

RESEARCH ARTICLE

# Structure determination of a major facilitator peptide transporter: Inward facing PepT<sub>St</sub> from *Streptococcus thermophilus* crystallized in space group P3<sub>1</sub>21

Esben M. Quistgaard<sup>1,2</sup>, Maria Martinez Molledo<sup>1</sup>, Christian Löw<sup>1,2\*</sup>

**1** Centre for Structural Systems Biology (CSSB), DESY and European Molecular Biology Laboratory Hamburg, Hamburg, Germany, **2** Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden

\* [christian.loew@embl-hamburg.de](mailto:christian.loew@embl-hamburg.de)



**OPEN ACCESS**

**Citation:** Quistgaard EM, Martinez Molledo M, Löw C (2017) Structure determination of a major facilitator peptide transporter: Inward facing PepT<sub>St</sub> from *Streptococcus thermophilus* crystallized in space group P3<sub>1</sub>21. PLoS ONE 12(3): e0173126. doi:10.1371/journal.pone.0173126

**Editor:** Dimitrios Fotiadis, University of Bern, SWITZERLAND

**Received:** December 15, 2016

**Accepted:** February 15, 2017

**Published:** March 6, 2017

**Copyright:** © 2017 Quistgaard et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files. The atomic coordinates and structure factors for inward open PepT<sub>St</sub> crystallized in space group P3<sub>1</sub>21 have been deposited in the Protein Data Bank with accession number PDB: 5MMT.

**Funding:** This research was funded by the European Union's Seventh Framework Programme (FP7/2007-2013) under BioStruct-X (grant agreement N°783), by the Horizon 2020 programme of the European Union (iNEXT grant,

## Abstract

Major facilitator superfamily (MFS) peptide transporters (typically referred to as PepT, POT or PTR transporters) mediate the uptake of di- and tripeptides, and so play an important dietary role in many organisms. In recent years, a better understanding of the molecular basis for this process has emerged, which is in large part due to a steep increase in structural information. Yet, the conformational transitions underlying the transport mechanism are still not fully understood, and additional data is therefore needed. Here we report in detail the detergent screening, crystallization, experimental MIRAS phasing, and refinement of the peptide transporter PepT<sub>St</sub> from *Streptococcus thermophilus*. The space group is P3<sub>1</sub>21, and the protein is crystallized in a monomeric inward facing form. The binding site is likely to be somewhat occluded, as the lobe encompassing transmembrane helices 10 and 11 is markedly bent towards the central pore of the protein, but the extent of this potential occlusion could not be determined due to disorder at the apex of the lobe. Based on structural comparisons with the seven previously determined P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> and C222<sub>1</sub> structures of inward facing PepT<sub>St</sub>, the structural flexibility as well as the conformational changes mediating transition between the inward open and inward facing occluded states are discussed. In conclusion, this report improves our understanding of the structure and conformational cycle of PepT<sub>St</sub>, and can furthermore serve as a case study, which may aid in supporting future structure determinations of additional MFS transporters or other integral membrane proteins.

## Introduction

Cellular uptake of dietary di- and tripeptides represents an important means of acquiring nitrogen and amino acids in many organisms [1–4]. In bacteria, this task is carried out by both ATP-binding cassette (ABC) transporters and major facilitator superfamily (MFS)

project no. 653706) and the Swedish Research Council (grant number 621-2013-5905).

**Competing interests:** The authors have declared that no competing interests exist.

transporters of the peptide transporter (PepT or PTR) family, also commonly referred to as the proton-dependent oligopeptide transporter (POT) family [5]. In contrast, humans use only PepTs for this task [5], of which the two best studied are the intestinal uptake transporter PepT1 and the renal reabsorption transporter PepT2 [6–10]. Notably, these proteins do not only transport dietary peptides, but also a variety of peptidomimetic drugs and amino acid-conjugated pro-drugs, and are therefore not only physiologically important, but also of great pharmacological relevance [11,12].

PepTs are believed to function by an alternate access mechanism involving gated transitions between inward open, occluded and outward open conformational states [13], which is essentially similar to that of any other MFS transporter [14,15]. Due to the inherent difficulty in crystallization and structure determination of integral membrane proteins, the first PepT crystal structure, which represents PepT<sub>So</sub> from *Shewanella oneidensis*, did not appear before 2011 [16]. However, in correlation with a general upsurge in MFS structure determinations [14,15], several additional structures soon followed [17–26]. Currently, structures are known for a total of six different bacterial PepTs, which were curiously all captured in inward facing conformations: PepT<sub>So</sub> and PepT<sub>So2</sub> from *Shewanella oneidensis*, PepT<sub>St</sub> from *Streptococcus thermophilus*, GkPOT from *Geobacillus kaustophilus*, YbgH from *Escherichia coli*, and YePEPT from *Yersinia enterocolitica* [17–26]. Although much still remains to be understood, these structures have aided greatly in understanding substrate recognition and key aspects of the transport mechanism.

We set out to determine a structure of PepT<sub>St</sub> before any other PepT structures were known. However, while our work was ongoing, several PepT structures were presented [17–26] including structures of inward facing PepT<sub>St</sub> crystallized in space groups P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> and C222<sub>1</sub> [17,21,25,26]. The structure presented here is likewise inward facing, but belongs instead to space group P3<sub>1</sub>21. We describe in detail how this structure was obtained, in order to provide a case report that may be of use to others who wish to pursue structure determinations of MFS transporters. In addition, we compare all three crystal forms with the aim of furthering our understanding of the structural plasticity of PepT<sub>St</sub> and its role in enabling the conformational cycle of the protein.

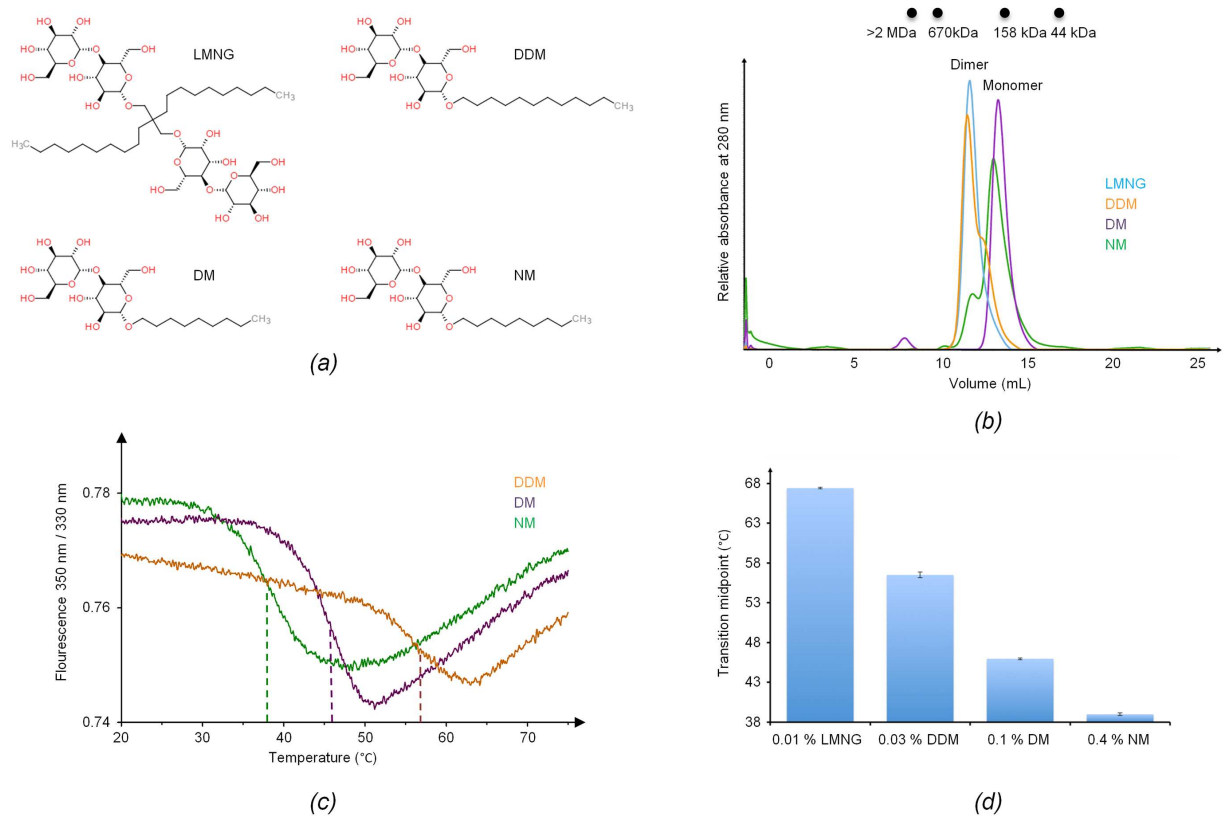
## Materials and methods

### Materials

The detergents used for purification were from Anatrace (Maumee, OH, USA), terrific broth (TB) was from Formedium (Norfolk, UK), kanamycin was from Duchefa Biochemie (Haarlem, NL), isopropyl β-D-1-thiogalactopyranoside (IPTG) was from Saveen Werner (Limhamn, S), and crystallization reagents were from Qiagen (Germantown, MD, USA). All other chemicals were of analytical grade and obtained from Sigma-Aldrich (St. Louis, MO, USA), unless otherwise stated.

### Expression and purification

We have previously reported the identification of PepT<sub>St</sub> as a suitable target for structural studies using a pipeline approach (referred to as Ce27 in that study) [27]. Here, we used the same C-terminally hexahistidine-tagged wild-type PepT<sub>St</sub> construct, and furthermore prepared two mutant variants, L226M and F338M, using standard Quikchange protocols. Protein expression in *Escherichia coli* C41, membrane solubilization and purification using IMAC and size exclusion chromatography (SEC), was carried out as described [27], except that a partially different set of detergents were used. It may be noted that the protein was eluted from the IMAC column by passing TEV protease over the beads, as the usage of high imidazole concentrations



**Fig 1. Detergent screening.** (a) Chemical structures of the four different detergents used for purification: LMNG (Lauryl Maltose Neopentyl Glycol), DDM (n-Dodecyl-β-D-Maltopyranoside), DM (n-Decyl-β-D-Maltopyranoside), and NM (n-Nonyl-β-D-Maltopyranoside). (b) Size exclusion chromatography (SEC) profiles for PepT<sub>St</sub> purified in LMNG (blue), DDM (orange), DM (purple), and NM (green). (c) Differential scanning fluorimetry results. Thermal stability curves of the protein in the presence of the detergents NM, DM and DDM are shown (same color code as in panel b). The fluorescence ratio (F<sub>350nm</sub>/F<sub>330nm</sub>) is represented as a function of temperature. Dashed lines indicate the melting temperatures (T<sub>m</sub>). (d) Comparison of the T<sub>m</sub> values for PepT<sub>St</sub> purified in different detergents. The error bars were calculated from three independent measurements.

doi:10.1371/journal.pone.0173126.g001

for the elution step were found to strongly destabilize the protein (Löw et al., 2013). The detergents used were four different alkyl maltosides: n-dodecyl-β-D-maltopyranoside (DDM), n-decyl-β-D-maltopyranoside (DM), n-nonyl-β-D-maltopyranoside (NM), and the recently developed lauryl maltose neopentyl glycol (LMNG) [28] (Fig 1a). For membrane solubilization, the detergent concentration was 1%. For purification, we used 0.03% of DDM, 0.2% of DM, 0.4% of NM, or 0.01% of LMNG. For each purification campaign, the same detergent was used from solubilization to the final SEC step, except for NM. For this detergent, PepT<sub>St</sub> was thus purified in DM up till the IMAC step, after which the detergent was exchanged to NM in the SEC step. For SeMet labeling, the protein was expressed in minimal medium supplemented with SeMet according to established protocols [29].

### Protein stability measurements

The stability of the different purified protein preparations was followed using a nanoDSF differential scanning fluorimeter (NanoTemper Technologies, GmbH). Here the intrinsic fluorescence at 330 and 350 nm after excitation at 280 nm is used to monitor the fluorescence change upon heat unfolding. Up to 48 samples can be measured in parallel without the addition of a dye. Typically, 10 μL of protein solution at a concentration of 0.5 mg/mL

**Table 1. Crystallization condition.**

Method	Sitting-drop vapor diffusion
Temperature (K)	277
Protein concentration	8.5 mg/mL
Buffer composition of protein solution	20 mM HEPES pH 7.5, 150 mM NaCl, 5% glycerol, 0.4% NM, 0.75 mM TCEP, 10 mM Penicillin G
Composition of reservoir solution	0.05 M HEPES pH 8.0, 30% PEG550 MME, 1.2% Fos-Choline 10

doi:10.1371/journal.pone.0173126.t001

was loaded in a capillary (in triplicates), and the unfolding was then measured at a heating rate of 1°C/min. The first derivative of the unfolding curves was used to determine the transition midpoint. As control experiment, the concentration dependence of the transition midpoint was determined in the range of 0.2–5 mg/mL for each protein batch. Resulting transition midpoints were within 1°C.

### Crystallization procedures

Crystallization trials were carried out by vapor diffusion in 96-well sitting-drop plates at both 277 K and 293 K. The volume of crystallant added to the reservoir was 50 µL, while the drops had a total volume of 300 nL and were composed of the PepT<sub>St</sub> sample in a concentration of 5–10 mg/mL, and the crystallant in ratios of 1:2, 1:1 or 2:1. Initial crystal hits were readily obtained regardless of whether the protein had been purified in DDM, DM or NM, but the best diffraction was attained when using NM, a temperature of 277 K, and a crystallant containing small PEGs as well as a buffer at near neutral to slightly alkaline pH (Table 1). Shifting to a 24-well format and scaling up the drop size was not found beneficial, which is an experience that we have also had with other integral membrane proteins [19,30,31]. For optimization, a wide range of conventional additives e.g. divalent cations were tested, along with numerous secondary detergents. Here, Fos-Choline-10 was found to improve diffraction significantly, and was therefore included in the crystallant (Table 1). We also tested the addition of peptides and other putative substrates. However, none of these improved the diffraction, and none could be identified in the electron density maps. This also includes the penicillin G compound that was present in the protein sample, which gave rise to the crystal used for refinement (Table 1). The crystals were typically flash frozen in liquid nitrogen without prior soaking in cryobuffer.

### Data processing and experimental phasing

Four data sets were used for structure determination: a native one used for refinement, and another native as well as two derivatives used for MIRAS phasing (Table 2). All data sets were processed using the XDS package [32]. As the diffraction displayed severe anisotropy, the native data sets were furthermore subjected to anisotropic scaling and truncation using the Diffraction Anisotropy Server [33] (Table 2). Initial phasing was carried out using a trimmed version of the PepT<sub>So</sub> structure (PDB: 2XUT) as a molecular replacement search model in Phaser [34] from the PHENIX suite [35], but the map was not of high quality, and we therefore decided to pursue experimental phases. The derivatives used for phasing were a SeMet labeled F338M crystal and a wild-type crystal soaked with KAu(CN)<sub>2</sub>. Soaking was carried out by adding 1.5 µL of a solution containing 90% of the original crystallant and 1 mM KAu(CN)<sub>2</sub> to a crystallization drop with a well-formed crystal, which was then removed and flash frozen in liquid nitrogen after an incubation time of nine minutes. Scaling, site finding and initial

**Table 2. Data collection and processing.** The Values for the outer shell are given in parentheses. The native data sets were subjected to anisotropic scaling and truncation (see main text for details). Without truncation  $\langle I/\sigma(I) \rangle$  of the native data set used for refinement falls below 2.0 between a maximum resolution of 3.80 and 3.75 Å at an overall completeness of over 99%.

Crystal	Native 1 (for refinement)	Native 2 (for phasing)	Derivative 1 (L338M SeMet)	Derivative 2 (KAu(CN) <sub>2</sub> )
<b>Diffraction source</b>	<b>Diamond I04</b>	<b>Diamond I24</b>	<b>Diamond I24</b>	<b>Diamond I02</b>
Wavelength (Å)	0.9795	1.0000	0.9788	0.8920
Temperature (K)	100	100	100	100
Space group	P3 <sub>1</sub> 2 <sub>1</sub>	P3 <sub>1</sub> 2 <sub>1</sub>	P3 <sub>1</sub> 2 <sub>1</sub>	P3 <sub>1</sub> 2 <sub>1</sub>
<i>a, b, c</i> (Å)	80.19, 80.19, 293.69	79.51, 79.51, 293.64	81.37, 81.37, 294.57	81.10, 81.10, 293.25
$\alpha, \beta, \gamma$ (°)	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120
Resolution range (Å)	44.85–3.40 (3.49–3.40)	44.68–3.60 (3.69–3.60)	45.20–4.10 (4.21–4.10)	48.88–4.20 (4.31–4.20)
Total no. of reflections	83693 (601)	148684 (6829)	179388 (13126)	185686 (13620)
Completeness (%)	74.6 (5.8)	97.1 (65.1)	99.6 (99.4)	99.2 (98.6)
Redundancy	5.3 (0.5)	11.2 (7.2)	10.5 (10.5)	11.8 (11.4)
$\langle I/\sigma(I) \rangle$	13.67 (3.13)	19.40 (2.49)	17.75 (2.95)	12.31 (3.35)
<i>R</i> <sub>meas</sub>	8.9 (67.9)	7.2 (106.5)	7.1 (124.9)	8.5 (104.4)
Wilson <i>B</i> -factor (Å <sup>2</sup> )	96.4	129.16	156.5	174.2

doi:10.1371/journal.pone.0173126.t002

MIRAS phasing was carried out with AutoSHARP [36], while final sub-structure refinement, phasing and density modification was carried out with SHARP [37] (Table 3). We used an initial resolution range of 48.94–4.10 Å, corresponding to the full range of the SeMet data set, and also close to that of the KAu(CN)<sub>2</sub> data set. This was however extended to the full range of the native data set, 48.94–3.60 Å, during density modification. FOM was 0.401 for acentric reflections and 0.258 for centric reflections before density modification, and respectively 0.281 and 0.204, after.

### Refinement and structural analysis

Model building and rebuilding was carried out in Coot [38]. We initially relied only on the experimental map, but later also used the P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> and C222<sub>1</sub> structures for guidance as they became available (Solcan *et al.*, 2012; Lyons *et al.*, 2014). Refinement was carried out with PHENIX refine [35] using one translation libration screw (TLS) group (Table 4). In order to ensure good geometry of the model, both secondary structure and reference restraints were used. This was found to greatly improve the Ramachandran plot and to reduce the number of rotamer outliers. For the last rounds of refinement, the X-ray/stereochemistry and X-ray/ADP weights were furthermore optimized, which helped improve RMS bonds and angles, as well as the Molprobit clash score [39]. The final clash score was 6.8 and the overall Molprobit score was 1.4 (both are in the 100<sup>th</sup> percentile of structures with similar resolution). Structural

**Table 3. MIRAS phasing statistics.**

Derivative	SeMet		KAu(CN) <sub>2</sub>	
Resolution (Å)	48.94–4.10		48.94–4.10	
Number of sites found	20		4	
Type of reflections	Acentric	Centric	Acentric	Centric
Phasing power (isomorphous)	0.801	0.700	1.312	1.859
Phasing power (anomalous)	1.784	0.000	0.325	0.000
<i>R</i> -Cullis (isomorphous)	0.403	0.478	0.200	0.159
<i>R</i> -Cullis (anomalous)	0.624	0.000	0.987	0.000

doi:10.1371/journal.pone.0173126.t003

**Table 4. Refinement statistics.**

Resolution range (Å)	44.85–3.40
No. of reflections, working set	11292
No. of reflections, test set	589
Final $R_{\text{cryst}}$	0.265
Final $R_{\text{free}}$	0.286
No. of non-H atoms	
Protein	3283
Other	0
R.m.s. deviations	
Bonds (Å)	0.005
Angles (°)	1.188
Average $B$ factors (Å <sup>2</sup> )	127.1
Ramachandran plot	
Most favoured (%)	99.3
Allowed (%)	0.7

doi:10.1371/journal.pone.0173126.t004

alignments were made using the DALI server ([http://ekhidna.biocenter.helsinki.fi/dali\\_server/start](http://ekhidna.biocenter.helsinki.fi/dali_server/start)) [40], and structure figures were generated using PyMol [41].

## Results and discussion

### Characterization of the oligomeric state and thermal stability in different detergents

In a previous high-throughput screening study, we found that maltoside detergents are suitable for purification of PepT<sub>St</sub> [27]. Here we screened four maltosides with aliphatic chains of different lengths: DDM, DM and NM, which have chain-lengths of twelve, ten and nine, respectively, as well as LMNG, which chemically represents a fusion of two DDM molecules (Fig 1a). These detergents all supported good purification yields and did not cause significant aggregation, as evaluated by SEC (Fig 1b). Using blue native PAGE and chemical cross-linking, we have previously shown that PepT<sub>St</sub> forms dimers in DDM [19]. The SEC results provided here show that LMNG also supports the dimeric form, while the shorter-chain detergents DM and NM disrupt it (Fig 1b). This is not a unique case. The higher oligomeric forms of rhodopsin were thus likewise found to be disrupted when substituting long-chain maltoside detergents with shorter-chain variants [42]. We next measured the thermal stability of PepT<sub>St</sub>, using intrinsic fluorescence measurements. Protein stability generally decreases upon shortening of the alkyl chain of the detergent used for purification, and this was also found to be the case for PepT<sub>St</sub>. Indeed, the effect was quite dramatic, with transition midpoints for thermal unfolding spanning a temperature range of 39°C in NM to almost 69°C in LMNG (Fig 1c and 1d). Finally, we also analyzed the thermal stabilities of SeMet labeled variants of the wild-type protein and two mutants, L226M and F338M, and found them to be very similar to unlabeled wild-type PepT<sub>St</sub> (data not shown).

### Obtaining the P3<sub>1</sub>21 crystal form

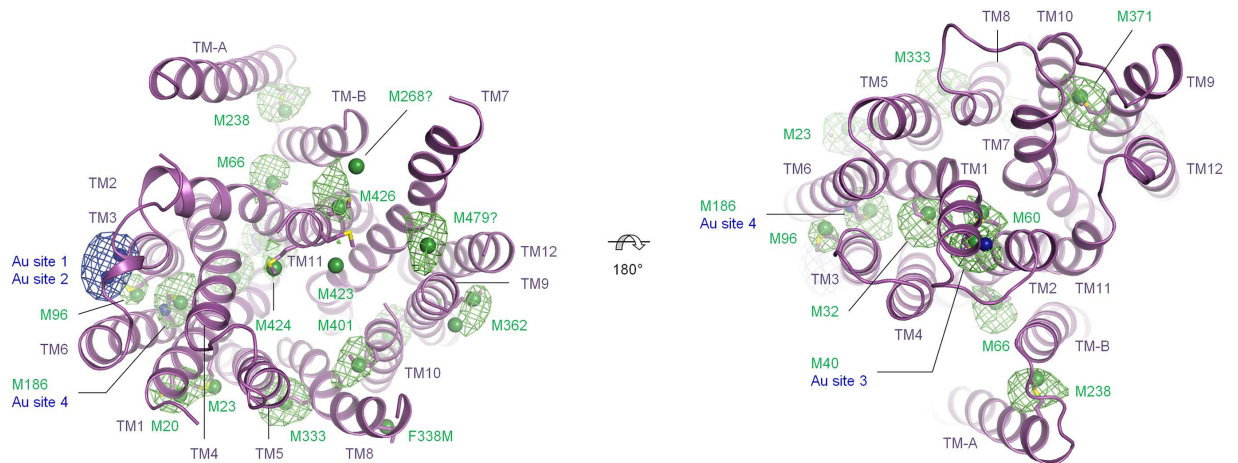
Crystallization trials were carried out for protein purified in DDM, DM and NM, and yielded crystal hits in all three cases. We obtained the best initial diffraction when using NM, though others have obtained similar resolution for a different crystal form, P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub>, which was grown using DDM [17]. As mentioned above, NM is highly destabilizing compared to the longer-

chain maltoside detergents, and it may therefore seem counterintuitive that it yields good crystals. However, although the stabilizing effect of detergents generally decreases with shorter chain-length, the micelle size also decreases, which can potentially support tighter crystal packing in some directions [43]. Indeed, a recent survey of the alpha-helical membrane protein structures in the Protein Data Bank (PDB) confirms that the alkyl chain-length of the detergent used for crystallization is on average inversely correlated with diffraction potential [44]. It is therefore generally advisable to test short-chain detergents for crystallization, even when they are highly destabilizing compared to longer-chain variants, provided that they do not cause significant aggregation during purification or concentration of the purified sample. The use of secondary detergents for purification or as crystallization additives can have a crucial effect on diffraction quality, as is for example well illustrated by the case of the structure determination of cc3b cytochrome c oxidase [45], and their use is therefore becoming increasingly common [44]. For this reason we tested numerous detergent additives, and found that Fos-Choline-10 significantly improved the diffraction potential.

## Structure determination

Native crystals diffracted to 3.4 Å maximum resolution, but in a severely anisotropic fashion. The data were therefore truncated and scaled anisotropically. As cut-off values for ellipsoidal truncation, we used the values suggested by the Diffraction Anisotropy Server [33], which were 3.9 Å in the *a*-direction, 3.7 Å in the *b*-direction and 3.4 Å in the best diffracting *c*-direction of the crystal. The resolution is similar to what was obtained for the P<sub>2</sub><sub>1</sub>2<sub>1</sub>2<sub>1</sub> form (3.3 Å anisotropic) [17], while the C222<sub>1</sub> form diffracted considerably better (up to 2.3 Å isotropic) [21,25]. At the time the native crystals were optimized, the first structure of PepT<sub>So</sub> had recently been published, and was therefore tested as search model for molecular replacement. A solution could indeed be found, which was essentially correct, as evidenced by anomalous difference Fourier maps calculated using preliminary SeMet data. However, as the map was of poor quality, and as model bias can be severe at the resolution we were able to attain, we decided to move on to experimental phasing using SeMet and heavy metal derivatives. In order to ensure accurate sequence assignment, we originally planned to generate a series of methionine mutants, and started out by introducing L226M and F338M point mutations in regions that were poorly defined in preliminary maps. However, when the P<sub>2</sub><sub>1</sub>2<sub>1</sub>2<sub>1</sub> and C222<sub>1</sub> PepT<sub>St</sub> structures were published, sequence assignment became a trivial matter and the strategy was therefore abandoned. Nonetheless, the best SeMet data set was obtained using the F338M mutant (Table 2). For heavy metal derivatization, numerous Hg, Au, Pt, Pb and lanthanide salts were tested in both co-crystallization and soaking experiments. However, we were only successful when soaking with KAu(CN)<sub>2</sub>, which, incidentally, is one of the most commonly used heavy metal salts for phasing of both soluble and membrane imbedded proteins [46].

AutoSHARP [36] was found to exhibit impressive sensitivity with respect to site finding. SeMet sites were thus identified for all methionines modeled in the structure, M20, M23, M32, M40, M60, M66, M96, M186, M238, M333, F338M, M362 (two sites), M371, M401, M423 and M424, as well as for what appears to be the unmodeled methionines, M268 and M479 (Fig 2). AutoSHARP uses log-likelihood gradient maps in the site finding procedure, which are more sensitive than anomalous difference Fourier maps. Indeed, for some of the identified sites, no difference Fourier peak could be detected at a threshold of 3  $\sigma$  in maps calculated using model phases—even when these model phases were derived from the final refined structure (Fig 2). We suggest bearing this in mind for other projects where SeMet is to be used for chain tracing. As for the Au sites, two of them share a substantial anomalous difference peak in a



**Fig 2. Site finding.** The structure is magenta and the methionine side chains are shown in sticks and labeled, green and blue spheres designate AutoSHARP SeMet and Au sites, respectively, and the similarly colored wire meshes designate the model phased SeMet and Au anomalous difference Fourier maps contoured at  $3\sigma$ . Two views are shown: cytoplasmic (left) and periplasmic (right).

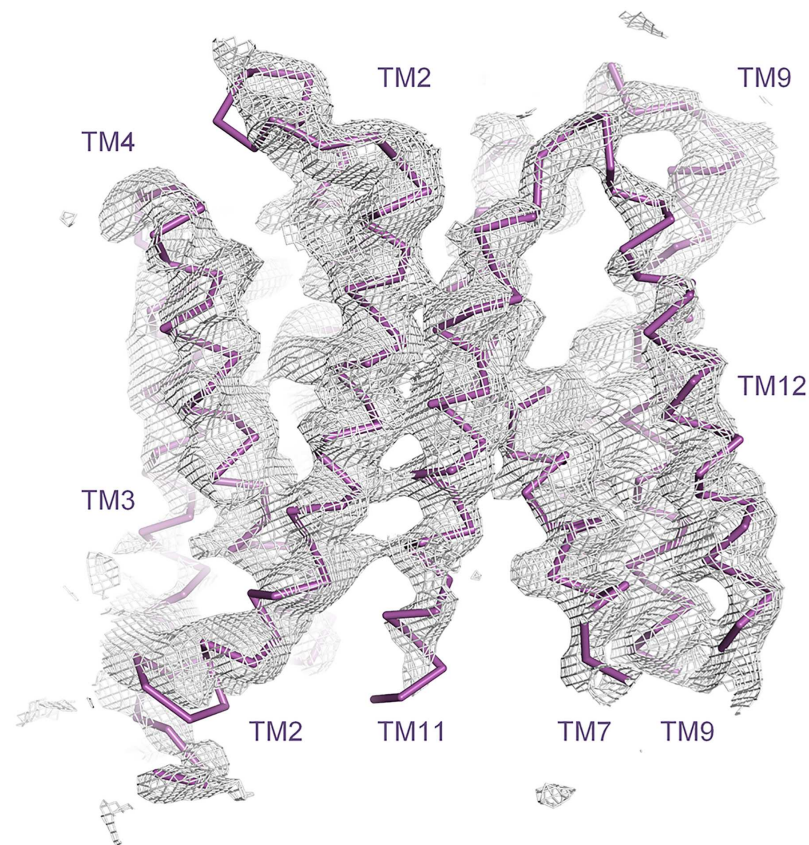
doi:10.1371/journal.pone.0173126.g002

hydrophobic pocket located between the cytoplasmic ends of transmembrane helices 3 and 6 (TM3 and TM6), and two more are found near M40 and M186, respectively, where, in contrast, very little difference density is observed (Fig 2). The SeMet and Au data sets were used together with an appropriate native data set to generate MIRAS phases in SHARP (Table 3) [37]. The transmembrane helices were generally well defined in the experimental map (Fig 3). There are in total fourteen of them, which are arranged in the same overall pattern as described previously [17]. Briefly, TM1–TM6 forms an N-terminal MFS domain, termed the N-domain, TM7–TM12 forms a similar C-terminal MFS domain, termed the C-domain, and the last two helices, TM-A and TM-B, are found in the linker between the two domains, where they adopt a V-shaped arrangement (Figs 2 and 3). As mentioned above, an anisotropically corrected data set was used for refinement. This was found to perform well (Table 4), in spite of exhibiting low completeness as measured within circular resolution shells (Table 2).

### Crystal packing

The packing modes in the three crystal forms are very different from each other (Fig 4). However, a common feature of the P3<sub>1</sub>21 and P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> forms is that they both exhibit ordered end-to-end packing through periplasmic-periplasmic and cytoplasmic-cytoplasmic interactions, while ordered lateral contacts are entirely missing (Fig 4a and 4b). Loose lateral packing is a typical feature of membrane protein crystals grown using traditional ‘*in surfo*’ vapor diffusion methods and relates to the sometimes prohibitive effect of the detergent belt on forming tight contacts [47]. A total lack of ordered lateral contacts is rather extreme, but not unheard of. Similar cases have thus been reported for other membrane proteins, such as the MFS transporter Xyle and the porin OmpF [48,49]. Notably, a lack of ordered lateral contacts typically manifests itself during data collection as severe diffraction anisotropy [48,50]. As mentioned above, such severe anisotropy was indeed also observed for the P3<sub>1</sub>21 and P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> forms of PepT<sub>St</sub>. Moreover, in both cases the directions exhibiting the most dramatic fallout of reflection intensity, do indeed correlate with the directions in the crystals where ordered packing interactions are missing (data not shown). The C222<sub>1</sub> form stands out by encompassing tight lateral interactions, causing the formation of continuous layers, which pack against each other in an overlapping fashion through head-to-tail interactions between periplasmic and



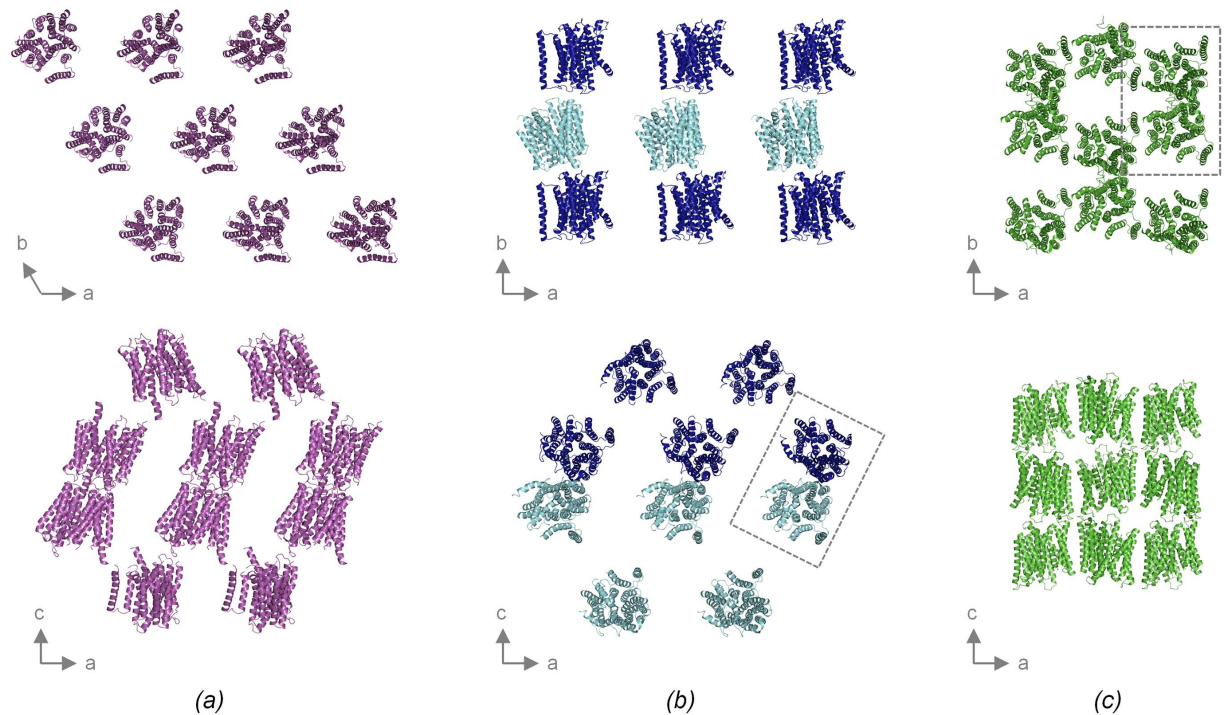


**Fig 3. Experimental electron density map.** The structure is shown in magenta, and the density modified MIRAS phased electron density map is shown at 1  $\sigma$  as a grey wire mesh.

doi:10.1371/journal.pone.0173126.g003

cytoplasmic parts (Fig 4c). This results in a quite compact packing pattern, which no doubt underlies the much better resolution and more isotropic diffraction of this crystal form. Tight packing is possible here, since crystallization was performed *in meso*, causing the detergent belt to be replaced with lipids, which better mimic the plasma membrane, and supports rather than inhibits packing interactions between transmembrane parts [51].

The packing interfaces in the three crystal forms are almost invariably different and are also formed by somewhat different subsets of structural elements (S1 Fig). The P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> and C222<sub>1</sub> forms do however share a single interface, though it is non-crystallographic rather than crystallographic in the former case. This interface is mainly formed by TM1, TM5 and TM6 (S1 Fig), and generates a specific side-by-side antiparallel dimer that is not seen in the P3<sub>1</sub>21 form (Fig 4). As mentioned previously, PepT<sub>St</sub> forms dimers in DDM, which was used for purification in the cases of the C222<sub>1</sub> and P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> forms, but not in NM, which was used for the P3<sub>1</sub>21 form (Fig 1b). It thus seems likely that the dimer seen in the crystal is the same as observed in solution. Some PepTs probably function as homooligomers [19], but so far, no attempts have been made to characterize the oligomeric state of PepT<sub>St</sub> in the membrane, and it is therefore unknown if a monomeric or oligomeric form is preferred *in vivo*. Yet, we find it rather unlikely that this particular dimer is physiologically relevant. An antiparallel arrangement with one protomer being flipped in the membrane to expose the cytoplasmic face to the extracellular side and the extracellular face to the cytoplasmic side does thus not seem likely to be functionally productive. Moreover, we are not aware of any functional antiparallel MFS



**Fig 4. Comparisons of crystal packing.** (a) Crystal packing in the P3<sub>1</sub>2<sub>1</sub> form. Top: slice of the crystal taken parallel to the (a-b)-plane and encompassing one layer of molecules. The PepT<sub>St</sub> molecules are colored magenta. Bottom: slice of the crystal taken parallel to the (b-c)-plane. (b) Crystal packing in the P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> form (PDB: 4APS). Depicted as in panel a, except that the two molecules in the asymmetric unit are colored light and dark blue, and that the dimer formed by these are boxed. (c) Crystal packing in the C222<sub>1</sub> form (PDB: 4D2C). Depicted as in panel a, except that the molecules are colored green, and that a dimer with the same arrangement as seen in the P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> form is similarly boxed, though this dimer is generated by crystal symmetry rather than being non-crystallographic. Crystal packing interfaces are shown in S1 Fig.

doi:10.1371/journal.pone.0173126.g004

dimers having ever been reported. However, regardless of the lack of an obvious biological relevance, the formation of the dimer may facilitate crystal nucleation, as homo oligomers tend to crystallize more readily than monomers [52], and would also in part dictate the overall packing pattern. In fact, it is likely that the difference in crystal packing between the three crystal forms are to a significant extent governed by the differences in oligomerization combined with the above mentioned differences in crystallization methods and their associated differential effects on packing between the transmembrane regions.

### Structural variation and flexibility

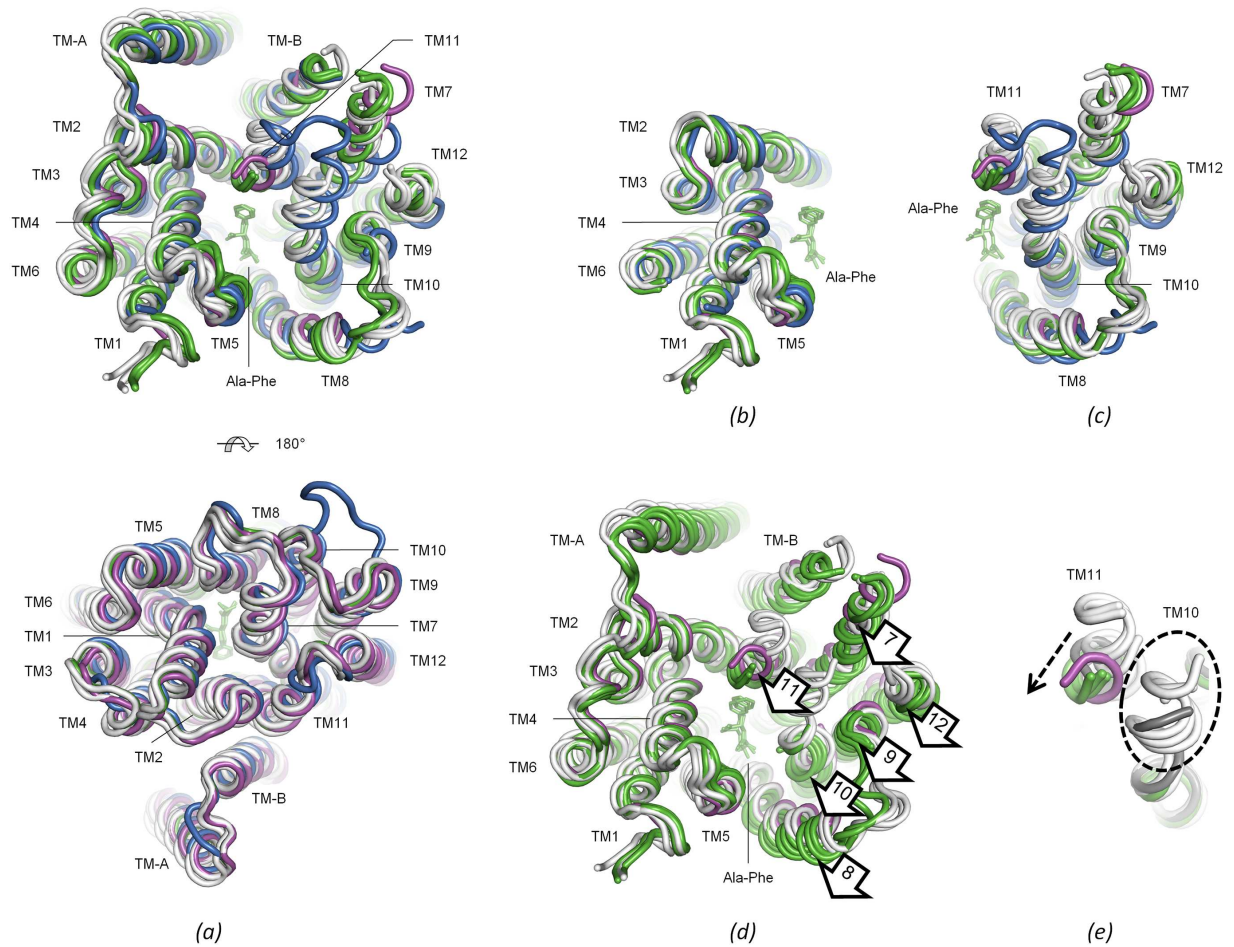
The P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> structure can be superimposed on the P3<sub>1</sub>2<sub>1</sub> structure with an RMSD value of 1.4 Å over 409 out of 425 C-alpha atoms (Table 5). For the individual domains, the values are 1.0 Å over 190 out of 190 C-alpha atoms for the N-domain, and 1.7 Å over 160 out of 174 C-alpha atoms for the C-domain (Table 5). The C-domain thus varies more than the N-domain between these two structures. The most obvious structural differences are in loop regions. For example, the loop between TM9 and TM10 differs markedly between the two structures, and the lobe encompassing the cytoplasmic tips of TM10 and TM11 as well as the connector loop between them is much better ordered in the P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> structure than in the P3<sub>1</sub>2<sub>1</sub> structure (Fig 5a–5c). However, there are also some smaller differences in position or bending of TM7, TM9 and TM12 in the C-domain (Fig 5c).

**Table 5. Overview of previously determined PepT<sub>St</sub> structures and their structural similarity to the P3<sub>1</sub>21 structure reported in this study.**

Crystal form	PDB	Ligand modeled	Data collection	Superimposition on the P3 <sub>1</sub> 21 structure					
				Full-length		N-domain		C-domain	
				RMSD (Å)	No. C-alphas	RMSD (Å)	No. C-alpha	RMSD (Å)	No. C-alphas
P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	4APS	None	Loop, 100 K	1.4	409	1.0	190	1.7	160
C222 <sub>1</sub>	4D2C	Dipeptide	Loop, 100 K	0.7	413	0.7	190	0.5	162
C222 <sub>1</sub>	5D58	Dipeptide	Plate, 100 K	0.9	423	0.7	190	0.7	172
C222 <sub>1</sub>	5D59	Dipeptide	Loop, 100 K	0.9	423	0.7	190	0.6	172
C222 <sub>1</sub>	4D2D	Tripeptide	Loop, 100 K	0.9	424	0.6	190	1.1	173
C222 <sub>1</sub>	4D2B	None	Loop, 100 K	1.0	420	0.6	190	1.4	170
C222 <sub>1</sub>	4XNI	None	Plate, 293 K	0.8	417	0.5	190	0.7	166
C222 <sub>1</sub>	4XNJ	None	Loop, 100 K	0.9	425	0.6	190	1.6	154

doi:10.1371/journal.pone.0173126.t005

Seven C222<sub>1</sub> structures have been published [21,25,26], which differ in which substrate were used for crystallization, in how the data were collected, and in which temperature was used during data collection (Table 5). They are all structurally rather similar, though as remarked by others [21], the structures with a dipeptide in the binding site differ somewhat from the other structures determined at cryogenic temperatures by exhibiting a slightly different position of the C-domain relative to the N-domain (Fig 5d), and a marked bending of lobe TM10–TM11 towards the bound substrate as well as increased disorder at its apex, permitting fewer residues to be modeled (Fig 5e). The single structure determined at room temperature was obtained without a dipeptide, but interestingly also displays some bending of TM11 (Fig 5e), as has likewise been remarked by others [25]. The P3<sub>1</sub>21 structure is more similar to any of the C222<sub>1</sub> structures than it is to the P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> structure. The full-length proteins can thus be superimposed with RMSD values ranging from 0.7–1.0 Å over 413–425 C-alpha atoms, the N-domains with values of 0.5–0.7 Å over 190 C-alpha atoms, and the C-domains with values of 0.5–1.6 Å over 154–173 C-alpha atoms (Table 5). The structural overlays reveal differences in loop TM3–TM4, loop TM4–TM5 and loop TM9–TM10, as well as in the bending of the cytoplasmic part of TM7 (Fig 5a–5b). Most of these differences could however be caused by packing effects, as all these structural elements are also implicated in crystal packing in one or both of the crystal forms (S1 Fig). Thus, although crystal contacts are not generally a major source of imprecision in crystal structures [53,54], smaller perturbations across such interfaces are not unlikely to occur [55,56]. Of more interest, the P3<sub>1</sub>21 structure exhibits a marked bending of TM11, which is highly reminiscent of the bending observed in the C222<sub>1</sub> structures with dipeptides (Fig 5e). Furthermore, the tip of TM10 is also bent in much the same way as in these structures, and is about equally disordered (Fig 5e). This is also reflected in the RMSD values (Table 5)—the C-domain of the P3<sub>1</sub>21 structure thus superimposes on the C222<sub>1</sub> structures with dipeptides with RMSD values of 0.5–0.7 Å over 162–172 C-alpha atoms, while it superimposes on the other C222<sub>1</sub> cryogenic structures with noticeably higher values of 1.1–1.6 Å over 154–173 C-alphas. On the other hand, the domain rotation observed in the C222<sub>1</sub> structures with dipeptide does not apply to the P3<sub>1</sub>21 structure (Fig 5d). The P3<sub>1</sub>21 structure thus adopts what can be viewed as a hybrid conformation, which shares the bending of lobe TM10–TM11 with the C222<sub>1</sub> structures with dipeptides, but the orientations of the domains relative to each other with the C222<sub>1</sub> cryo structures without dipeptides. Notably, the C222<sub>1</sub> structure determined at room temperature adopts a somewhat similar hybrid conformation, though lobe TM10–TM11 is less bent here (Fig 5e). Based on these observations, we infer that lobe TM10–TM11 can undergo bending independently of domain rotation, and that it

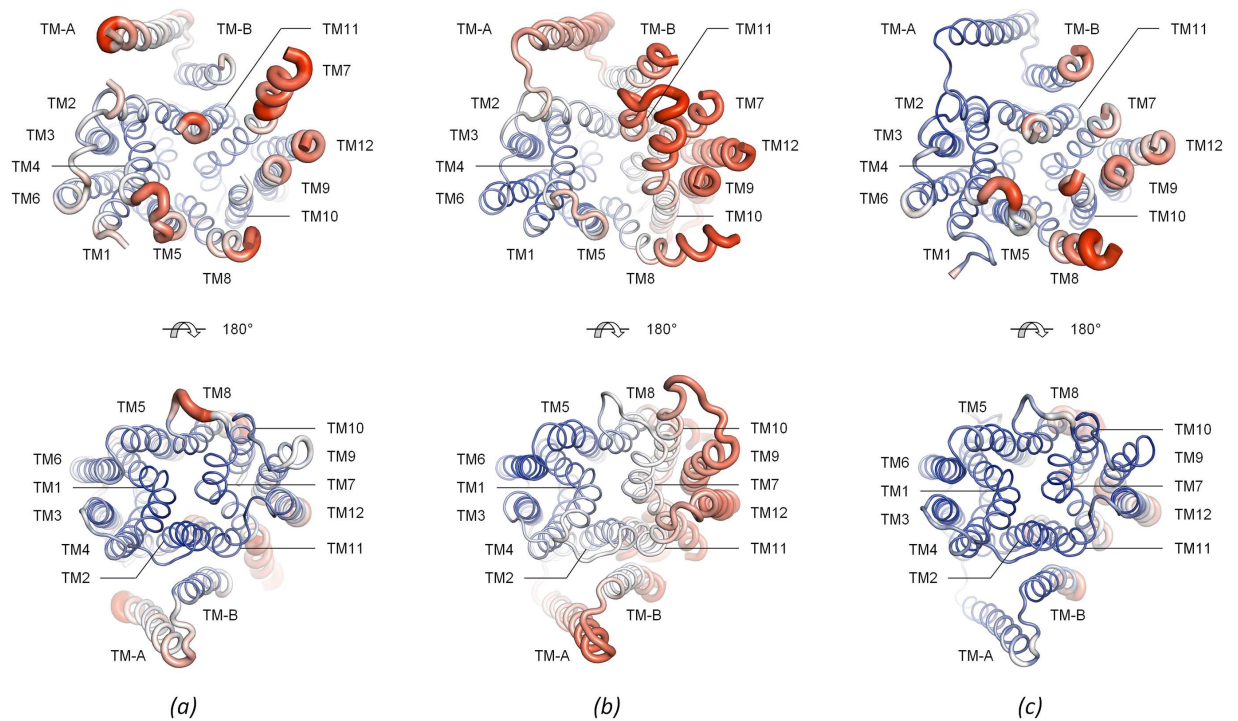


**Fig 5. Structural comparisons.** (a) All previously determined PepT<sub>St</sub> structures superimposed on the P3<sub>1</sub>21 structure (see also Table 5). The P3<sub>1</sub>21 structure is magenta, the P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> structure is blue the C222<sub>1</sub> structures with dipeptides are green (dipeptides are shown in sticks), and the C222<sub>1</sub> structures without dipeptides are white. Two views are shown—cytoplasmic (top) and periplasmic (bottom). (b) Superimposition of the N-domains of the same structures used in panel a. (c) Superimposition of the C-domains of the same structures used in panel a. (d) Same as in panel b, except that the C-domains are shown alongside the superimposed N-domains and that the P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> structure was omitted. A rotation of the C-domain relative to the N-domain can be recognized for the C222<sub>1</sub> structures with dipeptides (rotation marked with arrows). (e) Same as in panel c, except that only TM10 and TM11 are shown, that the C222<sub>1</sub> structure determined at room temperature is grey, and that the P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> structure was omitted. The C222<sub>1</sub> structures with dipeptides, the P3<sub>1</sub>21 structure and to some extent the C222<sub>1</sub> structure determined at room temperature, show bending of TM11 (indicated with a dashed arrow) and increased disorder in TM10 (indicated with a dashed oval) relative to the other C222<sub>1</sub> structures determined at cryogenic temperatures.

doi:10.1371/journal.pone.0173126.g005

samples a range of differently bent positions, rather than simply snap back and forth between discrete bent and unbent positions.

Analyzing how the crystallographic *B*-factors vary along the sequence is a commonly used procedure for identifying putative flexible regions. However, one should bear in mind that *B*-factors are sensitive to modeling errors [57], and can be heavily influenced by crystal contacts [58,59]. It is therefore desirable to include as many crystal forms as possible in such an analysis. Based on *B*-factor analysis, structural comparisons, and molecular dynamics studies, it has previously been inferred that the helices in the C-domain of PepT<sub>St</sub> are more flexible than those in the N-domain [17,21,23]. Here we show that an expanded *B*-factor analysis, which includes the monomeric P3<sub>1</sub>21 crystal form, further supports this assertion (Fig 6). Specifically, the cytoplasmic side of the C-domain exhibits particularly high *B*-factors in all three crystal forms



**Fig 6. B-factor analysis of inward facing PepT<sub>St</sub>.** (a) The P3<sub>1</sub>2<sub>1</sub> structure. The structure is shown in a putty tube representation where the diameter of the tube is correlated with the magnitude of the C-alpha B-factors. The structure is furthermore also colored by C-alpha B-factors: colors are ramped from blue over white to red, with blue designating low values and red designating high values. Two views are shown: cytoplasmic (top) and periplasmic (bottom). (b) The P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> structure (PDB: 4APS) depicted as in panel a. (c) The C222<sub>1</sub> structure (PDB: 4D2C) depicted as in panel a. A similar B-factor analysis on PepTs from other organisms is provided in S2 Fig. Higher B-factors are an indication of higher flexibility, but it should be noted that the regions with highest flexibility are those that could not be modeled at all, for example lobe TM10–TM11 in the C222<sub>1</sub> and P3<sub>1</sub>2<sub>1</sub> structures, and the cytoplasmic tip of TM7 in the P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> and C222<sub>1</sub> structures.

doi:10.1371/journal.pone.0173126.g006

of inward facing PepT<sub>St</sub> (Fig 6). This is of course well in line with the results from the structural comparisons of the different crystal forms, which as mentioned above identified the C-domain as the source of most of their structural differences. It should however be mentioned that the B-factor distribution is considerably less consistent if also including the other PepTs in the analysis (S2 Fig). Zones with particularly high B-factors are for example basically absent in GkPOT, YbgH and YePEPT, if exempting a couple of connectors (S2 Fig). At present, it is therefore unclear to which extent flexible regions may be conserved across the PepT family.

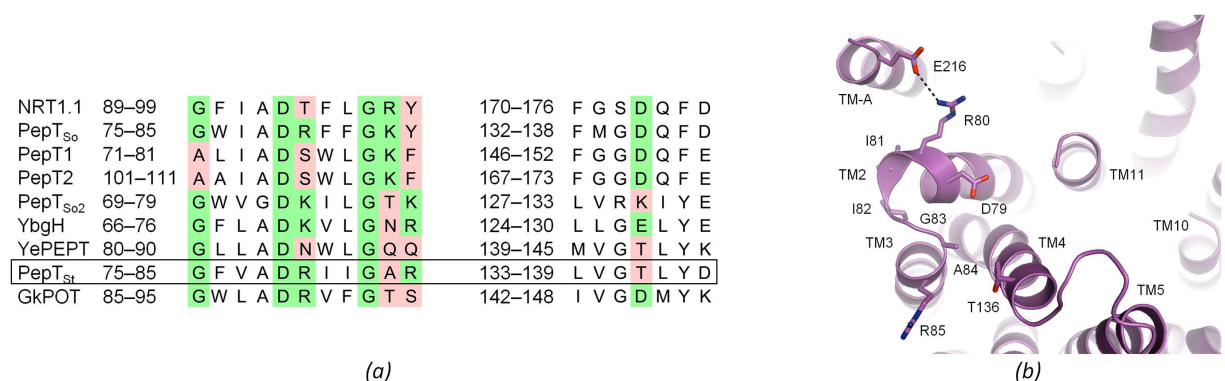
### Role of lobe TM10–TM11 in forming the inward facing occluded state

Two different occluded conformations have been identified for MFS transporters—outward facing occluded and inward facing occluded [15]. These states are both formed by bending of one or more pore lining helices towards the binding site, thus forming a constriction (typically referred to as a ‘gate’) that prevents substrate from entering or leaving [15]. Specifically, the outward facing occluded state is formed by bending of one or more of TM1, TM2, TM7 and TM8 on the extracellular side, and the inward facing occluded state by bending of one or more of TM4, TM5, TM10 and TM11 on the cytoplasmic side [15]. The bending of lobe TM10–TM11 towards the binding site in the P3<sub>1</sub>2<sub>1</sub> and the C222<sub>1</sub> structures with bound dipeptides thus indicates that these structures are either inward facing occluded states or intermediates between occluded and fully inward open, though it cannot presently be ascertained which of

these options applies due to much the lobe being disordered. Notably, a central role of lobe TM10–TM11 in occluding the binding site from the cytoplasmic side is not a novel proposition, but has also been suggested previously based on structural comparisons between PepT<sub>St</sub> and PepT<sub>So</sub> (characterized as inward facing occluded) [17], comparisons of the individual C222<sub>1</sub> PepT<sub>St</sub> structures [21], and through molecular dynamics studies carried out on PepT<sub>St</sub> and PepT<sub>So</sub> [23]. Bending of lobe TM10–TM11 may however not act alone in forming the fully occluded inward facing state. As mentioned above, a slight domain rotation can thus be observed in the C222<sub>1</sub> structures with dipeptides, which is likely also significant. Furthermore, structural changes in lobe TM4–TM5 and the TM-A and TM-B helix pair could potentially also be involved, as discussed below.

### Functional roles of the A-motif and the TM-A and TM-B helix pair

The A-motif represents the most well conserved motif in MFS transporters [60,61]. It has the consensus sequence G-X-X-X- (D/E) - (R/K) -X-G-[X] - (R/K) - (R/K) and typically spans loop TM2–TM3 and/or loop TM8–TM9 on the cytoplasmic side [60,61]. It functions in controlling the inward-outward transition, with the acidic residue capping a helix on the opposite domain in the outward open state, while being disengaged in the inward open state [62]. Indeed, mutating the acidic residue in the peptide transporter YbgH resulted in reduced activity [22], echoing similar results obtained for other MFS transporters [62–64]. In both these states, the second basic residue furthermore interacts with a co-conserved acidic residue, which is found in TM4 (A-motif in loop TM2–TM3) or TM10 (A-motif in loop TM8–TM9) [62]. We recently found that a third state of the A-motif can be recognized in several inward facing occluded structures, in which this interaction is broken, thus allowing TM4 (A-motif in loop TM2–TM3) or TM10 (A-motif in loop TM8–TM9) to bend towards the pore and occlude the binding site [15]. In the PepTs, an A-motif is found in loop TM2–TM3 (Fig 7a). It differs however somewhat from the consensus sequence in having an extra residue inserted prior to the second conserved glycine [22], and by often missing one or more of the basic residues as well as the acidic residue in TM4 (Fig 7a). The second basic residue and the acidic residue in TM4 are for example both missing in PepT<sub>St</sub> (Fig 7). The salt bridge normally found between these residues is thus absent here, which may suggest that regulated bending of TM4 does not



**Fig 7. The A-motif.** (a) Sequence alignment. Two fragments are shown of a sequence alignment that includes PepT<sub>St</sub> (boxed), the other PepTs for which structures are known, the related nitrate transporter NRT1.1 from *Arabidopsis thaliana*, for which structures are also known, and human PepT1 and PepT2. The left alignment fragment covers the A-motif in loop TM2–TM3, and the right fragment covers the region around the co-conserved acidic residue in TM4. Residues that match the consensus sequence G-X-X-X- (D/E) - (R/K) -X-G-[X] - (R/K) - (R/K) are highlighted in green, while those that do not are highlighted in pink. (b) Structure of the A-motif and the surrounding region, as modeled in the P3<sub>1</sub>21 structure of PepT<sub>St</sub>. Residues constituting the A-motif are labeled, and a salt bridge between R80 in the A-motif and E216 in TM-A is indicated with dashes.

doi:10.1371/journal.pone.0173126.g007

take place in PepT<sub>St</sub>. However, it is possible that other interactions may substitute for this salt bridge, as for example seen in the nitrate transporter NarK [15,65]. Indeed, the second basic residue is also missing in GkPOT, and yet, molecular dynamics studies suggested that the inward facing occluded form of this protein is generated through bending of lobe TM4–TM5 [18]. It may be mentioned though that although an inward facing occluded MFS structures has been described where only lobe TM10–TM11 is bent, namely XylE [31], we are not aware of any where only lobe TM4–TM5 is bent [15]. Thus, although lobe TM4–TM5 is quite bent in the inward facing occluded structures of EmrD and PiPT, these structures also feature apparent bending of lobe TM10–TM11 [66,67]. We therefore speculate that if bending of lobe TM4–TM5 takes place in some or all PepTs, it probably happens in addition to rather than instead of bending of lobe TM10–TM11.

It has been suggested that the TM-A and TM-B helix pair characteristic of bacterial PepTs acts as a sensor of conformational changes in concert with the A-motif [22,62]. In this respect it is notable that these two structural elements are in close proximity to each other (Fig 7b). Indeed, the A-motif is generally in direct contact with loop TM6–TM-A, though the specific interactions formed are quite variable among the different PepTs (not shown). A salt bridge between the A-motif and TM-A could furthermore be identified in PepT<sub>St</sub> (Fig 7b). However, although this interaction is also found in GkPOT, it is absent in all the other PepT structures, suggesting that it is either of little functional significance or that any function it may have in mediating communication between the A-motif and TM-A is replaceable. In conclusion, the transition of PepT<sub>St</sub> from inward open to inward facing occluded involves bending of lobe TM10–TM11 and a slight domain rotation, as discussed above. However, bending of lobe TM4–TM5 may also play a role in at least some PepTs. Finally, the A-motif probably functions in regulating transitions between the inward and outward facing states, but may also have additional roles in regulating the conformational changes in PepTs.

## Conclusions and outlook

Here we have provided a report of the structure determination of PepT<sub>St</sub> in space group P3<sub>1</sub>21, which may aid in guiding others aiming to determine structures of MFS transporters or other membrane proteins. Three different maltoside detergents were tested for crystallization. The best crystals obtained were grown using NM, which has the shortest chain-length, and is by far the most destabilizing of the three. This underlines that short-chain detergents should be tested for crystallization even if they are highly destabilizing compared to their longer-chain counterparts, as long as they do not cause significant aggregation. It should be mentioned though, that impressive results on PepT<sub>St</sub> and other integral membrane proteins have been achieved using the increasingly popular *in meso* method [21,51]. For future projects we therefore suggest to use this technique in parallel to the traditional *in surfo* vapor diffusion method employed in this study.

The three different crystal forms now available for PepT<sub>St</sub> exhibits very different packing patterns, which is probably in large part due to the fact that different methods were used for crystallization: *in surfo* for the P3<sub>1</sub>21 and P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> forms, and *in meso*, which enables tighter packing of transmembrane parts, for the C222<sub>1</sub> form, combined with the fact that the protein was purified in different detergents, favoring a monomeric form in the case of the P3<sub>1</sub>21 crystal form, but a likely unphysiological antiparallel dimer in the cases of the P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> and C222<sub>1</sub> forms. The P3<sub>1</sub>21 form is more similar to the C222<sub>1</sub> structures than to the P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> structure. It adopts an inward facing conformation that is either occluded or intermediary between inward open and occluded, sharing the bent position of lobe TM10–TM11 with the C222<sub>1</sub> structures with peptides and the relative domain orientation with those without, thus indicating that

bending of lobe TM10–TM11 can occur independently of domain rotation upon formation of the inward facing occluded state.

Much still remains to be understood about conformational changes in PepTs, which relates in part to the fact that all known PepT structures are inward facing. Obtaining structures of outward facing forms would thus clearly be highly valuable. Furthermore, spectroscopic methods such as single molecule Förster resonance energy transfer (FRET) or double electron-electron resonance (DEER) also hold great potential for further elucidating the conformational changes. Indeed, DEER spectroscopy applied to PepT<sub>So</sub> and combined with *in silico* modeling and molecular dynamics studies has already helped in characterizing the outward open form of this protein [23]. It may be noted that PepT<sub>St</sub> was found to be unsuitable for such a spectroscopic approach, as the dimeric nature of the purified protein would complicate data interpretation [23]. The protocol we have established here for purification of the monomeric form may thus prove useful for carrying out future spectroscopic studies and potentially other *in vitro* experiments on PepT<sub>St</sub>.

## Supporting information

**S1 Fig. Crystal packing interfaces.** (a) The P3<sub>1</sub>2<sub>1</sub> structure. Coloration is by residue count with colors ramped from white (N-terminus) over wheat, golden and orange to red (C-terminus). Side chains within 4 Å of a symmetry related molecule are shown in sticks with transparent surfaces, and the structural elements in which they occur are labeled. Two side views are shown. The contact areas on the periplasmic side are much more extensive than on the cytoplasmic side. Specifically, the former involves loop TM1–TM2, loop TM5–TM6, loop TM7–TM8, loop TM9–TM10 and loop TM11–TM12, and the latter TM7 and loop TM4–TM5. (b) The P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> structure (PDB: 4APS). Shown as in panel a, except that residues within 4 Å of the other molecule in the asymmetric unit are shown in light blue sticks with transparent surfaces. The contact surfaces encompass fewer residues here, and they are more evenly distributed between the periplasmic and cytoplasmic sides. Specifically, they are found in loop TM1–TM2 and loop TM11–TM12 on the former side and loop TM6–TM-A on the latter (c). The C222<sub>1</sub> form (PDB: 4D2C). Shown as in panel a. Two lateral interfaces are found. One is formed by the N-terminal tail, TM1, TM5, TM6, TM8 and loop TM7–TM8 packing against the equivalent elements of a symmetry-related mate to generate an antiparallel dimer, which is structurally equivalent to the non-crystallographic dimer seen in the P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> form. These dimers are in turn linked into a continuous planar layers (Fig 4c, bottom) by the second lateral interface, encompassing TM-A, TM9 and TM12 on each of two symmetry-related mates. Direct end-to-end interactions between layers are mediated by a single though tight interface formed by TM-A, loop TM6–TM-A on the cytoplasmic side, and TM9 and loop TM9–TM10 on the periplasmic side. (DOCX)

**S2 Fig. Expanded B-factor analysis on PepTs from various organisms.** (a) P3<sub>2</sub> crystal form of PepT<sub>So</sub> from *Shewanella oneidensis* (PDB: 2XUT). Shown in a putty tube representation where the diameter of the tube is correlated with the magnitude of the C-alpha B-factors, and also colored by C-alpha B-factors: colors are ramped from blue over white to red, with blue designating low values and red designating high values. Two views are shown: cytoplasmic (top) and periplasmic (bottom). (b) P4<sub>1</sub>2<sub>1</sub>2 crystal form of PepT<sub>So</sub> (PDB: 4UVM). (c) PepT<sub>So2</sub> –a different PepT from *Shewanella oneidensis* (PDB: 4LEP). (d) GkPOT from *Geobacillus kaustophilus* (PDB: 4IKV). (e) YbgH from *Escherichia coli* (PDB: 4Q65). (f) YePEPT from *Yersinia enterocolitica* (PDB: 4W6V). In the case of the P3<sub>2</sub> crystal form of PepT<sub>So</sub>, most of the C-terminal MFS domain (TM7–TM12) as well as TM-A and TM-B are characterized by very high B-



factors. The P4<sub>1</sub>2<sub>1</sub>2 form of this protein is structurally fairly similar, but here the *B*-factors are not as high overall. Nonetheless, the highest *B*-factors are still mainly in the C-terminal domain. In PepT<sub>So2</sub>, most of the high C-alpha *B*-factors are on the cytoplasmic side of the C-terminal MFS domain, but also the cytoplasmic side of the N-terminal MFS domain (TM1–TM6) features regions with high *B*-factors. In GkPOT, only TM-A and two loops on the periplasmic side of the C-terminal domain feature stretches with high C-alpha *B*-factors. In YbgH, regions with high *B*-factors are rather few and mainly include a few loops found on the cytoplasmic side of the N-terminal domain. In YePEPT, regions with high *B*-factors are even fewer, with loop TM1–TM2 sticking out the most. We conclude that the distributions of *B*-factors in the different PepTs are rather varied, and do not fully echo the scenario observed for PepT<sub>St</sub>. (DOCX)

## Acknowledgments

We thank the Diamond Light Source for access to beamlines I02, I04, and I24. We furthermore acknowledge the Protein Science Facility at Karolinska Institutet for providing crystallization infrastructure as well as the Sample Preparation and Characterization (SPC) Facility of EMBL for support with nanoDSF measurements. This research was funded by the European Union's Seventh Framework Programme (FP7/2007–2013) under BioStruct-X (grant agreement N° 783), by the Horizon 2020 programme of the European Union (iNEXT grant, project No. 653706) and the Swedish Research Council (grant number 621-2013-5905).

## Author Contributions

**Conceptualization:** EQ CL.

**Funding acquisition:** CL.

**Investigation:** EQ MM CL.

**Supervision:** CL.

**Writing – original draft:** EQ.

**Writing – review & editing:** EQ CL.

## References

1. Juillard V, Le Bars D, Kunji ERS, Konings WN, Gripon JC, Richard J. Oligopeptides are the main source of nitrogen for *Lactococcus lactis* during growth in milk. *Appl Environ Microbiol.* 1995; 61: 3024–3030. PMID: [7487034](#)
2. Aredes Fernández PA, Saguir FM, Manca De Nadra MC. Effect of dipeptides on the growth of *Oenococcus oeni* in synthetic medium deprived of amino acids. *Curr Microbiol.* 2004; 49: 361–365. doi: [10.1007/s00284-004-4367-7](#) PMID: [15486711](#)
3. Webb KE. Intestinal absorption of protein hydrolysis products: a review. *J Anim Sci.* 1990; 68: 3011–3022. PMID: [2211428](#)
4. Komarova NY, Thor K, Gubler A, Meier S, Dietrich D, Weichert A, et al. AtPTR1 and AtPTR5 transport dipeptides in planta. *Plant Physiol.* 2008; 148: 856–869. doi: [10.1104/pp.108.123844](#) PMID: [18753286](#)
5. Daniel H, Spanier B, Kottra G, Weitz D. From bacteria to man: archaic proton-dependent peptide transporters at work. *Physiol.* 2006; 21: 93–102.
6. Fei YJ, Kanai Y, Nussberger S, Ganapathy V, Leibach FH, Romero MF, et al. Expression cloning of a mammalian proton-coupled oligopeptide transporter. *Nature.* 1994; 368: 563–566. doi: [10.1038/368563a0](#) PMID: [8139693](#)

7. Liang R, Fei Y, Prasad PD, Ramamoorthy S, Han H, Yang-Feng TL, et al. Human intestinal H<sup>+</sup>/peptide cotransporter. Cloning, functional expression, and chromosomal localization. *J Biol Chem.* 1995; 270: 6456–6463. PMID: [7896779](#)
8. Liu W, Liang R, Ramamoorthy S, Fei YJ, Ganapathy ME, Hediger MA, et al. Molecular cloning of PEPT 2, a new member of the H<sup>+</sup>/peptide cotransporter family, from human kidney. *BBA—Biomembr.* 1995; 1235: 461–466.
9. Boll M, Herget M, Wagener M, Weber WM, Markovich D, Biber J, et al. Expression cloning and functional characterization of the kidney cortex high-affinity proton-coupled peptide transporter. *Proc Natl Acad Sci U S A.* 1996; 93: 284–289. PMID: [8552623](#)
10. Smith DE, Cléménçon B, Hediger MA. Proton-coupled oligopeptide transporter family SLC15: Physiological, pharmacological and pathological implications. *Mol Aspects Med.* 2013; 34: 323–336. doi: [10.1016/j.mam.2012.11.003](#) PMID: [23506874](#)
11. Brandsch M. Transport of drugs by proton-coupled peptide transporters: pearls and pitfalls. *Expert Opin Drug Metab Toxicol.* 2009; 5: 887–905. doi: [10.1517/17425250903042292](#) PMID: [19519280](#)
12. Brandsch M. Drug transport via the intestinal peptide transporter PepT1. *Curr Opin Pharmacol.* 2013; 13: 881–887. doi: [10.1016/j.coph.2013.08.004](#) PMID: [24007794](#)
13. Newstead S. Molecular insights into proton coupled peptide transport in the PTR family of oligopeptide transporters. *Biochim Biophys Acta—Gen Subj.* 2015; 1850: 488–499.
14. Yan N. Structural Biology of the Major Facilitator Superfamily Transporters. *Annu Rev Biophys.* 2015; 44: 257–283. doi: [10.1146/annurev-biophys-060414-033901](#) PMID: [26098515](#)
15. Quistgaard EM, Löw C, Guettou F, Nordlund P. Understanding transport by the major facilitator superfamily (MFS): structures pave the way. *Nat Rev Mol Cell Biol.* 2016; 17: 123–132. doi: [10.1038/nrm.2015.25](#) PMID: [26758938](#)
16. Newstead S, Drew D, Cameron AD, Postis VLG, Xia X, Fowler PW, et al. Crystal structure of a prokaryotic homologue of the mammalian oligopeptide-proton symporters, PepT1 and PepT2. *EMBO J.* 2011; 30: 417–426. doi: [10.1038/emboj.2010.309](#) PMID: [21131908](#)
17. Solcan N, Kwok J, Fowler PW, Cameron A, Drew D, Iwata S, et al. Alternating access mechanism in the POT family of oligopeptide transporters. *EMBO J.* 2012; 31: 3411–3421. doi: [10.1038/emboj.2012.157](#) PMID: [22659829](#)
18. Doki S, Kato HE, Solcan N, Iwaki M, Koyama M, Hattori M, et al. Structural basis for dynamic mechanism of proton-coupled symport by the peptide transporter POT. *Proc Natl Acad Sci U S A.* 2013; 110: 11343–11348. doi: [10.1073/pnas.1301079110](#) PMID: [23798427](#)
19. Guettou F, Quistgaard EM, Trésaugues L, Moberg P, Jegerschöld C, Zhu L, et al. Structural insights into substrate recognition in proton-dependent oligopeptide transporters. *EMBO Rep.* 2013; 14: 804–810. doi: [10.1038/embor.2013.107](#) PMID: [23867627](#)
20. Guettou F, Quistgaard EM, Raba M, Moberg P, Löw C, Nordlund P. Selectivity mechanism of a bacterial homolog of the human drug-peptide transporters PepT1 and PepT2. *Nat Struct Mol Biol.* 2014; 21: 728–731. doi: [10.1038/nsmb.2860](#) PMID: [25064511](#)
21. Lyons JA, Parker JL, Solcan N, Brintha A, Li D, Shah ST, et al. Structural basis for polyspecificity in the POT family of proton-coupled oligopeptide transporters. *EMBO Rep.* 2014; 15: 886–893. doi: [10.15252/embr.201338403](#) PMID: [24916388](#)
22. Zhao Y, Mao G, Liu M, Zhang L, Wang X, Zhang XC. Crystal structure of the *E. coli* peptide transporter YbgH. *Structure.* 2014; 22: 1152–1160. doi: [10.1016/j.str.2014.06.008](#) PMID: [25066136](#)
23. Fowler PW, Orwick-Rydmark M, Forrest LR, Correspondence SN, Radestock S, Solcan N, et al. Gating Topology of the Proton-Coupled Oligopeptide Symporters. *Structure.* 2015; 23: 290–301. doi: [10.1016/j.str.2014.12.012](#) PMID: [25651061](#)
24. Boggavarapu R, Jeckelmann J-M, Harder D, Ucurum Z, Fotiadis D. Role of electrostatic interactions for ligand recognition and specificity of peptide transporters. *BMC Biol.* 2015; 13.
25. Huang CY, Olieric V, Ma P, Panepucci E, Diederichs K, Wang M, et al. In meso in situ serial X-ray crystallography of soluble and membrane proteins. *Acta Crystallogr Sect D Biol Crystallogr.* International Union of Crystallography; 2015; 71: 1238–1256.
26. Huang C-Y, Olieric V, Ma P, Howe N, Vozeley L, Liu X, et al. In meso in situ serial X-ray crystallography of soluble and membrane proteins at cryogenic temperatures. *Acta Crystallogr Sect D Biol Crystallogr.* 2016; 72: 93–112.
27. Löw C, Moberg P, Quistgaard EM, Hedrén M, Guettou F, Frauenfeld J, et al. High-throughput analytical gel filtration screening of integral membrane proteins for structural studies. *Biochim Biophys Acta—Gen Subj.* 2013; 1830: 3497–3508.

28. Chae PS, Rasmussen SGF, Rana RR, Gotfryd K, Chandra R, Goren MA, et al. Maltose-neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins. *Nat Methods*. 2010; 7: 1003–1008. doi: [10.1038/nmeth.1526](https://doi.org/10.1038/nmeth.1526) PMID: [21037590](https://pubmed.ncbi.nlm.nih.gov/21037590/)
29. Van Duyne GD, Standaert RF, Karplus PA, Schreiber SL, Clardy J. Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J Mol Biol*. 1993; 229: 105–124. doi: [10.1006/jmbi.1993.1012](https://doi.org/10.1006/jmbi.1993.1012) PMID: [7678431](https://pubmed.ncbi.nlm.nih.gov/7678431/)
30. Löw C, Yau YH, Pardon E, Jegerschöld C, Wählin L, Quistgaard EM, et al. Nanobody Mediated Crystallization of an Archeal Mechanosensitive Channel. *PLoS One*. 2013; 8: e77984. doi: [10.1371/journal.pone.0077984](https://doi.org/10.1371/journal.pone.0077984) PMID: [24205053](https://pubmed.ncbi.nlm.nih.gov/24205053/)
31. Quistgaard EM, Löw C, Moberg P, Trésaugues L, Nordlund P. Structural basis for substrate transport in the GLUT-homology family of monosaccharide transporters. *Nat Struct Mol Biol*. 2013; 20: 766–768. doi: [10.1038/nsmb.2569](https://doi.org/10.1038/nsmb.2569) PMID: [23624861](https://pubmed.ncbi.nlm.nih.gov/23624861/)
32. Kabsch W. XDS. *Acta Crystallogr Sect D Biol Crystallogr*. 2010; 66: 125–132.
33. Strong M, Sawaya MR, Wang S, Phillips M, Cascio D, Eisenberg D. Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A*. 2006; 103: 8060–8065. doi: [10.1073/pnas.0602606103](https://doi.org/10.1073/pnas.0602606103) PMID: [16690741](https://pubmed.ncbi.nlm.nih.gov/16690741/)
34. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr*. 2007; 40: 658–674. doi: [10.1107/S0021889807021206](https://doi.org/10.1107/S0021889807021206) PMID: [19461840](https://pubmed.ncbi.nlm.nih.gov/19461840/)
35. Adams PD, Afonine P V, Bunkóczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr Sect D Biol Crystallogr*. 2010; 66: 213–221.
36. Vonrhein C, Blanc E, Roversi P, Bricogne G. Automated structure solution with autoSHARP. *Methods Mol Biol*. 2007; 364: 215–230. doi: [10.1385/1-59745-266-1:215](https://doi.org/10.1385/1-59745-266-1:215) PMID: [17172768](https://pubmed.ncbi.nlm.nih.gov/17172768/)
37. Bricogne G, Vonrhein C, Flensburg C, Schiltz M, Paciorek W. Generation, representation and flow of phase information in structure determination: Recent developments in and around SHARP 2.0. *Acta Crystallogr—Sect D Biol Crystallogr*. 2003; 59: 2023–2030.
38. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr Sect D Biol Crystallogr*. 2010; 66: 486–501.
39. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, et al. MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr Sect D Biol Crystallogr*. 2010; 66: 12–21.
40. Holm L, Rosenström P. Dali server: Conservation mapping in 3D. *Nucleic Acids Res*. 2010; 38: W545–W549. doi: [10.1093/nar/gkq366](https://doi.org/10.1093/nar/gkq366) PMID: [20457744](https://pubmed.ncbi.nlm.nih.gov/20457744/)
41. DeLano WL. The PyMOL Molecular Graphics System. Schrödinger LLC. <http://www.pymol.org>
42. Jastrzebska B, Fotiadis D, Jang GF, Stenkamp RE, Engel A, Palczewski K. Functional and structural characterization of rhodopsin oligomers. *J Biol Chem*. 2006; 281: 11917–11922. doi: [10.1074/jbc.M600422200](https://doi.org/10.1074/jbc.M600422200) PMID: [16495215](https://pubmed.ncbi.nlm.nih.gov/16495215/)
43. Tate CG. Practical considerations of membrane protein instability during purification and crystallisation. *Methods Mol Biol*. 2010; 601: 187–203. doi: [10.1007/978-1-60761-344-2\\_12](https://doi.org/10.1007/978-1-60761-344-2_12) PMID: [20099147](https://pubmed.ncbi.nlm.nih.gov/20099147/)
44. Parker JL, Newstead S. Membrane protein crystallisation: Current trends and future perspectives. *Adv Exp Med Biol*. 2016; 922: 61–72. doi: [10.1007/978-3-319-35072-1\\_5](https://doi.org/10.1007/978-3-319-35072-1_5) PMID: [27553235](https://pubmed.ncbi.nlm.nih.gov/27553235/)
45. Buschmann S, Richers S, Ermiler U, Michel H. A decade of crystallization drops: Crystallization of the cbb3 cytochrome c oxidase from *Pseudomonas stutzeri*. *Protein Sci*. 2014; 23: 411–422. doi: [10.1002/pro.2423](https://doi.org/10.1002/pro.2423) PMID: [24488923](https://pubmed.ncbi.nlm.nih.gov/24488923/)
46. Pike ACW, Garman EF, Krojer T, von Delft F, Carpenter EP. An overview of heavy-atom derivatization of protein crystals. *Acta Crystallogr Sect D, Struct Biol*. 2016; 72: 303–318.
47. Schulz GE. A new classification of membrane protein crystals. *J Mol Biol*. 2011; 407: 640–646. doi: [10.1016/j.jmb.2011.02.003](https://doi.org/10.1016/j.jmb.2011.02.003) PMID: [21315729](https://pubmed.ncbi.nlm.nih.gov/21315729/)
48. Quistgaard EM, Löw C, Moberg P, Nordlund P. Metal-mediated crystallization of the xylose transporter XylE from *Escherichia coli* in three different crystal forms. *J Struct Biol*. 2013; 184: 375–378. doi: [10.1016/j.jsb.2013.09.009](https://doi.org/10.1016/j.jsb.2013.09.009) PMID: [24060988](https://pubmed.ncbi.nlm.nih.gov/24060988/)
49. Chaptal V, Kilburg A, Flot D, Wiseman B, Aghajari N, Jault JM, et al. Two different centered monoclinic crystals of the *E. coli* outer-membrane protein OmpF originate from the same building block. *Biochim Biophys Acta—Biomembr*. 2016; 1858: 326–332.
50. Chaptal V, Kilburg A, Flot D, Wiseman B, Aghajari N, Jault JM, et al. Diffraction anisotropy falloff in the direction of the detergent belt for two centered monoclinic crystals of OmpF. *Data Br*. 2016; 7: 726–729.

51. Caffrey M. A comprehensive review of the lipid cubic phase or in meso method for crystallizing membrane and soluble proteins and complexes. *Acta Crystallogr Sect F Struct Biol Commun.* 2015; 71: 3–18.
52. Banatao DR, Cascio D, Crowley CS, Fleissner MR, Tienson HL, Yeates TO. An approach to crystallizing proteins by synthetic symmetrization. *Proc Natl Acad Sci U S A.* 2006; 103: 16230–16235. doi: [10.1073/pnas.0607674103](https://doi.org/10.1073/pnas.0607674103) PMID: [17050682](https://pubmed.ncbi.nlm.nih.gov/17050682/)
53. Andreć M, Snyder DA, Zhou Z, Young J, Montelione GT, Levy RM. A large data set comparison of protein structures determined by crystallography and NMR: Statistical test for structural differences and the effect of crystal packing. *Proteins Struct Funct Genet.* 2007; 69: 449–465. doi: [10.1002/prot.21507](https://doi.org/10.1002/prot.21507) PMID: [17623851](https://pubmed.ncbi.nlm.nih.gov/17623851/)
54. Sikic K, Tomic S, Carugo O. Systematic Comparison of Crystal and NMR Protein Structures Deposited in the Protein Data Bank. *Open Biochem J.* 2010; 4: 83–95. doi: [10.2174/1874091X01004010083](https://doi.org/10.2174/1874091X01004010083) PMID: [21293729](https://pubmed.ncbi.nlm.nih.gov/21293729/)
55. Eyal E, Gerzon S, Potapov V, Edelman M, Sobolev V. The limit of accuracy of protein modeling: Influence of crystal packing on protein structure. *J Mol Biol.* 2005; 351: 431–442. doi: [10.1016/j.jmb.2005.05.066](https://doi.org/10.1016/j.jmb.2005.05.066) PMID: [16005885](https://pubmed.ncbi.nlm.nih.gov/16005885/)
56. Rapp CS, Pollack RM. Crystal packing effects on protein loops. *Proteins Struct Funct Genet.* 2005; 60: 103–109. doi: [10.1002/prot.20492](https://doi.org/10.1002/prot.20492) PMID: [15852307](https://pubmed.ncbi.nlm.nih.gov/15852307/)
57. Stroud RM, Fauman EB. Significance of structural changes in proteins: expected errors in refined protein structures. *Protein Sci.* 1995; 4: 2392–2404. doi: [10.1002/pro.5560041118](https://doi.org/10.1002/pro.5560041118) PMID: [8563637](https://pubmed.ncbi.nlm.nih.gov/8563637/)
58. Carugo O, Argos P. Protein-protein crystal-packing contacts. *Protein Sci.* 1997; 6: 2261–2263. doi: [10.1002/pro.5560061021](https://doi.org/10.1002/pro.5560061021) PMID: [9336849](https://pubmed.ncbi.nlm.nih.gov/9336849/)
59. Song G, Jernigan RL. vGNM: A Better Model for Understanding the Dynamics of Proteins in Crystals. *J Mol Biol.* 2007; 369: 880–893. doi: [10.1016/j.jmb.2007.03.059](https://doi.org/10.1016/j.jmb.2007.03.059) PMID: [17451743](https://pubmed.ncbi.nlm.nih.gov/17451743/)
60. Henderson PJ, Maiden MC. Homologous sugar transport proteins in *Escherichia coli* and their relatives in both prokaryotes and eukaryotes. *Philos Trans R Soc Lond B Biol Sci.* 1990; 326: 391–410. PMID: [1970645](https://pubmed.ncbi.nlm.nih.gov/1970645/)
61. Pao SS, Paulsen IT, Saier MH. Major facilitator superfamily. *Microbiol Mol Biol Rev.* 1998; 62: 1–34. PMID: [9529885](https://pubmed.ncbi.nlm.nih.gov/9529885/)
62. Jiang D, Zhao Y, Wang X, Fan J, Heng J, Liu X, et al. Structure of the YajR transporter suggests a transport mechanism based on the conserved motif A. *Proc Natl Acad Sci U S A.* 2013; 110: 14664–14669. doi: [10.1073/pnas.1308127110](https://doi.org/10.1073/pnas.1308127110) PMID: [23950222](https://pubmed.ncbi.nlm.nih.gov/23950222/)
63. Yamaguchi A, Someya Y, Sawai T. Metal-tetracycline/H<sup>+</sup> antiporter of *Escherichia coli* encoded by transposon Tn10: The role of a conserved sequence motif, GXXXXRXGRR, in a putative cytoplasmic loop between helices 2 and 3. *J Biol Chem.* 1992; 267: 19155–19162. PMID: [1326546](https://pubmed.ncbi.nlm.nih.gov/1326546/)
64. Jessen-Marshall AE, Paul NJ, Brooker RJ. The conserved motif, GXXX(D/E)(R/K)XG[X](R/K)(R/K), in hydrophilic loop 2/3 of the lactose permease. *J Biol Chem.* 1995; 270: 16251–16257. PMID: [7608191](https://pubmed.ncbi.nlm.nih.gov/7608191/)
65. Fukuda M, Takeda H, Kato HE, Doki S, Ito K, Maturana AD, et al. Structural basis for dynamic mechanism of nitrate/nitrite antiport by NarK. *Nat Commun.* 2015; 6: 7097. doi: [10.1038/ncomms8097](https://doi.org/10.1038/ncomms8097) PMID: [25959928](https://pubmed.ncbi.nlm.nih.gov/25959928/)
66. Yin Y, He X, Szewczyk P, Nguyen T, Chang G. Structure of the multidrug transporter EmrD from *Escherichia coli*. *Science.* 2006; 312: 741–744. doi: [10.1126/science.1125629](https://doi.org/10.1126/science.1125629) PMID: [16675700](https://pubmed.ncbi.nlm.nih.gov/16675700/)
67. Pedersen BP, Kumar H, Waight AB, Risenmay AJ, Roe-Zurz Z, Chau BH, et al. Crystal structure of a eukaryotic phosphate transporter. *Nature.* 2013; 496: 533–536. doi: [10.1038/nature12042](https://doi.org/10.1038/nature12042) PMID: [23542591](https://pubmed.ncbi.nlm.nih.gov/23542591/)