The structure of enzyme IIA^{lactose} from *Lactococcus lactis* reveals a new fold and points to possible interactions of a multicomponent system

Piotr Sliz[1], Roswitha Engelmann[2], Wolfgang Hengstenberg[2] and Emil F Pai[1,3]*

Background: The bacterial phosphoenolpyruvate: sugar phosphotransferase system (PTS) is responsible for the binding, transmembrane transport and phosphorylation of numerous sugar substrates. The system is also involved in the regulation of a variety of metabolic and transcriptional processes. The PTS consists of two non-specific energy coupling components, enzyme I and a heat stable phosphocarrier protein (HPr), as well as several sugar-specific multiprotein permeases known as enzymes II. In most cases, enzymes IIA and IIB are located in the cytoplasm, while enzyme IIC acts as a membrane channel. Enzyme IIA^{lactose} belongs to the lactose/cellobiose-specific family of enzymes II, one of four functionally and structurally distinct groups. The protein, which normally functions as a trimer, is believed to separate into its subunits after phosphorylation.

Results: The crystal structure of the trimeric enzyme IIA^{lactose} from *Lactococcus lactis* has been determined at 2.3 Å resolution. The subunits of the enzyme, related to each other by the inherent threefold rotational symmetry, possess interesting structural features such as coiled-coil-like packing and a methionine cluster. The subunits each comprise three helices (I, II and III) and pack against each other forming a nine-helix bundle. This helical bundle is stabilized by a centrally located metal ion and also encloses a hydrophobic cavity. The three phosphorylation sites (His78 on each monomer) are located in helices III and their sidechains protrude into a large groove between helices I and II of the neighbouring subunits. A model of the complex between phosphorylated HPr and enzyme IIA^{lactose} has been constructed.

Conclusions: Enzyme IIA^{lactose} is the first representative of the family of lactose/cellobiose-specific enzymes IIA for which a three-dimensional structure has been determined. Some of its structural features, like the presence of two histidine residues at the active site, seem to be common to all enzymes IIA, but no overall structural homology is observed to any PTS proteins or to any other proteins in the Protein Data Bank. Enzyme IIA^{lactose} shows surface complementarity to the phosphorylated form of HPr and several energetically favourable interactions between the two molecules can be predicted.

Addresses: [1]Department of Biochemistry, University of Toronto, 1 King's College Circle, Toronto, ON M5S 1A8, Canada, [2]Fakultät Biologie, Arbeitsgruppe Physiologie der Mikroorganismen, Ruhr-Universität Bochum, D-44780, Bochum, Germany and [3]Departments of Medical Biophysics, Molecular and Medical Genetics and the Protein Engineering Centre of Excellence, University of Toronto, 1 King's College Circle, Toronto, ON M5S 1A8, Canada.

*Corresponding author.
E-mail: pai@hera.med.utoronto.ca

Key words: enzyme IIA, helical bundles, histidine phosphorylation, modelling, PTS, X-ray crystallography

Received: 17 March 1997
Revisions requested: 11 April 1997
Revisions received: 27 April 1997
Accepted: 30 April 1997

Structure 15 June 1997, 5:775–788
http://biomednet.com/elecref/0969212600500775

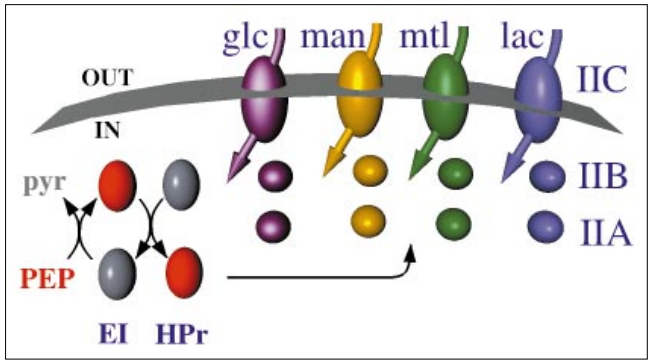
Introduction

In homofermentative lactic acid bacteria, many carbohydrates are taken up by the bacterium via a phosphoenolpyruvate: sugar phosphotransferase system (PTS). The PTS is responsible for binding, transmembrane transport, and phosphorylation of numerous sugar substrates. It is also involved in several regulatory events including catabolite repression, catabolite inhibition and chemotaxis, thereby acting as a primitive bacterial signal transduction system (for a review, see [1]).

The PTS is composed of two non-specific, energy-coupling components, enzyme I and a heat stable phosphocarrier protein (HPr), as well as several sugar specific enzymes II. Depending on the system, the enzymes II consist of three independent folding domains that exist either as individual subunits or as fused polypeptides. Commonly, two of the domains, IIA and IIB, are hydrophilic and located in the cytoplasm, while IIC is an integral membrane protein (Fig. 1).

Enzyme I is a histidine kinase which uses phosphoenolpyruvate (PEP) as a substrate for autophosphorylation. The phosphoryl group is then transferred to a histidine residue of HPr. The phosphorylated form of HPr (P–HPr) in turn passes the phosphoryl group onto histidine residues in a variety of sugar-specific enzymes IIA (or IIA domains in the case of fused peptide chains). The phosphoryl group is subsequently handed to a cysteine residue in the corresponding hydrophilic IIB enzyme or domain. The proper

Figure 1



The phosphoenolpyruvate (PEP)-dependent sugar transport system (PTS) in bacteria. Enzyme I (EI) is autophosphorylated by PEP and consequently phosphorylates the heat stable phosphocarrier protein (HPr). HPr transfers phosphate to several different enzymes IIA, which can be classified into four different families: glucose- (glc), mannose- (man), mannitol- (mtl) and lactose- (lac) specific, respectively. Enzymes IIA 'hand' the phosphate on to enzymes IIB from the same family. These enzymes in turn phosphorylate sugars as they are translocated through the membrane with the help of the transmembrane proteins, enzymes IIC. In the figure both unphosphorylated (grey) and phosphorylated (red) states of EI and HPr are depicted, but for simplicity only one sphere for both unphosphorylated and phosphorylated states of enzymes IIA and IIB are drawn.

carbohydrate binds to the IIC protein, the membrane-spanning part of the PTS, where it is simultaneously transported into the cell, converted to the phosphate ester and channelled directly into the glycolytic pathway.

In milk fermenting bacteria, such as *Lactococci* and *Lactobacilli*, the disaccharide lactose and galactoside analogues are imported into the cell mainly via the PTS. The corresponding enzymes II belong to the lactose/cellobiose family. There are five individual steps in the overall import and phosphorylation reaction:

$$\text{PEP} + \text{enz I} \rightarrow \text{P–enz I} + \text{pyruvate} \quad (1)$$

$$\text{P–enz I} + \text{HPr} \rightarrow \text{P–HPr} + \text{enz I} \quad (2)$$

$$3\ \text{P–HPr} + \text{enz IIA}^{lac} \rightarrow \text{P}_3\text{–enz IIA}^{lac} + 3\text{HPr} \quad (3)$$

$$1/3\ \text{P}_3\text{–enz IIA}^{lac} + \text{enz IIBC}^{lac} \rightarrow 1/3\ \text{enz IIA}^{lac} + \text{P–enz IIBC}^{lac} \quad (4)$$

$$\text{P–enz IIBC}^{lac} + \text{galactoside} \rightarrow \text{galactoside-6-P} + \text{enz IIBC}^{lac} \quad (5)$$

In the case of *Staphylococci*, *Lactococci* and *Lactobacilli* enzymes IIA^{lac} and $IIBC^{lac}$ are induced by adding galactose to the growth medium. The IIA^{lac} component was first isolated from *Staphylococcus* and later from milk fermenting bacteria [2]. The corresponding genes have been cloned and overexpressed in *Escherichia coli* [3].

So far, the three-dimensional structure of the N-terminal domain of enzyme I has been determined by X-ray crystallography [4] and the structures of HPr–proteins from a variety of microorganisms have been solved using both crystallographic [5–9] and NMR methods [10–13].

Based on differences in their amino acid sequences [14] and on limited structural information [15], enzymes II have been classified into four families. The prototype representatives of these families are the glucose (glc), mannose (man), mannitol (mtl), and lactose-specific (lac) proteins, respectively. The crystallization of the lactose-specific enzymes IIA from *Staphylococcus aureus* [16] and *Lactococcus lactis* has been reported [17] and the secondary structure of *E. coli* IIA^{mtl} [18] and IIA^{man} [19] have been assigned by heteronuclear NMR spectroscopy. High-resolution X-ray structures are known for IIA^{glc} from *E. coli* [20] and from *Bacillus subtilis* [21] as well as for IIA^{man} from *E. coli* [22]. While an antiparallel β-sandwich fold is adopted by the IIA^{glc} domains and α/β folds are found in enzymes IIA^{man} and IIA^{mtl}, the lactose-specific soluble IIA domains have been predicted to be highly helical [23–25].

Structural information on the IIB domains is even more restricted: the three-dimensional IIB^{glc} structure has been solved by multidimensional NMR [26]; and a high-resolution X-ray structure [27] as well as a solution NMR structure [28] have just been published for the cellobiose-specific enzyme $IIB^{cel.}$ IIB^{cel} is also a member of the lactose family, which makes the determination of the enzyme IIA^{lac} structure reported here a timely event.

The enzyme IIA^{lac} of *S. aureus* as isolated is a homotrimer; all its subunits can be phosphorylated as is indicated in Reactions 1–5 [24,29]. Upon its phosphorylation the trimer separates into subunits and a significant change is observed in its circular dichroism (CD) spectrum. This has been interpreted as a major conformational change undergone by the protein [24]. The phosphorylated and non-phosphorylated forms of *S. aureus* enzyme IIA^{lac} differ in their behaviour towards detergents: in the presence of a positively charged detergent only P–IIA^{lac} interacts with the detergent micelle and migrates towards the cathode during charge shift electrophoresis [24].

This paper describes the high-resolution X-ray structure of enzyme IIA^{lac} from *Lactococcus lactis*. This protein is homologous to enzyme IIA^{lac} from *S. aureus* and shows very similar biochemical characteristics. Enzyme IIA^{lac} from *L. lactis* crystallizes as the functional trimer forming a nine-helix bundle. With the exception of the active site histidine residues, no similarities to other known enzymes IIA are observed but the complementarity of parts of its surface to HPr seems to be maintained. The results of our analysis are described in light of existing biochemical and molecular biological data.

Results and discussion

Structure determination and model quality

Unless stated otherwise, in the following report enzyme IIA will refer to *Lactococcus lactis* IIA^{lac}; enzymes from other sources or of different specificity will be labelled accordingly. Crystals of enzyme IIA resisted derivatization for a long time. Finally, trimethyllead acetate (TMLA) was successfully bound when crystals were transferred into a solution containing a 20 mM concentration of this reagent and soaked for seven days [17]. Diffraction data from native and derivative crystals were collected using synchrotron radiation at a wavelength of 0.94 Å. Measuring close to an X-ray absorption maximum for lead provided a combination of isomorphous and anomalous signals with sufficient phase information to solve the structure.

The results show that in the crystal each asymmetric unit contains one copy of the biochemically functional trimer of enzyme IIA. The individual subunits A, B and C are related to each other through a noncrystallographic threefold rotation axis. For all three subunits, our model contains residues 1–101. The electron density for the C-terminal residues 102–105 was very weak and could not be reliably interpreted. The root mean square (rms) deviations between the positions of all mainchain atoms of the symmetry related molecules are 0.52 Å for molecules A and B, 0.32 Å for molecules A and C and 0.48 Å for molecules B and C, respectively. The largest differences between individual subunits are observed for amino acids 51–70, which also show high temperature factors for mainchain as well as sidechain atoms. In addition, our current model contains 55 water molecules and one metal ion, most probably Mg^{2+} or Mn^{2+}. The Cα trace of the enzyme IIA monomer is illustrated in Figure 2a; the fold and topology of the trimer structure are shown in Figure 2b.

R_{cryst} and R_{free} (as defined in the Materials and methods section) calculated using all measured reflections in the resolution range 8 Å–2.3 Å are 18.1% and 24.8%, respectively. The average individual B factor for all atoms in the trimer is 34.4 $Å^2$. The geometry of the present molecular model is consistent with the values generally accepted for X-ray structures determined at this resolution [30]. The rms deviations of bond lengths and bond angles from the standard geometry are 0.012 Å and 1.35°, respectively. The Ramachandran plot [31] of the backbone torsion angles (ϕ,ψ) shows no residues outside the allowed regions of the plot, with 97.5% of them located in the most favoured regions.

Structure of the protein: secondary structure

In enzyme IIA, 83% of the amino acids are in an α-helical conformation. The three helices I (residues 3–32), II (residues 37–65) and III (residues 74–100) are between 27 and 30 residues long, which corresponds to approximately eight helical turns and a length of between 39–45 Å. Helices I and II are connected by residues 34–36 (Gly-Asp-Phe), which adopt an inverse γ-turn conformation [32,33]. The C terminus of helix II and the N terminus of the following helix III are 20 Å apart and linked by residues 66–73, which are in a random coil conformation.

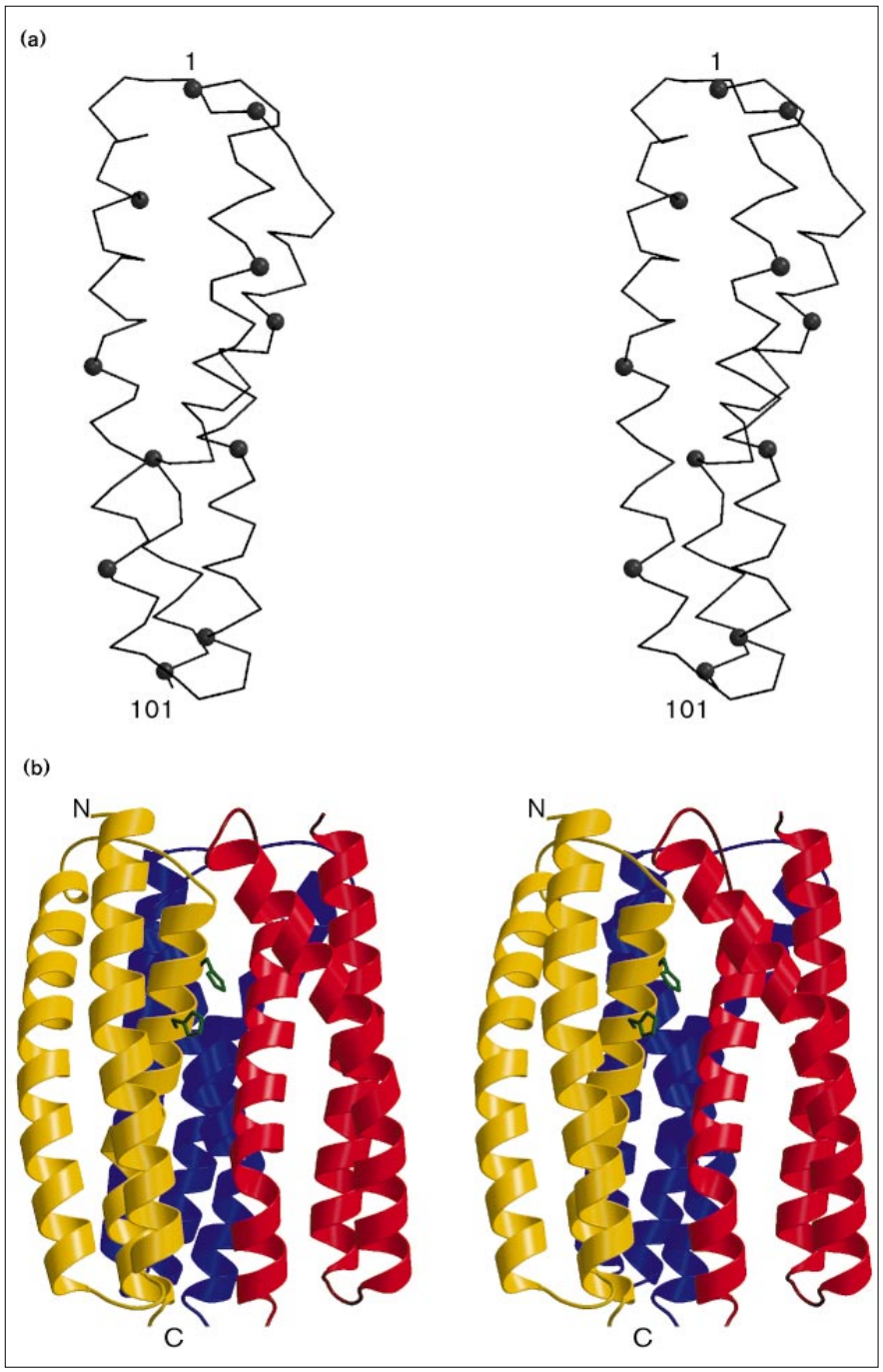
The crystallographically determined secondary structure of enzyme IIA is in good agreement with the results of the PHDsec program, a profile network-based secondary structure prediction package [34,35], which calculated an α helix content of 75%; the corresponding value for the closely related (72% amino acid identity) enzyme IIA^{lac} from *S. aureus* was 86.4%. No further experimental data regarding secondary structure assignments have been published for the *L. lactis* enzyme IIA. The secondary structure of enzyme IIA^{lac} from *S. aureus*, however, has been studied extensively. Based on CD studies, which estimated the α-helical content of *S. aureus* enzyme IIA^{lac} to lie between 50 and 53% [23,24], and using Chou and Fasman rules [36,37] Stuber *et al.* [25] proposed that each subunit of the *S. aureus* enzyme is built up from seven α-helical segments. Given the close relationship between the *L. lactis* and *S. aureus* proteins, our model suggests that the α-helical content of IIA^{lac} was previously underestimated. Some of the predicted non-helical regions in the seven α-helix model, when compared to the X-ray structure, are found to be located in the helical areas where the helices I, II and III show deviations from linearity.

Tertiary structure

Each subunit of *L. lactis* enzyme IIA forms a triple α-helical bundle, with the N and C termini located at opposite ends of the molecule and both of them exposed to solvent. If the top part of the molecule is defined as the one containing the N terminus, then helices I and III run downwards while helix II points upwards. Three-helical bundles with a down-up-down topology have been observed in proteins like spectrin [38], lysin [39] and the synthetic peptide coil-ser [40].

The helical bundle of enzyme IIA is stabilized by hydrophobic, electrostatic and hydrogen-bonding interactions. These interactions involve 21 residues between helices I and II, 33 residues between helices I and III and 18 residues contributing to contacts between helices II and III. Hydrophobic forces are clearly the dominating contributor to the formation of the three-helical bundle. The buried surface area corresponds to 1600 $Å^2$, equivalent to 26% of the total accessible surface area of the three helices combined. All three helices deviate from perfect linearity: helices I and II are bent by 15° and 35°, respectively; helix III is kinked at Val92, tipping the helix axis by 40°. As a result, the helices wrap around each other with two distinct packing patterns for the

Figure 2



The overall structure of *Lactococcus lactis* enzyme IIA^{lac}. (a) Stereo view of the Cα trace of the monomer; every tenth residue is marked with a black sphere. (b) Ribbon diagram of the trimer. The three monomers are shown in yellow, blue and red; the N and C termini are indicated. Sidechains (in green) of the phosphorylation site, His78 (top), and its neighbour His82 (bottom) are depicted. (Figure prepared using MOLSCRIPT [70].)

N-terminal (or 'top') and the C-terminal (or 'bottom') parts of the bundle.

Methionine cluster

In the top part of the enzyme IIA monomer, helices I and III pack closely together. Helix II, however, is positioned 20 Å away, connected via an eight residue extended chain. The resulting gap is filled by the sidechains of the hydrophobic core. These sidechains include a cluster of methionines, Met1, Met6, Met60 and Met76. The cluster extends to Met77 and Met84 and stabilizes the interactions between the three subunits of enzyme IIA (Fig. 3). In general, methionines occur relatively infrequently in proteins, but in enzyme IIA their abundance (5.8% of all

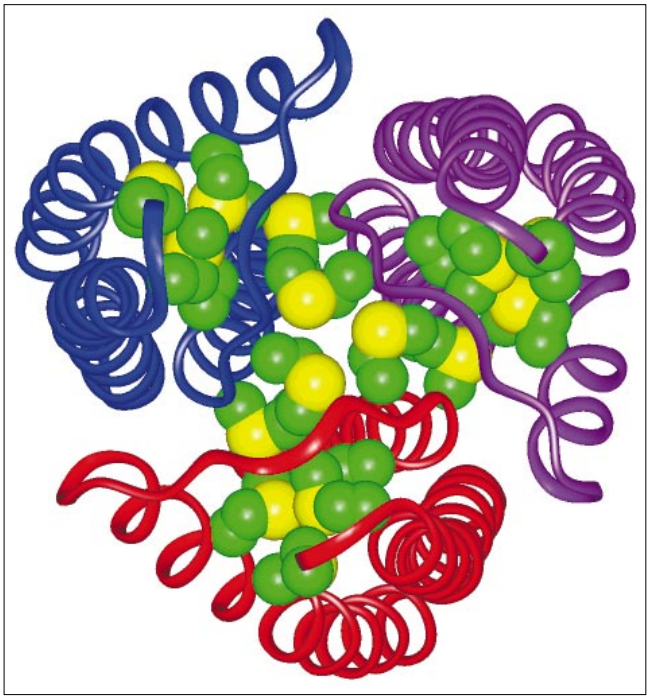
residues) is higher than the 3.8% average found in *E. coli* [41]. Although amino acids with long sidechains are required in the top part of enzyme IIA to stabilize the widely separated helices, methionine sidechains occupy only roughly the same volume as other frequently observed core residues like leucine, isoleucine or phenylalanine. The property, that sets methionines apart from the other large hydrophobic residues is their higher flexibility. Methionine mutants of T4 lysozyme, for example, showed an increase in their conformational repertoire compared to the wild-type enzyme [42]. In the case of enzyme IIA, the high content of methionine residues in the top portion of the molecule is most likely to be the reason why the C-terminal parts of helices II display high rms deviations between the subunits of the trimer. The methionine cluster, however, is not conserved in other enzymes IIA of the lactose/cellobiose family (Fig. 4) making any general functional role for it unlikely.

Coiled-coil packing of the enzyme IIA bundle

Closer to the C terminus, after the bend in helix II, the three-helix bundle converges. For their last five turns, the three helices are closely packed and the angles between their axes are 19°, 28° and 16° for helices I and II, II and III, I and III, respectively. This leads the helices to wrap around a superhelical axis, forming approximately one-eighth of a turn of a left-handed supercoil with a pitch of approximately 220 Å. The wrapping is mostly caused by bending of helix II and the 40° kink in helix III, while the small bend in helix I has little effect. This contrasts with typical α-helical bundles, in which all helices involved show a similar degree of bending (for a review of coiled-coils see [43]).

Another characteristic of coiled-coils, a repeating pattern of seven amino acids, is also present in *L. lactis* enzyme IIA. In such a heptad repeat, amino acid positions are conventionally labelled by the letters *a* to *g* with residues at positions *a* and *d* engaging in hydrophobic interactions between helices. In enzyme IIA, the *a* and *d* residues from

Figure 3



Backbone representation of the *Lactococcus lactis* enzyme IIA^{lac} trimer. View down the trimer axis with the top, N-terminal part of the molecule facing the viewer. The monomers are coloured in red, blue and purple. All the methionine residues in enzyme IIA are CPK rendered and coloured by atom type; carbon is in green and sulphur in yellow.

Figure 4

A structure-based sequence alignment of enzymes IIA from the lactose/cellobiose family. The secondary structure of IIA^{lac} from *Lactococcus lactis* is indicated above the primary sequence; cylinders represent helices. The alignment was created with the program Multiple Sequence Alignment [71]; identical residues are shown as white letters on a black background, conserved residues as white letters on a grey background. The organisms, their accession codes (SwissProt unless stated otherwise) and their percentage identity to the sequence of *L. lactis* IIA^{lac} are: *L. lactis*, P23532; *Staphylococcus aureus*, P02909 (72% identity); *Lactobacillus casei*, P11502 (51%); *Streptococcus mutans,* P26426 (69%); *Escherichia coli*, P17335 (35%); *Bacillus subtilis,* P46319 (37%); *Bacillus stearothermophilus*, PIR locus E49898 (41%).

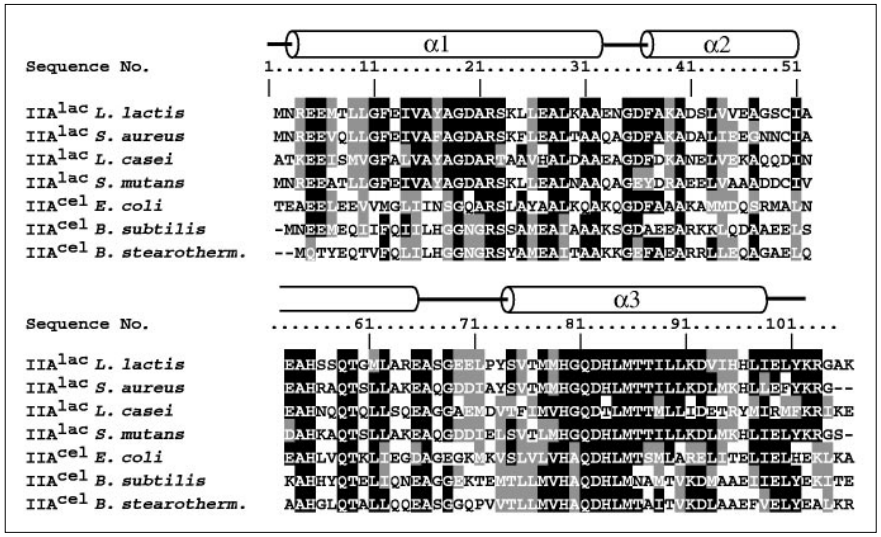
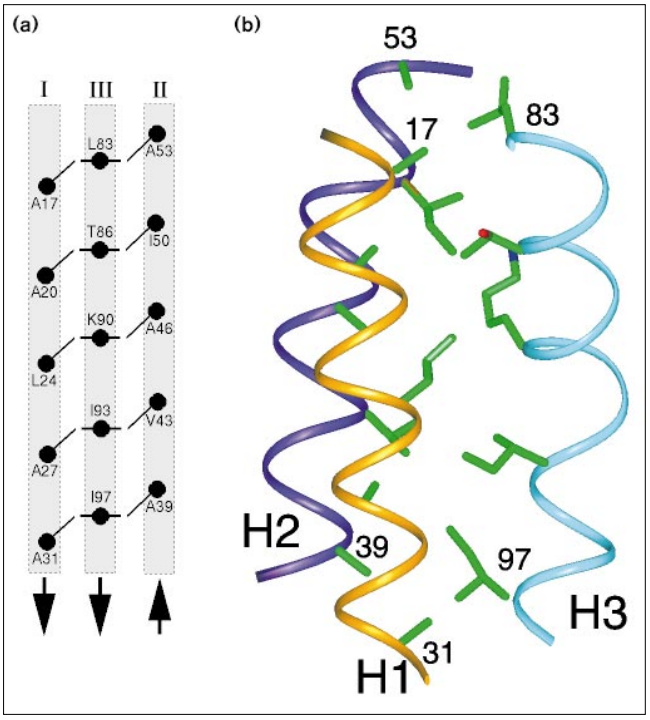


Figure 5



Packing of the helical bundle of *Lactococcus lactis* enzyme IIA^{lac}. The helical bundle packing in each monomer of enzyme IIA resembles a coiled-coil. Even though Cα atoms are on three different levels, the sidechains interact in the same planes. (a) Schematic diagram of interactions in the bottom, C-terminal part of the enzyme IIA monomer. Helices I, II and III are labelled at the top, and their direction is indicated at the bottom. Only the residues in positions *a* and *d* (see text) are depicted. The Cα atoms of labelled residues are marked with dots and the orientations of their sidechains are shown with short lines. (b) Cα trace of the coiled-coil region of the IIA^{lac} monomers. Sidechains of residues in positions *a* and *d* are rendered in the stick representation and coloured by atom type. The first and last residues involved in interactions from each of the helices H1–H3 are numbered.

helices I, II and III interact, forming five hydrophobic clusters: Ala31-Ala39-Ile97, Ala27-Val43-Ile93, Leu24-Ala46-Lys90, Ala20-Ile50-Thr86 and Ala17-Ala53-Leu83. Thr86 and Lys90 participate through the hydrophobic parts of their sidechains, all other residues occupying one of the positions *a* and *d* are hydrophobic.

The third characteristic of coiled-coils, the hallmark 'knobs-into-holes' packing, is also found in enzyme IIA, although it is accomplished in a rather unique way. Unlike in spectrin, coil-ser or other more regular coiled-coil proteins [43], the Cα atoms of the interacting residues *a* or *d* are on three different levels (Fig. 5). Keeping in line with our original definition of down-up-down topology, the Cα atoms of residues in helix I are the lowest with their sidechains pointing upwards. The Cα atoms in helix II are the highest with their sidechains pointing downwards. Cα atoms from helix III are located in between those of the two other helices and the hydrophobic parts of their sidechains are inserted between their contact partners from helices I and II. As a result, in-register packing of sidechains in enzyme IIA is similar to that found in other coiled-coil proteins. Most of the interacting residues *a* and *d* from helices I and II have short sidechains; four of the interacting residues in helix I and three of those in helix II are highly conserved alanines.

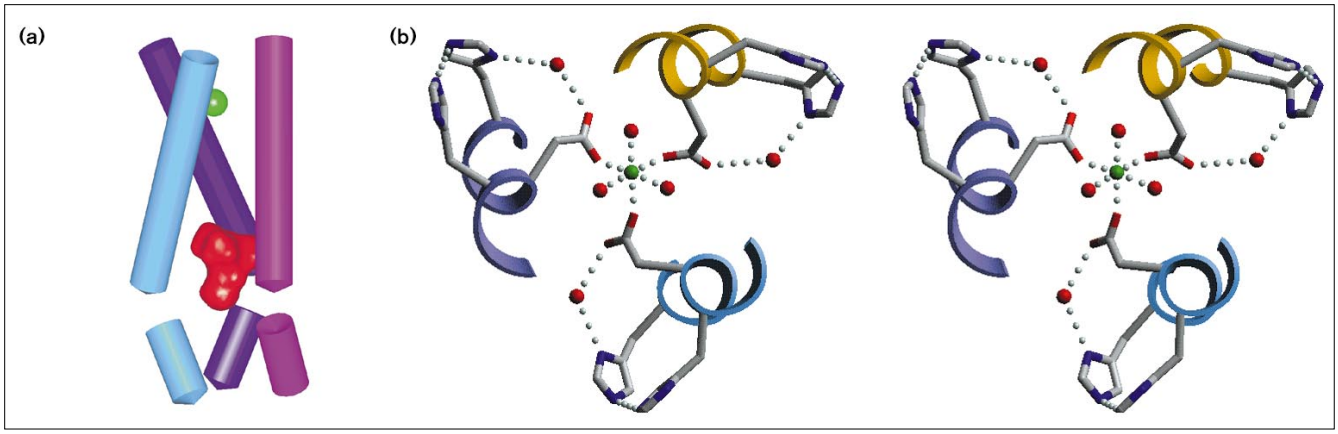
Several coiled-coil substructures have been identified in proteins but only one other example of the antiparallel three-stranded coiled-coil topology found in enzyme IIA has been published, that of the synthetic peptide coil-ser [40]. The distances between equivalent points on the helix axes in the C-terminal part of enzyme IIA vary between 7.5 Å and 11.5 Å. These distances are shorter than in coil-ser, where the corresponding values are 11.3 Å to 12.6 Å. This could be explained by the fact that in coil-ser Cα atoms *a* and *d* are located on the same level (in-register) in all three helices creating a less tightly packed bundle. In addition and contrasting with the large number of alanine residues in enzyme IIA, leucine and tryptophan residues are mostly found in positions *a* and *d* in coil-ser, amino acids whose bulky sidechains will also cause a loosening of the packing.

Quaternary structure

Studies in solution have postulated that the functional entity of enzyme IIA^{lac} is a homotrimer [44]. This is also the form which is found in the asymmetric unit of *L. lactis* enzyme IIA^{lac} crystals [17]. The subunits of the trimer are packed against each other forming a nine-helix bundle quaternary structure with each of the separate subunits A, B and C contributing three helices (Fig. 2b). The complex has the overall shape of a triangular prism with rounded edges, which has been twisted around its long axis. The height of the prism is 53 Å and its cross-section resembles an equilateral triangle with sides of 40 Å each. During oligomerization, a significant part of the 4919 $Å^2$ solvent-exposed surface area of each monomer is buried, therefore, the accessible surface area of the entire trimer is only 10 903 $Å^2$. For each individual monomer 1248 $Å^2$, or 26%, of the surface area becomes buried in the trimeric complex.

The intermolecular interactions within the trimer are mediated mostly by the symmetry-equivalent helices AIII, BIII and CIII, which make close contacts around the threefold symmetry axis. Starting from the N-terminal part of the molecule, the first five turns of the central helices are packed in a right-handed bundle with a crossing angle of –35°. For the last three turns of helix III, after its kink, the bundle becomes left-handed with a +32° crossing angle (Fig. 6a). Most of the intermolecular interactions between the inner helices obviously take place

Figure 6



Structural elements stabilizing the IIA^{lac} trimer. (a) Helices III from monomers A, B and C. In the top, N-terminal part of the bundle, helices are packed with a right-handed twist (crossing angle –35) and are stabilized by the divalent metal ion (green). The presence of the metal ion might be necessary in order to stabilize the trimeric structure of enzyme IIA and/or to regulate its assembly. After kinks at Val92, the bundle becomes left-handed (crossing angle +32). Kinks in the helices create a hydrophobic cavity which accommodates trimethyllead acetate (red). (b) Stereo view representation of the coordination geometry of the divalent metal ion (green sphere) and the hydrogen-bonding network involving bound water molecules (red spheres), Asp81, His78 and His82 of the threefold-related subunits. View down the trimer axis with the C-terminal, bottom part of the enzyme IIA trimer at the front. The metal ion binds to three water molecules and to the sidechains of three symmetry-related Asp81 residues. Asp81 acts as a hydrogen-bond donor to a structural water molecule, which in turn acts as a hydrogen-bond donor to the Nδ atom of His78. Hydrogen bonds and presumed metal–oxygen bonds are depicted with dotted grey lines. (Figure prepared using the program SETOR [72].)

around the trimers central axis of symmetry. The majority of these interactions are hydrophobic in nature but there are also other distinctive structural elements which stabilize the trimer. The sidechains of the three symmetry-equivalent copies of Tyr72 contribute by means of weakly polar interactions. The three carboxylates of the symmetry-related Asp81 residues all bind to a metal ion located in the centre of the trimer on the symmetry axis (Fig. 6b). Helices III are crucial for the assembly of the trimeric complex: they interact not only with each other but also make contacts with residues from two other helices. The contact area between helix AIII and helix CI involves 20 residues, the one between helix AIII and helix BII involves 11 residues.

Metal ion on the threefold symmetry axis

The discovery of a structural metal ion in enzymes IIA was unexpected as there was no previous evidence of its presence. The ion is positioned on the threefold axis and has octahedral coordination symmetry consistent with divalent metals. Its ligands are carboxylate oxygens from all three Asp81 residues as well as three buried water molecules (Fig. 6). In order to probe the chemical nature of the metal, we have refined its position without distance restraints, which resulted in average metal–oxygen distances of 2.09 Å. Although this result still does not allow an unequivocal identification of the metal, the geometry and ligand–ion distances are consistent with those of the divalent ions Mg^{2+} or Mn^{2+} [45]. As no divalent metals were added to our crystallization solutions, this metal ion is most likely to be an integral part of the trimeric structure. Conservation of Asp81 in homologues of enzyme IIA implies that this metal-binding site is preserved in these proteins as well. A very similar Mg^{2+}-binding site with octahedral geometry is also present on the threefold axis of dUTP pyrophosphatase from feline immunodeficiency virus [46]. In this and possibly other proteins, the presence of divalent metals could be necessary to regulate assembly and/or to stabilize the resulting trimeric structure.

Hydrophobic cavity

Oligomerization of enzyme IIA creates a hydrophobic cavity around the trimer symmetry axis in the lower part of the molecule (Fig. 6a). This cavity is lined with hydrophobic residues including Leu88, Leu89 and Val92 and is separated from the actual C terminus by two hydrophobic layers, comprising the sidechains of Leu96 and Leu99 and their counterparts in the other subunits. The volume of the cavity is 230 $Å^3$. The electron-density map does not show any evidence for the presence of structurally ordered water molecules or other compounds in the cavity. It is in this deeply buried cavity that the trimethyllead ions of the only heavy-atom derivative bind. The large size of the heavy metal compound, with an approximate volume of 140 $Å^3$, combined with the high concentration of TMLA and the long time of exposure necessary to achieve a reasonable occupancy, suggests molecular breathing as a mechanism to explain the

entry of the ion into the cavity, especially as an obvious entry channel is lacking.

Structural classification and comparison to other proteins

Enzyme IIA^{lac} is the first representative of the superfamily of lactose/cellobiose specific enzymes IIA whose three-dimensional structure is known. According to the Structural Classification of Proteins Database [47], it can be described as a member of the all α class, with a new fold of a multihelical trimer of identical three-helical subunits. It does not share structural homology with any of the other three known enzyme IIA families of the PTS system (i.e. the ones specific for glucose, mannose or mannitol transport). Enzyme IIA from the glucose-specific family displays β topology, while both the mannose and mannitol families have mixed α/β topologies [15]. While enzyme IIA^{man} forms a functional dimer, neither enzyme IIA^{glc} nor enzyme IIA^{mtl} require oligomerization for activity. A search of the Protein Data Bank [48] with the program DALI [49] using the enzyme IIA trimer as a search model revealed no proteins with any significant degree of structural homology, identifying enzyme IIA as unique. To the best of our knowledge, the fold of the nine-helix bundle has not been characterized previously. Searches were also performed using the enzyme IIA monomer and resulted only in rather low similarity scores. The highest score was obtained in a comparison with the ligand-binding domain of the bacterial aspartate receptor (PDB code 2LIG; [50]), which forms a dimer of two identical four-helix subunits. A superimposition of the two proteins showed that 81 Cα atoms from three of the receptor helices can be superimposed with Cα atoms from the three helices of enzyme IIA with an rms deviation of 3.0 Å.

The active site and phosphate transfer

Based on sequence alignments, site-specific mutagenesis [51] and NMR studies [52], it has previously been suggested that HPr phosphorylates enzyme IIA^{lac} on the Nε atom of His78. The crystal structure reveals that His78 is part of helix III, its sidechain protruding into a large groove between helices I and II of a neighbouring subunit (Fig. 7a). The trimer of enzyme IIA has three phosphorylation sites, related to each other by the threefold rotational symmetry and therefore 120° apart. Each His78 is in close proximity to two other histidines, His82 from the same subunit (3 Å away) and His54 from another subunit (7 Å away) (Fig. 7b). The close proximity between two histidines may not be so surprising as most of the other proteins from the PTS pathway seem to require an auxiliary charged residue, either histidine or arginine, to stabilize the histidine-bound phosphate. The His78–His82 pair of enzyme IIA can be fairly well superimposed onto the His83–His68 pair of IIA^{glc} (PDB code 1GPR; [21]), while the rest of the two structures are completely unrelated. It is not clear why His82 is not conserved in IIA^{lac} from *Lactobacillus casei* but in this organism its function might be replaced by residues corresponding to Arg21 or His54, residues also close to His78.

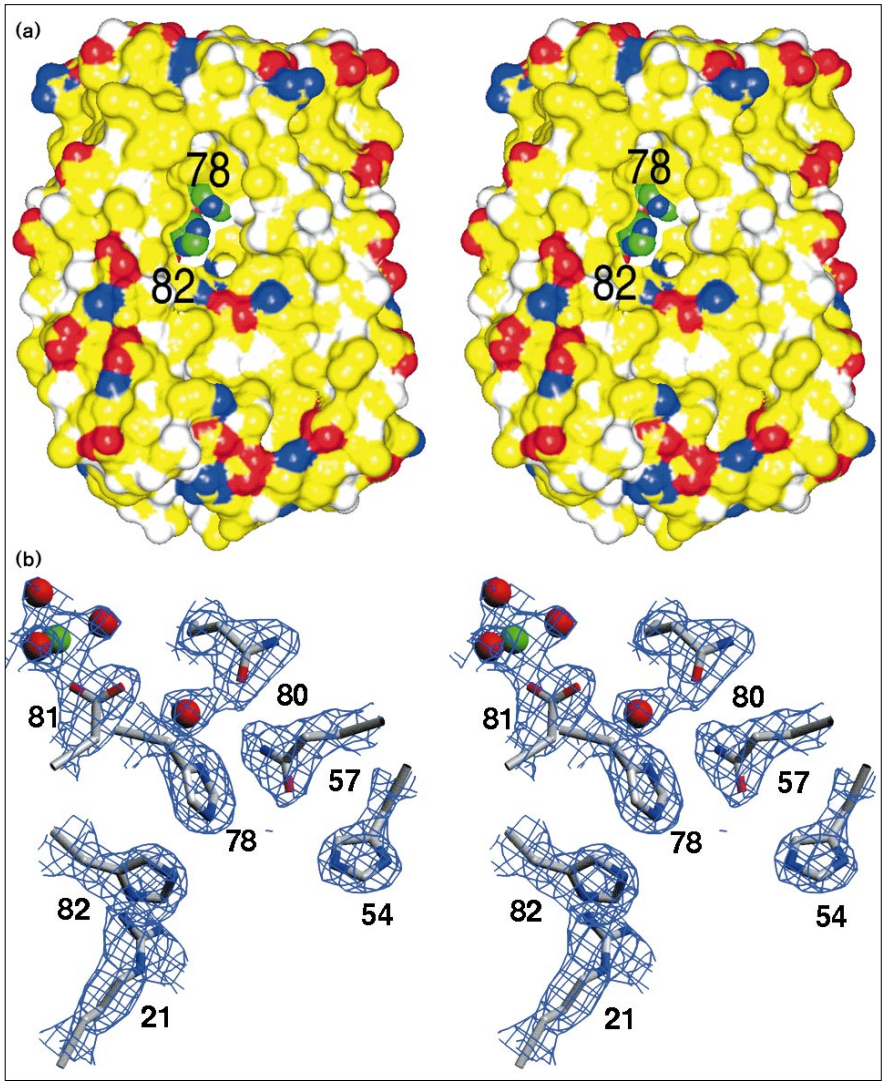
Although an electron-density map of 2.3 Å resolution will not allow us to locate the nitrogen atoms in the histidine sidechains, their positions can be deduced based on possible hydrogen-bonding patterns. In the case of enzyme IIA, there is a water molecule located deep inside the structure and within hydrogen-bonding distance of His78, Asp81 and Gln80 (Figs 6b,7b). The presence of this highly anchored water molecule in the active site provides additional rigidity to the phosphorylation site and fixes the imidazol ring of His78 in a distinct orientation. As His78 has a p*K* value of 6.0 [52] and is therefore mostly deprotonated at physiological pH, its Nε atom will not carry a hydrogen. It is ready to undergo a nucleophilic attack on the phosphoryl group of an approaching P–HPr molecule. This suggests that the protonated Nδ atom of His78 is hydrogen bonded to the active-site water and the water molecule then acts as a hydrogen-bond donor to Asp81 as well as to Gln80. The orientations of Asp81 and Gln80 are then stabilized further: the three Asp81 residues of subunits A, B and C are ligands of the central metal ion and Gln80 is part of a network of hydrogen bonds involving Gln57 and His54.

Protein interactions involving enzyme IIA

The structure of enzyme IIA can be of help in elucidating the mechanism of phosphate transfer in the lactose/cellobiose PTS family. As a first step in this direction, docking between enzyme IIA^{lac} and HPr was attempted. Once the coordinates of IIB^{cel} become available, docking studies involving enzyme IIA^{lac} and enzyme IIB^{cel} will also be possible. No conclusive results might be obtained, however, as upon phosphorylation enzyme IIA shows a large change in its CD spectrum and might even separate into its individual subunits [24].

All crystallographically determined HPr structures are of the free form of the protein; only one representation of its phosphorylated form is available (PDB entry code 1PFH; [53]), that of a family of 20 NMR structures of *E. coli* HPr. This assembly provides a picture of the phosphorylated state of the protein where the major structural differences are concentrated in the active site. It is justified to use the *E. coli* HPr model in a docking exercise with enzyme IIA as the PTS of *E. coli* and *L. lactis* are homologous (Fig. 4); *E. coli* HPr can even phosphorylate enzymes IIA^{lac} from gram-positive bacteria *in vitro* (WH, unpublished results). The docking procedure used for the enzyme IIA^{lac}–HPr complex was similar to the one described for the enzyme IIA^{glc}–HPr complex [54]. A family of phosphorylated HPr molecules, determined by NMR methods, was linked to the enzyme IIA^{lac} trimer with the phosphorus atom positioned ~2.0 Å away from the Nε atom of the active-site His78. The orientation of

Figure 7



Stereo view of the active site of enzyme IIA^{lac}. (a) Colour-coded solvent-accessible surface area of the enzyme IIA trimer: red, negatively charged; blue, positively charged; yellow, hydrophobic; and white, polar groups. His78 and His82 are CPK rendered with carbon atoms in green and nitrogen in blue. (b) The 2.3 Å $2F_o–F_c$ electron-density map displayed for residues close to the phosphorylation site, including His78, the metal ion (green) and the three coordinating water molecules (red). The electron density is contoured at 1.5σ. (Figure generated in SETOR [72].)

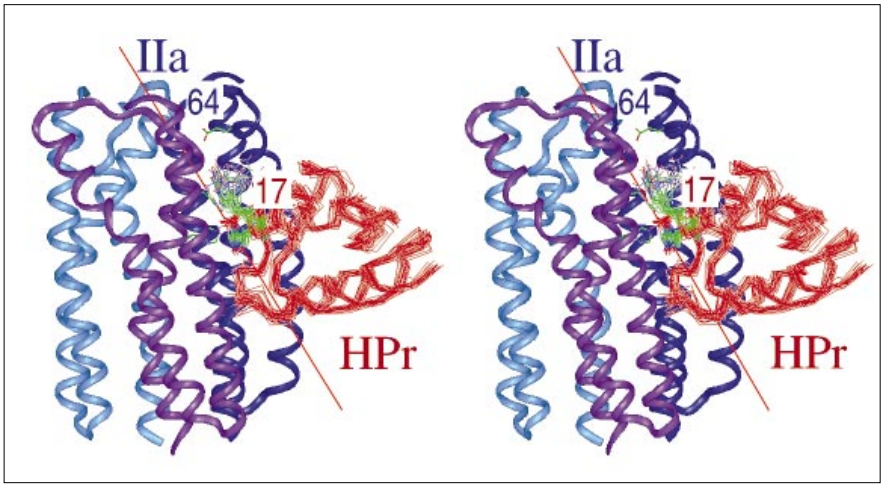
both molecules was further adjusted so that the HPr nitrogen atom carrying the phosphate, the phosphate atom and Nε of $His78_{IIA}$ were connected in a straight line. The family of HPr structures was then rotated around the N–P–N axis to minimize clashes and short contacts with enzyme IIA. In the actual docking, only the complementarity between the protein backbones, an important element in protein–protein recognition [55], was used to evaluate the best orientation of the molecules. Visual inspection of the complex revealed that within approximately 20°, the range of orientations does not lead to major overlaps between the two proteins.

In the modelled complex (Fig. 8), the interaction surface of enzyme IIA is contributed mostly by the region between helices I and II of two neighbouring subunits. The HPr residues which contact this area include two turns: residues 10–13 and 56–59. The surface of HPr between turns 37–40 and 56–59 as well as between residues 15–16 and 51–55 forms a large groove, which is filled by the C-terminal part of helix II of enzyme IIA. The surface area occluded by the interactions between enzyme IIA and the average structure of P–HPr (~1700 $Å^2$) is larger than the 1100 $Å^2$ of the enzyme IIA^{glc}–HPr complex [54] and close to the values observed in antigen–antibody complexes [56,57].

The interactions between enzyme IIA^{lac} and HPr are stabilized by electrostatic interactions between $Lys23_{IIA}$ and $Glu85_{HPr}$, and between $His78_{IIA}$ and $His–P_{HPr}$. In addition, further stabilization results from a few small hydrophobic patches involving contacts between $Leu25_{IIA}$ and $Pro10_{HPr}$ and between $Ala62_{IIA}$ and $Ala65_{IIA}$ and $Phe48_{HPr}$. Glu85 of HPr has not been implicated in electrostatic interactions of

Figure 8

Stereo view representation of a modelled enzyme IIA–HPr complex. With knowledge of the phosphorylation site, the family of NMR structures of P–HPr (PDB code 1PFH, coloured in red) was docked into the crystallographic model of enzyme IIA. The P–HPr ensemble of structures was rotated around the apical axis (red line) to minimize clashes and short contacts between the two proteins. Residues possibly involved in electrostatic interaction are $Glu64_{IIA}$ and $Arg17_{HPr}$. Sidechains of $His78_{IIA}$, $His82_{IIA}$ and $P–His15_{HPr}$ are displayed and coloured by atom type.



other modelled HPr complexes [22,54], in addition, Lys23 is not conserved in homologous enzymes IIA^{lac} making it most likely that their electrostatic interaction is coincidental and not essential for complex formation. $Pro10_{HPr}$ makes hydrophobic contacts in the modelled enzyme IIA^{man}–HPr complex [22] and $Met48_{HPr}$ of *B. subtilis*, which is replaced by Phe48 in *E. coli* HPr, makes several hydrophobic contacts in the enzyme IIA^{glc}–HPr model. The rest of the interacting surfaces of enzyme IIA and HPr are covered by polar residues which can stabilize the complex by means of hydrogen bonding.

It is surprising that the characteristic electrostatic interaction observed in other enzyme IIA–HPr complexes, between a negatively charged residue on enzyme IIA and the highly conserved $Arg17_{HPr}$, is most likely missing in the enzyme IIA^{lac}–HPr complex. The closest residue in enzyme IIA^{lac} which might possibly make such a contact is the highly conserved Glu64, but it is slightly buried inside the enzyme IIA^{lac} molecule and located 6–10 Å from the various conformations of Arg17 in the ensemble of HPr structures. Of course, it is conceivable that complex formation could induce conformational changes which would expose Glu64 on the surface making it available for interaction with $Arg17_{HPr}$. In such a case, however, this interaction could not be a part of the initial driving force of complex formation, but would contribute to the stabilization of the preformed complex.

The proposed model is in agreement with what would be expected for an enzyme IIA–HPr interaction. As enzyme IIA and HPr molecules have to readily separate after the phosphate transfer, the enzyme IIA–HPr complexes should not be characterized by too high an affinity. In the modelled complex between IIA^{lac} and HPr, the role of hydrophobic patches and electrostatic bridges is not very profound, despite the larger interaction surface as compared to other enzyme IIA–HPr complexes. In addition, our current model would make it difficult to explain how the IIA^{lac}–HPr interaction could be directly influenced by Ser46 of HPr, a residue whose phosphorylation state has been implicated in the lactose-specific PTS regulation [58]. Ser46 of HPr is not located near the binding interface and our results therefore support the interpretation that phosphorylation of Ser46 interferes with the phosphoryl transfer from enzyme I to HPr [59].

Structural implications for phase catalysis

The model of phosphate transfer in enzymes IIA from the lactose-specific PTS system suggested by Deutscher *et al.* [24], designates enzyme IIA^{lac} as a phase transfer catalyst. After phosphorylation, the hydrophobicity of enzyme IIA increases. This fact lead them to postulate that phosphorylation causes conformational changes, which lead to the separation of the subunits when they contact the membrane. The phosphorylated subunits would then 'float' in the membrane before transferring their phosphate groups to enzyme IIB, which in this specific case is fused to the channel component of the PTS system, the transmembrane protein enzyme IIC.

Several aspects of the enzyme IIA crystal structure can be interpreted in favour of the Deutscher model. The conformation of enzyme IIA is stabilized mostly by hydrophobic interactions and the protein's internal cavity decreases the stability of the complex. The strongest force holding the subunits together is the interaction between the three Asp81 residues and the central metal ion. If, however, these ion–ligand interactions are disturbed, then the negative charges on the aspartates would lead to repulsion between the subunits. Approximately 490 $Å^2$ of hydrophobic surface area per subunit buried in the enzyme IIA

trimer, would become exposed after subunit separation. The hydrophobic sidechains previously engaged in trimer contacts would now become available for interactions with the plasma membrane.

The question which remains unanswered is how the phosphorylation of enzyme IIA or the formation of the enzyme IIA–HPr complex could translate into a signal for subunit dissociation. Based on the current structure, we find that a mechanism that incorporates the liberation of the central metal ion as an early step would be a very appealing possibility. Additional insight into the nature of the phosphorylated state of enzyme IIA might be obtained from docking experiments involving enzyme IIA^{lac} and enzyme IIB^{cel} [27]. If the phosphorylated form of enzyme IIA separates into subunits the surface of IIB^{cel} would not be expected to show complementarity to the trimeric structure; even the short 105 amino acid peptide chain of the monomer could undergo major structural changes once it is removed from the stabilizing influence of the neighbouring subunits. On the other hand, if enzyme IIA remains in the trimeric form after phosphorylation, changes in the protein will be more local making it more feasible to construct a model of phosphorylated enzyme IIA docked to enzyme IIB. To address these questions from the experimental point, we currently aim our studies at the stabilization of the phosphorylated form of enzyme IIA^{lac}. In parallel, we are pursuing the crystallization of several complexes involving enzyme IIA^{lac}.

Biological implications

The phosphoenolpyruvate:sugar phosphotransferase system (PTS) is a multicomponent bacterial system which invokes a series of phosphoryl transfers between proteins. Its purpose is to import and phosphorylate sugars; it is also involved in several regulatory events, acting as a primitive signal transduction system. This additional, regulatory function of the PTS seems to be especially intriguing as several structures from the PTS reveal similarities and links to other bacterial systems. Recent structural studies of the proteins from the PTS are focused on determining the structures of missing members of the system, so that mechanistic models of phosphate transfer can be proposed.

The PTS consists of two non-specific proteins, enzyme I and HPr, which act in conjunction with a set of sugar-specific enzymes II. The latter enzymes are multiprotein permeases composed of the intracellular enzymes IIA and IIB and a transmembrane channel, enzyme IIC. Four different enzyme II families are known and the members of these families are specific for different sugars: glucose, lactose/cellobiose, mannose and mannitol. Many HPr structures are known and the first high-resolution structure of an enzyme IIB was recently published. Enzyme IIB belongs to the lactose/cellobiose family, and therefore provides opportunities for studying the spatial arrangements of possible IIA–IIB complexes.

The structure of enzyme $IIA^{lactose}$ reveals that it is a functional trimer with a novel topology of a nine-helical bundle. There is no resemblance to the β or α/β topologies of other enzymes IIA and the only common structural motif remaining is two neighbouring histidine residues at the phosphorylation site. The trimer is stabilized by a centrally located metal ion, which suggests that the presence of the metal might be involved in trimer assembly and regulation.

Based on the current structure of enzyme IIA^{lac} one can further explore the hypothesis put forward by Deutscher *et al.*, which proposes that after phosphorylation and upon contact with the bacterial membrane enzyme IIA separates into subunits. In favour of this hypothesis, a hydrophobic cavity was observed in the centre of the trimer which could decrease the stability of the complex. In addition, the binding of a large heavy-atom group in this cavity suggests that the protein can undergo breathing motions, which would also facilitate the separation of the subunits.

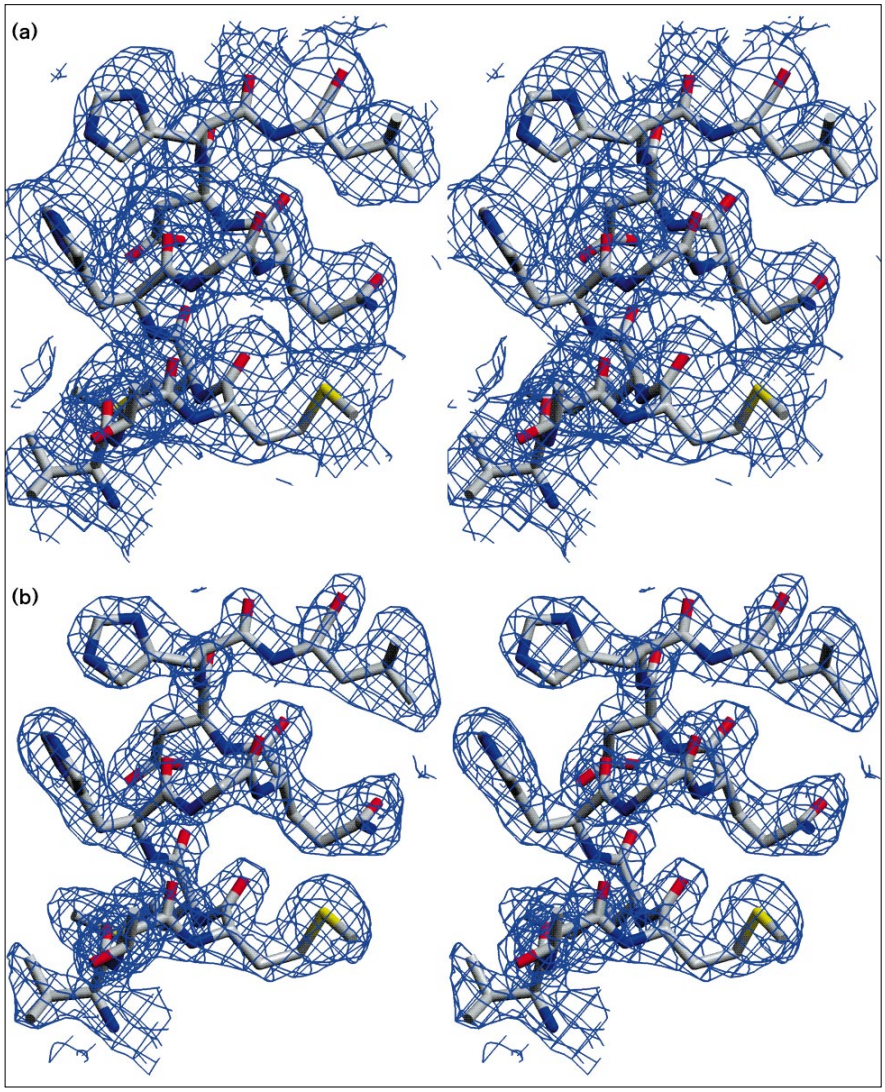
Although structural information about the phosphorylated forms of enzymes IIA and IIB is still lacking, it is now at least possible to start to study the interactions between the various components of a single pathway in structural terms. Surface complementarity, as described here for P–HPr and enzyme IIA^{lac}, should be a good indicator of the probability of the structures of complexes. Ultimately, however, experimental determination of the molecular structures of these complexes will be necessary.

Materials and methods

Protein purification and crystallization

Enzyme IIA^{lac} was cloned into *E. coli,* overexpressed and purified according to the procedure previously described [3]. The crystallization of this protein has been reported previously [17]. Crystals were grown by the sitting-drop method at 25°C, equilibrating 10 mg ml^{-1} of protein in 150 mM sodium/potassium phosphate buffer, pH 6.4, against a reservoir of 400 mM sodium/potassium phosphate buffer of the same pH. An isomorphous heavy-atom derivative was obtained by soaking native crystals for seven days in 400 mM sodium/potassium phosphate buffer, pH 6.4, containing 20 mM TMLA. Native and derivative X-ray diffraction data were collected at room temperature from a single crystal each on the X31 beamline of the EMBL outstation at the Deutsches Elektronen-Synchrotron (DESY) source in Hamburg, Germany, using the MAR image-plate detector. A wavelength of $\lambda = 0.94$ Å was used to increase the anomalous signal from the lead derivative. Weak diffraction was observed to 2 Å, but as R_{sym} values increased to more than 40% at resolutions higher than 2.3 Å, data collection and processing were truncated at this resolution. Analysis of the data using the program package DENZO [60] indicated that the crystals of *L. lactis* enzyme IIA^{lac} belong to the tetragonal space group $P4_12_12$ or its enantiomorph $P4_32_12$ (distortion index 0.19% versus 8.83% for primitive cubic) with unit-cell parameters $a = b = 90.9$ Å, $c = 82.4$ Å.

Figure 9



Electron-density maps calculated for a segment of helix III and contoured at 1.5σ levels. The final model of enzyme IIA is superimposed onto the electron density. **(a)** Solvent-flattened SIRAS experimental electron-density map. **(b)** Final $2F_o–F_c$ electron-density map.

Phase determination

The structure was solved using a combination of single isomorphous replacement (SIR), single-wavelength anomalous scattering (SAS) and solvent-flattening techniques. The presence and location of the heavy atom binding site was determined using the program package HEAVY [61]. The HASSP routine of this program uses a space group symmetry minimum method to obtain sets of atomic sites consistent with a Patterson function. Only one heavy atom binding site was located from this program, with the height of the peak equal to 6.6 times the standard deviation of the map. The probability for a chance occurrence of this interpretation was given as 0.000. The next highest peak was only 1.3 times the standard deviation.

All subsequent phase calculations were carried out using the program package PHASES [62]. The presence of an anomalous signal in the diffraction data was confirmed using Patterson methods. An anomalous difference Patterson map showed peaks at the same locations as the isomorphous map. The electron-density map calculated based on SIR/SAS signals was subjected to phase improvement techniques. Sixteen rounds of protein-boundary determination, solvent flattening and phase combination with experimental phases were carried out. The calculated map contained clear solvent channels and well defined electron density for the helical parts of the protein (Fig. 9a).

To further improve the map, we also attempted to include symmetry-averaging in the phase calculations. The noncrystallographic symmetry (NCS) of the trimer, which was initially determined using a self-rotation function, was evaluated using real space techniques. A solvent flattened electron-density map was skewed to the orientation of NCS and centered around the heavy atom binding site. After the trimer mask was created, the symmetry of the molecule was refined using the lsqrot function [62]. A shift of the NCS axis of 5° was observed. The refined position of the NCS axis was used in symmetry-averaging phase calculations, but did not result in an increase of the mean figure of merit from that of the solvent flattened map.

Model building, refinement and structure analysis.

The initial model of enzyme IIA was built using the graphics program O [63,64]. The location of the trimer in the unit cell was determined from the position of the TMLA-binding site, which was surrounded by the electron density of several α helices. Already in the early stages of map

Table 1

Data collection, phasing statistics and refinement.

	Native	TMLA derivative
Resolution (Å)	2.3	2.3
No. of unique reflections	15 009	15 417
Multiplicity	6.1	4.8
$< I > / \sigma(I)$	18.2/4.9*	14.9/5.0*
Completeness (%)	97.7/98.9*	99.8/99.9*
$R_{sym\,(I)}$ †	8.6/36.3*	10.9/34.0*
$R_{deriv\,(I)}$ ‡	N/A	15.0/21.7
Phasing statistics		
Number of sites		1
R_{cullis}(%)		56
R_{kraut}(%)		82
Phasing power		2.77
Mean figure of merit		0.783
Refinement statistics		
R factor (%)§		18.1
Free R factor#		24.8
Rms bond length (Å)		0.012
Rms bond angle (°)		1.35
Number of water molecules		55
Number of metal ions		1
Average B factor (Å^2)		34.4

* In highest resolution range (2.4–2.3 Å). †$R_{sym(I)} = (\Sigma|I - < I > |/\Sigma I) \times 100$, where I is the observed intensity, and $< I >$ is the average intensity obtained from multiple observations of symmetry-related reflections. ‡$R_{deriv(I)} = (\Sigma|I_D - I_N|/\Sigma I_N) \times 100$, where I_N and I_D are the observed reflection intensities for the native and heavy-atom substituted proteins, respectively, calculated using unique reflections. §R factor $= \Sigma|F_o - F_c|/\Sigma F_o$. #10% of reflections were used for R_{free}.

interpretation, it became obvious that the heavy atom binding site was located in the centre of the trimeric complex and that the structure of the protein monomer consisted of three helices. The mainchain for one of the molecules in the asymmetric unit showed continuous electron density, nevertheless superimposed maps as well as averaged maps at different resolution cut-offs were used for the initial chain tracing.

The initial model was refined using the program package X-PLOR [65]. The refinement protocol employed slow cooling of the system from 4000K to room temperature, followed by group B-factor refinement. Strict noncrystallographic constraints were used in these cycles of refinement. The molecular mask was calculated using the program MAMA from the RAVE suite of programs [66] and the model was rebuilt using electron-density maps which were symmetry-averaged using the program AVE (RAVE [66]). In the subsequent cycles of refinement, non-crystallographic restraints were used with decreasing force constants until the restraints were finally eliminated. The R_{free} factor based on 1109 reflections (representing 10% of the data between 8–2.3 Å) was used in all cycles of refinement in order to prevent over-fitting. The final refinement statistics are presented in Table 1 and an example of the final $2F_o–F_c$ electron-density map is shown in Figure 9b.

Assignment of the secondary structure of enzyme IIA and analysis of protein packing was carried out using the program PROMOTIF [67]. Rms deviations between symmetry-related molecules were calculated with the program LSQMAN (package DEJAVU [68]). Surface area and cavity volume calculations were performed using the program GRASP [69].

Accession numbers

The coordinates of enzyme IIA^{lac} have been deposited with the Brookhaven Protein Data Bank under the entry code 1E2A.

Acknowledgements

This work was supported by grants from the Medical Research Council of Canada, the National Science and Engineering Research Council of Canada and Connaught Laboratories, Inc. (to EFP). We are thankful to Klaus Schröter for his contributions in the initial phase of the project and to the staff of the EMBL laboratory at DESY for assistance with data collection.

References

1. Postma, P.W., Lengeler, J.W. & Jacobson, G.R. (1993). Phosphoenolpyruvate: carbohydrate phosphotransferase systems of bacteria. *Microbiol. Rev.* **57**, 543–594.
2. Hengstenberg, W., *et al*., & Kalbitzer, H.R. (1993). Structure and function of proteins of the phosphotransferase system and of 6-phospho-beta-glycosidases in gram-positive bacteria. *FEMS Microbiol. Rev.* **12**, 149–163.
3. de Vos, W.M., Boerrigter, I., van Rooyen, R.J., Reiche, B. & Hengstenberg, W. (1990). Characterization of the lactose-specific enzymes of the phosphotransferase system in *Lactococcus lactis. J. Biol. Chem.* **265**, 22554–22560.
4. Liao, D.I., Silverton, E., Seok, Y.J., Lee, B.R., Peterkofsky, A. & Davies, D.R. (1996). The first step in sugar-transport : crystal-structure of the amino-terminal domain of enzyme-I of the *Escherichia coli* pep: sugar phosphotransferase system and a model of the phosphotransfer complex with HPr. *Structure* **4**, 861–872.
5. Herzberg, O., Reddy, P., Sutrina, S., Saier, M.H., Jr., Reizer, J. & Kapadia, G. (1992). Structure of the histidine-containing phosphocarrier protein HPr from *Bacillus subtilis* at 2.0 Å resolution. *Proc. Natl. Acad. Sci.* **89**, 2499–2503.
6. Jia, Z., Quail, J.W., Waygood, E.B. & Delbaere, L.T. (1993). The 2.0 Å resolution structure of *Escherichia coli* histidine-containing phospho-carrier protein HPr. A redetermination. *J. Biol. Chem.* **268**, 22490–22501.
7. Jia, Z., Vandonselaar, M., Quail, J.W. & Delbaere, L.T. (1993). Active-centre torsion-angle strain revealed in 1.6 Å resolution structure of histidine-containing phosphocarrier protein. *Nature* **361**, 94–97.
8. Jia, Z., Vandonselaar, M., Hengstenberg, W., Quail, J.W. & Delbaere, L.T. (1994). The 1.6 Å structure of histidine-containing phosphotransfer protein HPr from *Streptococcus faecalis. J. Mol. Biol.* **236**, 1341–1355.
9. Pieper, U., Kapadia, G., Zhu, P.P., Peterkofsky, A. & Herzberg, O. (1995). Structural evidence for the evolutionary divergence of mycoplasma from gram-positive bacteria: the histidine-containing phosphocarrier protein. *Structure* **3**, 781–790.
10. Hammen, P.K., Scholtz, J.M., Anderson, J.W., Waygood, E.B. & Klevit, R.E. (1995). Investigation of a side-chain–side-chain hydrogen bond by mutagenesis, thermodynamics, and NMR spectroscopy. *Protein Sci.* **4**, 936–944.
11. Kalbitzer, H.R. & Hengstenberg, W. (1993). The solution structure of the histidine-containing protein (HPr) from *Staphylococcus aureus* as determined by two-dimensional 1H-NMR spectroscopy. *Eur. J. Biochem.* **216**, 205–214.
12. van Nuland, N.A., *et al*., & Robillard, G.T. (1994). The high-resolution structure of the histidine-containing phosphocarrier protein HPr from *Escherichia coli* determined by restrained molecular dynamics from nuclear magnetic resonance nuclear Overhauser effect data. *J. Mol. Biol.* **237**, 544–559.
13. Wittekind, M., Rajagopal, P., Branchini, B.R., Reizer, J., Saier, M.H., Jr. & Klevit, R.E. (1992). Solution structure of the phosphocarrier protein HPr from *Bacillus subtilis* by two- dimensional NMR spectroscopy. *Protein Sci.* **1**, 1363–1376.
14. Lengeler, J.W., Titgemeyer, F., Vogler, A.P. & Wohrl, B.M. (1990). Structures and homologies of carbohydrate: phosphotransferase system (PTS) proteins. *Philos. Trans. R. Soc. Lond. B [Biol.]* **326**, 489–504.
15. Herzberg, O. & Klevit, R. (1994). Unraveling a bacterial hexose transport pathway. *Curr. Opin. Struct. Biol.* **4**, 814–822.
16. Celikel, R., *et al*., & Reizer, J. (1991). Crystallization and preliminary X-ray analysis of the lactose-specific phosphocarrier protein IIA^{lac} of the phosphoenolpyruvate: sugar phosphotransferase system from *Staphylococcus aureus. J. Mol. Biol.* **222**, 857–859.
17. Sliz, P., Schröter, K.H., de Vos, W. & Pai, E. (1996). Crystallization and preliminary structural studies of lactose-specific enzyme IIA from *Lactococcus lactis. Acta Cryst. D* **52**, 1199–1201.
18. Kroon, G.J., Grotzinger, J., Dijkstra, K., Scheek, R.M. & Robillard, G.T. (1993). Backbone assignments and secondary structure of the *Escherichia coli* enzyme-II mannitol A domain determined by hetero-nuclear three-dimensional NMR spectroscopy. *Protein Sci.* **2**, 1331–1341.

19. Seip, S., *et al.*, & Erni, B. (1994). Mannose transporter of *Escherichia coli*. Backbone assignments and secondary structure of the IIA domain of the $IIAB^{Man}$ subunit. *Biochemistry* 33, 7174–7183.
20. Worthylake, D., Meadow, N.D., Roseman, S., Liao, D.I., Herzberg, O. & Remington, S.J. (1991). Three-dimensional structure of the *Escherichia coli* phosphocarrier protein III^{glc}. *Proc. Natl. Acad. Sci. USA* 88, 10382–10386.
21. Liao, D.I., Kapadia, G., Reddy, P., Saier, M.H., Jr., Reizer, J. & Herzberg, O. (1991). Structure of the IIA domain of the glucose permease of *Bacillus subtilis* at 2.2 Å resolution. *Biochemistry* 30, 9583–9594.
22. Nunn, R.S., et al., & Erni, B. (1996). Structure of the IIA domain of the mannose transporter from *Escherichia coli* at 1.7 Ångstroms resolution. *J. Mol. Biol.* 259, 502–511.
23. Simoni, R.D., Nakazawa, T., Hays, J.B. & Roseman, S. (1973). Sugar transport. IV. Isolation and characterization of the lactose phosphotransferase system in *Staphylococcus aureus*. *J. Biol. Chem.* 248, 932–940.
24. Deutscher, J., Beyreuther, K., Sobek, H.M., Stuber, K. & Hengstenberg, W. (1982). Phosphoenolpyruvate-dependent phosphotransferase system of *Staphylococcus aureus*: factor III^{lac}, a trimeric phospho-carrier protein that also acts as a phase transfer catalyst. *Biochemistry* 21, 4867–4873.
25. Stuber, K., Deutscher, J., Sobek, H.M., Hengstenberg, W. & Beyreuther, K. (1985). Amino acid sequence of the amphiphilic phosphocarrier protein factor III^{lac} of the lactose-specific phosphotransferase system of *Staphylococcus*. *Biochemistry* 24, 1164–1168.
26. Eberstadt, M., Grdadolnik, S.G., Gemmecker, G., Kessler, H., Buhr, A. & Erni, B. (1996). Solution structure of the IIB domain of the glucose transporter of *Escherichia coli*. *Biochemistry* 35, 11286–11292.
27. van Monfort, R.L., *et al.*, & Dijkstra, B.W. (1997). The structure of an energy-coupling protein from bacteria, $IIB^{cellobiose}$, reveals similarity to eukaryotic protein tyrosine phosphatases. *Structure* 5, 217–225.
28. Eiso, A.B., *et al.*, & Robillard, G.T. (1997). The NMR side-chain assignment and solution structure of enzyme $IIB^{cellobiose}$ of the phosphoenolpyruvate-dependent phosphotransferase system of *Escherichia coli*. *Protein Sci.* 6, 304–314.
29. Hays, J.B., Simoni, R.D. & Roseman, S. (1973). Sugar transport. V. A trimeric lactose-specific phosphocarrier protein of the *Staphylococcus aureus* phosphotransferase system. *J. Biol. Chem.* 248, 941–956.
30. 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.
31. Ramachandran, G.N. & Sasisekharan, V. (1968). Conformation of polypeptides and proteins. *Adv. Protein Chem.* 23, 283–438.
32. Milner-White, E., Ross, B.M., Ismail, R., Belhadj- Mostefa, K. & Poet, R. (1988). One type of gamma-turn, rather than the other gives rise to chain-reversal in proteins. *J. Mol. Biol.* 204, 777–782.
33. Rose, G.D., Gierasch, L.M. & Smith, J.A. (1985). Turns in peptides and proteins. *Adv. Protein Chem.* 37, 1–109.
34. Rost, B. & Sander, C. (1993). Prediction of protein secondary structure at better than 70% accuracy. *J. Mol. Biol.* 232, 584–599.
35. Rost, B. & Sander, C. (1994). Combining evolutionary information and neural networks to predict protein secondary structure. *Proteins* 19, 55–72.
36. Chou, P.Y. & Fasman, G.D. (1974). Prediction of protein conformation. *Biochemistry* 13, 222–245.
37. Chou, P.Y. & Fasman, G.D. (1977). Beta-turns in proteins. *J. Mol. Biol.* 115, 135–175.
38. Yan, Y., Winograd, E., Viel, A., Cronin, T., Harrison, S.C. & Branton, D. (1993). Crystal structure of the repetitive segments of spectrin. *Science* 262, 2027–2030.
39. Shaw, A., McRee, D.E., Vacquier, V.D. & Stout, C.D. (1993). The crystal structure of lysin, a fertilization protein. *Science* 262, 1864–1867.
40. Lovejoy, B., Choe, S., Cascio, D., McRorie, D.K., DeGrado, W.F. & Eisenberg, D. (1993). Crystal structure of a synthetic triple-stranded alpha-helical bundle. *Science* 259, 1288–1293.
41. Schulz, G.E. & Schirmer, R.H. (1985). Principles of protein structure. Springer-Verlag, New York, NY, USA.
42. Gassner, N.C., Baase, W.A. & Matthews, B.W. (1996). A test of the jigsaw puzzle model for protein folding by multiple methionine substitutions within the core of T4 lysozyme. *Proc. Natl. Acad. Sci. USA* 93, 12155–12158.
43. Lupas, A. (1996). Coiled coils: new structures and new functions. *Trends Biochem. Sci.* 21, 375–382.
44. Hengstenberg, W. (1977). Enzymology of carbohydrate transport in bacteria. *Curr. Topics in Microbiol. Immunol.* 77, 97–126.
45. Glusker, J.P. (1991). Structural aspects of metal liganding to functional groups in proteins. *Adv. Protein Chem.* 42, 1–76.
46. Prasad, G.S., *et al.*, & Stout, C.D. (1996). Crystal structure of dUTP pyrophosphatase from feline immunodeficiency virus. *Protein Sci.* 5, 2429–2437.
47. Murzin, A.G., Brenner, S.E., Hubbard, T. & Chothia, C. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* 247, 536–540.
48. Bernstein, F.C., *et al.*, & Tasumi, M. (1977). The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* 112, 535–542.
49. Holm, L. & Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* 233, 123–138.
50. Milburn, M.V., *et al.*, & Kim, S.H. (1991). Three-dimensional structures of the ligand-binding domain of the bacterial aspartate receptor with and without a ligand. *Science* 254, 1342–1347.
51. Finkeldei, U., Kalbitzer, H.R., Eisermann, R., Stewart, G.C. & Hengstenberg, W. (1991). Enzyme III^{lac} of the *Staphylococcal* phosphoenolpyruvate-dependent phosphotransferase system: site-specific mutagenesis of histidine residues, biochemical characterization and 1H-NMR studies. *Protein Eng.* 4, 469–473.
52. Kalbitzer, H.R., Deutscher, J., Hengstenberg, W. & Rosch, P. (1981). Phosphoenolpyruvate-dependent phosphotransferase system of *Staphylococcus aureus*: 1H nuclear magnetic resonance studies on phosphorylated and unphosphorylated factor III^{lac} and its interaction with the phosphocarrier protein HPr. *Biochemistry* 20, 6178–6185.
53. van Nuland, N.A., Grotzinger, J., Dijkstra, K., Scheek, R.M. & Robillard, G.T. (1992). Determination of the three-dimensional solution structure of the histidine-containing phosphocarrier protein HPr from *Escherichia coli* using multidimensional NMR spectroscopy. *Eur. J. Biochem.* 210, 881–891.
54. Herzberg, O. (1992). An atomic model for protein-protein phosphoryl group transfer. *J. Biol. Chem.* 267, 24819–24823.
55. Vakser, I.A. (1996). Main-chain complementarity in protein–protein recognition. *Protein Eng.* 9, 741–744.
56. Janin, J. & Chothia, C. (1990). The structure of protein–protein recognition sites. *J. Biol. Chem.* 265, 16027–16030.
57. Davies, D.R., Padlan, E.A. & Sheriff, S. (1990). Antibody–antigen complexes. *Annu. Rev. Biochem* 59, 439–473.
58. Ye, J.J., Reizer, J., Cui, X. & Saier, M.H., Jr. (1994). Inhibition of the phosphoenolpyruvate: lactose phosphotransferase system and activation of a cytoplasmic sugar-phosphate phosphatase in *Lactococcus lactis* by ATP-dependent metabolite- activated phosphorylation of serine 46 in the phosphocarrier protein HPr. *J. Biol. Chem.* 269, 11837–11844.
59. Napper, S., Anderson, J.W., Georges, F., Quail, J.W., Delbaere, L.T. & Waygood, E.B. (1996). Mutation of serine-46 to aspartate in the histidine-containing protein of *Escherichia coli* mimics the inactivation by phosphorylation of serine-46 in HPrs from gram-positive bacteria. *Biochemistry* 35, 11260–11267.
60. Otwinowski, Z. (1993). Oscillation data reduction program. In *Proceedings of the CCP4 Study Weekend*. (Sawyer, L., Isaacs, N. & Bailey, S., eds), pp. 56–62, SERC Daresbury Laboratory, Warrington, UK.
61. Terwilliger, T.C., Kim, S.-H. & Eisenberg, G. (1987). Generalized method of determining heavy-atom positions using difference Patterson function. *Acta Cryst. A* 43, 1–5.
62. Furey, W. (1991). PHASES, a program for phasing in macromolecular crystallography. University of Pittsburg, Pittsburg, PA, USA.
63. Kleywegt, G.J. & Jones, T.A. (1994). Halloween .... Masks and Bones. In *From First Map to Final Model*. (Bailey, S., Hubbard, R. & Waller, D., eds), pp. 59–66, SERC Daresbury Laboratory, Warrington, UK.
64. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for the building of protein models in electron-density maps and the location of errors in these models. *Acta Cryst. A* 47, 110–119.
65. Brünger, A.T. (1992). X-PLOR manual, version 3.1. *A System for X-ray Crystallography and NMR*. Yale University Press, New Haven, CT, USA.
66. Kleywegt, G.J. (1996). Making the most of your search model. *ESF/CCP4 Newsletter.* 32, 32–36.
67. Hutchinson, E.G. & Thornton, J.M. (1996). PROMOTIF a program to identify and analyze structural motifs in proteins. *Protein Sci.* 5, 212–220.
68. Kleywegt, G.J. & Jones, T.A. (1994). A super position. *ESF/CCP4 Newsletter.* 31, 9–14.
69. Nicholls, A.J. (1993). *GRASP Manual.* Columbia University, NY, USA.
70. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* 24, 946–950.
71. Lipman, D.J., Altschul, S.F. & Kececioglu, J.D. (1989). A tool for multiple sequence alignment. *Proc. Natl. Acad. Sci. USA* 86, 4412–4415.
72. Evans, S.V. (1993). SETOR: hardware-lighted three-dimensional solid model representation of macromolecules. *J. Mol. Graph.* 11, 134–138.