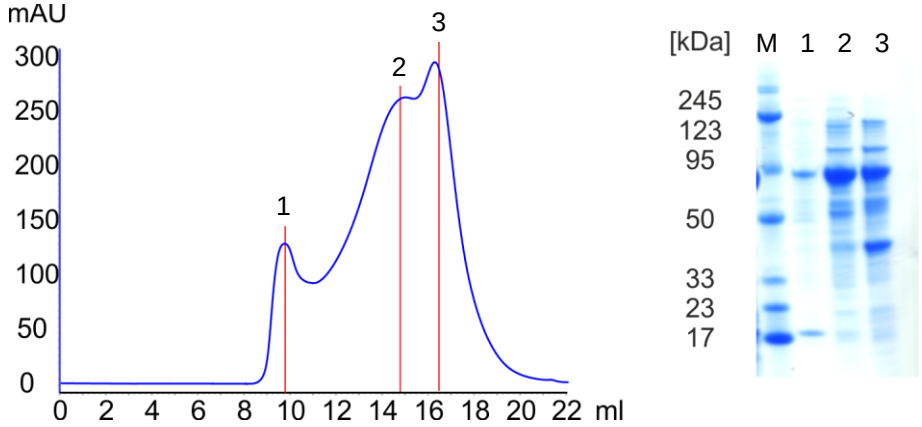


Figure 23: Gel filtration chromatogram (Superose 6 10/300 GL) of CtPOT with the cleaved N-terminal His-tag and the according SDS-gel with the marked fractions in the chromatogram. M) Marker, 1) sample peak 1, 2) sample peak 2, 3) sample peak 3.The protein is eluting in all three peaks. In the main peaks some impurities are still visible on the gel.

CtPOT is eluting from the column in three peaks. Fractions of these peaks were analyzed on the gel. In the second and third peak in this gel filtration run CtPOT is eluting and well

concentrated compared to the other column (figure 22). Also there are still a few impurities visible in the fractions of these peaks they were pooled (12.5 ml to 19 ml) and concentrated to about 14 mg/ml. One sample was directly used for crystallization and the rest was frozen at -80 °C.

To see whether the purity is increasing with a second gel filtration the third peak (16 ml to 19 ml) was pooled, concentrated and again loaded on the Superose 6 column. The chromatogram for this gel filtration is shown in the following figure.

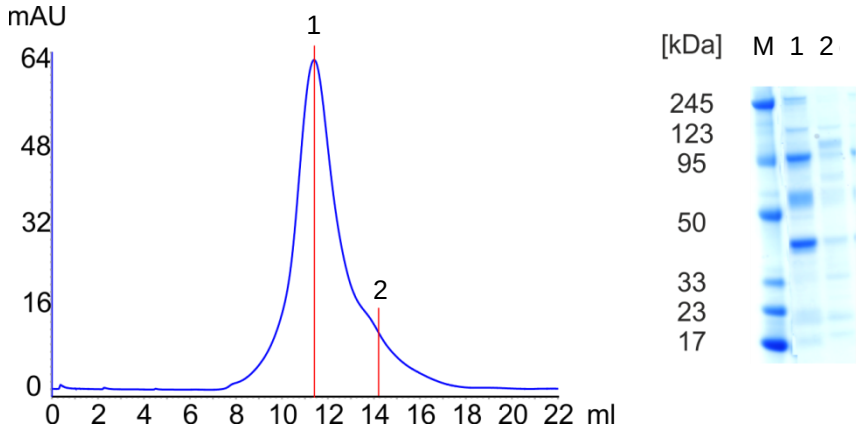


Figure 24: Second gel filtration chromatogram (Superose 6) of CtPOT with the cleaved N-terminal His-tag. On the right side the according gel with the fractions is shown. M) Marker, 1) sample peak 1, 2) sample peak 2. The protein is eluting in one main peak followed by a small peak. The same impurities compared to the first run are still visible.

The third peak of the first gel filtration (figure 23) elutes in the second gel filtration as one main peak followed by a small overlapped peak. The purity of the CtPOT is still not that good and compared to the first run the impurities and degradation products are increasing. The peak was concentrated to 6.36 mg/ml and the sample was used for crystallization.

The results of the Ni-NTA of PepT$_{So2}$-BRIL-insert/Cterm are shown in the following gel pictures.

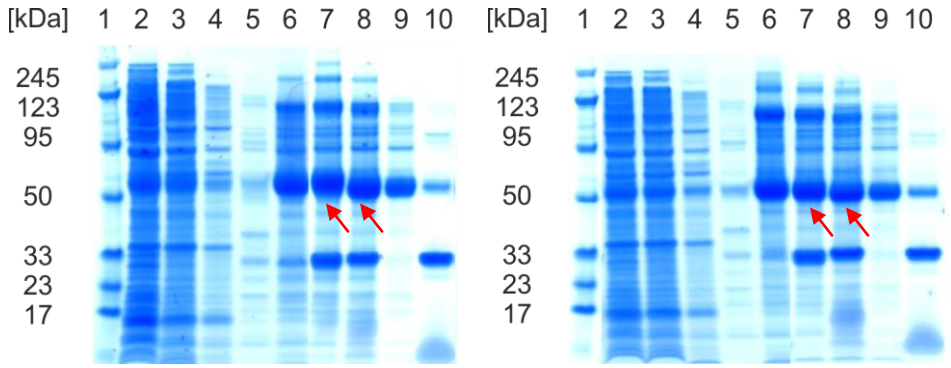


Figure 25: Affinity chromatography of PepT$_{So2}$-BRIL-Cterm (left) and PepT$_{So2}$-BRIL-insert (right). 1) Marker, 2) Supernatant after ultracentrifugation, 3) Flow through, 4)Wash 1, 5) Wash 2, 6) resuspended beads with protein solution, 7) resuspended beads with protein solution and TEV, 8) cleaved protein solution with beads and TEV, 9) PepT$_{So2}$ elution, 10) His-tag elution. Both cleaved proteins eluted from the column and in the elution fraction almost no impurities are visible.

After the cleavage over night with the TEV protease (lane 7 compared to lane 8) a slight difference in the molecular weight of both proteins is observed. This shift indicates that the cleavage of the His-tag was successful. In the elution fraction of the proteins (lane 9) almost no impurities are visible and the proteins are already well concentrated. These fractions were concentrated and loaded on a gel filtration column (Superdex 200 16/600 GL) for further purification. The chromatogram of PepT$_{So2}$-BRIL-insert is shown in the following figure 26 with the respective SDS-gel.

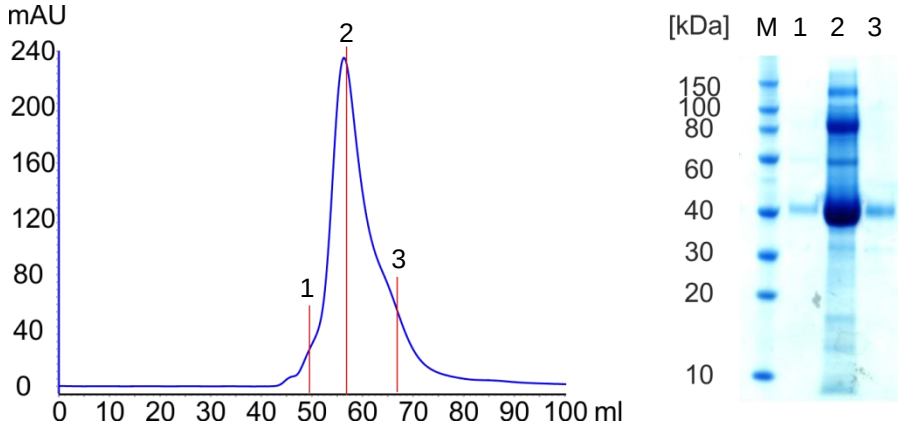


Figure 26: PepT$_{So2}$-BRIL-insert analyzed on a SDS-gel after gel filtration (Superdex 200 16/300) with the according chromatogram. M) Marker, 1) sample peak 1, 2) sample peak 2, 3) sample peak 3. The protein is eluting in a main peak preceded and followed by a small peak. In the sample of the main peak the protein is accumulated and only few impurities are visible. In the shoulders of the peak the transporter is present as well.

PepT$_{So2}$-BRIL-insert is eluting from the column in three overlaid peaks, one main peak and two smaller peaks, one right before and one right after the main peak. These three peaks

were analyzed on a SDS-gel. In all three fractions is the target protein present. The main peak was collected from 55 ml to 65 ml and concentrated to 12 mg/ml. One part of the protein sample was directly used for crystallization, the rest was flash frozen in liquid nitrogen and stored frozen at -80 °C.

The gel filtration chromatogram and the according gel picture of $PepT_{So2}$-BRIL-Cterm are shown in the following figure.

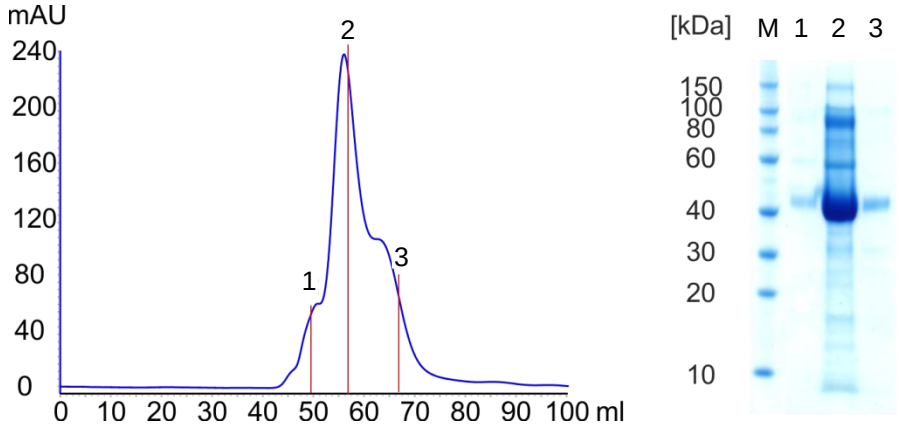


Figure 27: $PepT_{So2}$-BRIL-Cterm analyzed on a SDS-gel after gel filtration (Superdex 200 16/300) with the according chromatogram. M) Marker, 1) sample peak 1, 2) sample peak 2, 3) sample peak 3.The transporter is eluting in one main peak with one small peak in front and after the main peak. Compared to $PepT_{So2}$-BRIL-insert the small peaks are more pronounced. In all samples the transporter is present as seen on the gel. In the main peak the transporter is accumulated and only some impurities are visible.

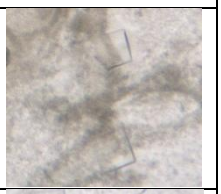
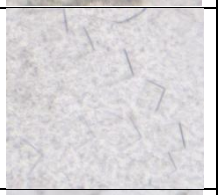
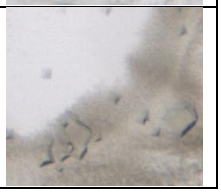
$PepT_{So2}$-BRIL-Cterm is eluting as well as $PepT_{So2}$-BRIL-insert in one main peak and two smaller peaks. The smaller peaks are compared to $PepT_{So2}$-BRIL-Cterm more pronounced. Fractions of the peaks were analyzed on a SDS-gel. In all three fractions the transporter is present. In the fraction of the top peak, the protein is well concentrated and only some impurities can be seen. Therefore the main peak was pooled from 55 ml to 65 ml and concentrated to 12 mg/ml and was directly used for crystallization. The rest of the protein sample was flash frozen in liquid nitrogen and stored frozen at -80 °C.

3.5. Crystallization of CtPOT and $PepT_{So2}$-BRIL-Cterm/insert

For the CtPOT small crystals were obtained with the initial MemGold HT-96 screen and JBScreen Membrane HTS. Due to time problems it was not possible to send them to the beamline for collecting data sets. Therefore it is not clear whether the crystals are protein or salt crystals. In case of protein crystals refinement screens would be necessary to optimize crystal quality.

For the protein $PepT_{So2}$-BRIL-insert little crystals were obtained. Because these crystals were quite small the condition for crystal growth was tried to optimize. All fished crystals and the conditions are shown in the following table.

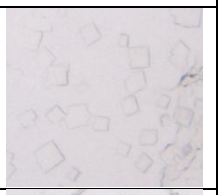
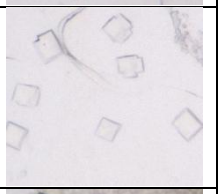
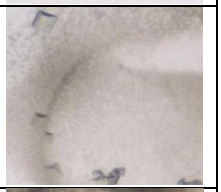
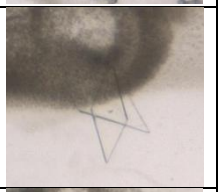
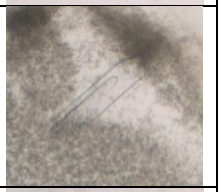
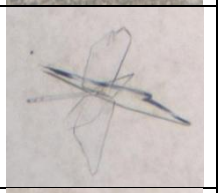
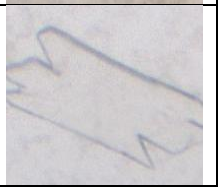
Table 11: Crystal conditions for $PepT_{So2}$-BRIL-insert and crystal pictures. In four conditions grew crystals for the transporter which were sent for collecting x-ray data sets.

Screen	Well	Concentration	Condition	Picture
MemGold-HT	D09	12.4 mg/ml	0.1 M MES, pH 7, 28.571 % PEG 400, 0.001 M cadmium chloride	
	F04	12.4 mg/ml	0.1 M MES, pH 6, 34.286 % PEG 400, 0.03 M magnesium chloride	
	F12	6.2 mg/ml	0.05 M sodium citrate, pH 4.5, 22 % PEG 400, 0.07 M sodium chloride	
	H11	12.4 mg/ml	0.1 M MES, pH 6.5, 30% PEG 400, 0.03 M magnesium chloride, 0.001 M cadmium chloride	

The crystals diffracted to 45 Å (well F04), 17 Å (well D09) and 18 Å (well F12) best. The crystal of well H11 showed no diffraction which indicates, that either the crystal was too small or that it is no protein crystal.

To improve the crystallization for $PepT_{So2}$-BRIL-insert the protein was incubated with peptides (Tyr-Ala, Met-Ala-Ser). All pictures and conditions for the fished crystals which were exposed to x-rays are shown in the following table.

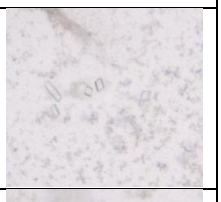
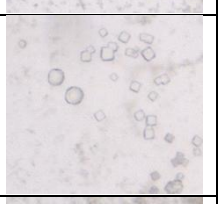
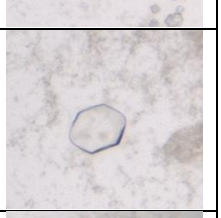
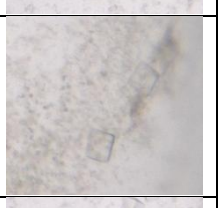
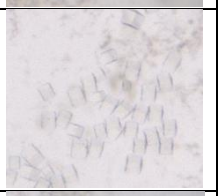
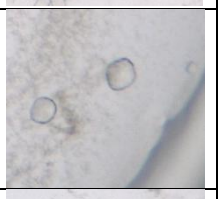
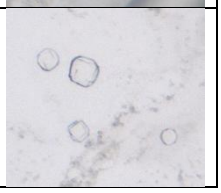
Table 12: Fished crystals for $PepT_{So2}$-BRIL-insert with peptides, their conditions and pictures. For crystallization with the tripeptide crystals grew in five conditions, for the dipeptide in three conditions. From these crystals x-ray data were collected.

Screen	Well	Concentrations	Condition	Picture
MemGold-HT	F08	12.4 mg/ml protein + 5 mM Met-Ala-Ser	0.1 M Tris, pH 8.5, 30 % PEG 2000 MME, 0.1 M sodium chloride, 0.005 M magnesium chloride	
	F08	12.4 mg/ml protein + 5 mM Tyr-Ala	0.1 M Tris, pH 8.5, 30 % PEG 2000 MME, 0.1 M sodium chloride, 0.005 M magnesium chloride	
	H11		0.1 M MES, pH 6.5, 30 % PEG 400, 0.03 M magnesium chloride, 0.001 M cadmium chloride	
Optimization screen 2	B04	12.57 mg/ml protein + 5 mM Met-Ala-Ser	0.1 M MES, pH 5.67, 35 % PEG 350 MME, 0.027 M magnesium chloride, 0.9 mM cadmium chloride	
	B05		0.1 M MES, pH 5.67, 25 % PEG 400	
	B06		0.1 M MES, pH 5.67, 28.333 % PEG 400, 0.2 M calcium chloride	
	A06		0.1 M MES, pH 5, 28.333 % PEG 400, 0.2 M calcium chloride	
	A02	12.57 mg/ml protein + 5 mM Tyr-Ala	0.1 M MES, pH 5, 35 % PEG 350 MME, 0.2 M calcium chloride	

The $PepT_{So2}$-BRIL-insert crystals with the tripeptide Met-Ala-Ser of well F08 didn't diffract. The crystals with the dipeptide Tyr-Ala of well F08 diffracted to around 19 Å and of H11 to 30 Å, respectively.

For $PepT_{So2}$-BRIL-Cterm no crystals were obtained with the MemGold-HT screen and JBScreen Membrane HTS. The obtained crystals with the peptides in the MemGold-HT and optimization screen which were analyzed are shown in the following table.

Table 13: $PepT_{So2}$-BRIL-Cterm-peptide crystals with the according conditions and crystal pictures. In one condition for the dipeptide crystals were fished. For the tripeptide crystals were frozen from six conditions.

Screen	Well	Concentrations	Condition	Picture
MemGold-HT	H07	12 mg/ml protein + 5 mM Tyr-Ala	0.1 M MES, pH 6.5, 33 % PEG 400, 0.2 M calcium chloride	
	C03	12 mg/ml protein + 5 mM Met-Ala-Ser	0.04 M Tris, pH 8.0, 27 % PEG 350 MME, 0.04 M sodium chloride	
	E12		0.05 M glycine, pH 9.5, 33 % PEG 300, 0.1 M sodium chloride	
	F04		0.2 M sodium chloride, 0.1 M HEPES, pH 7.0, 22 % PEG 550 MME	
	H01		0.1 M sodium dihydrogen phosphate-disodium hydrogen phosphate, pH 7.0, 33 % PEG 300, 0.1 M sodium chloride	
Optimization screen 1	D05	12 mg/ml protein + 5 mM Met-Ala-Ser	0.1 M HEPES, pH 7.14, 25 % PEG 400	
	E01		0.1 M HEPES, pH 7.36, 25 % PEG 350 MME	

The small crystals of $PepT_{So2}$-BRIL-Cterm with the dipeptide Tyr-Ala showed no diffraction at all as well as the crystals with the tripeptide of well C03 and H01. The initial crystals of the MemGold-HT screen of well E12 and F4 had diffraction limits to around 18 Å.

The crystal of well E01 showed hardly any diffraction and the crystal of well D05 of the optimization plate diffracted to very low resolution.

The diffraction images of the crystals $PepT_{So2}$-BRIL-insert with 5 mM Tyr-Ala (well F08) and $PepT_{So2}$-BRIL-Cterm with 5 mM Met-Ala-Ser (well E12) are shown in figure 28.

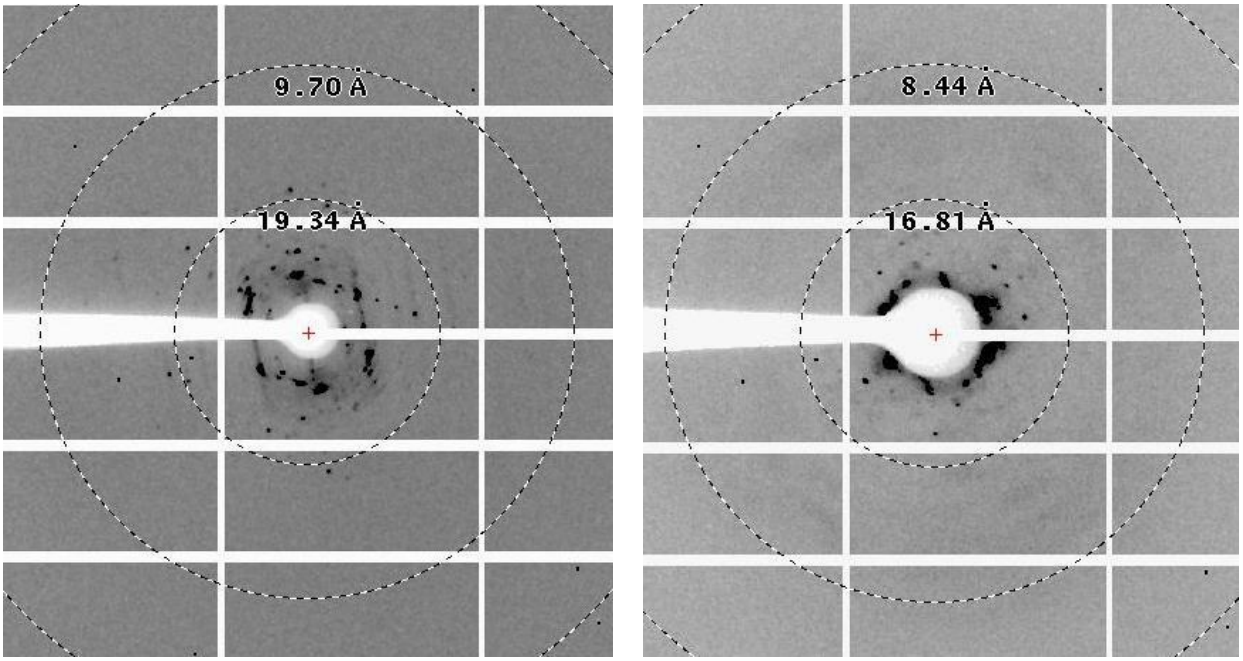


Figure 28: Diffraction images of $PepT_{So2}$-BRIL-insert with 5 mM Tyr-Ala (left) and $PepT_{So2}$-BRIL-Cterm with 5 mM Met-Ala-Ser (right). Both crystals showed only poor diffraction. The spectrum is free from ice rings.

The diffraction images are without ice rings, meaning that the crystals were well frozen. In both images the spots are separated but the crystals showed only limited diffraction. Therefore no structure model could be calculated with the current data. To improve the crystal quality crystal growth could be tried at 4 °C to slow down the crystal growth rate. For diffracting crystals refinements screens of the initial crystallization hit could be tried to improve crystal packing. As mentioned in the paper Guettou *et al.* (2013) also several rounds of optimization were necessary to get crystals diffracting beyond 4 Å. Another option could be the purity of the proteins. In the paper Guettou *et al. (*2013) the protein is also just purified with Ni-NTA affinity chromatography and gel filtration. Also the transporter are well concentrated after the gel filtration some impurities are still visible. To increase the purity of the transporter and therefore improve the crystal quality another purification step

could help. For example an anion- or cation exchange chromatography step could be included between the affinity chromatography and the gel filtration.

3.6. DLS

For the analysis of the DLS results only the acquisitions were taken in account where the measured values of the normalized intensity auto-correlation curve had a maximum deviation of ±0.003 from the theoretical baseline value of 1.000.

For the preliminary test detergents were measured in buffer with and without 5 % glycerol at different concentrations ranging from 0 % to 4 %. The MW and CMC of these detergents found in literature are listed in the following table.

Table 14: CMC and MW values found in literature for the different tested detergents.

Detergent	**MW [kDa]**	**Reference MW**	**CMC [%]**	**Reference CMC**
DDM	72	Strop, Brunger, 2005	0.0087	VanAken *et al.*, 1986
DM	33	calculated	0.087	Alpes *et al*., 1988
DMNG	-	-	0.0034	Chae *et al*., 2010
LDAO	21.5	Strop, Brunger, 2005	0.023	Herrmann, 1962
OG	25	Sigma Aldrich, 2015	0.53	Lorber *et al.*, 1990

The measured molecular weights at the different concentrations are shown in the following figure.

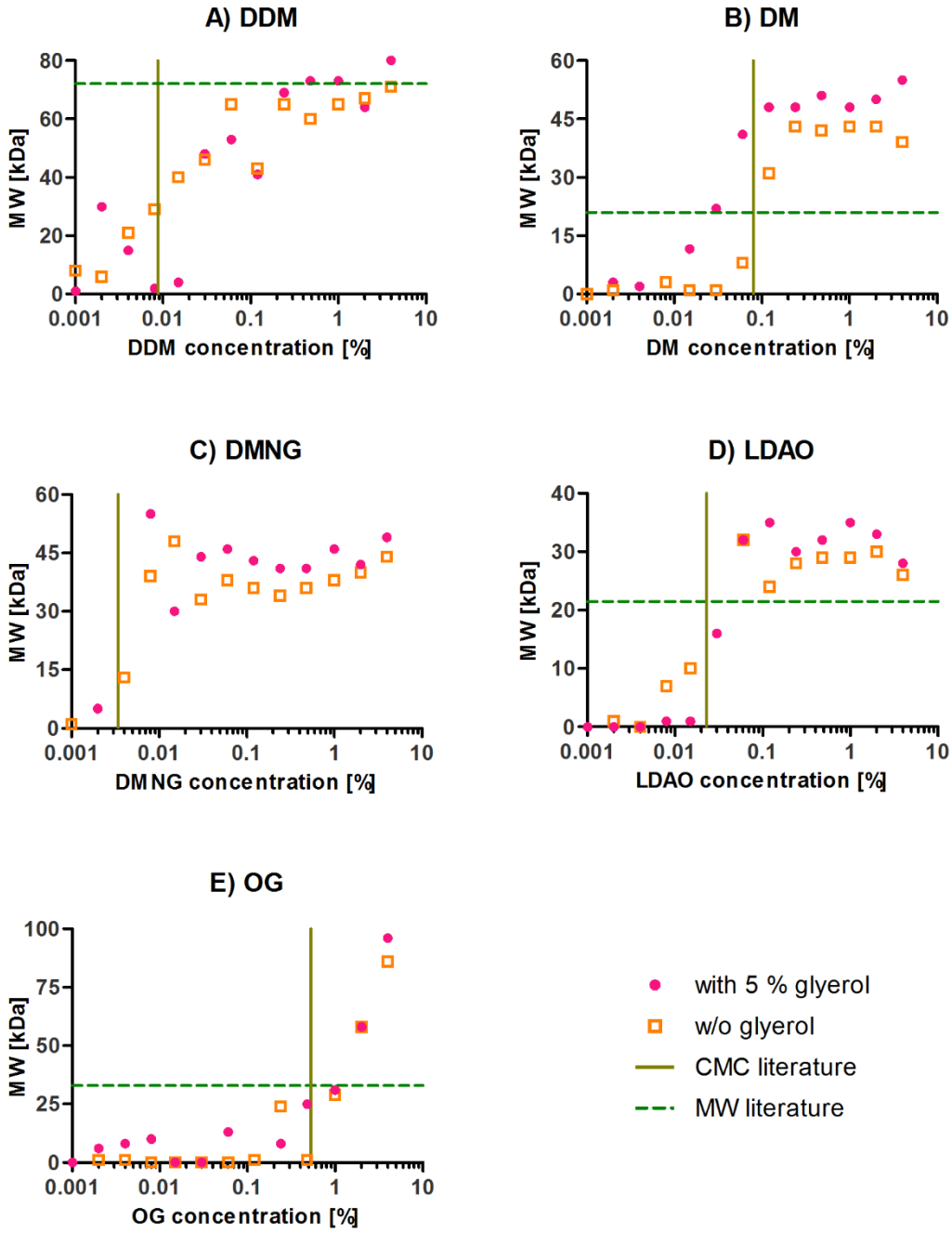


Figure 29: DLS results of the different measured detergents. Shown are the measured molecular weights of the detergent micelles in buffer with glycerol in closed circles and in buffer without glycerol in open squares against the detergent concentration. The CMC (green line) and MW (green dashed line) of detergent micelles which are found in literature are shown as well. A) DDM, B) DM, C) DMNG, D) LDAO, E) OG.

An increase in micelle size is seen for all detergents with increasing detergent concentration. Furthermore it seems that glycerol has an impact on micelle formation and micelle MW. In the graph for DM there is a trend seen that with glycerol the MW of the micelles is higher than in buffer without glycerol. Also the formation of micelles occurs at a lower concentration in buffer with glycerol, meaning, that glycerol affects the CMC of detergents. For DMNG and for LDAO a higher MW of micelles is seen in buffer with glycerol compared to the measurements in buffer without glycerol. For DDM which is often used for protein purification the micelle size is ranging from 60 kDa to 80 kDa. These data are similar to those found in literature with 72 kDa. The CMC for DDM is at a concentration of 0.0087 % (in water) but these micelle sizes occur at a concentration of 0.24 %. For DM the measured micelle MWs are in the range of 35-50 kDa. The MW which is found in literature is 33 kDa which is lower than the measured micelle molecular weight but it is still very similar. Micelles are formed at a concentration of 0.12 %, according to literature micelles can form at 0.087 % (in water). In literature only the CMC (critical micelle concentration) is found for DMNG. For the molecular weight no data are available. With DLS a micelle MW of around 40 kDa was measured. For LDAO a MW of around 30 kDa was measured which is higher than the MW found in literature with 21.5 kDa. OG has in literature a low MW of 33 kDa but a high CMC with 0.53% (in water). With DLS an increasing micelle MW was measured with increasing detergent concentrations. This was seen for the detergent measurements with and without glycerol. Theoretically there should have been no particles in the solution beyond the CMC concentration. Because either micelles have formed and are measurable in the solution or there are none. But especially for DDM it seems that there are intermediate micelles which can be measured.

To see whether the DLS device can differentiate between empty detergent micelles and IMPs with bound detergent 2 mg/ml of $PepT_{So2}$-BRIL-Cterm were spiked with different concentrations of DDM ranging from 0.03 % to 4 %. The results are shown in the following figure 30.

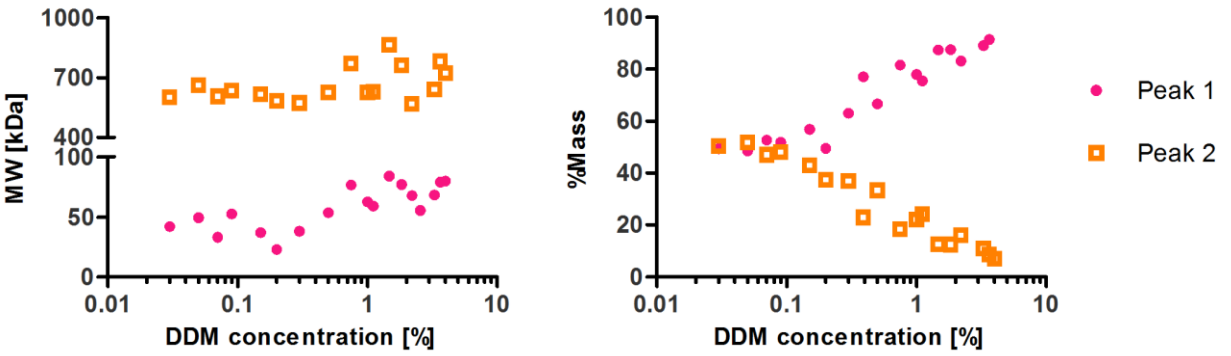


Figure 30: Measured protein solution of 2 mg/ml $PepT_{So2}$-BRIL-Cterm with spiked DDM in different concentrations. In the graphs are shown the two measured peaks of the histogram with the MWs (left) and the according %mass (right). Peak 1 refers to the detergent micelle and is shown in closed circles and the protein peak (peak 2) is shown in open squares. The MW of the detergent micelle is increasing with increasing DDM concentration and reaches at 4 % DDM a MW of about 80 kDa. The transporter mass ranges from about 550 kDa at 0.03 % DDM to 700 kDa at 4 % DDM. In the graph for the %mass the transporter %mass is decreasing with increasing DDM concentration whereas the detergent %mass is increasing.

In the left graph it can be seen that the DLS devise is capable to differentiate between two peaks. Peak 1 refers to the free detergent micelle and peak 2 to the protein bound to DDM. The MW of the IMP with the bound DDM complex ranges from 550 kDa to 700 kDa. The MW of the detergent micelle is increasing with increasing DDM concentration from around 45 kDa to 80 kDa at 4 % DDM. In these measurement again some intermediate MW of micelles can be measured which were also measured before in the detergent test for DDM. For the percentage mass of the peaks a decrease for the protein-detergent complex is seen with increasing DDM concentration from around 50 % mass in the beginning to around 10 % mass at 4 % DDM. The percentage mass of the micelles is increasing with increasing DDM concentration ranging from 50 % in the beginning to 90 % at 4 % DDM in the end.

With DLS it is therefore possible to differentiate between empty detergent micelles and IMPs bound to detergent. Due to measurement errors which can occur (e. g. pipetting errors, different handling of samples) it would be recommended to have a difference in the MW of the detergent micelles and the protein-detergent complex so that the error margins do not overlap. To use the DLS as a quality control for crystallization further tests would be necessary. First of all different transporter should be measured and also spiked with DDM.

If these proteins behave in a similar way it would be good to set up plates for crystallization with different percentage mass of detergent and MW of micelle sizes and evaluate crystal growth. The last step would be the determination of detergent percentage mass limits and micelle MWs for optimal crystallization.

4. Conclusion and outlook

For the development of new active pharmaceutical ingredients it is helpful to know the structure of the therapeutic relevant proteins and to get insights into the binding mechanism of ligands. Therefore the two human POTs, the eukaryotic CtPOT of *C. thermophilum* and different constructs of the bacterial $PepT_{So2}$ of *S. oneidensis* were studied in this master thesis. These POTs were successfully cloned into pET expression vectors, except for PepT1.

The human transporter PepT2 was expressed in BL21 Star(DE3) cells but it was not possible to solubilize it with the tested detergents. As this transporter is of great pharmacological interest and it was shown that it is possible to clone and express the vector it could be tried to express PepT2 in HEK cells and further characterize it.

CtPOT and the $PepT_{So2}$ constructs (except $PepT_{So2}$-PepT1/T2) were expressed in C41 (DE3) cells and solubilized with DDM. Purification was done with Ni-NTA affinity chromatography and His-tag cleavage on the beads followed by gel filtration.

For the CtPOT it was possible to get crystals in the initial screens but due to time problems it was not possible to collect x-ray data sets. Concerning the fact that so far only the structures of bacterial POTs are available it is worthwhile to keep on working on this transporter as the CtPOT as an eukaryotic transporter is more related to the human POTs than the bacterial transporters. Therefore this structure could give more precise information regarding human POTs.

Two $PepT_{So2}$ constructs, $PepT_{So2}$-BRIL-Cterm and -insert, were crystallized. For $PepT_{So2}$-BRIL-Cterm crystals were only obtained in co-crystallization with peptides. The other construct, $PepT_{So2}$-BRIL-insert, was crystallized with and without peptides. Due to limited diffraction of these crystals it was not possible to solve any structure. To get a better resolution of these crystals more refinement screens should be done. As it is shown that it is possible to insert proteins into the loop of this transporter one approach would be the insertion of dimer-forming proteins. Inserting such a protein into two loops on the cytoplasm

site of the $PepT_{So2}$ transporter possibly traps the transporter in the outward open conformation. So far no structure in this conformation is available but it is of great pharmacological interest as it gives information about the peptide binding to the POT transporter.

To improve crystallization dynamic light scattering could be implemented as quality control for the used protein samples. With DLS it is possible to differentiate between empty detergent micelles and protein-detergent complexes. It gives information about the MW as well as of the percentage mass of empty detergent micelles and protein-detergent complexes. These information are important for crystallization as it is unwanted to have a high amount of empty detergent micelles in the sample as they can disturb protein crystallization. Therefore DLS can be used to optimize protein crystallization.

5. Appendix

5.1. References

Aivaliotis, M., Samolis, P., Neofotistou, E., Remigy, H., Rizos, A. K., & Tsiotis, G. (2003). Molecular size determination of a membrane protein in surfactants by light scattering. Biochimica Et Biophysica Acta (BBA)-Biomembranes, 1615(1), 69-76.

Alpes, H., Apell, H., Knoll, G., Plattner, H., & Riek, R. (1988). Reconstitution of na /K -ATPase into phosphatidylcholine vesicles by dialysis of nonionic alkyl maltoside detergents. Biochimica Et Biophysica Acta (BBA)-Biomembranes, 946(2), 379-388.

Aslanidis, C., & de Jong, P. J. (1990). Ligation-independent cloning of PCR products (LIC-PCR). Nucleic Acids Research, 18(20), 6069-6074.

Bornhorst, J. A. & Falke, J. J. (2000,). Purification of Proteins Using Polyhistidine Affinity Tags. Methods in Enzymology, 326, 245-254.

Brown, T. A. (2002). Genomes (2nd edition) Wiley-Liss.

Canene-Adams, K. (2013). General PCR. Methods in Enzymology, 529, 291-298.

Casali, N., & Preston, A. (2003). *E. coli* plasmid vectors methods and applications (Volume 235) Humana Press.

Chae, P. S., Rasmussen, S. G., Rana, R. R., Gotfryd, K., Chandra, R., Goren, M. A., Pierre, Y. (2010). Maltose-neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins. Nature Methods, 7(12), 1003-1008.

Chayen, N. E. (2004). Turning protein crystallisation from an art into a science. Current Opinion in Structural Biology, 14(5), 577-583.

Chen, B., & Janes, H. W. (2002). PCR cloning protocols Springer.

Clark, D. P., & Pazdernik, N. J. (2010). Biotechnology (Academic Cell Update Edition) Elsevier.

Crommelin, D. J., Sindelar, R. D., & Meibohm, B. (2013). Pharmaceutical biotechnology fundamentals and applications (4th edition) Springer.

Dumon-Seignovert, L., Cariot, G., & Vuillard, L. (2004). The toxicity of recombinant proteins in *Escherichia coli*: A comparison of over-expression in BL21 (DE3), C41 (DE3), and C43 (DE3). Protein Expression and Purification, 37(1), 203-206.

Gad, S. C. (2005). Drug discovery handbook John Wiley & Sons.

Gravel, P., & Golaz, O. (1996). Protein blotting by the semidry method. The protein protocols handbook (pp. 249-260) Springer.

Griffiths, A. J., Miller, J. H., Suzuki, D. T., Lewontin, R. C., & Gelbart, W. M. (2004). An introduction to genetic analysis (8th edition) Palgrave Macmillan.

Guettou, F., Quistgaard, E. M., Raba, M., Moberg, P., Löw, C., & Nordlund, P. (2014). Selectivity mechanism of a bacterial homolog of the human drug-peptide transporters PepT1 and PepT2. Nature Structural & Molecular Biology, 21(8), 728-731.

Guettou, F., Quistgaard, E. M., Trésaugues, L., Moberg, P., Jegerschöld, C., Zhu, L., Löw, C. (2013). Structural insights into substrate recognition in proton-dependent oligopeptide transporters. EMBO Reports, 14(9), 804-810.

Han, F., Le, G. W., & Shi, Y. H. (2003). Molecular characteristics and tissue distribution of peptide transporters. Sheng Li Ke Xue Jin Zhan [Progress in Physiology], 34(3), 222-226.

Herrmann, K. (1962). Non-ionic—cationic micellar properties of dimethyldodecylamine oxide1. The Journal of Physical Chemistry, 66(2), 295-300.

Holtzhauer, M. (2006). Basic methods for the biochemical lab Springer.

Kato, T., Thompson, J. R., & Park, E. Y. (2013). Construction of new ligation-independent cloning vectors for the expression and purification of recombinant proteins in silkworms using BmNPV bacmid system. PloS One, 8(5), 64007.

Lodish, H., Berk, A., Zipursky, S. L., Matsudaira, P., Baltimore, D., & Darnell, J. (Eds.). (2000). Molecular cell biology (4th edition) WH Freeman.

Lorber, B., Bishop, J. B., & DeLucas, L. J. (1990). Purification of octyl β-D-glucopyranoside and re-estimation of its micellar size. Biochimica et Biophysica Acta (BBA)-Biomembranes, 1023(2), 254-265.

Luckey, M. (2014). Membrane structural biology: With biochemical and biophysical foundations Cambridge University Press.

McPHERSON, A. (1991). Review current approaches to macromolecular crystallization. EJB reviews 1990 (pp. 49-71) Springer.

Moraes, I., Evans, G., Sanchez-Weatherby, J., Newstead, S., & Stewart, P. D. S. (2014). Membrane protein structure determination - the next generation. Biochimica et Biophysica Acta (BBA)-Biomembranes, 1838(1), 78-87.

Narhi, L. O. (2013). Biophysics for therapeutic protein development (Volume 4) Springer.

Newstead, S. (2011). Towards a structural understanding of drug and peptide transport within the proton-dependent oligopeptide transporter (POT) family. Biochemical Society Transactions, 39(5), 1353-1358.

Newstead, S., Drew, D., Cameron, A. D., Postis, V. L., Xia, X., Fowler, P. W., McPherson, M. J. (2011). Crystal structure of a prokaryotic homologue of the mammalian oligopeptide–proton symporters, PepT1 and PepT2. The EMBO Journal, 30(2), 417-426.

Pelicic, V., Reyrat, J.-M., Gicquel, B. (1995). Expression of the *Bacillus subtilis sacB* Gene Confers Sucrose Sensitivity on Mycobacteria. Journal of Bacteriology, 178(4), 1197-1199.

Robinson, A. S. (Ed.). (2011). Production of membrane proteins: strategies for expression and isolation. John Wiley & Sons.

Rupp, B. (2010). Biomolecular crystallography Garland Science New York.

Seddon, A. M., Curnow, P., & Booth, P. J. (2004). Membrane proteins, lipids and detergents: Not just a soap opera. Biochimica Et Biophysica Acta (BBA)-Biomembranes, 1666(1), 105-117.

Sigma Aldrich (2015). Website: http://www.sigmaaldrich.com/catalog/product/sial/o9882?lang=de&region=DE. 12.03.2015, 15:03 pm.

Smyth, M. S., & Martin, J. H. (2000). X ray crystallography. Molecular Pathology: MP, 53(1), 8-14.

Solcan, N., Kwok, J., Fowler, P. W., Cameron, A. D., Drew, D., Iwata, S., & Newstead, S. (2012). Alternating access mechanism in the POT family of oligopeptide transporters. The EMBO Journal, 31(16), 3411-3421.

Strop, P., & Brunger, A. T. (2005). Refractive index-based determination of detergent concentration and its application to the study of membrane proteins. Protein Science, 14(8), 2207-2211.

VanAken, T., Foxall-VanAken, S., Castleman, S., & Ferguson-Miller, S. (1986). Alkyl glycoside detergents: Synthesis and applications to the study of membrane proteins. Methods in Enzymology, 125, 27-35.

Wagner, S., Baars, L., Ytterberg, A. J., Klussmeier, A., Wagner, C. S., Nord, O., de Gier, J. W. (2007). Consequences of membrane protein over-expression in *Escherichia coli*. Molecular & Cellular Proteomics: MCP, 6(9), 1527-1550.

Wagner, S., Klepsch, M. M., Schlegel, S., Appel, A., Draheim, R., Tarry, M., de Gier, J. W. (2008). Tuning *Escherichia coli* for membrane protein over-expression. Proceedings of the National Academy of Sciences of the United States of America, 105(38), 14371-14376.

Warren, T. D., Coolbaugh, M. J., & Wood, D. W. (2013). Ligation-independent cloning and self-cleaving intein as a tool for high-throughput protein purification. Protein Expression and Purification, 91(2), 169-174.

Zhao, Y., Mao, G., Liu, M., Zhang, L., Wang, X., & Zhang, X. C. (2014). Crystal structure of the *E. coli* peptide transporter YbgH. Structure, 22(8), 1152-1160.

5.2. Table of tables

5.3. Table of figures

5.4. Nucleotide sequences and primers of used vectors and POTs

Table 15: Partial vector sequences of pNIC-CTHF and pNIC28-Bs4 with the SacB gene and His-tag.

pNIC-CTHF	ACTTAGTATTTGAAGCAAACACTGGAACTGAAGATGGCTACCAAGGCGAAGAATC TTTATTTAACAAAGCATACTATGGCAAAAGCACATCATTCTTCCGTCAAGAAAGTCA AAAACTTCTGCAAAGCGATAAAAAACGTACGGCTGAGTTAGCAAACGGCGCTCTC GGTATGATTGAGCTAAACGATGATTACACACTGAAAAAAGTGATGAAACCGCTGAT TGCATCTAACACAGTAACAGATGAAATTGAACGCGCGAACGTCTTTAAAATGAACG GCAAATGGTACCTGTTCACTGACTCCCGCGGATCAAAAATGACGATTGACGGCAT TACGTCTAACGATATTTACATGCTTGGTTATGTTTCTAATTCTTTAACTGGCCCATA CAAGCCGCTGAACAAAACTGGCCTTGTGTTAAAAATGGATCTTGATCCTAACGATG TAACCTTTACTTACTCACACTTCGCTGTACCTCAAGCGAAAGGAAACAATGTCGTG ATTACAAGCTATATGACAAACAGAGGATTCTACGCAGACAAACAATCAACGTTTGC GCCTAGCTTCCTGCTGAACATCAAAGGCAAGAAAACATCTGTTGTCAAAGACAGC ATCCTTGAACAAGGACAATTAACAGTTAACAAATAAAAACGCAAAAGAAAATGCCG ATATCCTATTGGCATTGACGTCAGGTGGCACACCTGCAGAGAACCTCTACTTCCA ATCGCACCATCATCACCACCATGATTACAAGGATGACGACGATAAGTGAGGATCC GAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACCACCA CCAC
pNIC28-Bsa4	ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTA CTTCCAATCCATGGAGACCGACGTCCACATATACCTGCCGTTCACTATTATTTAGT GAAATGAGATATTATGATATTTTCTGAATTGTGATTAAAAAGGCAACTTTATGCCCA TGCAACAGAAACTATAAAAAATACAGAGAATGAAAAGAAACAGATAGATTTTTTAGT TCTTTAGGCCCGTAGTCTGCAAATCCTTTTATGATTTTCTATCAAACAAAAGAGGAA AATAGACCAGTTGCAATCCAAACGAGAGTCTAATAGAATGAGGTCGAAAAGTAAAT CGCGCGGGTTTGTTACTGATAAAGCAGGCAAGACCTAAAATGTGTAAAGGGCAAA GTGTATACTTTGGCGTCACCCCTTACATATTTTAGGTCTTTTTTTATTGTGCGTAAC TAACTTGCCATCTTCAAACAGGAGGGCTGGAAGAAGCAGACCGCTAACACAGTAC ATAAAAAAGGAGACATGAACGATGAACATCAAAAAGTTTGCAAAACAAGCAACAGT ATTAACCTTTACTACCGCACTGCTGGCAGGAGGCGCAACTCAAGCGTTTGCGAAA GAAACGAACCAAAAGCCATATAAGGAAACATACGGCATTTCCCATATTACACGCCA TGATATGCTGCAAATCCCTGAACAGCAAAAAAATGAAA

Table 16: CtPOT DNA sequence with the used primer sequences for LIC.

CtPOT	ATGGCGTCAGAACCGACCAAAGTCGCCACTGGCGACGATGTTCAAACCATCTCGA AGGAGAAGGATGCCGGTTTGGCCCATGTGCCCTCTCGTCCCGAAGATGATAAACT GGGAATTACGCCCGATGGCGGCGAACCGACTGAAGAAGAAATGCAGACGCTGCG TCGTGTTCCCGACAAGATTCCGCTGAACATCTACACTATCGCATTTATCGAGCTGT GCGAGCGATTCTCGTACTACGGTACCACTGCCGTATTTACCAACTTCATTCAACAG CCTCTCCCCCCCAACTCCTCGACCGGCGCCAGCGGTGACATCAACCATGGGCAA GCTGGTGCTCTTGGCCTTGGCCAGCGCACTGCCTTCTCCCTGACGACTTTCAACT CGTTCTGGCAGTATACAATGCCTCTGTTTGGTGCTTACGTTGCCGACTCTATTCTT GGCCGTTACCGAACCATCTGTACCGCCCTTCTTGTCGACATCATTGGCCATGTTC TCTTCATCATCTGCGGTCTTCCTCCTGTTATTAAGAACCCTAATGGCTCGCTTGCT GCTTTCTGCTTGGGTGTTATCCTGGTCGGTGTCGGTACCGGCGGTTTCAAGCCCA ACGTCAACCCTCTGATTGTCGAGCAGCTTGACTTGGAAAAGATGGTGGTTCGAAC GCTGCCCACTGGCGAGCGTGTGATTGTTGATCCTGCGACTACTGCCTCCCGTGTC TACCACTACTTCTATCTATTCATCAACATTGGCGCTCTGCTGGGCCAGCTCGCCAT GGTCTACTGCGAGCACTATGTCGGCTTCTGGTTGTCTTTCACCCTGCCTACCATC CTGTTGTGCCTCTGTCCGCTTGTCATGCTCTGGGGCCGCCGCCGGTACCGTCGT GTCCCGCCTCAGCCGTCGACCCTCGGCAAGTCGTTCAAGCTGTTCTGGCTTGCC AATAAGGGCCGCTGGTCGCTCAATCCTGTAAAGACATGGAAGCAGCTACATGATG GAACCTTCTGGGAGAACGTCAAGCCCTCCAAGTTTGACGATGCTACCAGGCCTAA GTGGATGGATTTCAACGATGCTTGGGTTGACGAGGTCCGCCGTGGCTTCAATGC CTGCGCTGTATTCCTGTGGTATCCTCTCTTTTGGATCTGCTACAACCAAATCAACA ACAACCTCATATCCCAAGCTGCCACCATGCGTGTTGATGGTGTTCCAAACGACGT GTTGACCAACCTTAACCCCTTTGCTCTGATTATCATGATTCCCATCCTGGACACCG TCGTCTATCCTGCCCTCCGCAAGTTGCGTATCAACCCGACTCCTCTCAAGCGCAT CACTGTCGGTTACTTCGTCGCTTCCGCCGCCATGATCTGGGCCTGCGTCATCCAA TACTATATTTACAAGCGCTCTGAGTGCGGCCACTACGCCAACGACTGCGACCCCG TCAACATCAGCATGTGGGCTCAATCTGGCTCTTACATCCTAATCGCTTTCTCCGAG ATCATGGCTTCCATTACCAGTCTCGAGTACGCCCACTCGAAGGCTCCAGCGAACA TGCGTTCCATGGTCCAGTCAGTTGCGCTCTTTATGAACGCCATCTCGTCGGCCAT CGGCTTCGCGCTGGCTCCCCTGGCCGACGACCCGAACCTGATCTGGAACTATGG CGTTGTCGCTGTTCTTGCTGCCATTGGTGGCACCATGTTCTGGATACAGTTCCGC CATCTTGATGAGCAGGAAGATGAGCTTAACATGCTTCCGACCGGCGAGTTTGGAA ACCAGGCTGCTGCGGTACCGGTTGAGACAGTTAAGGAGTCGGACAAGAGCCAGG CTTAA
Primer forward (pNIC-CTHF)	TTA AGA AGG AGA TAT ACT ATG GCG TCA GAA CCG ACC AAA GTC GCC
Primer reverse (pNIC-CTHF)	GAT TGG AAG TAG AGG TTC TCT GCA GCC TGG CTC TTG TCC GAC TCC
Primer forward (pNIC28-Bsa4)	TAC TTC CAA TCC ATG GCG TCA GAA CCG ACC AAA GTC GCC
Primer reverse (pNIC28-Bsa4)	TAT CCA CCT TTA CTG TTA AGC CTG GCT CTT GTC CGA CTC

Table 17: Primer sequences for LIC and PepT1 DNA sequence.

PepT1	ATGGGAATGTCCAAATCACACAGTTTCTTTGGTTATCCCCTGAGCATCTTCTTCAT CGTGGTCAATGAGTTTTGCGAAAGATTTTCCTACTATGGAATGCGAGCAATCCTG ATTCTGTACTTCACAAATTTCATCAGCTGGGATGATAACCTGTCCACCGCCATCTA CCATACGTTTGTGGCTCTGTGCTACCTGACGCCAATTCTCGGAGCTCTTATCGCC GACTCGTGGCTGGGAAAGTTCAAGACCATTGTGTCGCTCTCCATTGTCTACACAA TTGGACAAGCAGTCACCTCAGTAAGCTCCATTAATGACCTCACAGACCACAACCA TGATGGCCCCCCGACAGCCTTCCTGTGCACGTGGTGCTGTCCTTGATCGGCCTG GCCCTGATAGCTCTCGGGACTGGAGGAATCAAACCCTGTGTGTCTGCGTTTGGT GGAGATCAGTTTGAAGAGGGCCAGGAGAAACAAAGAAACAGATTTTTTTCCATCT TTTACTTGGCTATTAATGCTGGAAGTTTGCTTTCCACAATCATCACACCCATGCTC AGAGTTCAACAATGTGGAATTCACAGTAAACAAGCTTGTTACCCACTGGCCTTTG GGGTTCCTGCTGCTCTCATGGCTGTTGCCCTGATTGTGTTTGTCCTTGGCAGTGG GATGTACAAGAAGTTCAAGCCACAGGGCAACATCATGGGTAAAGTGGCCAAGTG CATCGGTTTTGCCATCAAAAATAGATTTAGGCATCGGAGTAAGGCATTTCCCAAGA GGGAGCACTGGCTGGACTGGGCTAAAGAGAAATACGATGAGCGGCTCATCTCCC AAATTAAGATGGTTACGAGGGTGATGTTCCTGTATATTCCACTCCCAATGTTCTGG GCCTTGTTTGACCAGCAGGGCTCCAGGTGGACACTGCAGGCAACAACTATGTCC GGGAAAATCGGAGCTCTTGAAATTCAGCCCGATCAGATGCAGACCGTGAACGCC ATCCTGATCGTGATCATGGTCCCGATCTTCGATGCTGTGCTGTACCCTCTCATTG CAAAATGTGGCTTCAATTTCACCTCCTTGAAGAAGATGGCAGTTGGCATGGTCCT GGCCTCCATGGCCTTTGTGGTGGCTGCCATCGTGCAGGTGGAAATCGATAAAAC TCTTCCAGTCTTCCCCAAAGGAAACGAAGTCCAAATTAAAGTTTTGAATATAGGAA ACAATACCATGAATATATCTCTTCCTGGAGAGATGGTGACACTTGGCCCAATGTCT CAAACAAATGCATTTATGACTTTTGATGTAAACAAACTGACAAGGATAAACATTTCT TCTCCTGGATCACCAGTCACTGCCGTAACTGACGACTTCAAGCAGGGCCAACGC CACACGCTTCTAGTGTGGGCCCCCAATCACTACCAGGTGGTAAAGGATGGTCTTA ACCAGAAGCCAGAAAAAGGGGAAAATGGAATCAGATTTGTAAATACTTTTAACGA GCTCATCACCATCACAATGAGTGGGAAAGTTTATGCAAACATCAGCAGCTACAAT GCCAGCACATACCAGTTTTTTCCTTCTGGCATAAAAGGCTTCACAATAAGCTCAAC AGAGATTCCGCCACAATGTCAACCTAATTTCAATACTTTCTACCTTGAATTTGGTA GTGCTTATACCTATATAGTCCAAAGGAAGAATGACAGCTGCCCTGAAGTGAAGGT GTTTGAAGATATTTCAGCCAACACAGTTAACATGGCTCTGCAAATCCCGCAGTATT TTCTTCTCACCTGTGGCGAAGTGGTCTTCTCTGTCACGGGATTGGAATTCTCATAT TCTCAGGCTCCTTCCAACATGAAGTCGGTGCTTCAGGCAGGATGGCTGCTGACC GTGGCTGTTGGCAACATCATTGTGCTCATCGTGGCAGGGGCAGGCCAGTTCAGC AAACAGTGGGCCGAGTACATTCTATTTGCCGCGTTGCTTCTGGTCGTCTGTGTAA TTTTTGCCATCATGGCTCGGTTCTATACTTACATCAACCCAGCGGAGATCGAAGCT CAATTTGATGAGGATGAAAAGAAAAACAGACTGGAAAAGAGTAACCCATATTTCAT GTCAGGGGCCAATTCACAGAAACAGATG
Primer forward (pNIC-CTHF)	TTAAGAAGGAGATATACTATGGGAATGTCCAAATCACACAGT
Primer reverse (pNIC-CTHF)	GATTGGAAGTAGAGGTTCTCTGCCATCTGTTTCTGTGAATTGGCC
Primer forward (pNIC28-Bsa4)	TAC TTC CAA TCC ATG GGA ATG TCC AAA TCA CAC AGT
Primer reverse (pNIC28-Bsa4)	TAT CCA CCT TTA CTG TTA CAT CTG TTT CTG TGA ATT GGC

Table 18: PepT2 gene sequences and used primers for LIC.

PepT2	ATGAATCCTTTCCAGAAAAATGAGTCCAAGGAAACTCTTTTTTCACCTGTCTCCATT GAAGAGGTACCACCTCGACCACCTAGCCCTCCAAAGAAGCCATCTCCGACAATCT GTGGCTCCAACTATCCACTGAGCATTGCCTTCATTGTGGTGAATGAATTCTGCGA GCGCTTTTCCTATTATGGAATGAAAGCTGTGCTGATCCTGTATTTCCTGTATTTCCT GCACTGGAATGAAGATACCTCCACATCTATATACCATGCCTTCAGCAGCCTCTGTT ATTTTACTCCCATCCTGGGAGCAGCCATTGCTGACTCGTGGTTGGGAAAATTCAA GACAATCATCTATCTCTCCTTGGTGTATGTGCTTGGCCATGTGATCAAGTCCTTGG GTGCCTTACCAATACTGGGAGGACAAGTGGTACACACAGTCCTATCATTGATCGG CCTGAGTCTAATAGCTTTGGGGACAGGAGGCATCAAACCCTGTGTGGCAGCTTTT GGTGGAGACCAGTTTGAAGAAAAACATGCAGAGGAACGGACTAGATACTTCTCAG TCTTCTACCTGTCCATCAATGCAGGGAGCTTGATTTCTACATTTATCACACCCATG CTGAGAGGAGATGTGCAATGTTTTGGAGAAGACTGCTATGCATTGGCTTTTGGAG TTCCAGGACTGCTCATGGTAATTGCACTTGTTGTGTTTGCAATGGGAAGCAAAATA TACAATAAACCACCCCCTGAAGGAAACATAGTGGCTCAAGTTTTCAAATGTATCTG GTTTGCTATTTCCAATCGTTTCAAGAACCGTTCTGGAGACATTCCAAAGCGACAGC ACTGGCTAGACTGGGCGGCTGAGAAATATCCAAAGCAGCTCATTATGGATGTAAA GGCACTGACCAGGGTACTATTCCTTTATATCCCATTGCCCATGTTCTGGGCTCTTT TGGATCAGCAGGGTTCACGATGGACTTTGCAAGCCATCAGGATGAATAGGAATTT GGGGTTTTTTGTGCTTCAGCCGGACCAGATGCAGGTTCTAAATCCCCTTCTGGTT CTTATCTTCATCCCGTTGTTTGACTTTGTCATTTATCGTCTGGTCTCCAAGTGTGGA ATTAACTTCTCATCACTTAGGAAAATGGCTGTTGGTATGATCCTAGCATGCCTGGC ATTTGCAGTTGCGGCAGCTGTAGAGATAAAAATAAATGAAATGGCCCCAGCCCAG CCAGGTCCCCAGGAGGTTTTCCTACAAGTCTTGAATCTGGCAGATGATGAGGTGA AGGTGACAGTGGTGGGAAATGAAAACAATTCTCTGTTGATAGAGTCCATCAAATC CTTTCAGAAAACACCACACTATTCCAAACTGCACCTGAAAACAAAAAGCCAGGATT TTCACTTCCACCTGAAATATCACAATTTGTCTCTCTACACTGAGCATTCTGTGCAG GAGAAGAACTGGTACAGTCTTGTCATTCGTGAAGATGGGAACAGTATCTCCAGCA TGATGGTAAAGGATACAGAAAGCAGAACAACCAATGGGATGACAACCGTGAGGTT TGTTAACACTTGCATAAAGATGTCAACATCTCCCTGAGTACAGATACCTCTCTCAA TGTTGGTGAAGACTATGGTGTGTCTGCTTATAGAACTGTGCAAAGAGGAGAATAC CCTGCAGTGCACTGTAGAACAGAAGATAAGAACTTTTCTCTGAATTTGGGTCTTCT AGACTTTGGTGCAGCATATCTGTTTGTTATTACTAATAACACCAATCAGGGTCTTC AGGCCTGGAAGATTGAAGACATTCCAGCCAACAAAATGTCCATTGCGTGGCAGCT ACCACAATATGCCCTGGTTACAGCTGGGGAGGTCATGTTCTCTGTCACAGGTCTT GAGTTTTCTTATTCTCAGGCTCCCTCTGGCATGAAATCTGTGCTCCAGGCAGCTTG GCTATTGACAATTGCAGTTGGGAATATCATCGTGCTTGTTGTGGCACAGTTCAGTG GCCTGGTACAGTGGGCCGAATTCATTTTGTTTTCCTGCCTCCTGCTGGTGATCTG CCTGATCTTCTCCATCATGGGCTACTACTATGTTCCTGTAAAGACAGAGGATATGC GGGGTCCAGCAGATAAGCACATTCCTCACATCCAGGGGAACATGATCAAACTAGA GACCAAGAAGACAAAACTC
Primer forward (pNIC-CTHF)	TTAAGAAGGAGATATACTATGAATCCTTTCCAGAAAAATGAGTCC
Primer reverse (pNIC-CTHF)	GATTGGAAGTAGAGGTTCTCTGCGAGTTTTGTCTTCTTGGTC
Primer forward (pNIC28-Bsa4)	TAC TTC CAA TCC ATG AAT CCT TTC CAG AAA AAT GAG TCC
Primer reverse (pNIC28-Bsa4)	TAT CCA CCT TTA CTG TTA GAG TTT TGT CTT CTT GGT CTC

Table 19: $PepT_{So2}$-PepT1 and primer forward and reverse sequence.

$PepT_{So2}$-PepT1	ATGACTCTTGGTACCAATCAGGTAAGCAAGACACATTCGTTTATGACTGTGTCACT TATCGAATTATGGGAACGTTTTGGTTATTACGGCATGCAGGCGCTGATCGTGTAC TTTATGGTGCAGCGTCTTGGCTTTGATGACTCACGTGCAAACTTAGTGTGGAGCG CCTGTGCGGCGTTAATTTACGTATCACCCGCAATTGGGGGCTGGGTCGGGGATA AAATTCTTGGCACCAAACGCACTATGTTGCTCGGTGCGGGGATTTTATCCGTCGG TTATGCACTCATGACGGTGCCAACCGAAAACACTTGGTTTATGTTCTCCGCCCTT GGGGTGATTGTGGTCGGTAACGGCCTGTTCAAACCCAATGCCGGTAACTTAGTC CGCAAGATCTACGAAGGCGATGACTCTAAAATCGACAGTGCTTTTACCATCTATTA TATGGCGGTAAACGTAGGTTCGACCTTCTCTATGCTGTTAACCCCTTGGATAAAG GACTATGTTAACGCTCAATACGGTAATGAATTTGGCTGGCATGCCGCCTTTGCCG TGTGCTGTGTGGGTATTCTTGTCGGTTTAGGCAACTATGCTCTGATGCATAAAAG CCTCGCGAACTATGGTTCTGAGCCAGATACGCGCCCAGTCAACAAGAAGAGCCT TGCTATTGTATTAGCATTAGCAGCACTTTCCGTGGTTGCTTCGGCGATTATTCTAG AATATGAAGATGTTGCCCGTGTATTCGTTTACGCTGCTGGTGTTGCAGTATTAGG CATTTTCTTCCATTTGATTCGCACAAGTGAGCCAAGTGAGCGTGCCGGACTTATC GCGGCGCTTATTCTCACCGTACAAACGGTATTCTTCTTTATCTTCTATCAGCAAAT GTCGACTTCGCTGGCTCTGTTTGCGCTACGTAACGTCGATTGGGATTTCCAAGTC TTTGGCACCCATTTATGGACTTGGTCACCTGCGCAGTTCCAAGCACTTAACCCAA TCTGGATCATGGTATTAAGCCCAGTTCTGGCATGGAGTTATTCCTGGGCTGGCCG CAATAATAAGGACTTTTCGATAGCAGCTAAGTTTGCCTTGGGCTTTGCTGTCGTG GCGATTGGTTTCTTTATCTACGGTTTTGCAGGCCAATTTGTGGTGGCGGCGATTG TGCAGGTGGAAATTGATAAAACCCTGCCGGTGTTTCCGAAAGGCAACGAAGTGCA GATTAAAGTGCTGAACATTGGCAACAACACCATGAACATTAGCCTGCCGGGCGAA ATGGTGACCCTGGGCCCGATGAGCCAGACCAACGCGTTTATGACCTTTGATGTG AACAAACTGACCCGCATTAACATTAGCAGCCCGGGCAGCCCGGTGACCGCGGTG ACCGATGATTTTAAACAGGGCCAGCGCCATACCCTGCTGGTGTGGGCGCCGAAC CATTATCAGGTGGTGAAAGATGGCCTGAACCAGAAACCGGAAAAAGGCGAAAAC GGCATTCGCTTTGTGAACACCTTTAACGAACTGATTACCATTACCATGAGCGGCA AAGTGTATGCGAACATTAGCAGCTATAACGCGAGCACCTATCAGTTTTTTCCGAG CGGCATTAAAGGCTTTACCATTAGCAGCACCGAAATTCCGCCGCAGTGCCAGCC GAACTTTAACACCTTTTATCTGGAATTTGGCAGCGCGTATACCTATATTGTGCAGC GCAAAAACGATAGCTGCCCGGAAGTGAAAGTGTTTGAAGATATTAGCGCGAATGG TAAAACCTCTTCTTGGGTAATGATTTGGGGTTACGCTTCTTACTCTTTGGGGGAGT TACTAGTAAGTGGCCTTGGTTTAGCGATGATTGCCCGCTACGTGCCTGCACGTAT GGGCGGCTTTATGATGGGCGCGTATTTCGTGGCATCGGGCATTTCCCAGTACTTA GGCGGCGTTGTGGCCAACTTTGCTAGCGTGCCACAGGACTTAGTCGATCCACTG CAAACACTCCCTGTTTATACCAATCTATTCAATAAATTAGGGGTTGCGGCCGTGGT TTGTACCATTATCGCTCTCGCCGTATTACCGCTGATGCGCCGTTTGACAGAAAGC CATCATGCCCATAGCAGTATTGAGAATAATGCAGCAGCCTCCTTGCGCGATGTAA AAGCTGAGCAA
Primer forward (pNIC-CTHF)	TTA AGA AGG AGA TAT ACT ATG ACT CTT GGT ACC AAT CAG GTA AGC
Primer reverse (pNIC-CTHF)	GAT TGG AAG TAG AGG TTC TCT GCT TGC TCA GCT TTT ACA TCG CGC AAG

Table 20: $PepT_{So2}$-PepT2 DNA sequence and primer sequences for LIC.

$PepT_{So2}$-PepT2	ATGACTCTTGGTACCAATCAGGTAAGCAAGACACATTCGTTTATGACTGTGTCAC TTATCGAATTATGGGAACGTTTTGGTTATTACGGCATGCAGGCGCTGATCGTGT ACTTTATGGTGCAGCGTCTTGGCTTTGATGACTCACGTGCAAACTTAGTGTGGA GCGCCTGTGCGGCGTTAATTTACGTATCACCCGCAATTGGGGGCTGGGTCGGG GATAAAATTCTTGGCACCAAACGCACTATGTTGCTCGGTGCGGGGATTTTATCC GTCGGTTATGCACTCATGACGGTGCCAACCGAAAACACTTGGTTTATGTTCTCC GCCCTTGGGGTGATTGTGGTCGGTAACGGCCTGTTCAAACCCAATGCCGGTAA CTTAGTCCGCAAGATCTACGAAGGCGATGACTCTAAAATCGACAGTGCTTTTAC CATCTATTATATGGCGGTAAACGTAGGTTCGACCTTCTCTATGCTGTTAACCCCT TGGATAAAGGACTATGTTAACGCTCAATACGGTAATGAATTTGGCTGGCATGCC GCCTTTGCCGTGTGCTGTGTGGGTATTCTTGTCGGTTTAGGCAACTATGCTCTG ATGCATAAAAGCCTCGCGAACTATGGTTCTGAGCCAGATACGCGCCCAGTCAAC AAGAAGAGCCTTGCTATTGTATTAGCATTAGCAGCACTTTCCGTGGTTGCTTCG GCGATTATTCTAGAATATGAAGATGTTGCCCGTGTATTCGTTTACGCTGCTGGTG TTGCAGTATTAGGCATTTTCTTCCATTTGATTCGCACAAGTGAGCCAAGTGAGCG TGCCGGACTTATCGCGGCGCTTATTCTCACCGTACAAACGGTATTCTTCTTTATC TTCTATCAGCAAATGTCGACTTCGCTGGCTCTGTTTGCGCTACGTAACGTCGATT GGGATTTCCAAGTCTTTGGCACCCATTTATGGACTTGGTCACCTGCGCAGTTCC AAGCACTTAACCCAATCTGGATCATGGTATTAAGCCCAGTTCTGGCATGGAGTT ATTCCTGGGCTGGCCGCAATAATAAGGACTTTTCGATAGCAGCTAAGTTTGCCT TGGGCTTTGCTGTCGTGGCGATTGGTTTCTTTATCTACGGTTTTGCAGGCCAATT TGCCGTAGCGGCGGCGGTGGAAATTAAAATTAACGAAATGGCGCCGGCGCAGC CGGGCCCGCAGGAAGTGTTTCTGCAGGTGCTGAACCTGGCGGATGATGAAGTG AAAGTGACCGTGGTGGGCAACGAAAACAACAGCCTGCTGATTGAAAGCATTAAA AGCTTTCAGAAAACCCCGCATTATAGCAAACTGCATCTGAAAACCAAAAGCCAG GATTTTCATTTTCATCTGAAATATCATAACCTGAGCCTGTATACCGAACATAGCG TGCAGGAAAAAAACTGGTATAGCCTGGTGATTCGCGAAGATGGCAACAGCATTA GCAGCATGATGGTGAAAGATACCGAAAGCCGCACCACCAACGGCATGACCACC GTGCGCTTTGTGAACACCCTGCATAAAGATGTGAACATTAGCCTGAGCACCGAT ACCAGCCTGAACGTGGGCGAAGATTATGGCGTGAGCGCGTATCGCACCGTGCA GCGCGGCGAATATCCGGCGGTGCATTGCCGCACCGAAGATAAAAACTTTAGCC TGAACCTGGGCCTGCTGGATTTTGGCGCGGCGTATCTGTTTGTGATTACCAACA ACACCAACCAGGGCCTGCAGGCGTGGAAAATTGAAGATATTCCGGCGAATGGT AAAACCTCTTCTTGGGTAATGATTTGGGGTTACGCTTCTTACTCTTTGGGGGAGT TACTAGTAAGTGGCCTTGGTTTAGCGATGATTGCCCGCTACGTGCCTGCACGTA TGGGCGGCTTTATGATGGGCGCGTATTTCGTGGCATCGGGCATTTCCCAGTACT TAGGCGGCGTTGTGGCCAACTTTGCTAGCGTGCCACAGGACTTAGTCGATCCA CTGCAAACACTCCCTGTTTATACCAATCTATTCAATAAATTAGGGGTTGCGGCCG TGGTTTGTACCATTATCGCTCTCGCCGTATTACCGCTGATGCGCCGTTTGACAG AAAGCCATCATGCCCATAGCAGTATTGAGAATAATGCAGCAGCCTCCTTGCGCG ATGTAAAAGCTGAGCAA
Primer forward (pNIC-CTHF)	TTA AGA AGG AGA TAT ACT ATG ACT CTT GGT ACC AAT CAG GTA AGC
Primer reverse (pNIC-CTHF)	GAT TGG AAG TAG AGG TTC TCT GCT TGC TCA GCT TTT ACA TCG CGC AAG

Table 21: $PepT_{So2}$-BRIL-insert, primer forward and primer reverse DNA sequences which were used for cloning into pNIC-CTHF vector..

$PepT_{So2}$-BRIL-insert	ATGACTCTTGGTACCAATCAGGTAAGCAAGACACATTCGTTTATGACTGTGTCAC TTATCGAATTATGGGAACGTTTTGGTTATTACGGCATGCAGGCGCTGATCGTGT ACTTTATGGTGCAGCGTCTTGGCTTTGATGACTCACGTGCAAACTTAGTGTGGA GCGCCTGTGCGGCGTTAATTTACGTATCACCCGCAATTGGGGGCTGGGTCGGG GATAAAATTCTTGGCACCAAACGCACTATGTTGCTCGGTGCGGGGATTTTATCC GTCGGTTATGCACTCATGACGGTGCCAACCGAAAACACTTGGTTTATGTTCTCC GCCCTTGGGGTGATTGTGGTCGGTAACGGCCTGTTCAAACCCAATGCCGGTAA CTTAGTCCGCAAGATCTACGAAGGCGATGACTCTAAAATCGACAGTGCTTTTAC CATCTATTATATGGCGGTAAACGTAGGTTCGACCTTCTCTATGCTGTTAACCCCT TGGATAAAGGACTATGTTAACGCTCAATACGGTAATGAATTTGGCTGGCATGCC GCCTTTGCCGTGTGCTGTGTGGGTATTCTTGTCGGTTTAGGCAACTATGCTCTG ATGCATAAAAGCCTCGCGAACTATGGTGCTGATCTTGAAGACAATTGGGAAACC CTCAACGACAATTTAAAAGTGATCGAAAAAGCGGATAACGCGGCGCAAGTCAAA GACGCGTTAACGAAGATGCGCGCCGCAGCCCTGGATGCGCAAAAAGCAACGC CGCCGAAGCTCGAAGATAAATCACCGGACAGCCCGGAAATGAAAGATTTCCGC CACGGTTTCGACATTCTGGTCGGTCAGATTGACGACGCGCTGAAGCTGGCAAA TGAAGGTAAAGTAAAAGAAGCGCAGGCTGCTGCAGAGCAACTGAAAACGACCC GCAACGCCTATATTCAGAAGTATCTGGTCAACAAGAAGAGCCTTGCTATTGTATT AGCATTAGCAGCACTTTCCGTGGTTGCTTCGGCGATTATTCTAGAATATGAAGAT GTTGCCCGTGTATTCGTTTACGCTGCTGGTGTTGCAGTATTAGGCATTTTCTTCC ATTTGATTCGCACAAGTGAGCCAAGTGAGCGTGCCGGACTTATCGCGGCGCTT ATTCTCACCGTACAAACGGTATTCTTCTTTATCTTCTATCAGCAAATGTCGACTTC GCTGGCTCTGTTTGCGCTACGTAACGTCGATTGGGATTTCCAAGTCTTTGGCAC CCATTTATGGACTTGGTCACCTGCGCAGTTCCAAGCACTTAACCCAATCTGGAT CATGGTATTAAGCCCAGTTCTGGCATGGAGTTATTCCTGGGCTGGCCGCAATAA TAAGGACTTTTCGATAGCAGCTAAGTTTGCCTTGGGCTTTGCTGTCGTGGCGAT TGGTTTCTTTATCTACGGTTTTGCAGGCCAATTTGCCGTAAATGGTAAAACCTCT TCTTGGGTAATGATTTGGGGTTACGCTTCTTACTCTTTGGGGGAGTTACTAGTAA GTGGCCTTGGTTTAGCGATGATTGCCCGCTACGTGCCTGCACGTATGGGCGGC TTTATGATGGGCGCGTATTTCGTGGCATCGGGCATTTCCCAGTACTTAGGCGGC GTTGTGGCCAACTTTGCTAGCGTGCCACAGGACTTAGTCGATCCACTGCAAACA CTCCCTGTTTATACCAATCTATTCAATAAATTAGGGGTTGCGGCCGTGGTTTGTA CCATTATCGCTCTCGCCGTATTACCGCTGATGCGCCGTTTGACAGAAAGCCATC ATGCCCATAGCAGTATTGAGAATAATGCAGCAGCCTCCTTGCGCGATGTAAAAG CTGAGCAA
Primer forward (pNIC-CTHF)	TTA AGA AGG AGA TAT ACT ATG ACT CTT GGT ACC AAT CAG GTA AGC
Primer reverse (pNIC-CTHF)	GAT TGG AAG TAG AGG TTC TCT GCT TGC TCA GCT TTT ACA TCG

Table 22: DNA sequences of PepT$_{So2}$-BRIL-Cterm, primer forward and reverse for cloning into the vector pNIC-CTHF with the C-terminal His-tag.

PepT$_{So2}$-BRIL-Cterm	ATGACTCTTGGTACCAATCAGGTAAGCAAGACACATTCGTTTATGACTGTGTCAC TTATCGAATTATGGGAACGTTTTGGTTATTACGGCATGCAGGCGCTGATCGTGT ACTTTATGGTGCAGCGTCTTGGCTTTGATGACTCACGTGCAAACTTAGTGTGGA GCGCCTGTGCGGCGTTAATTTACGTATCACCCGCAATTGGGGGCTGGGTCGGG GATAAAATTCTTGGCACCAAACGCACTATGTTGCTCGGTGCGGGGATTTTATCC GTCGGTTATGCACTCATGACGGTGCCAACCGAAAACACTTGGTTTATGTTCTCC GCCCTTGGGGTGATTGTGGTCGGTAACGGCCTGTTCAAACCCAATGCCGGTAA CTTAGTCCGCAAGATCTACGAAGGCGATGACTCTAAAATCGACAGTGCTTTTAC CATCTATTATATGGCGGTAAACGTAGGTTCGACCTTCTCTATGCTGTTAACCCCT TGGATAAAGGACTATGTTAACGCTCAATACGGTAATGAATTTGGCTGGCATGCC GCCTTTGCCGTGTGCTGTGTGGGTATTCTTGTCGGTTTAGGCAACTATGCTCTG ATGCATAAAAGCCTCGCGAACTATGGTTCTGAGCCAGATACGCGCCCAGTCAAC AAGAAGAGCCTTGCTATTGTATTAGCATTAGCAGCACTTTCCGTGGTTGCTTCG GCGATTATTCTAGAATATGAAGATGTTGCCCGTGTATTCGTTTACGCTGCTGGTG TTGCAGTATTAGGCATTTTCTTCCATTTGATTCGCACAAGTGAGCCAAGTGAGCG TGCCGGACTTATCGCGGCGCTTATTCTCACCGTACAAACGGTATTCTTCTTTATC TTCTATCAGCAAATGTCGACTTCGCTGGCTCTGTTTGCGCTACGTAACGTCGATT GGGATTTCCAAGTCTTTGGCACCCATTTATGGACTTGGTCACCTGCGCAGTTCC AAGCACTTAACCCAATCTGGATCATGGTATTAAGCCCAGTTCTGGCATGGAGTT ATTCCTGGGCTGGCCGCAATAATAAGGACTTTTCGATAGCAGCTAAGTTTGCCT TGGGCTTTGCTGTCGTGGCGATTGGTTTCTTTATCTACGGTTTTGCAGGCCAATT TGCCGTAAATGGTAAAACCTCTTCTTGGGTAATGATTTGGGGTTACGCTTCTTAC TCTTTGGGGGAGTTACTAGTAAGTGGCCTTGGTTTAGCGATGATTGCCCGCTAC GTGCCTGCACGTATGGGCGGCTTTATGATGGGCGCGTATTTCGTGGCATCGGG CATTTCCCAGTACTTAGGCGGCGTTGTGGCCAACTTTGCTAGCGTGCCACAGGA CTTAGTCGATCCACTGCAAACACTCCCTGTTTATACCAATCTATTCAATAAATTAG GGGTTGCGGCCGTGGTTTGTACCATTATCGCTCTCGCCGTATTACCGCTGATGC GCCGTTTGACAGAAAGCCATCATGCCCATAGCAGTGCTGATCTTGAAGACAATT GGGAAACCCTCAACGACAATTTAAAAGTGATCGAAAAAGCGGATAACGCGGCG CAAGTCAAAGACGCGTTAACGAAGATGCGCGCCGCAGCCCTGGATGCGCAAAA AGCAACGCCGCCGAAGCTCGAAGATAAATCACCGGACAGCCCGGAAATGAAAG ATTTCCGCCACGGTTTCGACATTCTGGTCGGTCAGATTGACGACGCGCTGAAGC TGGCAAATGAAGGTAAAGTAAAAGAAGCGCAGGCTGCTGCAGAGCAACTGAAA ACGACCCGCAACGCCTATATTCAGAAGTATCTG
Primer forward (pNIC-CTHF)	TTA AGA AGG AGA TAT ACT ATG ACT CTT GGT ACC AAT CAG GTA AGC
Primer reverse (pNIC-CTHF)	GAT TGG AAG TAG AGG TTC TCT GCC AGA TAC TTC TGA ATA TAG GCG

Table 23: $PepT_{So2}$-LT4-insert DNA sequences as well as the sequences of forward and reverse primer for LIC.

$PepT_{So2}$-LT4-insert	ATGACTCTTGGTACCAATCAGGTAAGCAAGACACATTCGTTTATGACTGTGTCACT TATCGAATTATGGGAACGTTTTGGTTATTACGGCATGCAGGCGCTGATCGTGTAC TTTATGGTGCAGCGTCTTGGCTTTGATGACTCACGTGCAAACTTAGTGTGGAGCG CCTGTGCGGCGTTAATTTACGTATCACCCGCAATTGGGGGCTGGGTCGGGGATA AAATTCTTGGCACCAAACGCACTATGTTGCTCGGTGCGGGGATTTTATCCGTCGG TTATGCACTCATGACGGTGCCAACCGAAAACACTTGGTTTATGTTCTCCGCCCTT GGGGTGATTGTGGTCGGTAACGGCCTGTTCAAACCCAATGCCGGTAACTTAGTC CGCAAGATCTACGAAGGCGATGACTCTAAAATCGACAGTGCTTTTACCATCTATTA TATGGCGGTAAACGTAGGTTCGACCTTCTCTATGCTGTTAACCCCTTGGATAAAG GACTATGTTAACGCTCAATACGGTAATGAATTTGGCTGGCATGCCGCCTTTGCCG TGTGCTGTGTGGGTATTCTTGTCGGTTTAGGCAACTATGCTCTGATGCATAAAAG CCTCGCGAACTATGGTAATATATTTGAAATGTTACGTATAGATGAACGTCTTAGAC TTAAAATCTATAAAGACACAGAAGGCTATTACACTATTGGCATCGGTCATTTGCTT ACAAAAAGTCCATCACTTAATGCTGCTAAATCTGAATTAGATAAAGCTATTGGGCG TAATACTAATGGTGTAATTACAAAAGATGAGGCTGAAAAACTCTTTAATCAGGATG TTGATGCTGCTGTTCGCGGAATTCTGAGAAATGCTAAATTAAAACCGGTTTATGAT TCTCTTGATGCGGTTCGTCGCGCTGCATTGATTAATATGGTTTTCCAAATGGGAGA AACCGGTGTGGCAGGATTTACTAACTCTTTACGTATGCTTCAACAAAAACGCTGG GATGAAGCAGCAGTTAACTTAGCTAAAAGTATATGGTATAATCAAACACCTAATCG CGCAAAACGAGTCATTACAACGTTTAGAACTGGCACTTGGGACGCGTATCGCCCA GTCAACAAGAAGAGCCTTGCTATTGTATTAGCATTAGCAGCACTTTCCGTGGTTG CTTCGGCGATTATTCTAGAATATGAAGATGTTGCCCGTGTATTCGTTTACGCTGCT GGTGTTGCAGTATTAGGCATTTTCTTCCATTTGATTCGCACAAGTGAGCCAAGTGA GCGTGCCGGACTTATCGCGGCGCTTATTCTCACCGTACAAACGGTATTCTTCTTT ATCTTCTATCAGCAAATGTCGACTTCGCTGGCTCTGTTTGCGCTACGTAACGTCG ATTGGGATTTCCAAGTCTTTGGCACCCATTTATGGACTTGGTCACCTGCGCAGTT CCAAGCACTTAACCCAATCTGGATCATGGTATTAAGCCCAGTTCTGGCATGGAGT TATTCCTGGGCTGGCCGCAATAATAAGGACTTTTCGATAGCAGCTAAGTTTGCCTT GGGCTTTGCTGTCGTGGCGATTGGTTTCTTTATCTACGGTTTTGCAGGCCAATTT GCCGTAAATGGTAAAACCTCTTCTTGGGTAATGATTTGGGGTTACGCTTCTTACTC TTTGGGGGAGTTACTAGTAAGTGGCCTTGGTTTAGCGATGATTGCCCGCTACGTG CCTGCACGTATGGGCGGCTTTATGATGGGCGCGTATTTCGTGGCATCGGGCATT TCCCAGTACTTAGGCGGCGTTGTGGCCAACTTTGCTAGCGTGCCACAGGACTTA GTCGATCCACTGCAAACACTCCCTGTTTATACCAATCTATTCAATAAATTAGGGGT TGCGGCCGTGGTTTGTACCATTATCGCTCTCGCCGTATTACCGCTGATGCGCCGT TTGACAGAAAGCCATCATGCCCATAGCAGTATTGAGAATAATGCAGCAGCCTCCT TGCGCGATGTAAAAGCTGAGCAA
Primer forward (pNIC-CTHF)	TTA AGA AGG AGA TAT ACT ATG ACT CTT GGT ACC AAT CAG GTA AGC
Primer reverse (pNIC-CTHF)	GAT TGG AAG TAG AGG TTC TCT GCT TGC TCA GCT TTT ACA TCG

Table 24: Forward and reverse primer for $PepT_{So2}$-LT4-Cterm and transporter DNA sequence for LIC into the pNIC-CTHF vector.

$PepT_{So2}$-LT4-Cterm	ATGACTCTTGGTACCAATCAGGTAAGCAAGACACATTCGTTTATGACTGTGTCACT TATCGAATTATGGGAACGTTTTGGTTATTACGGCATGCAGGCGCTGATCGTGTAC TTTATGGTGCAGCGTCTTGGCTTTGATGACTCACGTGCAAACTTAGTGTGGAGCG CCTGTGCGGCGTTAATTTACGTATCACCCGCAATTGGGGGCTGGGTCGGGGATA AAATTCTTGGCACCAAACGCACTATGTTGCTCGGTGCGGGGATTTTATCCGTCGG TTATGCACTCATGACGGTGCCAACCGAAAACACTTGGTTTATGTTCTCCGCCCTT GGGGTGATTGTGGTCGGTAACGGCCTGTTCAAACCCAATGCCGGTAACTTAGTC CGCAAGATCTACGAAGGCGATGACTCTAAAATCGACAGTGCTTTTACCATCTATTA TATGGCGGTAAACGTAGGTTCGACCTTCTCTATGCTGTTAACCCCTTGGATAAAG GACTATGTTAACGCTCAATACGGTAATGAATTTGGCTGGCATGCCGCCTTTGCCG TGTGCTGTGTGGGTATTCTTGTCGGTTTAGGCAACTATGCTCTGATGCATAAAAG CCTCGCGAACTATGGTTCTGAGCCAGATACGCGCCCAGTCAACAAGAAGAGCCT TGCTATTGTATTAGCATTAGCAGCACTTTCCGTGGTTGCTTCGGCGATTATTCTAG AATATGAAGATGTTGCCCGTGTATTCGTTTACGCTGCTGGTGTTGCAGTATTAGG CATTTTCTTCCATTTGATTCGCACAAGTGAGCCAAGTGAGCGTGCCGGACTTATC GCGGCGCTTATTCTCACCGTACAAACGGTATTCTTCTTTATCTTCTATCAGCAAAT GTCGACTTCGCTGGCTCTGTTTGCGCTACGTAACGTCGATTGGGATTTCCAAGTC TTTGGCACCCATTTATGGACTTGGTCACCTGCGCAGTTCCAAGCACTTAACCCAA TCTGGATCATGGTATTAAGCCCAGTTCTGGCATGGAGTTATTCCTGGGCTGGCCG CAATAATAAGGACTTTTCGATAGCAGCTAAGTTTGCCTTGGGCTTTGCTGTCGTG GCGATTGGTTTCTTTATCTACGGTTTTGCAGGCCAATTTGCCGTAAATGGTAAAAC CTCTTCTTGGGTAATGATTTGGGGTTACGCTTCTTACTCTTTGGGGGAGTTACTAG TAAGTGGCCTTGGTTTAGCGATGATTGCCCGCTACGTGCCTGCACGTATGGGCG GCTTTATGATGGGCGCGTATTTCGTGGCATCGGGCATTTCCCAGTACTTAGGCGG CGTTGTGGCCAACTTTGCTAGCGTGCCACAGGACTTAGTCGATCCACTGCAAACA CTCCCTGTTTATACCAATCTATTCAATAAATTAGGGGTTGCGGCCGTGGTTTGTAC CATTATCGCTCTCGCCGTATTACCGCTGATGCGCCGTTTGACAGAAAGCCATCAT GCCCATAGCAGTATTAATATATTTGAAATGTTACGTATAGATGAACGTCTTAGACTT AAAATCTATAAAGACACAGAAGGCTATTACACTATTGGCATCGGTCATTTGCTTAC AAAAAGTCCATCACTTAATGCTGCTAAATCTGAATTAGATAAAGCTATTGGGCGTA ATACTAATGGTGTAATTACAAAAGATGAGGCTGAAAAACTCTTTAATCAGGATGTT GATGCTGCTGTTCGCGGAATTCTGAGAAATGCTAAATTAAAACCGGTTTATGATTC TCTTGATGCGGTTCGTCGCGCTGCATTGATTAATATGGTTTTCCAAATGGGAGAAA CCGGTGTGGCAGGATTTACTAACTCTTTACGTATGCTTCAACAAAAACGCTGGGA TGAAGCAGCAGTTAACTTAGCTAAAAGTATATGGTATAATCAAACACCTAATCGCG CAAAACGAGTCATTACAACGTTTAGAACTGGCACTTGGGACGCG
Primer forward (pNIC-CTHF)	TTA AGA AGG AGA TAT ACT ATG ACT CTT GGT ACC AAT CAG GTA AGC
Primer reverse (pNIC-CTHF)	GAT TGG AAG TAG AGG TTC TCT GCC GCG TCC CAA GTG CCA GTT

5.5. Buffers

Lysis buffer I

20 mM	NaP, pH 7.5
300 mM	NaCl
1 mg/ml	lysozyme
5 U/ml	DNaseI
1	Complete Protease Inhibitor Cocktail tablet (EDTA free) (Rocke Applied Science) per 100 ml lysis buffer

Lysis buffer II

20 mM	NaP, pH 7.5
300 mM	NaCl
1 mg/ml	lysozyme
5 U/ml	DNaseI
15 mM	imidazol
5 %	glycerol
0.5 mM	TCEP
1	Complete Protease Inhibitor Cocktail tablet (EDTA free) (Rocke Applied Science) per 100 ml lysis buffer

Solubilization buffer I

20 mM	NaP, pH 7.5
300 mM	NaCl
1	Complete Protease Inhibitor Cocktail tablet (EDTA free) (Rocke Applied Science) per 100 ml solubilization buffer

Solubilization buffer II

20 mM	NaP, pH 7.5
300 mM	NaCl
15 mM	imidazol
5 %	glycerol
0.5 mM	TCEP
1 %	DDM
1	Complete Protease Inhibitor Cocktail tablet (EDTA free) (Rocke Applied Science) per 100 ml solubilization buffer

Wash buffer I

20 mM	NaP, pH 7.5
300 mM	NaCl
15 mM	imidazol
5 %	glycerol
0.5 mM	TCEP
0.03 %	DDM

Wash buffer II

20 mM	NaP, pH 7.5
300 mM	NaCl
30 mM	imidazol
5 %	glycerol
0.5 mM	TCEP
0.03 %	DDM

Elution buffer

20 mM	NaP, pH 7.5
200 mM	NaCl
300 mM	imidazol
5 %	glycerol
0.5 mM	TCEP
0.03 %	DDM

Gel filtration buffer

20 mM	HEPES, pH 7.5
150 mM	NaCl
5 %	glycerol
0.5 mM	TCEP
0.03 %	DDM

5.6. Optimization screens for crystallization

Table 25: Optimization screen 1 for the co-crystallization of PepT$_{So2}$-BRIL-insert with peptides.

		PEG 350 MME 25 % → 35 %				PEG 400 25 % → 35 %				PEG 550 MME 25 % → 35 %			
0.1 M HEPES		1	2	3	4	5	6	7	8	9	10	11	12
pH 6.5	A	No salt	0.04 M sodium chloride	0.1 M sodium chloride	0.2 M cadmium chloride	No salt	0.04 M sodium chloride	0.1 M sodium chloride	0.2 M cadmium chloride	No salt	0.04 M sodium chloride	0.1 M sodium chloride	0.2 M cadmium chloride
↓	B												
	C												
	D												
	E												
	F												
	G												
pH 8.0	H												

Table 26: Optimization screen 2 for the crystallization of PepT$_{So2}$-BRIL-Cterm with peptides.

			PEG 350 MME 25 % → 35 %				PEG 400 25 % → 35 %				PEG 2000 MME 25 % → 35 %			
	pH		1	2	3	4	5	6	7	8	9	10	11	12
0.1 M MES	5	A	No salt	0.2 M $CaCl_2$	0.001 M $CdCl_2$ + 0.03 M $MgCl_2$	0.0009 M $CdCl_2$ + 0.027 M $MgCl_2$	No salt	0.2 M $CaCl_2$	0.001 M $CdCl_2$ + 0.03 M $MgCl_2$	0.0009 M $CdCl_2$ + 0.027 M $MgCl_2$	No salt	0.2 M $CaCl_2$	0.001 M $CdCl_2$ + 0.03 MMgCl_2	0.03 M $MgCl_2$
	6	B												
	6.5	C												
	7	D												
0.1 M Tris	7.5	E	No salt	0.005 M $MgCl_2$	0.1 M NaCl	0.005 M $MgCl_2$ + 0.1 M NaCl	No salt	0.005 M $MgCl_2$	0.1 M NaCl	0.005 M $MgCl_2$ + 0.1 M NaCl	No salt	0.005 M $MgCl_2$	0.1 M NaCl	0.005 M $MgCl_2$ + 0.1 M NaCl
	↓	F												
		G												
	9	H												