The crystal structures of the oligopeptide-binding protein OppA complexed with tripeptide and tetrapeptide ligands

Jeremy RH Tame[1]*, Eleanor J Dodson[1], Garib Murshudov[1], Christopher F Higgins[2] and Anthony J Wilkinson[1]

[1]Department of Chemistry, University of York, York YO1 5DD, UK and [2]Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, UK

Background: The periplasmic oligopeptide-binding protein OppA has a remarkably broad substrate specificity, binding peptides of two to five amino-acid residues with high affinity, but little regard to sequence. It is therefore an ideal system for studying how different chemical groups can be accommodated in a protein interior. The ability of the protein to bind peptides of different lengths has been studied by co-crystallising it with different ligands.

Results: Crystals of OppA from *Salmonella typhimurium* complexed with the peptides Lys–Lys–Lys (KKK) and Lys–Lys–Lys–Ala (KKKA) have been grown in the presence of uranyl ions which form important crystal contacts. These structures have been refined to 1.4 Å and 2.1 Å, respectively. The ligands are completely enclosed, their side chains pointing into large hydrated cavities and making few strong interactions with the protein.

Conclusions: Tight peptide binding by OppA arises from strong hydrogen bonding and electrostatic interactions between the protein and the main chain of the ligand. Different basic side chains on the protein form salt bridges with the C terminus of peptide ligands of different lengths.

Structure 15 December 1995, **3**:1395–1406

Key words: oligopeptide, periplasmic binding protein, transport

Introduction

Specific transmembrane translocation of solutes is fundamental to all life, and living systems have evolved a wide variety of transporters to move hydrophilic molecules across lipid membranes. The periplasmic binding protein-dependent systems of Gram-negative bacteria comprise a large family of transporters responsible for the uptake of many types of molecules including ions, amino acids, sugars and vitamins [1,2]. These transport systems share a common organisation; a substrate-specific ligand-binding protein in the periplasm captures extracellular ligand and delivers it to a cognate complex of four protein domains associated with the cytoplasmic membrane. These four domains, two of which span the membrane and two of which are ATPases at the cytoplasmic face of the membrane, are themselves a common feature of the larger ATP-binding cassette (ABC) family of transporters, that are widely distributed in eukaryotes as well as prokaryotes, and includes the cystic fibrosis transmembrane conductance regulator gene product (CFTR) and the multi-drug resistance P-glycoprotein [3,4]. Binding, protein-dependent transporters are also found in Gram-positive bacteria; as these have no periplasmic space the binding proteins are anchored to the cell membrane to prevent them from being lost to the medium [5,6].

In contrast to the lack of structural information that exists for the membrane components, a number of periplasmic binding protein structures have been solved crystallographically, some with and some without bound ligand [7]. These include arabinose-binding protein (ABP) [8], galactose/glucose-binding protein (GGBP) [9,10], ribose-binding protein (RBP) [11], sulphate-binding protein (SBP) [12], the leucine- and leucine–isoleucine–valine-binding proteins (LBP) and (LIVBP) [13,14], and histidine-binding protein (HBP) [15,16]. Although these proteins generally show very little sequence similarity they all have a number of features in common. All are bilobate molecules consisting of two domains, the polypeptide chain crossing between domains two or three times to form a hinge which opens and closes on ligand binding and release [7,17]. The ligand is bound between the two domains which close over it and enclose it completely, so that even highly charged molecules such as phosphate and sulphate ions are buried within the protein. Charges on the buried ligand are stabilized by hydrogen-bonding networks which dissipate the charge over a large number of groups. Lysine–arginine–ornithine-binding protein (LAOBP) and maltose-binding protein (MBP) have been crystallized in both closed (liganded) and open (unliganded) forms [18–20]. Comparisons between these forms show that the lobes of the proteins move relative to each other as rigid bodies, and the conformational change is brought about by changes in the ϕ and ψ angles of only a few residues in the strands connecting the two domains. Ligand binding is accompanied by a decrease in molecular radius [21]. Bound ligand stabilizes the closed form of the protein, which can then interact with the membrane-associated components of the transporter and deliver the ligand. Recently Flocco and Mowbray reported the structure of GGBP in the closed unliganded form, demonstrating that the protein is in equilibrium between open and closed forms in the absence of ligand [22].

*Corresponding author.

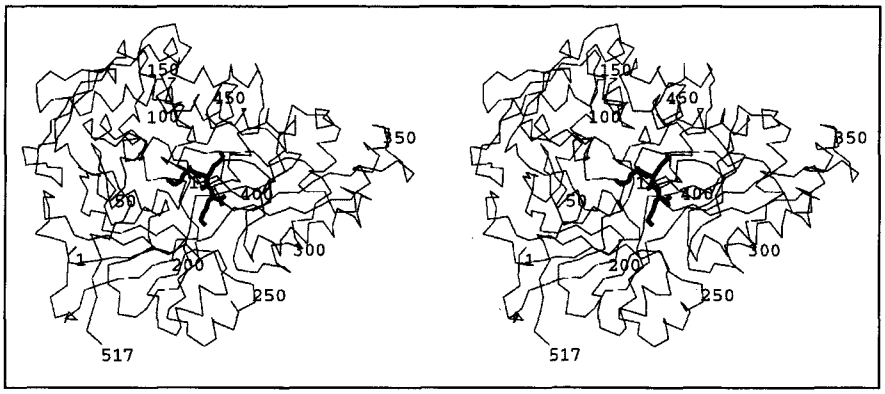

Fig. 1. Stereo Cα trace of OppA in the closed, ligand bound form. The tri-lysine ligand is shown in thicker lines.

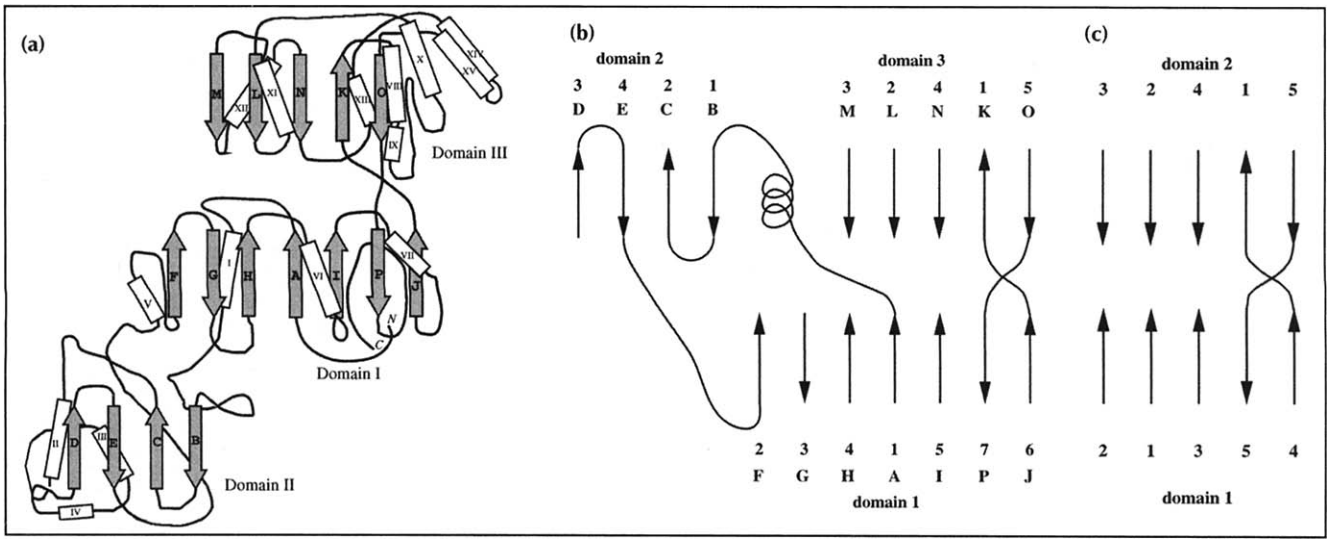

Fig. 2. Connectivity of secondary structure elements in OppA. **(a)** Topology diagram showing schematically the principal secondary structure elements in OppA. **(b)** Highly schematic diagram showing the connectivity of the three β sheets in OppA. Each strand is labelled by letter according to its position in the sequence and by number according to its position within the sheet. **(c)** β sheet connectivity in group II binding proteins. Comparison of (b) and (c) shows that strands H, A, I, P and J of domain I in OppA are in the same order as those in the first domain of a group II binding protein. The OppA subfamily of binding proteins could therefore have arisen from the insertion of an extra domain into a group II binding protein, or the group II binding proteins may have appeared through the loss of this domain from a larger ancestral protein.

Escherichia coli and *Salmonella typhimurium* have evolved three transport systems for the uptake of peptides: the dipeptide permease (Dpp), the tripeptide permease (Tpp), and the oligopeptide permease (Opp). Dpp and Opp are periplasmic binding protein-dependent transport systems whose periplasmic components, DppA and OppA, act as the initial receptors for transport. Dpp has a marked preference for dipeptides over amino acids and oligopeptides [23]. Opp transports peptides of between two and five residues in length, including cell-wall peptides that contain γ-linked and D-amino acids [24]. The oligopeptide-binding protein OppA is 58.8 kDa in size, making it one of the largest periplasmic binding proteins [25]. Its structure has been solved, revealing a new type of domain organisation among the binding proteins [26,27].

OppA and DppA are unusual binding proteins in that they accept an enormous variety of ligands; their binding affinity is affected relatively little by the sequence of the bound peptide. This behaviour contrasts strongly with other binding proteins which in general show marked ligand specificity. In the presence of uranyl ions OppA forms crystals of exceptional order for a molecule of its size. Here we describe the structure of OppA complexed to two different peptides, Lys–Lys–Lys (KKK) and Lys–Lys–Lys–Ala (KKKA), and discuss functional and evolutionary aspects of the protein.

Results and discussion

Overall structure

The most striking feature of the overall structure of OppA is the organisation of the polypeptide chain into three domains instead of two; however, it is apparent from Figure 2 that the relative organization of domains I and III with respect to each other and with respect to the ligand is similar to the arrangement of the two lobes of the other binding proteins. The binding proteins have been classified into two groups depending on the nature of the hinge region [19]. Group II proteins have a mixed β sheet in each lobe, and these exchange strands in a characteristic cross-over pattern also seen in OppA. The first domain of

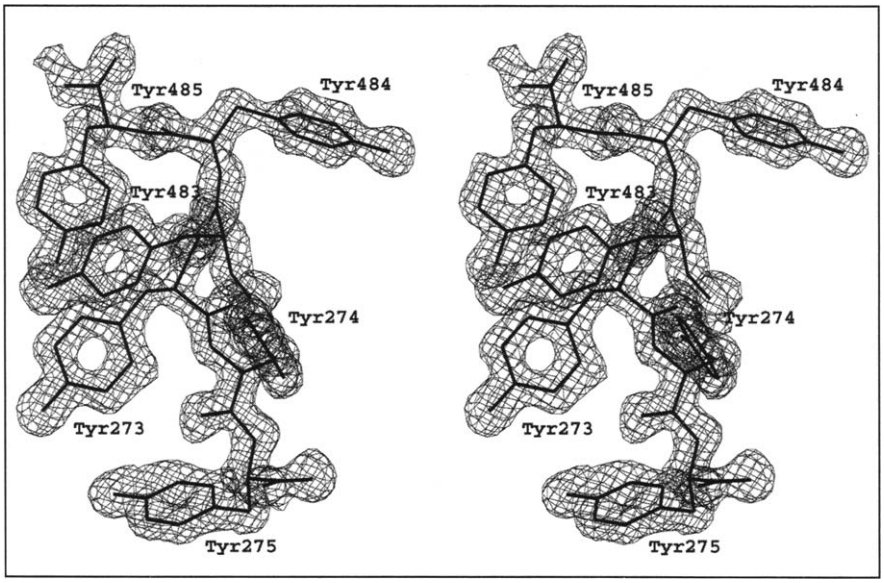

Fig. 3. The 1.4 Å $2F_o$–F_c electron-density map displayed over the two tri-tyrosine peptide segments, 273–275 and 483–485, that flank the sequence of domain III of OppA. The electron density is contoured at 1.2σ.

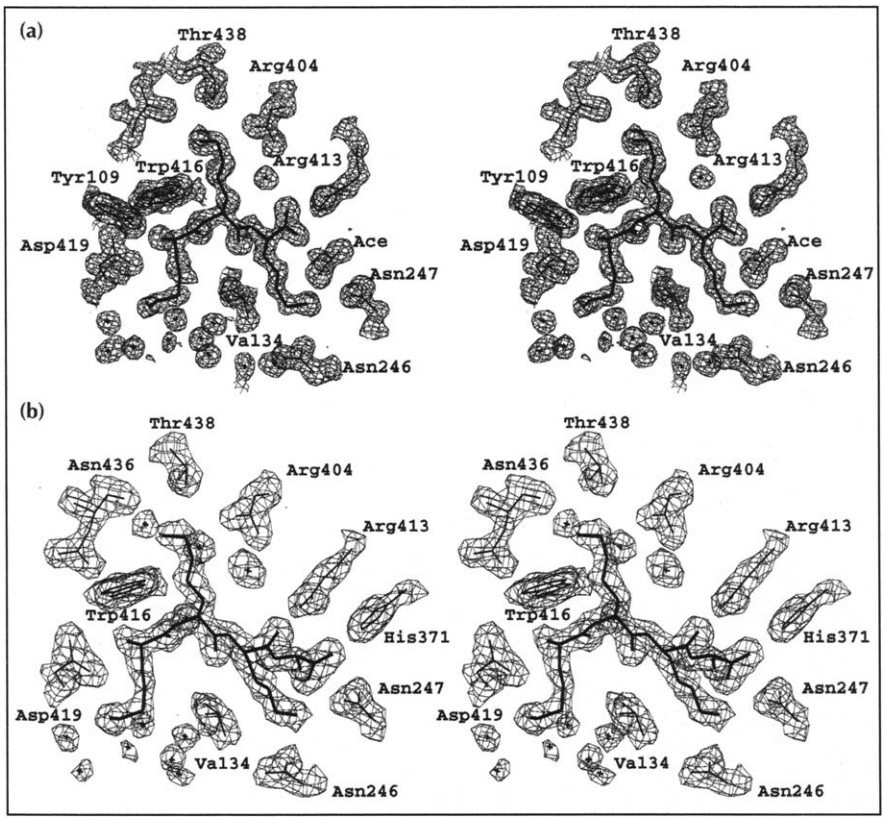

Fig. 4. Stereoviews of the ligand binding site in OppA. $2F_o$–F_c electron-density maps are displayed on the refined model, with protein bonds depicted as thin lines and ligand atoms as thick lines. **(a)** The OppA–KKK complex with electron density contoured at 1σ. 'Ace' denotes an acetate ion. **(b)** The OppA–KKKA complex with the electron density contoured at 1.5σ.

OppA consists of three separate segments of polypeptide chain, residues 1–44, 169–270 and 487–517, and contains a central seven-stranded β sheet. The N-terminal polypeptide segment forms a random coil for 14 residues before forming the central strand of sheet 1 which is followed by a further loop that includes residues 32–34 which form important contacts with the ligand (Fig. 1). The second domain is made up of a contiguous segment of polypeptide from residues 45–168. It comprises a four-stranded β sheet, one face of which is solvent exposed, and the opposite face is buried by two α helices and connecting segments in a well-ordered, highly hydrophobic core. The β sheet is not organised in the βαβ arrangement that is a common motif among the binding proteins, but instead this sheet consists of two pairs of strands connected by β hairpins. The topology of domain III is similar to lobe 2 of other group II binding proteins [19] and contains a five-stranded mixed β sheet. A disulphide bond connects Cys271 with Cys417. Domain III begins and ends with the triplet Tyr–Tyr–Tyr at residues 273–275 and 483–485, with the six tyrosines in close proximity as shown in Figure 3. The chain proceeds from the last β strand of domain III and almost immediately forms the final strand of the β sheet in domain I.

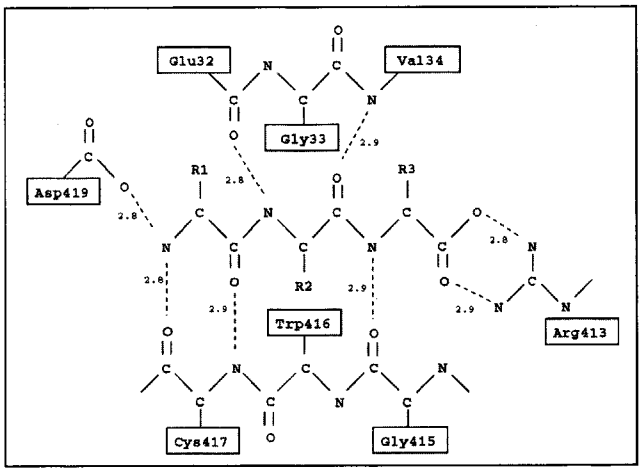

Fig. 5. Schematic diagram illustrating the interactions made by the main chain of the tri-lysine ligand with OppA. Protein residues are labelled. Hydrogen bonding and electrostatic interactions are indicated by the dotted lines. R1, R2 and R3 indicate the ligand side chains.

Ligand binding

In common with all other structures of liganded binding proteins, the ligand within OppA is completely buried and inaccessible to bulk solvent. The peptide adopts an extended conformation (Fig. 4). Tight peptide binding (K_D ~0.1–10 μM) is due to main-chain to main-chain contacts, the ligand forming antiparallel β-sheet-like interactions with an extended strand of sheet 3 (residues Gly415–Cys417) on one side and parallel β-sheet-like interactions with a loop of domain I (residues Glu32–Val34) on the other. All of the hydrogen bonding potential of the main-chain atoms of a tripeptide ligand is fulfilled. The majority of the ligand backbone contacts are provided by domain III. The structures of all liganded, closed form, binding proteins show one or other lobe interacting more strongly with the ligand than the other. It has been suggested that this may improve the efficiency of ligand capture by providing a good binding site in the open form of the protein [18]. Very well-ordered water molecules are found around the ligand, and relatively few interactions occur between the protein and the ligand side chains.

The main-chain interactions between the protein and a tripeptide are shown in Figure 5. The charges at the N and C termini of the peptide are countered in the complex by oppositely charged side chains, Asp419 and Arg413. LAOBP also uses an aspartate and an arginine to counter the charges on the α-ammonium and α-carboxyl groups of bound amino acids, in contrast to LIVBP which uses main-chain peptide bonds and side-chain hydroxyl groups for this purpose. The salt link formed between the ligand N terminus and Asp419 explains the low affinity of OppA for peptides with acetylated α-amino groups. From the structure of the OppA–KKK complex it is apparent that the protein has several charged side chains that can bind to the C terminus of peptides of different length. Two acetate ions found within the ligand-binding site in the OppA–KKK complex provide evidence that this is how longer peptides are accommodated (Fig. 6). The first of these is bonded to His371 and the second to Lys307, whose side chain points into the protein (Fig. 6). This second acetate ion presumably binds in the same place as the carboxyl group of a pentapeptide. The interactions made by the carboxyl group of a dipeptide are not known, though the side chain of Arg404 provides a potential counterion. No positively charged residue is found in a position to bind to the carboxyl group of an amino acid or a hexapeptide, which explains the preference of OppA for peptides of 2–5 residues.

In order to test the hypothesis that these different positively-charged groups can interact with the carboxyl group of ligands of different lengths, OppA was crystallized in the presence of a tetrapeptide, KKKA. The electron density around the tetrapeptide is shown in Figure 4b. It is clear that the first two residues of the KKKA and KKK peptides are closely superimposable and form identical interactions with the protein, but the ψ angle at the third residue is slightly different (Fig. 7). In the OppA–KKKA structure the salt bridge between the carboxyl group of tri-lysine and the guanidinium group of Arg413 is replaced by a hydrogen bond between this residue and the carbonyl group of Lys-3 of the ligand. In contrast to all other polar main-chain atoms of the ligand, the nitrogen atom of Ala-4 forms no hydrogen bonds. The C-terminal carboxyl group of KKKA forms the same interactions with the imidazolium side chain of His371 and the ε-amino group of Lys-3 as the acetate ion seen at this position in the

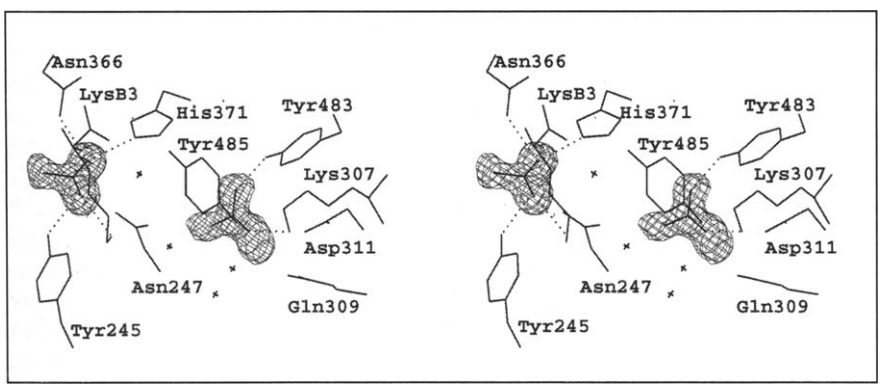

Fig. 6. Stereoview of the acetate-binding sites in the ligand-binding cavity in the OppA–tri-lysine complex. The F_o–F_c omit map is displayed on the acetate ions, contoured at 3σ. The calculated structure factors used to compute the map were generated from a coordinate set from which the two acetate ions were omitted and subjected to five cycles of least squares refinement in the program PROLSQ. Possible hydrogen bonding and electrostatic interactions made by these ions are indicated by the dotted lines. The third residue of the lignad is labelled

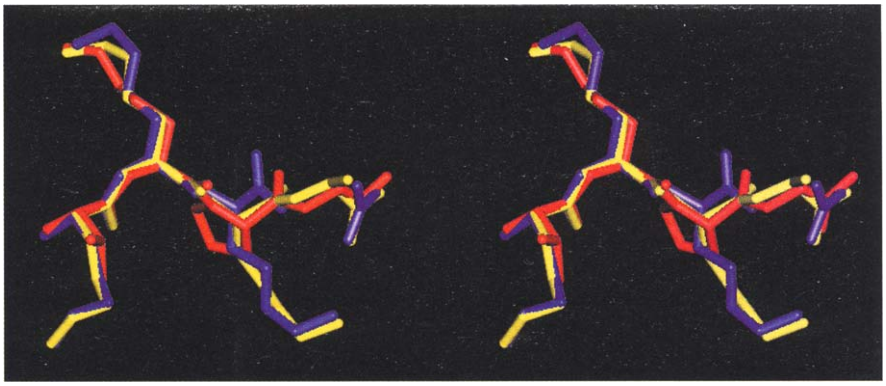

Fig. 7. Overlap of the peptide ligands in the OppA complexes with KKK (blue), KKKA (yellow) and the co-purified peptide modelled as Val-Lys-Pro-Gly (VKPG) (red) [28]. The ligands were superimposed by using least squares minimisation to overlap the main-chain protein atoms of residues 10–510 in the three structures. For the OppA–KKK structure acetate-1 is also shown.

OppA–KKK complex. It appears, therefore, that the positively charged side chains of Arg404, Arg413, His371 or Lys307 can form salt bridges with the C-terminal carboxyl group according to the whether the ligand is a di-, tri-, tetra- or pentapeptide; however, it should be noted that it is reported that OppA has a lower affinity for di- and pentapeptides than tri- and tetrapeptides [28]. Di-alanine binds OppA roughly 100 times less strongly than tri-alanine. As the affinity of tripeptides is of the order of 0.1–1.0 μM, dipeptides may bind with a K_D of perhaps 10–100 μM. This is still reasonably tight binding. Auxotrophic dpp^- strains of *E. coli* can survive using dipeptides as the source of essential amino acids, so it also appears physiologically relevant. Hexa-alanine is also reported to bind, but with very low affinity. OppA has so far failed to crystallize in the presence of the pentapeptide KKKAA, and only small crystals have been grown with the dipeptide KK. We are presently unable to confirm the predicted mode of binding of ligands containing two and five residues.

OppA and DppA are closely related in sequence, having identical residues or conservative substitutions at 234 of 507 positions [23]. The properties of these proteins reflect their different functional roles. OppA helps to recycle cell-wall peptides that would otherwise be lost to the medium as the bacteria grow and divide. The binding, by OppA, of cell-wall peptides containing γ-linked diaminopimelic acid may be mediated by arginines Arg17, Arg41 and Arg489, which lie in a water-filled cavity between domains I and III. DppA serves as a receptor for both peptide chemotaxis and dipeptide transport [29] and the sensing of peptide gradients would be impossible if this receptor also bound cell-wall peptides. Neither protein binds individual amino acids as these are transported by separate permeases, giving the bacteria greater flexibility of response to different nutrients [2]. Sequence alignment of these proteins shows that DppA has neutral residues at positions equivalent to Arg413, His371 and Lys307 of OppA consistent with its preference for dipeptide substrates. The basic sequence Lys(394)–Arg–Ala–Lys in DppA overlaps in the alignment with Arg404 of OppA. This has led to the suggestion of a role for one or more of these residues in binding the carboxyl group of a dipeptide ligand. DppA has been crystallized in the presence of several dipeptides [30] and the structure has recently been solved (P Dunten and S Mowbray, *Protein Sci.*, in press), revealing that in fact Arg355 is responsible for binding to the ligand C terminus (S Mowbray, personal communication).

The observation that the ligand is completely enclosed within the protein indicates that OppA has versatile side-chain pockets capable of accommodating groups of different size, shape, polarity and charge. The first three side-chain pockets are shown in Figure 8. Close to the ligand main chain the pockets are hydrophobic and there are no polar groups present on the protein that could compete for hydrogen bonds to these atoms. Val34 and the disulphide bridge between Cys271 and Cys417 lie against the main-chain atoms of the second ligand residue and form part of the first and third side chain pockets. The first ligand side chain (i.e. that of Lys-1) extends into the pocket and forms hydrogen bonds with three water molecules. A narrow channel filled with well-ordered water molecules connects this pocket to the surface of the protein. The first and third side-chain pockets are connected by another water-filled channel close to Val34. The second ligand side-chain pocket is lined by the side chains of Leu401, Arg404, Trp416 and Glu32 and capped by the side chain of Thr438. The Lys-2 side chain is flanked by two salt bridges between Glu32 and His405 and Arg404 and Glu276. The ε-amino group of Lys-2 is within hydrogen-bonding distance of a water molecule and the side chain of Glu32, though the strength of any bond with the latter is presumably reduced by its interaction with His405. The third pocket is enclosed by Asn246, Asn247 and Tyr245. The side chain of Lys-3 forms hydrogen bonds with the amide group of Asn247, the first of the buried acetate ions and a water molecule. The fourth ligand side chain pocket is less well-defined by the Ala-4 side chain whose methyl group has a similar location to that of the CH_3 group of the first acetate ion in the OppA–KKK structure (Fig. 7).

The exact manner in which various ligand side chains interact with OppA will be revealed by crystallographic studies of complexes with other ligands, though the structures of the pockets suggest they are able to accommodate any of the commonly occurring amino acid side chains with minimal conformational adaptation of the protein.

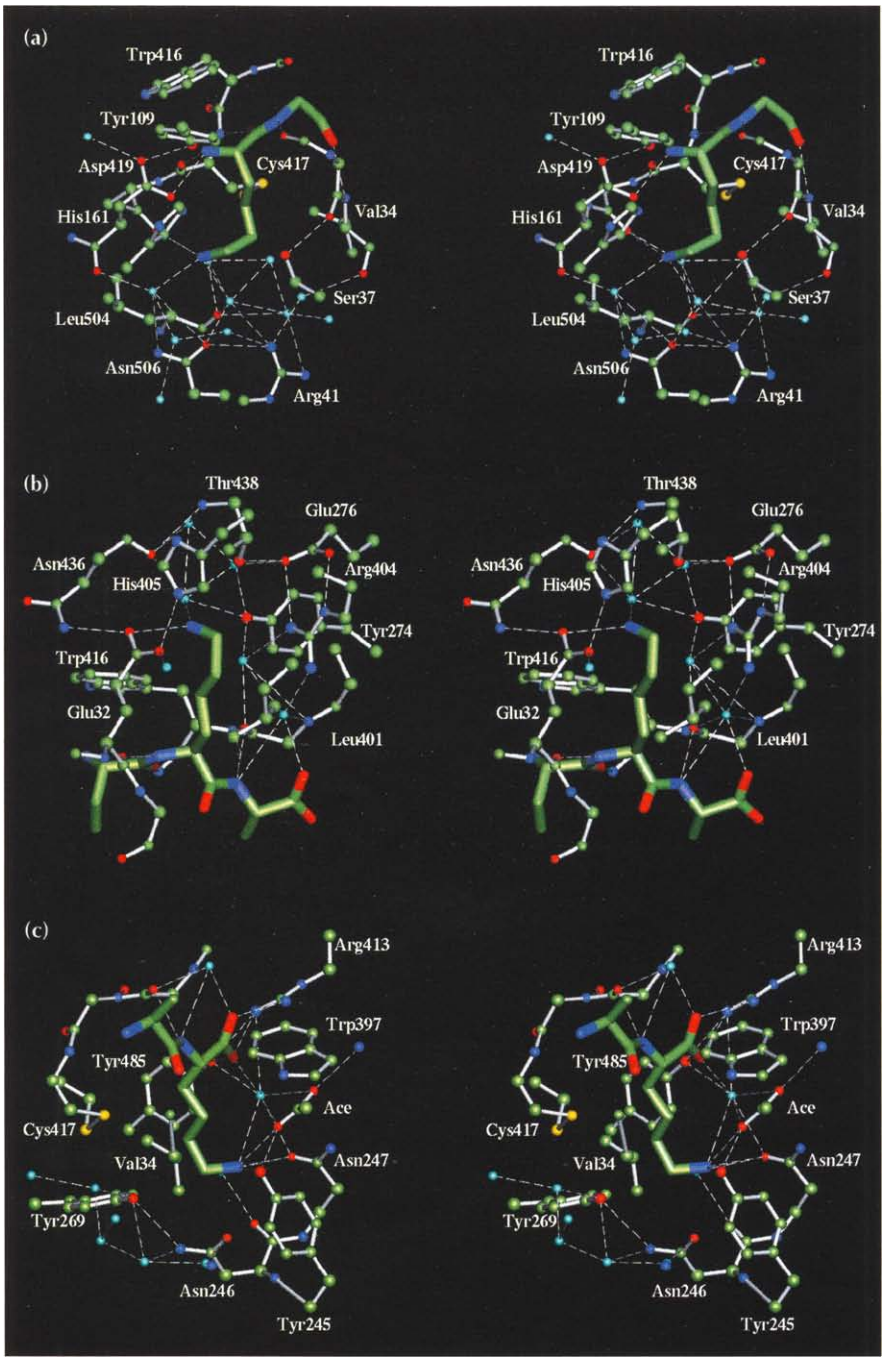

Fig. 8. Stereoviews of the three lysine side chains in the OppA–KKK complex: **(a)** Lysine-1; **(b)** Lysine-2; and **(c)** Lysine-3. The ligand is as a liquorice representation and the protein, whose residues are labelled, is represented as ball and stick, with the atoms coloured according to type carbon (white), oxygen (red) nitrogen (blue) and sulphur (yellow). Water molecules are coloured pale blue. 'Ace' denotes an acetate ion. Putative hydrogen bonding and electrostatic interactions are indicated by dashed lines. The figures were produced using QUANTA.

Ligand side chains of different size will clearly displace different numbers of water molecules. Although largely polar in nature, the pockets can accept large hydrophobic groups as the residues which form them are capable of hydrogen bonding to each other. This is most clearly seen in the second pocket which lies between two salt bridges. A lysine side chain in this position forms hydrogen bonds to Glu32 and a water molecule, yet a tryptophan residue would also fit as it is well known that salt bridges can be buried within the hydrophobic core of proteins.

Simply stated, the side-chain pockets of OppA have none of the characteristics that impose strong discrimination on ligand binding. They are not shaped to accept a unique ligand, nor are there polar or charged groups whose hydrogen-bonding potential can only be satisfied by particular ligand types. There will, of course, be differences in the nature and the extent of protein–ligand interactions with different peptide substrates. These factors result in small differences in ligand affinity. The specificity of Opp has been defined largely through genetic studies in which the capacity of amino acid auxotrophs to grow in the presence of defined peptide-containing media were examined. These assays demonstrate broad specificity but do not imply that all peptides are bound with equal affinity to OppA. Direct measurements of the ligand affinity of OppA from *E. coli* by equilibrium dialysis suggest that its affinity for different tripeptides can vary over a range greater than tenfold, depending on the peptide ligand sequence [28].

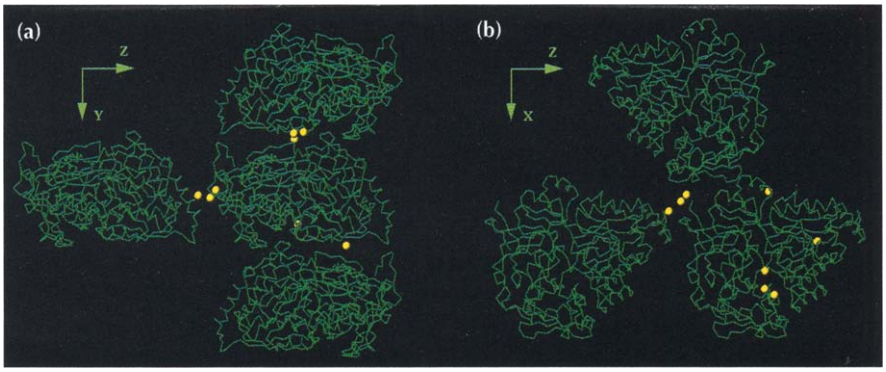

Fig. 9. Uranium atoms and the lattice structure of the crystal. **(a)** View down the x axis showing the Cα trace of an OppA molecule and some of its neighbours. Only the uranium atoms associated with the central OppA molecule are shown (as yellow spheres). **(b)** View down the y axis. The uranium atoms associated with the molecule on the bottom right are shown.

Crystal contacts; the role of uranium ions

The first crystals of OppA proved difficult to work with and an extensive search yielded no useful heavy atom derivatives [31]. Two subsequent crystal forms were grown when uranyl acetate was included in the crystallization mixture. OppA containing a heterogeneous mixture of co-purified peptides crystallizes under similar conditions to those for the OppA–KKK complex but in space-group $P2_12_12$ instead of $P2_12_12_1$, and contains one and a half rather than eight uranyl ions per asymmetric unit [26]. In space group $P2_12_12_1$, the crystal form described in this paper, the protein lies stacked with its flattened faces lying roughly in the xz plane. Along the y direction adjacent molecules therefore make a number of contacts, several of these mediated by uranium atoms 1 and 5. Uranium 1 is bonded to Asp323 and Asn394′ of a neighbouring protein molecule. A strong hydrogen bond is formed between the same pair of molecules by Asn199′ and the carbonyl group of Leu327, this interaction burying Val328. Residues in neighbouring molecules are indicated by prime (′). Along the z direction neighbouring molecules make only a single important contact, through uranium 8. This atom binds to Asp11 and Glu342′ of the adjacent molecule but there are no other intermolecular contacts at this point. The contact along the x axis is the most extensive and occurs between an external helix of domain I, lying against an edge strand of domain II in a symmetry related molecule. There are a number of very well defined water molecules in this region, and the two surfaces are sterically and electrostatically complementary. Arg237 and Asp141′ from the neighbouring molecule contact via a water molecule. The molecular packing is shown in Figure 9.

In the $P2_12_12$ crystal form there are three important crystal contacts. Uranium 1 holds two neighbouring molecules together through the carboxyl groups of Asp362 and Glu410 on one molecule and Glu150′ of its neighbour. Nearby Lys281 lies against residues Lys455′, Val456′ and Ala457′ in the adjacent molecule. The remaining contact occurs between Ala183 of one molecule and His75′ and Lys63′ of another. The histidine lies against the alanine side chain and the lysine forms a hydrogen bond to its carbonyl group. Uranium 2 lies on the twofold axis bridging the carboxyl groups of Glu329 of two adjacent OppA molecules. Although refined with an occupancy of 0.5 to account for symmetry, the temperature factor of this uranium atom is very high which suggests that the site is not fully occupied. Despite being grown in the presence of uranyl acetate, the $P2_12_12$ crystals are extremely sensitive to the addition of uranyl ions and must be mounted in mother liquor containing no uranium. It is possible that uranyl ions at high concentration pull the Glu329 residues of neighbouring molecules together and disrupt the crystal. The lattice contacts of the $P2_12_12$ crystals are less extensive than those of the $P2_12_12_1$ form, which is consistent with the greater physical strength of the latter.

Several protein crystal structures in which metal ions form important crystal contacts have now been described, and these crystals generally diffract to high resolution [16,32,33]. In the case of OppA this permitted phasing by multi-wavelength anomalous dispersion (MAD) methods [27] and also by 'reverse MIR' [26] as the occupancy of some uranyl ions could be reduced slightly by soaking crystals very briefly in a uranium-free mother liquor. EDTA destroys both crystal forms containing uranium.

Relationship to other periplasmic binding proteins

A number of proteins from a range of organisms show extensive sequence similarity to OppA (Gileadi *et al.*, unpublished data), including the *E. coli* nickel-binding protein NikA [34], the haem-binding protein from *Haemophilus influenzae* [35] and OppA from *Bacillus subtilis* [5]. The structure of OppA suggests no reason for the relatively large size of these proteins, and there appears to be no correlation between the size of a binding protein and the size of its ligand. In general the sequences of binding proteins are remarkably dissimilar. The sequence similarities between OppA and a number of other binding proteins suggests that these form a separate family, from which OppA is the first structure to be solved. DppA has four cysteine residues [23], none in positions corresponding to those in OppA, but the equivalent OppA residues form two close pairs (Asn18 and Ile241, Ser432 and Ser442), implying that the overall fold of the two proteins is very similar. (This has recently been confirmed [S Mowbray, personal communication].) Comparison with other sequences related to OppA shows that the pattern of conserved residues closely matches internal residues necessary for the protein fold. As expected, the residues involved in ligand binding are generally not

conserved. Asp419 is an exception and is only changed in NikA, which has alanine instead [34]. The haem-binding protein HbpA appears to provide a hydrophobic environment for its ligand similar to that found in cytochromes and globins, though it is not clear if the propionate groups are explicitly bonded or exposed at the surface. The sequence Arg–Lys–Arg–Ala–Lys occurs in a position close to the equivalent of His405 in OppA, lying over the enclosed ligand, and could be involved in stabilizing buried propionic acid groups; however, precise delineation of protein–ligand interactions within these complexes will require independent structural analysis.

Overall, domain II, which is missing from previously determined periplasmic binding protein structures, shows the highest degree of conservation and sequence identity. This suggests that it has no role in ligand binding but is involved in another function, for example interactions with the membrane complex, formed by OppB and OppC. Contacts made between domain II and domain III will clearly be broken as OppA switches to the open form, and this may allow the membrane complex to bind the closed form selectively. Further experiments are required to address this question. Deletion of domain II would leave a protein of similar size to the maltose-binding protein.

Comparison with other peptide-binding proteins

The crystal structures of a number of protein–peptide complexes have been solved [36,37]. These include Fabs [38–40], streptavidin [41], an SH2 domain from v-src oncogene product [42], and renins [43]. None of these proteins bury their ligand entirely, although among the Fab structures the surface area of the ligand removed from contact with the solvent correlates with the binding affinity. In general, peptide-ligand binding is highly specific, selectivity being brought about by the protein forming a number of interactions with the peptide side chains. Salt bridges are found to be quite rare in peptide binding, and hydrophobic interactions relatively important. Ligands bound to OppA form a number of main-chain to main-chain contacts, which are unusual among other peptide-binding proteins.

Non-specific ligand binding is achieved by class I and class II human leukocyte antigens (HLAs) in rather different ways to OppA. These molecules present peptide antigens to cytotoxic T-lymphocytes and T-helper cells, respectively, to elicit immune responses to viral infection or foreign antigens. This requires both classes of molecule to bind a very wide range of peptides extremely tightly. Class I and class II HLAs share a similar overall structure and bind peptides in an extended conformation, between two bent helices lying over a β sheet [44–46]. Class I HLAs show a preference for nonameric ligands, but 10mer and 11mer ligands will also bind. The crystal structures of a class I HLA complexed with a single peptide [47] and a mixture of peptides [48] show that binding is principally through clusters of conserved residues that bind the peptide termini and bury two 'anchor' peptide side chains. Ligands longer than nine residues are accommodated by the ligand bulging in the middle, which is quite distinct from the manner in which OppA accommodates ligands of different length. The importance of hydrogen bonds formed by the N and C termini of the peptides bound to class I HLA has been demonstrated by showing that their loss substantially destabilizes the protein–peptide complex [49]. Class II HLAs appear to bind main-chain atoms along the peptide ligand and make no contacts with the ligand termini [46]. All of these interactions are mediated via protein side chains. Although individual HLAs can bind a wide range of peptides, some polymorphic side-chain pockets will only accept certain amino acids on the ligand. This allelic restriction of ligands to particular sequence motifs is presumably to achieve very tight binding of a peptide that is necessarily exposed on the protein surface. In comparison, OppA need only bind small peptides which can pass through the bacterial outer membrane, and tight binding is accomplished solely by hydrogen bonding to the ligand main chain.

OppA is unusual both as a peptide-binding protein and as a periplasmic binding protein in that it will accept a wide variety of ligands. As much of the previous work in this area has been of a semi-quantitative nature, the true extent of the indifference of the protein to peptide sequence and length is unknown. Further experiments are underway to determine accurately the affinity of the protein for different peptides, and to analyse these data with regard to the structure of the protein–peptide complexes. The water structure in these pockets is clearly a crucial element in adapting the ligand-binding site to a wide variety of ligands. Crystals of OppA complexed to tri- and tetrapeptides can be grown that diffract to very high resolution, allowing the positions of the ligand atoms and buried water molecules to be determined very accurately. As ligand binding by OppA resembles the last step in protein folding processes it may prove a useful system for the analysis of structure/energy relationships among proteins in general.

Biological implications

A number of transmembrane transport systems found in Gram negative bateria involve a globular ligand-binding protein in the periplasmic space. These proteins, known as periplasmic binding proteins, share a common mode of ligand binding; the ligand is held in a deep cleft between two lobes of the protein and is completely buried. This enclosure of the ligand is normally associated with highly complementary interactions and therefore very specific ligand binding, a hallmark of this family of proteins. OppA, the periplasmic oligopeptide-binding component of oligopeptide permease (Opp), is an unusual periplasmic binding protein in that it will bind tightly to an enormous variety of ligands. Among peptide-binding proteins, the toleration of sequence variability of

the ligand is usually due to partial solvent exposure of the ligand, as is seen, for example, in the major histocompatibility complex (MHC) proteins. In contrast, OppA accommodates the ligand side chains in large hydrated pockets which impose little specificity on binding. Tight binding (K_D ~ 0.1–10 μM) is achieved by utilising the hydrogen-bonding potential of the main chain of the ligand, which remains invariant among α-linked peptides containing L-amino acids. Burial of the ligand within the body of the protein requires that any charges it carries be countered in some way. Water molecules and polar groups within the pockets dissipate any charge on ligand side chains, and the N and C termini of the peptide form salt bridges with charged residues found at the binding site of the protein. Four different positively charged protein side chains are found in suitable positions to bind to the C terminus of peptides that are two to five residues in length. The second domain of OppA has no counterpart in previously determined structures of binding proteins, but is found to be well conserved in the sequences of related proteins. It appears to have no role in ligand binding, and is possibly involved in contacts with the complex formed by OppB, C, D and F, components of the Opp transport system that sit within the cytoplasmic membrane.

Materials and methods

Crystallization and data collection

In common with other binding proteins, OppA has high affinity for its ligands and as a result co-purifies with bound peptides, a phenomenon referred as the retention effect [28,50]. Peptides can be released from OppA by extensive washing with sodium acetate at pH 5 on a Pharmacia MonoS cation exchange column. This obviates the need for protein denaturation which has been used to remove ligand from other binding proteins. Crystals were grown from 7–10% PEG 4000 with 50 mM sodium acetate pH 5.5 and 1 mM uranyl acetate. Ligand-free protein was mixed with a fivefold molar excess of KKK or KKKA immediately prior to setting up hanging drops. In the presence of these ligands and uranyl acetate, OppA crystallizes in space group $P2_12_12_1$. Crystals appear after a week and grow over a further 2–3 weeks, usually to a final size of about 0.1×0.1×0.2 mm^3. They are extremely well-ordered and amenable to freezing.

For OppA–KKK, X-ray diffraction data were collected from three crystals, all frozen to −150°C. The crystals were washed in a solution of three parts mother liquor to one part glycerol immediately prior to flash freezing in order to prevent ice formation. Data from the first crystal were collected to 1.8 Å spacing using a Siemens Xentronics multi-wire chamber detector and Cu Kα radiation from a rotating anode X-ray generator operated at 50 kV and 100 mA. These data were processed using XDS [51]. Data from two more crystals were collected at station 9.6 of the Daresbury Synchrotron Radiation Source using a 30 cm Mar Research image plate and 0.875 Å wavelength incident radiation. The crystal-to-detector distance was set to 195 mm, giving a nominal resolution of 1.36 Å at the plate edge. Image plate data were processed using MOSFLM version 5.2 [52]. The majority of spots below 3.0 Å resolution collected at the SRS were overloaded and a low-resolution cut-off at this limit was applied during processing. The Xentronics data (15.0–1.8 Å spacing) were scaled together with the image plate data using ROTAVATA/AGROVATA [52] to give a single dataset. The overall R_{merge} was 9.7% for all 101 686 reflections in the range of 10.0 Å–1.4 Å spacing. Approximately three-quarters of the spots measured between 1.40 Å and 1.42 Å resolution were stronger than 3σ. Reflections between 1.40 Å and 1.36 Å have a mean I/σ(I) of 2.8 (Table 1) but were not used in the refinement as the Wilson plot is not linear beyond 1.40 Å.

Table 1. Data collection and refinement statistics.

	OppA–KKKA	OppA–KKK
Space Group	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions (Å)	110.7,77.1,71.3	109.2,76.0,70.3
Temperature (°C)	18	−150
Resolution limits (Å)	10.0–2.1	10.0–1.4
Unique reflections	35 923	101 686
Average multiplicity	4.7	2.7
Completeness (%)		
Overall	99.9	88.3
Outermost data shell used	100.0	84.2
R_{merge}* (%)		
Overall	6.3	9.7
Outermost data shell used	17.6	20.4
Mean I/σ(I)		
Overall	9.6	5.5
Outermost measured data	3.8	2.8
No. of Protein and Ligand Atoms	4198	4193
No. of Solvent Atoms	337	549
R_{cryst}† (%)	14.4	18.3
Free R_{cryst} (%)	19.7	–
Rms_{bond} (Å)	0.018	0.016
Rms_{angle} (Å)	0.036	0.029
Rms_{planes} (Å)	0.040	0.034
Average temperature factor		
Main chain	16.9 Å^2	11.4 Å^2
Side chain	23.7 Å^2	17.0 Å^2
Solvent	33.4 Å^2	27.5 Å^2

*$R_{merge}=\Sigma|I_i-I_n|/\Sigma I_n$ where I_i is an observed intensity hkl and I_n is the average of the observed equivalents. †$R_{cryst}=\Sigma_{hkl}||F_{obs}|-|F_{calc}||/\Sigma_{hkl}|F_{obs}|$ where |Fobs| and $|F_{calc}|$ are the observed and calculated structure factor amplitudes of a reflection hkl respectively.

For OppA–KKKA, data were collected at room temperature using a Rigaku Raxis-II detector from a single crystal mounted in a glass capillary. Data were processed with DENZO [53] and processing statistics are given in Table 1.

Structure refinement procedure

The starting model for refinement was the 1.8 Å structure of OppA–trilysine described previously [26]. Five acetate ions were modelled into the 1.8 Å electron-density map, two of these within the protein itself and three on the protein surface close to the uranium sites. As density was relatively poor for the external acetate ions these were removed from the model before beginning refinement with the 1.4 Å dataset. Refinement was carried out with PROLSQ [52]. The model was refined to an R factor of 18.6% (all reflections between 10.0 and 1.4 Å) and a free R factor of 22.6% without any further manual adjustment of the structure. From this point all reflections were used in the refinement. Additional water molecules were added using the program ARP [54], and manual

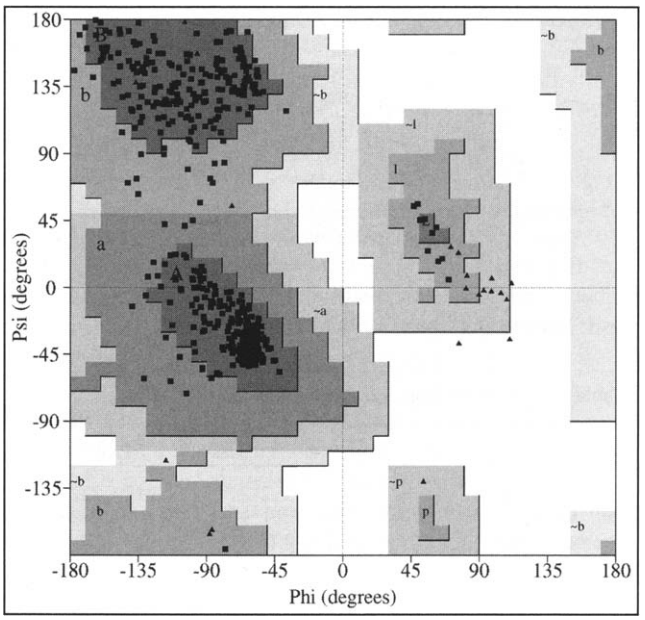

Fig. 10. Ramachandran plot of the OppA–KKK structure. This figure was produced using PROCHECK [57]. Glycine residues are shown as triangles.

adjustments were made using O [55]. The final model contains 549 water molecules, and has an R factor of 18.3%, calculated using all reflections between 10.0 and 1.4 Å resolution. Refinement of the OppA–KKKA complex was carried out in a similar manner, except that simulated annealing by X-PLOR [56] was used as an initial step so that a free R factor could be calculated. In both cases density around the uranium ions is unclear due to ripples in the electron density and anisotropic movement of the heavy atoms. The solvent model cannot be considered reliable in these regions. In the lower resolution OppA–KKKA map the acetate ions appear less clearly and are not included in the model.

Data collection and refinement statistics are presented for both structures in Table 1. The Ramachandran plot of OppA–KKK is given in Figure 10 and representative sections of the density maps for both structures are shown in Figure 4. The average B factor for all protein atoms in the OppA–KKK structure is 14.3 $Å^2$, and for all ligand atoms 10.1 $Å^2$. Although most of the protein atoms have clear density and can be positioned accurately, the maps show considerable distortion around the uranium ions, partly due to anisotropic movement of the heavy atoms which contribute very largely to the overall X-ray scattering of the crystals. The B factors for the uranium atoms have been refined anisotropically with MLPHARE [52] and the results are shown in Table 2.

Table 2. Anisotropic B factors calculated for uranium atoms at 1.4 Å resolution using MLPHARE.

	Only uranium atoms deleted from the model			Uranium and associated atoms deleted from the model		
U1 Eigen values	5.44	6.55	9.06	5.30	6.63	8.87
Eigen vectors	−0.933	−0.270	0.238	−0.956	−0.235	0.173
	0.116	−0.851	−0.512	0.101	−0.822	−0.561
	0.341	−0.450	0.825	0.274	−0.519	0.810
U2 Eigen values	6.97	10.11	15.60	6.81	9.97	15.58
Eigen vectors	−0.834	0.425	−0.351	−0.956	−0.235	0.173
	0.487	0.867	−0.108	0.101	−0.822	−0.561
	−0.258	0.261	0.930	0.274	−0.519	0.810
U3 Eigen values	7.75	9.92	15.27	7.78	9.94	15.07
Eigen vectors	−0.845	0.386	−0.370	−0.870	0.315	−0.379
	0.453	0.884	−0.113	0.397	0.904	−0.160
	−0.284	0.263	0.922	−0.292	0.290	0.911
U4 Eigen values	8.66	12.06	15.82	8.69	12.12	15.88
Eigen vectors	0.219	−0.937	−0.270	0.201	−0.942	−0.269
	−0.865	−0.315	0.391	−0.871	−0.297	0.392
	0.452	−0.148	0.880	0.449	−0.156	0.880
U5 Eigen values	8.82	11.27	17.94	7.85	10.19	22.26
Eigen vectors	−0.500	0.864	−0.061	−0.479	0.868	−0.129
	−0.413	−0.300	−0.860	−0.413	−0.352	−0.840
	−0.762	−0.404	0.506	−0.774	−0.350	0.527
U6 Eigen values	10.94	17.55	26.84	10.90	17.58	26.80
Eigen vectors	0.973	−0.215	−0.088	0.973	−0.215	−0.086
	0.073	−0.077	0.994	0.071	−0.080	0.994
	−0.220	−0.974	−0.059	−0.221	−0.973	−0.062
U7 Eigen values	14.60	22.52	31.31	14.52	22.29	31.11
Eigen vectors	0.982	−0.041	−0.184	0.980	−0.047	−0.192
	0.185	0.038	0.982	0.194	0.034	0.980
	−0.033	−0.998	0.045	−0.039	−0.998	0.042
U8 Eigen values	13.65	19.92	31.24	13.62	19.74	31.28
Eigen vectors	0.296	−0.158	0.942	0.295	−0.160	0.942
	−0.934	0.157	0.320	−0.935	0.153	0.319
	0.198	0.975	0.101	0.195	0.975	0.104

Coordinates and X-ray reflection data for both structures described in this paper have been deposited in the Brookhaven Protein Databank. No hold has been requested on release.

Acknowledgements: We would like to thank Dr Andrew Leslie for help with data processing. This work was supported by grants GR/H68864 from the SERC and G8908552CB from the MRC (AJW). CFH was supported by the Imperial Cancer Research Fund and is a Howard Hughes International Research Scholar. JRHT is a Royal Society University Research Fellow.

References

1. Ames, G.F.-L. (1986). Bacterial periplasmic transport systems: structure, mechanism and evolution. *Annu. Rev. Biochem.* **55**, 397–425.
2. Furlong, C.E. (1987). Osmotic-shock-sensitive transport systems. In Escherichia coli *and* Salmonella typhimurium*: Cellular and Molecular Biology.* (Neidhardt, F.C., ed), pp. 768–796, American Society for Microbiology, Washington DC.
3. Higgins, C.F. (1992). ABC transporters: from microorganisms to man. *Annu. Rev. Cell Biol.* **8**, 67–113.
4. Hyde, S.C., *et al.*, & Higgins, C.F. (1990). Structural model of ATP-binding proteins associated with cystic fibrosis, multi-drug resistance and bacterial transport. *Nature* **346**, 362–365.
5. Perego, M., Higgins, C.F., Pearce, S.R., Gallagher, M.P. & Hoch, J. A. (1991). The oligopeptide transport system of *Bacillus subtilis* plays a role in the initiation of sporulation. *Mol. Microbiol.* **5**, 173–185.
6. Gilson, E., Alloing, G., Schmidt, T., Claverys, J.P., Dindler, R. & Hofning, M. (1988). Evidence of high-affinity binding protein dependent transport systems in Gram positive bacteria and Mycoplasma. *EMBO J.* **7**, 3971–3974.
7. Quiocho, F.A. (1990). Atomic structures of periplasmic binding proteins and the high-affinity active transport systems in bacteria. *Phil. Trans. R. Soc. Lond. B Biol. Sci.* **326**, 341–352.
8. Quiocho, F.A. & Vyas, N.K. (1984). Novel stereospecificity of the L-arabinose binding protein. *Nature* **310**, 381–386.
9. Zou, J.Y., Flocco, M.M. & Mowbray, S.L. (1993). The 1.7 Å refined X-ray structure of the periplasmic glucose-galactose receptor from *Salmonella typhimurium. J. Mol. Biol.* **233**, 739–752.
10. Vyas, M.N., Vyas, N.K. & Quiocho, F.A. (1994). Crystallographic analysis of the epimeric and anomeric specificity of the periplasmic transport/chemosensory protein receptor for D-glucose and D-galactose. *Biochemistry* **33**, 4762–4768.
11. Mowbray, S.L. & Cole, L.B. (1992). 1.7 Å structure of the periplasmic ribose receptor from *Escherichia coli. J. Mol. Biol.* **225**, 155–175.
12. Pflugrath, J.W. & Quiocho, F.A. (1988). The 2 Å resolution structure of the sulfate-binding protein involved in active transport in *Salmonella typhiumurium. J. Mol. Biol.* **200**, 163–180.
13. Sack, J.S., Saper, M.A. & Quiocho, F.A. (1989). Periplasmic binding protein structure and function. Refined X-ray structure of the leucine/isoleucine/valine-binding protein and its complex with leucine. *J. Mol. Biol.* **206**, 171–191.
14. Sack, J.S., Trakhanov, S.D., Tsigannik, I.H. & Quiocho, F.A. (1989). Structure of the L-leucine-binding protein refined at 2.4 Å resolution and comparison with the Leu/Ile/Val binding protein structure. *J. Mol. Biol.* **206**, 193-207
15. Oh, B.-H., Kang, C.-H., De Bondt, H., Kim,S.-H., Nikaido, K., Joshi, A.K. & Ames, G.F.-L. (1994). The bacterial periplasmic histidine-binding protein. *J. Biol. Chem.* **269**, 4135–4143.
16. Yao, N., Trakhanov, S. & Quiocho, F.A. (1994). Refined 1.89 Å structure of the histidine-binding protein complexed with histidine and its relationship with many other active transport/chemosensory proteins. *Biochemistry* **33**, 4769–4779.
17. Quiocho, F.A. (1991). Atomic structures and function of periplasmic receptors for active transport and chemotaxis. *Curr. Opin. Struct. Biol.* **1**, 922–933.
18. Oh, B.-H., Pandit, J., Kang, C.-H., Nikaido, K., Gokcen, S., Ames, G.F.-L. & Kim, S.-H. (1993). Three-dimensional structures of the periplasmic lysine/arginine/ornithine-binding protein with and without a ligand. *J. Biol. Chem.* **268**, 11348–11355.
19. Spurlino, J.C., Lu, G.-Y. & Quiocho, F.A. (1991). The 2.3Å resolution structure of the maltose- or maltodextrin-binding protein, a primary receptor of bacterial active transport and chemotaxis. *J. Biol. Chem.* **266**, 5202–5219.
20. Sharff, A.J., Rodseth, L.E., Spurlino, J.C. and Quiocho, F.A. (1992). Crystallographic evidence of a large ligand-induced hinge-twist motion between the two domains of the maltodextrin binding protein involved in active transport and chemotaxis. *Biochemistry* **31**, 10657–10663.
21. Newcomer, M.E., Lewis, B.A. & Quiocho, F.A. (1981). The radius of gyration of L-arabinose-binding protein decreases upon ligation. *J. Biol. Chem.* **256**, 13218–13222.
22. Flocco, M.M & Mowbray, S.L. (1994). The 1.9Å X-ray structure of a closed unliganded form of the periplasmic glucose/galactose receptor from *Salmonella typhimurium. J. Biol. Chem.* **269**, 8931–8936.
23. Olson, E.R., Dunyak, D.S., Jurss, L.M. & Poorman, R.A. (1991). Identification and chracterisation of dppA, an *Escherichia coli* gene encoding a periplasmic dipeptide transport protein. *J. Bacteriol.* **173**, 234–244.
24. Goodell, E.W. & Higgins, C.F. (1987). Uptake of cell wall peptides by *Salmonella typhimurium* and *Escherichia coli. J. Bacteriol.* **169**, 3861–3865.
25. Hiles, I.D. & Higgins, C.F. (1986). Peptide uptake by *Salmonella typhimurium*. The oligopeptide-binding protein. *Eur. J. Biochem.* **158**, 561-567.
26. Tame, J.R.H., Murshudov, G.N., Dodson, E.J., Neil, T.K., Dodson, G.G., Higgins, C.F. & Wilkinson, A.J. (1994). The structural basis of sequence-independent peptide binding by OppA protein. *Science* **264**, 1578–1581.
27. Glover, I.D., Denny, R., Nguti, D., McSweeney, S.M., Thompson, A., Dodson, E.J., Wilkinson, A.J. & Tame, J.R.H. (1994). Structure determination of OppA at 2.3Å resolution using multiple-wavelength anomalous dispersion methods. *Acta Cryst. D* **51**, 39–47.
28. Guyer, C.A., Morgan, D.G. & Staros, J.V. (1986). Binding specificity of the periplasmic oligo-peptide binding protein from *Escherichia coli. J. Bacteriol.* **168**, 775–779
29. Abouhamad, W.N., Manson, M., Gibson, M.M & Higgins, C.F. (1991). Peptide transport and chemotaxis in *Escherichia coli* and *Salmonella typhimurium*: characterisation of the dipeptide permease (Dpp) and the dipeptide binding protein. *Mol. Microbiol.* **5**, 1035–1047.
30. Dunten, P.W., Harris, J.H., Feiz, V. & Mowbray, S.L. (1993). Crystallisation and preliminary X-ray analysis of the dipeptide binding protein from *Escherichia coli. J. Mol. Biol.* **231**, 145–147.
31. Tolley, S.P., Derewenda, Z., Hyde, S.C.Higgins, C.F. & Wilkinson, A.J. (1988). Crystallisation of the periplasmic oligopeptide-binding protein of *Salmonella typhimurium. J. Mol. Biol.* **204**, 493–494.
32. Leslie, A.G.W., Moody, P.C.E. & Shaw, W.V. (1988). Structure of chloramphenicol acetyl transferase at 1.75Å resolution. *Proc. Natl. Acad. Sci USA* **85**, 4133–4137.
33. Nordlund, P., Uhlin, U., Westergren, C., Sjoberg, B.-M. and Eklund, E. (1989). New crystal forms of the small subunit of ribonucleotide reductase from *Escherichia coli. FEBS Lett.* **258**, 251–254.
34. Navarro, C., Wu, L.-F. and Mandrand-Bertholet, M.-A. (1993). The *nik* operon of *Escherichia coli* encodes a periplasmic binding-protein-dependent transport system for nickel. *Mol. Microbiol.* **9**, 1181–1191.
35. Hanson, M.S., Slaughter, C. & Hansen, E.J. (1992). The hbpA gene of *Haemophilus influenzae* type b encodes a heme-binding lipoprotein conserved among heme-dependent *Haemophilus* species. *Infect. Immun.* **60**, 2257–2266.
36. Zvelebil, M.J.J.M. & Thornton, J.M (1994). Peptide–protein interactions: an overview. *Quart. Rev. Bio.* **26(3)**, 333–363
37. Stanfield, R.L. & Wilson, I.A. (1995). Protein-peptide interactions. *Curr. Opin. Struct. Biol.* **5**, 103–113.
38. Stanfield, R.L., Fieser, T.M., Lerner, R.A. & Wilson, I.A. Crystal structure of an antibody to a peptide and its complex with peptide antigen at 2.8Å resolution. *Science* **248**, 712–719.
39. Rini, J.M., Schulze-Gahmen, U. & Wilson, I.A. (1992). Structural evidence for induced fit as a mechanism for antibody-antigen recognition. *Science* **255**, 959–965.
40. Garcia, K.C., Ronco, P.M., Verroust, P.J., Bruenger, A.T. & Amzel, L.M. (1992). Three-dimensional structure of an angiotensin II-Fab complex at 3Å: hormone recognition by an anti-idiotypic antibody. *Science* **257**, 502–507.
41. Weber, P.C., Pantoliano, M.W. & Thompson, L.D. (1992). Crystal structure and ligand-binding studies of a screened peptide complexed with streptavidin. *Biochemistry* **31**, 9350–9354.
42. Waksman, G., *et al.*, & Kuriyan, J. (1992). Crystal structure of the phosphotyrosine recognition domain SH2 of v-src complexed with tyrosine-phosphorylated peptides. *Nature* **358**, 646–653.
43. Dhanaraj, V., *et al.*, & Hoover, D.J. (1992). X-ray analyses of peptide-inhibitor complexes define the structural basis of specificity for human and mouse renins. *Nature* **357**, 466–472.
44. Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L. & Wiley, D.C. (1987). Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* **329**, 506–512.
45. Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L. & Wiley, D.C. (1987). The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature* **329**, 512–518.

46. Brown, J.H., Jardetsky, T.S., Gorga, J.C., Stern, L.J., Urban, R.G., Strominger, J.L. & Wiley, D.C. (1993). Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature* **364**, 33–39.
47. Silver, M.L., Guo, H.-C., Strominger, J.L. & Wiley, D.C. (1992). Atomic resolution of a human MHC molecule presenting an influenza virus peptide. *Nature* **360**, 367–369.
48. Guo, H.-C., Jardetsky, T.S., Garrett, T.P.J., Lane, W.S., Strominger, J.L. & Wiley, D.C. (1992). Different length peptides bind to HLA-Aw68 similarly at their ends but bulge out in the middle. *Nature* **360**, 364–366.
49. Bouvier, M. & Wiley, D.C. (1994). Importance of peptide amino and carboxyl termini to the stability of MHC class I molecules. *Science* **265**, 398–402.
50. Silhavy, T.J., Szmelcman. S., Boos, W. & Schwartz, M. (1975). On the significance of the retention of ligand by protein. *Proc. Natl. Acad. Sci. USA* **72**, 2120–2124.
51. Kabsch, W. (1988). Evaluation of single-crystal X-ray diffraction data from a position-sensitive detector. *J. Appl. Cryst.* **21**, 916–924.
52. Collaborative Computational Project Number 4. (1994). *Acta Cryst. D* **50**, 760–763.
53. Otwinowski, Z. (1990). *DENZO Data Processing Package*. Yale University, New Haven, CT.
54. Lamzin, V.S & Wilson, K.S. (1993). Automated refinement of protein models. Acta Cryst. D **49**, 129–147.
55. Jones, T.A. & Kjeldgaard, M. (1990) O — the manual, Uppsala University.
56. Brünger, A.T. (1990) X-PLOR Version 2.1, Yale University, New Haven, CT.
57. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291.

Received: **26 Jul 1995**; revisions requested: **4 Sep 1995**; revisions received: **20 Sep 1995**. Accepted: **25 Sep 1995**.