Erschienen in: Journal of molecular biology ; 414 (2011), 1. - S. 75-85
https://doi.org/10.1016/j.jmb.2011.09.043

Escherichia coli Peptide Binding Protein OppA Has a Preference for Positively Charged Peptides

M. M. Klepsch [1], M. Kovermann [2,3], C. Löw [4], J. Balbach [2,3], H. P. Permentier [5], F. Fusetti [5], J. W. de Gier [1]*, D. J. Slotboom [5]* and R. P. -A. Berntsson [5]

[1]*Center for Biomembrane Research, Department of Biochemistry and Biophysics, Stockholm University, SE-106 91 Stockholm, Sweden*
[2]*Institut für Physik, Biophysik, Martin-Luther-Universität Halle-Wittenberg, D-06099 Halle (Saale), Germany*
[3]*Mitteldeutsches Zentrum für Struktur und Dynamik der Proteine, Martin-Luther-Universität Halle-Wittenberg, D-06099 Halle (Saale), Germany*
[4]*Department of Medical Biochemistry and Biophysics, Karolinska Institutet, SE-171 77 Stockholm, Sweden*
[5]*Biochemistry Department, Groningen Biomolecular Sciences and Biotechnology Institute, Netherlands Proteomics Centre and Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands*

Keywords:
peptide binding protein;
oligopeptide transporter;
ABC transporter;
Escherichia coli;
substrate binding protein

The *Escherichia coli* peptide binding protein OppA is an essential component of the oligopeptide transporter Opp. Based on studies on its orthologue from *Salmonella typhimurium*, it has been proposed that OppA binds peptides between two and five amino acids long, with no apparent sequence selectivity. Here, we studied peptide binding to *E. coli* OppA directly and show that the protein has an unexpected preference for basic peptides. OppA was expressed in the periplasm, where it bound to available peptides. The protein was purified in complex with tightly bound peptides. The crystal structure (up to 2.0 Å) of OppA liganded with the peptides indicated that the protein has a preference for peptides containing a lysine. Mass spectrometry analysis of the bound peptides showed that peptides between two and five amino acids long bind to the protein and indeed hinted at a preference for positively charged peptides. The preference of OppA for peptides with basic residues, in particular lysines, was corroborated by binding studies with peptides of defined sequence using isothermal titration calorimetry and intrinsic protein fluorescence titration. The protein bound tripeptides and tetrapeptides containing positively charged residues with high affinity, whereas related peptides without lysines/arginines were bound with low affinity. A structure of OppA in an open conformation in the absence of ligands was also determined to 2.0 Å, revealing that the initial binding site displays a negative surface charge, consistent with the observed preference for positively charged peptides. Taken together, *E. coli* OppA appears to have a preference for basic peptides.

**Corresponding authors.* E-mail addresses: degier@dbb.su.se; d.j.slotboom@rug.nl.

Abbreviations used: SBP, substrate binding protein; ITC, isothermal titration calorimetry; ASU, asymmetric unit; MS/MS, tandem mass spectrometry; CV, column volume; PDB, Protein Data Bank.

Introduction

The natural habitat of many bacteria, such as the Gram-negative bacterium *Escherichia coli*, is rich in peptides. Bacteria can use peptides as a source of amino acids, carbon, nitrogen, and/or energy. A variety of peptide uptake systems can mediate the translocation of peptides across the bacterial cytoplasmic membrane.[1–3] In Gram-negative bacteria, the outer membrane places constraints on the size of the peptides that can be taken up; only peptides up to five or six amino acids long can freely diffuse across the outer membrane.[4]

E. coli has several systems that mediate the translocation of peptides across its cytoplasmic membrane.[5–8] One of these systems is the oligopeptide permease Opp.[8] The Opp system is an ATP-binding cassette transporter consisting of the following: two homologous integral membrane proteins OppB and OppC, which together form the translocation pore; two homologous nucleotide binding domains OppD and OppF, which drive transport by ATP binding and hydrolysis; and the receptor or substrate binding protein (SBP) OppA.

OppA belongs to a large superfamily of SBPs associated with ATP-binding cassette transporters involved in nutrient uptake in prokaryotes. SBPs generally consist of two domains with α/β-folds connected by a hinge.[9] A rotation of the domains around the hinge allows the proteins to adopt closed and open conformations. Substrates bind between the two domains and stabilize the closed state, a process often referred to as a Venus flytrap mechanism.[9] The closed liganded protein associates with the membrane-embedded pore and delivers the cargo for translocation. *E. coli* OppA is classified based on its structure as a type II SBP[10] belonging to cluster C.[11] It is composed of three domains instead of two, as found in most other clusters.[11,12]

OppA from *E. coli* shares an 85.7% sequence identity with its orthologue from *Salmonella typhimurium* (the mature forms of the proteins, excluding signal sequences). Based on combined structural and biochemical studies on OppA from *S. typhimurium*, it has been proposed that OppA binds peptides between two and five amino acids long, regardless of their sequence.[13] In the closed conformation, ligands are bound in a closed cavity deep inside the protein.[14] The *S. typhimurium* OppA binding cavity provides large hydrated pockets for the various side chains of the substrates, explaining how different side chains can bind in the same pocket. Water molecules are important in the binding pocket as they occupy the space not taken up by the substrate, solvate charged groups, and can be involved in hydrogen bonds with the side chains.[14,15] Direct interactions between the protein and the peptide ligand are mainly with the backbone of the peptide and the termini.

Another well-studied homologue is OppA from the Gram-positive bacterium *Lactococcus lactis* (<25% sequence identity with *S. typhimurium* and *E. coli* OppA). This protein can bind much longer peptides ranging from 4 to 35 residues.[16] Until recently, it was thought that OppA from *L. lactis* binds its ligands without any sequence preference. However, using a combination of structural biology and biochemical approaches, we showed that OppA from *L. lactis* has a preference for peptides containing at least one hydrophobic residue.[17] Thus, the substrate preference of this peptide binding protein is more complex than previously thought.

Here, we have further defined the substrate preferences of OppA from *E. coli* using a combination of crystallography, mass spectrometry, and ligand binding assays [by isothermal titration calorimetry (ITC) and intrinsic tryptophan fluorescence titration], and we show that *E. coli* OppA also displays sequence preference.

Results

The structure of OppA complexed with peptides

We expressed a C-terminally His-tagged version of OppA in the periplasm of *E. coli* strain C41 (DE3).[18] The recombinantly expressed protein could be readily isolated using a combination of immobilized metal affinity and size-exclusion chromatography, as shown by SDS-PAGE (Supplementary Fig. 1). Mass spectrometry analysis of the purified His-tagged OppA showed that the mature (i.e., periplasmic) form of the protein was isolated (Supplementary Fig. 2).

As is often the case for SBPs, tightly bound ligands were copurified with OppA. Such copurification has previously been described for OppA from *S. typhimurium* and *L. lactis*, as well as for AppA from *Bacillus subtilis* and for MppA from *E. coli*.[12,17,19,20] Static light-scattering experiments showed that OppA with bound ligands was monomeric in solution and suitable for crystallization (Supplementary Fig. 3).[21] OppA with the bound ligands was crystallized, and its structure was solved to 2.0 Å (Fig. 1, Table 1). The asymmetric unit (ASU) contained eight molecules of OppA, which had almost identical structures. The r.m.s.d. values were <0.4 Å when all C^{α} atoms were superimposed; therefore, only one of them (chain A) will be discussed below. All residues of OppA, including the last histidine that served as the first histidine of the His tag, were observed in electron density. Upon refinement of the structure, extra electron density, which could not be accounted for by the protein, was observed in the ligand binding cavity and could readily be modeled as a tripeptide. In some of the

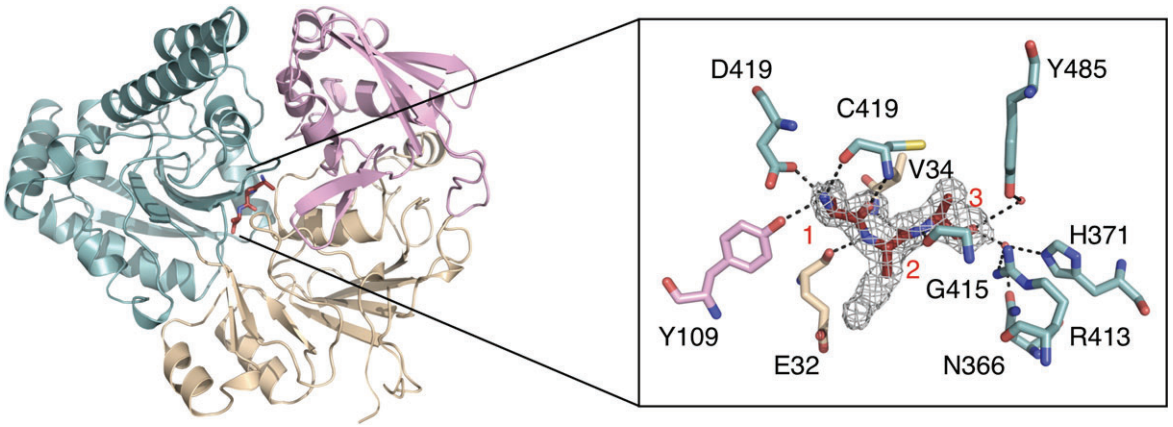


Fig. 1. Crystal structure of OppA (cartoon representation) complexed with an undefined mixture of copurified peptides. Domain I, beige; domain II, pink; domain III, teal; bound ligand, red. The close-up of the bound ligand (right) shows the main interaction partners between the bound ligand and OppA, with hydrogen bonds as broken lines. The electron density of the ligand ($2F_o - F_c$ map) is displayed as a gray mesh contoured at 1σ. Extra density can be seen for the side chain of the second amino acid of the bound peptide.

Table 1. Data collection and refinement statistics

	OppA structure		
	With endogenous ligands bound (closed)	With KGE bound (closed)	Unliganded (open)
Space group	$P2_1$	$P2_1$	$P2_12_12_1$
Cell dimensions			
a, *b*, *c* (Å)	62.5, 201.3, 206.9	62.5, 201.3, 206.9	57.4, 82.3, 125.9
α, β, γ (°)	90.0, 95.6, 90.0	90.0, 95.6, 90.0	90.0, 90.0, 90.0
Wavelength (Å)	0.972	0.972	0.972
Resolution range (Å)	48.3 2.0	48.9 2.0	37.4 2.0
Unique reflections	346,753	341,284	41,615
Completeness (%)	98.9 (94.2)	99.4 (97.5)	98.5 (92.6)
R_{meas}	0.09 (0.65)	0.12 (0.67)	0.08 (0.39)
I/σ (I)	12.4 (2.2)	10.2 (2.6)	14.4 (4.2)
Redundancy	3.7	2.0	4.7
Refinement			
Resolution range	48.3 2.0	48.9 2.0	37.4 2.0
Number of reflections	328,136	324,217	39,534
R_{work}/R_{free}	0.23/0.26	0.20/0.24	0.18/0.22
Number of atoms			
Protein	33,201	33,319	4176
Water	2408	2959	480
Average *B* factors (Å^2)			
Protein	30.0	24.5	30.8
Water	28.4	27.0	38.2
r.m.s.d.			
Bond lengths (Å)	0.013	0.010	0.011
Bond angles (°)	1.3	1.2	1.3

OppA molecules in the ASU, weak additional electron density in the binding pocket indicated that tetrapeptides might be present, albeit at low occupancy. The additional density was not well defined; therefore, a tripeptide was used in all cases in the final model. Neither the first residue nor the last residue of the tripeptide electron density was observed at the side-chain positions beyond the C^{β} atom. The absence of electron density for the side chains is consistent with the proposed lack of specificity of the peptide's sequences, since a mixture of peptide ligands with different sequences in the crystals will result in a low occupancy for each of the side chains. In contrast, the middle residue of the tripeptide had clear side-chain density up to the C^{δ} atom. It was not possible to unambiguously assign the density, but occupancy refinement pointed towards a large fraction (~50%) of the tripeptides with a lysine at the second position. The side-chain density observed at the second position of the peptide is an indication that *E. coli* OppA may have a preference for lysine-containing substrates. Since the sequence of the bound peptide could not be defined unambiguously, the deposited coordinates contain trialanine peptides.

The binding cavity of OppA from *E. coli* is virtually identical with that of OppA from *S. typhimurium* and has a similar volume of ca 1000 Å^3.[17] The interactions of the peptide with OppA from *E. coli* also resemble those seen in OppA from *S. typhimurium* (see, e.g., Tame *et al.*[12]). There are hydrogen bonds between the protein and the backbone CO and NH groups of the peptide. The N-terminus and the C-terminus of the peptide are fixed in place with a salt bridge to Asp445 and Arg439, respectively. Similar to OppA from *S. typhimurium*, the pockets in the protein that hold the side chains of the peptide appear to allow the presence of chemically diverse side chains.

Characterization of peptides complexed with OppA

Peptides that had been copurified with the protein were extracted and separated with a C_{18} column to further explore the substrate preference of OppA. Subsequently, the separated peptides were analyzed by matrix-assisted laser desorption/ionization tandem mass spectrometry (MS/MS). To facilitate the mass-spectrometry-based characterization of the separated peptides, we used iTRAQ (*i*sobaric *t*ag for *r*elative and *a*bsolute *q*uantitation) labeling.[22] The iTRAQ label, which is covalently linked to the N-terminus and side-chain amines of peptides, improves the ionization of peptides, and the extra mass added to the peptides also greatly facilitates the separation of peptides from matrix signals[23] (Fig. 2).

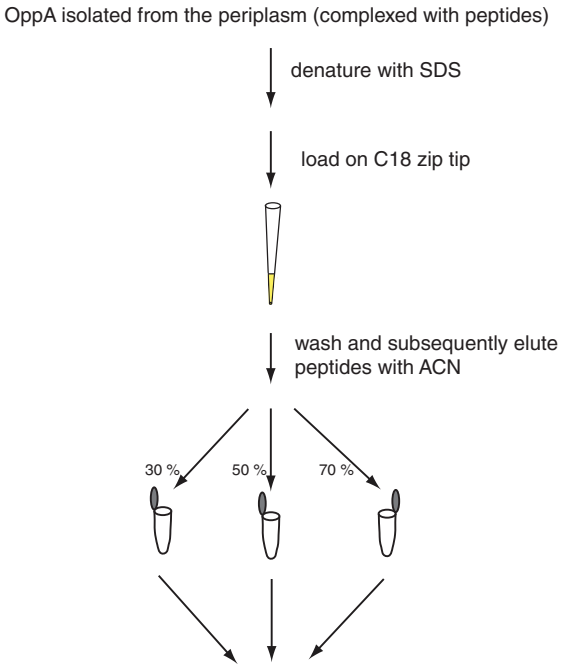
The observed masses obtained from the mass spectrometry analysis were compared to lists of masses of dipeptides, tripeptides, tetrapeptides, pentapeptides, and hexapeptides consisting of 19 natural amino acids (Supplementary Table 1). Since the masses of isoleucine and leucine are the same, the list contains leucine only. It should be noted that we did not determine the actual peptide sequences or their relative abundance; rather, they revealed the length and amino acid composition of the peptides complexed with OppA (Fig. 2). The MS/MS data were used only to distinguish between peptides and matrix adducts/contaminations.

Peptides that range in size from two to five amino acids were found. Notably, almost 50% of the identified peptides copurified with *E. coli* OppA contained at least one positively charged residue. The peptides were rich in lysine (8.5% of amino acids), arginine (8.2%), and histidine (5.0%) residues. This bias does not seem to be a result of the experimental setup used because a similar analysis of the peptides copurified with OppA from *L. lactis* revealed much lower percentages of basic residues (2.8% lysine, 6% arginine, and 1% histidine).[17] In contrast, acidic amino acids were less frequently observed in the peptides copurified with *E. coli* OppA (4.5% aspartate and 3.9% glutamate) compared to those bound to the *L. lactis* protein (7.1% aspartate and 10.6% glutamate). The bias can be considered as an indication that OppA has a preference for positively charged peptides.

Characteristics of the binding of model peptides to OppA

To further characterize the substrate preferences of OppA, we performed quantitative peptide binding studies with peptides of defined sequence using ITC and intrinsic protein fluorescence titration.

Substrate-free OppA, which is required for these binding studies, was prepared by partially denatur-



	numbers of peptides identified	numbers of peptides with basic residues
dipeptide	45	15
tripeptide	108	43
tetrapeptide	140	77
pentapeptide	28	24
total	321	159
total [%]	100	49.4

Fig. 2. Characterization of peptides complexed with OppA. Work flow of the isolation and mass spectrometry analysis of OppA complexed with peptides (for details, see Materials and Methods). This approach gives information about the length of peptides and their amino acid composition (see "output"; Supplementary Table 1).

Table 2. Dissociation constants of peptides to OppA

Peptide	K_d (M) (M) ITC	
	Fluorescence titration	ITC
GK		
LP		ND
KGE	$(1.18 \pm 0.31) \times 10^{-6}$	$(1.36 \pm 0.03) \times 10^{-6}$
PFK	$(0.33 \pm 0.2) \times 10^{-9}$	0.342×10^{-9}
PFG		$(5.65 \pm 0.32) \times 10^{-6}$
PKF	$(1.05 \pm 0.23) \times 10^{-6}$	$(2.17 \pm 0.03) \times 10^{-6}$
AKA	$<9.7 \times 10^{-9}$	ND
AGA		$(3.17 \pm 0.2) \times 10^{-6}$
KGGK	$(0.58 \pm 0.04) \times 10^{-6}$	$(0.43 \pm 0.02) \times 10^{-6}$
KYGK	$(0.32 \pm 0.02) \times 10^{-6}$	ND
YGGFL		ND
RDMPIQAF		ND
SLSQSKVLP		ND

The binding of the listed peptides to OppA was measured using ITC and intrinsic protein fluorescence titration.
( ) No binding observed; ND, not determined.

ing the protein with guanidine hydrochloride on the nickel affinity column, extensively washing it to remove any copurified peptides, and, finally, refolding the protein. The peptides used in the binding studies were of different lengths (i.e., two to nine amino acids long) and had been designed such that the apparent preference of OppA for lysines could be monitored in more detail (Table 2, Fig. 3a and b). Tripeptides and tetrapeptides containing lysines bound well to OppA. Neither the dipeptides nor the peptides longer than four amino acid residues showed any detectable binding to OppA, but it is likely that dipeptides and pentapeptides with sequences different from the ones tested bind to OppA, as indicated by the mass spectrometry analysis. Overall, OppA displays a large range of affinities for its substrates; dissociation constants (K_d) between 330 pM and 5 μM were determined.

Two peptides, AKA and PFK, stood out for their extremely high-affinity binding to OppA (K_d values of <9.7 nM and ~0.33 nM, respectively). The affinity of AKA for OppA was too high to allow for an accurate determination of a dissociation constant, even when competition experiments were performed. A direct comparison between peptides with very similar sequences, PFK *versus* PFG and AKA *versus* AGA, showed a pronounced preference of OppA for peptides with positively charged amino acid residues; the affinities of PFK and AKA are a thousand fold higher than those of PFG (K_d ~5.6 μM) and AGA (K_d ~3.2 μM), respectively.

Structures of OppA complexed with the tripeptide KGE and in the open unliganded conformation

Crystal structures of OppA in the presence of the tripeptide KGE, as well as in an open unliganded conformation, were also determined (both to 2.0 Å; Fig. 4). The structure of OppA complexed with KGE was virtually identical with the structure of OppA with the mixture of peptide ligands (r.m.s.d. of 0.29 Å over all C^{α} atoms), but there was now a well-defined side-chain density for the entire tripeptide. KGE was bound in the same way as the mixture of peptides.

In the open unliganded conformation of OppA, the domains had rotated by ~33° (as measured *via* the angle between Val60, Leu296, and Trp423) with regard to the closed structures. A rotation of 33° is roughly 10° more than what was measured for the opening of OppA from *S. typhimurium* or *L. lactis*[12,17] using equivalent residues. However, other SBPs that have been crystallized in both the open conformation and the closed conformation displayed domain rotations by as much as 70°.[11] The crystallization conditions and crystal contacts likely play major roles in the degree of opening seen in the different structures. It is therefore difficult to compare the amounts of opening between the structures, as it is not known whether the degree of opening is the same when the protein is in solution. In-solution NMR showed that the chemical shifts of unliganded OppA were dramatically different from those of OppA, with the bound tripeptide KGE supporting the extensive domain movements observed by X-ray crystallography (Supplementary Fig. 4).

Surface charge distribution of the peptide-binding pocket of OppA

The surface charge distribution of the peptide-binding pocket was determined in the unliganded open form of OppA (Fig. 5). Domains I and II create a negatively charged surface inside the binding cleft that appears to be an attractive binding site for basic peptides. On the opposite side of the binding cleft, domain III contributes a positively charged patch, which matches the position of the C-terminus of the tripeptide substrates (Fig. 1).

Discussion

Based on studies on its orthologue from *S. typhimurium*, it has been proposed that *E. coli* OppA binds peptides between two and five amino acids long, with no apparent sequence selectivity.[13] We determined the structure of OppA complexed with an undefined mixture of peptides. A tripeptide could be modeled into the electron density found in the binding site; however, because a heterogeneous mixture of peptides was bound, a specific peptide sequence could not be assigned to the electron density. Unexpectedly, the density at the middle position of the tripeptide indicated that a large fraction of the bound peptides had a lysine residue in this position. This observation is inconsistent with the absence of sequence preference by OppA.

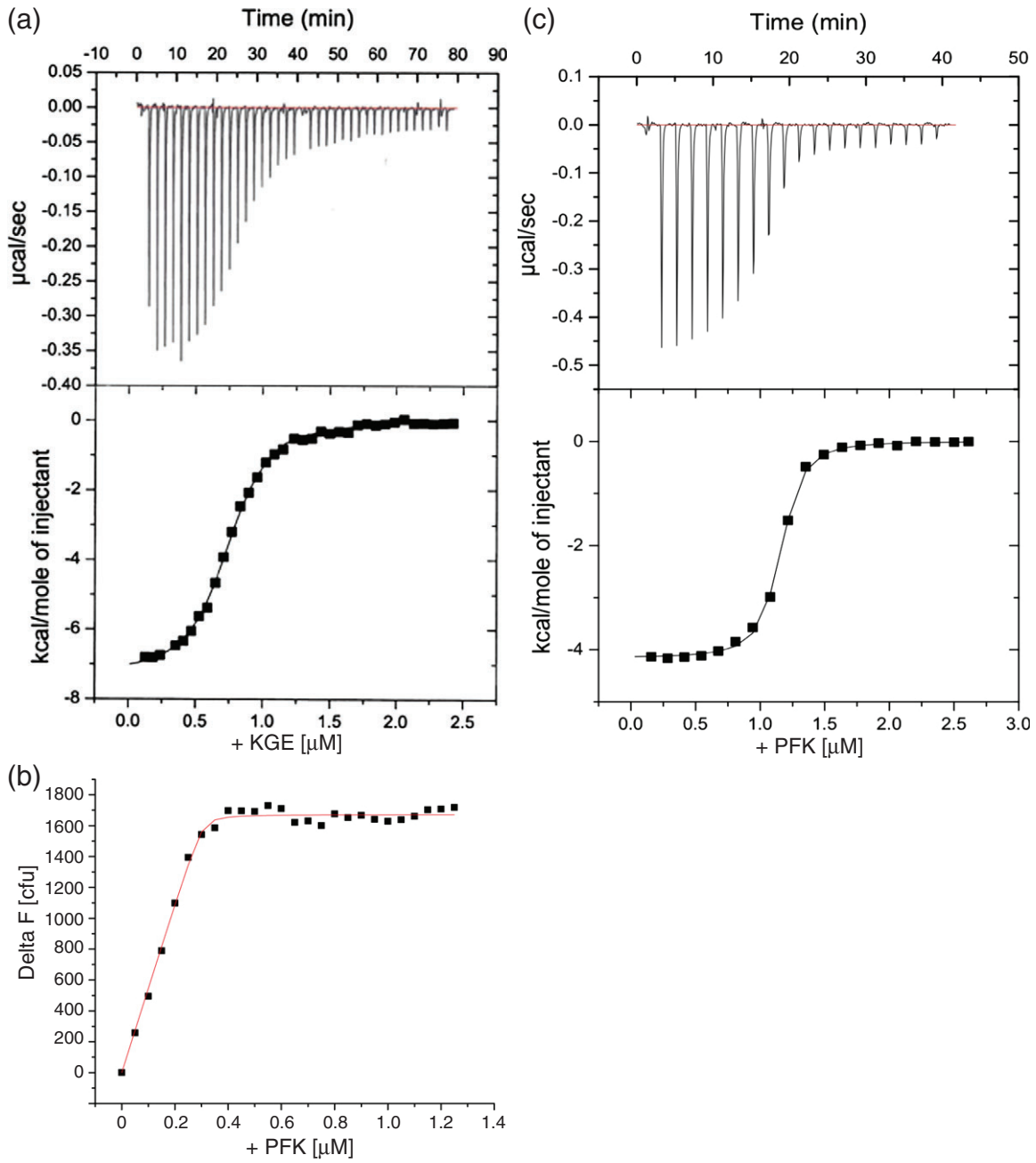


Fig. 3. Examples of ITC and fluorescence titration measurements. (a) Representative ITC measurement of the binding of the tripeptide KGE to OppA. The upper graph shows the heat released upon the binding of KGE to the protein. The number of peptide binding sites (*N*) was ~1 per protein. (b) Fluorescence titration of OppA with the tripeptide PFK. (c) Competition experiment using ITC. OppA was first preincubated with 1 mM KGE and subsequently titrated with PFK.

The modeled tripeptide was anchored within the binding cavity with a salt bridge to the C-terminus and Arg439. No other amino acids in the binding pocket are in a position to form salt bridges to the C-termini of longer peptides. It is, however, likely that Arg439 can adopt a different conformation to form a salt bridge with the C-terminal carboxylate of a bound tetrapeptide. We found no residue within the binding cavity that could easily form a salt bridge to a bound pentapeptide. In the case of OppA from *S. typhimurium*, Lys307 has the possibility to form a salt bridge to a pentapeptide; however, in our *E. coli* structure, the backbone of the protein is slightly different in that region. The corresponding lysine residue, Lys333, points outwards, away from the binding cavity, and cannot form a salt bridge to a

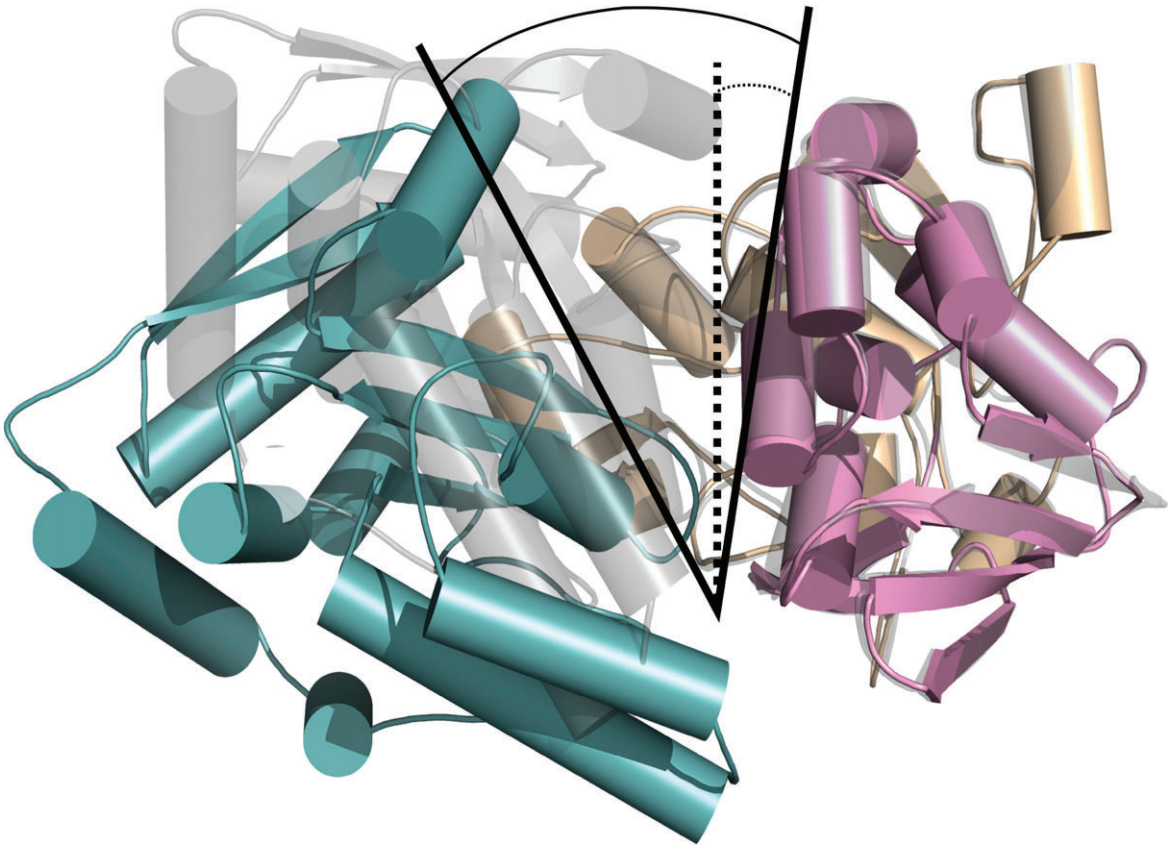


Fig. 4. A comparison between the open unliganded conformation (colored as in Fig. 1) and the closed conformation of OppA (gray). The open conformation displays an opening of ~33°, as visualized by the axis lines. The closed conformation is shown with broken lines, and the open conformation is shown with full lines.

bound pentapeptide. The absence of a potential ionic interaction with pentapeptides could explain OppA's lower affinity for them.

Characterization of the amino acid composition of peptides complexed with OppA, using mass spectrometry, supported the hypothesized preference of OppA for positively charged peptides. Binding experiments with a series of model peptides, with lysine as a positively charged amino acid residue, further corroborated that OppA has a preference for

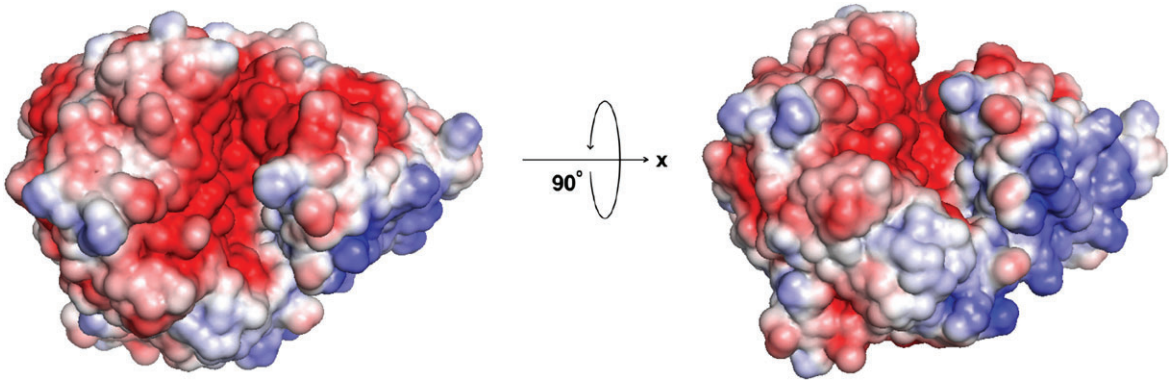


Fig. 5. Surface charge representation of OppA in its open unliganded conformation. Negative charge is visualized in red, and positive charge is visualized in blue. As indicated by the marking, the initial binding site of OppA has a negative charge. Images of the surface charge distribution of open unliganded OppA were created with the PyMOL Molecular Graphics System program and the APBS plug-in in PyMOL.[24]

positively charged peptides. The binding experiments also indicated that OppA has a preference for tripeptides and tetrapeptides over dipeptides and pentapeptides. Consistently, most of the peptides from the mixture that had copurified with OppA isolated from the periplasm were tripeptides and tetrapeptides, indicating that *E. coli* OppA indeed has a preference for positively charged peptides that are three or four amino acids long. Can we understand the preference of OppA for positively charged peptides? Determining the surface charge distribution of the peptide-binding pocket of OppA revealed a negatively charged surface. Domains I and II create a negatively charged surface inside the binding cleft that could attract basic peptides.

Why did the studies on *S. typhimurium* OppA, which are almost identical with the studies on *E. coli* OppA, indicate that OppA handles peptides with no regard for their sequence? A likely explanation is that—by coincidence—in the biochemical and structural characterization of OppA from *S. typhimurium*, almost all peptides that were used contained lysine residues.[14,15,25] In fact, the most extensive study on the lack of sequence specificity made use of tripeptides with the sequence Lys-X-Lys, in which X was varied.[14] The use of these peptides likely obscured the substrate preference of *S. typhimurium* OppA. In fact, there was a clue of a preference for lysine from the first structure that was solved, where the protein was crystallized together with a mixture of undefined peptides [Protein Data Bank (PDB) code: 1OLA]. The tetrapeptide Val-Lys-Pro-Gly was modeled to best describe the electron density.[12] Indeed, the second position of the peptide in that structure also contained a lysine residue, similar to our structure of OppA with bound ligands. Also, the surface charge distribution of the peptide-binding pocket of *S. typhimurium* OppA is almost identical with that of *E. coli*. These observations make it very likely that our findings also apply to OppA from *S. typhimurium*.

Is there a physiological explanation for the peptide preference of *E. coli* OppA, and most likely also of *Salmonella* OppA? Because positively charged peptides are a rich nitrogen source, it is tempting to speculate that the preference relates to the organism's nitrogen needs.

Our conclusions on the substrate preference of *E. coli* OppA show that peptide uptake in *E. coli* is more complex than previously thought. *E. coli*, just like many other bacteria, has a variety of peptide uptake systems. It has been generally accepted that the length of peptides was the major characteristic that determined the substrate preference of the various uptake systems. However, our study, together with the recent study on OppA from *L. lactis*, shows that the substrate preferences of peptide uptake systems are more complex. Future research will focus on the link between the uptake of specific peptides and the physiology of the cell.

Materials and Methods

Expression and purification of His-tagged OppA for structure determination

The gene from *E. coli* BL21(DE3) encoding OppA was fused at the 3′ end to a sequence coding for seven additional histidine residues and amplified by PCR. The OppA N (CGCCCATATGACCAACATCACCAAGAG) and OppA C (GGATCCCTAGTGGTGGTGGTGGTGGTGGTGGTGCTTCACAATGTACATTTCC) primers were used to amplify the *oppA* gene. The PCR product was cloned into the pET11b vector using NdeI and BamHI restriction sites (NdeI and BamHI restriction sites in the primers are underlined). The OppA-His-tag coding region of the overexpression vector was sequenced. OppA-His-tag was overexpressed in *E. coli* strain C43(DE3).[18] Cultures were grown at 37 °C in lysogeny broth (LB; Difco) supplemented with 100 μg/ml ampicillin and induced at an optical density at 600 nm of 0.6 (UV/VIS Spectrometer Lambda 2S; Perkin Elmer), with 0.4 mM IPTG for 4 h. After the cells had been harvested, all steps were carried out at 4 °C or on ice. For the isolation of His-tagged OppA for structure determination, the cell pellet was washed with lysis buffer [50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole (pH 8), 5 μg/ml DNase I, 2 mM PMSF, and 200 μg/ml lysozyme] before lysis by a bench-top cell disrupter (Constant Systems Ltd.) at 25,000 psi. The lysed cells were centrifuged for 20 min at 20,442*g* (Avanti Centrifuge J-20XP; Beckman Coulter). The supernatant was incubated (with shaking) for 45 min with Ni-NTA Superflow (Qiagen) before being loaded onto a gravity column. After OppA-His-tag had been washed with 30 column volumes of wash buffer [50 mM NaH_2PO_4, 300 mM NaCl, and 20 mM imidazole (pH 8)], it was eluted in four steps with elution buffer [50 mM 4-morpholineethanesulfonic acid, 300 mM NaCl, and 500 mM imidazole (pH 6)]. The protein fractions were combined and loaded on a gel-filtration column; the gel-filtration buffer was composed of 20 mM 4-morpholineethanesulfonic acid (pH 6.0) and 150 mM NaCl.

Protein concentrations were determined by measuring absorbance at a wavelength of 280 nm using a nanodrop spectrophotometer (ND-1000; Isogen Life Science).

SDS-PAGE and immunoblotting

Expression of His-tagged OppA and purity of His-tagged OppA were assessed by SDS-PAGE and immunoblotting. Whole cells (0.05 0.1 A_{600} units) or protein samples (0.5 2 μg) were solubilized and separated by standard SDS-PAGE on 12% polyacrylamide/SDS gels. Proteins were either stained with Coomassie brilliant blue R-250 (Serva) or transferred from the gel to a polyvinylidene fluoride membrane (Millipore). After protein transfer, membranes were blocked and decorated with an antiserum against His tags (Thermo Scientific). Proteins were visualized using the ECL system (GE Healthcare).

Crystallization and structure determination

OppA crystals were grown by hanging-drop vapor diffusion. The drops consisted of 1 μl of protein (14 mg/ml OppA) + 1 μl of reservoir solution (0.02 mM potassium

dihydrogen phosphate, 17% polyethylene glycol 8000, and 23% glycerol) and 1 ml of reservoir solution. For cocrystallization with defined peptides and crystallization of the open unliganded conformation of OppA (without bound peptides), OppA was partially denatured using guanidinium hydrochloride and refolded by washes with buffer with decreasing guanidinium hydrochloride concentrations [4 column volume (CV) 2 M, 4 CV 1.5 M, 4 CV 1 M, 4 CV 0.5 M, and 10 CV buffer alone] before being eluted. The peptide solution (KGE; 10 mM stock in MilliQ water) was mixed 1:10 with protein solution (final protein concentration, 14 mg/ml OppA). Crystals were grown at 18 °C.

Diffraction-quality crystals appeared after 2 3 weeks. Data collection was performed at Deutsches Elektronen-Synchrotron (Hamburg) and at beamline ID23-1 of the European Synchrotron Radiation Facility (Table 1) at 70 K. Crystals of OppA in the open conformation, with the ligand and the defined ligand KGE, all diffracted to 2.0 Å. Data processing and reduction were carried out using XDS and programs from the CCP4 package.[26,27] The structure of OppA with ligand was solved by molecular replacement with Phaser,[28] using the main chain of the previously determined structure of *S. typhimurium* OppA (PDB code: 1OLA) as search model (no ligand was present in the search model). The two other structures were solved using the refined model of *E. coli* OppA as search model. The output from Phaser was further refined with Refmac5 and PHENIX.refine,[29,30] using noncrystallographic symmetry with loose restraints, interspersed with manual building rounds in Coot.[31] Open unliganded OppA contained one molecule per ASU, whereas the two liganded structures contained eight molecules per ASU.

Characterization of peptides complexed with OppA

His-tagged OppA was expressed as described above. For a coverage of the full sequence range of possible substrates, cells were grown not only in lysogeny broth (LB) with tryptone (Bacto; casein proteins CASK, Q52GT1, Q52ET1, CASB, CASA1, and CASA2; LACB from *Bos taurus*) but also in LB with soy peptone (A3 SC; Organotechnie SAS) or casein peptone (N1; Organotechnie SAS) as peptide source†.

OppA complexed with peptides was isolated from the periplasmic fraction. After the cells had been spun down, the cell pellet was resuspended in osmotic shock buffer [500 mM sucrose, 100 mM Tris, and 1 mM ethylenediaminetetraacetic acid (pH 8)]. Every 5 min, the suspension was vigorously mixed for 10 s for a total period of 30 min, followed by centrifugation at 20,442*g* for 30 min (Avanti Centrifuge J-20XP; Beckman Coulter). Peptides complexed with His-tagged OppA were purified and identified as described by Berntsson *et al.*[17] Briefly, the protein was denatured by SDS and loaded into a C18 zip tip. Increasing concentrations of acetonitrile eluted the first peptides and, subsequently, the protein. Peptides were labeled with iTRAQ (4plex and 8plex; Applied Biosystems) to increase their mass. Further separation was achieved by nano-LC, and peptides were subsequently spotted on a target plate together with a matrix solution of 2 mg/ml α-cyano-4-hydroxycinnamic acid. Mass spectrometry data acquisition was performed in positive ion mode. During acquisition, peptides with a signal-to-noise level of Œ50 were selected for MS/MS analysis. For identification of peptides, the precursor ions were matched to a list of masses of single amino acids, dipeptides, tripeptides, tetrapeptides, and pentapeptides with and without iTRAQ label. The corresponding MS/MS spectra were checked for the presence of the iTRAQ label to validate the hits and to distinguish real peptides from contaminations.

Fluorescence titration and ITC

Fluorescence titration experiments were conducted using a Fluorolog fluorimeter (Horiba Scientific). In the fluorescence spectra (excitation, 295 nm), unliganded OppA showed a maximum at 345 nm, the intensity of which was quenched by peptide binding. Peptides were titrated to 800 μl with 0.5 μM protein in steps of 0.5 μl. Binding constants were calculated as described previously.[32]

ITC experiments were performed using an ITC200 calorimeter (MicroCal). Peptides (250 400 μM) were titrated into the thermally equilibrated ITC cell filled with 200 μl of OppA (20 50 μM). Temperature was kept constant at 25 °C. Data were analyzed using the ORIGIN-based software provided by MicroCal.

Constants for the high-affinity binding peptide were determined by displacement titration in accordance with Sigurskjold.[33]

Surface charge calculations

Images were created with the PyMOL Molecular Graphics System program and the APBS plug-in in PyMOL.[24] Negative charges are shown in red, and positive charges are shown in blue.

Accession numbers

Coordinates and structure factors have been deposited in the PDB as follows: OppA complexed with endogenous peptides (PDB code: 3TCF), OppA complexed with KGE (PDB code: 3TCG), and OppA in the open unliganded conformation (PDB code: 3TCH).

Acknowledgements

This research was supported by grants from the Swedish Research Council and the National Institutes of Health (5R01GM081827-03). M. M. Klepsch was supported by a European Molecular Biology Organization short-term fellowship and by travel stipends from Boehringer Ingelheim and Klas-Bertil och Margareta Augustinssons Stiftelse. D. J. Slotboom was supported by a VIDI grant from the Netherlands Organisation for Scientific Research (NWO).

† www.organotechnie.com

References

1. Abouhamad, W. N., Manson, M., Gibson, M. M. & Higgins, C. F. (1991). Peptide transport and chemotaxis in *Escherichia coli* and *Salmonella typhimurium*: characterization of the dipeptide permease (Dpp) and the dipeptide-binding protein. *Mol. Microbiol.* **5**, 1035 1047.
2. Hiles, I. D., Gallagher, M. P., Jamieson, D. J. & Higgins, C. F. (1987). Molecular characterization of the oligopeptide permease of *Salmonella typhimurium*. *J. Mol. Biol.* **195**, 125 142.
3. Higgins, C. F. (2001). ABC transporters: physiology, structure and mechanism an overview. *Res. Microbiol.* **152**, 205 210.
4. Nikaido, H. (2003). Molecular basis of bacterial outer membrane permeability revisited. *Microbiol. Mol. Biol. Rev.* **67**, 593 656.
5. Saurin, W., Hofnung, M. & Dassa, E. (1999). Getting in or out: early segregation between importers and exporters in the evolution of ATP-binding cassette (ABC) transporters. *J. Mol. Evol.* **48**, 22 41.
6. Paulsen, I. T. & Skurray, R. A. (1994). The POT family of transport proteins. *Trends Biochem. Sci.* **19**, 404.
7. Harder, D., Stolz, J., Casagrande, F., Obrdlik, P., Weitz, D., Fotiadis, D. & Daniel, H. (2008). DtpB (YhiP) and DtpA (TppB, YdgR) are prototypical proton-dependent peptide transporters of *Escherichia coli*. *FEBS J.* **275**, 3290 3298.
8. Pearce, S. R., Mimmack, M. L., Gallagher, M. P., Gileadi, U., Hyde, S. C. & Higgins, C. F. (1992). Membrane topology of the integral membrane components, OppB and OppC, of the oligopeptide permease of *Salmonella typhimurium*. *Mol. Microbiol.* **6**, 47 57.
9. Mao, B., Pear, M. R., McCammon, J. A. & Quiocho, F. A. (1982). Hinge-bending in L-arabinose-binding protein. The "Venus's-flytrap" model. *J. Biol. Chem.* **257**, 1131 1133.
10. Fukami-Kobayashi, K., Tateno, Y. & Nishikawa, K. (1999). Domain dislocation: a change of core structure in periplasmic binding proteins in their evolutionary history. *J. Mol. Biol.* **286**, 279 290.
11. Berntsson, R. P., Smits, S. H., Schmitt, L., Slotboom, D. J. & Poolman, B. (2010). A structural classification of substrate-binding proteins. *FEBS Lett.* **584**, 2606 2617.
12. Tame, J. R., Murshudov, G. N., Dodson, E. J., Neil, T. K., Dodson, G. G., Higgins, C. F. & Wilkinson, A. J. (1994). The structural basis of sequence-independent peptide binding by OppA protein. *Science*, **264**, 1578 1581.
13. Guyer, C. A., Morgan, D. G. & Staros, J. V. (1986). Binding specificity of the periplasmic oligopeptide-binding protein from *Escherichia coli*. *J. Bacteriol.* **168**, 775 779.
14. Sleigh, S. H., Seavers, P. R., Wilkinson, A. J., Ladbury, J. E. & Tame, J. R. (1999). Crystallographic and calorimetric analysis of peptide binding to OppA protein. *J. Mol. Biol.* **291**, 393 415.
15. Tame, J. R., Sleigh, S. H., Wilkinson, A. J. & Ladbury, J. E. (1996). The role of water in sequence-independent ligand binding by an oligopeptide transporter protein. *Nat. Struct. Biol.* **3**, 998 1001.
16. Doeven, M. K., Abele, R., Tampe, R. & Poolman, B. (2004). The binding specificity of OppA determines the selectivity of the oligopeptide ATP-binding cassette transporter. *J. Biol. Chem.* **279**, 32301 32307.
17. Berntsson, R. P., Doeven, M. K., Fusetti, F., Duurkens, R. H., Sengupta, D., Marrink, S. J. *et al.* (2009). The structural basis for peptide selection by the transport receptor OppA. *EMBO J.* **28**, 1332 1340.
18. Miroux, B. & Walker, J. E. (1996). Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J. Mol. Biol.* **260**, 289 298.
19. Levdikov, V. M., Blagova, E. V., Brannigan, J. A., Wright, L., Vagin, A. A. & Wilkinson, A. J. (2005). The structure of the oligopeptide-binding protein, AppA, from *Bacillus subtilis* in complex with a nonapeptide. *J. Mol. Biol.* **345**, 879 892.
20. Maqbool, A., Levdikov, V. M., Blagova, E., Herve, M., Horler, R. S., Wilkinson, A. J. & Thomas, G. H. (2011). Compensating stereochemical changes allow murein tripeptide to be accommodated in a conventional peptide binding protein. *J. Biol. Chem.* **286**, 31512 31521.
21. Slotboom, D. J., Duurkens, R. H., Olieman, K. & Erkens, G. B. (2008). Static light scattering to characterize membrane proteins in detergent solution. *Methods*, **46**, 73 82.
22. Unwin, R. D., Pierce, A., Watson, R. B., Sternberg, D. W. & Whetton, A. D. (2005). Quantitative proteomic analysis using isobaric protein tags enables rapid comparison of changes in transcript and protein levels in transformed cells. *Mol. Cell. Proteomics*, **4**, 924 935.
23. Wiese, S., Reidegeld, K. A., Meyer, H. E. & Warscheid, B. (2007). Protein labeling by iTRAQ: a new tool for quantitative mass spectrometry in proteome research. *Proteomics*, **7**, 340 350.
24. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl Acad. Sci. USA*, **98**, 10037 10041.
25. Sleigh, S. H., Tame, J. R., Dodson, E. J. & Wilkinson, A. J. (1997). Peptide binding in OppA, the crystal structures of the periplasmic oligopeptide binding protein in the unliganded form and in complex with lysyl lysine. *Biochemistry*, **36**, 9747 9758.
26. Kabsch, W. (2010). XDS. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66**, 125 132.
27. Collaborative Computational Project, Number 4. (1994). The CCP4 Suite: programs for protein crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **50**, 760 763.
28. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). Likelihood-enhanced fast translation functions. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **61**, 458 464.
29. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **53**, 240 255.

30. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W. *et al.* (2002). PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **58**, 1948 1954.
31. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2126 2132.
32. Lanfermeijer, F. C., Picon, A., Konings, W. N. & Poolman, B. (1999). Kinetics and consequences of binding of nona- and dodecapeptides to the oligopeptide binding protein (OppA) of *Lactococcus lactis*. *Biochemistry*, **38**, 14440 14450.
33. Sigurskjold, B. W. (2000). Exact analysis of competition ligand binding by displacement isothermal titration calorimetry. *Anal. Biochem.* **277**, 260 266.