Crystal structure of the dipeptide binding protein from *Escherichia coli* involved in active transport and chemotaxis

PETE DUNTEN AND SHERRY L. MOWBRAY

Department of Molecular Biology, BioMedical Center, Swedish University of Agricultural Sciences,
75124 Uppsala, Sweden

Abstract

The *Escherichia coli* periplasmic dipeptide binding protein functions in both peptide transport and taxis toward peptides. The structure of the dipeptide binding protein in complex with Gly-Leu (glycyl-L-leucine) has been determined at 3.2 Å resolution. The binding site for dipeptides is designed to recognize the ligand's backbone while providing space to accommodate a variety of side chains. Some repositioning of protein side chains lining the binding site must occur when the dipeptide's second residue is larger than leucine. The protein's fold is very similar to that of the *Salmonella typhimurium* oligopeptide binding protein, and a comparison of the structures reveals the structural basis for the dipeptide binding protein's preference for shorter peptides.

Keywords: chemotaxis; dipeptide; periplasmic binding protein; structure; transport

The dipeptide permease transports dipeptides across the *Escherichia coli* cytoplasmic membrane with high affinity, allowing a wide variety of dipeptides to enter (Payne & Bell, 1979; Perry & Gilvarg, 1984). Dipeptides as structurally and chemically diverse as glycyl-glycine, lysyl-lysine, and phenylalanyl-phenylalanine are taken up at about the same rates (Payne & Bell, 1979). A 57-kDa periplasmic binding protein, the product of the *dppA* gene, acts as the initial receptor for dipeptides during their uptake through this system. The dipeptide binding protein also plays a role in chemotaxis toward dipeptides, acting in concert with the inner membrane-bound Tap protein to initiate the chemotactic swimming response to dipeptides (Manson et al., 1986).

A number of binding protein structures have been solved, and some common themes have emerged (reviewed in Quiocho, 1990). The ligand binding site is invariably located between two domains that are joined by two or three connecting strands functioning as a hinge. The bound ligand is typically buried by the protein and inaccessible to solvent. This form of the protein is termed the "closed" form. In the "open" form of the protein, relative rotation of the two domains makes the binding site accessible to solvent, allowing entry and exit of ligand. Residues

from the two domains on the face of the protein opposite the hinge are brought close together by the closing of the protein, and only in the closed form do they have the proper positioning for interaction with the membrane components of the transport machinery (and the membrane-bound chemotaxis transducers, for those binding proteins with roles in both transport and taxis).

The dipeptide binding protein, with 507 residues in the mature form, is larger than most periplasmic binding proteins. It is a member of a family of binding proteins responsible for the uptake of peptides, heme, and ions such as nickel (Hanson et al., 1992; Tam & Saier, 1993). Only one structure of this family has been reported to date, that of the oligopeptide binding protein from *Salmonella typhimurium* (Tame et al., 1994). This structure was the first example of a periplasmic binding protein constructed from three domains rather than the usual two.

The di- and oligopeptide binding proteins differ in several respects. The dipeptide binding protein binds dipeptides and some tripeptides, whereas the oligopeptide binding protein can bind di-, tri-, tetra-, and pentapeptides. The dipeptide binding protein is involved in taxis toward dipeptides, whereas the oligopeptide binding protein plays a role in recycling of cell wall peptides, which precludes any involvement in sensing peptide gradients (Goodell & Higgins, 1987). We discuss here the 3.2-Å structure of the dipeptide binding protein in complex with the dipeptide Gly-Leu (glycyl-L-leucine) and consider some of the implications of the structure for the protein's function.

Results and discussion

Description of the structure

The overall dimensions of the dipeptide binding protein are 70 × 30 × 55 Å. The Cα trace of the molecule appears in Figure 1 and Kinemage 1, and a representative view of the experimental electron-density map is shown in Figure 2. The protein is organized in three structural domains, all members of the α/β class of folds. Domain I includes residues 1–33, 183–260, and 479–507; domain II includes residues 34–181; and domain III includes residues 261–478. The overall fold of the dipeptide binding protein is similar to that of the oligopeptide binding protein, as discussed below. Most other periplasmic binding proteins are constructed from only two domains corresponding to domains I and III of the dipeptide binding protein. The topology of these domains is most similar to that of the group II bidomain binding proteins, which are responsible for the uptake of sulfate, phosphate, maltodextrins, lysine/arginine/ornithine, and histidine (Spurlino et al., 1991; Louie, 1993).

Figure 3 shows the topology of the β-sheets in domains I–III, and Table 1 introduces the nomenclature used to describe the secondary structure elements in each of the domains (Kinemage 1). Domain I forms a seven-stranded, twisted β-sheet. The β-sheet in this domain is similar to the five-stranded, doubly wound β-sheet in the N-terminal domain of the group II bidomain binding proteins, with a β-hairpin from residues 185–200 contributing two additional strands to the sheet. As in the bidomain binding proteins, there are two β/α units that partially cover one face of the sheet with α-helices. The sheet is of mixed topology, with five parallel strands. The C-terminal ends of these strands are located at the interface with domain III. In the bidomain binding proteins, the ligand binding site is formed in part by the C-terminal ends of these strands. In the dipeptide binding protein, the loop following strand β1-I contributes residues 20–22 to the binding site.

Domain III is also built around a twisted, doubly wound β-sheet. The sheet is of mixed topology, with the C-terminal ends of four of the six strands located at the interface with domain I. The sense of winding of the β/α units changes between strands

β1-III and β3-III of the sheet, and the loops connecting the C-terminal ends of these strands to the following α-helices contribute residues to the binding site. Arg 355 in the loop following strand β1-III forms an ion pair with the ligand's carboxy terminus, and Asp 408 in the loop following strand β3-III interacts with the ligand's amino terminus (Fig. 4). The last hydrogen bond between strands β1-III and β3-III is between the carbonyl oxygen of residue 349 in strand β1-III and the peptidyl nitrogen of residue 402 in strand β3-III. Strand β3-III continues to residue 406; this extension and the dipeptide ligand are antiparallel with respect to one another and form two hydrogen bonds of the type present in antiparallel β-sheets. The interface between domain III and domains I/II is extensive, totaling 1,400 Å², or roughly 15% of domain III's solvent-accessible surface area.

Domain II has no equivalent in the bidomain binding proteins. The endpoints of domain II were chosen based on the observation that residues 33 and 183 of domain I are close enough in space (4.2 Å Cα to Cα distance) to consider insertion or deletion of the sequence encoding domain II as a possible event relating the dipeptide binding protein to the bidomain binding proteins. Domain II includes a four-stranded antiparallel β-sheet with α-helices packed against its hydrophobic face and a hydrophilic face exposed to solvent. It forms interfaces of largely hydrophilic character with both domains I and III. Domain II contributes only two residues to the binding site: the hydroxyl group of Tyr 114 provides a hydrogen bond to the amino group of the ligand, and the side chain of Asp 153 is located at the end of a pocket that accepts the side chain of the first residue of the dipeptide ligand.

Tam and Saier (1993) aligned sequences of seven binding proteins specific for peptides and nickel. The sequences are most similar in domain II, and Tam and Saier identified a signature sequence specific for this family. The five invariant residues in the signature sequence are Ala 51, Asp 59, Thr 64, Arg 68, and Lys 72 (Kinemage 1). Ala 51 is located two residues before the beginning of the first strand of the β-sheet in domain II, and its side chain forms part of the hydrophobic core of the domain. Packing of the core requires a small hydrophobic residue at this position. Arg 68 is located at the C-terminal end of strand β2-II,
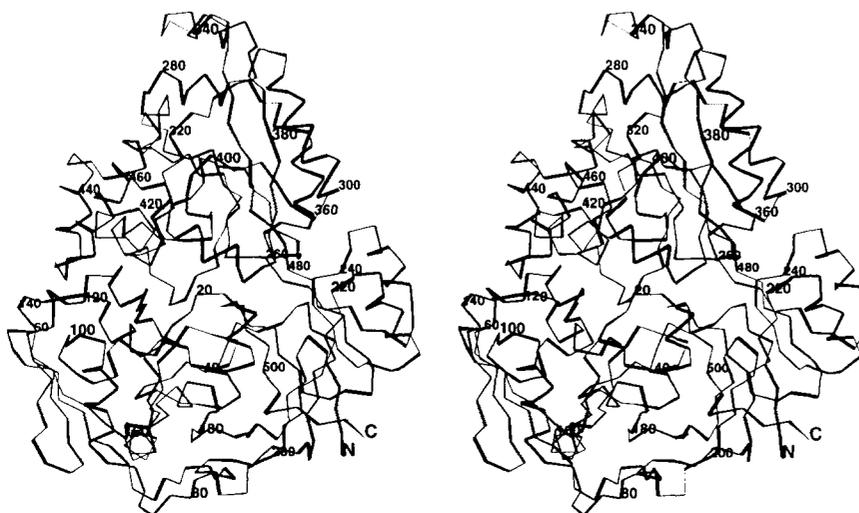
Fig. 1. Cα trace of the dipeptide binding protein. Domain I is at the bottom right, domain II is at the bottom left, and domain III is at the top of the figure. To simplify following the path of the polypeptide chain, disulfide bridges have not been shown. Disulfide bridges join Cys 6 with Cys 234 and Cys 422 with Cys 435.

A

B
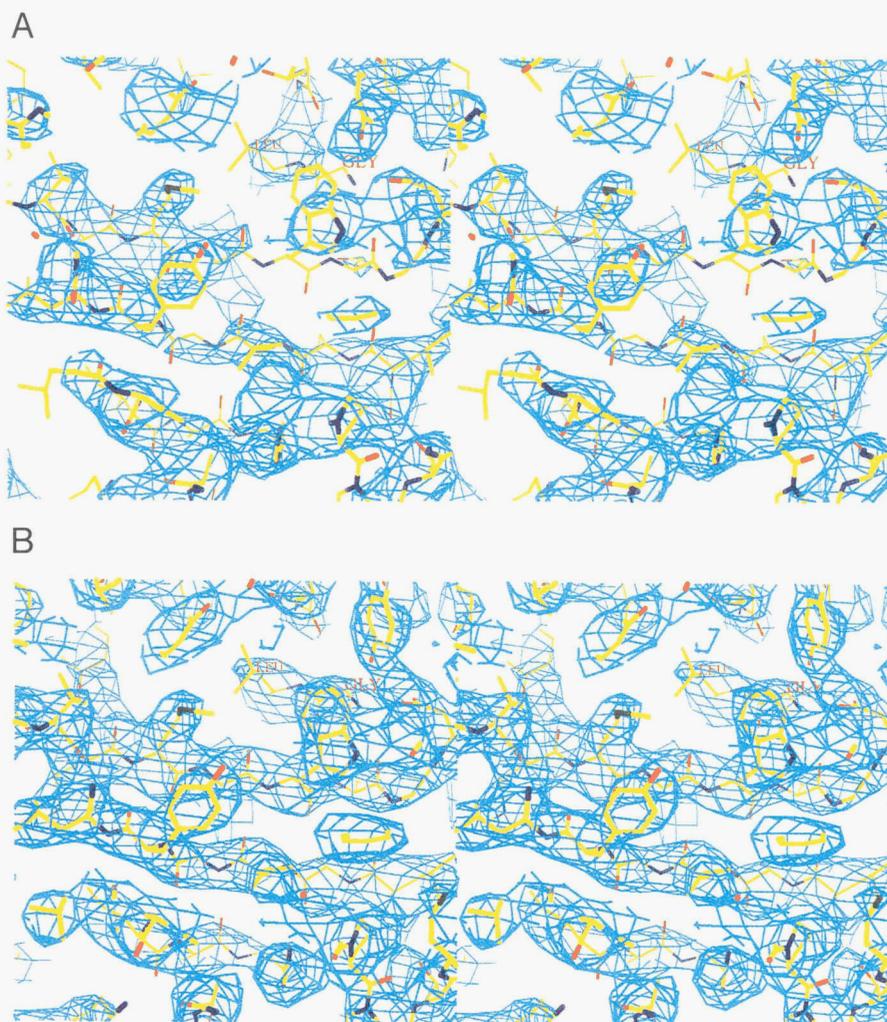
Fig. 2. Part of the β-sheet in domain III and the dipeptide ligand are shown together with the electron-density contoured at 1.2 times the sigma level. **A:** The electron-density map calculated with multiple-isomorphous replacement phases to 3.5 Å resolution. **B:** The map calculated with phases improved by density modification techniques, including fourfold averaging.

and its side chain hydrogen bonds to the main chain O of Gly 49 in the turn preceding the start of strand β1-II. The side chain stacks favorably with the side chain of Trp 205 from domain I. The corresponding Arg in the oligopeptide binding protein (Arg 77) makes similar interactions. The other three invariant residues in the signature sequence are all exposed at the surface of domain II, and there appear to be no structural requirements for a particular residue at these positions.

The dipeptide binding site

The peptide-binding site is located between domains I and III (Kinemage 2). These domains are connected by two strands that presumably function as a hinge (residues 260–262 and 478–480). In the crystal structure, the protein is "closed," with the dipeptide ligand completely buried. Figure 4 shows some of the interactions the protein makes with the ligand Gly-Leu. Specific interactions are made with the ligand's backbone, providing hydrogen bonds to the carbonyl O and peptidyl N of the peptide linkage, as well as both hydrogen bonds and salt linkages to the charged termini. Transport studies have shown that any changes to the ligand's terminal amino or carboxy groups that affect their charge are incompatible with binding (Payne & Gilvarg, 1968).

The binding site is clearly designed to select ligands with charged termini, because both Asp 408 and Arg 355 prefer oppositely charged partners. The spacing of these groups favors binding of dipeptides as opposed to free amino acids. The ability of the dipeptide binding protein to accept some tripeptides (Gly-Gly-Ile, for example [Manson et al., 1986]) suggests that Arg 355 is free to reposition its side chain to form a favorable interaction with the carboxy terminus of tripeptide ligands.

Two pockets opening off the binding site are positioned to accept the side chains of the dipeptide ligand, as shown in Figure 5. Both pockets can be visualized by calculating the solvent-accessible surface of the protein (Voorintholt et al., 1989). A large, empty pocket is present near the Cα of the dipeptide's first residue (a Gly residue in the crystal structure). This pocket appears to be large enough to accept any of the 20 naturally occurring side chains, and is probably occupied by water in the complex with Gly-Leu. The walls of the pocket are formed by residues from all three domains: Gly 22, Asp 153, Pro 356, Tyr 357, Thr 406, Leu 496, Gly 497, and Lys 498. The Leu side chain of the dipeptide Gly-Leu fits into a second, smaller pocket lined by the side chains of Thr 20 and Ser 21 (from domain I), Trp 386, Tyr 389, Leu 390, Met 403, Trp 405, Ser 429, and Tyr 431 (all from domain III). The main-chain N and O of

A

B

C
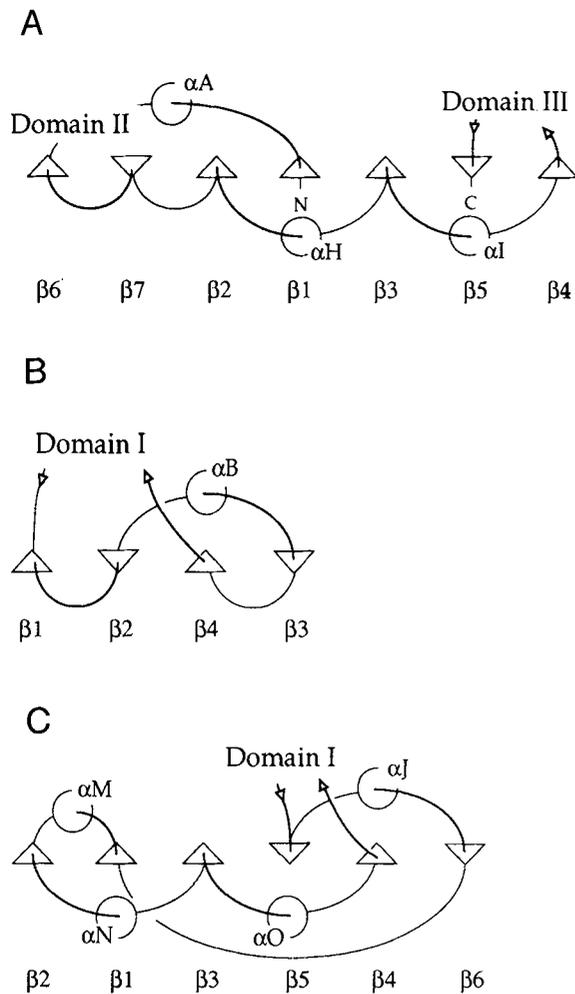
Fig. 3. Simplified topology of the β-sheets in domains I–III. **A:** Domain I. **B:** Domain II. **C:** Domain III. Strand numbering in domains I and III follows the convention established for describing the bidomain binding proteins. Only α-helices that can be considered part of a $\beta\alpha\beta$ motif are shown. β-Strands are diagrammed as triangles and α-helices as circles. Strands and helices run perpendicular to the plane of the paper, toward the viewer, in this two-dimensional simplification of the protein fold. For strands parallel to one another, triangles point in the same direction; for antiparallel strands, triangles point in opposite directions. Thick lines indicate connections at the top of the sheet, and thin lines indicate connections at the bottom.

Table 1. *Secondary structure*[a]

Strand	Residues	Helix	Residues
β1-I	2–7	A-I	22–32[b]
	37–39		
	46–48		
β1-II	51–56		
β2-II	62–67	B-II	87–98
		C-II	113–117
		D-II	120–123
β3-II	124–131		
β4-II	134–139	E-II	146–150
		F-II	160–168
		G-II	173–177
β6-I	185–191		
β7-I	195–200		
β2-I	213–218	H-I	222–230
β3-I	235–238[c]	I-I	242–250
β4-I	254–259		
β5-III	263–269	J-III	279–287
		K-III	291–297
β6-III	303–305	L-III	328–337
β1-III	344–349	M-III	361–373
β2-III	377–382	N-III	386–394
β3-III	400–406	O-III	413–420
		P-III	422–427
		Q-III	437–448[d]
		R-III	452–469
β4-III	472–478		
β5-I	480–485		

[a] Secondary structural elements were defined by the program DSSP (Kabsch & Sander, 1983), except as noted. The domain to which each belongs is indicated.

[b] Distorted at residue 28, due to the presence of Pro 31; there is an insertion near residue 29, relative to OppA.

[c] This strand is visually part of the sheet, but is slightly distorted and missing hydrogen bonds, due to Pro 238.

[d] Distorted at residue Pro 445.

Gly 404 also form part of the pocket's wall. The side chain of Met 403, located at the end of the pocket, appears free to move in order to accommodate ligands with larger side chains. The availability of crystals of the protein complexed to different dipeptides (Dunten et al., 1993) should help settle this question.
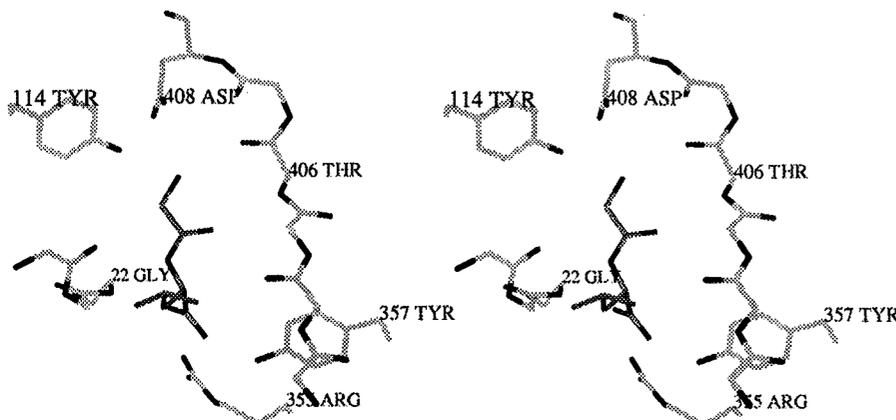
Fig. 4. Bound dipeptide Gly-Leu and residues involved in binding its backbone are shown. The β-strand (residues 400–406) and the dipeptide are antiparallel with respect to one another. Side chains of residues 20–21 and 403–406 have been omitted for clarity.
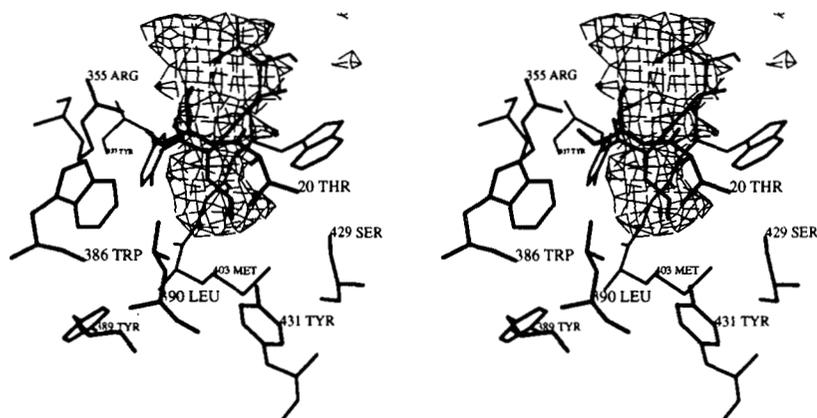
Fig. 5. Residues involved in forming the pocket that accepts the side chain of the dipeptide's second residue. The contour surface is the solvent-accessible surface calculated from the protein coordinates without dipeptide. The empty pocket that accepts the side chain of the dipeptide's first residue is visible at the top. The pocket that accepts the Leu side chain of the dipeptide Gly-Leu is at the bottom.

Similarity to the oligopeptide binding protein from S. typhimurium

The sequences of the dipeptide and oligopeptide binding proteins are 23% identical, with the oligopeptide binding protein having 12 additional residues at the N-terminus and 2 additional residues at the C-terminus. The two proteins can be aligned with an RMS deviation (RMSD) of 1.85 Å for 444 structurally equivalent Cα atoms. Each of the domains can be aligned separately, which results in a better fit for the aligned partners, while leaving the unaligned domains offset with respect to one another. If the dipeptide binding protein is first broken into three domains and then these are aligned with the oligopeptide binding protein, the RMSD is 1.55 Å for 441 Cα atoms. Thus, the do-

mains of the two proteins are oriented in a slightly different fashion.

Alignment of domain III reveals an interesting feature of the dipeptide binding protein that is tied to its preference for peptides shorter than those bound by the oligopeptide binding protein. The loop connecting strand $\beta1$-III to the following α-helix is five residues longer in the dipeptide binding protein than in the oligopeptide binding protein (Fig. 6 and Kinemage 2). In the dipeptide binding protein, the loop reaches into the binding site and positions the side chain of Arg 355 in proximity to the carboxy terminus of the bound dipeptide. In the oligopeptide binding protein, the connection between strand $\beta1$-III and the following α-helix is shorter, allowing for a larger binding site. The side chain of Arg 413 is positioned to interact with the car-
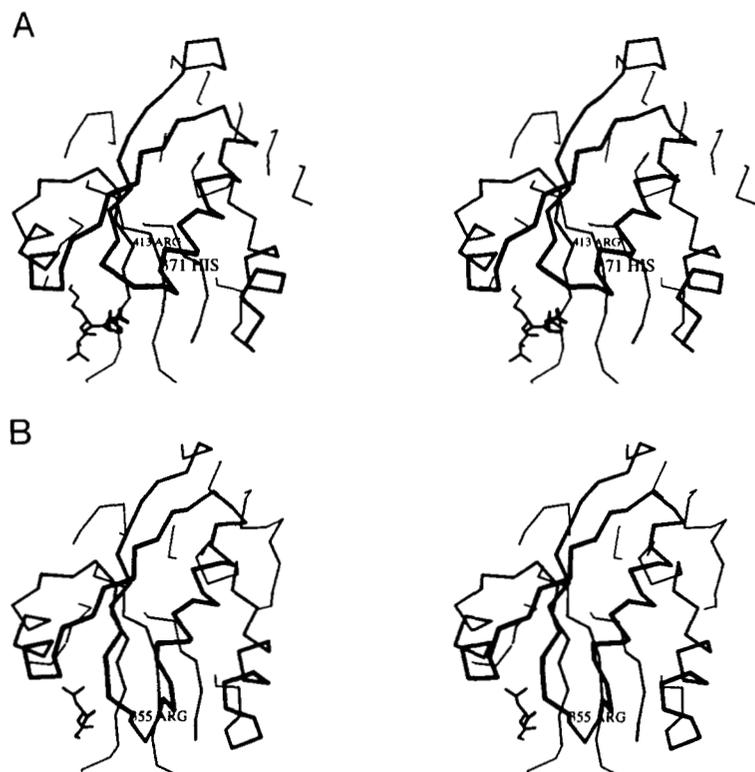
Fig. 6. Cα trace of domain III from the oligopeptide binding protein (**A**, pdb code 1ola, with ligand Val-Lys-Pro-Gly) and the dipeptide binding protein (**B**). Bound peptides are shown as stick models, with their carboxy terminus closest to the viewer when viewed in stereo. The view is clipped at the back for clarity.

boxy terminus of tripeptides, and the side chain of His 371 interacts with the carboxy terminus of the tetrapeptide present in the oligopeptide binding protein crystal structure. The corresponding residues in the dipeptide binding protein are Met 402 and Ala 361.

Domain II plays a similar role in peptide binding in both proteins. As was found for the dipeptide binding protein, domain II of the oligopeptide binding protein contributes residues to the lining of the pocket which accepts the side chain of the oligopeptide's first residue (Val 160, His 161, Pro 162). The side chain of Tyr 109 hydrogen bonds the oligopeptide's N-terminal amino group, filling the same role as Tyr 114 in the dipeptide binding protein.

Dipeptide chemotaxis

The dipeptide binding protein plays a role in both peptide transport and taxis toward peptides. The binding protein with bound dipeptide forms a complex with the inner membrane-bound Tap chemotactic transducer, initiating the chemotactic swimming response. In order to discriminate between peptide-bound and peptide-free binding protein, it is likely the Tap transducer recognizes residues from both domains I and III, given that these domains must be in different orientations with respect to one another in the open and closed forms of the protein. Mutations in the dipeptide protein affecting chemotaxis have not yet been characterized. Mutations affecting chemotaxis toward maltose have been isolated in the maltose binding protein (Zhang et al., 1992), and these mutations identify two areas on the surface of the protein recognized by the Tar chemotactic transducer (Tar and Tap belong to a family of structurally related chemotaxis transducers). We can predict which residues on the surface of the dipeptide binding protein are likely to be involved in the interaction with Tap by comparing the structure with that of the maltose binding protein (Spurlino et al., 1991). The secondary structure elements β1-I, β2-I, β3-I, and α-helix H (all in domain I) are oriented almost identically in the two proteins. Mutations in the maltose binding protein's first domain, which affect chemotaxis but not transport, are located in the equivalent to α-helix H and the turn leading into strand β3-I. These are residues 45–55 of the maltose binding protein, corresponding to residues 223–233 in the dipeptide binding protein. The similarity of the two proteins is not as great in domain III (the C-terminal domain of the maltose binding protein), and we cannot predict the location of residues likely to be involved in the interaction with Tap with as much confidence. α-Helices K and M are appropriately located on the surface of domain III, and may be involved in the protein–protein interaction with Tap.

Future perspectives

The dipeptide binding protein's role in peptide taxis can be exploited in future studies of peptide binding. For example, the effects of mutating binding site residues on binding a variety of peptides can be quickly assayed without the need to purify each mutant protein. The combination of genetics and now structure paves the way for a fuller understanding of how a wide variety of peptides are bound with high affinity. The binding site is also an attractive target for structure-based drug design, because antibiotics attached to a dipeptide backbone are known to enter the cell via the dipeptide permease (Hammond et al., 1987). Designing antibiotics for uptake by the dipeptide permease has the ad-

vantage that the bacteria should respond to the antibiotic as if it were an attractant. And finally, knowledge of the structure can be applied to modeling studies (Dunten & Mowbray, 1995 [companion paper]) of related proteins in order to understand how a binding site can be tailored to handle ligands as diverse as peptides, heme, and the nickel ion.

Methods

Data collection

The crystallization conditions have been described (Dunten et al., 1993). Briefly, equal volumes of protein and reservoir solution were equilibrated against reservoir solution in hanging-drop vapor phase diffusion set-ups. The protein concentration was 15 mg/mL, the dipeptide concentration was 1 mM, and the reservoir contained 1.49 M sodium citrate, pH 6.2, 1% ethanol. The protein crystallizes in space group P6$_1$ with $a = b = 182.57$ Å, $c = 211.88$ Å. Heavy-atom soaks were done in 1.5 M sodium citrate, pH 6.2, with saturating concentrations of heavy-atom reagents for 2–4 days. Native data and the Hg derivative data were collected at Daresbury Synchrotron Radiation Source, station 9.6, on the R-Axis image plate and 30-cm MAR image plate, respectively (Table 2). The wavelength of the synchrotron radiation was 0.89 Å. Reflections were integrated with DENZO (Otwinowski, 1993) and merged using the CCP4 suite of programs (Collaborative Computational Project, Number 4, 1994). The Pt derivative data were collected on an R-Axis image plate with a rotating anode generator providing CuKα radiation, and processed with the R-Axis software. The low-resolution Hg data were collected on a Xentronics area detector using CuKα radiation and processed with XDS (Kabsch, 1988).

Phasing

Consideration of the Matthews coefficient (Matthews, 1968) and the 622 symmetry of the diffraction pattern at low-resolution suggested there were four or six molecules in the asymmetric unit. A four-site solution to the Hg difference Patterson was found with RSPS (Knight et al., 1990) using data to 5 Å. Additional sites with lower occupancy were picked from difference Fourier maps. Inclusion of the anomalous Hg signal from the synchrotron data in the phasing calculations using MLPHARE (Otwinowski, 1991) indicated the space group was P6$_1$ and not P6$_5$. Ncs (noncrystallographic symmetry) operators were deduced by inspection of the heavy-atom sites at a graphics display using O (Jones et al., 1991). The initial ncs operators were determined by superpositioning three heavy-atom sites (later identified as His 105, Met 257, and His 467) from each molecule. A solvent-flattened map was calculated based on the assumption that the asymmetric unit contained four protein molecules and therefore a solvent content of 69%. This map was skeletonized to produce "bones," and a rough mask was created around the bones atoms of one molecule using MAMA (Kleywegt & Jones, 1993). At this point, it was clear that the asymmetric unit contained only four molecules. The mask and ncs operators were subsequently used to improve the intial MIR phases via electron-density averaging. Solvent flattening, histogram matching, and averaging were applied in MAGICSQUASH (Cowtan & Main, 1993; Schuller et al., 1995) and DM (Cowtan, 1994) starting with data at 5 Å and extending to 3.2 Å. The map calculated using

Table 2. *Data collection statistics*

Data set (source)	Resolution (Å)	No. observations (no. crystals)	No. unique	Completeness (%)	$\langle I/\sigma \rangle$ (3.28–3.20 Å)	$R_{sym}{}^a$ (%)	$R_{iso}{}^b$ (%)	Number of sites	Phasing power[c]
Native (synchrotron)	3.2	170,663 (2)	57,483	88	2.6[d]	10.5	–	–	–
CH_3HgCl (synchrotron)	3.5	119,674 (4)	46,223	92	–	8.5	14.4	8	1.2
CH_3HgCl (home)	5.0	32,274 (1)	17,177	97	–	13.1	18.1	9	1.3
cis-$Pt(NH_3)_2Cl_2$ (home)	4.0	55,740 (1)	29,541	87	–	10.9	20.7	19	1.3
Overall figure of merit	0.422 (3.5 Å)								

[a] $R_{sym} = \Sigma |I_j - \langle I_j \rangle| / \Sigma I_j$, where I_j is the intensity of an observation of reflection j and the sum is over reflections observed more than once.

[b] $R_{iso} = \Sigma \| F_{PH}| - |F_P \| / \Sigma |F_P|$, where $|F_P|$ is the native structure factor amplitude and $|F_{PH}|$ is the heavy-atom derivative structure factor amplitude.

[c] Phasing power = $RMS(|F_H|/E)$ for acentric reflections, where $|F_H|$ is the heavy atom structure factor amplitude and E is the residual lack of closure.

[d] The native data are 89% complete in this resolution shell.

these improved phases was of sufficient quality to trace the entire polypeptide chain.

Model building/refinement

The amino-acid sequence of the protein was taken from Abouhamad et al. (1991) and Olson et al. (1991). An initial model consisting of residues 1–507 was built from an edited "bones" skeleton using the baton commands in O. Only the loop from residues 74–79 is not well defined by the density. The initial model placed side chains in a rotamer conformation, as tabulated in Ponder and Richards (1987). The four independent molecules in the asymmetric unit were constrained to be identical during refinement in X-PLOR (Brünger et al., 1987) in order to maintain a favorable observation to parameter ratio (Table 3). The initial R-factor for the model was 0.456 (R-free 0.455) for data in the 12–3.2-Å range. After rigid-body refinement and positional refinement to relieve bad contacts, the R-factor was 0.322 (R-free 0.332). The model was examined in the regions where PROCHECK (Laskowski et al., 1993) indicated Ramachandran violations and in the regions where O reported high pep_flip values, and adjusted where needed. Adjustments to the model, simulated annealing refinement, and further positional refinement lowered the R-factor to 0.265 (R-free 0.285). A solvent correction was applied in X-PLOR using data between 25 and 3.2 Å, producing better agreement between observed and calculated structure factor amplitudes for reflections in the 12–6-Å range. The solvent correction was applied throughout

the remainder of the refinement, which made use of all data between 25 and 3.2 Å. The ligand was fit to the electron density in the binding site, and then group B factors were introduced. The R-free indicated that grouped B factors consisting of main-chain N, C, O, $C\alpha$, $C\beta$ in one group, and all remaining side-chain atoms in a second group, were preferable to the default grouping in X-PLOR with N, C, $C\alpha$ as main chain and all others as side chain. Regions of the molecule with above-average B factors (residues 75–83 and 168–171) are those that are not involved in any crystal contacts. Note that this B factor model (with strict ncs applied) effectively averages across all four molecules in the asymmetric unit. In some cases, side chains on the surface of the protein are involved in crystal contacts in one or more molecules, but not all four. In these instances, the B factor for the side chain atoms includes contributions from both the well-defined side chains involved in crystal contacts and the side chains of the molecules for which there are no crystal packing effects. The final $2F_o - F_c$ maps indicate breakdown of the ncs due to crystal packing for the side chains of Met 259 and Arg 362: we have modeled these in a single conformation. The N-terminal Lys is visible in only one of the four molecules, where it is held in place by crystal packing. The current model has no Ramachandran violations. At this stage of refinement, no water molecules have been added to the model. The coordinates have been submitted to the Brookhaven Protein Data Bank and will be available as entry 1dpp.

Table 3. *Quality of the model*

No. non-hydrogen atoms per monomer	4,061
No. reflections total	57,347
Resolution	25–3.2 Å
R-factor	0.223
No. reflections for R-free	2,915
R-free	0.240
RMSD bond distances	0.008 Å
RMSD bond angles	1.4°
Average B factor	21.5 Å²

References

Abouhamad WN, Manson M, Gibson MM, Higgins CF. 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.

Brünger AT, Kuriyan J, Karplus M. 1987. Crystallographic R factor refinement by molecular dynamics. *Science 235*:458–460.

Collaborative Computational Project, Number 4. 1994. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D 50*:760–763.

Cowtan KD. 1994. "dm": An automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM Newsl Protein Crystallogr 31*:34-38.

Cowtan KD, Main P. 1993. Improvement of macromolecular electron-density maps by the simultaneous application of real and reciprocal space constraints. *Acta Crystallogr D 49*:148-157.

Dunten PW, Harris JH, Feiz V, Mowbray SL. 1993. Crystallization and preliminary X-ray analysis of the periplasmic dipeptide binding protein from *Escherichia coli. J Mol Biol 231*:145-147.

Goodell EW, Higgins CF. 1987. Uptake of cell wall peptides by *Salmonella typhimurium* and *Escherichia coli. J Bacteriol 169*:3861-3865.

Hammond SM, Claesson A, Jansson AM, Larsson LG, Pring BG, Town CM, Ekström B. 1987. A new class of synthetic antibacterials acting on lipopolysaccharide biosynthesis. *Nature 327*:730-732.

Hanson MS, Slaughter C, Hanson EJ. 1992. The hbpA gene of *Haemophilus influenzae* type b encodes a heme-binding lipoprotein conserved among heme-dependent *Haemophilus* species. *Infect Immunol 60*:2257-2266.

Jones TA, Zou JY, Cowan SW, Kjeldgaard M. 1991. Improved methods for the building of protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A 47*:110-119.

Kabsch W. 1988. Evaluation of single crystal X-ray diffraction data from a position sensitive detector. *J Appl Crystallogr 21*:916-924.

Kabsch W, Sander S. 1983. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers 22*:2577-2637.

Kleywegt GJ, Jones TA. 1993. Masks made easy. *Joint CCP4 and ESF-EACBM Newsl Protein Crystallogr 28*:56-59.

Knight S, Andersson I, Brändén CI. 1990. Crystallographic analysis of ribulose 1,5-bisphosphate carboxylase from spinach at 2.4 Å resolution. *J Mol Biol 215*:113-160.

Laskowski R, MacArthur M, Moss D, Thornton J. 1993. Procheck: A program to check the stereochemical quality of protein structures. *J Appl Crystallogr 26*:91-97.

Louie GV. 1993. Porphobilinogen deaminase and its structural similarity to the bidomain binding proteins. *Curr Opin Struct Biol 3*:401-408.

Manson MD, Blank V, Brade G, Higgins CF. 1986. Peptide chemotaxis in *E. coli* involves the Tap signal transducer and the dipeptide permease. *Nature 321*:253-256.

Matthews BW. 1968. Solvent content of protein crystals. *J Mol Biol 33*:491-497.

Olson ER, Dunyak DS, Jurss LM, Poorman RA. 1991. Identification and characterization of dppA, an *Escherichia coli* gene encoding a periplasmic dipeptide transport protein. *J Bacteriol 173*:234-244.

Otwinowski Z. 1991. Maximum likelihood refinement of heavy atom parameters. In: Wolf W, Evans PR, Leslie AGW, compilers. *Proceedings of the CCP4 study weekend: "Isomorphous replacement and anomalous scattering."* Daresbury, UK: SERC Daresbury Laboratory. pp 80-86.

Otwinowski Z. 1993. Oscillation data reduction program. In: Sawyer L, Isaacs N, Bailey S, compilers. *Proceedings of the CCP4 study weekend: "Data collection and processing."* Daresbury, UK: SERC Daresbury Laboratory. pp 56-62.

Payne JW, Bell G. 1979. Direct determination of the properties of peptide transport systems in *Escherichia coli*, using a fluorescent-labeling procedure. *J Bacteriol 137*:447-455.

Payne JW, Gilvarg C. 1968. The role of the terminal carboxyl group in peptide transport in *Escherichia coli. J Biol Chem 243*:335-340.

Perry D, Gilvarg C. 1984. Spectrophotometric determination of affinities of peptides for their transport systems in *Escherichia coli. J Bacteriol 160*:943-948.

Ponder JW, Richards FM. 1987. Tertiary templates for proteins: Use of packing criteria in the enumeration of allowed sequences for different structural classes. *J Mol Biol 193*:775-791.

Quiocho FA. 1990. Atomic structures of periplasmic binding proteins and the high-affinity active transport systems in bacteria. *Philos Trans R Soc Lond B 326*:341-351.

Schuller DS, Grant GA, Banaszak LJ. 1995. The allosteric ligand site in the Vmax-type cooperative enzyme phosphoglycerate dehydrogenase. *Nature Struct Biol 2*:69-74.

Spurlino JC, Lu GY, Quiocho FA. 1991. The 2.3-Å resolution structure of the maltose- or maltodextrin-binding protein, a primary receptor of bacterial active transport and chemotaxis. *J Biol Chem 266*:5202-5219.

Tam R, Saier MH Jr. 1993. Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. *Microbiol Rev 57*:320-346.

Tame JRH, Murshudov GN, Dodson EJ, Neil TK, Dodson GG, Higgins CF, Wilkinson AJ. 1994. The structural basis of sequence-independent peptide binding by OppA protein. *Science 264*:1578-1581.

Voorintholt R, Kosters MT, Vegter G, Vriend G, Hol WGJ. 1989. A very fast program for visualizing protein surfaces, channels and cavities. *J Mol Graphics 7*:243-245.

Zhang Y, Conway C, Rosato M, Suh Y, Manson MD. 1992. Maltose chemotaxis involves residues in the N-terminal and C-terminal domains on the same face of maltose-binding protein. *J Biol Chem 267*:22813-22820.