Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of the *Thermus thermophilus* putative periplasmic glutamate/glutamine-binding protein

Hitomi Takahashi, Eiji Inagaki, Chizu Kuroishi and Tahir H. Tahirov*

Highthroughput Factory, RIKEN Harima Institute, 1-1-1 Kouto, Mikazuki-cho, Sayo-gun, Hyogo 679-5148, Japan

Correspondence e-mail: tahir@spring8.or.jp

Received 16 June 2004
Accepted 6 August 2004

PDB References: native *Tt*GluBP, 1us4, r1us4sf; selenomethionyl *Tt*GluBP, 1us5, r1us5sf.

As part of a structural genomics project, the crystal structure of a 314-amino-acid protein encoded by *Thermus thermophilus* HB8 gene TT1099 was solved to 1.75 Å using the multiple-wavelength anomalous dispersion (MAD) method and a selenomethionine-incorporated protein. The native protein structure was solved to 1.5 Å using the molecular-replacement method. Both structures revealed a bound ligand, L-glutamate or L-glutamine, and a fold related to the periplasmic substrate-binding proteins (PSBP). Further comparative structural analysis with other PSBP-fold proteins revealed the conservation of the predicted membrane permease binding surface area and indicated that the *T. thermophilus* HB8 molecule is most likely to be an L-glutamate and/or an L-glutamine-binding protein related to the cluster 3 periplasmic receptors. However, the geometry of ligand binding is unique to the *T. thermophilus* HB8 molecule.

1. Introduction

Structural genomics projects are often aimed towards revelation of the possible functions of unknown proteins by analysis of their three-dimensional structures (Yokoyama *et al.*, 2000). The target of current studies was a 314-amino-acid protein encoded by the *Thermus thermophilus* HB8 gene TT1099. A sequence-homology search revealed that the protein shares low sequence homology with putative immunogenic proteins of unclear function, while a structure-based *DALI* (Holm & Sander, 1995) search of the PDB using the solved structure revealed that the protein has significant structural similarity to PSBPs (Oh *et al.*, 1993) and the ligand-binding cores of glutamate receptors (Armstrong & Gouaux, 2000; Mayer *et al.*, 2001).

The bacterial periplasmic transport system, which transports a wide variety of substrates, consists of a PSBP (initial receptor) and a membrane permease that actually translocates the substrate from the periplasm to the cytoplasm (Ames *et al.*, 1990; Furlong, 1987). The role of the PSBP is primarily that of providing a substrate to the membrane permease. Substrates first bind to the binding proteins with high affinity rather than binding directly to the membrane component (Quiocho, 1990; Treptow & Shuman, 1985). After binding their respective substrates, the binding proteins interact with the membrane-bound complex and translocation of the substrates from the binding proteins to the cytoplasm takes place with concomitant ATP or GTP hydrolysis by the membrane-associated proteins (Ames *et al.*, 1992). The substrates transported by these permeases include amino acids, peptides, monosaccharides, oligosaccharides, inorganic ions *etc.* (Ames *et al.*, 1990). The periplasmic receptors consist of two similar

domains connected by a two- or three-stranded hinge. In the absence of substrate, the molecules are in an open conformation (Björkman & Mowbray, 1998). Small-angle X-ray scattering results have suggested that these proteins undergo large conformational changes in association with substrate binding (Newcomer *et al.*, 1981; Shilton *et al.*, 1996). The structure of the protein complexed with a ligand (L-glutamate or L-glutamine) from *T. thermophilus* is most similar to the structure of an amino-acid-binding protein, lysine/arginine/ornithine-binding protein from *Salmonella typhimurium* (*St*LAOBP) complexed with lysine (Oh *et al.*, 1993), although there is no significant homology between the amino-acid sequences of these two proteins.

The ancient PSBP fold is widely distributed in eukaryotes, archaebacteria and prokaryotes and is observed in proteins with diverse functions, *e.g. lac* transcriptional repressors, ionotropic and metabotropic glutamate receptors, G-protein-coupled receptors, calcium-sensing and pheromone receptors *etc*. (Felder *et al.*, 1999). The mammalian PSBP-fold proteins are often implicated in various dysfunctions such as Alzheimer's, Parkinson's and Huntington's diseases, schizophrenia *etc*. and are attractive targets for structure-based drug design.

The protein structure from *T. thermophilus* also reveals homology with the ligand-binding cores of GluR0, a prokaryotic glutamate-receptor ion channel from *Synechocystis* PCC 6803 (Mayer *et al.*, 2001), and GluR2, a rat AMPA-subtype neurotransmitter receptor (Armstrong & Gouaux, 2000). These glutamate receptors (GluRs) are ligand-gated ion channels activated by the amino acid L-glutamate and mediate excitatory synaptic transmission in the brain (Dingledine *et al.*, 1999; Mayer & Westbrook, 1987). An evolutionary and functional relationship between these neurotransmitter receptors and bacterial PSBPs suggests that the precursors of neurotransmitter receptors were created by gene-fusion events in bacteria involving a primitive ion channel similar to KcsA, the pH-sensitive potassium channel of *Streptomyces lividans*, and PSBPs (Chen *et al.*, 1999; Chiu *et al.*, 1999; Doyle *et al.*, 1998; Schrempf *et al.*, 1995).

Here, we determined a 1.5 Å resolution crystal structure of protein encoded by the *T. thermophilus* HB8 gene TT1099 and identified it as a putative L-glutamate and/or L-glutamine-binding protein (*Tt*GluBP) related to the cluster 3 periplasmic receptors (Tam & Saier, 1993).

2. Material and methods

2.1. Protein expression and purification

The gene was amplified by the polymerase chain reaction (PCR) using *T. thermophilus* HB8 genomic DNA as the template. The PCR product was ligated with pT7blue (Novagen). The plasmid was digested with *Nde*I and *Bgl*II and the fragment was inserted into the expression vector pET-11a linearized with *Nde*I and *Bam*HI. The recombinant plasmid was transformed into *Escherichia coli* BL21 (DE3) cells and grown at 310 K in LB medium containing 50 μg ml^{-1} ampicillin for 20 h. The cells were harvested by centrifugation at 6500 rev min^{-1} for 5 min, suspended in 20 m*M* Tris–HCl pH 8.0 (buffer *A*) containing 0.5 *M* NaCl and 5 m*M* 2-mercaptoethanol and disrupted by sonication. The supernatant was heated at 343 K for 11.5 min. After heat treatment, the cell debris and denatured proteins were removed by centrifugation (14 000 rev min^{-1}, 30 min) and the supernatant solution was used as the crude extract for purification. The crude extract was desalted using a HiPrep 26/10 desalting column (Amersham Biosciences) and applied onto a SuperQ Toyopearl 650M column (Tosoh) equilibrated with buffer *A*. The protein was eluted with a linear gradient of 0–0.2 *M* NaCl. The fraction containing the protein was desalted with HiPrep 26/10 with buffer *A* and subjected to a Resource S6 column

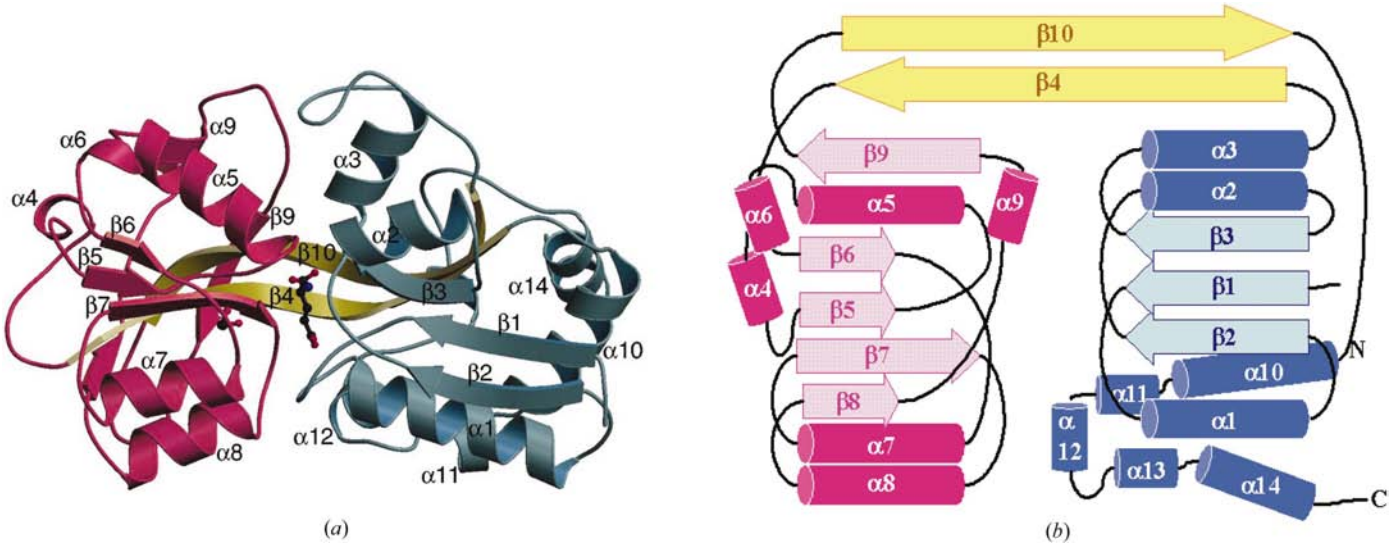

Figure 1
(*a*) Ribbon diagram of the *Tt*GluBP structure colour-coded by domain organization, with domain I coloured cyan, domain II pink and the interdomain strands yellow. L-Glutamate and an ethylene glycol are presented as a ball-and-stick model, with C atoms coloured black, N atoms blue and O atoms red. (*b*) Topology diagram of the *Tt*GluBP coloured as in (*a*), with arrows representing β-strands and cylinders representing α-helices.

(Amersham Biosciences) equilibrated with buffer *A*. The protein was eluted with a linear gradient of 0–0.2 *M* NaCl. The fraction containing the protein was desalted with HiPrep 26/10 containing 10 m*M* phosphate pH 7.0 and applied onto a Bio-Scale CHT-20-I column (Bio-Rad) equilibrated with the same buffer. The protein was eluted with a linear gradient of 10–80 m*M* phosphate pH 7.0. The fractions containing the protein were pooled, concentrated by ultrafiltration (Amicon, 10 kDa cut) and loaded onto a HiLoad 16/60 Superdex 75 column (Amersham Biosciences) equilibrated with buffer *A* containing 0.2 *M* NaCl. The purified protein was homogeneous on SDS–PAGE.

A selenomethionyl derivative was expressed in B834(DE3)plysS *E. coli* grown in M9 medium containing 50 mg l^{-1} D,L-selenomethionine and purified as described above for native protein.

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the last resolution shell.

Crystal type	Selenomethionine derivative			Native
Crystal data				
Unit-cell parameters (Å)				
a	43.44			51.52
b	69.48			68.09
c	103.80			81.38
Space group	$P2_12_12_1$			$P2_12_12_1$
Data collection				
Wavelength (Å)	0.9 (remote)	0.9789 (peak)	0.9793 (edge)	1.0000
Resolution (Å)	50–1.75 (1.81–1.75)	50–1.9 (1.97–1.9)	50–1.9 (1.97–1.9)	20–1.5 (1.53–1.5)
Measured reflections	123941	99093	99495	215880
Unique reflections	28082	22069	22197	50003
Completeness (%)	95.2 (98.0)	95.2 (97.90	95.1 (96.0)	97.7 (94.7)
$I/\sigma(I)$	25.1 (13.6)	27.2 (21.4)	27.3 (20.9)	20.7 (3.1)
R_{merge} (%)	6.7 (15.3)	7.1 (13.1)	7.1 (12.7)	6.5 (36.2)
MAD phasing (1.90 Å)				
R_{cullis}, dispersive/anomalous	—/0.72	0.68/0.62	0.56/0.58	
Phasing power	—/1.12	1.33/1.68	1.85/1.94	
Figure of merit	—/0.25	0.29/0.35	0.40/040	
For global phase set	0.70			
Refinement				
Resolution (Å)	29–1.75 (1.86–1.75)			20–1.5 (1.59–1.5)
R (%)	17.8 (17.8)			18.4 (26.8)
R_{free} (%)	20.6 (20.9)			20.5 (28.9)
Protein atoms	2259			2290
Heterogen atoms	14			14
Water molecules	235			286
R.m.s. deviations				
Bonds (Å)	0.005			0.005
Angles (°)	1.3			1.3
Average *B* factors (Å^2)				
Models	13.0			16.9
Wilson plot	9.4			13.3
Ramachandran plot (%)				
Favoured	91.4			89.8
Allowed	7.8			9.4
Generous	0.4			0.4
Disallowed	0.4			0.4

2.2. Crystallization and data collection

Native crystals of the protein in a complex with glutamate were obtained at 291 K using the oil-batch method by TERA (Sugahara & Miyano, 2002). A 0.5 µl aliquot of protein solution (20 mg ml^{-1} protein, 20 m*M* Tris–HCl pH 8.0, 0.2 *M* NaCl) was mixed with an equal volume of solution containing 22.5% PEG 4000, 1 *M* lithium chloride and 0.1 *M* sodium citrate pH 5 and covered with 15 µl paraffin oil. Within one month, crystals had grown to full size (0.03 × 0.03 × 0.15 mm) in the form of parallelepipeds. Glutamate-bound crystals of selenomethionyl protein were obtained by sitting-drop vapour diffusion using 10 mg ml^{-1} protein and a reservoir solution consisting of 20.3% PEG 4000, 0.09 *M* sodium citrate pH 5. Within 10 d, the crystals had grown to full size (0.05 × 0.05 × 0.15 mm) and had the same shape as the native crystals. Native and SeMet crystals belong to space group $P2_12_12_1$, with a monomer in the asymmetric unit, solvent contents of 42 and 45% and specific volumes V_M of 2.1 and 2.3 Å^3 Da^{-1} (Matthews, 1968), respectively. For data collection, these crystals were flash-cooled in a 100 K dry nitrogen stream; 18% ethylene glycol was added to the crystallization solution as a cryoprotectant. The data sets were collected using synchrotron radiation at beamline BL26B1, SPring-8, Japan. Multi-wavelength anomalous diffraction data sets were collected from a selenomethionyl crystal at 0.9789, 0.9793 and 0.9000 Å using 1° oscillations. High-resolution data sets were collected to 1.5 Å resolution from a native crystal. The data were processed with the programs *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). The X-ray crystallographic data are summarized in Table 1.

2.3. Structure determination and refinement

The structure was solved by the MAD method using three wavelengths of the Se-MAD experiment. Determination and refinement of the selenium sites, phase calculation and density modification were carried out with the program *CNS* (Brünger *et al.*, 1998). The initial map was of high quality (Fig. 4). Automatic model building was performed with the program *ARP*/*wARP* (Perrakis *et al.*, 1999) using the improved phases. 94% of the residues were automatically built. Four residues which remained were built manually using the program *TURBO-FRODO* (Roussel & Cambillau, 1996). The model was completed after two rounds of manual adjustments, addition of solvent molecules and refinement using standard protocols in *CNS* (Brünger *et al.*, 1987). The native molecule

structure has been solved by the molecular-replacement method using the coordinates of the SeMet structure and refined at 1.5 Å. The data-collection and refinement statistics are summarized in Table 1.

3. Results and discussion

3.1. The overall structure

The *Tt*GluBP structure was solved by the MAD method using an SeMet-substituted recombinant protein crystal (Table 1). The coordinates obtained were used to solve the structure of the native protein, the crystals of which belonged to a different crystal form and diffracted to 1.5 Å resolution. In both crystals the refined molecules were highly similar, with a root-mean-square deviation (r.m.s.d.) of 0.2 Å, contained residues 17–313 and revealed a two-domain molecule with dimensions of 60 × 45 × 45 Å and a bound ligand, L-glutamate or L-glutamine. One domain consists of residues 17–102 and 253–313 (domain I) and the other domain consists of residues 118–241 (domain II). The domains are separated by a deep cleft, the floor of which is formed by a pair of trans-domain antiparallel β-strands, β4 and β10 (Fig. 1*a*). Domain I, the larger of the two, includes eight α-helices and three β-strands and domain II includes six α-helices and five β-strands (Fig. 1*b*). The two domains exhibit a similar folding pattern consisting of a central core of pleated sheets surrounded by helices, which has been described as a

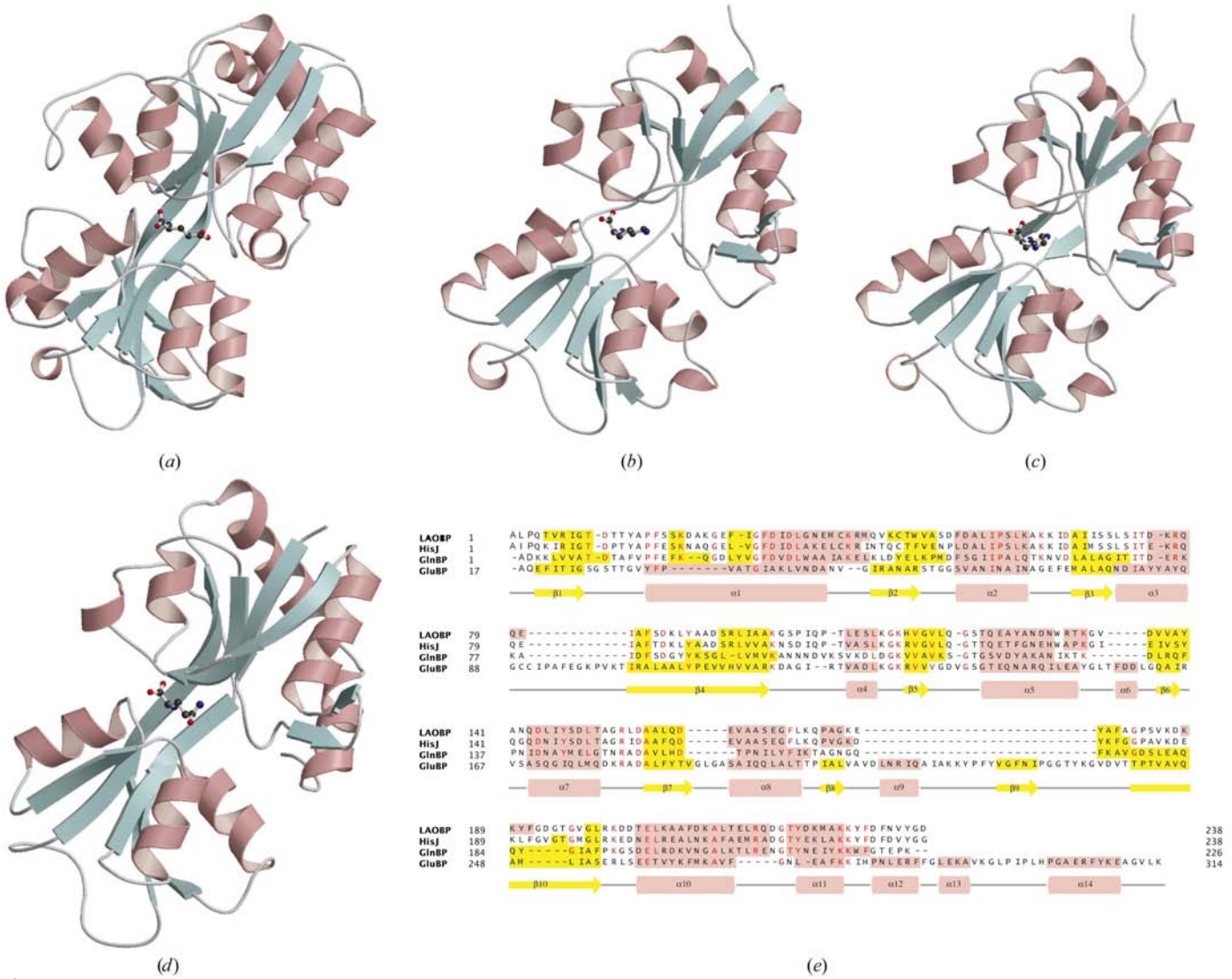

Figure 2
Structural homology of small amino-acid binding proteins. Ribbon diagrams of the crystal structures of (*a*) *Tt*GluBP complexed with L-glutamate (this study), (*b*) *St*LAOBP complexed with L-lysine, (*c*) *St*HisJ complexed with L-histidine and (*d*) *Ec*GlnBP complexed with L-glutamine. (*e*) Multiple sequence alignments and secondary-structure assignment by *PROCHECK* (Laskowski *et al.*, 1993). Secondary structure in *Tt*GluBP with ligand is displayed under the *St*LAOBP, *St*HisJ, *Ec*GlnBP and *Tt*GluBP sequences: helices are shown as pink cylinders and strands as yellow arrows, with the same labelling as in Fig. 1. The secondary structures for *St*LAOBP, *St*HisJ and *Ec*GlnBP are coloured identically to that for *Tt*GluBP. Red characters show the 55 conserved residues among *St*LAOBP, *St*HisJ and *Ec*GlnBP. Red characters and blue characters in the *Tt*GluBP sequence reveal the corresponding conserved residues and similar residues to the conserved residues among *St*LAOBP, *St*HisJ and *Ec*GlnBP.

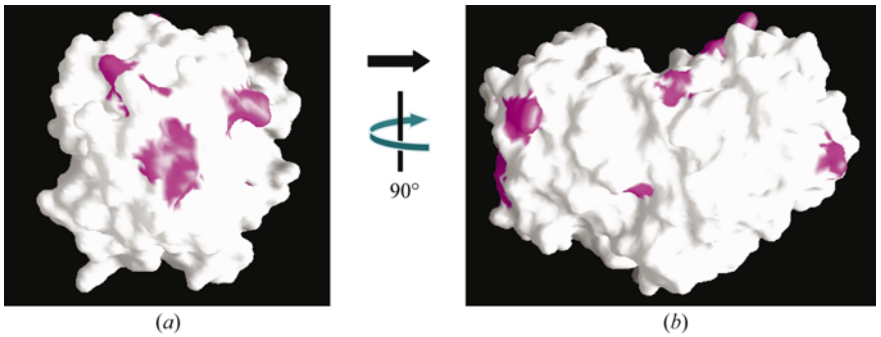

Figure 3
Molecular surface of *Tt*GluBP. (*b*) is seen from the same orientation as Fig. 1(*a*). (*a*) is rotated by 90° from the view in (*b*). The surface of exposed conserved residues is in pink.

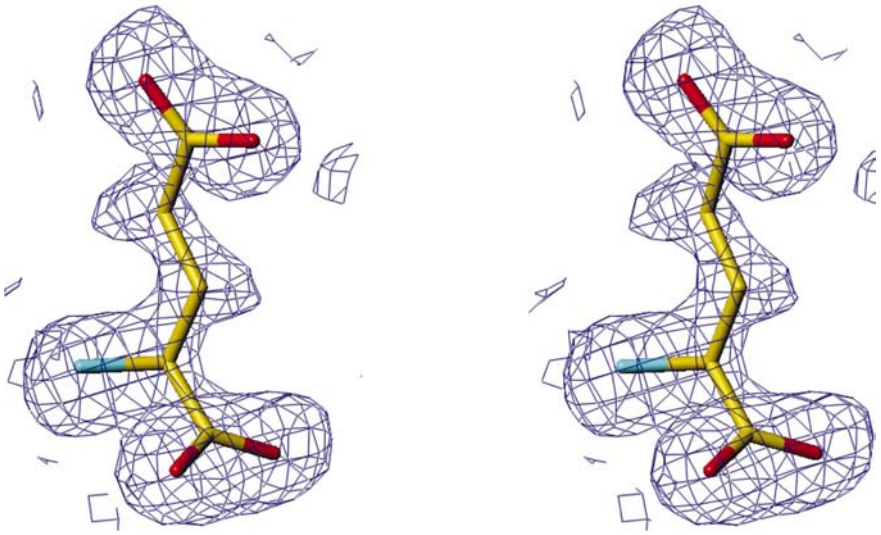

Figure 4
Electron-density map of the bound ligand in the crystal structure of SeMet protein. The map was calculated using the MAD phases after density modification. The contour level of the map is 1σ.

nucleotide-binding fold (Quiocho *et al.*, 1977; Rossmann & Argos, 1975), although domain I has more composition elements. Both ends of the long trans-domain antiparallel β-strands, which consists of β4 and β10, participate in formation of the β-sheets in each domain (Fig. 1*a*).

3.2. Comparison with PSBPs

The PSBPs can be classified into eight clusters based on their sequence similarities and these groupings were generally found to correlate with the molecular sizes and ligand-binding specificities of the included proteins (Tam & Saier, 1993). Of the PSBPs with known structures, five are amino-acid-binding proteins. Three of them, *St*LAOBP (Oh *et al.*, 1993), histidine-binding protein from *S. typhimurium* (*St*HisJ; Oh *et al.*, 1994; Yao *et al.*, 1994) and glutamine-binding protein from *E. coli* (*Ec*GlnBP; Sun *et al.*, 1998), belong to cluster 3, members of which are specific for polar amino acids and opins (Tam & Saier, 1993), while the other two, leucine/isoleucine/valine-binding protein from *E. coli* (*Ec*LIVBP; Sack, Saper *et al.*, 1989) and leucine-specific binding protein from *E. coli* (*Ec*LSBP; Sack, Trakhanov *et al.*, 1989), belong to cluster 4, members of which are specific for aliphatic hydrophobic amino acids. *St*LAOBP is 238 residues long and similar in size and structure to *St*HisJ and *Ec*GlnBP (Fig. 2*e*). *Ec*LIVBP and *Ec*LSBP, with 344 and 346 residues, respectively, are much larger. Alhough *Tt*GluBP, consisting of 314 residues, is closer to *Ec*LIVBP and *Ec*LSBP in size, the structural homologies between *Tt*GluBP and these proteins are much lower. The folding of *Tt*GluBP resembles the structures of the cluster 3 proteins (Fig. 2).

The three small periplasmic amino-acid-binding proteins *St*LAOBP, *St*HisJ and *Ec*GlnBP are structurally similar to each other. Structure-based least-squares superpositions of *St*LAOBP with *St*HisJ and *Ec*GlnBP gave r.m.s.d.s of 1.1 and 1.6 Å, respectively, while those of *Tt*GluBP (except the C-terminal extension, α13 and α14) with *St*LAOBP or *St*HisJ and *Ec*GlnBP gave r.m.s.d.s of 3.1 and 3.2 Å, respectively (Figs. 2*a*–2*d*).

3.3. Surface properties

*St*LAOBP and *St*HisJ show good sequence identity (70%), while *Ec*GlnBP has less than 30% identity with either *St*LAOBP or *St*HisJ. A multiple sequence-alignment program (Feng & Doolittle, 1987) identified 55 conserved residues among these three proteins (Fig. 2*e*). A sequence alignment containing *Tt*GluBP based on the secondary structures suggests that in spite of there being no significant homologies with other three proteins, *Tt*GluBP has equivalent residues to 21 of the 55 conserved residues in the other three proteins (Fig. 2*e*). Of these 21 residues, none participate in interaction with the ligand; however, 13 of the 21 residues are exposed on the molecular surface (Fig. 3). Some of these 13 exposed residues are concentrated on α4 and β5 and the loop between these elements and the loop between α7 and β7 which is located near β5 at the one end of the molecule (Fig. 2*e*). The fact that *St*LAOBP and *St*HisJ, which share the same membrane transporter, exhibit a high sequence similarity (∼90%) in this region suggests that this may be related to recognition of a PSBP by its target membrane transporter (Ames, 1986). Moreover, this region, which connects with the membrane-embedded domains, may be important for signal transmission in the ligand-binding cores of the ionotropic glutamate receptors (Armstrong & Gouaux, 2000; Armstrong *et al.*, 1998; Sun *et al.*, 2002). The conformations of the helix and the loop between β4 and β5 in *Tt*GluBP are similar to the conformations of corresponding elements in *St*LAOBP and *St*HisJ;

however, the corresponding elements assume a different conformation in *Ec*GlnBP. Molecular-dynamics simulations of *Ec*GlnBP show that this part of the molecule has a high degree of flexibility (Pang *et al.*, 2003) and we speculate that the interaction with the membrane transporter may induce a fold observed in the other three proteins.

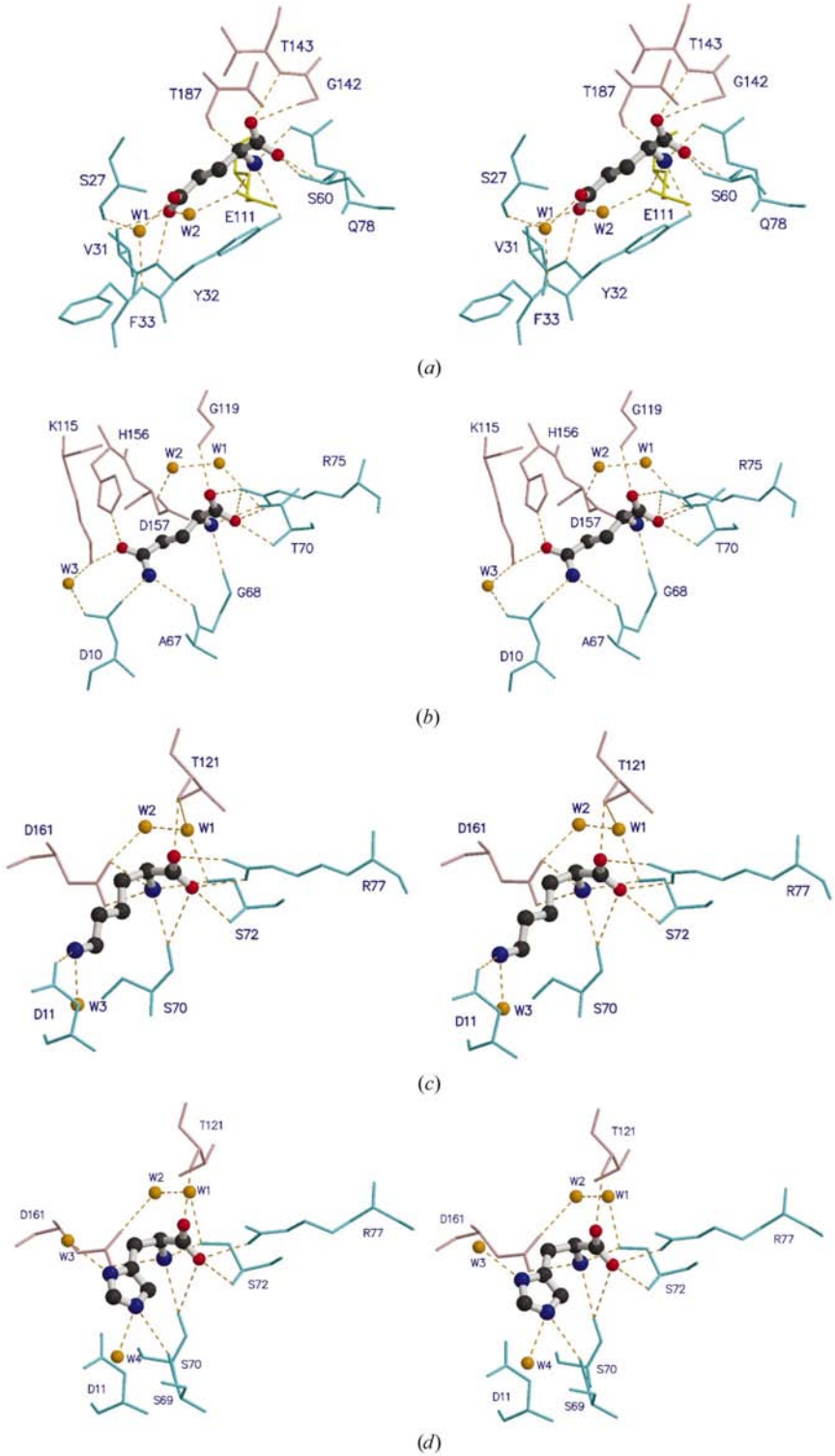

Figure 5
Stereoview of the ligand-binding sites of (*a*) the *Tt*GluBP complex with L-glutamate, (*b*) the *Ec*GlnBP complex with L-glutamine, (*c*) the *St*LAOBP complex with L-lysine and (*d*) the *St*HisJ complex with L-histidine. As in Fig. 1, domain I and II residues are coloured cyan and pink, respectively, and a residue in a trans-domain strand of the *Tt*GluBP complex is coloured yellow. Ligands are depicted in ball-and-stick representation coloured by domain with associated C, N and O atoms in black, blue and red, respectively. Water molecules are shown as orange balls and hydrogen bonds are represented by orange dotted lines.

3.4. The ligand binding

The initial electron-density map calculated with modified MAD phases for the SeMet crystal showed the existence of a well defined ligand in the structure (Fig. 4). The $2F_o - F_c$ and $F_o - F_c$ Fourier maps confirmed the presence of ligand in native protein crystal. Although both L-glutamate and L-glutamine fit well to this density, in subsequent discussions the ligand is referred to as L-glutamate. The fact that the ligand was not released from the protein during the various steps of the purification protocol indicates that it is very tightly bound. As in all other ligand-bound 'closed-cleft' structures, the glutamate is completely buried in a pocket formed between the two domains (Lee & Richards, 1971) and makes contacts with both domains (Fig. 1*a*). The trans-domain β-strand $\beta4$ also participates in binding to the ligand (Fig. 5*a*). In the *Tt*GluBP–Glu complex a total of ten hydrogen bonds, six from domain I, three from domain II and one from the trans-domain strand, are observed between the protein and the ligand. The side chains participating in glutamate binding include charged, polar and non-polar side-chains, as well as atoms from the peptide backbone.

The α-carboxyl group of the ligand interacts with Ser60, Gly142 and Thr143. The main-chain N atoms of Ser60, Gly142 and Thr143 and the hydroxyl group of Ser60 form hydrogen bonds. The α-amino group of the ligand is stabilized by Tyr32, Gln78, Glu111 and Thr187: the NH_3^+ moiety of the ligand forms hydrogen bonds with the hydroxyl group of Tyr32, the side chains of Gln78 and Glu111 and the main-chain O atom of Thr187. Long-range ionic interactions can also be found between the NH_3^+ moiety of the ligand and the carbonyl groups of Glu111 and Gln78. The side chain of the glutamate ligand was anchored and stabilized by Ser27, Val31, Tyr32 and Phe33. The main-chain N atoms of Val31 and Tyr32 directly form hydrogen bonds with $O^{\varepsilon1}$ of the ligand, while the hydroxyl group of Ser27 and the main chain of Phe33 form water-mediated hydrogen bonds with $O^{\varepsilon2}$ of

the ligand (Fig. 5*a*). Although the hydrogen bonds dominate the protein–ligand interactions, the hydrophobic interactions within the ligand-binding pocket also play an important role. The aliphatic portion of the glutamate ligand is sandwiched in a hydrophobic pocket formed between Tyr32 and Tyr186.

3.5. Comparison with ligand binding by *St*LAOBP, *St*HisJ and *Ec*GlnBP

Great similarities exist in ligand binding among the *St*LAOBP–Lys (Oh *et al.*, 1993), *St*HisJ–His (Oh *et al.*, 1994; Yao *et al.*, 1994) and *Ec*GlnBP–Gln (Sun *et al.*, 1998) complexes. Three of the 55 conserved residues among these proteins, Asp10, Arg75 and Asp157 in the *Ec*GlnBP complex, are involved in ligand binding by forming hydrogen bonds or ionic interactions (Fig. 5*b*). The corresponding residues are Asp11, Arg77 and Asp161 in the *St*LAOBP and *St*HisJ complexes (Figs. 5*c* and 5*d*). On the other hand, *Tt*GluBP has no corresponding residues. The binding interactions between the protein and the α-amino and α-carboxyl groups of the ligand are nearly the same in all the three proteins. In *Ec*GlnBP, the α-amino group of the ligand is stabilized by interacting with residues Gly68, Thr70 and the conserved Asp157. The NH_3^+ moiety of the ligand makes hydrogen bonds with the main-chain carboxyl O atom of Gly68, the hydroxyl group of Thr70 and the side chain of Asp157. An ionic interaction can also be found between the side chain of Asp157 and the NH_3^+ moiety of the ligand. The α-carboxyl group of the ligand interacts with Thr70, the conserved Arg75 and Gly119. The main-chain N atoms of Thr70 and Gly119 form a hydrogen bond, while the side chain of the conserved Arg75 neutralizes the opposite charges and forms salt bridges with the COO^- group of the ligand (Fig. 5*b*). In the *St*LAOBP and *St*HisJ complexes the corresponding residues interact with the ligands in a similar manner. The ligand binding in these three PSBPs differs appreciably from that of *Tt*GluBP, in which the ligand α-carboxyl group is mainly hydrogen bonded to main-chain NH groups (Fig. 5). However, Gly142, whose main-chain N atom interacts with the ligand α-carboxyl group, Thr187 whose main-chain O atom interacts with the ligand α-amino group, and Gln78, whose side-chain $O^{\varepsilon 1}$ atom forms a hydrogen bond and ionic interaction with the ligand α-amino group, in *Tt*GluBP are located in a highly conserved positions to Gly119, Asp157 and Gly68 in *Ec*GlnBP (Figs. 2*e*, 5*a* and 5*b*). In the *Ec*GlnBP complex, the Gln side chain is sandwiched between two aromatic residues, Phe13 and Phe50, in a manner similar to that observed in the *St*LAOBP and *St*HisJ complexes. The corresponding residues, Tyr14 and Phe52 in the *St*LAOBP complex and Tyr14 and Leu52 in the *St*HisJ complex, are in the same locations as those in *Ec*GlnBP, respectively. Although the Glu side chain in *Tt*GluBP is also involved in hydrophobic stacking with Tyr32 and Tyr186, the locations of these residues are different from those in the other three proteins (Fig. 2*e*).

Although the three periplasmic ligand-binding proteins use similar interactions to stabilize the α-ammonium and α-carboxyl groups of the ligand zwitterions, the ligand side-chain interactions are quite varied owing to the difference in length, shape and charge of the respective ligand side chains. Analysis of the ligand binding suggested that residues 11 and 52 play an important role in determining the ligand specificity of the *St*LAOBP and *St*HisJ complexes. In *St*LAOBP, Asp11 is involved in hydrogen bonding with the ligand side chain, while in *St*HisJ Asp11 does not interact with the ligand. Residue 52 differs in both size and charge between these two proteins and represents the only substitution found within the binding pocket. The corresponding residues (Asp10 and Phe50) in the *Ec*GlnBP–Gln complex are conserved (Fig. 2*e*) and involved in ligand interactions similar to those observed for Asp11 and

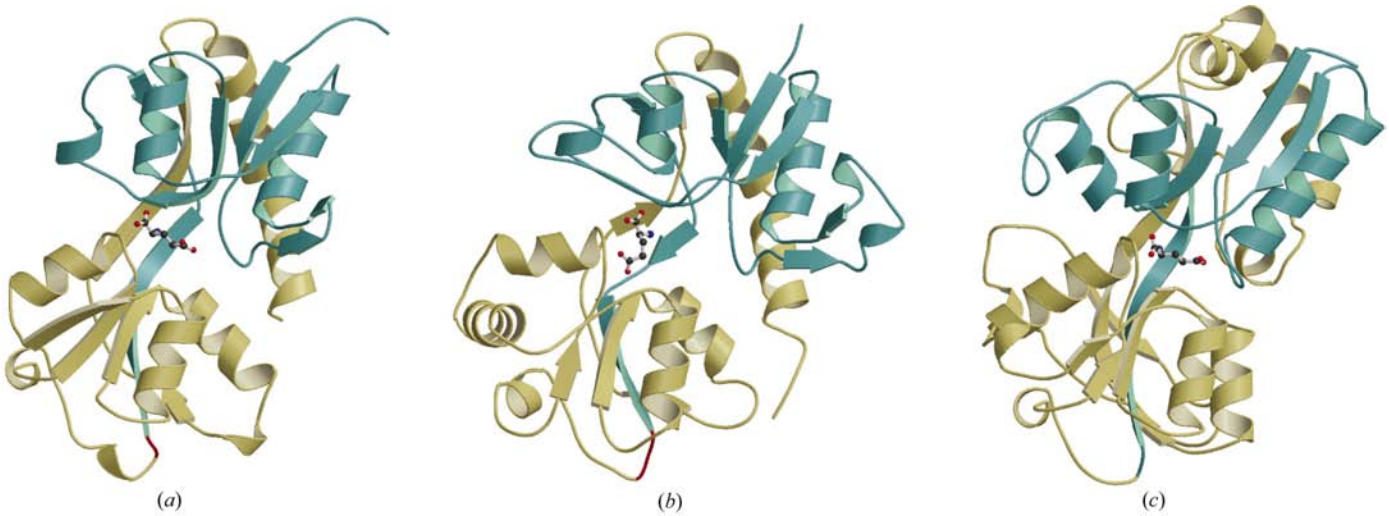

Figure 6
Structural homology of *Tt*GluBP with prokaryotic and eukaryotic glutamate receptor ligand-binding cores. (*a*) and (*b*) Ribbon diagrams of the crystal structures of the ligand-binding cores of GluR0 and GluR2 in complex with L-glutamate, coloured to show the contributions of segments S1 and S2 to domains I and II. The S1 and S2 segments that form a two-domain ligand-binding core are coloured green and yellow, respectively; the dipeptide linkers are shown in red. (*c*) Ribbon diagram for the *Tt*GluBP (this study) crystal structure complex with L-glutamate coloured to show the similar domain organization to that for the GluR0 and GluR2 ligand-binding cores.

Phe52 in the *St*LAOBP–Lys complex. Moreover, Lys115 and His156, which are unique to *Ec*GlnBP, interact with $O^{\varepsilon 1}$, which is unique to the neutral ligand side chain (Fig. 5*b*). In *Tt*GluBP, there are no corresponding residues to these residues in the other three protein complexes and Figs. 5(*a*) and 5(*b*) reveal that the orientation of the carboxyl group of the Glu (or possibly the amide group of Gln) side chain in *Tt*GluBP is different from that of the amide group of the Gln side chain in *Ec*GlnBP, although both ligands are in the similar extended conformation. In the *Tt*GluBP complex, one O atom of the ligand Glu side chain directly interacts with the main-chain N atoms of Val31 and Tyr32, while another O atom (or N atom) of the Glu (Gln) side chain only forms a water-mediated hydrogen bond (Fig. 5*a*).

3.6. Comparison with the ligand-binding cores of glutamate receptors

At present, structures of glutamate receptor-binding sites have been solved for three proteins: GluR0, a prokaryotic glutamate-receptor ion channel that shows structural homology to *Ec*GlnBP (Mayer *et al.*, 2001), GluR2, a eukaryotic AMPA selective ligand-gated ion channel which shows structural homology to *St*LAOBP (Armstrong & Gouaux, 2000), and mGluR1, a G-protein-coupled receptor which shows structural homology to *Ec*LIVBP (Kunishima *et al.*, 2000; Sack, Saper *et al.*, 1989).

The *Tt*GluBP resembles GluR0 and GluR2. GluR0 is a 397-amino-acid ion-channel protein containing a 19-residue signal peptide (Chen *et al.*, 1999). In glutamate-receptor ion channels the S1 and S2 peptide sequences that make up the two-domain ligand-binding core are interrupted by insertion of the first and second transmembrane ion-channel segments. A 233-residue GluR0 S1S2 construct, which included amino acids 44–140 from S1 and 256–385 from S2 linked by a GT peptide with a glutamate, has been reported (Mayer *et al.*, 2001). In GluR2, a 259-residue S1S2 construct, which included residues 95–506 from S1 and 632–776 from S2 linked by a GT peptide with a glutamate, has also been reported (Armstrong & Gouaux, 2000). Despite low amino-acid sequence identity between *Tt*GluBP and these two GluRs, domains I and II have similar secondary structures (Fig. 6).

As mentioned above, the binding interactions between proteins and the ligand α-carboxyl and α-amino groups are almost identical in the *Ec*GlnBP, *St*LAOBP and *St*HisJ complexes. Figs. 5, 7(*b*) and 7(*c*) show that in GluR0, GluR2 and *Ec*GlnBP complexes the bindings of ligand α-carboxyl and α-amino groups to proteins are quite similar. The amino-acid residues which interact with L-glutamate in GluR0 and GluR2 and with L-glutamine in *Ec*GlnBP occur in highly conserved locations; in particular, Arg residues which form hydrogen bonds and salt links with the α-carboxyl groups of ligands are highly conserved among glutamate receptors and periplasmic polar amino-acid-binding proteins. The binding of ligand α-carboxyl and α-amino groups to protein in *Tt*GluBP is different. The conserved Arg is not found in *Tt*GluBP. Thr143 and Ser60, which interact with ligand α-carboxyl group, and Glu111, which interacts with the ligand α-amino group, in *Tt*GluBP do not correspond to the amino acids that interact with ligand α-carboxyl and α-amino group in GluRs (Fig. 7).

In the *Ec*GlnBP complex with L-glutamine and in the GluR0 complex with L-glutamate, the ligands adopt a similar extended conformation and bind to residues near the end of

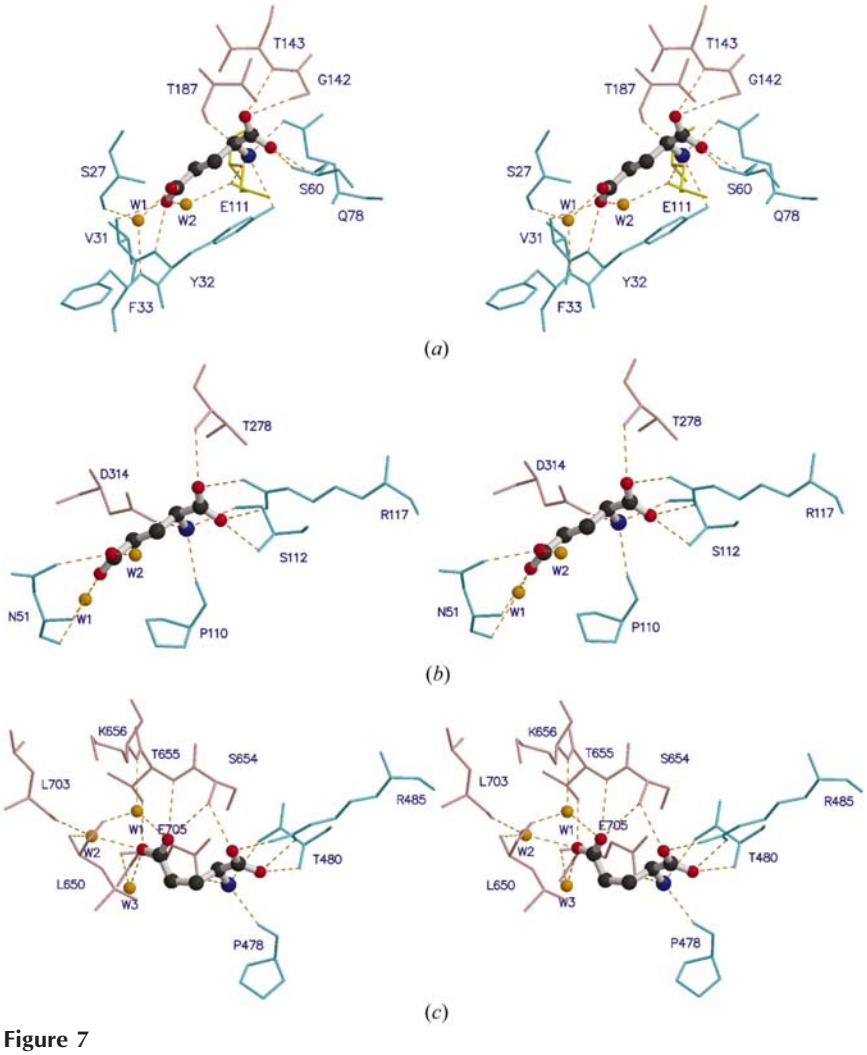

Figure 7
Stereoview of the ligand-binding sites of (*a*) the *Tt*GluBP complex with L-glutamate, (*b*) GluR0 complex with L-glutamate and (*c*) GluR2 complex with L-glutamate. As in Fig. 1, domain I and II residues are coloured cyan and pink, respectively, and a residue in a trans-domain strand of the *Tt*GluBP complex is colored yellow. Ligand, L-glutamate and L-glutamine are shown as ball-and-stick models, water molecules are shown as orange balls and hydrogen bonds are represented by orange dotted lines.

the first β-strand in domain I. In the *Tt*GluBP complex, the bound L-glutamate reveals a similar extended conformation to these and a water-mediated hydrogen bond to a residue, Ser27, near the end of the first β-strand in domain I. In contrast, in the GluR2 complex with L-glutamate the torsion angle of the ligand side chain has undergone a 105° rotation such that the γ-carboxyl group projects towards and interacts with the base of the helix in domain II corresponding to α5 in *Tt*GluBP (Fig. 7). Despite the similar conformations of L-glutamate and L-glutamine in the *Tt*GluBP, GluR0 and *Ec*GlnBP complexes, the binding mechanism by which these ligands interact with domains I and II are different. For *Ec*GlnBP, the ligand binding involves hydrogen bonds to residues in domain II, Lys115 and His156, which are absent in the *Tt*GluBP and the GluR0 complexes. Though the ligand-extended conformation in the *Tt*GluBP is most similar to that in GluR0, the residue which interacts with the ligand in the trans-domain strand, Glu111, is found only in the *Tt*GluBP (Fig. 7).

4. Conclusions

Crystallographic structural studies of the protein encoded by *T. thermophilus* HB8 gene TT1099 revealed a PSBP fold. The presence of a bound ligand, L-glutamate or L-glutamine, conservation of residues on the predicted membrane permease binding surface area (Oh *et al.*, 1993) and comparative structural analysis with other PSBP-fold proteins indicate that the *T. thermophilus* HB8 molecule is most likely an L-glutamate and/or L-glutamine-binding protein related to cluster 3 of periplasmic receptors (Tam & Saier, 1993). Although the exact function of *Tt*GluBP has not been confirmed, the reported structure with a unique ligand-binding geometry may provide some hints and ideas for better drug design

We are grateful to Shigeyuki Yokoyama and Seiki Kuramitsu for the plasmid vector, Masumi Maekawa for purification, and Naoko Takahashi for the amino-acid sequence verification of the proteins, and Yuki Nakamura and Mitsuaki Sugahara for the automated crystallization. This work was supported by the National Project on Protein Structural and Functional Analysis funded by MEXT of Japan (Project No. TT1099/HTPF00209).

References

Ames, G. F.-L. (1986). *Cell*, **47**, 323–324.
Ames, G. F.-L., Mimura, C. S., Holbrook, S. R. & Shyamala, V. (1992). *Adv. Enzymol. Relat. Areas Mol. Biol.* **65**, 1–47.
Ames, G. F.-L., Mimura, C. S. & Shyamala, V. (1990). *FEMS Microbiol. Rev.* **75**, 429–446.
Armstrong, N. & Gouaux, E. (2000). *Neuron*, **28**, 165–181.
Armstrong, N., Sun, Y., Chen, G. Q. & Gouaux, E. (1998). *Nature (London)*, **395**, 913–917.
Björkman, A. J. & Mowbray, S. L. (1998). *J. Mol. Biol.* **279**, 651–664.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Brünger, A. T., Kuriyan, J. & Karplus, M. (1987). *Science*, **235**, 458–460.
Chen, G. Q., Cui, C., Mayer, M. L. & Gouaux, E. (1999). *Nature (London)*, **402**, 817–821.
Chiu, J., DeSalle, R., Lam, H. M., Meisel, L. & Coruzzi, G. (1999). *Mol. Biol. Evol.* **16**, 826–838.
Dingledine, R., Borges, K., Bowie, D. & Taynelis, S. F. (1999). *Pharmacol. Rev.* **51**, 7–45.
Doyle, D. A., Cabral, J. M., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., Chait, B. T. & MacKinnon, R. (1998). *Science*, **280**, 69–77.
Felder, C. B., Graul, R. C., Lee, A. Y., Merkle, H.-P. & Sadee, W. (1999). *AAPS Pharm. Sci.* **1**, 1–20.
Feng, D. F. & Doolittle, R. F. (1987). *J. Mol. Evol.* **25**, 351–360.
Furlong, C. E. (1987). *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, edited by F. C. Neidhardt, pp. 768–796. Washington, DC: American Society of Microbiology.
Holm, L. & Sander, C. (1995). *Trends. Biochem. Sci.* **20**, 478–480.
Kunishima, N., Shimada, Y., Tsuji, Y., Sato, T., Yamamoto, M., Kumasaka, T., Nakanishi, S., Jingami, H. & Morikawa, K. (2000). *Nature (London)*, **407**, 971–977.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lee, B. & Richards, F. M. (1971). *J. Mol. Biol.* **55**, 379–400.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Mayer, M. L., Olson, R. & Gouaux, E. (2001). *J. Mol. Biol.* **311**, 815–836.
Mayer, M. L. & Westbrook, G. L. (1987). *Prog. Neurobiol.* **28**, 197–276.
Newcomer, M. E., Lewis, B. A. & Quiocho, F. A. (1981). *J. Biol. Chem.* **256**, 13218–13222.
Oh, B.-H., Kang, C.-H., Bondt, H. D., Kim, S.-H., Nikaido, K., Joshi, A. K. & Ames, G. F.-L. (1994). *J. Mol. Biol.* **269**, 4135–4143.
Oh, B.-H., Pandit, J., Kang, C.-H., Nikaido, K., Gokcen, S., Ames, G. F.-L. & Kim, S.-H. (1993). *J. Biol. Chem.* **268**, 11348–11355.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pang, A., Arinaminpathy, Y., Sansom, M. S. P. & Biggin, P. C. (2003). *FEBS Lett.* **550**, 168–174.
Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.
Quiocho, F. A. (1990). *Philos. Trans. R. Soc. London Ser. B*, **326**, 341–352.
Quiocho, F. A., Gilliland, G. L. & Phillips, G. N. Jr (1977). *J. Biol. Chem.* **252**, 5142–5149.
Rossmann, M. G. & Argos, P. (1975). *J. Biol. Chem.* **250**, 7525–7532.
Roussel, A. & Cambillau, C. (1996). *TURBO-FRODO Manual.* AFMB–CNRS, Marseille, France.
Sack, J. S., Saper, M. A. & Quiocho, F. A. (1989). *J. Mol. Biol.* **206**, 171–191.
Sack, J. S., Trakhanov, S. D., Tsigannik, I. H. & Quiocho, F. A. (1989). *J. Mol. Biol.* **206**, 193–207.
Schrempf, H., Schmidt, O., Kummerlen, R., Hinnah, S., Muller, D., Betzler, M., Steinkamp, T. & Wagner, R. (1995). *EMBO J.* **14**, 5170–5178.
Shilton, B. H., Flocco, M. M., Nilsson, M. & Mowbray, S. L. (1996). *J. Mol. Biol.* **264**, 350–363.
Sugahara, M. & Miyano, M. (2002). *Tanpakusitsu Kakusan Koso*, **47**, 1026–1032.
Sun, Y., Olson, R., Homing, M., Armstrong, N., Mayer, M. & Gouaux, E. (2002). *Nature (London)*, **417**, 245–253.
Sun, Y.-J., Rose, J., Wang, B.-C. & Hsiao, C.-D. (1998). *J. Mol. Biol.* **278**, 219–229.
Tam, R. & Saier, M. H. (1993). *Microbiol. Rev.* **57**, 320–346.
Treptow, N. A. & Shuman, H. A. (1985). *J. Bacteriol.* **163**, 654–660.
Yao, N., Trakhanov, S. & Quiocho, F. A. (1994). *Biochemistry*, **33**, 4769–4779.
Yokoyama, S., Hirota, H., Kigawa, T., Yabuki, T., Shirouzu, M., Terada, T., Ito, Y., Matsuo, Y., Kuroda, Y., Nishimura, Y., Kyogoku, Y., Mioki, K., Masui, R. & Kuramitsu, S. (2000). *Nature Struct. Biol.* **7**, 943–945.