BIOLOGICAL CRYSTALLOGRAPHY
ISSN 1399-0047

Structural basis for amino-acid recognition and transmembrane signalling by tandem Per–Arnt–Sim (tandem PAS) chemoreceptor sensory domains

Yu C. Liu,[a]‡ Mayra A. Machuca,[a]‡ Simone A. Beckham,[b] Menachem J. Gunzburg[b] and Anna Roujeinikova[a,b]*

[a]Department of Microbiology, Monash University, Clayton, Victoria 3800, Australia, and [b]Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3800, Australia. *Correspondence e-mail: anna.roujeinikova@monash.edu

Received 3 June 2015
Accepted 21 July 2015

Edited by Z. Dauter, Argonne National Laboratory, USA

‡ YCL and MPM contributed equally to this work.

Keywords: chemotaxis; methyl-accepting protein; sensing domain.

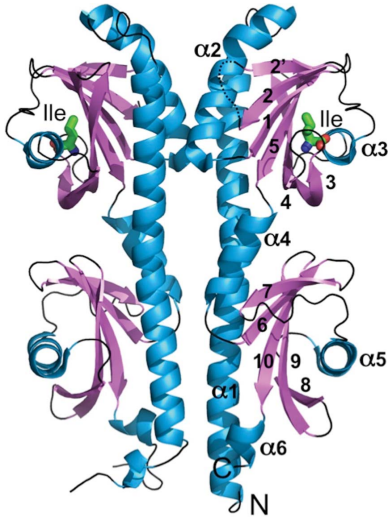
PDB references: Tlp3 PTPSD, 4xmq; complex with isoleucine, 4xmr

Supporting information: this article has supporting information at journals.iucr.org/d

Chemotaxis, mediated by methyl-accepting chemotaxis protein (MCP) receptors, plays an important role in the ecology of bacterial populations. This paper presents the first crystallographic analysis of the structure and ligand-induced conformational changes of the periplasmic tandem Per-Arnt-Sim (PAS) sensing domain (PTPSD) of a characterized MCP chemoreceptor. Analysis of the complex of the *Campylobacter jejuni* Tlp3 PTPSD with isoleucine (a chemoattractant) revealed that the PTPSD is a dimer in the crystal. The two ligand-binding sites are located in the membrane-distal PAS domains on the faces opposite to the dimer interface. Mutagenesis experiments show that the five strongly conserved residues that stabilize the main-chain moiety of isoleucine are essential for binding, suggesting that the mechanism by which this family of chemoreceptors recognizes amino acids is highly conserved. Although the fold and mode of ligand binding of the PTPSD are different from the aspartic acid receptor Tar, the structural analysis suggests that the PTPSDs of amino-acid chemoreceptors are also likely to signal by a piston displacement mechanism. The PTPSD fluctuates between piston (C-terminal helix) 'up' and piston 'down' states. Binding of an attractant to the distal PAS domain locks it in the closed form, weakening its association with the proximal domain and resulting in the transition of the latter into an open form, concomitant with a downward (towards the membrane) 4 Å piston displacement of the C-terminal helix. *In vivo*, this movement would generate a transmembrane signal by driving a downward displacement of the transmembrane helix 2 towards the cytoplasm.

1. Introduction

Most bacteria are motile. Chemotaxis, the chemically guided movement towards an attractant or away from a repellent, plays an important role in the ecology of bacterial populations. It underpins the ability of bacteria to colonize microenvironmental niches that serve as a supply of nutrients, a process central to symbiosis and pathogenesis. Chemotaxis is essential for the host colonization and virulence of many pathogenic bacteria associated with human, animal and plant diseases (Josenhans & Suerbaum, 2002; Rosenberg *et al.*, 2007). For example, chemotaxis towards chemicals released by corals and their symbionts at the endangered Great Barrier Reef is pivotal to the infection of corals by pathogenic bacteria (Rosenberg *et al.*, 2007). Renewable production of nitrogen for agriculture *via* symbiotic association between *Rhizobium* bacteria and legumes is dependent on chemotaxis towards legume roots (Fox *et al.*, 2007). Furthermore, bacterial chemotaxis plays a pivotal role in ocean-scale or global-scale biogeochemical fluxes, including carbon, nitrogen and sulfur cycling (Stocker & Seymour, 2012).



Chemical signals control the movement of bacteria *via* interaction with their membrane-embedded methyl-accepting chemotaxis protein (MCP) receptors. External chemical stimuli are detected by the periplasmic sensing domains (SDs) of the MCP receptors. Some ligands are sensed directly *via* binding to these domains and some indirectly (*e.g. via* periplasmic binding proteins), but it is the direct sensing that is of particular interest, since the specificity of the chemoreceptor ligand-binding site can, in these instances, be precisely investigated, exploited and redesigned to detect different chemical cues.

The structural basis behind the direct recognition of attractants has been systematically studied for only one family of SDs possessing a four-helix bundle fold, as exemplified by the *Salmonella typhimurium* aspartic acid receptor Tar. Previous studies have shown that binding of aspartate to the Tar SD results in a downward (towards the membrane) 1.5 Å piston displacement of the periplasmic helix. This in turn drives a piston-type sliding of the transmembrane helix 2, which links the SD to the signalling domain, towards the cytoplasm, thus transmitting the message to the cytoplasmic moiety of the receptor (Chervitz & Falke, 1996). It has since become clear that bacterial chemoreceptor SDs are extremely diverse in sequence and structure (Krell *et al.*, 2011). Two periplasmic SDs with distinctly different folds have recently been characterized: the bimodular, two four-helical bundle SD of *Pseudomonas putida* McpS (Pineda-Molina *et al.*, 2012) and the Per–Arnt–Sim (PAS)-like SD of *Helicobacter pylori* TlpB (Sweeney *et al.*, 2012). In addition, previous sequence analysis of MCPs containing the Cache (**ca**lcium channels and **che**motaxis receptors) motif (Anantharaman & Aravind, 2000) identified a new structural family of sensing modules, that of periplasmic tandem PAS SDs (PTPSDs). In the full-length receptor *in vivo*, both termini of the PTPSD are attached to transmembrane helices, one of which links the PTPSD to the cytoplasmic methyl-accepting (MA) signalling domain *via* the HAMP region (present in **h**istidine kinases, **a**denylyl cyclases, **M**CPs and **p**hosphatases). Mutagenesis studies on representative members of this family including *P. aeruginosa* PctA (Rico-Jiménez *et al.*, 2013), *Vibrio cholerae* McpX/Mlp24 (Nishiyama *et al.*, 2012), *Bacillus subtilis* McpB (Glekas *et al.*, 2010) and *Sinorhizobium meliloti* McpU (Webb *et al.*, 2014) mapped amino-acid residues in the distal PAS domain that are important for ligand recognition. However, the mechanism of signal transmission across the membrane used by SDs that differ from the well studied four-helix bundle remains to be established.

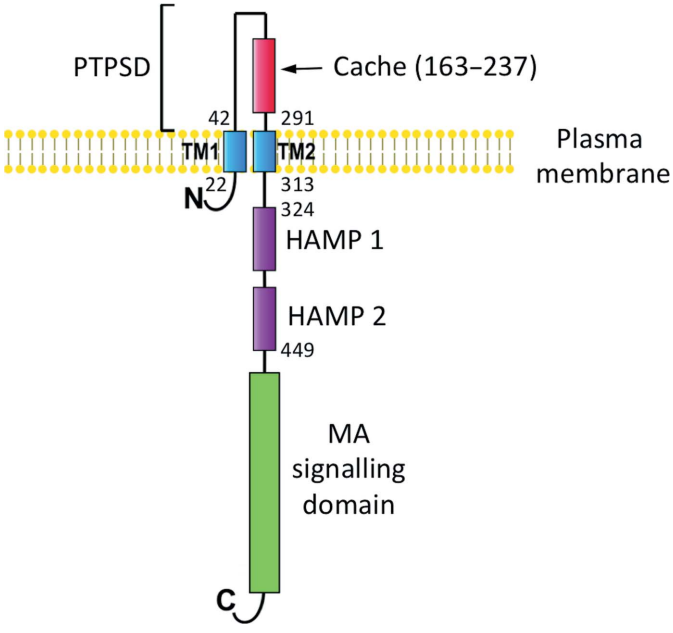


Figure 1
Overall topology of *C. jejuni* Tlp3. The PTPSD comprises residues 42–291.

Table 1
X-ray data-collection and phasing statistics.

Values in parentheses are for the outer shell.

	Gold derivative	Free form	Complex with isoleucine
Wavelength (Å)	1.04	0.95	0.95
Space group	$P2_1$	$P2_1$	$P2_1$
a, b, c (Å)	43.0, 138.2, 49.2	42.3, 137.5, 49.1	42.6, 138.0, 49.0
β (°)	94.5	94.5	94.3
Resolution range (Å)	69–1.35 (1.37–1.35)	68–1.50 (1.53–1.50)	40–1.30 (1.32–1.30)
Total No. of reflections	398772	228292	441329
No. of unique reflections	118130	75588	128260
Completeness (%)	95 (63)	85 (79)	93 (59)
Multiplicity	3.4 (2.1)	3.0 (3.1)	3.4 (2.6)
$\langle I/\sigma(I)\rangle$	9.3 (2.8)	10.5 (2.3)	12.2 (2.1)
$R_{\rm merge}$†	0.076 (0.390)	0.045 (0.281)	0.045 (0.295)

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the ith observation of reflection hkl.

As a step towards elucidating the detailed molecular mechanism by which ligand binding to PTPSDs is signalled across the membrane, the PTPSD of the transducer-like protein 3 (Tlp3; Fig. 1) chemoreceptor from the human pathogen *Campylobacter jejuni* (Rahman *et al.*, 2014) was expressed in *Escherichia coli*, refolded from inclusion bodies, purified and crystallized as described by Machuca *et al.* (2015). This article reports the determination of the high-resolution crystal structures of the Tlp3 PTPSD and its complex with isoleucine (an attractant) and presents a model for ligand-induced conformational change within the periplasmic and membrane-spanning domains.

2. Materials and methods

2.1. Crystallization and data collection

The crystal of the complex of Tlp3 PTPSD from *C. jejuni* serotype O:2 (strain NCTC 11168) with isoleucine was obtained as described previously (Machuca *et al.*, 2015). Crystals of free Tlp3 PTPSD were produced under similar conditions. X-ray diffraction data for the native cryocooled crystal and for the isoleucine complex were collected at 100 K on the MX1 and MX2 beamlines of the Australian Synchrotron (AS; McPhillips *et al.*, 2002) to 1.5 and 1.3 Å resolution, respectively. A single-wavelength anomalous dispersion

Table 2
Refinement statistics.

	Free form	Complex with isoleucine
Resolution range (Å)	68–1.5	20–1.3
Final R_{cryst}†	0.141	0.132
Final R_{free}‡	0.189	0.165
R.m.s. deviations		
Bonds (Å)	0.015	0.019
Angles (°)	1.4	1.7
No. of atoms		
Protein	4200	4189
Water	852	800
Average *B* factors (Å^2)		
Protein	25	20
Water	36	34
Ligand	—	19
Ramachandran plot		
Most favoured (%)	99	97
Allowed (%)	1	3
MolProbity§ clash score	4.5	5.9
PDB code	4xmq	4xmr

† $R_{cryst} = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$. ‡ The free *R* factor was calculated using 5% of the data omitted at random. § Chen *et al.* (2010).

(SAD) experiment was performed on a crystal of a potassium tetrabromoaurate derivative on the AS MX2 beamline. All data were processed and scaled using *iMosflm* (Battye *et al.*, 2011) and *AIMLESS* (Evans & Murshudov, 2013) (see Table 1).

2.2. Structure determination

The locations of the eight Au sites for the derivative were found using *AutoSol* in *PHENIX* (Adams *et al.*, 2010). An initial partial model was generated using *AutoBuild* in *PHENIX* and was then manually completed using *Coot* (Emsley & Cowtan, 2004) and refined against the 1.3 Å resolution isoleucine data set using *PHENIX* and later *REFMAC* (Murshudov *et al.*, 2011). The structure of free Tlp3 PTPSD was solved by molecular replacement using *Phaser* (McCoy *et al.*, 2007) using the isoleucine complex as a search model. Refinement statistics and stereochemistry are given in Table 2.

2.3. Size-exclusion chromatography coupled with multi-angle light scattering (SEC-MALS)

To determine the hydrated molecular mass of Tlp3 PTPSD in solution, a 100 µl protein sample at a concentration of 100 µ*M* was loaded onto a WTC-030S5 SEC column (Wyatt Technology Corporation) pre-equilibrated with buffer 1 (100 m*M* Tris pH 8.0, 150 m*M* sodium chloride) or buffer 2 (buffer 1 plus 10 m*M* isoleucine), and the eluant was passed through an inline DAWN HELEOS light-scattering detector, an Optilab T-rEX differential refractive-index detector and a quasi-elastic light-scattering detector (WyattQELS, Wyatt Technology Corporation). For calculations of the molecular weight and R_h, the light-scattered intensity and the refractive index were analysed using *ASTRA* 6.0 (Wyatt Technology Corporation) (Table 3). Theoretical calculations of R_h from the crystal structure were carried out using *HYDROPRO* (Ortega *et al.*, 2011).

Table 3
Dynamic light-scattering results.

Sample	Polydispersity	Molecular weight (kDa)	R_h (nm)
Tlp3 PTPSD	1.0	27.0	2.5
Tlp3 PTPSD + isoleucine	1.0	27.4	2.5
BSA	1.0	63.8	3.7

2.4. Site-directed mutagenesis

Single-point alanine substitutions were introduced at positions Tyr118, Val126, Lys149, Trp151, Tyr167, Asp169, Thr170, Val171, Asp196, His237 or Arg262 *via* the oligonucleotide-directed mutagenesis technique (QuikChange, Stratagene).

2.5. CD analysis

Tlp3 PTPSD and its single-point variants were dialysed exhaustively against 10 m*M* sodium phosphate pH 7.4, 150 m*M* NaCl. Far-UV CD spectra were recorded at a protein concentration of 0.06 mg ml^{-1} at 25°C using a Jasco J-815 spectropolarimeter over the wavelength range 200–260 nm with a scan rate of 20 nm min^{-1}.

2.6. Isothermal titration calorimetry (ITC)

Tlp3 PTPSD and its variants were dialysed against buffer 1. Solutions of isoleucine (3 and 15 m*M*) were prepared in the dialysis buffer. Measurements were performed at 25°C using a VP-ITC MicroCal calorimeter (Malvern Instruments, UK). The protein sample in a 1.45 ml reaction cell was injected with 25 successive 1 µl aliquots of ligand solution at a spacing of 300 s. Binding isotherms were generated by plotting the heat change evolved per injection *versus* the molar ratio of isoleucine to Tlp3 PTPSD variant. The data were fitted to a single-site binding model using nonlinear least-squares regression, fixing the stoichiometry (*N*) as 1 and allowing all other fitting parameters to float (*Origin* 7, OriginLab, USA).

2.7. Bioinformatic analysis

The overall topology of Tlp3 PTPSD was obtained by searching against the SMART protein-domain database (Letunic *et al.*, 2015). To identify homologous MCP SDs, the TrEMBL database was searched for the consensus motif D*XXX*(R/K)WY*XX*A using the Quick Matrix Method in *SCANSITE* (Obenauer *et al.*, 2003) with keyword 'chemotaxis' and a molecular-weight range of 60–85 kDa. The hits were used to generate a phylogenetic tree using *PhyloT* (http://phylot.biobyte.de/) based on the NCBI taxonomy, which was visualized with the *Interactive Tree of Life* (*iTOL*; Letunic & Bork, 2011). Computational predictions of the fold of SDs were based on the detection of remote homology to proteins of known structure (*HHpred*; Söding *et al.*, 2005).

2.8. Normal-mode analysis (NMA) and morphs

NMA was performed using *elNémo* (Suhre & Sanejouand, 2004). Morphs were produced using the *UCSF Chimera* package developed by the Resource for Biocomputing,

Visualization and Informatics at the University of California, San Francisco, USA (Pettersen *et al.*, 2004).

2.9. PDB references

The atomic coordinates and structure factors for free and isoleucine-bound *C. jejuni* Tlp3 PTPSD have been deposited in the Protein Data Bank (http://www.rcsb.org; PDB entries 4xmq and 4xmr).

3. Results and discussion

3.1. Overall structure and comparison to other periplasmic SDs

The structure of *C. jejuni* Tlp3 PTPSD (residues 42–291 plus an additional N-terminal GIDPFT sequence introduced by the cloning procedure) co-crystallized with isoleucine was determined to 1.3 Å resolution using the SAD method with a gold derivative. The asymmetric unit contains a dimer with the twofold axis perpendicular to the putative membrane plane (Fig. 2*a*). The Tlp3 PTPSD subunit comprises membrane-distal and membrane-proximal PAS domains and a long stalk helix, the N-terminal and the C-terminal halves of which form part of the proximal and distal domains, respectively. Fig. 2(*b*) shows the secondary-structure topology of Tlp3 PTPSD. Although the two PAS domains do not display any significant sequence similarity, they have a very similar fold (r.m.s. deviation of 2.4 Å for the superimposition of 92 C^{α} atoms showing 10% sequence identity over equivalent positions).

The core of the distal domain (residues 63–197) comprises a central six-stranded antiparallel β-sheet with strand order 2′ 2 1 5 4 3. This β-sheet is flanked on one side by an antiparallel two-helix bundle formed by helix α2 and the C-terminal half of helix α1, and on the other side by helix α3 (secondary-structure elements are numbered to highlight deviations from the canonical PAS fold, which lacks the peripheral strand β2′). The proximal domain (residues 42–62 and 199–291) contains a central five-stranded antiparallel β-sheet with strand order 7 6 10 9 8. This β-sheet is flanked by an antiparallel two-helix bundle formed by helix α4 and the N-terminal half of helix α1 on one side and by helix α5 on the other side. In the full-length receptor, the N-terminus of helix α1 is connected to transmembrane helix 1, and the C-terminus of helix α6 is connected to transmembrane helix 2. The proximal and distal domains are intimately associated with one another, with about 17% (1170 Å^2) of the accessible surface area (ASA) of each domain buried at the interface.

This is the first crystal structure of the PTPSD of a characterized MCP receptor. It is distinctly different from the structures of the SDs of the MCPs McpS, Tar and TlpB (Fig. 3*a*). However, we note that the dimerization mode of Tlp3 PTPSD in the crystal is similar to that observed for the single PAS SD of TlpB, with the α1 and α2 helices forming a four-helix bundle at the dimerization interface. This dimerization is likely to be weak, as PTPSD is monomeric in solution according to SEC-MALS analysis (Fig. 3*b*; Table 3). Tlp3 PTPSD eluted as a single monodisperse peak both in the absence and the presence of isoleucine. The derived molecular weight was ∼27 kDa, which is close to that calculated from the amino-acid sequence of a monomer (28.7 kDa). The apparent hydrodynamic radius R_h of the particles in this peak was

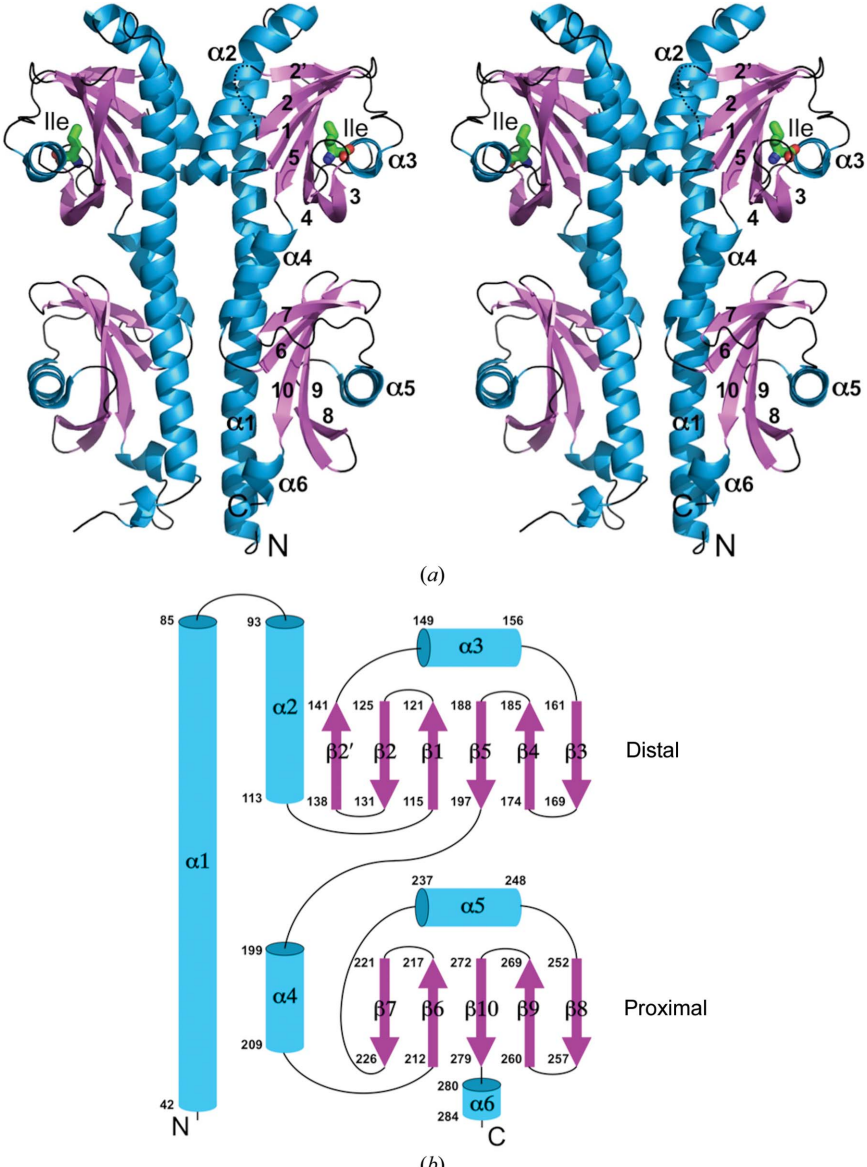


Figure 2
(*a*) Wall-eyed stereo representation of the structure of the *C. jejuni* Tlp3 PTPSD dimer. (*b*) The topology of the secondary-structure elements of Tlp3 PTPSD. The α-helices are represented by rods and β-strands by arrows. The distal and proximal PAS domains are labelled.

25 Å, which is the same as the R_h value calculated from the crystal structure of the single Tlp3 PTPSD subunit. The monomeric state of Tlp3 PTPSD in solution is in agreement with the very low relative value (6.6% or 900 Å^2) of the subunit ASA buried at the dimer interface in the crystal.

In a comparison of the atomic coordinates of Tlp3 PTPSD against the structures in the Protein Data Bank that have been characterized in the literature, using *PDBeFold* (Krissinel & Henrick, 2004), significant similarities were found with the SDs of bacterial family 1 histidine kinases (HK1$_s$s) HK1$_s$-Z2 and HK1$_s$-Z3 from *Methanosarcina mazei* (Fig. 3*c*) and HK1$_s$-Z8 from *V. parahaemolyticus* (Zhang & Hendrickson, 2010). Tlp3 PTPSD and HK1$_s$-Z2, HK1$_s$-Z3 and HK1$_s$-Z8 adopt a very similar fold [root-mean-square (r.m.s.) deviations of 2.5, 2.7 and 2.8 Å for the pairwise superimposition of 199, 195 and 200 C$^\alpha$ atoms from Z2, Z3 and Z8, respectively], despite the limited sequence homology (<18% identity for pairwise comparisons of Tlp3 PTPSD with HK1$_s$). Furthermore, a

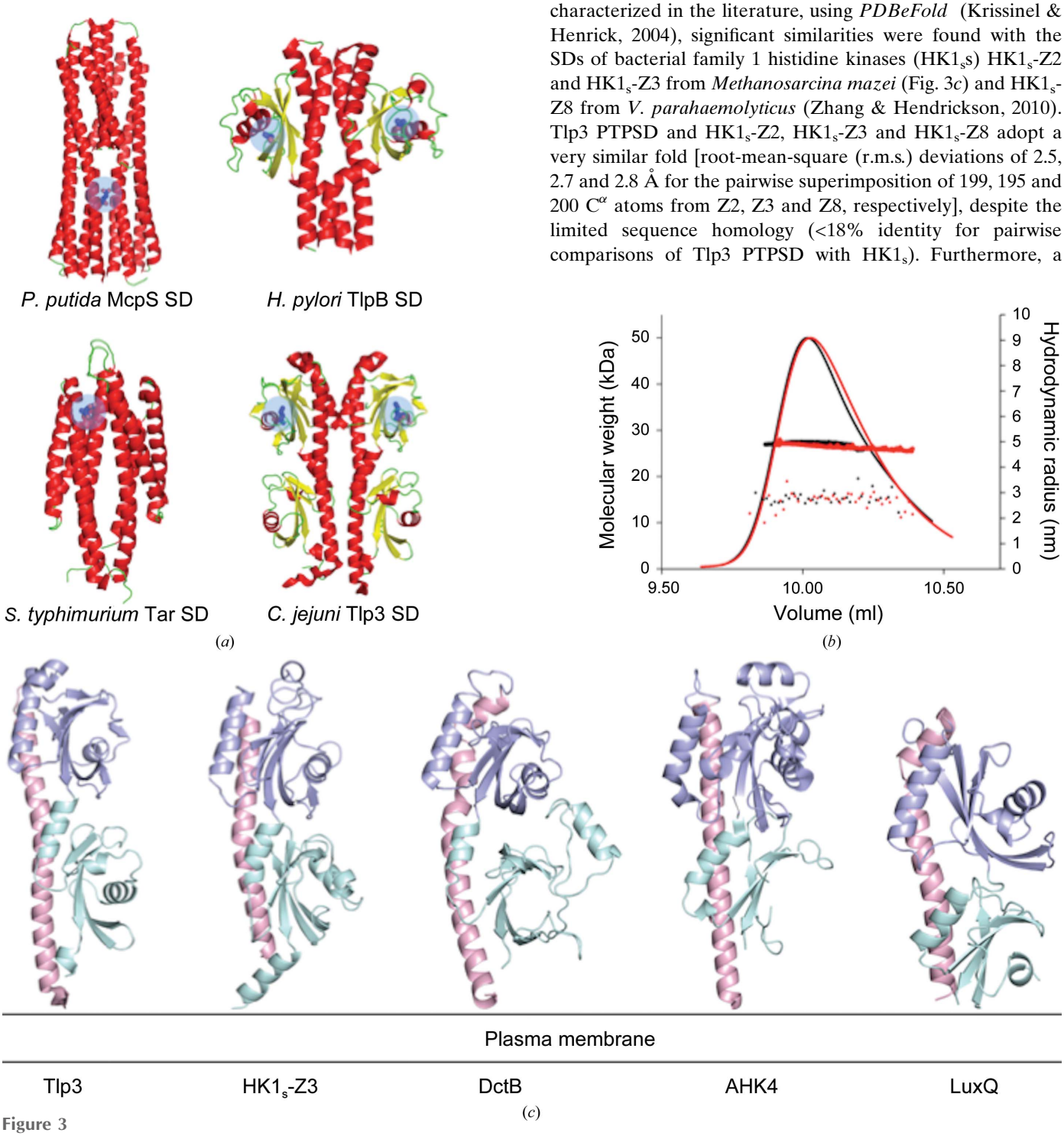


Figure 3
(*a*) Comparison of the structures of SD dimers for *P. putida* McpS (in complex with succinate; PDB entry 2yfb; Pineda-Molina *et al.*, 2012), *H. pylori* TlpB (in complex with urea; PDB entry 3ub6; Sweeney *et al.*, 2012), *S. typhimurium* Tar (in complex with aspartate; PDB entry 2lig; Milburn *et al.*, 1991) and *C. jejuni* Tlp3 (in complex with isoleucine). The bound ligands are shown in stick mode and the locations of the ligand-binding sites are highlighted with a blue circle. (*b*) Size-exclusion chromatography and molecular-weight and hydrodynamic radius (R_h) determination of Tlp3 PTPSD in the absence (black) and presence (red) of isoleucine. A bold solid line superimposed on the peak indicates the molecular weight as shown on the left-hand *y* axis. Dots represent the calculated hydrodynamic radius; its values are shown on the right-hand *y* axis. (*c*) Comparison of the structures of the periplasmic SDs of Tlp3, HK1$_s$-Z3 (PDB entry 3lib; Zhang & Hendrickson, 2010), DctB (PDB entry 3by9; Zhou *et al.*, 2008), AHK4 (PDB entry 3t4k; Hothorn *et al.*, 2011) and LuxQ (PDB entry 1zhh; Neiditch *et al.*, 2005).

similar fold comprising a long stalk helix followed by two PAS domains has previously been observed in the SDs of eukaryotic HKs, including *Arabidopsis thaliana* HK4 (AHK4; Hothorn *et al.*, 2011) and bacterial HKs from other families, including the *Sinorhizobium meliloti* C4-dicarboxylate transport sensory HK DctB (Zhou *et al.*, 2008) and the *V. harveyi* luminescence (lux) system HK LuxQ (Neiditch *et al.*, 2005), despite a very low degree of sequence conservation between those receptors and Tlp3 (Fig. 3*c*).

3.2. Isoleucine-binding site in the membrane-distal PAS domain

Analysis of the molecular surface of free Tlp3 PTPSD using *CASTp* (Dundas *et al.*, 2006) with a probe radius of 1.4 Å revealed putative ligand-binding pockets in both the membrane-distal and proximal PAS domains, with solvent-accessible volumes of 630 and 350 Å^3, respectively. The near-atomic resolution (1.3 Å) electron-density maps of the isoleucine complex of Tlp3 PTPSD are of high quality (Fig. 4*a*) and clearly show an almost identical mode of binding of isoleucine to the distal PAS domains of both subunits of the dimer in the asymmetric unit. The aliphatic side chain of isoleucine is in a largely hydrophobic environment, making van der Waals contacts with the side chains of Tyr118, Val126, Trp151 and Val171 (Fig. 4*a*) and approaching within 4.6 and 4.0 Å of the side chains of Leu128 and Leu144, respectively. The amino group forms hydrogen bonds to Tyr167 O$^{\eta}$, Asp169 O$^{\delta1}$ and Asp196 O$^{\delta2}$. The complex is further stabilized by hydrogen bonds between the carboxyl O atoms of isoleucine and Lys149 N$^{\zeta}$, Trp151 N$^{\varepsilon}$, Asp169 O$^{\delta2}$ and Thr170 O$^{\gamma}$. ASA calculations show that isoleucine is fully shielded from solvent upon binding to Tlp3.

To assess the contribution of individual amino-acid residues of Tlp3 to isoleucine binding, comprehensive alanine-scanning mutagenesis of the binding pocket was undertaken and the

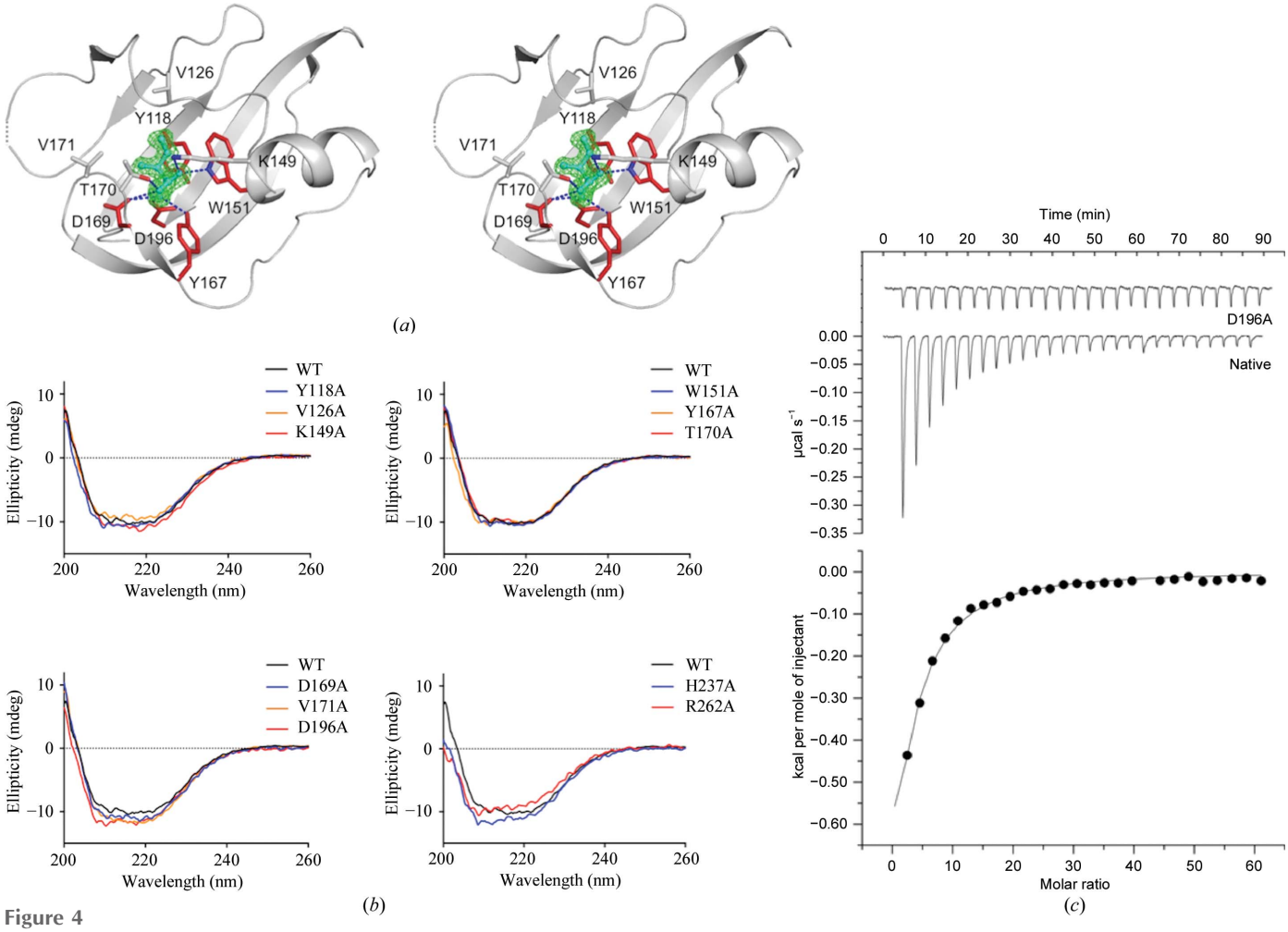
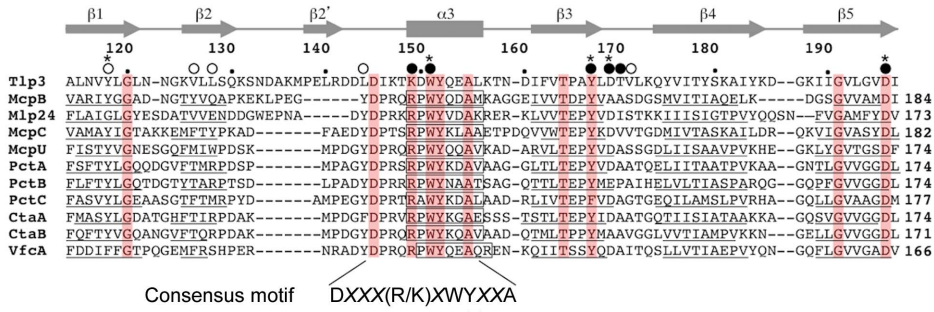
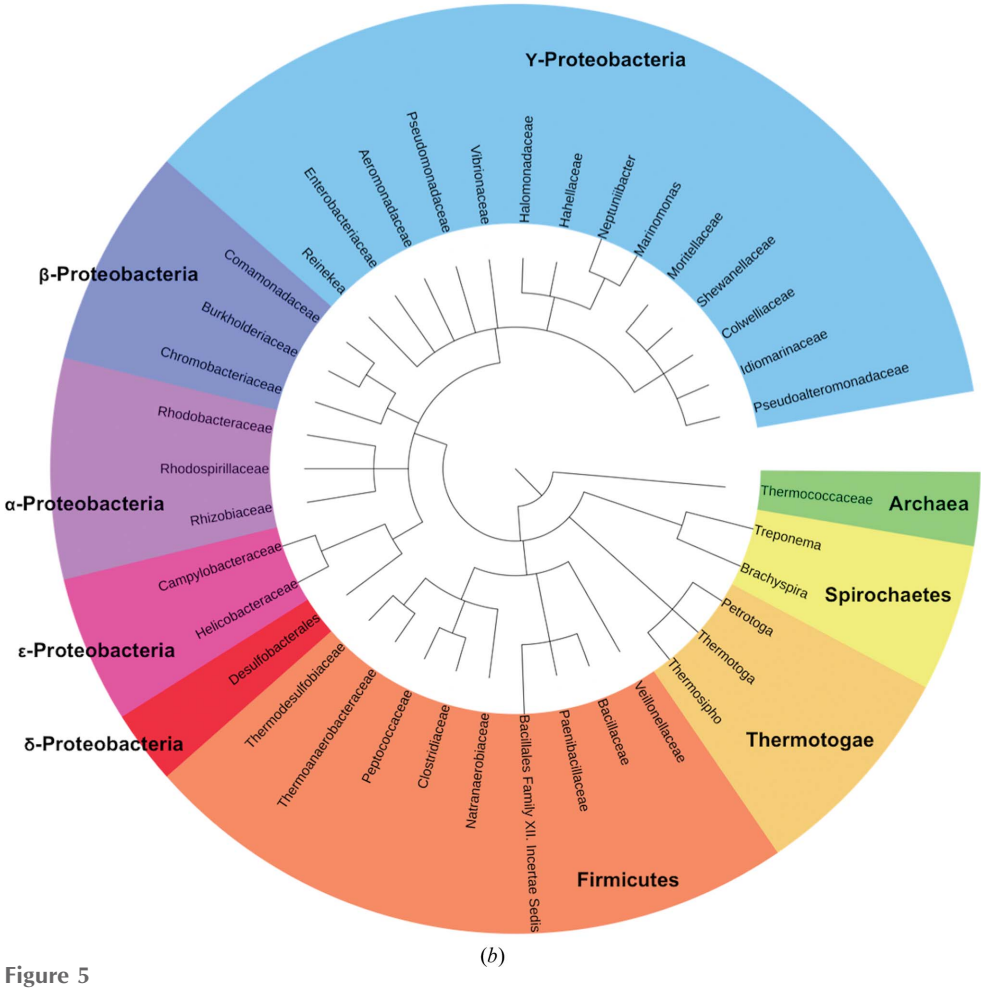


Figure 4
(*a*) The architecture of the ligand-binding site in the distal PAS domain. The ($mF_o - DF_c$) σ_A-weighted (Winn *et al.*, 2011) electron density for isoleucine is shown in green. The map was calculated at 1.3 Å resolution and contoured at the 3.0σ level. The isoleucine molecule is shown in all-atom ball-and-stick representation with C atoms coloured green. The protein side chains that form direct contacts with isoleucine are shown in stick representation. Amino-acid residues for which alanine substitutions abolished isoleucine binding are shown in red. (*b*) CD spectra of wild-type Tlp3 PTPSD (WT) and its Y118A, V126A, K149A, W151A, Y167A, D169A, T170A, V171A, D196A, H237A and R262A variants. (*c*) ITC titrations of native Tlp3 PTPSD and its D196A variant with isoleucine. Each peak in the top figure corresponds to the injection of 10 µl of 3 m*M* isoleucine into a 1.45 ml reaction cell containing protein at a concentration of 10 µ*M*. The cumulative heat of reaction is displayed in the bottom figure as a function of the ligand:protein molar ratio. The solid line is the least-squares fit of the experimental data to a single-site binding model.

effect of the substitutions on the binding affinity was examined by ITC. A total of 11 variants were generated, including two with substitutions in the proximal PAS pocket (H237A and R262A) as a negative control. The far-UV CD spectra of all of the variants were similar to that of the wild-type protein (Fig. 4*b*), indicating that they were folded and had wild-type-like secondary structure. The substitutions Y118A, W151A, Y167A, D169A and D196A in the distal PAS domain reduced the affinity of isoleucine binding by at least 35-fold (Table 4), as exemplified by the ITC titration curve for Tlp3 PTPSD D196A (Fig. 4*c*). In contrast, the substitutions V126A, K149A, T170A, V171A (distal PAS), H237A and R262A (proximal PAS) had little effect on the affinity.



(*a*)



(*b*)

Figure 5
(*a*) Local sequence alignment of a representative subset of the characterized PTPSDs of MCP receptors for amino acids showing a consensus motif D*XXX*(R/K)*X*WY*XX*A in the ligand-binding site. The sequences are of Tlp3 from *C. jejuni*, McpB and McpC (Glekas *et al.*, 2012) from *B. subtilis*, Mlp24/McpX from *V. cholerae*, McpU from *S. meliloti*, PctA, PctB and PctC from *P. aeruginosa* (Rico-Jiménez *et al.*, 2013), CtaA and CtaB from *P. fluorescens* (Oku *et al.*, 2012) and VfcA from *V. fischeri* (Brennan *et al.*, 2013). Conserved residues are highlighted in pink. The positions of the Ala substitutions associated with a loss of isoleucine binding in Tlp3 are shown by asterisks. Residues that stabilize the invariant (main-chain) moiety and the side chain of the amino-acid ligand are marked with a filled and an open circle, respectively. (*b*) Phylogenetic distribution of species in the TrEMBL database with putative PTPSD-containing chemoreceptors for amino acids identified using the search for the consensus motif (see also Supplementary Table S1).

3.3. Comparison to other PTPSDs recognizing amino acids

To gain further insight into the structural basis of the ligand specificity of Tlp3 PTPSD, the amino-acid sequence of its ligand-binding region has been aligned with those of other characterized PTPSDs recognizing various amino acids, using their predicted secondary structures as an additional guide (Fig. 5*a*). This comparison revealed that, with the exception of Thr170, all of the protein side chains in Tlp3 PTPSD that form interactions with the amino and carboxyl groups of the ligand are strongly conserved in other amino-acid receptors containing a PTPSD. This finding suggests that the receptors of this structural family share a common mechanism of recognition of the invariant moiety of an amino acid. In contrast, the Tlp3 PTPSD residues that interact with the side chain of the ligand but do not form contacts with its invariant part are not conserved. For the known receptors with a narrow ligand specificity, there is an apparent correlation between the chemical nature of the amino acids lining the pocket for the side chain of the ligand and the physical properties of the latter. In Tlp3 PTPSD, for example, the pocket for the highly hydrophobic side chain of the isoleucine molecule is lined by the aliphatic side chains of Val126, Leu128, Leu144 and Val171. In the sequence of *B. subtilis* McpB

Table 4
Thermodynamic parameters of isoleucine binding to wild-type Tlp3 PTPSD and its variants derived from ITC measurements.

Data are means and standard deviations from two experiments.

Mutant	K_d (μ*M*)	Enthalpy, Δ*H* (cal mol^{-1})	Entropy, Δ*S* (cal mol^{-1} K^{-1})
Native	86 ± 10	−4415 ± 230	3.6 ± 1
Y118A	>3000	ND	ND
V126A	117 ± 19	−5838 ± 148	−1.6 ± 0.8
K149A	130 ± 1	−4530 ± 105	2.6 ± 0.3
W151A	>3000	ND	ND
Y167A	>3000	ND	ND
D169A	>3000	ND	ND
T170A	51 ± 2	−6745 ± 37	−2.9 ± 0.1
V171A	129 ± 10	−9127 ± 684	−12.8 ± 2.4
D196A	>3000	ND	ND
H237A	161 ± 6	−12460 ± 42	−24.5 ± 0.2
R262A	103 ± 3	−9083 ± 338	−12.2 ± 1

(Glekas *et al.*, 2010), which is specific for asparagine and, to a lesser degree, aspartate, glutamine and histidine, these residues are substituted by those with larger, polar side chains (Tyr121, Gln123, Tyr133 and Ser161, respectively), which creates hydrogen-bonding possibilities that favour a polar amino-acid ligand smaller than isoleucine. In the sequence of the proline receptor *S. meliloti* McpU (Webb *et al.*, 2014), the side chains of Val126 and Leu128 are substituted by Met110 and Trp112, respectively, the bulky hydrophobic side chains of which are likely to protrude into the ligand-binding pocket, reducing its size while retaining its hydrophobic nature. Thus, analysis of the sequence differences between PTPSDs of Tlp3 and other MCP receptors recognizing amino acids appears to be fully consistent with the differences in their specificity.

Furthermore, analysis of the sequence alignment of a representative subset of the characterized PTPSDs of MCP receptors for amino acids identified a consensus motif in the ligand-binding site, D*XXX*(R/K)*X*WY*XX*A, that comprises helix α3 and the loop N-terminal to it. A search for this motif in the TrEMBL database, containing approximately 20 000 protein sequences, with keyword 'chemotaxis' and a molecular-weight range of 60–85 kDa identified approximately 1100 putative MCP receptors from 824 different bacteria and archaea (Fig. 5*b*; Suppementary Table S1). Analysis of the computational predictions of the membrane topology and the fold of the sensory domain confirmed that all of the hits possess a PTPSD, suggesting that the motif D*XXX*(R/K)-*X*WY*XX*A may be used as a signature for the identification of PTPSD-containing amino-acid receptors for more systematic studies.

3.4. Isoleucine-induced conformational change

Crystallization of the free protein yielded a crystal form isomorphous to that for the isoleucine complex (Table 1). The crystal structure was determined by molecular replacement using the protein monomer of the latter as a search model with the ligand and water molecules removed. Superposition of the structures of free and isoleucine-bound Tlp3 PTPSD (Fig. 6) revealed a substantial conformational change in the distal PAS domain. Its ligand-binding pocket is relatively open and accessible in the structure of the free protein. In the complex the loop connecting β3 and β4 closes over the ligand-binding site, bringing residues Tyr167, Asp169 and Thr170 into contact with the isoleucine molecule. This movement breaks a hydrogen bond and van der Waals interaction of Tyr175, residing on this loop, with the loop connecting β6 and β7, thus weakening the association between the distal and proximal domains.

3.5. Comparison of the two subunits in the asymmetric unit, normal-mode analysis and a model for the molecular mechanism of transmembrane signalling

Superposition of the structures of the two halves of the Tlp3 PTPSD dimer revealed that the proximal domain of one subunit adopts a more open form than the other, resulting in a downward ∼4 Å piston displacement of the C-terminal helix α6 towards the membrane (Fig. 6). This difference in conformation was observed to the same extent in the crystals of both the free protein and its complex with isoleucine, and may be owing to different crystal-packing contacts. However, in the physiological scenario, where helices α1 at the dimer interface remain fixed perpendicular to the membrane, the different

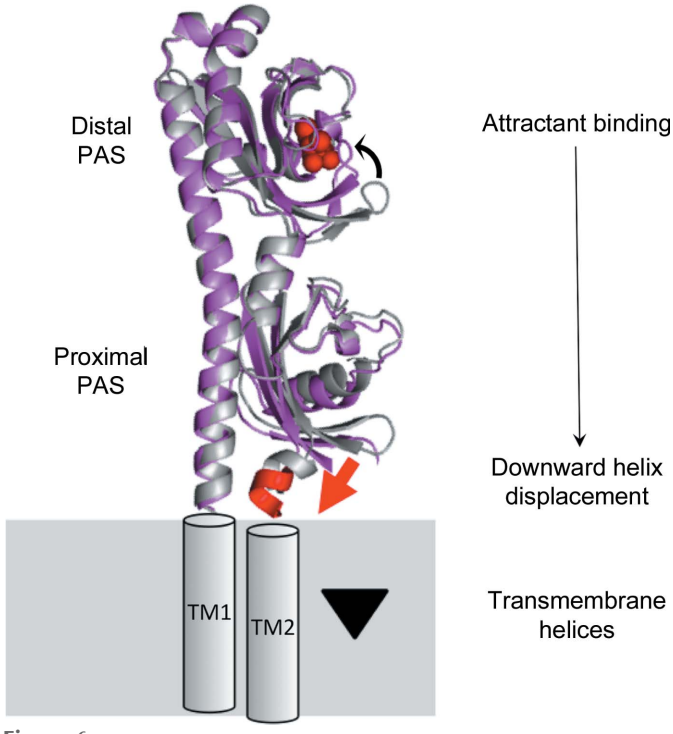


Figure 6
Piston model for transmembrane signalling by PTPSDs. The model is based on the superimposition of the two extreme conformational states of Tlp3 PTPSD observed in subunit *A* of the free protein (shown in grey) and subunit *B* of the isoleucine complex (shown in magenta/red), respectively. Attractant binding to the distal PAS domain locks it in the closed form, weakening its association with the proximal domain, which results in the transition of the latter into an open form, concomitant with a downward ∼4 Å displacement of the C-terminal helix towards the membrane (see also Supplementary Movies S1 and S2). *In vivo*, this movement would be coupled to a piston-type downward displacement of the TM2 towards the cytoplasm, generating a transmembrane signal.

conformations of the proximal PAS domains provide a likely mechanism by which it can transmit the signal across the membrane *via* a downward displacement of helix α6 towards the membrane.

Analysis of the conformational changes in Tlp3 PTPSD induced by isoleucine binding suggests that the transition of the proximal domain into a more open form *in vivo* is likely to occur when its association with the distal domain is weakened by disruption of the Tyr175-mediated link. The structures of subunit *A* in the free protein (distal PAS open, proximal PAS closed, piston up) and subunit *B* in the isoleucine complex (distal PAS closed around the ligand, proximal PAS open, piston down) therefore represent the extreme conformational states of free and ligand-bound Tlp3 PTPSD in solution. To test this hypothesis, patterns of concerted movements in the Tlp3 PTPSD subunit were studied using normal-mode analysis (NMA). NMA calculations indicated that movement of one extreme conformation (free, subunit *A*) along the fourth lowest-frequency mode generated a structure similar to the other extreme conformation (isoleucine-bound, subunit *B*) (Supplementary Movies S1and S2). NMA of the movements of Tlp3 PTPSD in solution was therefore consistent with a model of signalling in which the closure of the distal PAS domain around the ligand is accompanied by the opening of the proximal domain and the displacement of its C-terminal helix towards the membrane (Fig. 6).

4. Discussion

This paper presents the first X-ray crystallographic analysis of the structure of and the ligand-induced conformational changes in the PTPSD of a characterized MCP chemoreceptor performed at near-atomic resolution (1.3 Å). Analysis of the crystal structure of the complex of *C. jejuni* Tlp3 PTPSD with the chemoattractant isoleucine reveals that upon interaction with the receptor the ligand is completely engulfed by the distal PAS domain, with extensive interactions formed between the protein and both side-chain and main-chain moieties of the ligand. Crystallographic localization of the isoleucine-binding site in the distal domain is supported by ITC experiments on the pocket mutants and is in line with previous mutagenesis studies on PTPSDs of other MCP receptors recognizing amino acids (Rico-Jiménez *et al.*, 2013; Nishiyama *et al.*, 2012; Glekas *et al.*, 2010; Webb *et al.*, 2014). Alanine substitutions of the five strongly conserved residues that stabilize the invariant (main-chain) moiety of the amino-acid ligand (Lys149, Trp151, Tyr167, Asp169 and Asp196) abrogated ligand binding both in Tlp3 PTPSD (this study) and its homologue *P. aeruginosa* PctA PTPSD (Rico-Jiménez *et al.*, 2013). This finding suggests that the mechanism of amino-acid recognition by this family of chemoreceptors is highly conserved and that the five conserved residues are important determinants that define specificity towards an amino acid, as opposed to a different small molecule.

The significance of the identification of the consensus motif in MCP receptors for amino acids is that it has now become possible to find new putative amino-acid chemoreceptors in a wide range of bacteria and archaea, which paves the way to the systematic study of this family. As one of the outcomes of this analysis, the identification of Tlp3 homologues in eubacteria (*Yersinia* spp., *Dickeya* spp.; Supplementary Table S1) dispels the notion that all enterobacterial amino-acid receptors contain four-helix bundle SDs (Rico-Jiménez *et al.*, 2013).

Our crystallographic analysis shows that although PTPSD contains a four-helix bundle at the dimer interface, like all previously characterized chemoreceptor SDs with a different fold, the bundle does not interact with the ligand. The two ligand-binding sites of the PTPSD dimer are located on the molecular faces opposite to the dimer interface. Furthermore, dimerization is not required for ligand binding, as both free and isoleucine-bound Tlp3 PTPSD are monomeric in solution, in line with previous studies on PctA PTPSD, which also binds its amino-acid ligands as a monomer (Rico-Jiménez *et al.*, 2013). Thus, the mode of ligand recognition by PTPSD is distinctly different from that of Tar, in which the ligand-binding site is located at the dimer interface (Milburn *et al.*, 1991). It is therefore particularly interesting that, despite the different fold and mode of ligand recognition, our structural analysis suggests that like the four-helix bundle SDs, PTPSDs of amino-acid chemoreceptors signal by a piston-displacement mechanism. NMA and structural comparison of the two subunits in free and isoleucine-bound Tlp3 PTPSD indicate that PTPSD may fluctuate between two conformations that correspond to the piston (C-terminal helix) 'up' and piston 'down' states. Binding of an attractant to the distal PAS domain locks it in the form closed around the ligand, which is likely to weaken its association with the proximal domain, resulting in its transition into an open form, concomitant with a downward (towards the membrane) $\sim$4 Å piston displacement of the C-terminal helix α6. In the full-length receptor, this movement could generate a transmembrane signal by driving a piston-type downward displacement of the transmembrane helix 2 (TM2) towards the cytoplasm.

Previous studies have shown that the two binding pockets in Tar display a negative cooperativity: the attractant binds to only one of the two, and upon binding TM2 of one monomer shifts with respect to TM2′ of the other (Chervitz & Falke, 1996). Although we have observed structural asymmetry in the two halves of the Tlp3 PTPSD dimer in the crystal, both binding pockets were occupied, possibly owing to the concentration of isoleucine in the crystallization mixture exceeding the K_d by 100-fold. It remains to be established whether binding of the attractant induces a symmetric or asymmetric change in the two halves of the Tlp3 PTPSD dimer *in vivo*. We note that earlier cross-linking studies on *B. subtilis* McpB detected no substantial change of the TM2–TM2′ interface upon asparagine binding (Szurmant *et al.*, 2004), which is consistent with a symmetric model of signalling by McpB PTPSD.

The presented structural analysis provides a foundation for more systematic mutagenesis and biochemical studies. Structural studies on the new putative amino-acid receptors of this type, as identified in this study, will be valuable. In addition, the presence of a putative 350 Å^3 ligand-binding pocket in the

proximal PAS domain, which opens and closes in counter-phase to the distal domain, offers a possible mechanism for modulating the receptor activity by small-molecule ligands (*e.g.* repellents).

5. Related literature

The following references are cited in the Supporting Information for this article: Alexiev *et al.* (2014), Beesley *et al.* (2010), Di Franco *et al.* (2002), Gilad (2007) and Mesbah *et al.* (2007).

Acknowledgements

We thank Victoria Korolik (Griffith University, Australia) for her contribution to the conception of the study and reagents. Part of this research was undertaken on the MX1 and MX2 beamlines of the AS, Victoria, Australia. We thank the AS staff for their assistance with data collection. This work was supported by an Australian Research Council Research Fellowship to AR (DP1094619).

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Alexiev, A., Coil, D. A., Badger, J. H., Enticknap, J., Ward, N., Robb, F. T. & Eisen, J. A. (2014). *Genome Announc.* **2**, e00470-14.
Anantharaman, V. & Aravind, L. (2000). *Trends Biochem. Sci.* **25**, 535–537.
Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* D**67**, 271–281.
Beesley, C. A., Vanner, C. L., Helsel, L. O., Gee, J. E. & Hoffmaster, A. R. (2010). *FEMS Microbiol. Lett.* **313**, 47–53.
Brennan, C. A., DeLoney-Marino, C. R. & Mandel, M. J. (2013). *Appl. Environ. Microbiol.* **79**, 1889–1896.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Chervitz, S. A. & Falke, J. J. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 2545–2550.
Di Franco, C., Beccari, E., Santini, T., Pisaneschi, G. & Tecce, G. (2002). *BMC Microbiol.* **2**, 33.
Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y. & Liang, J. (2006). *Nucleic Acids Res.* **34**, W116–W118.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* D**69**, 1204–1214.
Fox, J. E., Gulledge, J., Engelhaupt, E., Burow, M. E. & McLachlan, J. A. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 10282–10287.
Gilad, J. (2007). *Recent Pat. Anti-Infect. Drug Discov.* **2**, 233–241.
Glekas, G. D., Foster, R. M., Cates, J. R., Estrella, J. A., Wawrzyniak, M. J., Rao, C. V. & Ordal, G. W. (2010). *J. Biol. Chem.* **285**, 1870–1878.
Glekas, G. D., Mulhern, B. J., Kroc, A., Duelfer, K. A., Lei, V., Rao, C. V. & Ordal, G. W. (2012). *J. Biol. Chem.* **287**, 39412–39418.
Goers Sweeney, E., Henderson, J. N., Goers, J., Wreden, C., Hicks, K. G., Foster, J. K., Parthasarathy, R., Remington, S. J. & Guillemin, K. (2012). *Structure*, **20**, 1177–1188.
Hothorn, M., Dabi, T. & Chory, J. (2011). *Nature Chem. Biol.* **7**, 766–768.
Josenhans, C. & Suerbaum, S. (2002). *Int. J. Med. Microbiol.* **291**, 605–614.
Krell, T., Lacal, J., Muñoz-Martínez, F., Reyes-Darias, J. A., Cadirci, B. H., García-Fontana, C. & Ramos, J. L. (2011). *Environ. Microbiol.* **13**, 1115–1124.
Krissinel, E. & Henrick, K. (2004). *Acta Cryst.* D**60**, 2256–2268.
Letunic, I. & Bork, P. (2011). *Nucleic Acids Res.* **39**, W475–W478.
Letunic, I., Doerks, T. & Bork, P. (2015). *Nucleic Acids Res.* **43**, D257–D260.
Machuca, M. A., Liu, Y. C., Beckham, S. A. & Roujeinikova, A. (2015). *Acta Cryst.* F**71**, 211–216.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
McPhillips, T. M., McPhillips, S. E., Chiu, H.-J., Cohen, A. E., Deacon, A. M., Ellis, P. J., Garman, E., Gonzalez, A., Sauter, N. K., Phizackerley, R. P., Soltis, S. M. & Kuhn, P. (2002). *J. Synchrotron Rad.* **9**, 401–406.
Mesbah, N. M., Hedrick, D. B., Peacock, A. D., Rohde, M. & Wiegel, J. (2007). *Int. J. Syst. Evol. Microbiol.* **57**, 2507–2512.
Milburn, M. V., Privé, G. G., Milligan, D. L., Scott, W. G., Yeh, J., Jancarik, J., Koshland, D. E. Jr & Kim, S.-H. (1991). *Science*, **254**, 1342–1347.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Neiditch, M. B., Federle, M. J., Miller, S. T., Bassler, B. L. & Hughson, F. M. (2005). *Mol. Cell*, **18**, 507–518.
Nishiyama, S., Suzuki, D., Itoh, Y., Suzuki, K., Tajima, H., Hyakutake, A., Homma, M., Butler-Wu, S. M., Camilli, A. & Kawagishi, I. (2012). *Infect. Immun.* **80**, 3170–3178.
Obenauer, J. C., Cantley, L. C. & Yaffe, M. B. (2003). *Nucleic Acids Res.* **31**, 3635–3641.
Oku, S., Komatsu, A., Tajima, T., Nakashimada, Y. & Kato, J. (2012). *Microb. Environ.* **27**, 462–469.
Ortega, A., Amorós, D. & García de la Torre, J. (2011). *Biophys. J.* **101**, 892–898.
Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *J. Comput. Chem.* **25**, 1605–1612.
Pineda-Molina, E., Reyes-Darias, J.-A., Lacal, J., Ramos, J. L., Garcia-Ruiz, J. M., Gavira, J. A. & Krell, T. (2012). *Proc. Natl Acad. Sci. USA*, **109**, 18926–18931.
Rahman, H., King, R. M., Shewell, L. K., Semchenko, E. A., Hartley-Tassell, L. E., Wilson, J. C., Day, C. J. & Korolik, V. (2014). *PLoS Pathog.* **10**, e1003822.
Rico-Jiménez, M., Muñoz-Martínez, F., García-Fontana, C., Fernandez, M., Morel, B., Ortega, A., Álvaro, O., Ramos, J. L. & Krell, T. (2013). *Mol. Microbiol.* **88**, 1230–1243.
Rosenberg, E., Koren, O., Reshef, L., Efrony, R. & Zilber-Rosenberg, I. (2007). *Nature Rev. Microbiol.* **5**, 355–362.
Söding, J., Biegert, A. & Lupas, A. N. (2005). *Nucleic Acids Res.* **33**, W244–W248.
Stocker, R. & Seymour, J. R. (2012). *Microbiol. Mol. Biol. Rev.* **76**, 792–812.
Suhre, K. & Sanejouand, Y.-H. (2004). *Nucleic Acids Res.* **32**, W610–W614.
Szurmant, H., Bunn, M. W., Cho, S. H. & Ordal, G. W. (2004). *J. Mol. Biol.* **344**, 919–928.
Webb, B. A., Hildreth, S., Helm, R. F. & Scharf, B. E. (2014). *Appl. Environ. Microbiol.* **80**, 3404–3415.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Zhang, Z. & Hendrickson, W. A. (2010). *J. Mol. Biol.* **400**, 335–353.
Zhou, Y.-F., Nan, B., Nan, J., Ma, Q., Panjikar, S., Liang, Y.-H., Wang, Y. & Su, X. D. (2008). *J. Mol. Biol.* **383**, 49–61.