






Crystal structure of the sugar acid-binding protein CxaP from a TRAP transporter in *Advenella mimigardefordensis* strain DPN7^T

Lukas Schäfer¹ , Christina Meinert-Berning¹ , Stefanie Kobus², Astrid Höppner² , Sander H. J. Smits^{2,3}  and Alexander Steinbüchel^{1,4} 

¹ Institute of Molecular Microbiology and Biotechnology, Westfälische Wilhelms University Münster, Münster, Germany

² Center for Structural Studies, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

³ Institute of Biochemistry, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

⁴ Environmental Science Department, King Abdulaziz University, Jeddah, Saudi Arabia

Keywords

ligand binding; mutagenesis; substrate binding; sugar acid; TRAP transporter

Correspondence

S. H. J. Smits, Center for Structural Studies, Heinrich Heine University Düsseldorf, Universitätsstrasse 1, Düsseldorf 40225, Germany

Tel: +49 211 81 12647

E-mail: sander.smits@uni-duesseldorf.de and

A. Steinbüchel, Institut of Molecular Microbiology and Biotechnology, Westfälische Wilhelms University Münster, Corrensstrasse 3, Münster 48149, Germany
Tel: +49 251 83 38388

E-mail: steinbu@uni-muenster.de

(Received 13 January 2021, revised 19 February 2021, accepted 24 February 2021)

doi:10.1111/febs.15789

Recently, CxaP, a sugar acid substrate binding protein (SBP) from *Advenella mimigardefordensis* strain DPN7^T, was identified as part of a novel sugar uptake strategy. In the present study, the protein was successfully crystallized. Although several SBP structures of tripartite ATP-independent periplasmic transporters have already been solved, this is the first structure of an SBP accepting multiple sugar acid ligands. Protein crystals were obtained with bound D-xylonic acid, D-fuconic acid D-galactonic and D-gluconic acid, respectively. The protein shows the typical structure of an SBP of a tripartite ATP-independent periplasmic transporter consisting of two domains linked by a hinge and spanned by a long α -helix. By analysis of the structure, the substrate binding site of the protein was identified. The carboxylic group of the sugar acids interacts with Arg175, whereas the coordination of the hydroxylic groups at positions C2 and C3 is most probably realized by Arg154 and Asn151. Furthermore, it was observed that 2-keto-3-deoxy-D-gluconic acid is bound in protein crystals that were crystallized without the addition of any ligand, indicating that this molecule is prebound to the protein and is displaced by the other ligands if they are available.

Database

Structural data of CxaP complexes are available in the worldwide Protein Data Bank (<https://www.rcsb.org>) under the accession codes **7BBR** (2-keto-3-deoxy-D-gluconic acid), **7BCR** (D-galactonic acid), **7BCN** (D-xylonic acid), **7BCO** (D-fuconic acid) and **7BCP** (D-gluconic acid).

Introduction

Nutrient uptake is essential for growth of all living organisms. Versatile strategies have evolved using different types and mechanisms for the uptake of nutrients. The well-studied phosphotransferase system is often used for carbohydrate uptake in bacteria [1]. Other strategies include permeases of the major

facilitator superfamily or ATP-binding cassette transporters [2] for a broad range of substrates. Furthermore, a tripartite ATP-independent transport (TRAP) system mediates the uptake of D-gluconate in *Sinorhizobium meliloti* [3]. Subsequent to their discovery in the 1990s [4], several TRAP transport systems have

Abbreviations

K_d , dissociation constant; KDG, 2-keto-3-deoxy-D-gluconate; PDB, Protein Data Bank; SBP, substrate binding protein; TRAP, tripartite ATP-independent periplasmic.

been described accepting various ligands. The only limitation appears to be the need of a carboxylic acid in the structure of the ligand [5–8]. TRAP transport systems consist of three components: the two transmembrane components, DctQ and DctM, and the substrate binding protein (SBP), DctP [9]. The transport of the ligands starts with an interaction of the ligand with the SBP, which subsequently interacts with the complex of the transmembrane proteins DctQ and DctM. The translocation of the ligand is then mediated by DctM, the larger of both membrane proteins, consisting of twelve transmembrane-spanning helices that form a barrel structure [6].

Besides the TRAP transporter from *S. meliloti*, *Bacillus halodurans* also possesses an SBP of a TRAP transporter, which was also shown to accept D-glucuronate, D-xylonate and D-galactonate as ligands, although the corresponding transport system has not been characterized [8]. Recently, a TRAP transporter involved in the uptake of sugar acids was described in *Advenella mimigardefordensis* strain DPN7^T [10] as part of a novel strategy of carbohydrate uptake in this Gram-negative bacterium. The deletion of the individual components of the TRAP transporter led to restricted growth of the bacterium with the monosaccharides D-glucose, D-galactose, D-fucose and L-arabinose, as well as to poor growth with D-xylose. Additionally, intracellular accumulation of ¹⁴C-labeled D-glucose was also shown to be dependent on the identified TRAP transporter, demonstrating that the same single transporter is crucial for the uptake of carbohydrates [10]. This was unexpected because the substrate range of so far described TRAP transporters is limited to organic acids [5,8].

Further investigations of the SBP in *A. mimigardefordensis* DPN7^T revealed that not the sugar molecules, but the corresponding acids are the ligands of this protein [10]. *In silico* analyses revealed that the strain possesses putative membrane-bound quinone glucose dehydrogenases, and the oxidation of the respective sugars was observed in the membrane fraction of cells of *A. mimigardefordensis*. This indicated that, prior to transport, the carbohydrates are oxidized to their corresponding sugar acids in the periplasmic space of the bacterial cell, which explains the restricted growth of the strain with D-glucose and the lack of ¹⁴C-labeled D-glucose accumulation in the deletion mutants.

Because the protein is not a binding protein for dicarboxylates, the name DctP_{Am} could lead to confusion regarding its function. To reflect the substrates of this particular SBP and thus the function with respect to the transport of sugar acids in *A. mimigardefordensis*

DPN7^T, it is renamed CxaP. This new name is used throughout the present study.

During the screening of ligands of CxaP, it became obvious that the terminal carboxylic acid group and the hydroxylic groups at position C2 and C3 are important for ligand binding. In general, the overall sequence similarity of TRAP type SBPs is relatively low. Structurally, however, they appear to be conserved. DctP proteins, similar to CxaP, usually consist of two domains linked by a hinge and spanned by a long α -helix, as is typical for SBPs of TRAP transporters [5,11,12]. Generally, the ligand binding in TRAP type SBPs involves a salt bridge formed by a highly conserved arginine and the carboxylic acid of the ligand [5,13]. This was also identified in CxaP at position Arg175 [10]. Because a large variety of ligands are to be bound by CxaP [5,14], it would be of interest to investigate the binding site of the protein in great detail by X-ray crystallography in combination with mutagenesis analysis. The present study shows high-resolution crystal structures of CxaP in complex with different ligands and thereby provides information about the architecture of the binding pocket and coordination of the ligands in the protein. This is the first crystal structure with sugar acid ligands reported for a TRAP type SBP.

Results and Discussion

The uptake of carbohydrates in *A. mimigardefordensis* strain DPN7^T by a novel strategy has been described previously [10]. It was shown that the monosaccharides D-glucose, D-fucose, D-galactose, D-xylose and L-arabinose are oxidized in the periplasm and the resulting corresponding sugar acids are subsequently transported into the cell by the characterized TRAP transport system [10]. The SBP displayed wide specificity with similar affinity toward five different ligands, which is unusual for TRAP type SBPs. Usually, these transporters exhibit high specificities toward one or two structurally very similar ligands, as well as modest affinities, with this being reduced by an approximate factor of ten toward a few other structurally similar ligands [13–15].

Heterologous expression and crystallization of CxaP

For crystallization, the protein was heterologously expressed. After the exchange of the elution buffer to sodium phosphate buffer (20 mM, pH 7.2, 100 mM NaCl), crystallization experiments were performed as described in the [Materials and methods](#). Crystals were

obtained with a pure CxaP protein solution and CxaP, which was supplemented with 10 mM D-galactonic, D-gluconic, D-fuconic or D-xylonic acid, respectively, prior to crystallization at similar conditions (see [Materials and methods](#)).

Crystals of the CxaP protein diffracted to a maximum resolution of 1.3 Å. This high-resolution dataset was phased using the automated AUTORICKSHAW pipeline employing the automated molecular replacement mode (<http://www.embl-hamburg.de/Auto-Rickshaw>) and the CxaP protein sequence as input. After several rounds of manually model building using COOT [16] and subsequent refinement using REFMAC5 (<https://www.ccp4.ac.uk>)/PHENIX (<https://www.phenix-online.org>), the structure of the full-length CxaP protein was modeled in terms of the electron density. The CxaP structure was refined to 1.3 Å resolution and the resulting R_{work} and R_{free} values were 13.9% and 16.1%, respectively. Interestingly, within the protein, a ligand was found to be bound later identified as 2-keto-3-deoxy-D-gluconic acid (see below). This ligand was not added to the protein and apparently was copurified with the protein during the whole purification and crystallization procedure. Therefore, we termed this structure CxaP:2-keto-3-deoxy-D-gluconic acid. Once the refinement of the 1.3 Å structure of the CxaP:2-keto-3-deoxy-D-gluconic acid protein was finished, it was used to phase the other datasets obtained from crystals of the CxaP protein co-crystallized with D-galactonic acid, D-xylonic acid, D-fuconic acid or D-gluconic acid. Adding the ligands to the CxaP protein also yielded high-resolution crystal structures that were subsequently refined. The CxaP:D-galactonic crystal was solved at 2.0 Å (R_{work} and R_{free} values were 20.3% and 25.7%, respectively); CxaP:D-xylonic acid crystal was solved at 1.7 Å (R_{work} and R_{free} values were 18.2% and 20.2%, respectively); CxaP:D-fuconic crystal was solved at 2.0 Å (R_{work} and R_{free} values were 20.8% and 26.7%, respectively); and CxaP:D-gluconate crystal was solved at 2.0 Å (R_{work} and R_{free} values were 18% and 23.8%, respectively). A summary of the data collection statistics, refinement details and model content of these different CxaP structures is provided in Table 1.

Interestingly, the crystal parameters, especially the unit cell dimensions and space group, differ between the crystals (Table 1), resulting in a different symmetry of the CxaP D-fuconic acid co-crystals. Here, the symmetry was found to be $P2_1$, whereas all other datasets display a $P2_12_12_1$ symmetry. All crystals contained a monomer of CxaP in the asymmetric unit except the CxaP:D-fuconic acid, where two monomers were found.

Overall structure of CxaP

We compared the individual structures of CxaP with its substrates and determined rmsd values that range from 0.15 to 0.3 Å over 264 Cα atoms, indicating that there were no large conformational changes present. Because the monomers of the different DctP crystal structures are almost identical in overall shape, we describe the general monomeric fold for the highest resolution structure, namely CxaP:2-keto-3-deoxy-D-gluconic acid, which was solved at 1.3 Å in detail as outlined below.

CxaP comprises 335 amino acids and, as expected for a TRAP transporter binding protein, it is composed of two domains, which are colored blue (domain I, residues 1–143 and 247–260) and light brown (domain II, residues 159–236 and 292–335), respectively (Fig. 1). Each of these domains is composed of a central antiparallel five-stranded β-sheet. Furthermore, domain I contains six α-helices, whereas domain II harbors eight α-helices and one additional β-strand (Fig. 1), which flank the central sheet (green in Fig. 1) on both sites. This strand order is found in many other SBPs of TRAP transporters, as well as in class II ATP-binding cassette-dependent SBPs [11,12]. The domains are located non-collinearly on the gene sequence and helix α13 (amino acids 256–291) and β-strands β4 (amino acids 143–158) and β9 (amino acids 237–247) connecting both domains together (Fig. 1).

To identify the structurally closest homologs of the CxaP protein, we performed a DALI search [17,18], which revealed a variety SBP of TRAP transporters, of which the top 10 are listed in Table 2. Interestingly, the Z-scores are quite high, ranging from 29.9 to 38.4, despite the relatively low sequence identity, which ranges from 15% to 26%. This highlights that SBP from TRAP transporters adopt highly similar three-dimensional structures, suggesting that the venus fly trap mechanism for capturing the ligand remains the same. The overall fold of the CxaP protein places this protein in class IIIa of SBPs, which is a subgroup containing mainly SBPs from TRAP transporters [11,12]. Among the 10 homologous structures, five are bound with a rather linear ligands, similar to the CxaP protein homolog from *Pseudomonas aeruginosa* with bound L-malate [Protein Data Bank (PDB) code: 4NF0] [8], from *Chromohalobacter salexigens* with bound D-alanine-D-alanine (PDB code 4N8G) [8], from *Desulfovibrio salexigens* with bound diglycerolphosphate (PDB code 4N6K) [8], from *Rhodoferrax ferrireducens* with bound malonate (PDB code 4MCO) and from *C. salexigens* with bound (S)-(+)-2-amino-1-propanol (PDB code 4XF5). The others contain a ring-like ligand structure similar to UehA from *Silicibacter pomeroyi* DSS-3 [19] with bound ectoine (PDB code 3XFB), as well as the

Table 1. Data collection and refinement statistics.

	2-keto-3-deoxy-D-gluconic acid	D-galactonic acid	D-xylonic acid	D-fuconic acid	D-gluconic acid
PDB code	7BBR	7BCR	7BCN	7BCO	7BCP
Data collection					
Wavelength (Å)	0.967	0.967	0.967	0.967	0.967
Space group	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>P</i> 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2 ₁
Cell dimensions					
<i>a</i> , <i>b</i> , <i>c</i> (Å)	59.76 61.85 88.25	59.03 62.21 88.64	58.11 62.35 97.64	58.16 92.08 62.37	57.93 62.31 98.19
α , β , γ (°)	90.0 90.0 90.0	90.0 90.0 90.0	90.0 90.0 90.0	90. 90.7. 90.0	90.0 90.0 90.0
Resolution (Å)	42.98–1.3 (1.34–1.3)	49.14–2.0 (2.07–2.0)	31.17–1.7 (1.76–1.7)	31.24–2.0 (2.07–2.0)	32.1–2.0 (2.07–2.0)
<i>R</i> _{merge} ^a	0.05 (0.58)	0.09 (0.33)	0.05 (0.42)	0.03 (0.29)	0.04 (0.29)
<i>R</i> _{pim}	0.03 (0.33)	0.04 (0.13)	0.03 (0.22)	0.028 (0.21)	0.023 (0.14)
CC _{1/2}	99.9 (85.4)	0.99 (0.96)	0.99 (0.86)	0.99 (0.87)	0.99 (0.95)
Mean <i>I</i> / σ (<i>I</i>)	14.9 (3.15)	12.49 (4.22)	14.09 (2.85)	13.57 (2.47)	22.17 (5.99)
Completeness (%)	98.44 (96.78)	99.18 (98.78)	96.55 (97.43)	92.70 (97.64)	95.73 (99.75)
Multiplicity	5.8 (6.1)	7.0 (7.2)	4.1 (4.4)	2.3 (2.3)	4.9 (5.1)
Total reflections	599 917	156 928	159 256	95 881	15 539
<i>R</i> _{work} ^b	0.13	0.20	0.18	0.21	0.18
<i>R</i> _{free}	0.16	0.26	0.20	0.27	0.24
Number of atoms					
Protein	2420	2404	2404	4808	2404
Ligands	12	13	22	24	13
Solvent	562	291	474	370	287
Protein residues	310	310	310	620	310
<i>B</i> -factors					
Macromolecules	16.63	28.83	24.57	34.97	28.46
Ligands	10.98	24.10	17.18	27.61	22.09
Solvent	30.74	36.32	35.87	40.74	35.80
Wilson <i>B</i>	19.26	29.62	26.36	35.34	29.21
rmsd					
Bond lengths (Å)	0.007	0.007	0.006	0.009	0.011
Bond angles (°)	0.97	0.81	0.82	1.11	1.24
Ramachandran					
Favored (%)	98.70	97.08	98.70	97.40	98.38
Outliers (%)	0.00	0.00	0.00	0.00	0.00

Values for the highest resolution shell are shown in parenthesis.

^a $R_{\text{merge}} = \sum hkl \sum i |I(hkl) - \langle I(hkl) \rangle| / \sum hkl \sum i I(hkl)$, where $I(hkl)$ is the intensity of the i th observation of reflection hkl and $\langle I(hkl) \rangle$ is the average of overall observations of reflection hkl ; ^b $R_{\text{factor}} = \sum hkl |F_o - F_c| / \sum hkl F_o$ for all data excluding the 5% that comprised the R_{free} used for cross-validation.

DctP homologs from *Ruegeria pomeroyi* with bound 3,4-dihydroxybenzoate (PDB code 4PAF), from *R. ferrireducens* with bound benzoate (PDB code 5IM2) and from *Bordetella bronchiseptica* with bound picolinic acid (PDB code 4YIC) [8,19]. This suggests that although the overall structure is closely related, the binding site of the ligand discriminates between a linear and a ring-like ligand.

The ligand binding site of CxaP

In the SBP of TRAP transporters, the ligands are bound in the cleft flanked by the two domains of the protein. In the CxaP protein, which was crystallized in

the apo-form, a ligand was bound that could be identified as 2-keto-3-deoxy-D-gluconate (KDG) as a result of the high resolution. This was unexpected and suggests that KDP was copurified after expression in *Escherichia coli* and tightly bound to the CxaP protein (Fig. 2).

KDG is bound in the binding cleft in between and via interaction with the two domains. Interaction with the domain I is mediated via the side chains of Arg175, Arg154, Asn215, and Asn151 (Fig. 2, shown in light brown). Other interactions arise from domain II and are mediated by the side chains of Ser96 and Asp78, as well as via the backbone interaction of

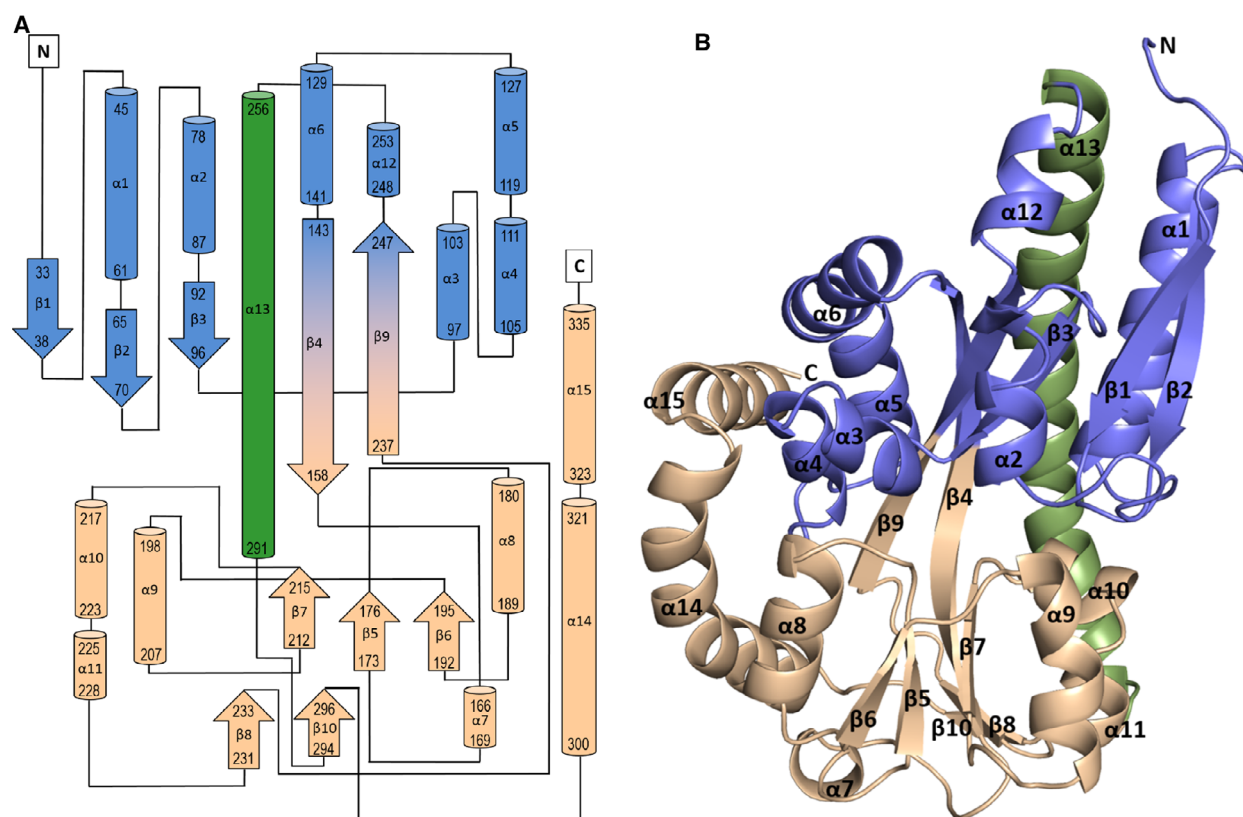


Fig. 1. Overall structure of CxaP. (A) A topology diagram of CxaP where domain I is colored in blue, domain II is colored in beige and the alpha helix is shown in green. (B) Overall structure of CxaP. The N and C termini are labeled. The structures were generated using PYMOL 2.4.0.

Table 2. TopTen of structures found using the DALI server using the CxaP structure as template (<http://ekhidna2.biocenter.helsinki.fi/dali/oldstyle.html>).

Protein	Organism	Bound ligand	Z-score	rmsd	Alignment of amino acids	Percentage (%) seq identity	PDB code
UehA _{Sp}	<i>Silicibacter pomeroyi</i> DSS-3	Ectoine	38.4	1.7	302	23	3XFB
DctP _{Pa}	<i>Pseudomonas aeruginosa</i>	L-malate	37.8	1.7	290	26	4NF0
DctP _{CS}	<i>Chromohalobacter salexigens</i>	D-alanine-D-alanine	37.8	1.8	303	26	4N8G
DctP _{Rp}	<i>Ruegeria pomeroyi</i>	3,4-Dihydroxybenzoate	37.4	2.0	398	22	4PAF
DctP _{DS}	<i>Desulfovibrio salexigens</i>	Diglycerolphosphate	36.4	2.0	301	18	4N6K
DctP _{Rf}	<i>Rhodoferrax ferrireducens</i>	Malonate	35.3	2.0	301	18	4MCO
DctP _{Bb}	<i>Bordetella bronchiseptica</i>	Benzoate	33.6	2.2	301	19	5IM2
DctP _{Bb}	<i>Bordetella bronchiseptica</i>	Picolinic acid	31.3	2.7	305	18	4YIC
DctP _{Chs}	<i>Chromohalobacter salexigens</i>	(S)-(+)-2-amino-1-propanol	31.0	2.6	300	12	4XF5
DctP _{Bbr}	<i>Bordetella bronchiseptica</i>	–	29.9	3.4	297	15	4N4U

Tyr38, Gly39 and Leu40. These amino acids also play a role in the binding of the ligands (Fig. 3).

The presence of the KDG intermediate of the Entner–Doudoroff pathway KDG in the binding pocket was unexpected because the ligand structure differs from the known sugar acid ligands of CxaP. The Entner–Doudoroff pathway is used for gluconate

metabolism in *E. coli* and KDG is most probably formed during cultivation for protein expression [20]. This is not the first time that an unexpected molecule was identified in SBP protein crystals. Binding of glycerol to an α -keto acid-binding SBP of *Rhodobacter sphaeroides* has been reported previously [15].

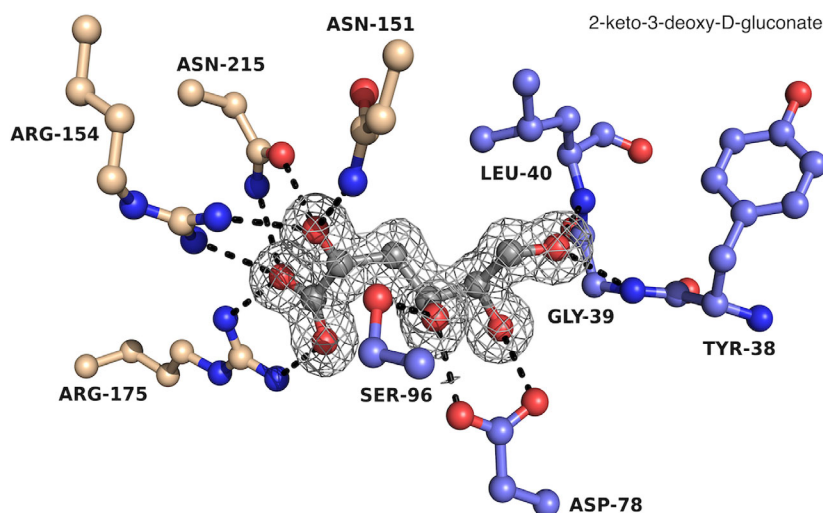


Fig. 2. Binding site of CxaP 2-keto-3-deoxy-D-gluconate binding site shown as ball and sticks. Residues of domain I (Ser96, Asp78, Gly39, Leu40) are colored in blue and those of domain II (Asn151, Asn215, Arg154, Arg175, Met177) are colored in beige. The electron density is shown in gray and is contoured at 1.0 sigma. The structures were generated using PYMOL 2.4.0.

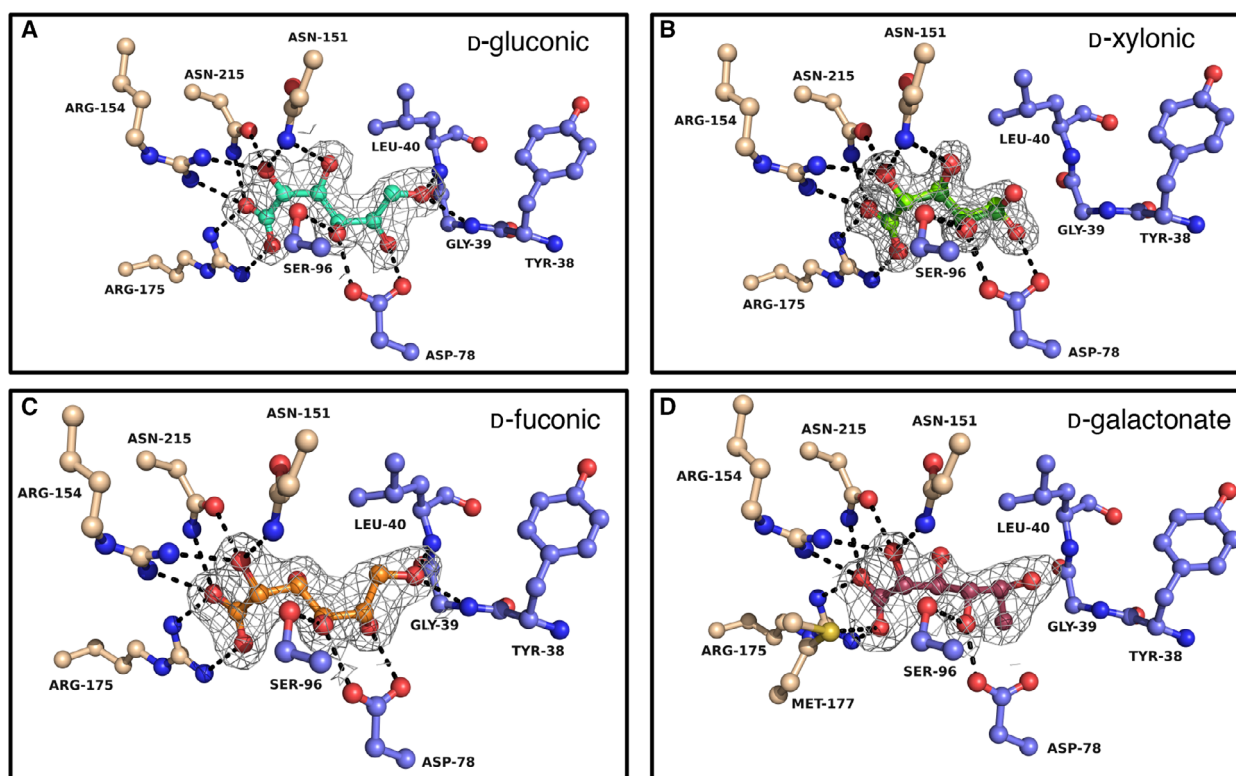


Fig. 3. Substrate binding of CxaP The binding site for D-gluconic (cyan) (A), D-xylonic (green) (B), D-fuconic (orange) (C) and D-galactonic (dark red) (D) acid are shown as ball and sticks. Residues of domain I (Ser96, Asp78, Gly39, Leu40) are colored in blue and those of domain II (Asn151, Asn215, Arg154, Arg175, Met177) are colored in beige. The electron density is shown in gray and is contoured at 1.0 sigma. The structures were generated using PYMOL 2.4.0.

By comparison of the interactions of KDG with the amino acid residues in the binding pocket of CxaP and the interactions observed for the other ligands, it is obvious that KDG shows interactions with amino acid residues similar to those of sugar acids. Approximate dissociation constants for the interaction of CxaP and

the sugar acid ligands have been described previously, with the values appearing to be approximately 10-fold higher compared to other TRAP transporters [6,10–12]. Because the protein used for determination of the dissociation constant (K_d) was synthesized by *E. coli* in complex medium and prebound ligands

potentially have not been removed [10], the observation of higher K_d values is most probably a result of prebound KDG in the protein samples. The prebound KDG is thereby most likely displaced by the sugar acid ligands. To our knowledge, this is the first report of an interaction of KDG with an SBP of a TRAP transporter. By contrast to the known ligands of the CxaP, KDG possesses a keto group at position C2 and no functional group at position C3. This indicates that the substrate range of CxaP could include further ligands, which probably have lower affinities but still might be accepted by the protein.

It was reported earlier that the selectivity of CxaP is based on a carboxylic group and a hydroxyl group at C2 in *R*-configuration and a hydroxylic group positioned at C3 in *S*-configuration [10]. In general, the architecture of the binding site of CxaP confirmed this observation (Fig. 2). However, it was observed that KDG was also accepted as a ligand in the binding pocket, as discussed further below.

We set out to crystallize CxaP with different bound ligands and succeeded to crystallize the CxaP in complex with D-gluconic, D-xylonic, D-fuconic and D-galactonic acid. Crystals were grown under similar conditions as the wild-type crystals. Datasets were obtained at high resolution, and the already determined structure of CxaP in complex with KDG was used as a template to solve the structures with the different ligands (Table 1).

The carboxylic group of the different sugar acids interacts with a highly conserved Arg175 (Fig. 3), which is typical for TRAP SBPs and is present in all known structures. Furthermore, the Arg residue has been shown to be crucial for the high selectivity toward carboxylic acid ligands [21]. The carboxylic group also interacts with another arginine residue (Arg154) and an asparagine (Asn215). These interactions are not exclusively present in CxaP. Similar interactions including two Arg and an Asn have also been reported for SiaP, a sialic acid-binding protein from *Haemophilus influenzae* [13]. Arg154 and Asn215 additionally interact with the hydroxylic group at position C2. In addition, in the case of the C6-ligands, another interaction of Asn151 at position C2 could be observed.

The ligands are stabilized at position C3 by an interaction with either Ser96 (Fig. 3). Additionally, all sugar acid ligands are stabilized by hydrogen bonds between the backbone of the ligand in position C4 and Asp78.

Activity and CxaP variants

To confirm the function of the amino acids in the binding pocket of CxaP, single-exchange mutants were

generated via site-directed mutagenesis. All proteins were purified resulting in pure protein as confirmed by SDS/PAGE. Especially, the variant Asp78Leu appeared to be less stable and resulted in significantly less protein compared to the others. The targeted amino acids were Gly39, Asp78, Ser96, Arg154, Arg175, Phe198 and Asn215. Three different mutants were generated for Gly39. This was changed to alanine, leucine and valine to determine whether the size of the amino acid at this site has an impact on the activity of the protein toward differently sized ligands. Asp78 was changed to leucine and Ser96 was converted to alanine and valine, respectively. The Arg154 residue was mutated to alanine and Arg175 was changed to lysine. Isoleucine was introduced for Phe198 and Asn215 was changed to leucine. The mutations were chosen based on sequence alignment using sequence of different homologous proteins (Fig. 4). Here, the different CxaP homolog sequences that were aligned appear to have rather narrow substrate specificity.

K_d values of wild-type CxaP and the generated variants were determined to analyze the impact of the different amino acid exchanges on protein and substrate binding via a fluorescence binding assay. The analyzed ligands were D-gluconic, D-galactonic, D-fuconic and D-xylonic acids as ligand for CxaP. All calculated dissociation constants are shown in Table 3. The wild-type protein binds D-gluconic acid with a K_d value of 8.0 μM ; for D-galactonic acid, the wild-type K_d value was 6.6 μM . D-fuconic acid was bound to CxaP with a K_d value of 1.9 μM and, with D-xylonic acid, a dissociation constant of 3.9 μM could be determined. To determine the K_d value, the measured fluorescence decrease (%) is plotted against the used substrate concentration (μM). Fits for the K_d value determination for all of the tested ligands were generated using ORIGIN (OriginLab, Northampton, MA, USA). In Fig. 5, the curves for the wild-type CxaP are shown. All of the calculated K_d values for the CxaP variants were determined accordingly. In summary, the K_d values calculated for the wild-type protein CxaP were in the range 2–8 μM for the different ligands which is in line with previous studies [10]. In most cases, the K_d value of the generated variants is higher than the K_d of the wild-type CxaP, indicating a lower affinity (Table 3). Especially, the variants Arg154Ala, Arg175Lys, and Phe198Ile led to an increase of the dissociation constant with most ligands or even to a lack of interaction, which can be explained by the important interactions being observed in the wild-type structures. For Gly39Val, the K_d values of the different ligands were two- to six-fold higher compared to the wild-type

The FEBS Journal **288** (2021) 4905–4917 © 2021 The Authors. The *FEBS Journal* published by John Wiley & Sons Ltd on behalf of Federation of European Biochemical Societies.

K_d . The K_d values of Arg154Ala were approximately five-fold higher compared to the wild-type with D-gluconic and D-xylonic acid and 10-fold higher using D-galactonic acid as the ligand. No interaction between the protein and the ligand was detectable either for Arg175Lys with D-gluconate and D-xylonic acid or for Phe198Ile with D-fuconic and D-xylonic acid. The K_d value of Phe198Ile with D-gluconic acid was approximately three-fold higher compared to the K_d of the wild-type with the same ligands (Table 3). In summary, however, the observed changes of the K_d values are similar. This is likely a result of the multiple interactions present for binding a sugar acid ligand. A loss of a single interaction does not lead to a dramatic loss of affinity. This again highlights that CxaP has a wide substrate range and the binding site appears to be optimized with the ability to accommodate several different ligands, although these are of the same nature. This is striking because CxaP is unique in being able to bind a variety of ligands with a very similar affinity. Usually, other proteins of the TRAP transporter family are optimized for one specific substrate and can bind other ligands with a lower affinity.

Function of Gly39

An interesting amino acid in the binding pocket of CxaP is Gly39. This residue directly interacts with the hydroxylic groups of D-gluconic acid and KDG and also with the methyl group of D-fuconic acid, in each case at position C6. The shorter pentanoic acid D-xylonic acid does not interact with Gly39 at all. This indicates that the Gly39 determines the size of accepted ligands with regard to the chain length. This can also be deduced from the sequence alignment with protein displaying a narrow substrate spectrum (Fig. 4). Here, the Gly residue is not conserved throughout the proteins. Instead, it appears to be exchanged to other amino acids, such as Asp, Asn, Tyr, Leu and Val. By intruding any other amino acid, a side chain would point to within the binding site, emphasizing that this might alter the substrate specificity. We confirmed this hypothesis by introducing three variations at the Gly39 position. Here, glycine was mutated to alanine (Gly39Ala), leucine

(Gly39Leu) or valine (Gly39Val). After expression and purification, the K_d values for the substrates D-gluconic, D-galactonic, D-fuconic and D-xylonic acid were measured (Table 3). Especially, the introduction of alanine and leucine appeared to have a stabilizing effect on D-xylonate. Here, the affinity toward this substrate increased. The wild-type protein displayed a K_d of $3.9 \pm 0.5 \mu\text{M}$. Whereas, in the variant Gly39Ala this was reduced to $2.4 \pm 1.5 \mu\text{M}$. The strongest change, however, was observed for the Gly39Leu mutation where this affinity was even further strengthened resulting in a further drop of the K_d to $1.3 \pm 0.1 \mu\text{M}$. It appears that the side chains introduced are interacting with the D-xylonic acid substrate giving a stabilizing effect. A further increase of the side chain appears to hinder the binding capacity for this substrate as observed by the lowered K_d for the Gly39Val mutation and the D-xylonic acid substrate. Here, the K_d drops even below the K_d measured for the wild-type protein. A similar effect was observed when galactonic acid was used as substrate with the corresponding mutations, although to a smaller extent. Because the other used substrates were already interacting with the backbone of Gly39, the introduction of another amino acid (e.g. side chain) resulted in a sterically unfavored binding, as observed by lower affinity for the Gly39 mutation toward these substrates.

Conclusions

CxaP was described as the SBP of a TRAP transport system with a selectivity filter for sugar acids involving particular characteristics. By analysis of the crystal structure, the binding of the ligands depending on a carboxylic group at position C1, as well as the position of the hydroxylic groups at C2 and C3 in *R*- and *S*-configuration [10], respectively, was confirmed. However, the substrate range must be expanded by a ligand with a less similar structure, KDG, which is most probably accepted by the SBP with modest affinity. Further analysis of the binding pocket of CxaP is desirable to confirm these observations.

The present study is the first description of the structure of a TRAP type SBP accepting sugar acid ligands with the described structural properties. The

Fig. 4. Multiple sequence alignment of CxaP with different DctP proteins. The identical amino acids are given as a percentage (%). The ligand of the respective proteins is given in the lower right corner. The red boxes highlight the amino acid residues probably involved in ligand binding. If available, the PDB code is given in front of the genus. Sequence are derived from the UniProt database (<https://www.uniprot.org>) with the accession numbers: SMa0252: Q930R1; 4N17 (PDB entry): Q0B2F6 (uniprot entry); 4OVR: A7IKQ4; 4OVS: D1AZL7; 4MHF: Q128M1; 4MIJ: Q128M; 4OVQ: Q160Z9; 4PBQ: HICG_00826; 4N8Y: A5E8D2; 4PF8: A3T0C3; 4P1L: Q1QUN2; 4O94: Q2IUT5; 4PCD: Q16BC9; 4PC9: Q16BC9; 4OA4: A3QCW5; 4OVT: A6X5V3; 4O8M: A6VKP1; 4N91: C7RDZ3. The alignment was made with CLUSTAL OMEGA (<https://www.ebi.ac.uk/Tools/msa/clustalo>).

Table 3. K_d values of CxaP wild-type in comparison with the K_d values of the generated CxaP mutants using different substrates. All K_d values are given in μM and represent data from at least three independent biological experiments. For some variant and ligand combinations, substrate binding was not determined (ND).

Mutant	D-gluconic	D-galactonic	D-fuconic	D-xylonic
WT	8.0 ± 2.7	6.6 ± 1.3	1.9 ± 0.4	3.9 ± 0.5
G39A	14.7 ± 5.0	4.4 ± 0.4	2.2 ± 0.4	2.4 ± 1.5
G39L	17.1 ± 6.6	5.2 ± 3.1	3.7 ± 1.9	1.3 ± 0.1
G39V	16.0 ± 8.4	18.5 ± 2.7	12.6 ± 5.6	12.1 ± 6.1
D78L	15.7 ± 5.1	13.2 ± 2.8	7.1 ± 1.7	6.9 ± 0.7
S96A	1.6 ± 0.5	12.6 ± 9.2	6.2 ± 1.0	17.8 ± 6.6
S96V	33.3 ± 7.5	9.5 ± 4.3	14.3 ± 6.3	15.2 ± 11.5
R154A	42.6 ± 17.2	ND	4.8 ± 0.8	26.4 ± 13.4
R175K	ND	9.9 ± 5.7	21.8 ± 10.4	ND
F198I	25.6 ± 12.7	3.5 ± 0.5	ND	ND
N215L	5.1 ± 0.6	6.1 ± 4.9	3.8 ± 1.3	6.9 ± 0.3

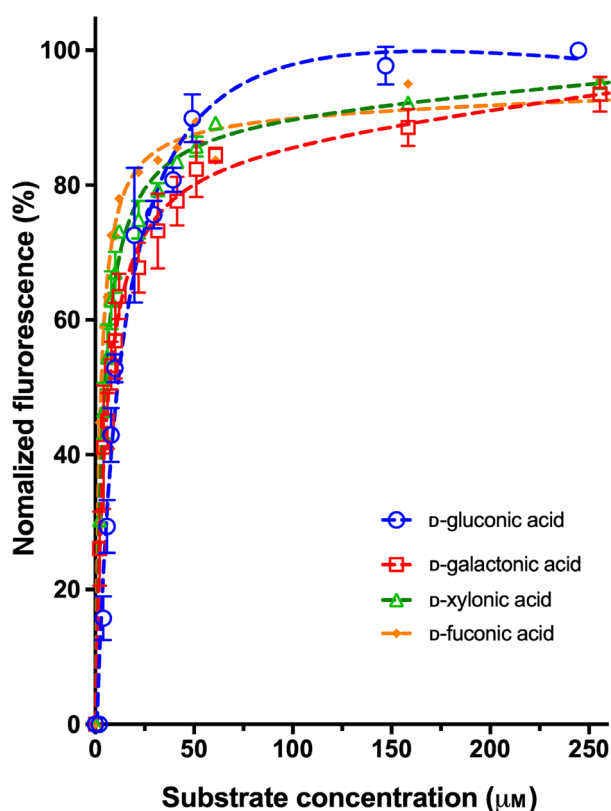


Fig. 5. Exemplary fits for the wild-type K_d for each substrate used. The measured fluorescence decrease was plotted against the substrate concentration used. The fit was generated using ORIGIN (OriginLab). The K_d values for the CxaP variants were determined accordingly. The curves are colored as: D-gluconic acid (blue), D-galactonic acid (red), D-xylonic acid (green) and D-fuconic acid (orange). The error bars displayed are the SDs calculated from three (n) independent biological triplicates.

majority of the SBPs of TRAP transporters occur as monomers, although stable dimers have also been reported [15]. CxaP has a monomeric structure. The interaction of the ligand with a highly conserved arginine residue is essential for the coordination and binding of the solute in SBPs of TRAP transporters [21] and has been reported for all characterized TRAP transporters and confirmed in the obtained crystals of CxaP. A special feature of CxaP is the glycine residue in position 39, which allows multiple ligands to bind to the substrate binding pocket. This is not common for other proteins of the TRAP transporter family and was achieved via the function of Gly39 in the binding pocket of CxaP. Mutational analysis of this amino acids revealed changes in the affinity toward some of the ligands.

To our knowledge, the interaction of an SBP of a TRAP transporter with its sugar acid ligands has only been reported in such detail three times to date [3,8,10]. Only amino acid sequences for which experimental data are available have been chosen for the preparation of the alignment. By comparison of the amino acid sequences, it is striking that the amino acid residues, which are directly involved in the coordination of the ligand, are highly conserved in the sequences of the D-gluconic acid accepting SBPs. To our knowledge, the uptake strategy, consisting of oxidation in the periplasm and subsequent import via CxaP and its corresponding transport system, has only been described for *A. mimigardefordensis* DPN7^T [10,22].

Materials and methods

Cultivation of bacterial strains

Cells of *E. coli* strains (Table 4) were cultivated in lysogeny-broth medium at 30 °C or 37 °C supplemented with addition of ampicillin (Amp, $75 \mu\text{g}\cdot\text{mL}^{-1}$), if necessary. Liquid cultures were incubated in baffled flasks on a rotary shaker at 130 r.p.m. *E. coli* Top10 was used for cloning procedures and plasmid propagation (Table 4). *E. coli* BL21 (DE3) pLysS was used for heterologous expression of *cxaP* and the respective mutants. Solid medium contained 1.8% (w/v) agar.

Expression and purification of CxaP

Heterologous expression of CxaP was accomplished in *E. coli* BL21 (DE3) by applying pET19b:*cxaP* as described before [10]. For this, a preculture of 20 mL of LB medium containing ampicillin in a concentration of $75 \mu\text{g}\cdot\text{mL}^{-1}$ was grown overnight. A main culture of 100 mL ZYP-

Table 4. Strains and plasmids used in the present study.

Strain	Description	Reference
<i>E. coli</i> Top 10	F ⁻ <i>mcrA</i> Δ(<i>mrr-hsdRMS-mcrBC</i>) <i>rpsL</i> <i>nupG</i> ϕ 80 <i>lacZ</i> Δ <i>M15</i> Δ <i>lacX74</i> <i>deoR</i> <i>recA1</i> <i>araD139</i> Δ(<i>ara-leu</i>)7697 <i>galU</i> <i>galK</i> <i>endA1</i>	Invitrogen (Carlsbad, CA, USA)
<i>E. coli</i> BL21 (DE3)	F ⁻ <i>ompT</i> <i>hsdS</i> _B (<i>r</i> _B ⁻ <i>m</i> _B ⁻) <i>gal</i> <i>dcm</i> (DE3)/pLysS (Cm ^r)	Novagen (Madison, WI, USA)
pET-19b	pBR22 <i>ori</i> . Ap ^r . T7 <i>lac</i> . His ₆ -N- terminal tag	Novagen (Madison, WI, USA)
pET-19b:: <i>dctP</i> _{Am}	Template for site-directed mutagenesis	[10]

autoinduction medium was inoculated with 1 mL of this preculture [23] and grown for 16–20 h on a rotary shaker at 130 r.p.m. and 30 °C. Cells were harvested by centrifugation (10 min, 4 °C, 7000 g) and washed in 20 mL of sodium phosphate buffer (20 mM, pH 7.2) containing 20 mM imidazole and 500 mM NaCl. After washing, the cell pellet was resuspended in 5 mL of sodium phosphate buffer and then cells were disrupted with three passages through a French press. Insoluble components of the lysate were removed by centrifugation for 30 min at 21 000 g and 4 °C. CxaP was purified as described previously [10]. Prior to the crystallization experiments, the elution buffer from the purification was exchanged to 20 mM sodium phosphate buffer containing 100 mM NaCl using Vivaspin 500 tubes (GE Healthcare, Freiburg, Germany).

Table 5. Oligonucleotides used for generation of *cxap* variants.

Oligonucleotide	Sequence (5'- to 3')	Amino acid substitution
G39A_for	CTTTGGCTACGCGCTGGCCGATG	G39A
G39A_rev	CGGATAATGCGCGGCTTGATGTCC	
G39L_for	CTTTGGCTACGTACTGGCCGATG	G39L
G39L_rev	CGGATAATGCGCGGCTTGATGTCC	
G39V_for	CTTTGGCTACGTACTGGCCGATG	G39V
G39V_rev	CGGATAATGCGCGGCTTGATGTCC	
D78L_for	GGGACCTCTGGAACAGTTGATC	D78L
D78L_rev	AGGGCGCCATTGCCGAAAGTC	
S96A_for	CACATTCGTGGCGACCGCACC	S96A
S96A_rev	CGGTGCGGTCGCCACGAATGTG	
S96V_for	CACATTCGTGGTAACCGCACCG	S96V
S96V_rev	CGGTGCGGTTACCACGAATGTG	
R154A_for	GAAAATGGCTTCGCGAATATCACCAAC	R154A
R154A_rev	GTTGGTGATATTCGCGAAGCCATTTTC	
R175K_for	GCATTAAACTGAAAGTGATGCAG	R175K
R175K_rev	CGCCAATATCATCCAGTTTGAG	
F198I_for	GATTCCAATGCCGATTACGGAAC	F198I
F198I_rev	CGCATTGGCACCCAGTCCC	
N215L_for	GATGGTCAGGAAC TGCCGCTGTC	N215L
N215L_rev	TACGGTTTTGGTTTCAGCGCAGTGAAC	

The protein concentration was determined via a Bradford assay [24] and the purity of the protein was analyzed via SDS/PAGE.

Site-directed mutagenesis

Site-directed mutagenesis was used for amino acid substitutions of the respective CxaP variants. For this, a pair of phosphorylated primers was generated (Table 5) to substitute a certain amino acid DNA triplet via point mutation. The forward primer introduced the desired mutations into the nucleotide sequence. The reverse primer was located adjacently. In a polymerase chain reaction using pET-19b::*cxap* as template, a linear fragment consisting of the complete plasmid sequence was amplified. The polymerase chain reaction was performed using Phusion polymerase (Fermentas, St Leon-Rot, Germany) and Phusion High Fidelity buffer supplemented with 3% (v/v) dimethylsulfoxide. Afterward, the parental plasmid was digested using *DpnI* (Thermo Scientific, Darmstadt, Germany) and the linear fragments were ligated using T4-Ligase (Thermo Scientific). Subsequently, *E. coli* Top10 was transformed with the mutated plasmid. All plasmids were verified by sequencing. The following protein variants were generated: G39A, G39L, G39V, D78L, S96A, S96V, R154A, R175K, F198I and N215L.

Crystallization of CxaP

Protein samples for crystallization were initially screened by applying a commercial screen (Molecular Dimensions,

Newmarket, UK) in sitting drop 96-well plates at 12 °C by applying the sitting drop vapor diffusion method. Therefore, 0.1 µL of a CxaP protein solution (18 mg·mL⁻¹) was mixed with 0.1 µL of reservoir solution and equilibrated against 40 µL of reservoir solution. The most promising conditions were found with a solution containing 0.3 M magnesium nitrate hexahydrate, 0.1 M Tris (pH 7.0–7.4) or 0.002 M zinc sulfate heptahydrate, 0.08 M Hepes (pH 7) and 25% Jeffamine. The crystals of the initial screening obtained in this condition were shredded and used in microbatch set-ups (1 µL + 1 µL drops, 300 µL of reservoir solution) as seed stock in the optimization of the crystallization conditions in grid screens around the initial conditions, varying the concentrations. Thereby, protein crystals were obtained by incubation of CxaP with the 100 µM D-fuconic acid, D-xylonic acid and 1 mM D-gluconic acid and D-galactonic acid for 1 h on ice prior to crystallization. The protein crystals formed were fished after overlaying the drop with mineral oil and then flash-frozen in liquid nitrogen.

Data processing and structure determination

Data sets were collected from a single crystal of either CxaP or CxaP in complex with the different ligands at the ID29 beamline (European Synchrotron Radiation Facility, Grenoble, France) at 100 K. Initial images were collected and a strategy was calculated using the beamline software [25]. These data sets were processed using the XDS package and scaled with XSCALE [26,27]. Initial phases were obtained by molecular replacement using the web server autorickshaw employing solely the sequence of CxaP as template. Model building and refinement were performed using COOT [16] and PHENIXREFINE [28] as part of the PHENIX software. Data refinement statistics and model content are summarized in Table 3. The atomic coordinates and structure factors have been deposited in the Worldwide PDB ([https://www.wwpdb.org](https://www ww p d b . o r g)) under the accession codes: CxaP (DctP_{Am}):2-keto-3-deoxy-D-gluconic acid – PDB code 7BBR; CxaP (DctP_{Am}):D-galactonic acid – PDB code 7BCR; CxaP (DctP_{Am}):D-xylonic acid – PDB code 7BCN; CxaP (DctP_{Am}):D-fuconic acid – PDB code 7BCO; and CxaP (DctP_{Am}):D-gluconic acid – PDB code 7BCP. Illustrations of the crystal structures of the CxaP (DctP_{Am}) protein were prepared using PYMOL (<https://pymol.org>).

Fluorescence spectroscopy

Fluorescence spectroscopy was used to determine the K_d values of CxaP and the generated mutants of CxaP with gluconate, galactonate, fuconate and xylonate. For this, 1 µM protein was used per measurement starting with a substrate concentration of 0.2 µM, which was added to the protein. To achieve the desired substrate concentrations, four different stock solutions were used for each ligand containing 0.1, 1, 5 or 50 mM of the respective substrate. In

every titration step, 2 µL was used, starting with the 1 mM stock solution up to the 50 mM stock solution, to cover the whole concentration range. Fluorescence was measured at 328 nm. After volume correction and normalization, the K_d values were determined using values for the decrease in fluorescence (%) plotted against the substrate concentration (µM). For that, software ORIGIN (OriginLab) was used and the equation $F_0 + F_{amp} \times 0.5 \times (1/P) \times ((P + K_d + x) - ((P + K_d + x)^2 - 4 \times P \times x)^{0.5})$ was applied. F_0 is set as the lowest value for the fluorescence decrease and F_{amp} is set as the amplitude of values for the fluorescence decrease. The used protein concentration is indicated by P (1 µM). The K_d value determination was calculated from at least biological triplicates. As a control, the same experiment was performed with buffer was added instead of one of the substrates and used for background correction.

Acknowledgements

We cordially acknowledge the staff of beamline P14 (Deutsches Elektronen-Synchrotron, European Molecular Biology Laboratory, Hamburg, Germany) for their kind and helpful support during crystal screening and we are equally thankful to the staff of the ID29 beamline of the European Synchrotron Radiation Facility (Grenoble, France) for expertly providing synchrotron radiation facilities. The Center for Structural studies is funded by the Deutsche Forschungsgemeinschaft (grant number 417919780, INST 208/740-1 FUGG). Alexander Steinbüchel is very grateful to the Rahn-Quade-Stiftung, which supported the employment of one PhD student.

Conflict of interests

The authors have no conflicts of interest.

Author contributions

AH and SK performed the crystallization studies and collected the X-ray data. AH and SHJS refined the structure. LS and CM-B isolated, cloned, expressed and purified the protein, and performed the characterization studies. SHJS and AS supervised the study. All authors contributed to the writing of the manuscript.

References

- 1 Postma PW & Lengeler JW (1985) Phosphoenolpyruvate:carbohydrate phosphotransferase system of bacteria. *Microbiol Rev* **49**, 232–269.
- 2 Saier MH Jr (2000) Families of transmembrane sugar transport proteins. *Mol Microbiol* **35**, 699–710.

- 3 Steele TT, Fowler CW & Griffiths JS (2009) Control of gluconate utilization in *Sinorhizobium meliloti*. *J Bacteriol* **191**, 1355–1358.
- 4 Forward JA, Behrendt MC, Wyborn NR, Cross R & Kelly DJ (1997) TRAP transporters: a new family of periplasmic solute transport systems encoded by the dctPQM genes of *Rhodobacter capsulatus* and by homologs in diverse gram-negative bacteria. *J Bacteriol* **179**, 5482–5493.
- 5 Fischer M, Zhang QY, Hubbard RE & Thomas GH (2010) Caught in a TRAP: substrate-binding proteins in secondary transport. *Trends Microbiol* **18**, 471–478.
- 6 Mulligan C, Fischer M & Thomas GH (2011) Tripartite ATP-independent periplasmic (TRAP) transporters in bacteria and archaea. *FEMS Microbiol Rev* **35**, 68–86.
- 7 Mulligan C, Leech AP, Kelly DJ & Thomas GH (2012) The membrane proteins SiaQ and SiaM form an essential stoichiometric complex in the sialic acid tripartite ATP-independent periplasmic (TRAP) transporter SiaPQM (VC1777-1779) from *Vibrio cholerae*. *J Biol Chem* **287**, 3598–3608.
- 8 Vetting MW, Al-Obaidi N, Zhao S, San Francisco B, Kim J, Wichelecki DJ, Bouvier JT, Solbiati JO, Vu H, Zhang X *et al.* (2015) Experimental strategies for functional annotation and metabolism discovery: targeted screening of solute binding proteins and unbiased panning of metabolomes. *Biochemistry* **54**, 909–931.
- 9 Kelly DJ & Thomas GH (2001) The tripartite ATP-independent periplasmic (TRAP) transporters of bacteria and archaea. *FEMS Microbiol Rev* **25**, 405–424.
- 10 Meinert C, Senger J, Witthohn M, Wubbeler JH & Steinbuchel A (2017) Carbohydrate uptake in *Advenella mimigardefordensis* strain DPN7(T) is mediated by periplasmic sugar oxidation and a TRAP-transport system. *Mol Microbiol* **104**, 916–930.
- 11 Berntsson RP, Smits SH, Schmitt L, Slotboom DJ & Poolman B (2010) A structural classification of substrate-binding proteins. *FEBS Lett* **584**, 2606–2617.
- 12 Scheepers GH, Lycklama ANJA & Poolman B (2016) An updated structural classification of substrate-binding proteins. *FEBS Lett* **590**, 4393–4401.
- 13 Muller A, Severi E, Mulligan C, Watts AG, Kelly DJ, Wilson KS, Wilkinson AJ & Thomas GH (2006) Conservation of structure and mechanism in primary and secondary transporters exemplified by SiaP, a sialic acid binding virulence factor from *Haemophilus influenzae*. *J Biol Chem* **281**, 22212–22222.
- 14 Salmon RC, Cliff MJ, Rafferty JB & Kelly DJ (2013) The CoupSTU and TarPQM transporters in *Rhodopseudomonas palustris*: redundant, promiscuous uptake systems for lignin-derived aromatic substrates. *PLoS One* **8**, e59844.
- 15 Gonin S, Arnoux P, Pierru B, Lavergne J, Alonso B, Sabaty M & Pignol D (2007) Crystal structures of an extracytoplasmic solute receptor from a TRAP transporter in its open and closed forms reveal a helix-swapped dimer requiring a cation for alpha-keto acid binding. *BMC Struct Biol* **7**, 11.
- 16 Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D* **66**, 486–501.
- 17 Holm L (2020) Using Dali for protein structure comparison. *Methods Mol Biol* **2112**, 29–42.
- 18 Holm L (2020) DALI and the persistence of protein shape. *Protein Sci* **29**, 128–140.
- 19 Lecher J, Pittelkow M, Zobel S, Bursy J, Bonig T, Smits SH, Schmitt L & Bremer E (2009) The crystal structure of UehA in complex with ectoine-A comparison with other TRAP-T binding proteins. *J Mol Biol* **389**, 58–73.
- 20 Peekhaus N & Conway T (1998) What's for dinner?: Entner-Doudoroff metabolism in *Escherichia coli*. *J Bacteriol* **180**, 3495–3502.
- 21 Fischer M, Hopkins AP, Severi E, Hawkhead J, Bawdon D, Watts AG, Hubbard RE & Thomas GH (2015) Tripartite ATP-independent periplasmic (TRAP) transporters use an arginine-mediated selectivity filter for high affinity substrate binding. *J Biol Chem* **290**, 27113–27123.
- 22 Thomas GH (2017) On the pull: periplasmic trapping of sugars before transport. *Mol Microbiol* **104**, 883–888.
- 23 Studier FW & Moffatt BA (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* **189**, 113–130.
- 24 Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248–254.
- 25 Gabadinho J, Beteva A, Guijarro M, Rey-Bakaikoa V, Spruce D, Bowler MW, Brockhauser S, Flot D, Gordon EJ, Hall DR *et al.* (2010) MxCuBE: a synchrotron beamline control environment customized for macromolecular crystallography experiments. *J Synchrotron Radiat* **17**, 700–707.
- 26 Kabsch W (2010) Xds. *Acta Crystallogr D* **66**, 125–132.
- 27 Kabsch W (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D* **66**, 133–144.
- 28 Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH & Adams PD (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D* **68**, 352–367.